Single Molecule Fluorescence Studies of Viral Transcription



Francisco Javier Periz Coloma

St Cross College

University of Oxford

A thesis submitted for the partial fulfilment of the requirements for the degree of *Doctor of Philosophy*, Michaelmas Term 2013

Disclaimer

I declare that the work in this thesis is that of the candidate alone, with the exceptions described below.

Seamus J. Holden wrote the general data processing software that includes image registration and particle auto-detection; Kris Gryte wrote the software to plot the two dimensional E*/S histogram and time traces; Justin Pinkney wrote the software to plot the total fluorescence intensity from the red channel; Bo Jing wrote the data analysis software to correct non-uniform illumination in the field of view and calculate the number of colocalisations observed with respect to the number expected due to random coincidence; Ludovic Le Reste and Johannes Hohlbein built the microscope setups that I used to perform my research. Cristina Celma from LSHTM provided the viral and synthetic RNA and the DLP samples.

Francisco Javier Periz Coloma

October 2013

Acknowledgements

In a well known essay by Isaiah Berlin, he used a verse by Archilochus "The fox know many things, but the hedgehog knows one big thing" [1] to distinguish two types of thinkers and human beings in general; the hedgehogs which entertain an unified coherent vision of the world, and the foxes which by nature search for multiple paths of enquiry in the view that perhaps there is not such a unified view. My heterogeneous biological background makes me the latter kind of animal, so I would like to start this acknowledgement with a thanks and justification to Professor Achilles Kapanidis for accepting me as DPhil student in his lab of "Single-molecule Spectroscopy of Gene Machines", and for providing me with guidance and the financial support for my research. I did this DPhil in the belief that this research would help me to shape my own research vision. I would also like to thank the members in the lab that before and during my time here, facilitated my research. Special thanks to DPhils, Ludo Le Reste, Justin Pinkney, Kris Gryte, and Mr Bo Jing that built the microscopes and wrote data analysis software that I used in this research or were available for queries and discussion. Thanks to Professor Polly Roy and Dr Cristina Celma from LSHTM for providing invaluable DLP and ssRNA samples for this research. I am also grateful to the MRC for funding my studentship, St Cross College in Oxford for giving me financial support to attend a research conference and providing a much needed quiet space for writing my thesis. I must thank also my parents for being on my side along these years and my wife, Frances for her love and support during the writing up of this thesis when our baby daughter, Emma was born. I hope I can spend with you both more time now. I dedicate this thesis to you Emma, newborn, and still without words as I am when I see you, but perhaps years later, this thesis lets you understand the path that I was following.

1. Berlin, I. (2009) The Hedgehog and The Fox. An Essay on Tolstoy's View of History. *1st ed Isaiah Berlin 1953, Hogarth Press, ed Hardy H and Kelly A. Paperback edition. Phoenix Paperback..Orion House Ltd.*, 1-85.

Single Molecule Fluorescence Studies of Viral Transcription

Francisco Javier Periz Coloma, St Cross College

A thesis submitted for the degree of Doctor of Philosophy
University of Oxford, Michaelmas Term 2013

Abstract

Rotaviruses are the single most common cause of fatal and severe childhood diarrhoeal illness worldwide (>125 million cases annually). Rotavirus shares structural and functional features with many viruses, such as the presence of segmented double-stranded RNA genomes selectively and tightly packed with a conserved number of transcription complexes in icosahedral capsids. Nascent transcripts exit the capsid through 12 channels, but it is unknown whether these channels specialise in specific transcripts or simply act as general exit conduits; a detailed description of this process is needed for understanding viral replication and genomic organisation. To test these opposing models, a novel single-molecule assay was developed for the capture and identification (CID) of newly synthesised specific RNA transcripts. CID combines the hybridisation of transcripts with biotinylated and FRET compatible labelled ssDNAs with the implementation of recent developments in single molecule fluorescence such as alternating laser excitation (ALEX) and total internal reflection fluorescence (TIRF) microscopy. CID identifies and quantifies specific transcripts of rotavirus based on a FRET/Stoichiometry (E*/S) value of the hybridised labelled probes. I used CID to pull down the capsid on the surface slide and identify partially extruded transcripts of three different segments 2, 6 and 11. The findings presented in this thesis support a model in which each channel specialises in extruding transcripts of a specific segment, that in turn is linked to a single transcription complex. The method can be extended to study other transcription systems including *E.coli*, and can be further developed as a potential diagnostic tool.

Contents

Chapter 1. Introduction 1

1.1. Rotavirus - a case model for the study of transcript export in dsRNA viruses 2

1.2. Outline of the thesis 3

Chapter 2. The biology of rotavirus 5

2.1. Rotavirus is a member of the *Reoviridae* 6

2.2. Organisation of the capsid in rotavirus 9

2.2.1. DLP structure 13

2.2.2. The TLP 16

2.3. The genome organisation 17

2.4. Overview of the life cycle 18

2.5. Structural study of transcription and replication of rotavirus 23

2.6. Translation 26

2.7. Packaging assembly and virus release 27

2.8. Model for packaging and assortment 28

Chapter 3. Materials and methods 35

3.1. Basic concepts in single molecule FRET 36

3.2. Single molecule Förster Resonance energy transfer (FRET) 40

3.3. TIRF-ALEX optical set up 42

3.3.1. Laser light source (excitation module) 43

3.3.2. TIRF illumination module 44

3.3.3. The concept of ALEX 46

3.3.4. Camera and image splitter device (emission module) 49

3.3.5. Detection and linkage of G and R fluorescence particles 50

3.3.6. Plotting of fluorescence emission intensity of G and R particles 51

3.3.7. Correction of non-uniform illumination in the field of view 51

3.3.8. Number of colocalisations observed vs. the number expected due to random coincidence 52

3.4. Confocal ALEX microscopy 53

3.5. TIRF slide preparation 53

3.5.1. Vectabond coating 54

3.5.2. PEG coating 54

3.5.3. Neutravidin coating and DNA incubation 55

3.6. DNA design and hybridisation. 55

3.6.1. Polymerase chain reaction (PCR) 57

3.6.2. Measuring DNA concentrations 58

3.6.3. DNA labelling 58

3.6.4. DNA purification protocol 59

3.7. Molecular virology methods 59

3.7.1. Synthetic RNA control: T7 synthetic RRV 2 and 11 segment 59

3.7.2. Double-layered particles (DLPs) purification 60

3.8. Capture and identification (CID) assays 61

3.8.1. Capture assay of transcripts in solution 61

3.8.2. Capture assay of extruded transcripts from DLPs 61

3.8.3. Capture of DLPs bearing extruded transcripts using capture antibodies. 63

Chapter 4. Design and development of an assay to capture and identify ssRNA from rotavirus 64

4.1. Introduction 65

4.2. Results 66

4.2.1. Experimental design of a capture assay based on single molecule hybridisation of ssRNAs 66

4.2.2. Initial experiment - Hybridisation of labelled probes to ssDNA complementary biotinylated probes in NaCl, Mg^{2+} 69

4.2.3. Detection of ssRNA transcripts released in solution from DLPs 70

using, fluorimetry and confocal microscopy
4.2.4. Description of CID assay. Proof of principle. Synthetic ssRNAs from segment11 74
4.2.5. Capture of released transcripts from DLPs using CID 77
4.3. Discussion 79

Chapter 5. Investigation of mRNA release models in double-layered viral particles of rotavirus 82
5.1. Introduction 83
5.2. Results 85
5.2.1. Capture of extruded transcripts (+)ssRNA 11 from DLPs using CID 85
5.2.2. Capture of extruded transcripts (+)ssRNA 6 from DLPs using CID 89
5.2.3. Capture of extruded transcripts (+)ssRNA 2 from DLPs using CID 92
5.2.4. Actively transcribing DLPs carry single extruded transcripts of segments 6 and 11 95
5.3. Discussion 98

Chapter 6. Detection of RNA transcripts in RNA polymerase elongation complexes using single-molecule hybridisation 101
6.1. Introduction 102
6.1.1. Bacteria transcription 103
6.1.2. RNAP structure 103
6.1.3. The bacteria transcription cycle 106
6.1.4. Transcription initiation 106
6.1.5. Early phase of transcription elongation 107
6.1.6. Mature transcription elongation 108
6.1.7. Single molecule ALEX to investigate transcription factor retention 109
6.2. Methods 109
6.2.1. Expression and purification of $\sigma 70_{596}$ 109

6.2.2. PCR and purification of LACUV5(-40Bio+64Cy5) and probe labelling 109
6.2.3. Formation of RPo 110
6.2.4. Standard transcription single molecule experiment 110
6.3. Results 111
6.3.1. Expression and purification of $\sigma 70_{596}$ and production of LACUV5(-40Bio+64Cy5) 111
6.3.2. RPo formation optimisation 111
6.3.3. Ensemble transcription experiments 115
6.3.4. Single-molecule transcription experiments 117
6.3.5. Single-molecule hybridisation in real time 121
6.4. Discussion 124

Chapter 7. General Discussion 128
7.1. General discussion 129
7.2. RNA packaging and transcription in RV 131
7.1.3. Prospects for the CID assay 132
7.2. Further research 135

References 140

List of figures and tables

Fig. 2.1. Architecture of rotavirus capsid 10

Fig. 2.2. Structure of the VP2 in the SLP and its interaction with VP1 15

Fig. 2.3. Rotavirus replication cycle 19

Fig. 2.4. Processing of VP4 and conformational changes associated with membrane penetration 20

Fig. 2.5. Structure of VP1 24

Fig. 2.6. Models of assortment and packaging of (+)RNA segments in RV 33

Fig. 3.1. Jablonsky diagram of the energy levels of a fluorophore 37

Fig. 3.2. Förster Resonance energy transfer (FRET) between two fluorescent molecules 40

Fig. 3.3. Relative orientation of the D emission and the A absorption dipoles 42

Fig. 3.4. TIRF-ALEX optical set up 43

Fig. 3.5. Schematic of TIRF illumination 45

Fig. 3.6. Alternating-laser excitation (ALEX) combined with TIRF microscopy 47

Fig. 4.1. Time course hybridisation of R_{11} to T_{11}-G_{11} 68

Fig. 4.2. Time course hybridisation of R_{11} to T_{11}-G_{11} construct detected using TIRF-ALEX 70

Fig. 4.3. Hybridisation of R_{11} and G_{11} to (+) RNA11 in solution detected using confocal-ALEX microscopy and fluorimetry 73

Fig. 4.4. CID detects synthetic transcripts from segment 11 76

Fig. 4.5. Detection of nascent transcripts from transcriptionally active DLPs 77

Fig. 4.6. CID assays for ssRNA11 and ssRNA6 transcripts produced by rotavirus DLPs 79

Fig. 5.1. Models for transcript exit in rotavirus 84

Fig. 5.2.Transcriptionally active DLPs carry a single extruded transcript for segment 11 86

Fig. 5.3. Standards analysed using E*-S histograms and fluorescence 88

intensities
Fig. 5.4. Transcriptionally active DLPs carry a single extruded transcript for segment 6 91
Fig. 5.5. Transcriptionally active DLPs carry a single extruded transcript for segment 2 92
Fig. 5.6. Antibody-captured transcriptionally active DLPs carry a single extruded transcript for segment 2 94
Fig. 5.7. Colocalisation of (+)ssRNA6 and (+)ssRNA11 on DLPs 96
Fig. 6.1. Transcription cycle of bacteria RNAP 104
Fig. 6.2. Anion exchange purification and labelling of $\sigma 70_{596C}$ and PCR of LacUV5-50Cy5 template 112
Fig. 6.3. Formation and optimisation of RPo open complex 114
Fig. 6.4. Formation of RDe. Single-molecule transcription experiments 117
Fig. 6.5. Hybridisation of complementary probe G(+5/+28) to staled RDe Single-molecule hybridisation in real time 119
Fig. 6.6. Real time hybridisation of RDe complexes immobilized on TIRF slides. A schematic of the single-molecule transcription experiment 122
Fig. 6.7. Hybridisation kinetics of G(+5/+28) probes to stalled RDe complex 123
Fig. 7. 1. Stepwise packaging model of ssRNAs in spatially conserved localisations within the RV capsid 134
Fig. 7.2. Study of spatial conservation of segments within the BTV viral capsid 137
Fig. 7.3. Study of spatial conservation of segments within the viral capsid using a CID assay 139

Table 2.1.Genera of the *Reoviridae* family 7
Table 2.2. RNA segments and proteins of rotavirus 11
Table 3.1. Oligonucleotides used for hybridisation in the viral studies 56
Table 3.2. Oligonucleotides used for studying transcription in *E. coli* RdRP 57

Nomenclature

ALEX alternating laser excitation

BTV bluetongue virus

C capture

CID capture and identification assay

DLP double-layered particle

DMSO dimethyl sulfoxide

DNA (P) DNA promoter

dsRNA double-stranded RNA

E* raw FRET

ELISA enzyme-linked immunosorbent assay

EMCCD electro multiplying charge-coupled device

ERGIC endoplasmic reticulum golgi intermediate compartment

FRET Förster resonance energy transfer

G green

Gloxy glucose oxydase

ICTV international committee for the Taxonomy of viruses

MOI multiplicity of infection

ORV orthoreovirus

R red

RDe elongation complex

RdRP RNA dependent RNA polymerase

RI replication intermediate

RNAP RNA polymerase

RNAP(i) intermediate complex

RP(c) closed complex

RPo open complex

RRV rhesus rotavirus

RV rotavirus

S stoichiometry

SA sialic acid

SLP single-layered particle

sm single molecule

TC transcription complex

TIRF total internal reflection fluorescence

TLP triple-layered particle

Trolox 6-hydroxyl-2,5,7,8-tetramethychroma-2-carboxylic acid

λ wave length

Chapter 1

Introduction

1.1. Rotavirus - a case model for the study of transcript export in dsRNA viruses

Double-stranded RNA (dsRNA) viruses are a diverse group, but they share certain unique features and structures reflecting parallels in their replication cycles; this is true in particular among the members of the *Reoviridae,* a family that includes many viruses of veterinary and biomedical importance. One of the best studied is rotavirus, a virus known as one of the most important causes of gastroenteritis in infants and young children worldwide.

Distinguishing features for the *Reoviridae* include a multicomponent capsid that crosses the host membrane, and transcription of dsRNA segments by enzymes integral to the structure. During cell entry, the outer protein layers of rotavirus (RV) are lost, while their inner capsids provide a compartment for the multiple genome segments (10-12 dsRNAs) to be transcribed repeatedly by capsid-associated transcription complexes.

Export of transcripts in RV occurs via channels at the 12 vertices of the icosahedral capsid. Although crucial for the establishment of infection, it is not clear whether these channels specialise in the extrusion of specific transcripts (channel specialisation model) or simply act as general exit conduits (channel generality model); progress in the characterisation of the channel function has been hindered by the lack of methodologies and the intrinsic difficulty in identifying specific single RNA transcripts at the time they are extruded from the viral capsid.

To address this question, I set two aims for this thesis, first to develop a single molecule hybridisation method to capture and identify the copy number of ssRNAs from rotavirus (hereafter CID assay) capture and identification assay ; second to

perform functional studies to test these opposing channel-utilization models. The findings in this thesis support the "channel-specialisation" model, in which an individual segment is transcribed by a single transcription complex and extruded through a specific channel. While these experiments were performed in rotavirus, a variant of this assay was also implemented in the *E. coli* in vitro transcription model, supporting its use in different transcription models. The CID assay also has the potential of being used in broader applications including the biosensing of ssRNA.

1.2. Outline of the thesis

Chapter 2 describes the biology of rotavirus, study model representative of the *Reoviridae* family. This chapter introduces the main aspects of its biology in the wider context of the *Reoviridae* family, and describes the structural and functional knowledge available on the capsid proteins and the genome; then it summarises the replication cycle of the virus and poses the questions still open on the packaging and assembly of the virus, describing two current models.

Chapter 3 describes the main fluorescence methods used in this thesis, which includes mainly single molecule Förster Resonance energy transfer (FRET), total internal reflection fluorescence (TIRF) and alternating laser excitation (ALEX). Then the chapter describes in detail the microscopy set up, the main methods in sample preparation and the methodology used to implement the data analysis.

Chapter 4 describes the design, development and implementation of a novel single molecule hybridisation assay for the capture and identification (CID) of ssRNA transcripts in RV. The results showed the reliability of the assay to detect and quantify the copy number of ssDNA, and synthetic RV ssRNA (both in solution and released from DLPs).

Chapter 5 describes two models of ssRNA export through channels in RV, called the "channel generality model" and the "specialisation model". The validity of these models is tested using the assay developed in the previous chapter. The results support the specialisation model, The biological implications of the model are discussed.

Chapter 6 expands the principle of single molecule hybridisation applied in the CID assay to investigate the detection of ssRNA from RdRP. The results show the detection of transcripts in real time, and demonstrate that the E. coli in vitro system is a useful tool to optimise the complexity of the transcription assay in RV.

Chapter 7 summarises the work of this thesis, discusses the channel specialisation of RV in the context of the packaging and assembly of the virus; and then discusses the CID assay in relation to the family of fluorescence in situ hybridisation assays (FISH). A discussion about further experiments raised from this thesis is presented.

Chapter 2

The Biology of Rotavirus

2.1. Rotavirus is a member of the *Reoviridae*

Double-stranded (dsRNA) viruses represent a large group of viruses that bear segmented genomes (1-12) within their capsid (1). One of the most relevant families due to its genome complexity and range of hosts is the *Reoviridae.* The *Reoviridae* family has 10-12 segments protected by a complex multilayered icosahedral capsid. The current classification of the *Reoviridae* by the ICTV (International Committee for the Taxonomy of Viruses) includes 15 genera that infect different hosts including vertebrates, invertebrates, plants and fungi (2) (Table 1.1.)

Despite the wide range of hosts, members of the *Reoviridae* share structural and functional properties. The capsid is complex and consists of concentric shells of proteins with outer, middle and inner layers. The outer layer mediates entry into the cell. The middle and inner layers are never fully disassembled and contain a conserved number of dsRNA segments and transcription enzymes that synthesise (+)ssRNA which are exported to the cytosol through 12 channels located at the five fold axis of the inner layer capsid (3).

There are two-subfamilies in the *Reoviridae*: turreted and non-turreted viruses, (turreted meaning the presence on the surface of the middle-layered capsid of protein projections ‘turrets’ or their absence). One of the most studied genus in the turreted group is the orthoreovirus (ORV). Non-turreted viruses include viruses of medical importance such as coltivirus and rotavirus (RV), and veterinary importance such as bluetongue virus (BTV) and african horse sickness (Table 2.1).

Table 2.1. Genera of the *Reoviridae* family

Virus species	Structure	No. of genome segments	No. of species; proposed species (+)	Host and vectors
Aquareovirus	Non-turreted	11(~23.7 Kbp)	6(+5)	Molluscs, finfish, Crustacea
Cardoreovirus	Non-turreted	12(~20.15 Kbp)	1(+2)	Crabs
Coltivirus	Non-turreted	12(~29 Kbp)	2(+1)	Mammals (including humans), ticks
Cypovirus	Turreted	10(~24.8-33.3 Kbp)	21(+2)	Insects (Lepidoptera, Diptera and Hymenoptera), a single isolate was reported from freshwater daphnid
Fijivirus	Turreted	10(~28.7 Kbp)	5	Plants (Graminae, Liliaceae), insects (Delphacidae, planthoppers)
Dinovernavirus	Turreted	9(~23.35 Kbp)	1	Mosquitoes
Idnoreovirus	Turreted	10+1 (25.1 Kbp)	5(+1)	Insects (Hymenoptera)
Mimoreovirus	Non-turreted	11(~25.56 Kbp)	1	Phytoplankton
Mycoreovirus	Turreted	11-12(~24.4 Kbp)	3	Fungi
Orbivirus	Non-turreted	10(~19.2 Kbp)	22+(+10)	Mammals (including humans), birds, *Culicoides*, mosquitoes, phlebotomines, ticks
Orthoreovirus	Turreted	10(~23.5 Kbp)	5	Birds, reptiles, mammals (including humans)
Oryzavirus	Turreted	10(~26.1 Kbp)	2	Plants (Graminae), insects (Delphacidae planthoppers)
Phytoreovirus	Non-turreted	12(~25.1 Kbp)	3(+1)	Plants, insects (Cicadellidae, leafhoppers)
Rotavirus	Non-turreted	11(~18.5 Kbp)	5(+2)	Birds, mammals
Seadornavirus	Non-turreted	12(~21 Kbp)	3	Mammals (including humans), mosquitoes
Unassigned reoviruses (potentially a new genera)		10,11 or 12	5	

Table adapted from ref. (2)

Although, the homology in the amino acid sequence in proteins along the *Reoviridae* family is very low, the overall viral architecture is highly conserved Phylogenetic studies suggest that the conservation in the capsid of the *Reoviridae* is originated from an ancestor infecting a marine species (4). The ancestor virus subsequently evolved together with the newly colonised hosts by a co-speciation mechanism which involves changes in the genomic sequences such as deletions, duplication and gene rearrangements.. It would explain why the inner layers which confine the transcription complexes of the virus have overall a conserved

architecture while the outer layers that determine cell receptor and host specificity vary largely along the family (4,5).

RV is the most medically important virus in the family and as a result it has been well studied at the structural molecular level. It causes severe diarrhoea in infants and is estimated to be responsible for 420,000-490,000 deaths in children younger than five years worldwide. RV counts for 37 % of the deaths caused by diarrhoea and 5% of the total mortality occurring in this age group (6).

All the main aspects of RV biology have been reviewed in ref. (7). Here, I summarise only those relevant to the understanding of the structure and function of the virus. RV was identified morphologically in 1973 by EM in a tissue section of duodenal mucosa from paediatric patients (8), the virus resembles a wheel-like shaped particle (rota meaning 'wheel' in latin). This initial discovery was later confirmed leading to rotavirus being established as the major causative agent of diarrhoea in children (9).

Other morphological features have been used to classify the rotavirus (7) as an independent genus such as: an icosahedral structure of 100 nm in diameter in the virion (mature form of the virus) which bears 60 spikes or protein projections around the surface, these projections are stabilised by calcium binding sites.

The capsid consists of three functional shells of proteins, the inner-layered, the middle-layered and the outer-layered particle, also called the single, double and triple layered-particle (SLP, DLP and TLP) respectively. Within the space formed by the protein shell, there are transcription complexes containing a RNA dependent RNA polymerase (RdRP), and 11 segments.

The infection of the TLP in vitro requires a treatment of the TLP particle with trypsin-like proteases. The replication of the virus takes places in the cytoplasm.

Other viruses of the same group can exchange genome segments in the infected cell in a process called reassortment (10). Although RVs are not enveloped viruses, they are transiently enveloped by the endoplasmic reticulum (ER) membrane during the process of assembly. After full assembly, a RV particle is released by two mechanisms: the cell lysis or by using a non-classical transport in the polarised enterocyte cell (7).

There are five recognised groups (2) (A-E) in RV and three proposed independent groups (F_H) (11) based on serological response of the protein forming the DLP (7) Group A includes viruses that infect children, young animals and bird species. Group A includes the main cause of gastroenteritis in children and is identified by distinct serotype response that blocks surface proteins of the TLP (VP7 and VP4). There are up to 15 different G (glycoprotein) serotypes to VP7. The complex processing of VP4 (also called P or protease-sensitive protein) has limited the use of the P serotype and instead the use of 26 different P genotypes is widely used as a classification criterium (12).

2.2. Organisation of the capsid in rotavirus

The *Reoviridae* has three functional viral structures (Fig.2.1): the SLP that consists of the inner-layered particle and is thought to be responsible for replication; the DLP which contains an additional middle-layer and is important in transcription, and the TLP which has an outer-layered shell that mediates cell attachment.

The proteins forming the structure of RV share amino acid homology with other well studied members of *Reoviridae* such as ORV and BTV (Table 2.2). I will focus briefly on the description of the DLP and the TLP for which the function is better

understood and I will summarise the current ideas about the role of the SLP during the life cycle of the virus. (Fig.2.1)

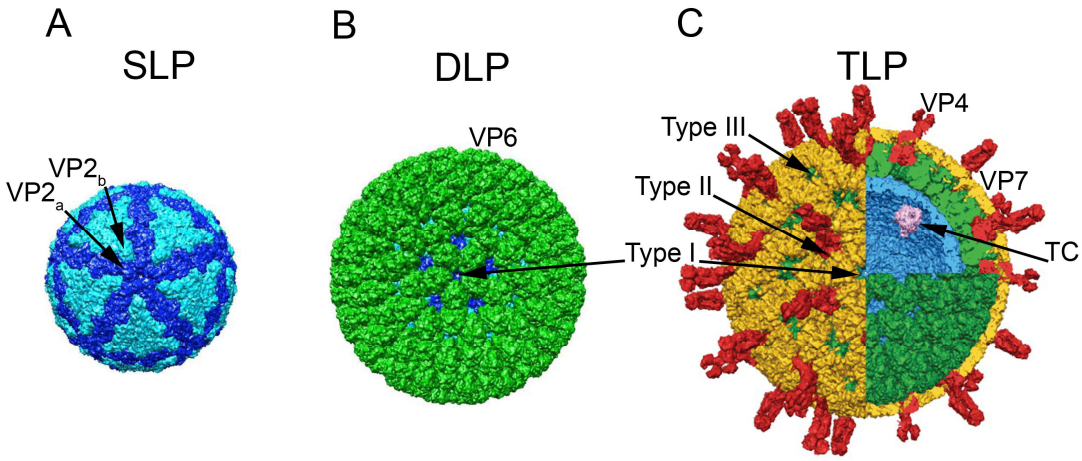


Fig. 2.1. Architecture of rotavirus capsid. *A)* The single-layered particle (SLP), ~50nm in diameter consists of a shell of VP2 protein in two different conformations (a and b). *B)* The double-layered particle (DLP), ~70 nm in diameter contains an additional shell of protein called VP6. *C*) The triple-layered particle (TLP), ~100 nm in diameter is the mature virion and is formed by proteins VP4 and VP7. Inside the capsid, it is hypothesised that there are 12 transcription complexes (TC) formed by VP1 and VP3 proteins and 11 dsRNA segments tightly packaged (~400 mg/ml molecular , mass ~10.0 × 10^6 Da) . Each dsRNA segment may be associated with the TC. Three aqueous channels (named I, II and III) connect the outside with the internal space of the virus and become active after the release of VP4 and VP7 in the infected cell. Adapted from ref. (13,14,15).

Initially, a combination of isolation with EM studies enabled the identification of three types of particles inside the cell, the triple-layered particle (TLP) that corresponds to the virion, consisting of an external layer of two types of protein, (VP4 and VP7) the middle layer and the inner layer (Fig 2.1); the double-layered particle (DLP), a viral particle that consists of a middle layer of protein and the inner layer particle (DLPs are transcriptionally active because they lack the outer layer), and the single–layered particle (SLP) which does not contain the genome inside (16, 17).

The virion consists of VP4 and VP7. There are 260 trimers of VP7 located at all threefold axes of a T=13 icosahedral shell. There are 132 channels at the five fold

and quasi sixfold axis of the shell. There are three types of channels (I, II and III) that connect the internal space of the capsid with outside (18). There are 12 type I channels located in the five fold axis of the icosahedral capsid, 60 type II channels located at the "six coordinated positions surrounding the five fold axis" (19); and 60 type III channels that form an array of six coordinated positions surrounding the three fold axis of the icosahedral lattice (19, 17).

Table 2. 2. RNA segments and proteins of rotavirus

Segment (Gene bank accession number)	Length (bp)	ORF	Protein	Copy (n_0)	Function	Homologous protein
1 (X16830)	3302	18-3282 (125 kDa)	VP1	12	RdRNApol	ORV13; BTV VP1
2 (X16831)	2690	17-2659 (102.4 kDa)	VP2	120	SLP	ORV11; BTV VP3
3 (X16062)	2591	50-2554 (98.1kDa)	VP3	12	Capping enzyme	ORV12; BTV VP4
4 (D16346, X14204)	2362	10 -2337 (86.7 kDa)	VP4	120 ?	TLP	
		(60 kDa)	VP5*			
		(28 kDa)	VP8*			
5 (L18944, X14914)	1611	31-1515 (58.6 kDa)	NSP1	0	associated to cytoskeleton	
6 (L15384, L33365, M27824)	1356	24-1214 (48.1 kDa)	VP6	780	DLP, structural	BTV VP7
7 (M87502)	1105	26-970 (34.6 kDa)	NSP3	0	Binds ElF4G1	
8 (L04531)	1059	47-997 (36.7 kDa)	NSP2	0	viroplasm (helix destabilization)	ORV sN2; BTV NS2
9 (K02028)	1062	49-1026 (73.6 kDa)	VP7	780	TLP,	
10 (AFO87678)	751	41-569 (20.9 kDa)	NSP4	0	ER protein	
11 (X07831, M28347)	667	22-615 (21.7 kDa)	NSP5	0	viroplasm	
		80-355 (11.1 kDa) (out of frame product)	NSP6	0	viroplasm	

Sequences from rotavirus SA11; SLP single-layered particle; DLP double-layered particle; TLP tripled-layered particle; * VP4 cleaved products. Adapted from ref. (20)

The dimensions in the DLP of type III channels are ~14 nm in depth and ~5.5 nm in the external diameter of the channel. Type I and II also have similar dimensions

with type II having a slightly narrower diameter of 4 nm at the outer side of the channel (17). The diameter of the type I channel is 1.8 nm (21) at its narrowest.

The function of the channels is to connect the internal space that contains the transcription complexes (TC) and a tightly packaged genome with the external environment.

This structural design is found in the *Reoviridae* and all the dsRNA viruses and confers several advantages: First the dsRNA is located with the viral molecular machinery for transcription-replication since the host cell replication machinery only recognises DNA templates, second, the double protein shell protects the free dsRNA from the defence mechanism in the host cell that degrades exogenous nucleic acids and triggers a strong apoptotic response. The degradation pathways have been partially described in some members of the *Reoviridae* family including RV and ORV. These pathways are triggered by cell viral sensors RIG-I and MDA5 proteins. Specifically, RIG-I detects uncapped ssRNAs, such as uncapped 5'-triphosphorylated (5' ppp) ssRNA, short dsRNAs but also 3'-monophosphorylated RNA. MDA5 has been shown to bind long dsRNA ~ 2kb and oligomer RNA structures (22). A precise description of the formation and function of these channels has started to be unveiled using structural molecular biology. EM studies in BTV and RV have identified that each channel displays a specific role. Cryo-EM studies in BTV have shown the entry and exit channels for NTPs, cations and by-products in the channels surrounding the five fold axis of the DLP (23), and have also confirmed observations in RV of the exit of RNA transcripts (24) through type I channels in the five fold axis of the DLP (23).

Although the confinement of the TCs in the capsid confers an advantage for controlling transcription, this strategy poses many questions that have been raised previously by others (25) and are still not well understood. These questions include:

a) How is the synthesis of dsRNA orchestrated within the core?
b) What is the mechanism by which this segmented genome is assorted and packaged selectively in the core?
c) Is the genome packaging mediated by RNA-RNA, RNA-protein interactions or both?
d) What is the order of interactions that lead to the packaging of the segments?
e) Is there any spatial organisation of the dsRNA segments and the proteins in the confined space of the capsid?
f) Does the RdRP structure change to distinguish between synthesis of (+)ssRNA transcription and (-)ssRNA
g) Where does the synthesis of the new (-)RNA take place to form a complete dsRNA?
h) How does the virus extrude nascent single stranded RNA, are there specialised channels within the capsid to release specific RNAs?

With the application of structural molecular techniques, the capsid architecture of some of the members of the family is starting to be unveiled. In particular, X-ray crystallography and single particle electron microscopy reconstruction techniques have led to a precise description on how the channels exchange precursors and by-products of transcription (21, 26).

2.2.1. DLP structure

In rotavirus, the DLP is the transcriptional unit in the cytoplasm that synthesises and releases the (+)ssRNA through type I channels (24). The DLP (Fig. 2.1*B*) consists of the inner-layer formed by VP2 (Fig.2.2*A*) and an outside layer of VP6 proteins. The VP2 (120 kDa) has two main domains: the N-terminal domain (~1 to 100 Aa) and the principal domain (~100 to 800 consisting of the central and dimerisation domain) (Fig.2.2*B*). The amino-terminal domain of VP2 has not been fully resolved. Nevertheless, the analysis of an X-ray based structure from bovine rotavirus double-layered particles (DLPs) (26) showed that the VP2 N-terminal regions form a hub-like cylindrical structure below each 5-fold axis. This 5-fold hub consists of the first ~100 residues of 10 VP2 monomers.

The VP2 adopts two conformations (A and B), that form 60 dimers which are organised in 12 decamers (Fig.2.2*C* and *D*) with the hub (formed by the N-terminal domains) in the centre around the five fold axis of the icosahedral capsid (26). Several roles are attributed to the hub, it binds the RdRP, VP1, the capping enzyme VP3, and coils the dsRNA segments (26). The VP1 is also thought to be associated with a RNA capping enzyme (VP3, 88 kDa) (27, 28). The main domain of VP2 forming the inner side of the capsid is also important in binding to VP1 and activating replication (29).

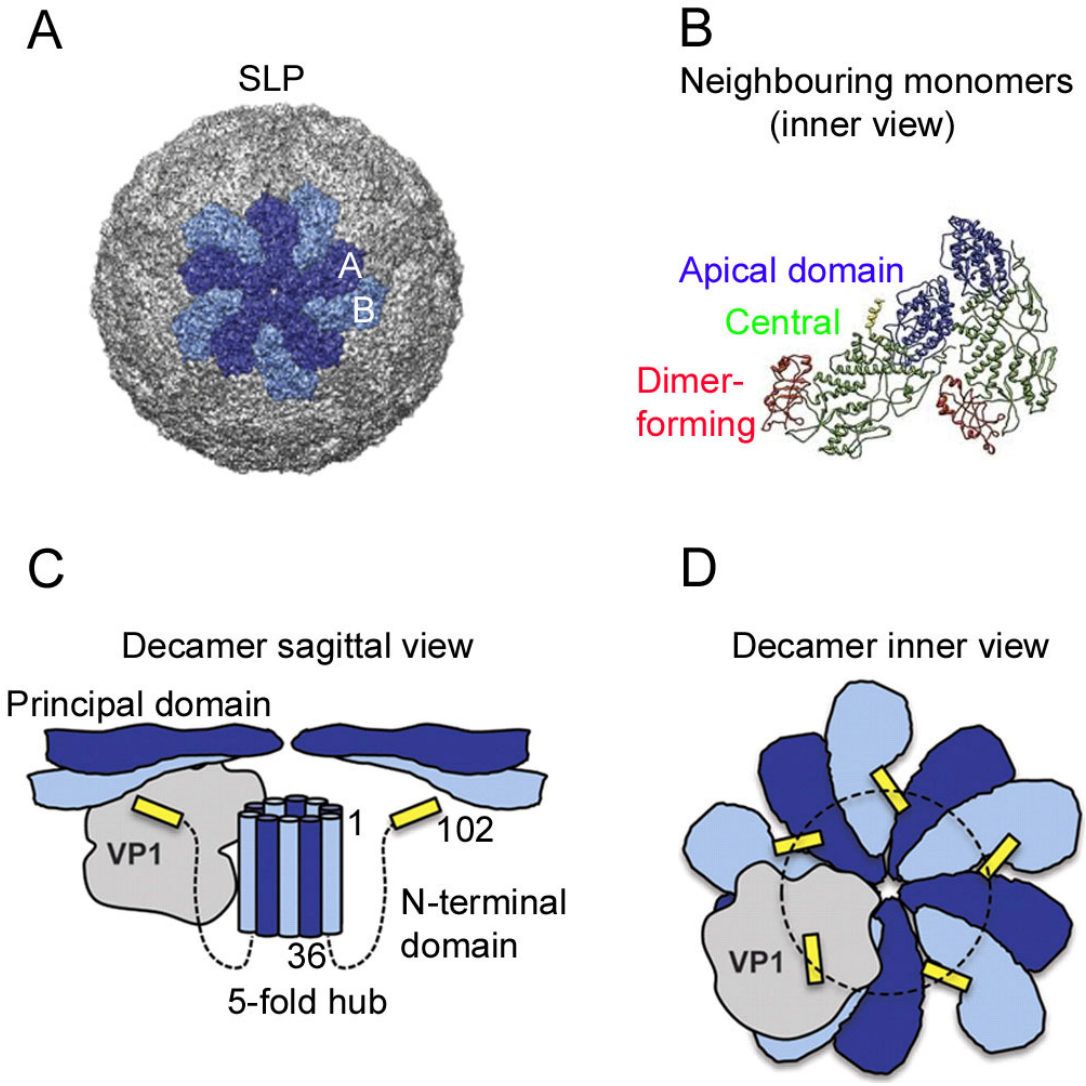


Fig. 2.2. Structure of the VP2 in the SLP and its interaction with VP1. *A*) VP2 adopts two different conformations called A (dark blue) and B (light blue) that form dimers which assemble in decamers in the single–layered particle (SLP). *B*) VP2 is a 120 kDa protein with two domains; the N-terminal domain (amino acids 1-102, the resolved structure from 81 to 100 Aa is shown in yellow) and a large domain, called the principal domain (103-882 Aa) formed by three distinctive subdomains called: the apical, central and the dimerisation domain. *C*) Cartoon model showing the interaction of VP2 with VP1 (sagittal view). The N-terminal domains of a decameric VP2 form a five-fold hub that interacts with the RV RNA dependent RNA polymerase, VP1. The interaction of the VP2 principal domain with the VP1, activates the replication of dsRNA in the SLP. *D*) Cartoon model (inner view) of the interaction between VP2 decamer and VP1. Adapted from ref. (29)

Inside the shell formed by the layer of VP2 proteins, there are 11 dsRNA segments (660-3300 bp), each segment is linked to the VP1 through a 5′ 7-MeG cap modification of the 5'end of each (+)ssRNA (30). In addition to this, the dsRNA

segments are in nearly crystalline form, which in addition to the small internal volume of the DLP 5.8×10^6 nm^3 limits their free diffusion within the capsid (26).

The current knowledge of the structural organization of the DLP leaves unexplained the question on whether there is one axis empty or one polymerase unused, and if so what the function of this empty channel is.

The external side of the VP2 layer is covered by a shell of VP6 (45 KDa) proteins which are organised forming 260 trimers. The VP6 proteins are organised in a five fold axis in register with the decamer VP2 axis resulting in the type I channels that extrude the newly synthesised transcripts (21, 26).

The loss of the VP6 shell abrogates the transcription as shown by removal of the VP6 layer and the observation that SLPs only replicate the dsRNA but are transcriptionally inactive (7).

2.2.2. The TLP

The TLP is the virion, the mature form of the RV; the TLP has an additional outer layered-protein formed by VP4 (~88 kDa) and VP7 (~37 kDa) that mediates the attachment and invasion of the host cell.

VP4 forms 60 trimeric spikes (depending on the strain) onto the type II channels (31) distributed surrounding pentamers of VP6 proteins. The position of VP4 is stabilised after the addition of a layer of plates formed by trimers of VP7 during the assembly (14, 32).

Intact, VP4 is a precursor protein responsible for the infection of the rotavirus particle. In order to became infective, VP4 has to be processed by trypsin (priming), and undergoes complex conformational changes associated with processing (33) which cleaves VP4 in two proteins VP8 (27 KDa) and VP5 (60 KDa) which are

covalently bound (31, 34). VP8 corresponds to the N-terminal domain of the intact VP4, and binds glycans and sialic acid receptors on the surface of the infected cell (35). VP5 (60 KDa) forms a trimer with a symmetry that is trimeric in its base and dimeric in the stalk of the structure (34). The base of the spike is located in the centre of the six fold symmetry formed by the layer of VP6 trimers in the DLP and the position restricted by the shell of VP7 proteins. Two of the VP5 units are folded away from the surface and bind with two subunits of VP8. The remaining VP5 is folded flat on the surface and lacks VP8 binding subunits (34).

VP7 (34 kDa) forms trimers and bears calcium binding domains that regulate the stability or the release of the VP7 shell structure (36, 37). The outer layer of the TLP is lost after the internalization of the virus in the endosome. The release of the DLP from the endosome into the cytoplasm is controlled by a depletion of calcium levels in the endosome which triggers the dissociation of the outer-layered proteins from the virion, and the perforation and rupture by VP4 of the endosome with the subsequent release of the DLP into the cytosol (38).

2.3. The genome organisation

The genomes of rotaviruses A have been sequenced. The length of the 11 dsRNA segments varies widely (39) and the dsRNA is highly folded within the capsid (15). Although the genome is tightly packed, the capsid can package up to an additional 2000 bp in a process called reassortment, that leads to changes in the segment size and shows that the capsid still has empty space to store a significant amount of RNA genome (40).

The base composition consists of 61% of A+T bases indicating that there is a codon biased structure of the segments (41). The analysis of the (+)ssRNA strand also

shows conserved features. The untranslated region (UTR) is initiated at the 5' with a G followed by a short consensus sequence 5'-GGC(A/U)-3'(30). The 5'end is capped with a sequence $m^7Gppp^{(m)}GPy$ (42).

The UTR is followed downstream by a starting codon and the open reading frame (ORF) coding for a protein which is followed by a stop codon at the end. In segments 7,9,10 there are additional starting codons in the same phase; in segment 11 there is an additional out of phase ORF (43). The segments finish with a 3'UTR without a polyadenylation signal but with a consensus sequence 5'-UGUGACC-3' with the sequence UGUG cassette working as a minimal promoter (30) in RV (A,C,D,F) but not in RV(B,G,H) (44).

2.4. Overview of the life cycle

Rotavirus infects in the small intestine, and it infects only mature enterocytes which are fully functional cells which have lost the capability to divide. Nevertheless, the establishment of cell culture to study RV infection shows that RV can infect a variety of cells suggesting the presence of common receptors spread along many cell types. The replication time of RV varies depending on the cell line and the temperature used (from 12 hours at 37°C (45) in spread cells to 24 hours in polarised cell lines (46)).

The entry of RV in the cell starts by attachment to the apical and basolateral membrane in polarised cells. The mechanism of exit from the cell is caused by cell lysis but also by a mechanism of transport that is independent from the Golgi (46,7).

Fig. 2.3

The entry of the cell by RV has been extensively reviewed (47), this process is mediated by the VP4 and VP7 proteins on the surface of the TLP. Some of the events leading to the translocation of the virus have been described (47) However, due to the absence of a good cell model the process of attachment, the way that RV is internalised and its release into the cytoplasm remains poorly understood.

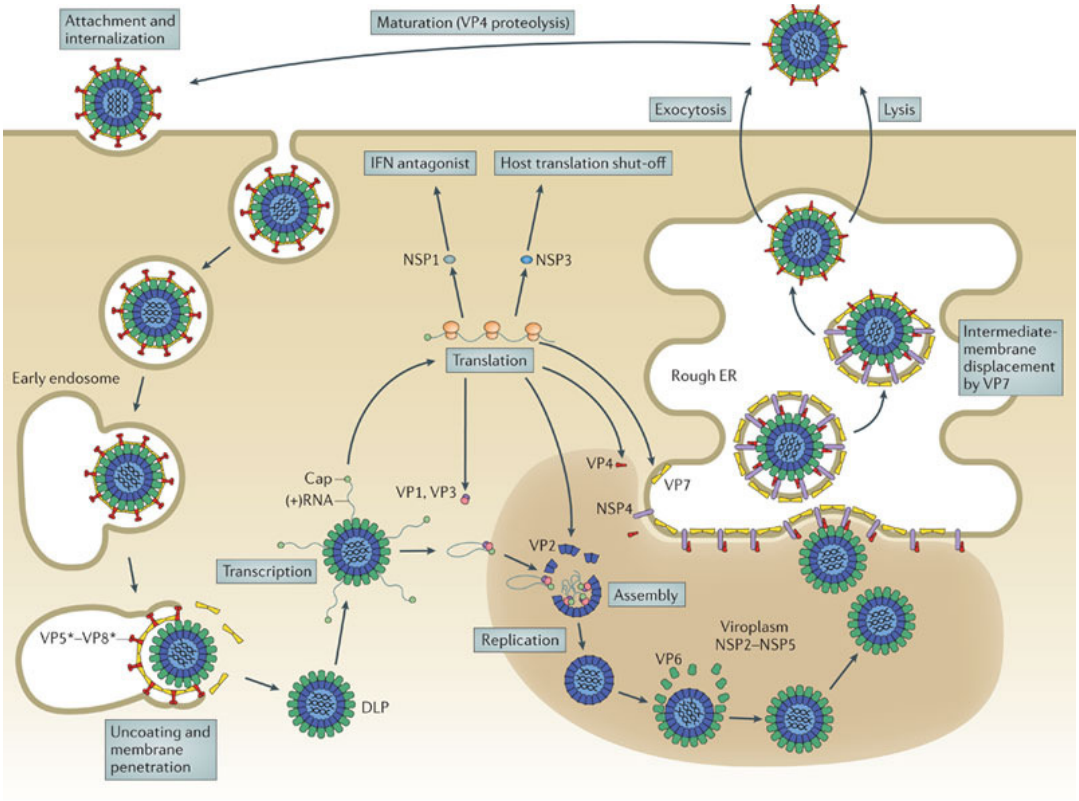


Fig. 2.3. Rotavirus replication cycle. RV attached to the tip of the mature enterocyte. After that, RV goes into the cell by endocytosis. A decrease of calcium causes the loss of VP5 and VP8, and the release of DLP into the cytoplasm. DLP becomes transcriptionally active, and performs a dual function, some (+)ssRNAs are translated and others are packaged and assembled in a new virus. Packaging and assembly starts in the viroplasm, a structure produced by the viral proteins and continues in the ER. The newly formed capsid is released by lysis or an exocytosis mechanism. Cartoon taken from ref. (48).

Several crucial steps have been described (47). First, the processing of VP4 by trypsin-like proteins (priming), which lead to a conformational change of VP4, which may be important to expose regions important for attachment (Fig. 2.4);

second, attachment of RV to what is possibly a variety of cell receptors; third, RV internalization takes place by a mechanism that is dynamin dependent; fourth, RV loses its outer-layered shell of VP4 and VP7 of the TLP triggered by VP4 that undergoes a second conformational change that exposes a fusion membrane perforation domain that leads to the release of the DLP into the cytoplasm.

The VP4 processing by trypsin-like proteases found in the intestinal lumen confers rigidity and stabilises the protein conformation (33, 34) that may be important for a sequential recognition of receptors in the cell.

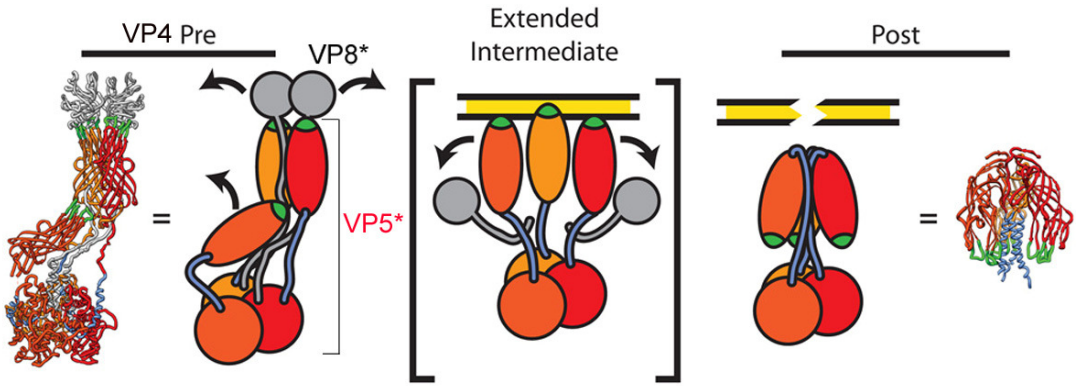


Fig. 2.4. Processing of VP4 and conformational changes associated with membrane penetration. VP4 proteins are processed by trypsin-like proteins resulting in a trimer of VP5 and a dimer of VP8 that remain associated. This processed VP5/VP8 structure binds the cell receptor; after that the host receptor binding domains are dissociated and hydrophobic peptide domains bind the membrane and fold back resulting in perforation of the membrane. Adapted from ref (48).

Processing occurs after arginines R247, R241 and makes RV competent to infect the cell host (49). The two processed fragments remain associated after trypsin processing. VP5 (amino acids 248-776) has functional homology to 'fusion membrane peptides' (50) and VP8 (amino acids 1-231) bears a hemagglutinin motif

with a galectin-like fold called the VP8 core (amino acids 65-224) motif that binds sialic acid (51).

In addition to trypsin like enzymes, VP4 is also affected by pH treatment which regulates the binding to SA receptors (52). Once the virus is within the cytoplasm the virion loses the outer capsid and becomes transcriptionally active (53).

The attachment to the cell by RV is characterised by several features: the binding to multiple binding receptors, their wide heterogeneity and the low affinity of the interaction. The binding to the cell host is carried out by VP5, VP8 and VP7. The first cell receptor identified was the hemaglutinin, sialic acid (SA, N-acetyl neuraminic acid) that is bound by VP8. The binding affinity is very low (Kd=1.2 mM) (54). The VP8 has the same folding as a galectin-like domain (35).

The exact localisation of SA post-translational modifications has not been assigned to a protein or a lipid, and may vary within the cell lines studied. Some experiments showed that it was allocated to the O-linked oligosaccharide of proteins distributed on specific tissues for RV infection (55). Alternatively, in other experiments the importance of gangliosides as sialic modified receptors for RV attachment was shown. (56). Due to the localization of the SA, the RV attachment to this receptor can determine the entry of RV through the apical side of polarised cells (57).

The importance of SA as the chief receptor for RV entry has been discussed but remains controversial (7). It is argued (7) that most of the strains of RV (including the majority of human rotavirus) are sialic binding independent (58), and that the binding to the SA receptor is unimportant in human infectivity, since the affinity of some Rhesus RV (RRV) strains for N-acetyl neuraminic acid (which is synthesised in human cells) is higher than the affinity for animal N-glycolyl neuraminic acid (54).

In addition to SA, it has been proposed that the outer layer of RV may use hetero-dimeric transmembrane receptors called integrins which play roles in attachment and cell signalling. VP5 is proposed to bind α2β1, α4β7 integrins (59);

The best characterised interaction of VP5 with an integrin is the binding of VP5 to α2β1. Solid phase assays showed that the interaction occurs through a DGE motif (60). However, its implication in RV attachment remains also controversial because α2β1 has a basolateral localization in the human crypt cells rather than the villi of the mature enterocyte, the true target for RV attachment (61).

The protein VP7 may also be responsible for binding integrins αxβ2,α4β1, and αvβ3 (60, 62, 63,64). In addition to integrins, the heat shock cognate protein (HSCP 70) (65) has also been described to be a ligand for RV. The localization of the HSCP70 is mainly intracellular suggesting a role in trafficking in the cytoplasm.

The mechanism that RV uses to enter the host cell is still a subject of investigation. Based on independent experiments, several models have been proposed. One model suggests that it is an active mechanism in which RV disrupts the cell membrane to gain access into the cytoplasm. A second model suggests that the virus is taken passively by endocytosis. A third model that reconciles these conflicting views is also possible (47).

RV entry varies in time. In primed RV (treated with trypsin-like proteases), it occurs very quickly in ~3-5 min but in untreated RV it takes ~50 min (66) which may suggest a passive mechanism of invasion.

One of the effects of priming in RV is the disruption of the cell membrane as is shown in experiments with cells pre loaded with chromium which released their contents after RV invasion (66). However, in addition to or as a part of the mechanism of entry, other experiments suggest that an endocytosis-based

mechanism may also be important, although the data is contradictory. Blocking of endocytosis mediated by dynamin blocked RV infection, pointing towards a mechanism of endocytosis during RV entry. However, this mechanism consumes energy and therefore the data was conflicting with experiments that has shown that ATP inhibitors such as sodium azide did not affect the entry using this pathway (47). Other endocytic pathways mediated by clathrin or caveolin are not used in RV entry. Sodium azide and cytochalasin that inhibits chlatrin mediated endocytosis and specific inhibitors of caveolin mediated endocytosis did not affect infectivity (67). After internalisation of the TLP, uncoating follows**.** The VP5 region (amino acids 385-404) causes small holes < 4KDa (68) that permeabilise the membrane, causing a depletion of calcium concentration (68). The loss of calcium dissociates VP7 trimers (69) which results in DLP release and disruption of the membrane (68). Other factors that are important in internalisation were also investigated. VP5 has exhibited lipophilic activity that is also important in cell entry as it is temperature dependent and entry is inhibited between 0-4°C (70). However, intra-endosomal acidic pH (71) is not important in the process by which the RV is released in the cell cytoplasm.

2.5. Structural study of transcription and replication of rotavirus

The synthesis of (+)ssRNA has been detected and visualised in DLPs of RV and other members of the *Reoviridae* using native EM and biochemical assays. Structural and biochemical analysis of the RdRP, VP1 in RV (30), together with the homologous proteins λ3 in ORV (72), the RdRP of bacteriophage ϕ6 (73) and the VP1 of BTV (74) have started to unveil the dual mechanism of transcription and replication by the RdRP in the *Reoviridae* (75). The VP1 (Fig.2.5) in RV binds the

N-terminus of the VP2 protein (29), and is located in close contact to the channel at the five fold axis (13). VP2 terminus is rich in positively charged amino acids that bind the VP1 and RNA template, suggesting that the *Reoviridae* could have a general mechanism for the assembly of the RdRP in the capsid and it is important to induce a conformational change that activates replication and the synthesis of the (-)ssRNA (30).

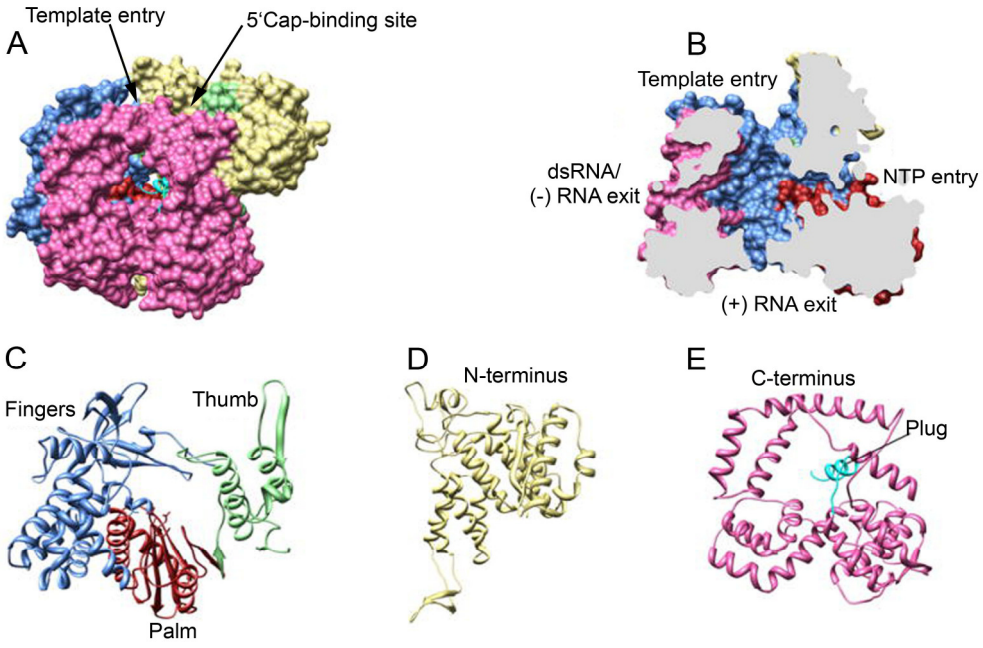


Fig. 2.5. Structure of VP1. *A*) Surface rendering. *B)* Section of image *A* rotated 90º to the left shows four channels and the central cavity. *C*) Ribbon diagram of the polymerase catalytic domains, fingers, (blue), palm (red) and thumb (green). *D)* Ribbon diagram of the VP1 N-terminus (yellow). *E)* Ribbon diagram of the C-terminal domain (purple) with a plug domain (light blue) which is hypothesised to control the RNA synthesis of VP1. Adapted from ref. (75, 30).

VP1 binds a consensus sequence (CS) 5'-UGUGACC-3' in the 3' end of the (+)ssRNA. The CS is conserved in the RV segments and works as a minimal promoter unit. The motif 5'-UGUG-3' is important for the interaction of the (+)ssRNA and the VP1; the downstream 5'-ACC-3' works as a recognition sequence for the synthesis of the (-)ssRNA strand and the formation of a complete

dsRNA segment (76). There is also a binding site for the capped 5'end of the (+)ssRNA strand that may be important in maintaining the association of the template to the VP1 and ensuring multiple rounds of synthesis (30).

The structure of the VP1 (125 kDa) is 7 nm in diameter and has conserved N and C domains (1-332) and (779-1089), respectively that form a cage that shapes the polymerase domain formed by the fingers (333-488, 524-595), palm (489-523, and 596-685) and thumb subdomains (686-778) (30). The N-terminal domain is located over the active site of the VP1 and keeps a closed conformation of the fingers and the thumb. The C-terminal domain closes the dsRNA exit tunnel (30).

VP1 is catalytically inactive, it only becomes active after the VP2 protein binds VP1. The interaction between VP1 and VP2 occurs during the process of assembly or uncoating of the virion and results in a conformational change that leads to dsRNA replication and synthesis of (+)ssRNA respectively (77, 78).

VP1 has four channels. One allocates the template (dsRNA during transcription or ssRNA in replication), a second provides the nucleotides during the synthesis, the third releases the template and the fourth releases the new (+)ssRNA. (30, 75).

In addition to synthesis of (+)ssRNA the virus modifies the transcript with a capping enzyme. In RV, VP3 is the capping enzyme (27) and functional studies show guanylation, N7-methylation and 2'O methylation.(30).

VP3 lacks triphosphatase activity that is present in other homologous capping enzymes such as VP4 from BTV (27). The location of the capping enzyme VP3 has not been formally demonstrated within the capsid (75).

EM studies using VP2 mutants lacking the N-terminus show the loss of this region abrogates the interaction of VP1/VP3 (79). In ORV, the capping enzyme is located on the outside of the capsid forming a turret like structure (hence turreted virus). In

BTV the capping enzyme, VP4, is located in close proximity with the polymerase VP1 within the capsid space.

The working model for the location of VP3 in RV, proposes that in order for VP3 to perform its catalytic activity, it must be located inside the capsid where VP1 is located and in close proximity to the exit channels at the five-fold axis. The close proximity of both enzymes is supported indirectly by data that showed that capping occurs after a short synthesis of 5-7 nucleotides in the capsid and the omission of capping substrate SAM (S-adenosyl methionine) blocks polymerase activity (28).

2.6. Translation

Translation occurs in the ribosomes and is mediated by NSP3, a non-structural RNA binding protein which has functional homology with the eukaryotic poly (A) binding protein (PABP). NSP3 forms dimers (80), the N-terminus binds the 3'consensus sequence of (+)ssRNAs, the C-terminus binds elF4G that participates in translation; NSP3 also binds protein RoXaN which may be also important in translation (81-83).

NSP3 displaces the binding of PABP to elF4G resulting in an enhancement of viral protein synthesis (82, 83). In addition to this direct role in translation, NSP3 binds the cytoskeleton and it is hypothesised to bring (+)ssRNAs as a cargo to the viroplasm where replication occurs and therefore regulates the dual role of the (+)ssRNA in translation and replication, processes that are compartmentalised in the cell (84). After translation the newly synthesised proteins are released in the cytoplasm with the exception of VP7 and VP4 that have signal peptide like sequences that target the ER, in the case of VP7 the N-terminal signal is further removed (84).

2.7. Packaging assembly and virus release

The assembly of the virus follows a sequential process that occurs in three distinct localisations. Protein localisation studies show that the assembly is initiated in the viroplasm where VP2, NSP2, NSP5, and NSP6 are found (85, 86). In the viroplasm the VP2 forms a single-layered viral particle that encloses the genome and the transcription complexes (VP1-VP3 protein complex). The next stage in the assembly is the addition of VP6. Experiments that look at the localisation of the VP6 (14) suggest that this step occurs outside the viroplasm and in close proximity to the ER. Once the formation of the DLP is completed, the DLP moves into the ER where NSP4, VP4 and VP7 proteins are found (32, 87) and where the TLP is fully assembled.

The contribution of each of these proteins has been investigated. VP4 appears to associate with microtubules (88) and also with lipid rafts. The association with lipid rafts (89) takes place early in the ER supporting the possibility that VP4 contributes to the transport of the virion from the ER to the cell membrane. Interestingly, VP4 is disposable for the assembly of VP7 and the abrogation of the protein with sRNAi does not affect the assembly of VP7 (14).

NSP4 is a transmembrane protein that targets the ER and has a well defined role in the assembly. The NSP4 is a tetrameric glycosylated protein that is processed to a mature form of 28 kDa (90). One of the functions of NSP4 is to bind VP6, bringing the DLP into the ER.

The relevance of this posttranslational modification is not well understood. Glycosylated NSP4 binds calnexin (91), an ER chaperone thought to be important in assembly; NSP4 exhibits other functions in the cell that are being unravelled. Inside the cell NSP4 mobilises calcium reservoirs (92) and it is hypothesised that

calcium release mediates the formation of heteroligomers with VP4 and VP7 suggesting a role of NSP4 in the assembly of the outer layer into the DLP and contributing to the function of VP7 in the removal of a transient membrane envelope in the DLP after the process of budding into the ER.

In addition to the localisation in the ER, NSP4 is found in the endoplasmic reticulum golgi intermediate compartment (ERGIC) and the cytoplasmic structures and includes an association with the autophagosomal protein LC3, the viroplasm and caveolin (14, 93) indicating that the protein may play a role in the release of the virus, the formation of the viroplasm and as a secreted protein, respectively.

Once the capsid is fully assembled, it is released from the cell. Electron microscopy studies have shown that the completion of the infectious cycle causes cell lysis and the release of viral proteins, DLPs and TLPs associated with the cytoskeleton (94). NSP4 association with lipid rafts also supports the notion that NSP4 may target the virion to the surface (14). However a detailed description of the mechanism by which the TLP is released is still lacking.

2.8. Model for packaging and assortment

RV and other segmented RNA viruses pack their genomic segments in a mature infectious virion that infects and replicates in a target cell. The strategy to package the genome inside the capsid differs among viruses and it is only partially understood.

On one hand, dsRNA viruses with one to two segments, such the *Birnaviridae,* pack the genome randomly (95) via a mechanism based on the recruitment of the VP1 polymerase and the associated genome by the inner-layered protein (96), resulting

in a large number of viral particles with polyploid genomes, heterogeneous type of segments and with a variable capability to infect cells.

On the other hand, viruses with three or more genomes such as *Cystoviridae* (3) or the *Reoviridae* family (10-12) including RV, pack selectively one single copy of each segment per capsid (97). This means that each segment that varies largely in length must be folded within the small volume of the capsid and implies that the RV segments must have a dynamic conformation in solution (39). In RV the folding of (+)ssRNAs takes place in conserved secondary structures such as long-range interactions at the 5' and 3' termini and stem-loop structures (39). Although RNAs are in several conformations in solution, the RNA folding is a highly organised process as shown in bacteriophage MS2 and the small plant virus satellite tobacco necrosis virus (STNV) where their capsid proteins can recognise and fold their specific genomes in the cytoplasm containing a pool of host mRNAs (98). Therefore, the capability to adopt different conformations may explain why the RNA transcripts bind many different ligands and take part in diverse functions of the virus such as transcription, replication and translation as well as the assortment and packaging of the virus capsid.

Studies in *Reoviridae* show that assortment and packaging relies on specific recognition sequences in the RNA transcripts which enables them to bind to the proteins in the capsid and possibly to other RNAs.

The molecular study of the proteins of RV and the detailed description of the life cycle of the virus has shown where the assortment takes place in the cell and which proteins participate in the process.

The recruitment of (+)ssRNAs takes place in the viroplasm. The viroplasm is a dynamic structure formed by building blocks of NSP2 and NSP5 that alters the

microtubule network and occupies a perinuclear localisation where assembly of the new virus takes place (99). Interestingly, the viroplasm contains replication intermediate (RI) cores which are not fully formed capsids, and consists of NSP2, NSP5, NSP6, VP1, VP3 and VP2. Treatment of RIs with RNAase resulted in RNA degradation suggesting that the RIs are not fully closed capsids (100). The RIs have packing and replicase activity. The main structural components of the viroplasm are NSP2 and NSP5. NSP2 binds with high affinity (+)ssRNA. In addition, NSP2 also binds capsid proteins VP1 and VP2. A detailed description of the sequential interaction between these proteins in the viroplasm is not available, although there are biochemical functional data of the individual proteins. The function of NSP2 has been studied in some detail. NSP2 is an octameric protein that bears a histidine-triad like domain that binds ssRNA; the NSP2 forms a protein complex with NSP5. The functions of NSP5 include NTPases, RNA triphosphatase (RTPase), kinase and unwinding activity in vitro (101). Interestingly, helix destabilising activity requires Mg2+ but it is ATP independent and non–directional suggesting a role in removing secondary structures during packaging (102). The hypothetical role of NSP2 as a helicase unwinding dsRNA template facilitating the polymerase activity of VP1 has never been shown. Direct comparison with similar proteins in the *Reoviridae* family is not possible since NSP2 does not share sequence homology with any protein of the family. Still, there are several proteins in the *Reoviridae* family with NTPase and helicase activity, examples of these proteins are λ1, λA, VP3 and VP6 found in mammalian orthoreovirus, avian reovirus, aquareoviruses and BTV respectively. In the case of λ1, it is well characterised, but it is unusual in that it has a dual role as a structural and enzymatic protein; VP6 is described as a typical helicase located within the core (103).

NSP2 has NTPase activity that hydrolyses the terminal phosphoanidride bond of any NTPs releasing NDP and free Pi (104). It has also NDP-kinase activity that produces a transient short-lived phosphorylated NSP2 that is rapidly dephosphorylated to form NTP in the presence of aceptor NDP (105). The RNTPase activity removes ɣ-phosphate in the 5'end of of the (-) ssRNA (105). The RNTPase activity is associated with conformational changes caused by nucleotide binding which could regulate the NSP2 interactions between NTP, RNA and proteins NSP5 and VP1.

Both the NTPase and RTPase activity can contribute to this short-lived autophosphorylation of NSP2. These complex and interrelated functions may be important for creating NTP reservoirs and regulating binding to multiple ligands.

Studies using RNAi show that the abrogation of NSP2 expression inhibits the synthesis of ssRNAs, protein, and dsRNAs (101). These findings are the basis to propose a model where the viroplasm traps a population of DLPs which synthesise the (+)ssRNAs which are recruited in the viroplasm as templates for dsRNAs and participate in the assembly of new capsids.

NSP5 is phosphorylated by NSP2 (101) and forms a complex with NSP2 (106), RNA (107) and VP1. The function of NSP5 is still unknown, but it is speculated that a phosphorylation state can regulate the binding of VP1 to VP2.

Therefore, a working model for RNA packaging shows where the packaging takes place and proposes that NSP2 in the presence of ATP participates in the assembly events taking place in the viroplasm. NSP2 can destabilise the secondary structure of the (+)ssRNA, making available the binding site of RNAs segments to allow the initial engagement to VP1 molecules and the change in affinity binding of VP1-NSP5 to VP1-VP2 interactions (101).

The interaction of RNA with VP1 could drive the initial incorporation of RNA segments in the VP2 shell since a site directed mutagenesis study (108) has shown that VP1 and VP3 bind to the hub-like structure formed by the VP2 five fold and this interaction mediates their encapsidation of proteins into the core (26).

The second question that remains unanswered is how only one copy of each segment is packaged in the viral particle. It is thought to be mediated by a specific and ordered interaction among the segments. This interaction relies on conserved genomic sequences rather than conservation in the codon usage, as shown by experiments in which the mixture of genomic sequences from different strains (reassortment) appears to be restricted (30, 109).

The packaging of exactly one copy of each segment may also depend on a unique secondary structure. The ssRNAs (39) that form panhandles and stem-loops with the 3'end are fixed in a conformation that is accessible for an interaction with the VP1 polymerase (97).

The signals for assortment and packaging are thought to be present in regions flanking the 5' and 3' end and not in the ORF as shown by the finding that:

a) strains lacking ORFs are still capable of assorting (109)

b) segments with duplications are detected in cell culture suggesting that the duplication of signals favours packaging (109)

c) studies in the related ORV for which a cDNA reverse genetic system is available also showed that the ORF is not needed in packaging. ORF was replaced with a sequence encoding a cloramphenicol acethytransferase as a reporter protein and was still able to produce an infectious progeny. In this study the 5'UTR sequences were shown to be important for packaging. The 5' UTR was found to bear packaging signals in three different segments. Modelling of the regions in segments m1, s2 and

11 showed that the packaging is associated to stem–loop formation in this area (110).

The localisation of the regions for packaging in RV and other *Reoviridae* are not yet defined. This set of evidence leads to the proposal of two packaging models (97) on how the assembly of the virion takes place (Fig. 2.6). The first model is called the concerted packaging model and is based on the influenza packaging model in which 8 (-)ssRNA segments form a genome complex with ribonucleoproteins (111) and the segments are selected by interaction between cis-acting sequences before the capsid assembly (112). The concerted model summarises the observations on where, how and what proteins interact with the (+)ssRNAs described above.

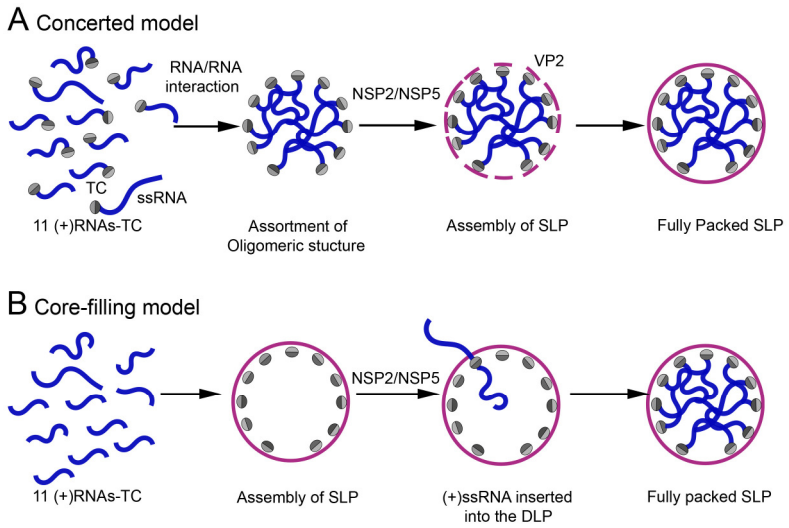


Fig. 2.6. Models of assortment and packaging of (+)RNA segments in RV. Adapted from ref. (97). *A*) Concerted model. Each (+)RNA (blue lines) form a complex with the transcription complex (TC). Assortment occurs by interaction between (+)ssRNA specific sequences. In the presence of NSP2/NSP5 the assembly of VP2 proteins (Purple line) takes place resulting in a single-layered particle (SLP). *B*) Core filling model. (+)ssRNAs are introduced into a fully assembled but empty SLP by the NSP2/NSP5 complexes.

In the RV packaging model, the (+)ssRNAs binds the TC using the interaction of the capped 5'end with the VP3 and the 3'end with the VP1. The proteins in the viroplasm NSP2 and NSP5 facilitate the assortment of exactly 11 segments with the

TC. Subsequently VP2 plates interact with VP1 and are assembled onto this RNA-protein structure forming a single-layered particle. This model is supported by RI data (100).

The association of VP2 induces conformational changes in VP1 that activate the replication of the genome (77). Only after the addition of VP2, dsRNA is detected. It is not degraded by RNAases and appears to be in a compacted conformation (113), suggesting that the replication must occur once the single-layered particle is formed.

An alternative model based on the packaging mechanism identified in bacteriophage ϕ6 (114) proposed that because there is a presence of empty capsids in the infected cells, they could be the container particles for ordered packaging of the ssRNAs.

In bacteriophage ϕ6, the segmented dsRNA virus has three segments (s, m, l), the packaging sequences are located in the 5'UTRs, in stem-loops, and RNAs are packed by a motor protein that introduces the segments s, l and m into an empty capsid (115). The packaging of the segments follows a step wise process. The packaging of a segment triggers a conformational change in the core, which reveals the specific recognition site to pack the next specific RNA segment (116). The studies of packaging of bacteriophage ϕ6 also showed that the size and the genomic composition of the genome is also important in determining the subsequent packaging of the next segment by triggering a conformational change (116).

The candidate motor protein responsible for inserting the segments would be NSP2 that has helicase and ATPase activity (107). However, there is very little supporting evidence for this model. There is no functional homology with the packaging motor ϕ6 P4, *in vitro* preformed single-layered particles cannot be filled with ssRNAs in

the presence of NSP2 and there are no descriptions in literature of partially packed capsids (97).

Chapter 3

General Materials and Methods

3.1. Basic concepts in single molecule FRET

In this thesis, I use a combination of single molecule fluorescence techniques that make use of Förster resonance energy transfer (FRET), total internal reflection fluorescence (TIRF) and alternating laser excitation (ALEX) microscopy. Several properties are unique to single molecule fluorescence techniques: they have high sensitivity so the sample concentration needed for the analysis is very low (in the range of nM to pM), these techniques enable the study of the behaviour of active individual molecules in real time, which leads to a characterization of the heterogeneity of the sample, a striking difference to the standard biochemical techniques that record the averaged behaviour of the population. The progressive implementation of single molecule techniques in the field of biology is rapidly advancing the field by revealing a detailed description of the cell structures and the intermediate steps in many biochemical processes.

Next, I will briefly present basic concepts underlying the fluorescence methods used in this thesis. A comprehensive description of the field of fluorescence can be found elsewhere (117).

Fluorescence is a form of emission of photons by a molecule in an electronically excited state caused by the absorption of light. One of the properties that characterises a fluorescent molecule is the phenomenon of the Stokes' shift by which a molecule is excited with a shorter (higher energy) wavelength which results into an excited singlet state, followed by the subsequent emission of a photon with a longer wavelength (lower energy) and the final return to the ground state. The life time in this process is very fast (10^{-8} s) (117).

The fluctuation in energy levels of a fluorescent molecule is described in a Jablonsky diagram (Fig.3.1). The energy levels of the fluorophores can be described

according to their singlet ground, first, second and higher excited states electronic levels (S_0, S_1, S_2). In singlet states the electrons are assembled in pairs with a spin S=0 and have different vibrational energy levels. If the fluorophore absorbs a photon with an energy $E = hv_{Abs}$ where h is the Plank's constant ($1.58\ 10^{-37}$ Kcal.s) and v is the frequency of the photon (s^{-1}= Hz), then the molecule is excited to the singlet state S1 or S2. After which, the molecule undergoes internal conversion or vibrational relaxation with the energy level returning to the lowest vibrational state of the singlet state S1 (life time, τ_r ~10^{-12} s). Here, the molecule releases a photon and returns to the ground state S_0 (117).

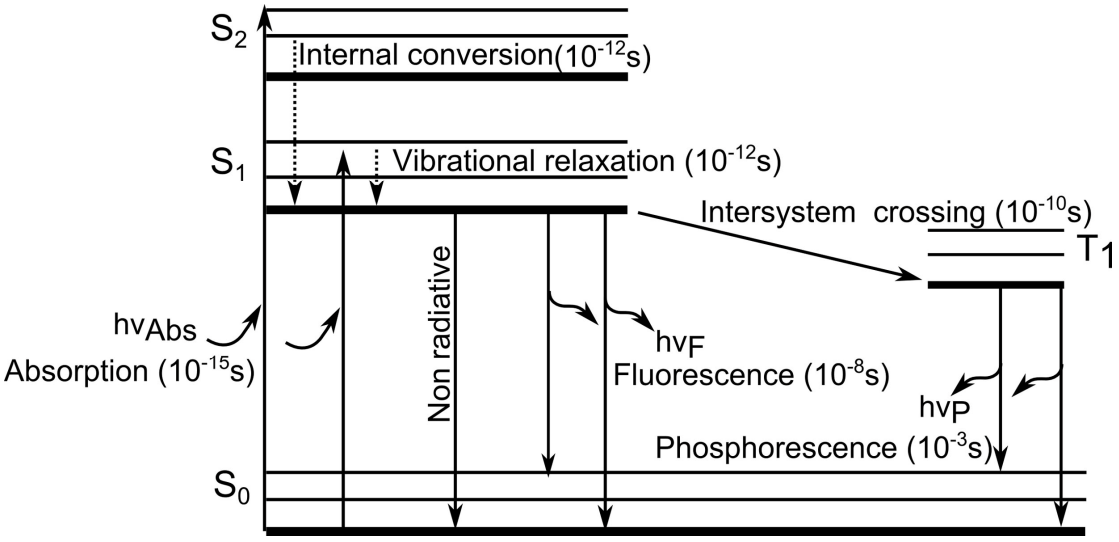


Fig. 3.1. Jablonsky diagram of the energy levels of a fluorophore. The Jablonsky diagram represents the changes in the electronic state of the fluorophore under different energy levels. A fluorophore in its ground energy level is represented by the singlet state S_0 with its different energy sublevels. After the absorption of photons the energy levels raise the fluorophore to a singlet state S1 or S2 in which the fluorophore will undergo internal conversion or vibrational relaxation to a lower energy level and very quickly return to the ground state releasing non radiative energy and spontaneous emission of photons (fluorescence). Alternatively, the fluorophore can undergo intersystem crossing to a triplet state from which it can return to the ground state (phosphorescence) in a very slow process or to the S1 after which return to the ground state (delayed fluorescence). The different transition rates (τ) from each state are shown in seconds (s). Adapted from ref. (117).

Alternatively, the molecule can undergo a phenomenon of intersystem crossing in which the S_1 has a spin conversion forming a triplet state. The phenomenon of vibrational relaxation from the triplet state to the ground state is called phosphorescence, it is very slow (τ_{rP} ~10^{-3} s) by comparison with fluorescence life and its transition to the ground state takes a longer time (117).

The formation of a triplet state has practical consequences since it affects the photophysics of the molecule producing a transient dark state of the fluorophore, known as blinking. In addition, the repeated excitation of the S_0 to a new S_1 and the effect of the buffer results in photo-induced chemical changes in the fluorophore that produce either transient or permanent bleaching of the fluorescence.

Therefore, the control of the photophysics of the fluorophore caused by blinking and the photobleaching determine the recording of consistent measurements. Both phenomena blinking and photobleaching are affected by the oxygen present in the buffer solution. The photobleaching of the fluorophores is increased by the effect of photoxidation (118). To increase the time and stability of the measurements, the imaging buffers include enzymes that remove oxygen from the solution (117).

In this thesis, the oxygen scavenger system used is Gloxy (119) that combines enzymes glucose oxydase, catalase and glucose. However, because the effect of oxygen is dual and is also shown to be very important in minimizing the amount of triplet states of the fluorophore; imaging buffers contain an alternative triplet state quencher that compensates the depletion of oxygen.

Triplet state quenchers used in single molecule include reducing agents such as β-mercaptoethanol or the antioxidant analog 6-hydroxyl-2,5,7,8-tetramethychroma-2-carboxylic acid (Trolox) (120, 121) which was used in this research. In order to be active, trolox must be solubilised and treated with UV that partially oxidises trolox

into trolox-quinone. The presence of the oxido-reducing component blocks the formation of triplet states of the fluorophore (120).

The fluorescence emission also depends on specific properties of each fluorophore (117,122) such as the quantum yield Φ and the fluorescence lifetime τ_f. Φ provides information about the brightness of a fluorophore by showing the efficiency of the fluorescence process, it is defined as the ratio of photons emitted out of the total amount of photons absorbed by a fluorophore. Φ can be quantified as the ratio between the radiative rate of the depopulation of S1 and the total depopulation rate (the sum of the radiative and non-radiative rate).

$$\Phi = \frac{\kappa_r}{\kappa_r + \kappa_{nr}}$$

Fluorophores with quantum yields ~1 indicates that the mechanism of depopulation of the excited state is mainly by photon emission. The fluorescence lifetime τ_f is the average time that the molecule stays in the singlet state S1 and is described by the following expression:

$$\tau_f = \frac{1}{\kappa_r + \kappa_{nr}}$$

τ_f is a characteristic value of each fluorophore and is affected by the variations caused by the environment in the depopulation rate constants of S1 depending on the radiative and non-radiative process. The value of τ_f, can be useful to know the temporal distribution of the photon emission of a fluorophore I(t) that are related by the following expression:

$$\mathrm{I}_t = \mathrm{I}_0 \exp\frac{-t}{\tau_f}$$

3.2. Single molecule Förster Resonance energy transfer (FRET)

Single molecule (sm)FRET was firstly described by Weiss (123) who implemented FRET to measure biological processes of individual molecules in real time as a way to avoid the averaging in ensemble measurements. FRET is now established as a robust ratiometric tool to measure interactions between molecules in a range distance of 3 to 8 nm so it is used as spectroscopic ruler (124).

The principle of FRET was initially described by Theodor Förster in 1948. FRET measures the transfer of non-radiative energy from a donor (D) fluorescent molecule to an acceptor (A) that is usually also a fluorescent molecule (although dark quenchers are also widely used) (125). In FRET, the transfer of energy occurs when the spectrum of D emission is overlapping the absorption spectrum of the A molecule (Fig. 3.2*A*).

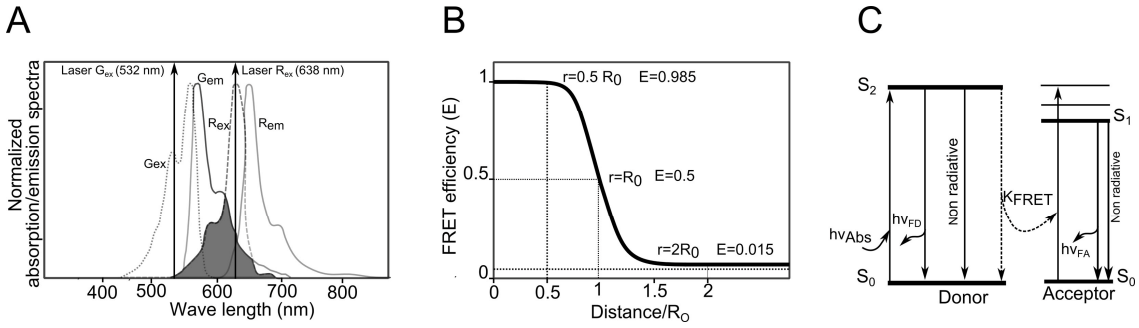


Fig. 3.2. Förster Resonance energy transfer (FRET) between two fluorescent molecules. *A*) Diagrammatic representation of the spectral overlap between the donor (D) emission and the acceptor (A) Absorption is shown in grey. *B*) FRET efficiency depends on the distance between the D and the A. *C*) Jablonsky diagram of the electronic transitions of the D and A during a FRET process. Adapted from ref. (117).

The energy transfer rate κ_T is described by the expression:

$$\kappa_T = \frac{1}{\tau_D}\left(\frac{R_O}{r}\right)^6$$

τ_D is the fluorescence lifetime; R_o is the Förster radius, r is the distance between D and A. The efficiency of energy transfer E between a single D-A pair is also described by the expression:

$$E = \frac{\kappa_T}{\kappa_D + \kappa_T}$$

Because $\kappa_D = \frac{1}{\tau_D}$ is the D fluorescence decay rate without the presence of the A.

The FRET efficiency is commonly shown as a value depending on r distance between A and D. In this formula R_o the Förster radius is the distance in which the E efficiency is 50% (Fig. 2.3*B,C*).

$$E = \frac{1}{1 + \left(\frac{r}{R_o}\right)^6}$$

The calculation of the Förster radius has been calculated using the formula $R_0^6 = 8.79 \text{x} 10^{-5} \kappa^2 n^{-4} \Phi_D J$; where n is the buffer refractive index, Φ_D is the D quantum yield, J is the integral of the region where D emission and A absorption spectra overlapped and is described by the formula:

$J = \int_0^\infty F_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda$ where $F_D(\lambda)$ is the normalised fluorescence emission intensity of the D, ε_A extinction coefficient of A, λ is the wave length.

κ^2 is an orientation factor and describe the relative orientation of D emission and A absorption transition dipoles (Fig. 2.3) which is calculated as:

$$\kappa^2 = (\cos\theta_T - 3\cos\theta_D \cos\theta_A)^2 = (\sin\theta_T \sin\theta_A \cos\theta_T - 2\cos\theta_D \cos\theta_A)^2$$

θ_T is the angle between in the emission transition dipole of D and the absorption transition dipole of A, θ_D and θ_A are the angles formed by the D and A dipoles

respectively with the vector formed by D and A. Φis the angle between D and A plane dipoles. Most studies assumed that the transition dipoles are rotating freely so κ^2 ranges from 0 to 4, resulting in a valid integrated value of all the angle combinations of κ^2=2/3. However, in some experiments, where fluorophores are coupled to DNA oligos this value is not an accurate estimation (126). Nevertheless, for the purpose of this research, a precise κ^2 value was not needed since the FRET value was not used to measure molecular distances.

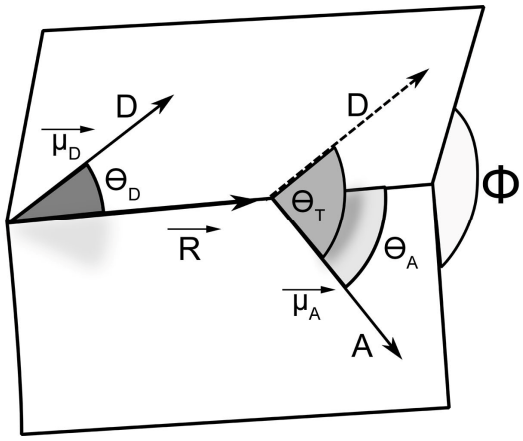


Fig. 3.3. Relative orientation of the D emission and the A absorption dipoles. The FRET rate is depending on the factor κ^2 which represents the relative orientation of the D emission dipole $\vec{\mu}_D$ versus the A absorption dipole $\vec{\mu}_A$. $\vec{R}$ is the vector formed by $\vec{\mu}_D$ and $\vec{\mu}_A$. θ_D, θ_A are the angles formed by $\vec{\mu}_D$ and $\vec{\mu}_A$ with $\vec{R}$ respectively. Φ is the angle between the planes formed by $\vec{\mu}_D$ and $\vec{\mu}_A$. Adapted from ref. (127).

3.3. TIRF-ALEX optical set up

In this research single molecule studies were performed using a custom built optical set up that combines TIRF with ALEX which together allow the visualisation of fluorophores with two different emission spectra. The schematic diagram of the microscopy set up is shown (Fig. 3.4). The detailed description of the set up has been shown previously (122).

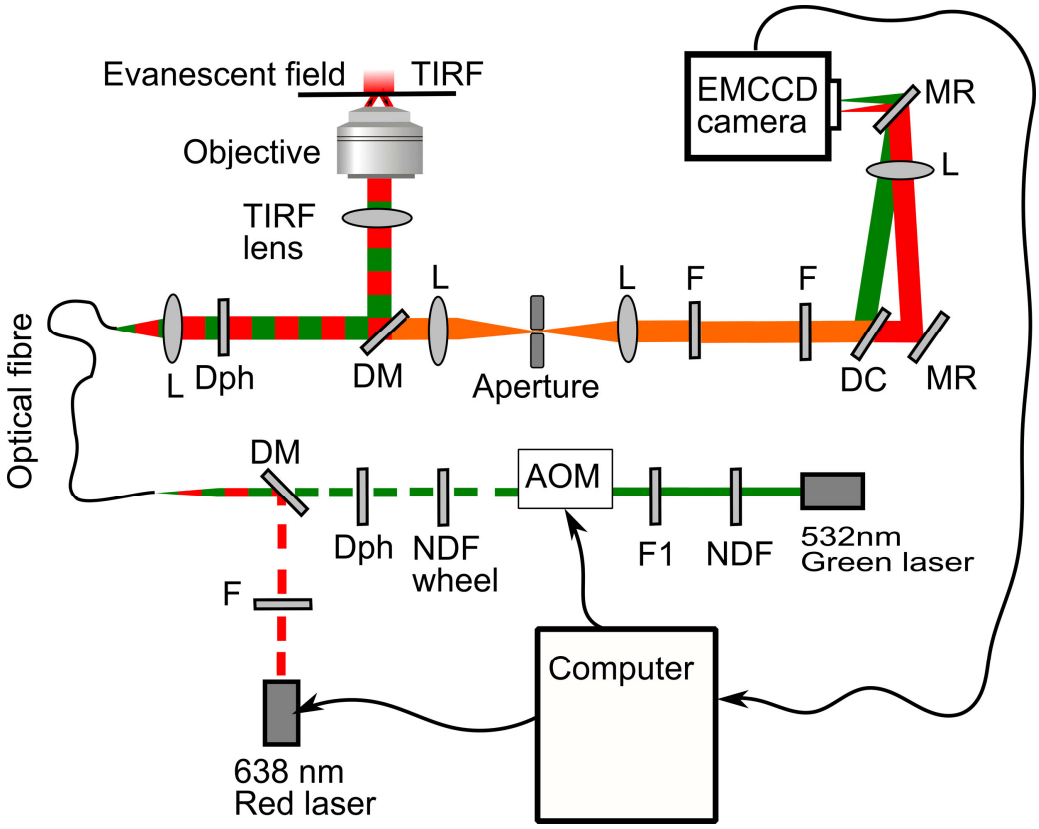


Fig. 3.4. TIRF-ALEX optical set up. Schematic of the total internal reflection microscope (TIRF) combined with alternating laser excitation (ALEX) and a two colour detection camera. The optical set up consists of three main modules, a laser light source (excitation module) that has green (532 nm Dex) and red (638 nm Aex) lasers which illuminate the sample in an alternating time period; the TIRF illumination module, a group of a lens and objective that are designed to produce TIRF illumination and, an excitation module that collects the photon counts in a EMCCD (electro multiplying charge-coupled device) camera. The three modules are coordinated with a computer that synchronises the recording of the signal with the alternation period of the lasers. NDF: neutral filter density, Dph: diaphragm, L: lens, F: filter, DM: dichroic mirror, MR: mirror. For simplicity, a detailed description about the location, direction and reflection of the laser beams in the objective lens is not drawn; the schematic is adapted from ref. (122) which describes fully the construction of the custom-built TIRF microscope used in this research.

3.3.1. Laser light source (excitation module)

The excitation module contains two lasers, the red (R) laser (Coherent Cube, USA; λ emission 638 nm), was used with an intensity ~0.75 mW and an alternation period modulated with the Cube software; the green (G) laser (Cobolt Samba,

Sweden) emits at 532 nm, and the intensity ~1.5 mW power was controlled by a density filter and the alternation with an acoustic optical modulator (AA Optics, France). The typical alternation period between G and R lasers was 100 ms. The laser beams were combined with a dichroic mirror, coupled with optical fibre and propagated to the TIRF illumination module.

3.3.2. TIRF illumination module

TIRF is a microscopy technique where the sample is illuminated in a region very close to the cover slide of ~100 nm in thickness. Because the illumination is restricted to a very small volume of sample, TIRF microscopy improves the signal-to-noise ratio and limits the exposure of the rest of the fluorescent molecules in the sample and as a result minimises the photobleaching and the background fluorescence.

The principle of TIRF illumination has already been described in detail in several reviews (128). Here I will summarise the main concepts of TIRF illumination and how TIRF was applied in this set up. The full detailed description on how the microscope was built has already been described (122).

After the beam is propagated by the optical fibre, the laser beam is delivered to the surface slide through the objective lens that collects the fluorescence from the sample. It allows an easy access to the sample and a versatile use of the objective lens (100x oil immersion objective, NA 1:4, Olympus, Japan).

In through out objective TIRF illumination which is implemented in this set up, the angle of the laser light beam is modified with a TIRF lens, so the beam is focused in the back of the focal plane of the objective lenses with the same angle (collimated).

The beam is propagated through the objective with a numerical aperture NA≥1.4. (the lens indicates the maximum angle in which the objective releases the excitation light and which must be higher than the refractive index of the sample).

Once the beam emerges from the objective, it crosses the immersion oil (Zeiss, Germany, refractive index of n = 1.45) and reaches the cover slip (No.1 0.13-0:16 mm) where it makes contact with the interface with an aqueous sample. TIRF occurs when the angle of incidence of the laser beam in the cover slip-sample interface is higher than the "critical angle" of the beam (Fig.3.5), and the refractive

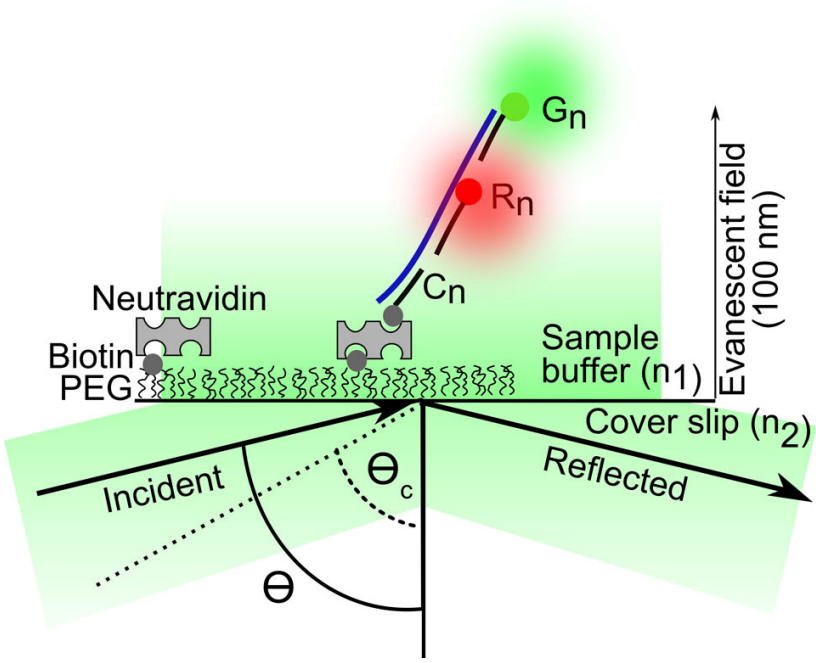


Fig. 3.5. Schematic of TIRF illumination. When the angle of the excitation laser beam is higher than the critical angle, the beam is reflected back from the cover slip-sample interface generating and evanescent field that illuminates approximately a distance of 100 nm from the surface. Θ_c: critical angle, Θ: TIRF angle, n_1 and n_2 are the refractive indexes for the cover slip and the sample respectively. Typical sample used in this research consists of a dsDNA formed by hybridised single-stranded DNAs labelled with green (G), red (R) fluorophores and biotin. The samples are attached with neutravidin to cover slip treated with biotinylated-PEG (polyethylene glycol). Adapted from ref. (129).

index in the sample (water n_1~1.33) is smaller than the one of the cover slip (n_2 ~1.5) here, the laser beam goes into a total internal reflection, creating a wave of transmitted evanescent light. The critical angle θ_c is described by Snell's law (129):

$$\theta_c = \sin^{-1}\left(\frac{n_1}{n_2}\right)$$

The intensity of the evanescent field is described by the formula:

$I_z = I_0^{-z/d}$, I_0 intensity of the evanescent field where the Z axis is equal to 0. d is the depth or the distance to the cover slip to which the intensity decreases to 37% and is expressed by: $d = (\lambda_0/4_\pi) * \left(n_2^2 \sin^2\theta - n_1^2\right)^{-1/2}$, θ, λ_0 being the incidence angle and the wavelength of the laser light beam respectively (129).

3.3.3. The concept of ALEX

In this research, TIRF microscopy is combined with two colour alternating laser excitation (ALEX). ALEX (130, 131, 132) is a method that records the fluorescence emission of two fluorophores with separated emission spectra, independently, by using in alternating time intervals, pulses of two lasers with different wavelengths (λ). A diagrammatic representation of the method is shown (Fig.3.6). The uniqueness of this method is that it can record the interaction between different fluorophores, in a time course by using two observables: a FRET (E) signal and the stoichiometry (S) also defined as the ratio of the fluorophores taking part in the interaction. In this work excitation was carried out with a green "G" (λ=532 nm) and a red "R" laser (λ=638 nm).

Emissions of G and R fluorophores (Gem, Rem respectively) were acquired in channels that detected an emission wavelength ranging from 540-630 nm for the G channel and a 650-800 nm for the R channel. The ALEX method records the photon

counts from G and R fluorophore emissions in four channels: GexGem emission of the G fluorophore (Gem) upon green excitation (Gex); GexRem emission of R fluorophore (Rem) upon Gex; RexGem, Gem upon red excitation (Rex); RexRem, Rem upon Rex.

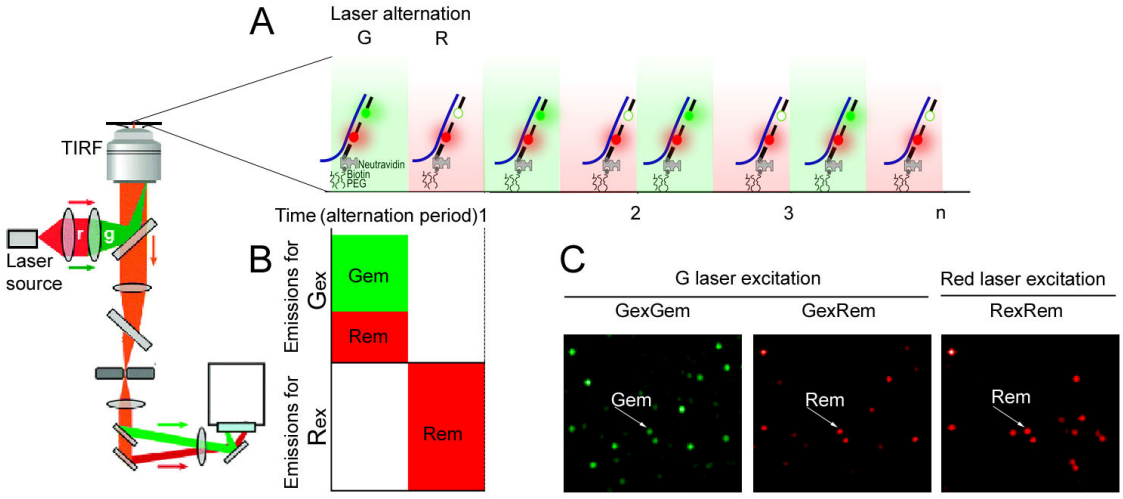


Fig.3.6. Alternating-laser excitation (ALEX) combined with TIRF microscopy. *A*) The TIRF microscope is built with a green (G) and a red (R) laser which can illuminate the sample in an alternating sequence. The sample in this cartoon represents a typical dsDNA construct used in this thesis and consists of a template ssDNA (blue line) hybridised with complementary G, R fluorescent ssDNAs (green ball and a red ball followed by a black line respectively) and a biotinylated ssDNA (grey ball representing a biotin and a black line representing the ssDNA). *B*) Each G-R laser alternation period yields three types of fluorescent emission; the excitation with the G laser (Gex) generates the emission of G labelled molecules (Gem) and the emission of red labelled molecules due to FRET (Rem); the excitation with R laser generates a fluorescent emission that it is independent from the G fluorophore. *C*) Frames generated from one alternating period. A doubly labelled dsDNA yields information about the fluorescent emission signals for the G, R labelled positions and the FRET between the two fluorophores. The information from each individual particle can be linked so each particle can be identified with a variation of E* and S in a period of time and time traces can be generated (adapted from ref. (131, 133)).

The amount of photons (F_{Yex}^{Xem}) detected in each channel are quantified using the following expressions:

$$F_{Gex}^{Gem} = I_{Gex}\sigma_{Gex}^{G}\phi_{G}\eta_{Gem}^{G}\left(1-E_{G\to R}\right)$$

$$F_{Gex}^{\mathrm{Re}m} = I_{Gex}\sigma_{Gex}^{G}\phi_{G}\eta_{\mathrm{Re}m}^{G}\left(1-E_{G\to R}\right)+I_{Gex}\sigma_{Gex}^{R}\phi_{R}\eta_{\mathrm{Re}m}^{R}$$
$$+I_{Gex}\sigma_{Gex}^{G}\phi_{R}\eta_{\mathrm{Re}m}^{R}E_{G\to R}=Lk+Dir+F^{FRET}$$

$$F_{\mathrm{Re}x}^{\mathrm{Re}m} = I_{\mathrm{Re}x}\sigma_{\mathrm{Re}x}^{R}\phi_{R}\eta_{\mathrm{Re}m}^{R}$$

$F_{Gex}^{Gem}, F_{Gex}^{\mathrm{Re}m}, F_{\mathrm{Rex}}^{\mathrm{Re}m}$ are defined as the photon count emissions, Gem or Rem from G and R fluorophores upon G or R laser excitation (Gex or Rex). It depends on: the intensity of the G(I_{Gex}) or R laser ($I_{\mathrm{Re}x}$) used; a set of properties which are characteristic of each fluorophore such as the absorption of the specific fluorophore upon G or R excitation ($\sigma_{Gex}^{G}, \sigma_{Gex}^{R}, \sigma_{\mathrm{Re}x}^{R}$), the quantum yield ϕ_{G}, ϕ_{R}; the detection efficiency of G and R fluorophores upon excitation with G or R laser ($\eta_{Gem}^{G}, \eta_{\mathrm{Re}m}^{G}, \eta_{\mathrm{Re}m}^{R}$); and the FRET efficiency $E_{G\to R}$. F^{FRET} is the photon count for R-emission exclusively due to a FRET interaction. The most complex photon emission occurs in $F_{Gex}^{\mathrm{Re}m}$ because it corresponds with the photon counts of three different sources: the *Lκ* which is the leakage of the Gem upon Gex in the red detection channel, the *Dir* is the Rem upon Rex by G laser wavelength. The photon counts of green fluorophores due to excitation by the red laser F_{Rex}^{Gem} is negligible and is not accounted for the analysis.

The raw photon counts from $F_{Gex}^{Gem}, F_{Gex}^{\mathrm{Re}m}, F_{\mathrm{Rex}}^{\mathrm{Re}m}$ are used in the ALEX technique to calculate the FRET (E*) and the stoichiometry (S) ratios of an interaction between G and R fluorophores:

$$E^{*} = \frac{F_{Gex}^{\mathrm{Re}m}}{F_{Gex}^{Gem}+F_{Gex}^{\mathrm{Re}m}} \qquad S_{raw} = \frac{F_{Gex}^{Gem}+F_{Gex}^{\mathrm{Re}m}}{F_{Gex}^{Gem}+F_{Gex}^{\mathrm{Re}m}+F_{Gex}^{\mathrm{Re}m}}$$

Taking together the E* and S raw values define each individual molecule of a population without ambiguity, and it is used in this research to characterise labelled molecules with G and R fluorophores by a code of E*/S.

Although for the purpose of this research accurate measurements of $E_{G\to R}$ efficiency and accurate S are not used, they can however, be calculated by the expression:

$$E_{G\to R} = \frac{F^{FRET}}{\gamma F_{Gex}^{Gem} + F^{FRET}} \qquad S = \frac{F_{Gex}^{Gem} + F^{FRET}}{F_{Gex}^{Gem} + F^{FRET} + F_{\mathrm{Re}x}^{\mathrm{Re}m}}$$

where the value of F^{FRET} is calculated by the formula

$$F^{FRET} = I_{Gex}\sigma_{Gex}^{G}\phi_{R}\eta_{\mathrm{Re}m}^{R}E_{G\to R}$$

and the value of accurate FRET can be corrected by γ, which is calculated by the expression: $\gamma = \frac{\Phi_R \eta_{\mathrm{Re}m}^{R}}{\Phi_G \eta_{Gem}^{G}}$

3.3.4. Camera and image splitter device (emission module)

The photon count emission of the G and R fluorophores are collected by the objective and separated from the incoming excitation laser beams with a dichroic mirror (450 nm/650 nm, Semrock, USA) and filters (545 nm LP, Chroma, USA; and 632/25 nm notch, Semrock, USA) and the G and R emission photons split and projected into two different regions using an image splitter device (dichroic mirror 630 nm DRLP, Omega, USA) in a EMCCD camera (electro multiplying charge-coupled device, Andor iXon, Ireland). The size of the field of view used is 512x256 pixels.

The magnification of the set up was 170x (94 nm x 94 nm per pixel). The alternation period of the green and red excitation was synchronised with the frame rate of the camera. 12.5% of the intensity detected in the green channel leaks into the red channel which generates an apparent FRET baseline of E*=0.11 from a Cy3B-only signal. This is due to the imperfect colour splitting and the natural spectral leakage of Cy3B into the red channel. The red channel intensity of ATTO647N when excited by the 532 nm laser is 3% of the intensity of Cy3B excited by the 532 nm laser (or ATTO647N excited by the 635 nm laser). Therefore, the apparent FRET is E*=0.13 when both donor and acceptor are present without actual FRET.

3.3.5. Detection and linkage of G and R fluorescence particles

The point spread functions (PSF) corresponding to the fluorophores (Fig. 3.6*C*) are analysed with Two Tone, which is custom software written in Matlab (134). The particles are detected and chosen for analysis as follows: particles with a positive signal in the first 10 frames are fitted with a Gaussian mask and only the particles above a threshold are selected for the next step of analysis. Then, the PSFs are fitted with a 2-dimensional Gaussian (2D Gaussian) and the centroid of each particle is localised. After that, the particles in the G and R channels are linked using a map generated using the leakage of the Rem of fluorescent beads into the Gem channel. Once the particle linkage is performed, the particles are filtered using a threshold for the nearest neighbour distance (96 nm/pixel) and the particle excentricity, E_{cc} is described by the formula:

$E_{CC} = \sqrt{1 - \frac{s_x^2}{s_y^2}}$, where s_x^2 and s_y^2 are the PSFs' width in the longest and smallest dimensions. In these experiments the ellipticity and nearest neighbour (*NN*) criteria were $E_{cc} < 0.7$, $NN > 5$ pixels.

3.3.6. Plotting of fluorescence emission intensity of G and R particles

To extract time traces from filtered PSFs, the PSF in each G and R frame are fitted with a circular 2D function, and the fluorescence emission intensity obtained from the integral of each function plotted in a single data time bin for the G and emission intensity. 2D histograms were plotted with the E* and S corresponding with the fluorescence emission intensity of the frames that define each particle.

Histograms display the combined data of movies recorded in identical conditions. Statistical analysis ANOVA (Tukey's test) and the t-test were performed and plotted using GraphPad (GraphPad Software Inc).

3.3.7. Correction of non-uniform illumination in the field of view

In TIRF experiments, there is a significant correlation between fluorescence intensity from a single molecule and the location of the molecule within the field of view. This is caused by the non-uniform excitation field across the field of view and can be approximated by an elliptical 2D Gaussian. The non-uniform excitation has been corrected (see $A_{ex}A_{em}$ intensity distributions in Fig. 3.3 and 3.5) calculating for each molecule that emitted at a stable level (no photoblinking and no photobleaching) the mean intensity counts per frame $\bar{F}$ (e.g. in the acceptor channel) and localisation position $(\bar{x}, \bar{y})$. Next a 2D Gaussian function of the form

$$f_F(x,y) = b + A\exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}\right) = b + A\,\phi(x,y)$$

was fitted to the set of experimental data points $\{\bar{F}(\bar{x},\bar{y})\}$, compiled from all single molecules detected in movies taken under identical conditions. Each single-frame intensities F of a molecule found at $(\bar{x},\bar{y})$ can be then corrected by

$$F^* = b + \frac{F-b}{\phi(\bar{x},\bar{y})},$$

corresponding to a mapping to a normalised intensity level $b + A$. Therefore, F^* corresponds to the intensity observed if the position of the molecule were at (μ_x,μ_y), the maximum of the 2D Gaussian profile f_F. Corrected intensities were designated with an asterisk (i.e., $A_{exc}A_{em}$*).

A molecule was defined to be stable when the single-molecule intensity time series $F = \{F_1,\ldots,F_{20}\}$ (20 x 100 ms ALEX frames) did not have more than 2 consecutive frames with intensity counts deviating from the median intensity more than 2 times the median absolute deviation

$$\mathrm{MAD}(F) = \mathrm{Median}(\{|F_i - \mathrm{Median}(F)|\})$$

in both detection channels. MAD is a robust estimator for the standard deviation σ, and for normally distributed data it can be shown that $\sigma \approx 1.48\,\mathrm{MAD}$.

3.3.8. Number of colocalisations observed vs. the number expected due to random coincidence

Random colocalisation of D and A probes can occur due to the density of probes deposited on the surface slide and the method used in Two Tone data analysis software used in this research. As a result random colocalisation was quantified in this research.

Colocalisation of a pair of donor and acceptor particles is defined as the finding of a pair when the donor-acceptor distance is less than the R_{search} ($R_{search} = 250nm$) and no further donor or acceptor particles are present within a vicinity of $R_{exclude} = 500nm$ from each of the particles analysed. Using the R_{search} and $R_{exclude}$ thresholds together with the observed densities of donor and acceptor particles ρ_D and ρ_A in each field of view, the random colocalisation density can be estimated as:

$$\rho_{coloc} = \pi R_{search}^2 \rho_D \rho_A (1 - \pi R_{exclude} \rho_D)(1 - \pi R_{exclude} \rho_A).$$

This formula overestimates the number of detected random colocalisations since additional filters are in place to exclude particles with elliptical point spread functions (arising from closely spaced particles) as well as particles that fall below an intensity threshold. We determined ρ_D and ρ_A by counting the number of donor and acceptor particles in the homogenously illuminated part of the field of view and dividing by the area.

The number of observed colocalisations was tested against the number of estimated random colocalisations using a one-sided paired-sample Wilcoxon signed-rank in MATLAB (MathWorks). A paired test was used of the observed number and estimated number from the same movie, since both depend on the density of donor and acceptor signals. The test is preferred to a paired t-test, as it is not assumed to be normally distributed data.

3.4. Confocal ALEX microscopy

Measurements of samples in solution were performed using sm confocal microscopy set up as described elsewhere (135). pM concentrations of molecules labelled with fluorophores were allowed to diffuse through out a femtolitre confocal spot and the signal burst detected with avalanche photodiodes (APD). In confocal microscopy the observations are restricted to the diffusion time in the

confocal spot (~1ms) but they are not affected by artefacts caused by the interaction between surface and the sample.

3.5. TIRF slide preparation

Cover slips (25x65 mm, thickness number 1, VWR) used in single molecule experiments were treated at high a temperature. Coverslips were placed in a furnace at 773 K at a rate of 90 K/hr, once the temperature was reached, it was kept for 1hr more followed by a temperature decrease gradient of 50 K/hr.

3.5.1. Vectabond coating

The coverslips were then coated with Vectabond solution (Vector Laboratories). The slides were immersed in a staining beaker containing 50 ml of acetone for 5 min. A 1 ml solution of Vectabond was then added to the acetone solution and the coverslips were incubated for another 5 min. After that coverslips were then rinsed thoroughly with water and dried under nitrogen.

3.5.2. PEG coating

Next, silicone gaskets (Grace Bio-Labs, USA) (136) were placed on top of the coated coverslip. 20 µl of PEG solution (400 µl of 100 mM Sodium Borate, pH 8.5, 80 mg of NHS-PEG (PEG-SVA mw 5000 Da, LAYSAN Bio), 1-3 mg biotin-PEG-NHS (mw 5000 Da, LAYSAN Bio) per chamber was used. Each chamber was incubated with the PEG solution for three hours at room temperature. The chambers were then rinsed three times with 100 µl of PBS buffer (pH 7.3). The slides could be kept in a humidity chamber at 4ºC for 1 week (making sure that the slide never dried).

3.5.3. Neutravidin coating and DNA incubation

Neutravidin solution (0.5 mg/ml) was prepared in PBS buffer.

20 µl of neutravidin solution was added to the chambers from which the PBS had been previously removed and incubated for 5 min at room temperature. The excess of neutravidin was removed and the chambers were washed three times with 50 µl of PBS buffer to remove the excess of neutravidin. After that 40 µl of ~50 pM biotinylated DNA was incubated in the chambers for ~5 min at room temperature. The solution was removed, the chambers rinsed with 150 PBS µl after which the chamber was filled with imaging buffer solution which consists of PBS containing 1.6% W/V glucose, an oxygen scavenging system (165 U/ml glucose oxidase, Sigma, 2170 U/ml catalase, Roche) and 2 mM Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2–carboxylic acid) sealed with a glass slide from air contact.

3.6. DNA design and hybridisation

Oligonucleotides to investigate rotavirus transcripts were designed using a positive-strand segment 2, 6, 10 or 11 as a template with amino-C6-dT modifications at the 5'-end. The approximate E* value of ~0.35 was calculated using the Förster radius of the donor-acceptor pair (134), the donor-acceptor separation (19-20 bp, ~6.75~7.02 nm) upon hybridisation to the targeted transcript, the size of the fluorophore and linker and the variation due to the free rotation of the fluorophores (137). The probes were relatively short (~40 and ~20 nt for the capture and FRET probes, respectively) to minimise formation of secondary structures, duplexes, or non-specific binding, but still generate the required proximity. The probes targeted regions close to the 5'-end of ssRNA of segment 2, 6, 10 or 11 (ssRNA2, ssRNA6 and ssRNA11, respectively), while avoiding the conserved consensus sequence in the 5'-UTRs that may serve as assortment signals (138).

All the sequences used for hybridisation of viral RNA are described in Table 3.1. Oligos $T_{2,}$ T_6, $T_{11,}$ $T_{6\text{-}11,}$ $T_{11}^{th,}$ C_2, C_6 and C_{10} were purchased as biotinylated forms (IBA, Gottingen, Germany), whereas C_{11} was biotinylated using sulfo-LC NHS biotin (Invitrogen). The other oligos were labelled at the 5'-end with either Cy3B or ATTO647N *N*-hydroxy-succinimidyl esters and purified using denaturing PAGE. Positive controls T_6-G_6-R_6, T_{11}-G_{11}-R_{11},$T_{(6\text{-}11)}$-G_6-R_{11} and G_{11}-Sp-R_{11}-R_{11}-$T(_{11}^{th})$ were prepared by mixing equimolar concentrations of oligos. For example T_{11}-G_{11}-R_{11} consisted of oligos T_{11}, G_{11} and R_{11}. The annealing buffer consisted of 20 mM Tris-HCl pH 8, 250 mM NaCl, 1 mM EDTA. The oligo mixture was heated at 95º C and left to cool slowly to 4º C. In TIRF experiments

Table 3.1. Oligonucleotide used for hybridisation in the viral studies.

Oligo	Segment sequence	Oligo sequence	Modification
G_6	(+17/+36)B	5'-ACAGGACATCCATGTTGAAG-3'	Cy3B 5'end
R_6	(+37/+56)B	5'-AAGAGTTTTTGACAAGGAGT-3'	ATTO647N 5'end
T_6	(+14/+56)T	5'-AGTCTTCAACATGGATGTCCTGTACTCCTTGTCAAAAA CTCTT-3'	Biotin 5'end
C_6	(+56/+95)B	5'-GTATAATGTGCCTTCGTCAATTTTGACTCTAGCATCTT TA-3'	Biotin 5'end
C_{10}	(+21/+40)B	5'-TTTCCGCACGCGCTCTCTCG-3'	Biotin 5' end
G_{11}	(+16/+35)B	5'-ATACTGAGAGACATCACTGT-3'	Cy3B 5'end
R_{11}	(+36/+54)B	5'-TGGAAGACTTGTCACGTCA-3'	ATTO647N 5'end
T_{11}	(+15+54)T	5'-TACAGTGATGTCTCTCAGTATTGACGTGACAAGTCTTC CA-3'	Biotin 5'end
C_{11}	(+55+94)B	5'-TGAAGATGATTCATGTTTATAAATGCTGGAGGAAATA GA-3'	Biotin 5'end
$T_{(6\text{-}11)}$	Hybrid template	5'-AGTCTTCAACATGGATGTCCTGTTCCAAGTTAAGTGAT CTAATTCAACTGACGTGACAAGTCTTCCA-3'	Biotin 5'end
$T_{(11^{th})}$	Hybrid template	5'-ACAGTGATGTCTCTCAGTATTCCAAGTTAAGTGATCTA ATTCAACTGACGTGACAAGTCTTCCATGACGTGACAAGT CTTCCA-3'	Biotin 5'end
Sp	(spacer sequence)	5'-GTTGAATTAGAAGATCACTTAACTTGGA-3'	
G_2	(+41+60)B	5'-TTTAAATTCGTCTCACGACG-3'	Cy3B 5'end
R_2	(+61+79)B	5'-TTGCATTCGATCATCTTGT-3'	ATTO647N 5'end
T_2	(+41+79)T	5'-CGTCGTGAGACGAATTTAAAACAAGATGATCGAATGC AA-3'	Biotin 5'end
C_2	(+21+40)B	5'-CGCTCCACGCTTTCTGTACG-3'	Biotin 5'end

Oligos are described with a capital letter that indicates the type of probe and labelling used followed by a subscript indicating the segment. The (+)ssRNA (sense strand) sequences are named top strand (T) and numbered from the 5' end (+1). Oligo sequences are designed using the (+)ssRNA as a template; B denotes bottom strand (anti-sense strand) complementary to a sequence in the T strand (sense strand). G is a green FRET probe labelled with Cy3B; R is a red FRET probe labelled with ATTO647N corresponding to the segment used as a template; C is a capture probe labelled with biotin. Sp is a spacer sequence based partially on a ssRNA6 genomic sequence, and is part of the long dsDNA control which consisted of one G_{11} and two R_{11} probes hybridised to complementary sequence $T(_{11}^{th})$. Sp minimises the flexibility and non-specific hybridisation of probes in the G_{11}-Sp-R_{11}-R_{11}-$T(_{11}^{th})$ construct.

freshly prepared solutions of control constructs were diluted to (~50 pM) using serial dilutions from stocks.

3.6.1. Polymerase chain reaction (PCR)

For transcription studies of *E. coli* RNA polymerase, all the oligos (Table 3.2) were purchased from (IBA). The promoter construct LacUV5-50(-40+64Cy5) designated as in ref. (139) was generated by standard PCR. The reaction mix consisted of: 0.5μM of each forward primer LacU5-50(-40/-20)TBio-40, reverse primer LacUV5-50(+44/+64)BCy5+64, 0.5 nM of template promoter LacUV5 spanning nucleotides (-40/+64)T, 250 μM dNTPs, 0.5 U taq polymerase (NEB),1x PCR buffer (ThermoPol). The reaction was performed in a thermocycler (PX2 Thermo Fisher, MA PCR) using 30x (94ºC 1m, 94ºC 20s, 55ºC 20s, 72ºC 15s, 72 ºC 3min) and 4 º C. Samples were purified by PAGE with a estimated final concentration of 500 nM as determined using a confocal microscopy. An additional oligo LacUV5(+5/+28)B-Cy3B labelled at the 5' with Cy3B was used to hybridise newly synthesised RNA transcripts.

Table 3.2. Oligonucleotides used for studying transcription in *E. coli* RdRP.

Oligo name	Segment sequence	oligo sequence	Modification
LacUV5(-40/-20)TBio-40	(-40/-20)T	5'- AGGCTTTACACTTTATGCTTC -3'	Biotin 5'end
LacUV5(+44/+64)BCy5+64	(+44/+64)B	5' - GAAATTGTTATCCGCTCTCAC -3'	Cy5 5'end
LacUV5(-40/+64)	(-40/+64)T	5'-AGGCTTTACACTTTATGCTTCCGGCTCG TATAATGTGTGGAATTGTGAGAGGTTAGT TGATTGTATTGAGTTGATTGGATTGAGTG AGAGCGGATAACAATTTC -3'	Template
G(+5/+28)	LacUV5 (+5+28)B	5'-CAATACAATCAACTAACCTCTCAC-3'	Cy3B 5'end

Oligos are described as in ref. (139); oligos are designed on LacUV5 promoter sequence. Specific region are in brackets. B denotes bottom strand (anti-sense strand); T top strand (sense strand); Bio stands for biotinylated probe. G is for "green" probe (labelled with Cy3B).

3.6.2. Measuring DNA concentrations

The concentration of labelled DNA probes was measured with a UV/Vis spectrophotometer (Nanodrop, USA); DNA concentration was calculated measuring its absorbance at 260 nm, and the molar extinction coefficient (ε) of each sequence. The fluorophore concentration of each DNA construct was measured using its maximum absorbance and the extinction coefficient ε for Cy3B = 130:000 $cm^{-1}M^{-1}$ (140), ATTO647N = 155000 $cm^{-1}M^{-1}$ (141) and Cy5 =250000 $cm^{-1}M^{-1}$ (142).

3.6.3. DNA labelling

Amino-modified DNAs (5-C6-Amino-2-dT) were labelled with NHS-ester-modified fluorophores. First, 5 nmoles of DNA were precipitated by adding 10 µl of 3 M NaCl and 250 µl of ice-cold 100% ethanol. The solution was mixed and placed at -80°C for 1 h, then centrifuged at 4°C and 16,000 g for 10 min. After this, the supernatant was removed. To remove the excess of salt the pellet was washed using 200 µl of 75% ice-cold ethanol and centrifuged at 4°C and 16,000 g for 5 min. This process was repeated twice after which the pellet was dried either at room temperature for ~15 min or until the liquid had completely evaporated.

The pellet was resuspended in 95 µl of labelling buffer (0.1 M sodium borate buffer, pH 8.5) and mixed with 50 nmoles of derivatised fluorophore, and dissolved in 5 µl of anhydrous DMSO. The sample was protected from light, the tube sealed with parafilm to avoid evaporation and then incubated at 50ºC for 1 hour. After this an additional 50 nmoles of the fluorophore were added to the sample and incubated overnight. The labelled DNA was then ethanol-precipitated by adding 10 µl of 3 M NaCl and 250 µl of ice-cold 100% ethanol to the reaction mix. The solution was mixed and placed at -80ºC for 1 h. Next, the solution was centrifuged at 4ºC and

16,000 g for 10 min, the supernatant was removed, and the DNA pellet rinsed with 200 µl of ice-cold 75% ethanol. The solution was centrifuged again at 4ºC, 16,000 g for 5 min, the supernatant was removed, and the DNA pellet was dried at room temperature as described above.

3.6.4 DNA purification protocol

After ethanol precipitation the oligos were resuspended in equal amounts of TE buffer and dimethylformamide (5µl of each solution) and separated from the unlabelled fluorophore using denaturing 20 % polyacrylamide gel electrophoresis, 6 M UREA that denatures the ssDNA. The mixture was heated at 95ºC for 5 minutes and then kept on ice. The gel was pre-run at 450 V for 30 minutes in TBE buffer (89 mM Tris-HCl-HCl, 89 mM boric Acid and 2 mM EDTA) before adding the sample and then running the gel at 550 V for about 6 hours. Then the band was cut out, crushed and dissolved in TE buffer over night. The ssDNA solution was separated from the gel fragments by centrifugation 16000 g 1 min and filtering with a 0:45µm pore size spin column. Samples were concentrated with a speed vac to ~70 µl, and the UREA buffer exchanged with miliQ water using micro-columns (Nanosep 3K spin columns , Pall, USA, or Micro Bio-Spin 6 column from Bio-Rad, USA). The DNA filtrate was then dried in then speed-vac and resuspended in TE buffer to a µM concentration.

3.7. Molecular virology methods

3.7.1. Synthetic RNA control: T7 synthetic RRV 2 and 11 segment.

cDNA copies of RRV 2 and 11 genomic segment were amplified from viral dsRNA in a sequence-independent manner using the method of full-length amplification of

cDNAs (FLAC), (143,144) Hairpin anchor primer C9 was ligated to viral dsRNA, followed by cDNA synthesis from gel-purified genome segments with RevertAid Premium reverse transcriptase (Fermentas) at a concentration of 10 U/µl and 55°C for 1.5 hr. PCR amplification was performed using FLAC 2 primer with KOD Hot Start DNA Polymerase (Novagen). PCR products were cloned and fully sequenced and the T7 promoter and a suitable restriction site at the end were introduced to generate correct ends. Transcripts with a 5′-cap analogue were generated from the digested T7 plasmid clones using a mMESSAGE mMACHINE T7 Ultra Kit (Ambion). Purified ssRNA were resuspended in nuclease-free water and stored at −80°C.

3.7.2. Double-layered particles (DLPs) purification

RRV were kindly provided by Prof H. Greenberg and the DLPs purified from rotavirus infected cells as described (145). MA104 cells infected with RRV were harvested at 100% cytopathic effect; we used MOI of 0.5 - 1 for growing stock viruses (propagation of the virus) and MOI of 3 for infection and purification of DLPs. Cells were disrupted by freezing and thawing twice, EDTA (10mM, pH 8) added, followed by incubation for 1 hr at 37°C. After centrifugation, the pellet was re-suspended in TNC buffer (10 mM Tris-HCl pH 7.4, 140 mM NaCl, 10 mM $CaCl_2$) with 0.1% NP-40 and 50 mM EDTA pH 8.0. Half volume of trichlorotrifluoethane was added to the lysate and mixed with vortex. Aqueous phase was separated by centrifugation and DLPs were isolated by equilibrium ultracentrifugation at 100,000 g in CsCl gradient for 18 hr; the DLPs band was collected, diluted with TNC buffer and pelleted by ultracentrifugation at 110,000 g for 2 hr.

3.8. Capture and identification (CID) assays

3.8.1. Capture assay of transcripts in solution

T7-derived positive-sense ssRNA2 or 11 (10 nM) were mixed with a hybridisation buffer (100 mM Tris-HCl pH 8.0, 9 mM $MgCl_2$, 150 mM NaCl, 0.5 U/µl RNase inhibitor), a FRET code that consisted of 10 nM each of G_{11} and R_{11} probes, and 10 nM capture probe C_{11}. The mixture was incubated for 1 hr at 37 ºC and then aliquots were transferred to a TIRF-ALEX microscope for analysis. The same hybridisation method was performed for T7-derived positive-sense ssRNA2, but using G_2 and R_2 FRET probes, and capture probe C_2. Transcription activity assays (145, 53) of DLPs were performed as above with the following modifications: DLPs (2 nM) were mixed with 20 nM of probes C6, G6, R6. The sample was mixed in a transcription buffer 100 mM Tris-HCl pH 8.0, 4 mM ATP, 2 mM GTP, 2 mM CTP, 2 mM UTP, 9 mM $MgCl_2$, 150 mM NaCl, 0.5 mM S-adenosyl methionine (SAM), 0.5 U/µl RNase inhibitor for 2 min at 37 ºC, after which 2 mM EDTA was added to the reaction. 20 µl aliquots were removed at time intervals of ~1, 15, 30, 60 or 90 min, added to a slide for 2 min, and imaged using a PBS buffer with 1 mM TROLOX (6-hydroxy-2,5,7,8-tetramethyl-chroman-2-carboxylic acid) (121) and an oxygen scavenger system (see sample preparation for details). The negative control consisted of an experiment in a transcription buffer but without nucleotides. The same hybridisation method was implemented to capture ssRNA2, and ssRNA11 in solution but using specific FRET and capture probes for each transcript.

3.8.2. Capture assay of extruded transcripts from DLPs

A suspension of DLPs in Tris-HCl pH 8 was exchanged in PBS buffer using a Micro Bio-Spin 6 column (Bio-Rad). After that, 50 µl DLPs (2 nM) were mixed with the transcription buffer (50 mM phosphate buffer pH 7.4, 1.4 mM KCl, 150 mM NaCl, 4 mM ATP, 2 mM GTP, 2 mM CTP, 2 mM UTP, 9 mM $MgCl_2$, 0.5 mM SAM, 0.5 U/µl RNase inhibitor) and incubated for 1 min at 37 ºC. The transcription was stopped by fixing the DLPs with acidified glutaraldehyde (GT) vapour for 5 min at room temperature (146), followed by 5 min incubation with 150 mM Tris-HCl (pH 8.0) to quench the fixation reaction. The GT-containing buffer was then exchanged using a Micro Bio-Spin 6 column equilibrated in sodium-chloride sodium-citrate (SSC) buffer.

Hybridisation to target ssRNAs templates was performed using a hybridisation buffer, and conditions reported previously (147), but with several modifications. DLPs were mixed with 10 nM of C_6, G_{11} and R_{11} and a hybridisation buffer consisting of 2×SSC, 9 mM citric acid (pH 6.0), 0.1% Tween-20, 50 μg/mL heparin for 2 hr at 50 ºC. The hybridisation buffer was then exchanged in SSC using a Micro Bio-Spin 6 column (Bio-Rad).

Capture was also performed to identify transcripts from segment 6; in this experiment, gentler hybridisation conditions were introduced. Transcription was performed using the same buffer conditions optimised to capture segments in solution, but decreasing the concentration of nucleotides (1 mM ATP, 500 µM UTP, 500 µM GTP, 500 µM CTP). The reaction mixture was incubated for 2 min and the reaction was then blocked using 4 mM EDTA. 20 nM of the probes G_6, R_6, C_{10} and C_{11} were used, and the hybridisation lasted for 1 hr at 37 ºC. Experiments were also performed to identify transcripts from segment 2 using the same conditions and probes G_2, R_2, C_6, and C_{11}. For the colocalisation of segments 6 and 11, the

experiment was performed as described above, but using capture probes C_6, C_{11} and reporter probes G_6, R_{11}. In colocalisation experiments using RNA as a template, we synthesised the RNA using the same transcription conditions used in DLP-capture experiments. The ssRNA was separated from the DLPs by 50 min centrifugation at 100,000 g in an ultracentrifuge (Optima TL, Beckman; rotor TLA 100.2).

3.8.3. Capture of DLPs bearing extruded transcripts using capture antibodies

The experiment was performed as described previously (148) with some modifications. DLPs were allowed to transcribe for 2 min and fixed with 2% paraformaldehyde for 5 min. DLPs were buffer-exchanged to SSC buffer using a micro Bio-Spin column (Bio-Rad) and interrogated using 20 nM of G_2 and R_2 FRET probes, which target (+)ssRNA2 for 2 hr at 37°C. DLPs were then buffer-exchanged in 10 mM Tris-HCl pH 8, 50 mM NaCl, and incubated for 1 hr with 66 nM sheep anti-VP6 (primary antibody ab35417, abCAM). Next, the DLPs were added for 5 min to a coverslip coated with biotinylated rabbit anti-sheep secondary antibody (ab6746, abCAM), prepared by incubating the coverslip with secondary antibody (26 nM in the same buffer conditions) for 30 min followed by a PBS wash. We used two negative controls to test non-specific interactions of antibodies to DLPs: the first had the primary antibody substituted by rabbit gamma-globulins (purified from non-immunised rabbit serum; 011-000-002, Jackson Immunoresearch); the second was performed in the absence of biotinylated secondary antibody.

Chapter 4

Design and Development of an Assay to Capture and Identify ssRNA from Rotavirus

4.1. Introduction

Double-stranded RNA (dsRNA) viruses comprise a wide variety of families which vary in genome complexity, yet they share unique features reflecting parallels in their replication; for the *Reoviridae,* such features include a multicomponent capsid that crosses the host cell membrane, and transcription of their dsRNA segments by capsid-attached enzymes. During cell entry, the outer layers of these viruses are lost, while their inner capsids forming the DLP provide a compartment for genome segments (10-12 dsRNAs).

Transcript export occurs via channels at the 12 vertices of the DLP; the release of new (+)ssRNA transcripts is crucial for establishing infection, since the transcripts act as templates for both translation and genomic dsRNA synthesis.

There are still, several questions that remain unclear regarding the synthesis of (+)ssRNA; first the mechanism of transcript export is unclear. In particular, it is unknown whether the nascent transcripts are selectively released through specialised channels, and if so, what the basis of selectivity is; second the mechanism by which these transcripts are recruited for translation or for capsid packaging is still unknown.

To address these questions, it is necessary to develop a technique that is able to identify and quantify the copy number of specific (+)RNAs from the DLPs.

In this chapter, I hypothesise that single molecule hybridisation combined with TIRF and ALEX microscopy techniques can be used to capture, identify and quantify specific ssRNAs in solution. To test this hypothesis I designed a capture and identification (CID) assay that implements these three developments, the single molecule hybridisation with the ALEX method and TIRF microscopy.

The first of these three combined methods, the single molecule hybridisation relies on a labelled complementary sequence to tag and identify specific sequences; the use of hybridisation is a powerful tool in biology for a wide range of applications from the initial application in PCR to microarrays (149), fluorescence in situ hybridisation of RNA (FISH) (150) or molecular beacons (151, 152); the ALEX method (131) may distinguish between specific hybridisation events because it identifies single molecules using two robust observables. The first is the raw uncorrected FRET, E*, the second is the stoichiometry S, defined as a ratio of G and R fluorophores present in the labelled molecule. The implementation of ALEX results in the independent recording of donor and acceptor excitations which can be used to measure a specific signal due to FRET between a single pair of fluorophores between 0 and 10 nm and the S that informs about the ratio between labelled fluorophores. The stoichiometry "*S*" ratio provides a signal to investigate molecular coincidence independently from the FRET signal; *S* classifies molecular complexes with multiple D-A fluorescence signals, S combined with FRET also enables the study of biomolecular structure by monitoring different molecular conformations (153). The observation of molecules with TIRF illumination means that only molecules bound to the surface of the slide are recorded, minimising the background signal and potentially extending the use of the method for measuring real time observations.

4.2. Results

4.2.1. Experimental design of a capture assay based on single molecule hybridisation of ssRNAs

To investigate the questions described above, it is necessary to know the identity and copy number of one or more ssRNA segments extruded from the capsid after the initiation of transcription by the DLPs. To achieve this, I developed a single-molecule fluorescence assay that relies on the hybridisation of RNA segments with short DNA probes that result in specific signals of fluorophore proximity (as measured using FRET) and fluorophore stoichiometry (as measured using an S ratio). This assay in its basic design consists of three ssDNA complementary probes that hybridise close to the 5' end of the positive strand of the target ssRNA. A set of two probes are designed as a FRET code of about 35 % and consists of two singly labelled ssDNA probes, one probe labelled with a donor fluorophore (Cy3B) designated "G" probe and the other probe labelled with an acceptor fluorophore (ATTO647N) designated "R" probe. A third probe (capture probe), designated as "C" probe is designed downstream at the region complementary to the FRET code, and it is biotinylated.

In order to identify RNA synthesis from DLPs, the DLPs were mixed with the probes in the transcription buffer (Fig. 4.1). Newly synthesised transcripts were released from the capsid and hybridised to the capture probe, which bound the RNA covalently to a slide treated with neutravidin. The microscope slides were first treated with a mixture of PEG and biotinylated PEG that passivated the surface minimising non specific absorption (154). Imaging was performed using TIRF-ALEX microscopy that detects only the signal in the limits of an evanescent field of

approximately 100 nm and limits the fluorescence of any background signal. Target ssRNA are identified as specific hybridisations with the expected FRET code and stoichiometry ratio.

The assay can be extended to identify ssRNA on the DLP surface. In this experimental design, the DLP synthesis is blocked after a few minutes to minimise an excess of transcripts in solution. I designed a biotinylated capture probe on RNA targets different from that identified with the FRET code so the capture probe would pull down the DLPs on the surface. In this way the FRET code identified only those ssRNAs on the DLP surface (Fig. 4.1). Data imaging analysis is subsequently applied to sort labelled molecules according to their FRET and stoichiometric ratios.

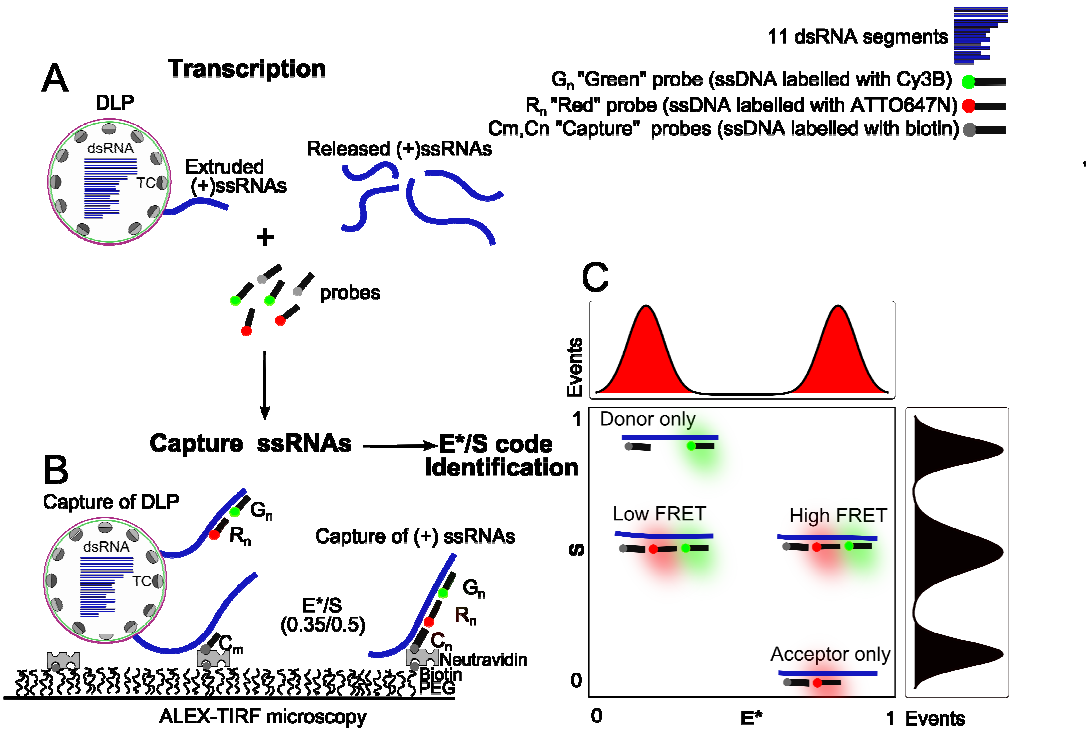


Fig. 4.1. Schematic of the CID assay. *A*) DLP are mixed in a hybridisation buffer at 37ºC in a test tube. Transcripts (either attached to DLPs or free) are incubated with complementary capture and FRET probes (donor shown in green; acceptor shown in red) *B*) The hybridisation of the probes results in the capture of transcripts to the neutravidin treated surface of the microscope slide. *C*) Captured transcripts are identified using ALEX-TIRF. Specific transcripts are linked to specific

combinations ("codes") of FRET efficiency E* and relative probe stoichiometry S; in most cases, an E*/S code of ~0.35/0.5 is used.

4.2.2. Initial experiment - Hybridisation of labelled probes to ssDNA complementary biotinylated probes in NaCl, Mg^{2+}

Next, I tested whether it is possible to identify ssRNAs using the ALEX method based strategy, I designed an experiment which measured a time course hybridisation of singly-fluorescent labelled ssDNA probe (R_{11}) to a construct T_{11}-G_{11} (Fig. 4.2*A*), formed by a complementary biotinylated ssDNA template T_{11} pre-hybridised with the probe G_{11} (Fig. 4.2*B*). Template T_{11} was a ssDNA designed on the 5' sequence of (+)ssRNA11, a segment which encodes NSP5 and NSP6 proteins. In this experiment 10 nM of R_{11} and T_{11}-G_{11} were mixed in 100 mM Tris-HCl pH 8, 9 mM, $MgCl_2$ buffer at room temperature for 1, 5 and 10min. The buffer conditions used mirrored the ionic strength of those used in RV transcription. After a variable hybridisation time (1 to 10 min), an aliquot was added to the slide, captured and imaged. No molecules hybridised with R_{11} are detected after 1 or 5 min (Fig. 4*C* and *D)*. However, after 10 min of incubation in hybridisation buffer, there is a population with an E*/S of 0.35/0.65 matching the E*/S signature of the positive control T_6-G_6-R_6 (Fig. 4.2 *H*), a fully hybridised T_6-G_6-R_6, construct was used which had been prepared according to standard annealing conditions.

Further experiments were also performed after an incubation time of 30 min under the same buffer conditions (Fig. 4.2*F*) or with the addition of 0.250 M NaCl (Fig. 4.2*G*). The results show a clear population with the same E*/S code as the positive control, the addition of NaCl to the buffer did not affect dramatically the amount of hybridisation of R_{11}. Taken together, the results of these experiments showed that the hybridised constructs can be detected after 10 min, the constructs had a

stoichiometry of ~0.65 slightly higher than the theoretical 0.5 but exactly the same as the fully hybridised positive control showing that hybridised constructs do not self-associate and the S value is most probably due to TIRF microscope settings.

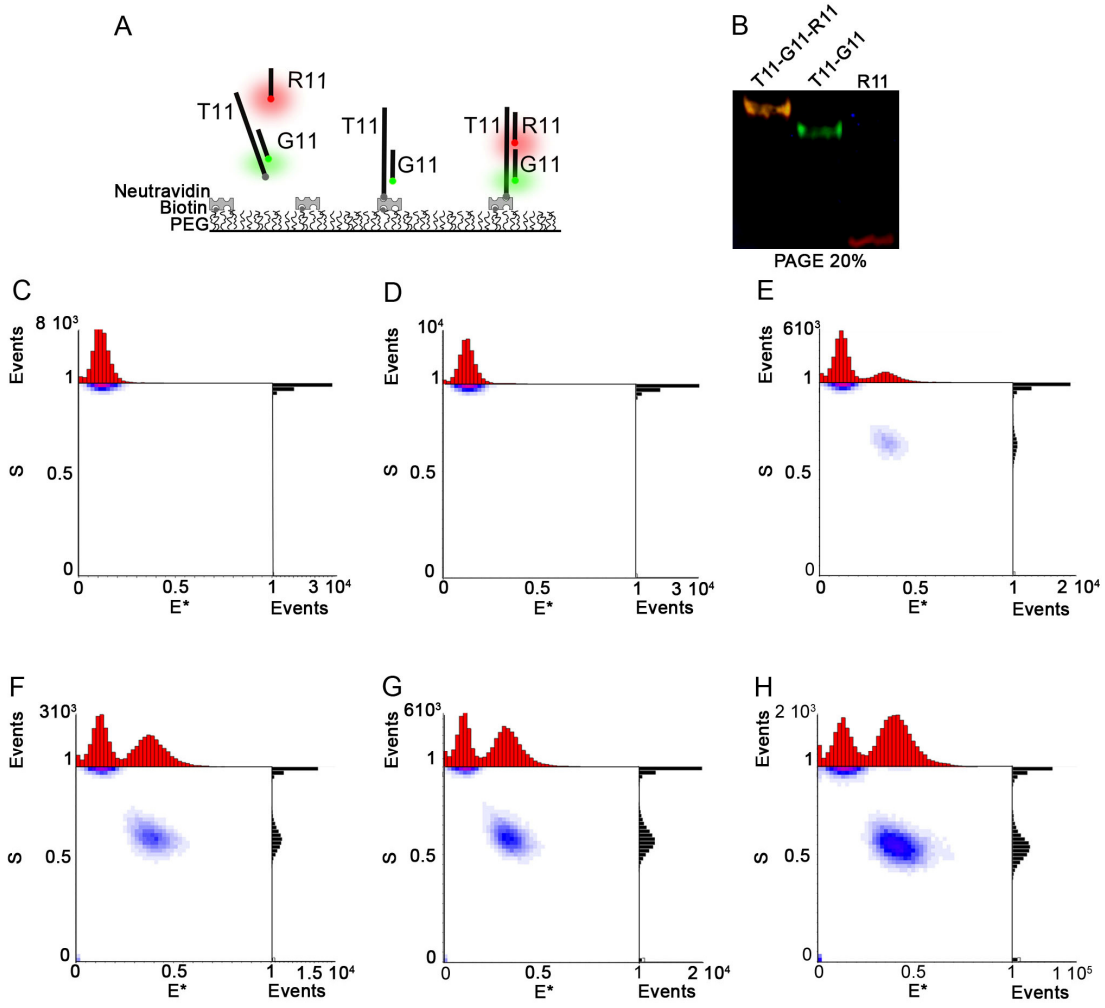


Fig. 4.2. Time course hybridisation of R_{11} to T_{11}-G_{11} construct detected using TIRF-ALEX.
A) Top left, schematic of capture of T_{11}-G_{11}-R_{11} in a TIRF slide. R_{11} were mixed with T_{11}-G_{11} in 100 mM Tris-HCl pH 8, 9 mM, $MgCl_2$ for 1 hour at room temperature. *B*) Top right. Fully hybridised T_{11}-G_{11}-R_{11}(orange), construct T_{11}-G_{11} (green) and probe R_{11}(red) were resolved in a 20% native PAGE. *C-E*) E*-S histograms for hybridisation for 1, 5, 10 min. The total number of particles displayed in the E*/S (*C-E*) histograms were 4949, 5536 and 4370 particles respectively. The expected E*/S code of ~0.35/0.65 is detected after 10 min. *F*) E*/S histogram of T_6-G_6-R_6 construct incubated for 30 min in the same buffer as C-E (total number of particles 3923) or *G* with an additional concentration of 250 mM NaCl (total number of particles 3540); after 30 min there is relative increase of hybridised constructs irrespective of NaCl present in the buffer. *H*) Positive control, T_{11}-G_{11}-R_{11} incubated under standard annealing and buffer conditions (see materials and methods Chapter 3). The E*/S code of

~0.35/0.65 for the control matches the signature for hybridisations in *E,F* and *G*. The E*-S histograms in all panels represent particles from ~40 movies.

4.2.3. Detection of ssRNA transcripts released in solution from DLPs using fluorimetry and confocal microscopy

The experiments described above proved that detection of specific hybridisations of R_{11} to a ssDNA construct is possible under concentrations and buffer conditions used in RV transcription experiments.

After this, to test whether ssDNA probes R_{11}, G_{11} hybridise to complementary (+)ssRNA$_{11}$ a set of single molecule and ensemble fluorescence experiments were performed.

In the first of this new set of experiments, ~10 nM of R_{11}, G_{11} and synthetic ssRNA$_{11}$ (see materials for description of synthetic ssRNA) were mixed with 100 mM Tris-HCl pH 8, 9 mM, $MgCl_2$, 150 mM NaCl, 6 mM DTT, for 1 hour at 37º C and aliquots diluted to ~50 pM was added to the slide and measured using confocal microscopy (Fig.4.3*A*).

Next, I tested whether R_{11}, G_{11} probes hybridise (+)ssRNA$_{11}$ in a sample containing the complete set of (+)ssRNA transcripts from DLPs from RV (Fig.4.3*B*).

The experiment was carried out in the same buffer conditions; the concentration of (+)ssRNA$_{11}$ was estimated to be approximately 10 nM, the same as the concentration of the synthetic (+)ssRNA template.

The results using a synthetic (+)ssRNA11 as a template showed one defined population with E*/S code of ~0.4/0.5. When the complete set of (+)ssRNA transcripts is used as a template, two main populations are detected (Fig. 4.3*B*), one with an E*/S code of ~0.4/0.5 like the one in Fig. 4.3*A* and an additional population

with an E* ~0.15-0.2 and S ~0.3-0.8 that corresponds to random coincidence of the labelled probes diffusing in the confocal spot of the microscope. This population without FRET is due to the high concentration of probes difussing in the confocal spot.

After this, an alternative ensemble experiment was carried out to test whether it is possible to detect hybridisations in real time of G_{11} and R_{11} to (+)ssRNA$_{11}$ transcripts being released by the DLP (Fig.4.3*C*).

For the ensemble experiment, I used slightly different conditions for the hybridisation experiments described above because of the less sensitive nature of the ensemble fluorimetry experiment. Specifically, I increased the concentration of probes; in this experiment 100 nM of each G_{11} and R_{11} probes were mixed with 4 nM DLPs in RV transcription buffer (53) (100 mM Tris-HCl pH 8.0, 4 mM ATP, 2 mM GTP, 2 mM CTP, 2 mM UTP, 9 mM $MgCl_2$, 150 mM NaCl, 0.5 mM S-adenosyl methionine (SAM), 0.5 U/µl RNase inhibitor).

The sample was illuminated with 532 nm laser light to excite the donor, and the emission spectrum recorded from 550 – 700 nm. Two main peaks were observed; the emission from the donor (corresponding to DexDem in the main peak at ~575 nm) and the emission from the acceptor (corresponding to DexAem) at ~660 nm. Then, the maximum height of the emission peak in each recorded spectrum was used to report on the progress of the reaction since this indicates the quenching of the DexDem due to FRET. The maximum height of the emission peak was fitted as an exponential curve (Fig.4.3*D*):

$$y = Ae^{-Kont} + c$$

where *A* is the initial absorbance; t is time (s); K_{on} is the on rate constant. (s^{-1}), c is a constant that corresponds to the size of the DexDem peak at very long times. In ideal conditions (if all the probes reacted), the value of C would correspond to that

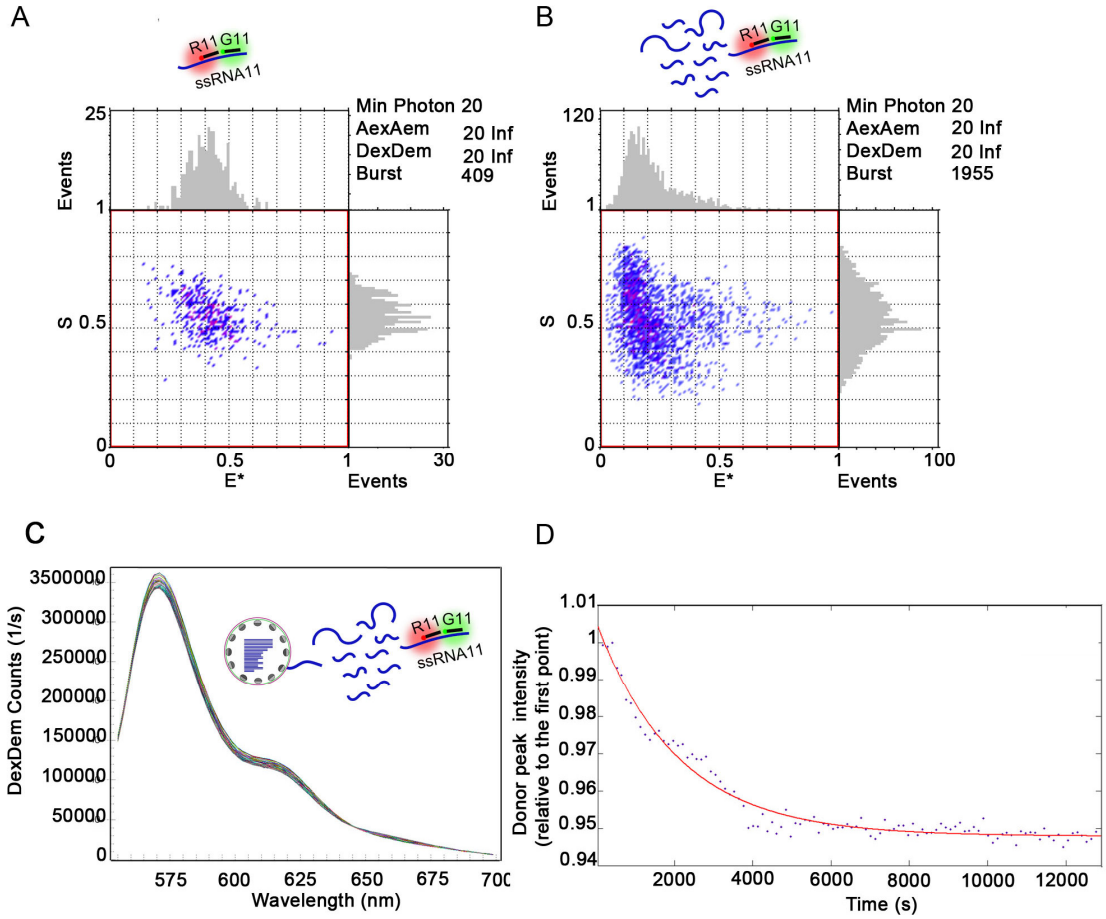


Fig. 4.3 Hybridisation of R_{11} and G_{11} to (+)RNA11 in solution detected using confocal-ALEX microscopy and fluorimetry.

A) G_{11}, R_{11} were mixed in a hybridisation buffer containing a synthetic ssRNA11 in 100 mM Tris-HCl pH 8, 9 mM, $MgCl_2$, 150 mM NaCl, 6 mM DTT, for 30 min at 37ºC. E*/S histogram shows a population (~409 particles assuming that each burst corresponds with a particle diffusing in the confocal spot only once) of ~0.4/0.5. *B*) G_{11}, R_{11} were mixed with a solution containing the complete set of (+)ssRNA transcripts released from transcribing DLPs. E*-S histogram shows two main populations (total number ~1955 particles), one population with an E*/S code of ~0.4/0.5; a second population with an E* ~0.15-0.2 and S ~0.3-0.8 that corresponds with random coincidence of the probes in the confocal volume. *C-D*) Hybridisations in real time of G_{11} and R_{11} to (+)ssRNA 11 transcripts being released by the DLP. For the single-molecule confocal microscopy experiment ~10 nM of each G_{11} and R_{11} probes were mixed with 10 nM of ssRNA11 in a transcription buffer and diluted to ~15 pM before being measured in the confocal microscope. For the ensemble experiment 100 nM of each G_{11} and R_{11} probes were mixed with 4 nM of DLPs in a transcription buffer (100 mM

Tris-HCl pH 8.0, 4 mM ATP, 2 mM GTP, 2 mM CTP, 2 mM UTP, 9 mM $MgCl_2$, 150 mM NaCl, 0.5 mM S-adenosyl methionine (SAM), 0.5 U/µl RNase inhibitor) at room temperature. Scans of fluorescence emission of spectrum were recorded every 2 min for 4 hours using a fluorimeter (FluoroMax-3 fluorimeter, Horiba Jobin Yvon). The saturation point is reached after ~34 min.

of a FRET of 0.4, although under empirical conditions it does not due to spectral cross talk, free dye etc.

Therefore, the decrease in DexDem is not due to photobleaching because it is concomitant with the emission at ~660 nm of the DexAem peak. A close examination of the spectra shows that the DexAem peaks go up as the other DexDem goes down as it is expected from FRET. Although FRET efficiency is not evaluated in this experiment, FRET measurement would have been possible using a method that takes the ratio of the DexDem peak and the DexAem peak. FRET measurements were not necessary for these experiments as the actual value of FRET is not very informative for the ensemble average measurements (leakage, spectral response of the fluorimeter, free dye were not corrected). Taking into account these considerations, the decrease in the DexDem peak (left peak) was a fair indicator to report on the progress of the hybridisation reaction. Under these experimental conditions, the saturation point is reached after ~34 min, it showed the characteristic time constant (tau), that corresponds with the inverse of K_{on}.

4.2.4. Description of CID assay. Proof of principle using synthetic ssRNAs from segment 11.

I have already shown supporting preliminary data of specific hybridisation to ssDNA, ssRNA using a single molecule hybridisation approach. Here I designed an experiment to fully validate the CID assay. To establish that CID identifies transcripts with high sensitivity and specificity, a synthetic RNA version of the

segment 11 transcript was targeted using capture probe C_{11} and FRET probes G_{11} and R_{11} (Fig. 4.4*A*). After hybridisation, surface-capture, and single-molecule imaging, the E*/S histograms showed a single population of molecules with the designed code of 0.35/0.5 (Fig. 4.4*A*). These results matched exactly those for a positive control, a dsDNA construct (T_{11}-G_{11}-R_{11}) wherein the ssRNA11-specific FRET probes hybridise to a complementary biotinylated DNA (T_{11}) to produce the 0.35/0.5 code (Fig. 4.4*B*). As a negative control, I tested for any non-specific binding of the transcript from segment 11 to a probe set (G_6, R_6 and C_6) complementary to a transcript from segment 6; only 2 particles were captured, and none with the correct code (Fig. 4.4*C*).

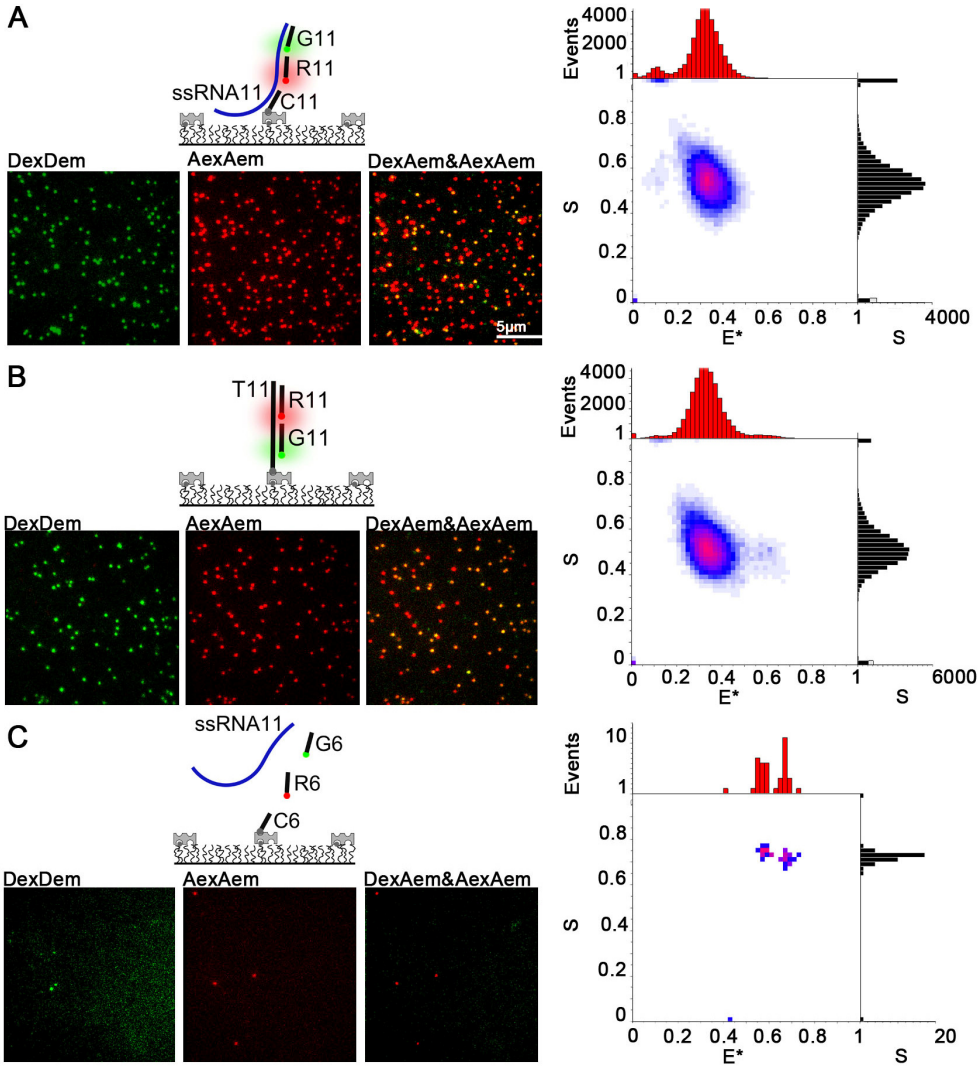


Fig. 4.4. CID detects synthetic transcripts from segment 11. *A*) Top left panel: schematic showing a synthetic (+)ssRNA$_{11}$ hybridised to FRET probes (G$_{11}$ and R$_{11}$) and capture probe C$_{11}$ on a neutravidin-coated slide. Bottom left panels (left to right): fields of view (scale bar 5µm) showing particles detected with donor emission upon donor excitation ($D_{ex}D_{em}$; green-coloured particles); second panel particles detected with acceptor emission upon acceptor excitation channel ($A_{ex}A_{em}$; red-coloured particles); third panel showing FRET ($D_{ex}A_{em}$) channel and acceptor emission upon acceptor excitation channel, referred to as the "$D_{ex}A_{em}$&$A_{ex}A_{em}$" panel; in this channel, a colocalised G and R probe appears yellow, where a free R probe appears red. Right panel: the synthetic (+)ssRNA11 using FRET probes (G$_{11}$ and R$_{11}$) and capture probe C$_{11}$ gives an E*/S histogram (n=1895 particles) with the expected E*/S code. *B)* Positive control using T$_{11}$-G$_{11}$-R$_{11}$, a dsDNA construct made of G$_{11}$ and R$_{11}$ hybridised to a complementary biotinylated ssDNA (T$_{11}$) matching a sequence in (+)ssRNA11. The E*/S histogram (n=2043) shows the E*/S code of 0.35/0.5. *C*) Negative control does not show a specific signal when incubating (+)ssRNA11 with FRET probes G$_6$, R$_6$ and capture probe C$_6$ targeting (+)ssRNA6. Representative FOVs of DexDem, AexAem and the overlay DexAem and AexAem channels are shown for each experiment. For display purpose only the contrast in the overlay image was increased in this figure.

4.2.5. Capture of released transcripts from DLPs using CID

I then tested whether our CID assay could detect both extruded and released transcripts from transcriptionally active DLPs. Specifically, I monitored the

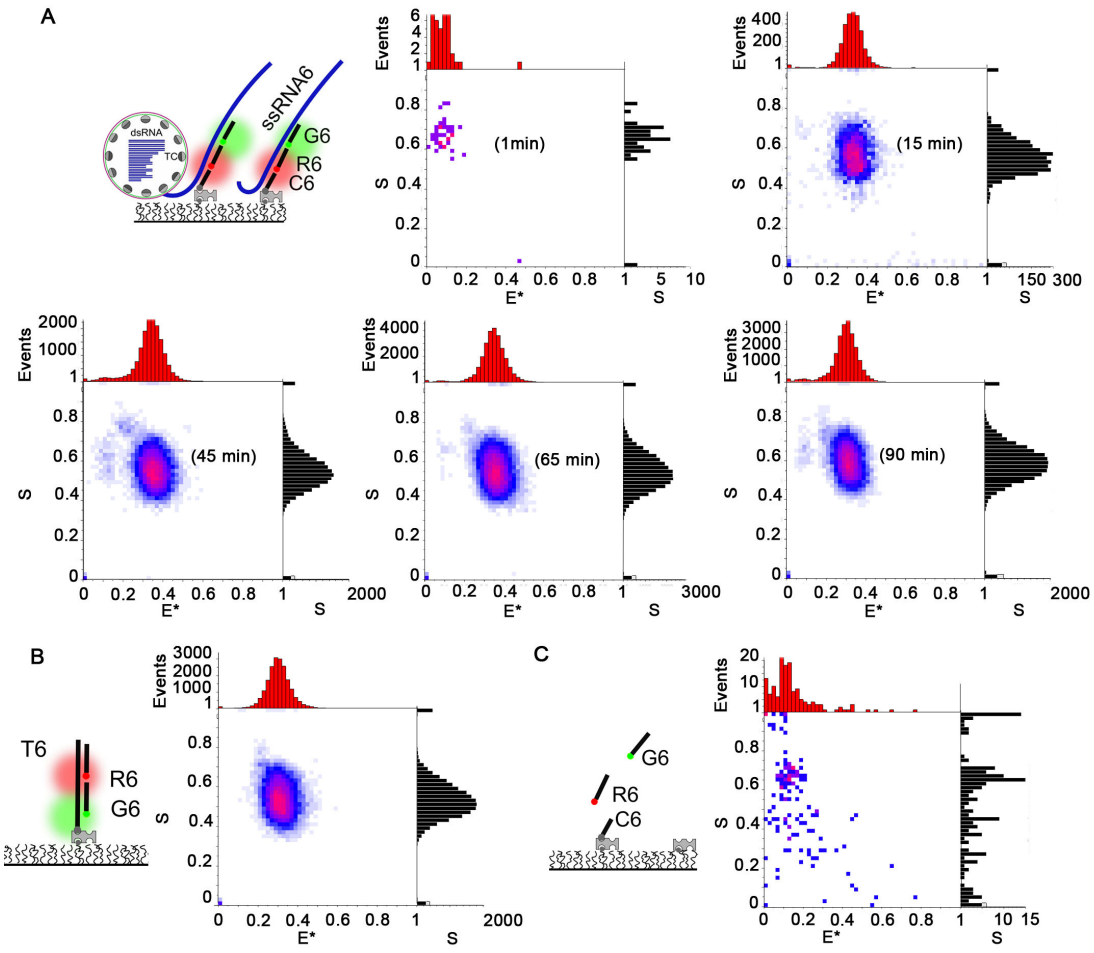


Fig. 4.5. Detection of nascent transcripts from transcriptionally active DLPs. *A*) Top left, schematic of the capture of (+)ssRNA6 transcripts either in solution or while extruded from DLPs. DLPs were mixed in a transcription buffer containing nucleotides, FRET probes G_6, R_6 and a capture probe C_6 for 2 min before transcription was blocked. RNA synthesis in the DLPs is very fast, so the majority of (+)ssRNA6 captured on the surface corresponds to (+)ssRNA6 in solution. Only a minimal number of DLPs are expected to be attached on the surface. Top right and bottom panels: E*-S histograms for hybridisation for 1, 15, 45, 60 and 90 min. The expected E*/S code of 0.35/0.5 is detected after 15 min (159 particles). The number of particles increased with time (45 min: 867 particles; 65 min: 1402 particles); the decrease at 90 min (1171 particles) is due to the surface saturation with transcripts and a corresponding increase in the number of elliptical particles rejected during image analysis (134). The E*-S histograms in all panels represent particles from 40 movies. *B*) Positive control using construct T_6-G_6-R_6 (designed as the positive control in Fig. 4.5*B*). The E*/S code of 0.35/0.5 for the control matches the signature for transcripts in Fig. 4.5*A*. *C*) Negative

control incubating probes G_6, R_6 and C_6 in the absence of (+)ssRNA6. No specific hybridisation is observed.

synthesis of (+)ssRNA6, which encodes VP6 (155); this was performed by mixing DLPs with a transcription buffer with probes specific for segment 6 (G_6, R_6 and C_6) and an E*/S code of 0.35/0.5.

After transcription for 2 min, the reaction was blocked, and transcripts were hybridised for up to 90 min (Fig. 4.5*A*), captured and imaged. After 15 min of hybridisation, the ssRNA6 transcripts were clearly observed (main population in Fig. 4.5*A*, 15-min panel), matching the signature of the positive control (T_6-G_6-R_6, Fig. 4.5*B*); the number of detected ssRNA6 transcripts increased with hybridisation time. All populations had a stoichiometry of ~0.5, showing that ssRNA transcripts do not self-associate to higher-order structures upon their release from DLPs.

To test for any probe association with DLPs in the absence of transcription, we performed the same hybridisation experiment but in the absence of nucleotides, in that experiment (negative control) the data analysis did not retrieve any particle with association of G and R labelled probes, demonstrating that the probes do not interact with the DLPs or their dsRNA segments in the absence of transcription; there was also no detectable probe self-association (Fig. 4.5*C*).

These results established that a high level of transcription can be detected from DLPs even after a short incubation in transcription buffer. Further experiments to capture synthesised ssRNA11 from RNA transcribed from DLPs also showed the specific E*/S code (Fig. 4.6).

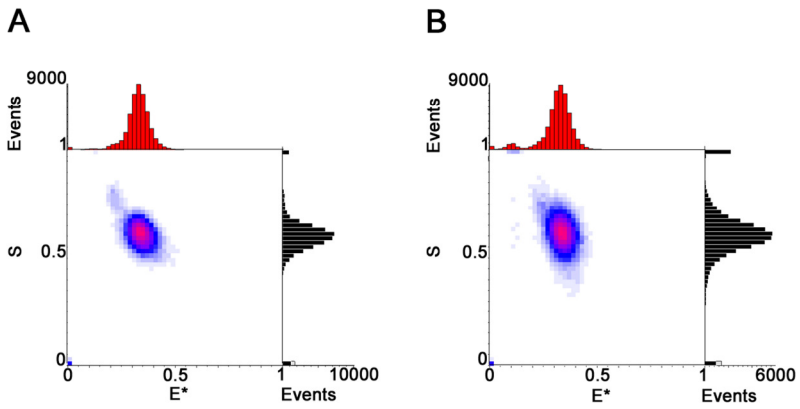


Fig. 4.6. CID assays for ssRNA11 and ssRNA6 transcripts produced by rotavirus DLPs. ssRNAs were incubated with capture and FRET probes, captured on the surface and detected using ALEX-TIRF. *A*) (+)ssRNA11 detected using G_{11}, R_{11} FRET probes, and capture probe C_{11}. Combined E*/S histogram data of 2553 detected particles with an E*/S signature of 0.35/0.55, matching well the code of 0.35/0.5. *B*) (+)ssRNA6 detected using G_6, R_6 FRET probes, and capture probe C_6. Combined E*/S histogram data of 2696 detected particles with an E*/S signature of 0.35/0.6, matching well the code of 0.35/0.5.

4.3. Discussion

The results from preliminary experiments showed in this chapter, support that this single molecule hybridisation approach using ssDNA enables the identification of ssDNA and ssRNA sequences by exploiting the complementarity of base pairing between the probe and the target sequence. This conclusion is based on the identification of a defined population with an E*/S code that matches the E*/S value of a fully hybridised positive control, and the results are further supported by confocal and fluorimetry data. Although slight discrepancies in E*/S code are observed between TIRF experiments and the confocal results, it can be explained by the differences between TIRF-ALEX and the confocal microscope set up. Under the experimental conditions used in those experiments, hybridisation of ssDNA with ssRNA is not hampered by the chemical differences in nucleotides and the differences in persistence length of the RNA and DNA sequences (156). Although the preliminary hybridisation studies were only designed as orientative results to

evaluate the feasibility of the CID assay, they provide some interesting observations, the magnesium based buffer was sufficient to facilitate efficient probe hybridisation, hybridisation was detected at 37 ºC suggesting that under these conditions the melting temperature of the hybrids is still stable at conditions used in transcription of DLPs, and the hybridisation rate may not have been affected substantially by temperature as reported in other studies (157). After this set of experiments, I validated the CID assay for detection of specific (+)ssRNA from RV. The CID assay is a novel approach for studying the synthesis of specific transcripts from Rotavirus particles. The capture and identification assay has a simple design and allows unambiguous identification and quantification of specific ssRNAs. The CID assay was applied to capture and quantify (in the experiment) the capture of released specific transcripts from DLPs, the active transcribing rotavirus particles. Quantification of total number of transcripts among experiments is potentially possible with the implementation of internal hybridisation standards. The implementation of biotinylated capture probes introduce the two fold advantage of concentrating the hybridised constructs on the slide, increasing the signal, and at the same time, it allowed me to wash and remove the excess of unbound probes present in the solution which minimises the fluorescence background. The assay relies on an E*/S code and can identify RNA targets with high specificity faster (10-15 min) than other detection tools such as PCR based methods, and with very low false positives, a task difficult for standard hybridisation techniques that are often characterised by non-specific hybridisation due to use of either multiple short probes or very long probes (158, 159). The CID assay also circumvents the need for complex multi-colour labelling (160) that may complicate signal interpretation due to crosstalk between overlapping emission spectra. The CID assay can be easily

extended to study other members of the *Reoviridae* family, RNA viruses and various transcription systems. Its simple design and reliability may also make it a useful diagnostic tool. The combination of this method with other tested techniques such as super-resolution switchable FRET enables the localisation and quantification of transcripts in the cell, if so this method may open an avenue to dissect the role of (+)ssRNA transcripts in translation and packaging.

Chapter 5

Investigation of mRNA Release Models in Double-layered Viral Particles of Rotavirus

5.1. Introduction

DLPs have a set of unique channels that connect the internal space of the DLP with the outside. The structure of the channels is known. In RV there are 12 type I channels, each placed at the 5-fold vertices of the icosahedral capsid. The channels traverse the layer and the VP6 and VP2 proteins form 12 small pores along the 5-fold axis of the icosahedron and serve as RNA-exit channels for the newly transcribed positive-sense single-stranded RNA (+ssRNA). Close to these channels, 12 TCs interact with the viral genomic segments and attach to a hub-like structure formed by VP2 (24, 27). This organisation implies that one of the 12 TCs could be unoccupied (14). `

Rotavirus RNA synthesis is thought to be asynchronous (79), with some transcripts produced earlier than others (161). It has also been shown that transcripts are released simultaneously through the 12 RNA-exit channels (162), but the assignment of specific segments to channels (if any) is unknown, since standard microscopy can only visualise the presence, but not the identity, of the extruded segments per single particle.

Currently the DLP structural organisation is compatible with two mechanistic models for transcript extrusion (Fig. 5.1). The channel-specialisation model (Fig. 5.1*A*) postulates that each channel is linked to a specific TC and to transcripts of only one specific genomic segment (138); this model predicts that, at any given time, only one transcript of a specific segment can be extruded per channel of a DLP. In contrast, another theoretical model called the channel-generality model can be proposed (Fig. 5.1*B*). The channel-generality model predicts that each channel can extrude transcripts of multiple (perhaps all) genomic segments, but only one at a

time. This model also predicts that, at any given time, a genomic segment may interact with several TCs, thus some particles carry two or more transcripts of a specific segment extruded from two or more channels. Hybrid versions of the two main models are also possible.

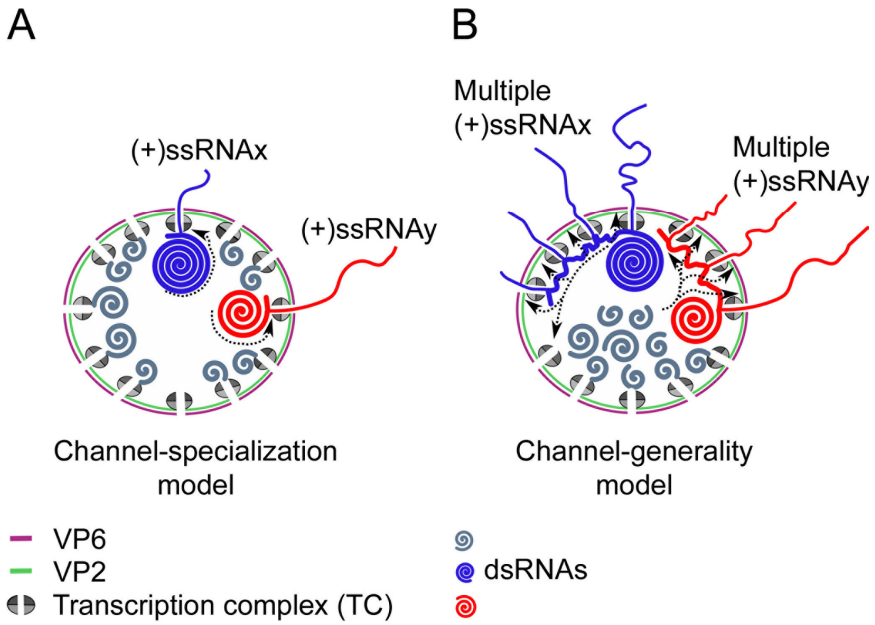


Fig. 5.1. Models for transcript exit in rotavirus. *A*) Channel-specialisation model: each RNA-exit channel is linked to a specific TC and transcripts (thin blue and red lines for segments x and y, respectively) of a single specific genomic segment (thick blue and red lines for segments x and y, respectively). This model predicts that only a single transcript of each viral genomic segment is extruded from each DLP at any given time. *B*) Channel-generality model: a segment can interact with multiple TCs and their corresponding RNA-exit channels during transcription. This model predicts that multiple transcripts from each genomic segment can be extruded from each DLP at any given time.

The answer to this question is important, since it will provide insight into our basic knowledge about the regulation of transcription in RV, how the segments are packed in the confined space of the DLP and as result contribute to new intervention approaches to block the replication cycle of the virus.

To test these models, the copy number of specific transcripts was determined during their extrusion from transcriptionally active DLPs using the CID assay validated in chapter 3. CID has the advantage of maintaining the transcriptional activity of viral particles, in contrast to approaches that may induce conformational changes in the channels, such as the use of anchor antibodies (163, 164) or labelling methods that target surface amines or cysteines. The findings presented in this chapter support the "channel-specialization" model and exclude alternative models in which an individual segment is transcribed by several complexes and extruded through multiple channels. The method can be implemented in the sensing of transcripts in other biological model systems.

5.2. Results

5.2.1. Capture of extruded transcripts (+)ssRNA 11 from DLPs using CID

Here, I applied the CID assay to investigate specific transcripts during their extrusion from DLPs (and prior to their release) by capturing DLPs with transcripts of segment 11 extruded simultaneously (Fig. 5.2). To detect such DLPs, transcription was allowed for 1 min (see materials and methods – Chapter 3) during which DLPs should produce 10-20 transcripts of each segment (assuming transcription at ~50 nt/s (79)), with some of the DLPs with transcripts still attached to a DLP. After fixing the transcribing DLPs, transcripts were hybridised to an ssRNA6 capture probe and ssRNA11 FRET probes; captured DLPs (311 particles) showed the expected code for a DLP carrying ssRNA11 transcripts (Fig. 5.2*A*), and matched exactly the code of the positive control (T_{11}-G_{11}-R_{11} construct; Fig. 5.2*B*).

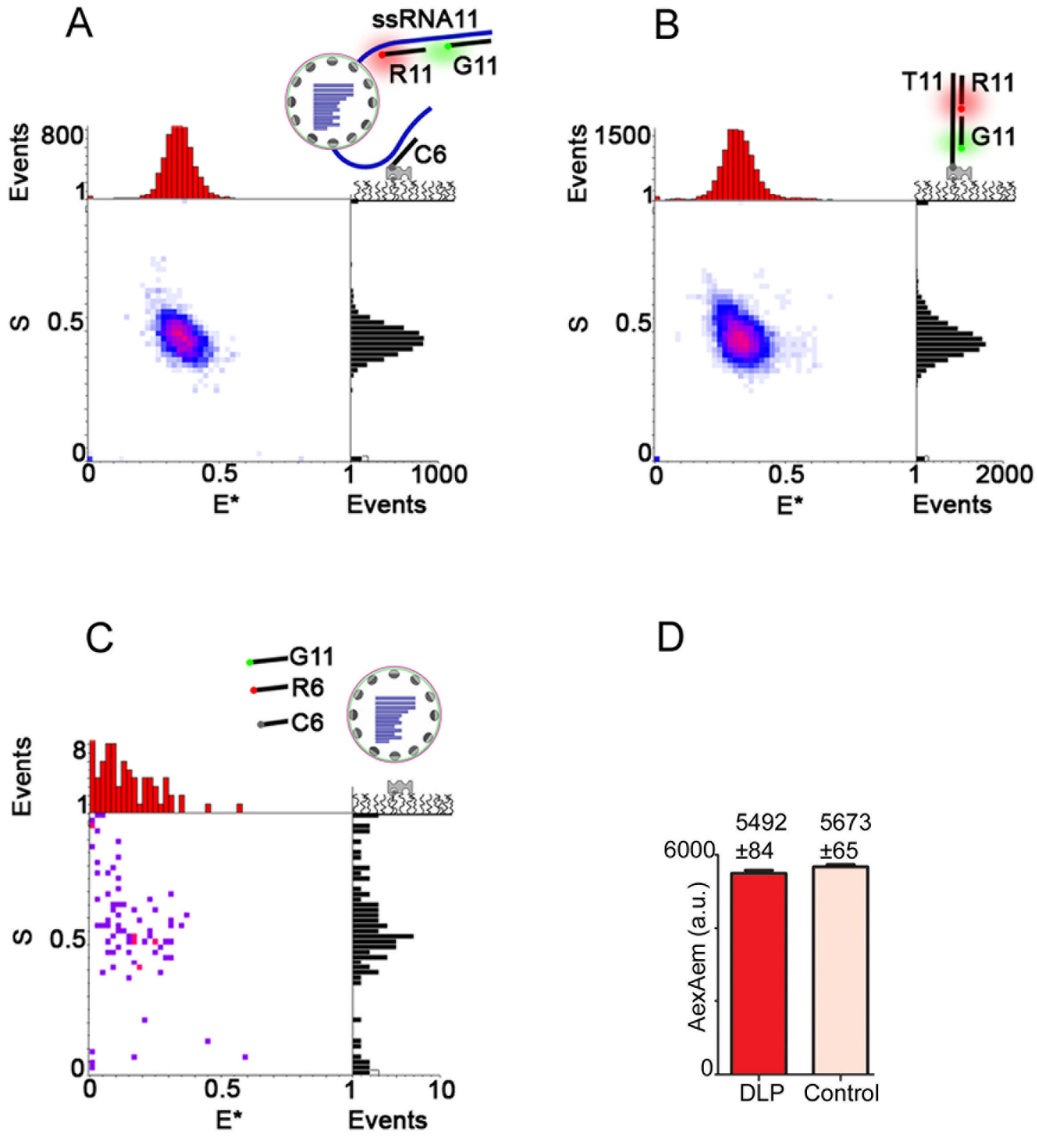


Fig. 5.2. Transcriptionally active DLPs carry a single extruded transcript for segments 6 and 11. *A*) CID of actively transcribing DLPs (allowed to transcribe for 2 min) interrogated using probe C_6 (which targets extruded (+)ssRNA6) and probes G_{11} and R_{11} (which target extruded (+)ssRNA11). The E*/S histogram (n=311) shows an E*/S population of 0.35/0.5, matching that of the positive control (panel *B*). *B*) Positive control. The E*/S histogram (n=640) for construct T_{11}-G_{11}-R_{11}. *C*) Negative control. E*/S histogram (n=4) from DLPs incubated under the same experimental conditions but without nucleotides show no specific signal. *D*) Comparison of $A_{ex}A_{em}$ fluorescence intensities (mean ± SEM from experiments in panels *A* and *B*). Hybridised DLPs, red bar; positive control using T_{11}-G_{11}-R_{11}, pink bar. A t-test comparison between the results shows no significant differences in the mean (p value > 0.10).

In a negative control (no nucleotides; Fig. 5.2*C*), only 4 molecules were captured and none with the code. These results showed that DLPs were captured through an

extruded ssRNA6 transcript, while carrying a yet undetermined number of extruded transcripts of segment 11. However, the S value of ~0.5 for the main population, and the absence of other populations of hybridised particles on the E*/S histogram (Fig. 5.2*A*), excluded relative stoichiometries of the two FRET probes other than 1:1, and further suggested the presence of only a single extruded ssRNA11 per DLP. There is a presence of a very small population with E* and S values which are slightly different from the main population but they are observed in both the experiment and the positive control. This can be attributed to different causes including particles attached to randomly close neutravidin particles, or a high density of the particles in the slide due to variation in pegylation efficiency among individual slides, or fluorophore blinking.

I tested independently the stoichiometry of two standards with different stoichiometries. I hypothesised that if a significant number of DLPs carried two extruded ssRNA11 transcripts, the bleaching or sub-stoichiometric hybridisation for some particles would have resulted in populations with 2:1 or 1:2 G:R stoichiometry (Fig. 5.3), and in broadening of the S distribution relative to the 1:1 positive control. In contrast, the S distribution for the captured DLPs matches exactly that of the 1:1 positive control (cf. black distributions in Figs. 5.2*A* and 5.2*B*). To confirm the presence of a single extruded transcript of ssRNA11 per DLP, we compared the DLP-based acceptor emission intensity upon acceptor excitation ($A_{ex}A_{em}$, a FRET-independent measure of the number of acceptor fluorophores per particle; using data from Fig. 5.2*A* with that of the positive control (T_{11}-G_{11}-R_{11}) construct, with an absolute G:R stoichiometry of 1:1; Fig. 5.2*B*). For a single extruded transcript per DLP, the mean $A_{ex}A_{em}$ intensity for the captured DLPs should match that of the 1:1 control; however, if the DLPs carry two (or more) extruded transcripts for

ssRNA11, the mean $A_{ex}A_{em}$ for DLPs should be 2-fold higher (or more) than that of the 1:1 control. (see Fig. 5.3*C* and 5.3*G* for the $A_{ex}A_{em}$ distribution for standards with G:R stoichiometries of 1:1 and 1:2, respectively; see also Fig. 5.3*J* for the mean $A_{ex}A_{em}$ of the same standards). The results showed no differences between the

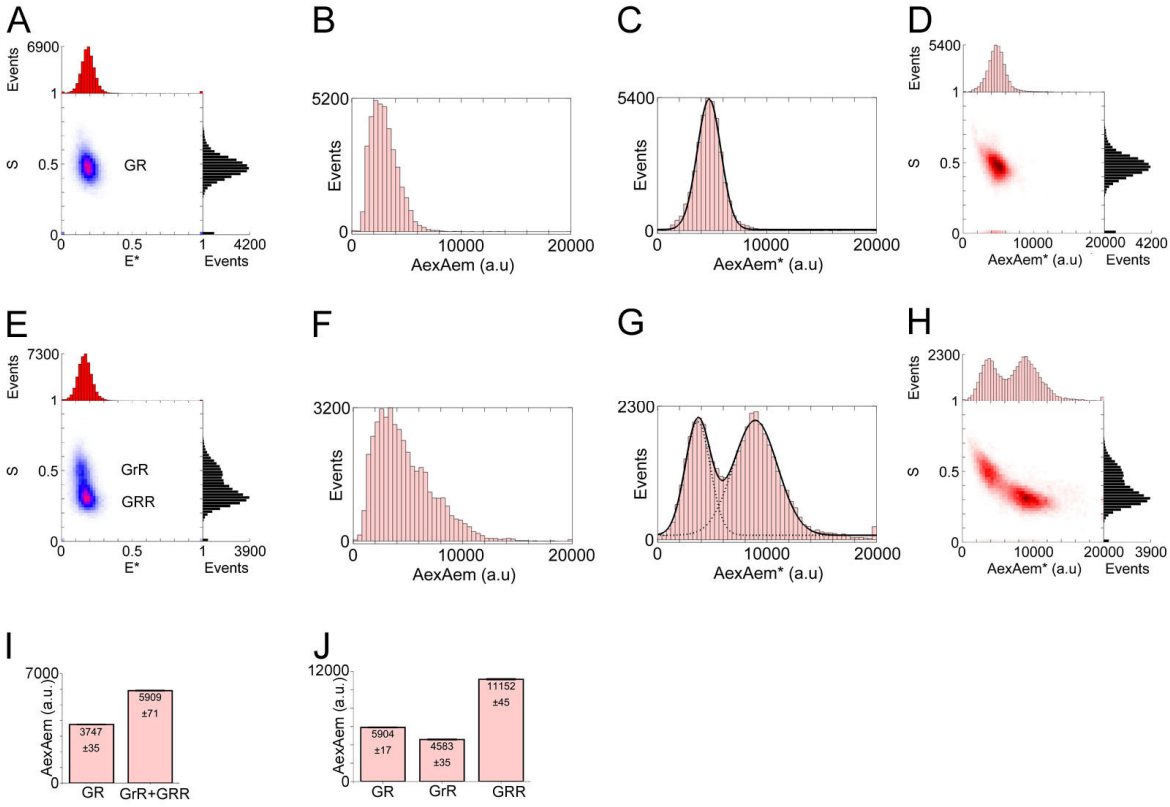


Fig. 5.3. Standards analysed using E*-S histograms and fluorescence intensities. *A*) E*/S histogram of dsDNA construct G_6-R_6-$T_{(6\text{-}11)}$ (a GR standard), compiled from 1959 particles. The long fluorophore separation (40 bp) leads to negligible FRET. The population at S≈0.5 has the signal expected from constructs carrying one red and one green fluorophore. *B*) Histogram of $A_{ex}A_{em}$ intensities for the GR standard (mean ± SEM: 3747±35 a.u.). *C*) As in panel B, but corrected for non-uniform illumination (corrected emission intensities, $A_{ex}A_{em}$*). The fitted Gaussian function was centred at 5904±17a.u.*D*) S/$A_{ex}A_{em}$* histogram of the GR constructs shows that the well-defined population of S≈0.5 corresponds to a well-defined population of $A_{ex}A_{em}$* intensity centred at 5909a.u. *E*) E*/S histogram of dsDNA construct G_{11}-Sp-R_{11}-R_{11}-$T_{(11^{th})}$ (a GRR standard). The long fluorophore separation (47 bp) leads to negligible FRET. The histogram (2247 particles) shows a major population at S≈0.3 and a minor at S≈0.5, consistent with a main population of fully hybridised constructs (GRR) and a minor population with a single R probe (GrR, photophysically equivalent to the GR standard). *F*) Distributions of $A_{ex}A_{em}$ intensities (5909±71 a.u.) from GRR samples with 26% of GRR particles showing intensities higher than 7500 a.u. (cf. to 1.5% of GR in panel B).*G*) Distributions of $A_{ex}A_{em}$* intensities, revealing two well-resolved populations centred at

4583±35 and 11152±45 a.u., both fit to Gaussian functions. The lower intensity for GRR compared to the GR standard may indicate fluorophore interactions with the non-hybridised portion of DNA. *H*) Two distinct $A_{ex}A_{em}$* populations can also be visualised on the S*/$A_{ex}A_{em}$* histogram. *I* and *J*) Mean values and Gaussian fit results with error estimates.

mean $A_{ex}A_{em}$ intensity per captured DLP, and the control (Fig. 5.2*D*; 5492±84 counts for DLPs; 5673±65 for the 1:1 control), leading me to conclude that only a single ssRNA11 transcript was present on the surface of each DLP during transcription. Overall, the comparison between the mean $A_{ex}A_{em}$ intensity per captured DLP and the control strongly suggest the presence of a single R fluorophore per DLP. In these experiments, the fluorescence background was very low. The accuracy of the AexAem mean intensity measurement could potentially be improved by using standard epifluorescence microscopy rather than TIRF illumination which depends on the z position in the evanescent field.

The results of the measurement using these hybridisation standards could also be improved by the analysis of more complex constructs consisting of GGRR. GGRR could have been a more realistic control, however, proof of principle experiments that demonstrate the capability of ALEX to resolve populations with different stoichiometries has already been published (165) and the GGRR hybridisation standard would have been easy to distinguish from the 1:1 stoichiometry because it would have resulted in a distinct pattern of multiple populations due to partial hybridisation.

5.2.2. Capture of extruded transcripts (+)ssRNA 6 from DLPs using CID

To examine whether the extrusion of a single transcript per segment on transcribing DLPs applied to other segments, DLPs were targeted by swapping the capture and FRET probes used previously, i.e. a capture probe we used for ssRNA11, and FRET

probes for ssRNA6. Hybridisation was performed under the same conditions as those carried out to capture ssRNA6. To improve DLP capture, an additional capture probe (C_{10}) was used, which targeted ssRNA10; the introduction of additional capture probes was used to increase the probability of detecting hybridisation events. The results also showed the same ssRNA6-specific signal (Fig. 5.4*A;* 322 particles). The presence of a single capture probe (C_{11}) under these hybridisation conditions resulted in a specific signal but less efficient capture of transcripts (38 particles, Fig. 5.4*B*). The detected signal in our experiment is specific, the E*/S value is the same as the positive control (matching the signal for released ssRNA6; Fig. 5.4*C;* 1832 particles); the negative control (FRET probe without capture probes) showed only 8 particles attached to the TIRF slide but they were not showing any specific E*/S signal of ~0.5/0.35; Fig. 5.4*D*. Finally, as with ssRNA11, (Fig. 5.2*E* the mean $A_{ex}A_{em}$ intensity for probe R_6 hybridised in DLPs matched that of the positive control (3915±72 counts for DLPs; 4000±37 for the 1:1 control), suggesting that the captured DLPs carry only a single extruded ssRNA6 transcript (Fig. 5.4*E*).

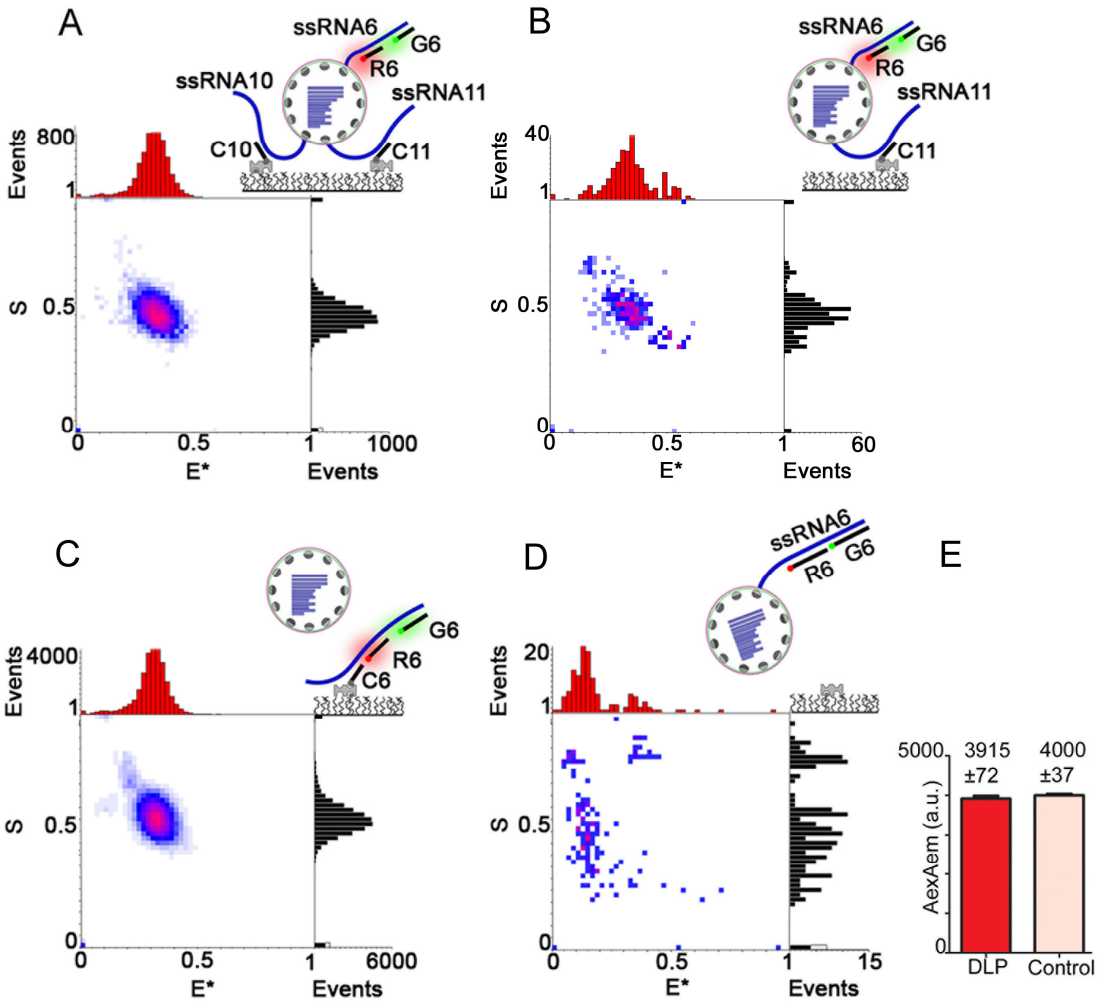


Fig. 5.4. Transcriptionally active DLPs carry a single extruded transcript for segments 6 and 11. *A*) CID of actively transcribing DLPs (allowed to transcribe for 2 min) interrogated using probes C_{10} and C_{11} (which target extruded (+)ssRNA10 and (+)ssRNA11, respectively) and probes G_6 and R_6 (which target extruded (+)ssRNA6). The E*/S histogram (n=322) shows an E*/S population of 0.35/0.5, matching that of the positive control (panel *C*). *B*) CID of actively transcribing DLPs interrogated using probe C_{11} only and probes G_6 and R_6 .The E*/S histogram (n=322) shows an E*/S population of 0.35/0.5, matching that of the positive control (panel *C*). *C*) Positive control. The E*/S histogram (n=1832) for captured (+)ssRNA6 (synthesised under the same conditions as in panel *E*, and detected using C_6, G_6, and R_6) shows the expected E*/S code of 0.35/0.5. *D*) Negative control performed with FRET probes but in the absence of a capture probe do not show any specific hybridisation (n=8). *E*) Comparison of $A_{ex}A_{em}$ fluorescence intensities (mean ± SEM from experiments) in panels *A* and *C*). Hybridised DLPs, red bar; positive control using (+)ssRNA6, pink bar. A t-test comparison between the results shows no significant differences in the mean (p value > 0.35).

5.2.3. Capture of extruded transcripts (+)ssRNA 2 from DLPs using CID

Additional experiments aimed at the detection of a third rotavirus segment (ssRNA2) also showed the extrusion of a single transcript (Fig. 5.5).

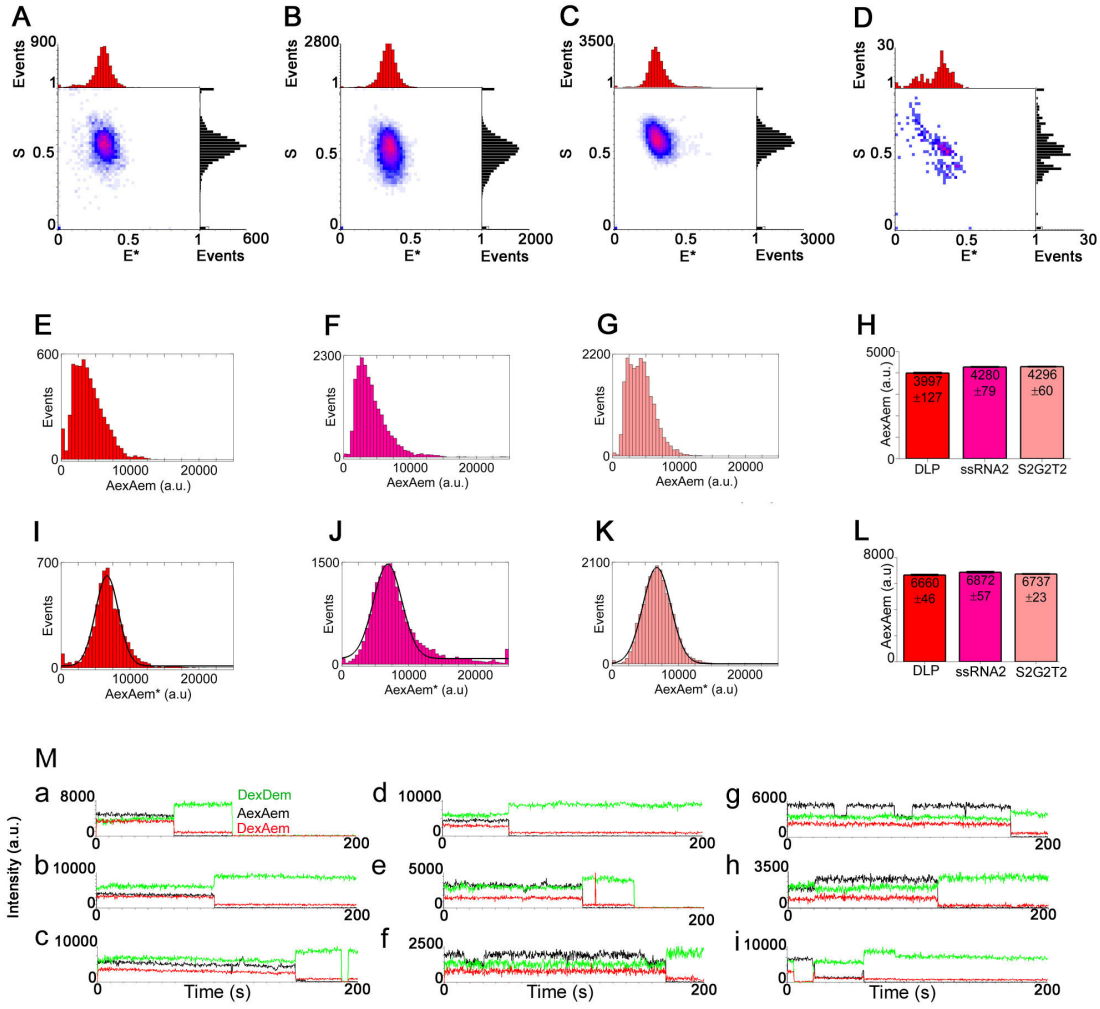


Fig. 5.5. Transcriptionally active DLPs carry a single extruded transcript for segment 2.

A) Actively transcribing DLPs (allowed to transcribe for 2 min) interrogated using capture probes C_6 and C_{11} and FRET probes G_2 and R_2 (which target extruded (+)ssRNA2). The E*/S histogram (292 particles) shows a population of 0.35/0.5, matching the positive control (panels *B,C*). *B*) Positive control. E*/S histogram (998 particles) of synthetic ssRNA2 hybridised with G_2 and R_2, and captured on the surface with capture probe C_2. *C*) Positive control. E*/S histogram (1079 particles) of G_2-R_2-T_2 dsDNA construct which consists of one G_2 and one R_2 probe hybridised to biotinylated T_2 ssDNA. *D*) Negative control. E*/S histogram (11 particles) experiment performed with DLPs, FRET probes but without capture probes shows very few particles attached to the surface. (*E-G*) Distributions of average $A_{ex}A_{em}$ intensities from the sample in panels *A (red)*, *B (dark pink, positive control)*, and *C (light pink, positive control)*. *H*) Comparison of mean $A_{ex}A_{em}$ from ssRNA2 extruded in DLPs

shown in *E* (red), captured ssRNA2 (*F,* dark pink, positive control) and dsDNA control (*G,* light pink). *I, J, K*) Corrected $A_{ex}A_{em}$* intensity distributions from panels *A, B,* and *C*, respectively, along with single-Gaussian fits. *L*) Mean $A_{ex}A_{em}$* intensity from the Gaussian fits for DLPs (panel *E*), ssRNA2 (panel *F*) and dsDNA control (panel *G*). *M*) Photobleaching examples of DLPs captured with extruded ssRNA2 (39 particles). CID was performed as in panel *A* but taking longer movies (2000 frames, 100 ms frames in a, b, c, d, e, i and 50 ms frames in f-h,). Approximately 90% of the particles showed a single photobleaching step in the acceptor channel, while 10% displayed complex photophysics (see panel i).

These results establish that for segments 2, 6 and 11, at most one single transcript of each of these rotavirus genes is being extruded from a DLP at any given time, and argue for a general mechanism in which at most one single transcript of any segment is synthesised and released from a DLP at any given time.

To confirm that the detected ssRNAs were colocalised with DLPs, we performed CID experiments using antibodies for DLP capture (Fig. 5.6) instead of DNA capture probes. Specifically, we captured transcribing DLPs using a sheep antibody targeting VP6 (the protein covering the DLP surface with 780 copies) and a rabbit biotinylated antibody against sheep IgG; the DLPs were probed for the presence of extruded ssRNA2 (Fig. 5.6*A*). The E*/S histogram for the captured DLPs (Fig. 5.6*B*) displays a single population (n=120) with the ssRNA2-specific E*/S code of ~0.35/0.5, thus matching the results for DLPs captured using DNA-based capture probes (Fig. 5.5). This single population matches the E*/S populations in two positive controls: probed ssRNA2 released from DLPs (Fig. 5.6*C*), and probed synthetic ssRNA2 (Fig. 5.6*D*), both captured using the C2 probe. Furthermore, the mean $A_{ex}A_{em}$ intensity for probe R_2 hybridised on antibody-captured DLPs matches

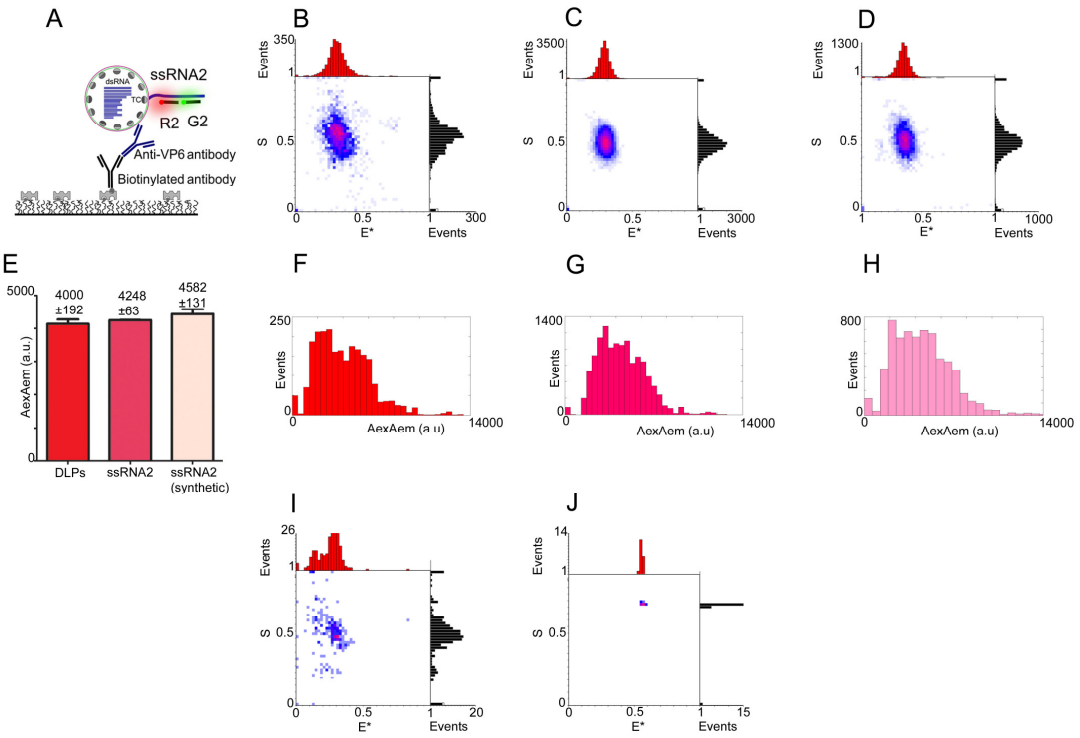


Fig. 5.6. Antibody-captured transcriptionally active DLPs carry a single extruded transcript for segment 2. *A*) Schematic of the antibody-based CID assay. Actively transcribing DLPs (allowed to transcribe for 2 min) were interrogated using FRET probes G_2 and R_2 (which target (+)ssRNA2). After incubation with a VP6-specific sheep antibody, the DLPs were added to a coverslip coated with a biotinylated rabbit anti-sheep antibody. For details, see materials and methods, chapter 3. *B*) CID of antibody-captured actively transcribing DLPs interrogated using probes G_2 and R_2. The E*/S histogram (n=120) shows a single E*/S population of 0.35/0.5, matching the main populations of the positive controls (panels *C* and *D*). *C*) Positive control. The E*/S histogram (n=879) for (+)ssRNA2 transcripts captured from transcribing DLPs using FRET probes R_2, G_2 and capture probe C_2 shows a single E*/S population of 0.35/0.5. *D*) Positive control. The E*/S histogram (n=414) for synthetic (+)ssRNA2 transcripts captured using the same probes as in *C* shows a single E*/S population of 0.35/0.5. *E*) Pairwise comparison of the $A_{ex}A_{em}$ fluorescence intensities (mean ± SEM) from experiments in panels *B* (red bar), *C* (dark pink bar) and *D* (light pink bar). A one-way ANOVA using a Tukey's test shows no significant differences between DLPs and the ssRNA control (p-value > 0.05), supporting that only one ssRNA2 is extruded from the DLP at any given time.

F,G,H) Distributions of $A_{ex}A_{em}$ fluorescence intensities shown in E from experiments in panels *B* (red), *C* (dark pink) and *D* (light pink). *I*) Negative control. A CID assay performed as in panels *A* and *B,* but with the primary antibody being substituted by rabbit IgG (which does not bind to VP6). The E*/S histogram shows only few particles (n=12) with an E*/S population of 0.35/0.5. *(J*) Negative control. A CID assay performed as in panels *A* and *B* but in the absence of a secondary antibody. Only a single particle is detected on the surface.

those of the positive controls (Fig. 5.6*E*), showing that only a single ssRNA2 is being extruded from a DLP at a given time. The average and the distributions of the the AexAem fluorescence intensity in captured DLPs do not suggest the presence of a DLP population with a stoichiometry significantly different from the two positive controls evaluated. The shape of the distribution observed in the captured DLPs followed the same shape as the positive controls only with a lower number of particles. The slight heterogeneity observed can be attributed to the photophysics in the hybridised probes and the the low number of captured particles compared with the relatively high contribution of non specific attachment (as shown in the negative control Fig. 5.6*I*).

Two negative controls were also performed. First, the VP6-specific antibody was substituted with rabbit gamma globulins and found only 12 colocalisations (some with the right code, possibly due to non-specific interactions between DLPs and gamma globulins). Second, immune-capture was performed in the absence of the biotinylated secondary antibody; this pulled down only one particle, without the right E*/S code. These results confirmed that the vast majority of probed ssRNA molecules are attached to transcriptionally active surface-captured DLPs.

5.2.4. Actively transcribing DLPs carry single extruded transcripts of segments 6 and 11.

To confirm that specific transcripts are extruded through single channels, the CID assay was modified to directly observe colocalisation of ssRNA6 and ssRNA11 transcripts at the single DLP level (Fig. 5.7). This was achieved by using a ssRNA6-specific probe pair (capture probe C_6 and reporter probe G_6) and a ssRNA11-

specific probe pair (capture probe C_{11} and reporter probe R_{11}). If the prediction was correct, the probing of actively transcribing DLPs using these 4 probes should capture particles with an E*/S value of 0.1/0.5, since independent transcripts extruded through independent channels should be beyond FRET range (>10 nm; the E* of ~0.1 is due to cross-talk between emission channels; see ALEX method in chapter 3 materials and methods). Indeed, the DLP sample (357 particles) showed E*/S values of ~0.1/0.5 (Fig. 5.7*A*). Due to the nature of the sample (which contains both DLPs with extruded RNAs, as well as released RNAs), the colocalisation is in part due to captured DLPs and in part due to random coincidence of released transcripts. The degree of ssRNA6-ssRNA11 colocalisation for the DLP sample was higher (p-value = 0.0004; Fig. 5.7*B*) than what was expected on the basis of random coincidence (see materials and methods, chapter 3), supporting the presence of significant specific transcript colocalisation on the captured DLPs.

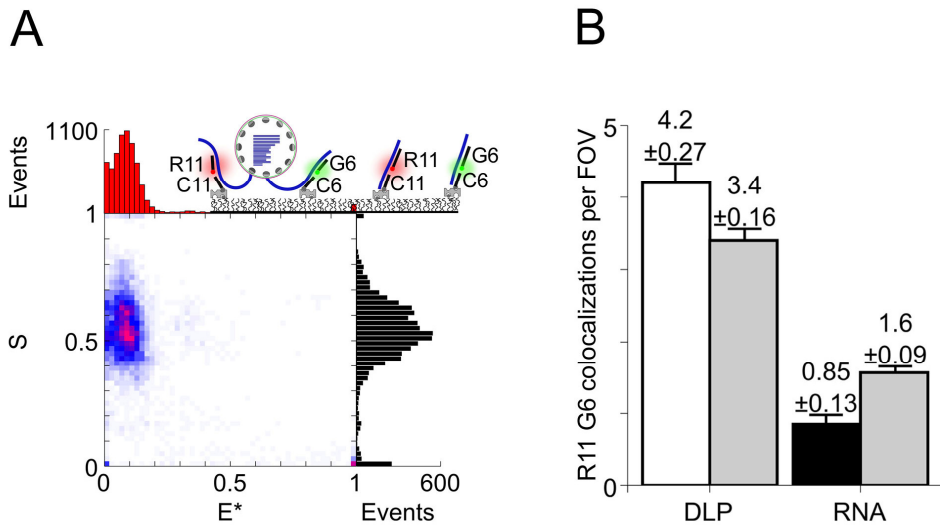


Fig. 5.7. Colocalisation of (+)ssRNA6 and (+)ssRNA11 on DLPs. *A*) CID of actively transcribing DLPs (allowed to transcribe for 2 min) interrogated using probes C_6 and G_6 (which target extruded (+)ssRNA6) and probes C_{11} and R_{11} (which target extruded (+)ssRNA11). The E*/S histogram (n=357, 80 movies) showed the expected E*/S value of ~0.1/0.5, matching that of the positive

control (panel *B*). *B*) Comparison of the mean number of observed colocalisations per 25 μm x 25 μm field of view with the mean estimated number of random colocalisations (gray columns; all error bars: SEM). For each field of view, the expected number of random colocalisations was calculated using the number of R_{11} and G_6 particles in each field of view. For the DLP experiment in (*A*) (white column), use of the one-sided paired-sample Wilcoxon signed-rank test showed that significantly more colocalisations were observed than expected by random colocalisation ($p=0.0004$; on average, 90 R_{11} and 183 G_6 particles). For the RNA experiment, fewer colocalisations were observed (black column; 34 colocalisations in 40 movies) than expected due to random colocalisation, therefore $p=1$ (on average, 66 R_{11} and 98 G_6 particles per field of view). A control experiment using the same hybridisation conditions as for the DLP sample but without the addition of nucleotides led to the capture of only 2 colocalised particles (20 movies), and no significantly higher colocalisation than expected by chance ($p>0.05$; on average, 1.8 R_{11} and 5.6 G_6 signals per field of view).

These colocalisation experiments provided an opportunity to test the hypothesis that rotavirus transcripts can associate independently and guide sequence assortment during encapsidation (14). Therefore, the CID assay was used to test for such an association between ssRNA6 and ssRNA11. We synthesised all transcripts (preparing a "total ssRNA" sample) by DLP transcription, followed by DLP removal (Fig. 5.7*B*). The number of colocalisations for total ssRNA was then compared to the colocalisations expected due to random coincidence; the absence of any statistically significant difference further supported that the DLP-dependent colocalisation arises due to specific colocalisation of transcripts on DLPs, as opposed to colocalisation due to capture of higher-order complexes of transcripts; it is also possible to conclude that, under the experimental conditions, there is no detectable direct or indirect (within a higher-order self-assembled structure) association of ssRNA6 and ssRNA11.

5.3. Discussion

The CID assay was applied to the study of three rotavirus transcripts that differ in sequence, type of proteins encoded, and length (S2, 2691 nt; S6: 1356 nt; S11: 667 nt, representing long, intermediate-length, and short segments, respectively). The results provide evidence for only a single (+)ssRNA copy from a specific segment being released at a given time; the data from these CID experiments do not show any evidence that support that multiple (+)ssRNAs copies from a specific segment are released at the same time from the DLP. The results from the CID assays are compatible with a channel-specialisation model in which rotavirus channels specialise in extruding specific transcripts from DLPs. Although there is no available super-resolution microscopy data of a specific channel being associated to one specific transcript in real time; my interpretation however, is supported by previous functional and structural data summarised in the following observations: a) transcription is very fast and occurs in multiple rounds (chapter 2) arguing against a model where segments diffuse and engage to different TCs and are released from different channels, b) structural data also showed that each ssRNA-TC is bound by the 5'end of the (+)ssRNA, a mechanism that facilitates multiple rounds of transcription by a single TC, making difficult the interaction of the segment with other TCs (13),(15), and c) the additional observation of tightly packed dsRNA segments in a semi-crystalline state (166) argues against the possibility that extrusion of a specific transcript occurs through different channels each time.

These RNA-TC interactions, mediated through the VP1 component of TC, may guide packaging of RNA segments in the capsid; the importance of these interactions suggests that our observations on ssRNA2, ssRNA6 and ssRNA11 are

likely to reflect the behaviour of all rotavirus transcripts. The results from the CID assays shown in this chapter, along with structural studies, also support the suggestion that one of the 12 TCs is not engaged with an RNA segment or alternatively, one of the channels is empty and lacks completely TC. The EM experiments do not allow one to resolve the exact stoichiometry of the TC within the DLP. However, a recently published model of RNA packaging in the capsid suggests that complexes of specific RNA segments and proteins are needed for the formation of a fully assembled and infectious capsid (98). Therefore, this assembly model would agree with a RV virion with 11 segments associated with exactly 11 TCs and an empty channel.

The results from CID assays also agree with the capsid organisation suggested in studies of BTV (167), another member of the *Reoviridae* family, which showed that the virus recruits a complete set of 10 RNA segments in the infectious capsid; notably, the absence of even a single ssRNA abrogated recruitment of the segments in the inner layered particle. The same report also suggested a hierarchy in the uncapped ssRNAs that leads to the formation of an RNA scaffold with a conserved set of RNA molecules per capsid. Although the mechanism by which rotavirus packages exactly 11 segments is not understood, it is thought that *cis*-acting sequences in the transcripts drive this process (48). This is the case for influenza virus, which has 8 negative-strand ssRNA segments and where it was shown (using FISH) that there is selective packaging driven by *cis*-acting regions located at the segment termini (148). In rotavirus, *in silico* modelling revealed the presence of regions with codon conservation that may act as structural motifs of *cis*-acting sequences (39). RNA sequences within viral segments are predicted to form conserved long-range interactions, stem loops, codon conservation, but also

dynamic ssRNA structures which may form intermolecular interactions or bind proteins to accommodate the segments in the confined space of the capsid. Under the experimental conditions used in the CID assays, no evidence of such scaffold was observed for total ssRNA, (although we cannot discard the possibility of smaller oligomeric subunits). One possibility is that the assembly requires different reaction conditions. For example, our experiments were performed in the absence of ancillary proteins, NSP2, NSP5 or VP1 that may stabilise or unveil cryptic binding sites in the ssRNAs. Furthermore, the low ssRNA concentration used might have been unable to drive the formation of RNA-RNA interactions that can take place in the viroplasm (where the local RNA concentration may be very high). Further insight on the assortment process should arise from future studies on the RNA sequence, structures or interactions that drive oligomerisation; such studies will no doubt benefit from the CID assay.

Chapter 6

Detection of RNA Transcripts in RNA Polymerase Elongation Complexes Using Single-molecule Hybridisation

6.1. Introduction

In this chapter, I use an in vitro transcription system from *E. coli* to test whether the CID assay can be applied to detect and investigate the RNA synthesis in a prokaryote. The RNA polymerase (RNAP) transcription model was chosen because it is a well established model to study transcription and provides several advantages to optimise the CID assay. For instance, by comparison with the DLP transcription system, the in vitro transcription model in *E. coli* is a simpler model to study; the proteins and DNA sequences are produced easily with standard molecular biology techniques, the interactions of the protein and DNA promoter are well characterised; there are already available data from studies pursued in the field with single molecule fluorescence techniques, the validity of these studies has also proven to be instrumental to unravel new interactions between DNA template and holoenzyme (168), sensing the phenomenon of transcription initiation (169, 170), promoter escape (171) and early elongation (139). The results of these studies have generated a wide range of reagents available for single molecule studies.

In this chapter, I formed RNA polymerase open complexes (RPo) which are the catalytically active transcription units. The RPo consisted of $\sigma 70_{596C}$, RNAP core, and promoter construct derivative of *pLacUV50* which was biotinylated and labelled with Cy5 fluorophores. *pLacUV50* is also modified to have the first cytosine in the transcribed region at position 50. Transcription was started by the addition of nucleotides ATP, CTP, UTP but in the absence of GTP that led to the formation of an elongation complex (RDe) with an extruded but not released transcript in the stalled RNAP.

RDe complexes were captured on a slide treated with neutravidin and identified with an ALEX–TIRF microscopy set up by colocalisation of the fluorophore

labelled probe and promoter which results in a E*\S code hybridisation of ~0/0.5 (no FRET and stoichiometry ratio of labelled molecules 1:1).

The findings of this chapter extend the use of the single-molecule hybridisation CID assay to capture, detect and quantify RNA transcripts from a biologically relevant prokaryote model. The time detection level is limited to an apparent hybridisation rate of 4.5 min. This research shows how this approach can be applied to gain an insight in the dissection of the components involved during transcription elongation in bacteria, and the detection and quantification of nucleic acids as a diagnostic tool.

6.1.1. Bacterial transcription

Transcription is the synthesis of a RNA molecule by a RNA polymerase (RNAP), from usually a DNA molecule so the genomic information is in a format that can be translated into a protein in the cell. However, in some virus families including the *Reoviridae* family, which is the central study of this thesis, RNA can be synthesised from a RNA template. Therefore, in this and the previous chapters I have used the definition of transcription in an inclusive and wide sense as a process responsible for the synthesis of (+)RNA regardless if the template is DNA or viral RNA.

In bacteria, RNAP is a multi domain protein structure that shares high homology with Archea and eukaryotic proteins which, in addition to its importance as phylum involved in a vast range of biological interactions, has made this system a standard model for studying transcription.

6.1.2. RNAP structure

The bacteria RNAP or "core" is 380 kDa and consists of five domains $\alpha^{I}\alpha^{II}\beta\beta'\omega\omega'$(Fig. 6.1). The RNAP is active only once it has formed a stable protein

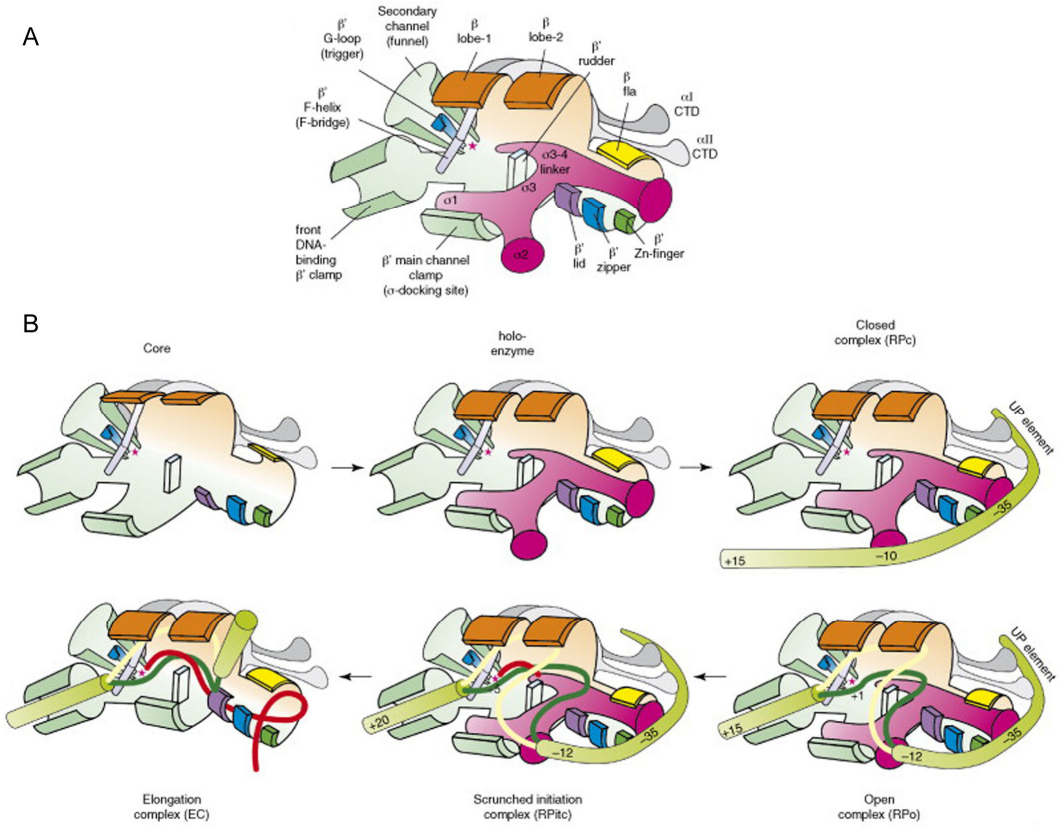


Fig. 6.1. Transcription cycle of bacteria RNAP. *A*) Schematic cartoon of the holoenzyme structure based on *Thermus aquaticus* RNAP core. *B*) Schematic cartoon of the main steps taking place during the transcription cycle. β'clamp recognises 15-20 bp downstream dsDNA during pausing and termination. The β lobes 1-2 undergoes conformational changes which regulate the opening of the main RNAP channel, and the binding and release of σ, binding the non-template strand in the transcription bubble and the RNA release in termination. The β fork loop-2 is responsible for directing the movement of downstream duplex DNA into the catalytic site and is also important for fixing the dsDNA in the transcription bubble. The β rudder binds the domains σ2, σ3, σ3-4 linker and provides stability to the DNA-RNA hybrid, β' lid domain is responsible for forming the end of the RNA exit channel. The β' lid domain controls the orientation of σ4 C-terminal portion, σ3-σ4 linker during the process of promoter recognition in the initiation of transcription. The β' lid directs the RNA into RNA exit channel. The β'zipper is also a component of the RNA exit channel and is responsible for regulating folding RNA hairpins which are determinants for termination and pausing of transcription. The β'N-terminal Zn-finger is located outside the exit channel; the functions attributed to this domain are pausing and regulating the termination. The β'N-terminal Zn-finger interacts with σ4 which suggests a function in promoter binding. β'flap is part of the RNA exit channel, It is part of the domains to be important for recognition of regions of -1/-35 promoter. The interaction of the β'flap with the RNA is important for controlling transcription, pausing and termination. It is also the domain that binds σ with the core and releases σ during elongation. The catalytic centre is formed by the β'F (bridge)-helix β'G (trigger)-loop which are integral elements of the active centre and ensure the fidelity in incorporation of nucleotides into the RNA chain, the translocation and the recognition of signal during termination. Taken from ref (172).

complex, called the holoenzyme (173), with one of the seven transcription factors of the σ family. The factor σ70 is one of the best characterised (174) and it is used in this study.

The RNAP core can be divided into 5 modules which form the three sides of the main channel of the polymerase with the fourth side formed by downstream dsDNA; the core module consists of the α N-terminal domain and regions of ω, β and β' that gives shape of the active site, the clamp module, β1 module, β2, and the β flap module.

The individual function of each module forming the holoenzyme, once bound to the promoter have been partially assigned to each stage of the transcription cycle (172). The structure of σ_{70} (174, 175) has five conserved regions σ_1, σ_2, σ_3, σ_4, the last two regions are linked with a conserved unstructured loop named $\sigma_{3/4}$. The functional role of each domain has been described (175). The domains σ_1, σ_3 and σ_4 form a globular structure that recognises a specific sequence promoter on the DNA template from which the transcription machinery starts. The modules σ_1 ,σ_3, σ_4 are responsible for binding specific DNA sequences and are involved in binding the domains -10, the extended -10 and the -35 element domain respectively (176). σ_2 is responsible for making contacts with the β' domain of the core. σ_4 establishes also an interaction with the β domain close to the exit of the RNA exit channel. Both domains σ_2 and σ_4 participate in binding conserved regions in the promoter (-35 and -10 elements) in the closed complex conformation of DNA unmelted (dsDNA) (176).

Upon interactions of aromatic amino acids of σ_2 the DNA melts, forming a transcription bubble (177), σ_1 and the $\sigma_{3/4}$ loop are flexible domains, unstructured and positively charged; $\sigma_{3/4}$ loop has the N-terminal region in the main RNAP

channel. It is thought to have an important regulatory function (175). The σ_3- σ_4 linker connects the σ_{70} domains that are separated by 45Å, the linker reaches on one side close to the active site which raises the possibility that it might affect binding of NTP and on the other side the linker interacts with the RNA channel exit where it is hypothesised to regulate abortive initiation and σ release (178, 179).

The structure of σ_1 is not fully resolved. However, based on the structure available from one of the sub-domains, it is shown that $\sigma_{1.2}$ has a role in coordinating three important functions a) abrogating the binding of the σ to the promoter in the absence of RNAP core, b) promoter recognition and c) opening the double-stranded DNA (DNA melting) (179).

6.1.3. The bacterial transcription cycle

Once the holoenzyme is formed, the transcription cycle is initiated with binding to a specific promoter, and it is followed by a transcription initiation, then it continues with a processive synthesis of RNA and finishes with a termination step (180).

In the initial step of the cycle, the σ70 recognises and binds the DNA promoter (P) forms a tight interaction called the RNAP-P closed complex RP(c), that progresses to an intermediate complex RNAP(i) which melts a DNA duplex sequence around the transcription initiation site (in the LacUV5 promoter it overlapped -11 to +4 nucleotide), forming the transcription bubble which leads to the formation of the RNAP-P promoter open complex RPo which is stable and catalytically active (180).

6.1.4. Transcription initiation

RNAP starts the synthesis of transcripts after the addition of the correct nucleotide triphosphate (NTPs) to the position (+)1; the incorporation of nucleotides takes

place in an active centre formed by the F-bridge and G-loop in the β domain that selects the correct NTP (181). The formation of a phosphodiester bond between substrates in the catalytic site results in the production of 7-15 nt transcript that move the RNA together with the DNA towards the RNA exit channel, while maintaining the core promoter fixed by the σ, while a similar sequence of downstream DNA is brought into the main channel. This process causes a "scrunching" (182) that results in the DNA-RNA hybrid displacing the template and non-template strand from the main channel.

The scrunching process may lead to two different outcomes: one is the appereance of several cyclic rounds of non-processive transcription which synthesises only short RNA transcripts (<9nt) which are released from the RPo, the phenomenon of cycles of truncated transcription is called abortive initiation (183).

The other is the rupture of the σ promoter interaction, while the transcription cycle continues into the elongation phase.

6.1.5. Early phase of transcription elongation

During the transcription elongation step, the RNA synthesis can oscillate between productive synthesis or blockage of transcription by backtracking (184) in which the space in the catalytic site is occupied by RNA. Our current understanding of the transcription elongation phase is based on a model that compiled biochemical and structural data from various organisms (185). In this model, the RDe does not show dramatic conformational differences with the RPo model (185).

After these initial rounds of short transcripts, there is a transition from initiation to elongation complexes resulting in transcripts longer than 9-11 nt. The elongation

complex is thought to keep a stable size of 12 nt in the transcription bubble, and about 8 nt of newly synthesised RNA-DNA hybrid (184, 172, 186).

6.1.6. Mature transcription elongation

Once the RNAP escapes from the promoter region, it transcribes processively at an approximately average rate of ~30–35 nt per second (187). The synthesis by the RNAP is regulated in the transcription cycle elongation phase by checkpoints such as pausing, backtracking (188), arrest, and editing. The multiple checkpoints ensure fidelity, and can synchronise RNA synthesis with translation and the cell cycle.

The precise description of the steps that lead to the formation of a mature elongation complex (up to 25 nt of newly synthesised RNA) and the kinetics of this process are currently being investigated. Initial observations have resulted in a model that postulates that σ_{70} was released after the synthesis of 9-11 nt (stable elongation complex or mature elongation complex (RDe)). These observations were based on standard biochemical techniques whose measurement of the release of σ_{70} relied on separation steps which causes an affinity decrease of σ_{70} and its loss from the RPo (189,190).

This view however, was challenged by experiments that showed that σ_{70} release is a non-obligatory step in the transition from early to mature elongation complex (191, 192). Three kinds of independent experiments support this view: a) experiments at the ensemble level that implement gentle separation steps of sigma transcription factor from phage promoters in the study of proximal pausing, b) identification of a fraction of σ_{70} in the elongation complexes using gentler purification steps, and c) identification of σ_{70} in a *pLac* promoter that bears a proximal pause in position -10 like sequence.

6.1.7. Single molecule ALEX to investigate transcription factor retention

Single molecule ALEX is a direct assay, it measures interprobe distance, stoichiometry of fluorophores that has been used to measure accurate quantification of σ_{70} retention during elongation (191).

The implementation of single molecule ALEX using *pLacUV5* DNA as a template to form RPo (191), showed that 80 % of the first stable elongation complexes (those containing 11nt of RNA) retain σ_{70} after synthesising 14 nt of RNA. Mature elongation complexes retain 25-55% of the σ_{70} factor. The effect of the template sequence in the σ_{70} retention has also been investigated (191).

Site directed mutagenesis of residues in σ_{70} factor and elimination of the pause inducing sequence reduces affinity for the remainder of the elongation complex but does not decrease the initial extent of σ_{70} retention. Removal of the pause-inducing sequence did not affect the properties of mature elongation complexes (191).

6.2. Methods

6.2.1. Expression and purification of $\sigma 70_{596}$

BL21(DE3) competent cells were transformed with PGEMD sigma plasmid that encodes $\sigma 70_{596}$. Expression and purification of $\sigma 70_{596}$ was prepared following a method described previously (193 , 191 , 194).

6.2.2. PCR and purification of LACUV5(-40Bio+64Cy5) and probe labelling

Construct LacUV5-50(-40+64Cy5) was generated by standard PCR using as a forward primer LacU5-50(-40/-20)TBio-40 and reverse primer LacUV5-50(+44/+64)BCy5+64 and using as a template the promoter of LacUV5 spanning

nucleotides (-40/+64)T (see materials and methods, chapter 3, table 3.2) . Samples were purified by PAGE with an estimated final concentration of ~500 nM as determined using single molecule confocal microscopy. Briefly, the construct concentration was calculated from a pM dilution visualised as an individual bursts in the single-molecule confocal spot. Labelling of LacUV5(+5/+28)B was performed as described for constructs in chapter 3.

6.2.3. Formation of RPo

The RPo open complex of *E.coli* was formed following published protocols (195 , 139,169) with several variations. Reconstituted holoenzyme was formed by mixing 100 nM RNAP core (Epicentre) with derivative $\sigma 70_{596C}$ (0.4µM) in T8 buffer (50mM Tris-HCl-HCl, pH8, 100 mM KCl, 10 mM MgCl2, 1 mM DTT, 5% glycerol) in a final volume of 30 µl and incubated for 33ºC for 30 min; RPo were formed by mixing the holenzyme complex with dsDNA template for 15 min at 37ºC. Next, the RPo were mixed with 2µl of resuspended heparin beads (1mg/ml GE Healthcare) and incubated for 30 seconds at 37ºC to disrupt non-specific interactions and centrifuged for 15 seconds after which 12 µl of the supernatant was transferred to a pre-warmed aliquot for its use in transcription experiments. RPo formation was evaluated by loading a mixture (5µl of sample and 1µl of 50% glycerol) in a 5% native polyacrylamide gel.

6.2.4. Standard transcription single molecule experiment

The standard transcription reaction consisted of 12 µl of RPo solution (containing 12 nM DNA template), mixed with a transcription buffer (nucleotides 2mM of ATP, CTP, UTP but lacking of GTP, 30 nM of G(+5/+28) probes which are Cy3B

labelled ssDNA probes in T8 buffer) in a final volume of 30 µl and incubated for different intervals of time at 37ºC. For single molecule transcription experiments, the sample was diluted to pM concentrations and added to the slide. For real time experiments, RPo complexes were preformed at 37ºC in solution, after which they were diluted (~50 pM) and transferred to a TIRF slide. Then the RPo complexes were mixed in the slide with a transcription buffer at 37ºC containing nucleotides, oxygen scavengers and Trolox (2mM). The experiment was recorded at room temperature.

6.3. Results

6.3.1. Expression and purification of $\sigma 70_{596}$ and production of LACUV5(-40Bio+64Cy5)

BL21(DE3) competent cells were transformed with PGEMD sigma. Expression and purification of $\sigma 70_{596}$ was prepared following a method described previously (193, 191,194).

$\sigma 70_{596C}$ and the promoter template LacUV5-50Cy5 have been described previously (139). $\sigma 70_{596C}$ is used for these experiments unlabelled. Briefly, sequence encoding σ70 was originally cloned in a PGEM-sigma plasmid and mutated with a site directed mutagenesis kit to introduce a change of amino acid Cys in position 596, $\sigma 70_{596C}$ (139). The plasmid was then used for transformation in BL21 cells. Once the protein was expressed, the recombinant was extracted as insoluble inclusion bodies. The protein was subsequently dissolved, renatured and refolded in a stepwise process. Refolded protein was purified using anion exchange chromatography (Fig. 6.2*A* and *B)* and buffer exchanged in a storage buffer for

further use. The concentration for this preparation was 7.1 µM (6 mg of total protein).

The promoter template LacUV5-50Cy5 used in this study is based on a construct described previously (196), (191) with several modifications.

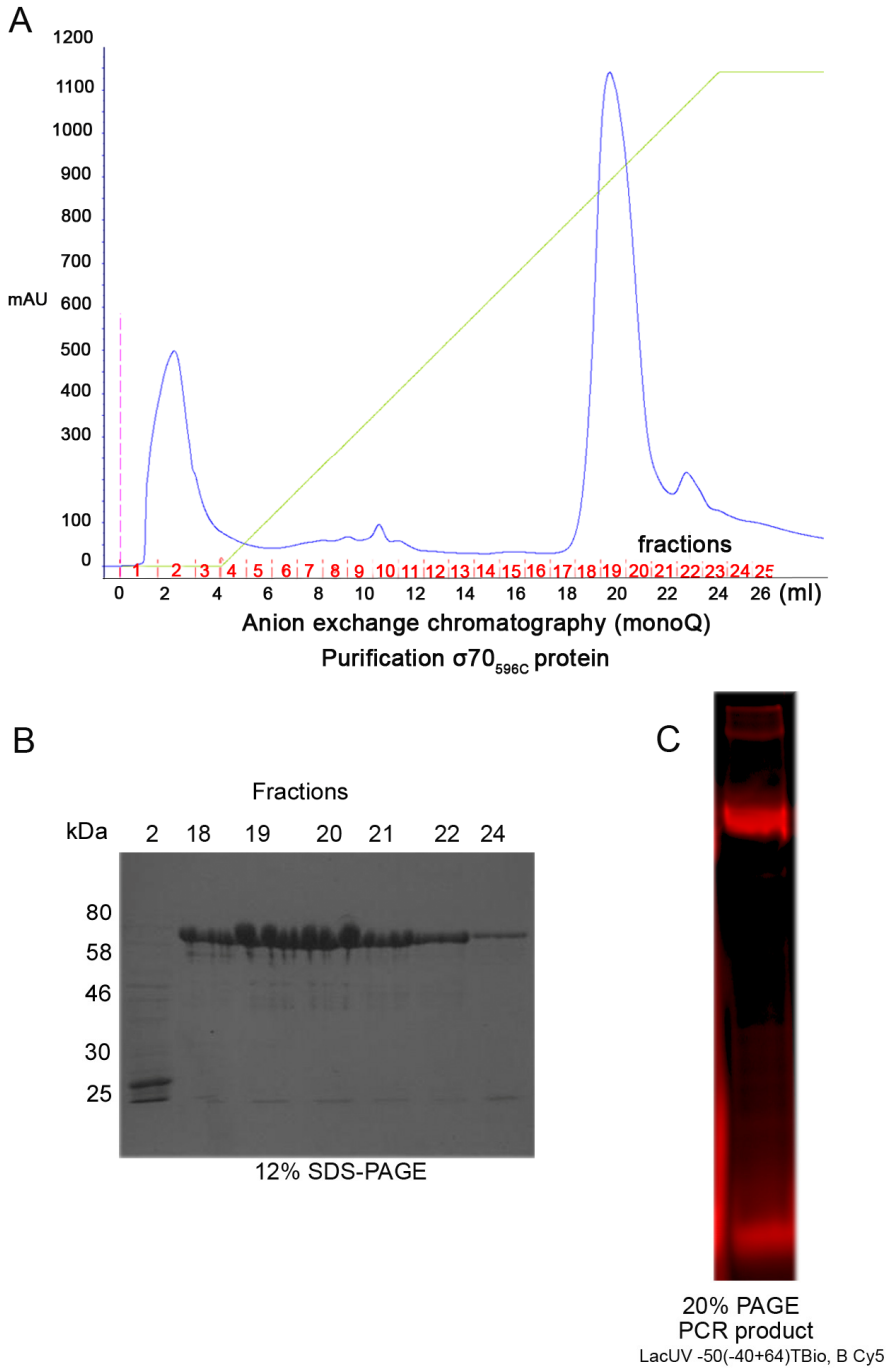


Fig. 6.2. Anion exchange purification and labelling of $\sigma70_{596C}$ and PCR of LacUV5-50Cy5 template. *A*) Anion exchange purification of refolded $\sigma70_{596C}$. Refolded $\sigma70_{596C}$ was separated and concentrated using anion exchange chromatography with a NaCl gradient of 500 mM. *B*) Fractions

from eluted peaks are shown in a 12% SDS-PAGE gel. *C*) LacUV5-50Cy5 was generated by PCR to introduce a biotin and a Cy5 label at the 5' end of the top strand (sense strand) and bottom strand (anti-sense strand) respectively.

The construct was generated by PCR (Fig. 6.2*C*) using a forward primer that introduced a biotin at the 5'end of the top strand (sense strand) and reverse primer that introduced a Cy5 at the 5' of the bottom (anti-sense) strand. PCR products were purified by 20% PAGE gel. The concentration after purification was approximately 500 nM.

6.3.2. RPo formation optimisation

Initial experiments were carried out to evaluate the formation of RPo and optimise the amount of RNAPo versus the amount of uncoupled LacUV5-50Cy5 template free in the solution. Before performing single-molecule hybridisation in halted RDe, preliminary experiments were done to optimise the formation of RPo with this batch of reagents; the optimisation step was important because the RPo formation is an inefficient process and an excess of free dsDNA template can potentially saturate the slide minimising the observations of hybridisation events.

σ70 (400 nM) and RNAP core (100 nM) were added in excess and mixed to form holoenzyme. The holoenzyme complexes were then incubated with serial dilutions of lacUV5-50Cy5 promoter Fig. 6.3*A* (left panel). The efficiency of RPo formation is evaluated in 5% PAGE gels and calculated as the percentage of fluorescence intensity of the shifted band corresponding to the formation of RPo complexes compared to the total intensity of the LacUV5-50Cy5. (Fig. 6.3*A*, right panel).

The effect of concentration of $\sigma 70_{596C}$ in the formation of RPo was also tested (Fig. 6.3*A* - left panel).

No dramatic changes in shifted bands are observed with concentrations of $\sigma 70_{596C}$ higher than 400 nM, suggesting that the initial amount is at saturating level. To evaluate the stability of the complexes, the amount of RPo was measured after 3 hours. In these experiments the 26, 44, and 22 % of the initial amount of RNAPo were still stable in three of the dilutions tested (Fig. 6.3*B*, right panel).

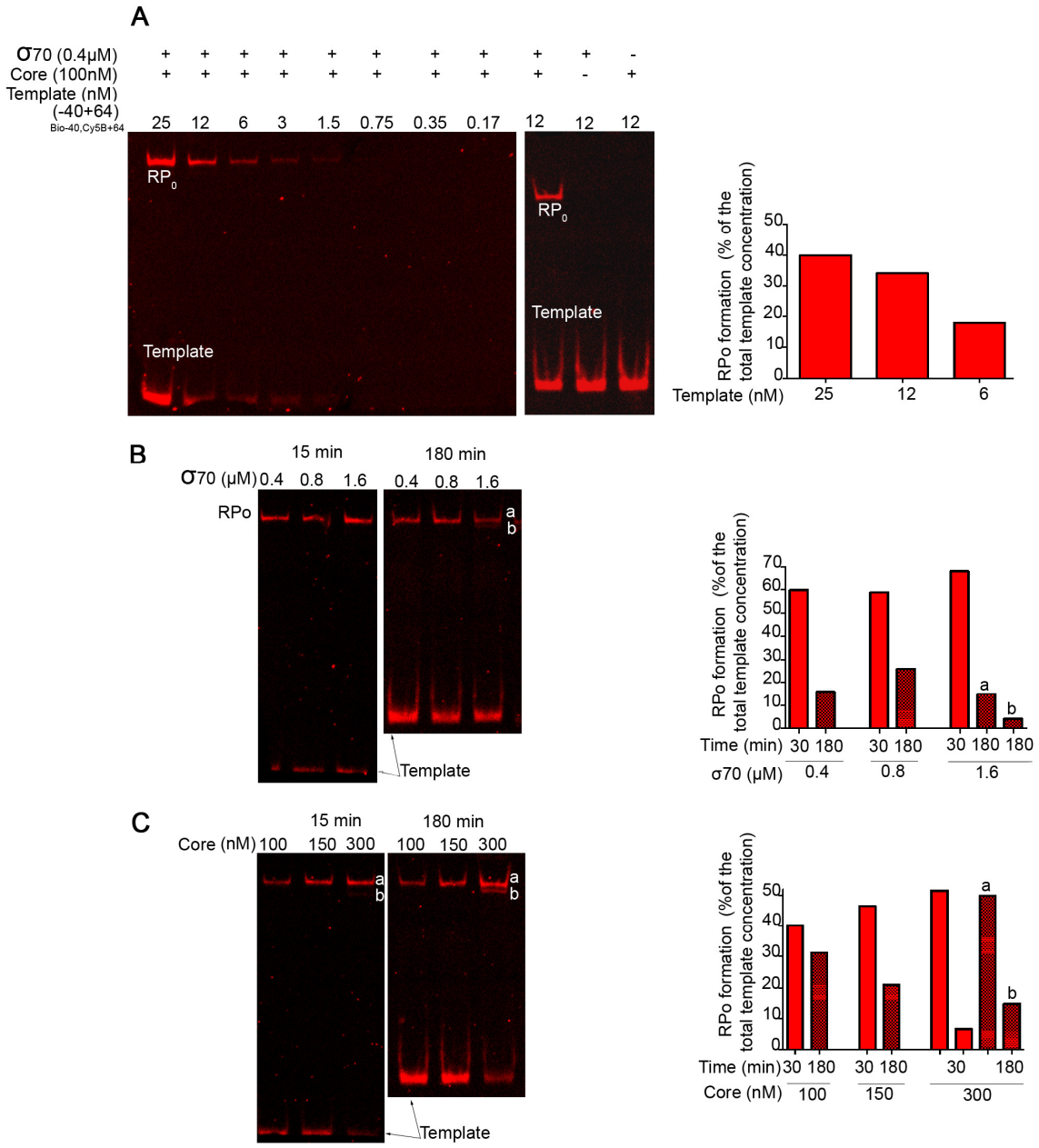


Fig. 6.3. Formation and optimisation of RPo open complex. *A*) Formation of RPo in variable concentration of LacUV5-50 promoter. Holoenzyme complexes were preformed by mixing 100 nM

RNAP core (Epicentre) with derivative $\sigma 70_{596C}$ (0.4 µM) in T8 buffer. Holenzyme complexes were incubated with variable concentrations of promoter template, treated with heparin and the RPo complexes visualised in 5% PAGE gel. RPo formation was calculated as the amount of average fluorescence intensity of RPo bands from the total fluorescence corrected for the background intensity in each gel. Fluorescence was measured with a gel scanner. Three promoter template concentrations were plotted. *B*) Formation and stability of RPo in variable concentration of $\sigma 70_{596C}$. RPo formation was evaluated with 100 nM core, 12 nM template and variable concentrations of $\sigma 70_{596C}$ .The stability of RPo at 37ºC was also measured with time in PAGE and the efficiency of RPo formation in the gels plotted as percentage of total template promoter. Band a corresponds with RPo; band b possibly corresponds with non-specific core-DNA complex *C*) Formation and stability of RPo in variable concentration of RNAP core. RPo formation was evaluated with 400 nM $\sigma 70_{596C}$, 12 nM promoter template and variable concentrations of RNAP core. The stability of RPo complexes at 37ºC was also measured with time. Based on the results from these experiments, subsequent RPo complexes were formed using 400nM $\sigma 70_{596C}$, 100nM RNAP core and 12 nM promoter template.

In a further attempt to increase the formation of RPo, the effect of RNAP core concentration in the formation of RPo was evaluated (Fig. 6.3*C*, right panel) and quantified (Fig. 6.3*C*, left panel) in a PAGE gel. No dramatic increase of RNAPo formation is detected after these conditions. In the presence of 300 nM of template an additional small band was observed that could be the result of non-specific binding of high concentrations of core with the dsDNA template LacUV5-50Cy5.

The experiments also showed that under the RNAP core concentrations tested (100, 150, 300 nM) after three hours there is still 13%, 45% and 77% of RPo remaining. Taking together these experiments, I used these results as the basis to prepare RPo complexes with 400 nM σ70, 100 nM RNAP core and 12 nM LacUV5-50. These concentrations show a good compromise between unbound promoter and formation of RPo complexes.

6.3.3. Ensemble transcription experiments

Next, I evaluated the activity of RPo in an ensemble experiment by performing a chase experiment. RPos were formed using the standard protocol optimised above, mixed with the four set of nucleotides, G(+5/+28) probes, and incubated for 1 hour at 37ºC. The addition of nucleotides should result in the formation of RDe where the RNAP translocated through the LacUV5-50Cy5 promoter, synthesised a full transcript of 64 nt and dissociates from the LacUV5-50Cy5.

The addition of nucleotides to RPo (in the absence of probes) resulted in the formation of a transient RDe, and its subsequent dissociation, which indicates that the RPos were active (Fig. 6.4.*A* - experiments without probes).

After the addition of nucleotides, the gels showed only a small amount of the RPo, approximately 3.75% (band a) and 12% (band b) of the shifted fluorescence band (Fig. 6.4.*B*.)

These bands corresponded with re-association of RPo and non-specific binding of $\sigma 70_{596C}$-LacUV50 respectively. When the experiment was repeated in the presence of G(+5/+28) probes (which are complementary to synthesised transcripts), the RPo was still dissociated after the addition of nucleotides, showing that the probes did not block the synthesis of transcripts (remaining shifted bands showing re-association of RPo complex were too faint to be measured accurately). Surprisingly, the experiment did not show any band that would be expected from the hybridisation between 24nt G(+5/+28) ssDNA probe and the released 64 nt ssRNA transcript.

The reasons to explain the absence of hybridisation to the transcripts can be attributed to the abundance of short transcripts produced during abortive initiation but also due to the sensitivity of the gel scanner, which could not detect the fluorescence intensity of a low concentration of hybridised constructs.

6.3.4. Single-molecule transcription experiments

Next, I use a single-molecule hybridisation experiment to investigate whether G(+5/+28) probes hybridised RNA transcripts stalled in RDe complexes under ionic strength conditions of the transcription buffer.

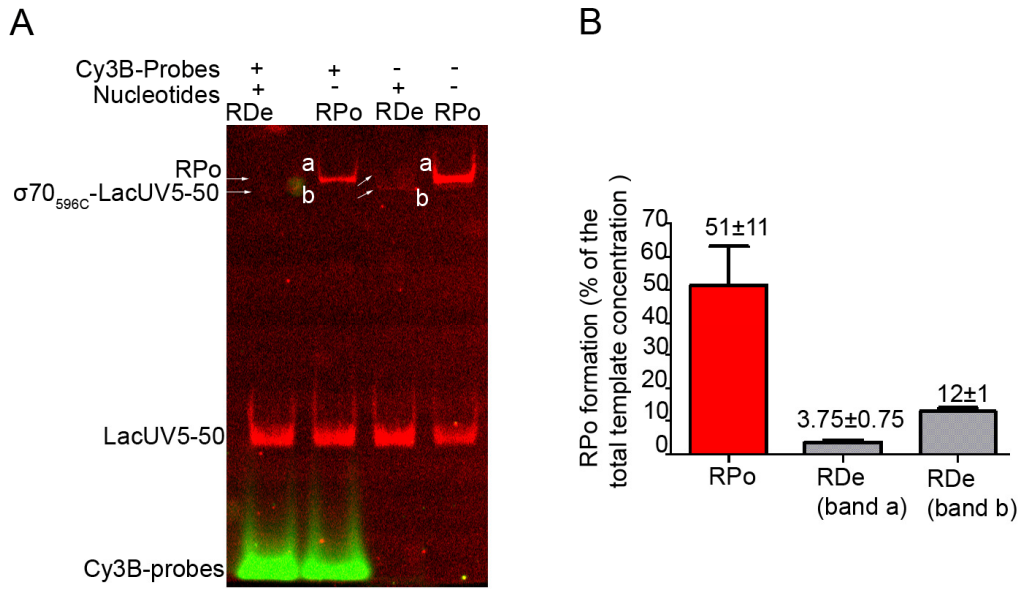


Fig. 6.4. Formation of RDe. *A*) RPo were mixed with ATP,CTP,UTP,GTP in the presence or absence of G(+5/+28) probes complementary to the sequence (+5+28) of the LacUV5-50 promoter. Incubation was performed for 1 hour at 37ºC. Negative controls included RPo in the presence and absence of probes. RPo and RDe experiments were resolved in a 5% PAGE gel.
B) RPo formation in the absence of nucleotides (red bar). After the addition of nucleotides RDe are formed resulting in the synthesis of 64 nt RNA transcript and further dissociation of RPo. Slow percentage of shifted bands are found after transcription due to RPo re-association (band a) and $\sigma 70_{596C}$ association to the LacUV5-50 promoter (band b). RPo and shifted band formation were calculated as the percentage of fluorescence intensity of shifted bands normalised to the total fluorescence intensity corrected for the background intensity in each gel.

To form stalled RDe complexes with defined transcripts length, I used RPo formed with the LacUV5-50Cy5 derivative promoter which has the following features. The construct is biotinylated at nucleotide position -40 and labelled with Cy5

fluorophore at nucleotide position +64 (see Table 5.1 materials and methods for details), allowing the capture and identification of the RDe at neutravidin-treated slides; the LacUV5-50 sequence is also modified so the first guanidine in the sequence susceptible of being transcribed is at position +50.

RPo were mixed with a 2 mM concentration of the nucleotides (ATP, CTP, UTP) with G(+5/+28) probes and incubated for 1 hour at 37ºC. Upon the addition of nucleotides, the transcription resulted in the formation of RDe with a transcript of 28 nt in length. The constructs were diluted to pM concentrations and immobilised onto a TIRF slide.

The E*/S histogram shows the presence of positive hybridisations of G(+5/+28) (Fig. 6.5*A*.) The E*/S code of 0.15/0.60 agrees well with a hybridisation where there is no expected FRET and the ratio of probes per RDe corresponds with a 1:1 ratio of G(+5/+28) and LacUV5-50Cy5 labelled molecules.

The amount of hybridisation was estimated of approximately 8% and was calculated as a percentage of hybridised RDe per field of view (FOV) from the total number of LacUV5-50Cy5 promoter immobilised on the surface. A T-test comparison shows significant differences in hybridisation between RDe complexes in the presence (green bar) and absence of nucleotides (grey bar). The experiment was performed in triplicate (p value ≤ 0.001) (Fig. 6.5*B*.)

Next, I test whether hybridisation of probes to extruded RNA transcripts can be improved by decreasing secondary structure or tertiary structure to non-specific self-binding between probes. Probes were incubated in 50% DMSO and 2.5 mM EDTA solution. The probe was then added to the transcription buffer to a final concentration of 30 nM probes in 1.6% DMSO, 80 µM EDTA.

The E*/S histogram shows the presence of positive hybridisations of G(+5/+28) showing 523 hybridisations with a similar E*/S code of 0.15/0.65 (Fig. 6.5*C*.)

A slight shift of stoichiometry is observed, the cause is not clear but it could be due to variations in the microscope settings and alignment between experiments or the effect of DMSO solvent traces in the fluorescence state of the fluorophore which

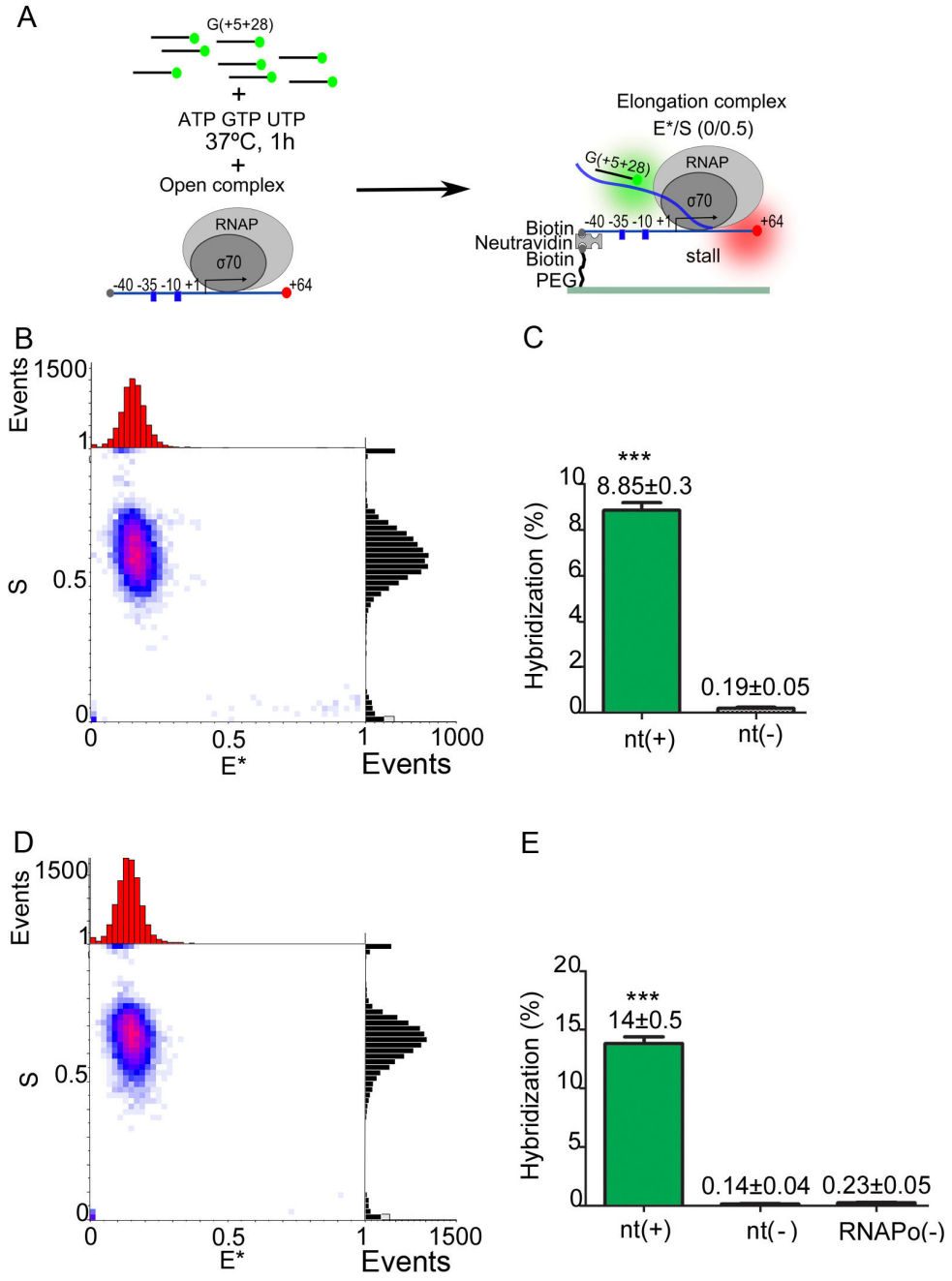


Fig. 6.5. Hybridisation of complementary probe G(+5/+28) to staled RDe

A) RPo were mixed with G(+5/+28) probes (30 nM), ATP,CTP, UTP (2mM) in the absence of GTP, so RNA synthesis was halted after the synthesis of 28 nt, and the transcript was not released due to

the absence of complementary guanine to continue the synthesis. The hybridisation was performed for 1 hour at 37ºC under standard transcription conditions. The E*/S histogram (n=431 hybridisations, one experiment, 40 movies recorded) shows an E*/S code of approximately 0.15/0.6 that agrees well with 0/0.5 code expected from fully hybridised RDe complex with 1:1 ratio of G(+5/+28) and (LacUV5-50) promoter. *B*) Comparison of hybridisation percentage (mean ± SEM) in the presence of nucleotides (9 %) and the absence of nucleotides (control). T-test comparison between treatments (≥150 movies) showed significant differences (***p-value <0.001). Hybridisation percentage of RDe per FOV was calculated out of the total of LacUV5-50 immobilised on the surface. *C*) The experiment was repeated in the presence of buffer containing DMSO (1.6%). The E*/S histogram shown (n=523 a E/S code of approximately 0.15/0.65 that mirror the experiment hybridisation performed in standard buffer conditions (A). *D*) Pairwise comparison of hybridisations (≥160 movies) percentages (mean ± SEM) in the presence of nucleotides (14 %), absence of nucleotides, or absence of RPo show significant differences (The symbol ***stands for a p-value <0.001). Hybridisation percentage of RDe per FOV was normalised to the total number of LacUV5-50 promoter immobilised on the surface. One-way ANOVA using a Tukey's test was used to compare different treatments. Experiments were performed in triplicate (≥160 movies in total).

may have changed the interaction of the fluorophore spacer or the interaction of the fluorophore itself with the protein and the DNA template.

The addition of probes treated with DMSO and EDTA increased the hybridisation per field of view to 15%. (Fig. 6.5*D* - green bar). The low concentration of DMSO did not appear to be detrimental for RNA synthesis or the complex stability since there is a 40% increase in hybridisation. Negative controls included an experiment performed in the same conditions but in the absence of nucleotides (Fig. 6.5*D* - grey bar nt-) which tests for non-specific binding to the template; a second negative control was performed in the absence of protein (Fig. 6.5*D* - grey bar RPo-) and tested for the effect of DMSO in unfolding the dsDNA template and making it accessible to hybridisation with the complementary probes. No binding was detected in any of the negative controls.

6.3.5. Single-molecule hybridisation in real time

Having shown that transcripts were detected in halted RDe complexes under standard transcription conditions, I conducted a single-molecule hybridisation experiment to investigate the time scale of hybridisation under transcription buffer conditions at room temperature.

RPo complexes were added to a TIRF slide, the excess of RPo complexes removed by washing, and the immobilised RPos mixed with transcription buffer at 37ºC containing G(+5/+28) probes. After which multiple field of views were recorded sequentially with long movies (100 ms, 2000 frames) consecutively (Fig. 6.6.) In this experiment a hybridisation is defined as the presence of time trace that included the appearance of any green probe G(+5/+28) with a defined intensity. I found differences in percentage of hybridisations per field of view 3.5% (RPo + nt) and 1 % (RPo – nt) but they were not statistically significant.

Possible explanations for the absence of significant hybridisation are: the way the particles are categorised in this experiment where we cannot discard non-specific interactions that will lead to the presence of false positives in the control experiment, the decrease in transcription activity of many RPos after being immobilised into the surface, the effect of the decrease in temperature that would slow down the interaction between molecules and can potentially disrupt the formation of RPo complexes, the low concentrations of labelled probes in real time experiments, the folding of RNA together with the association with the potential association with RdRP and the short period of time of the measurements.

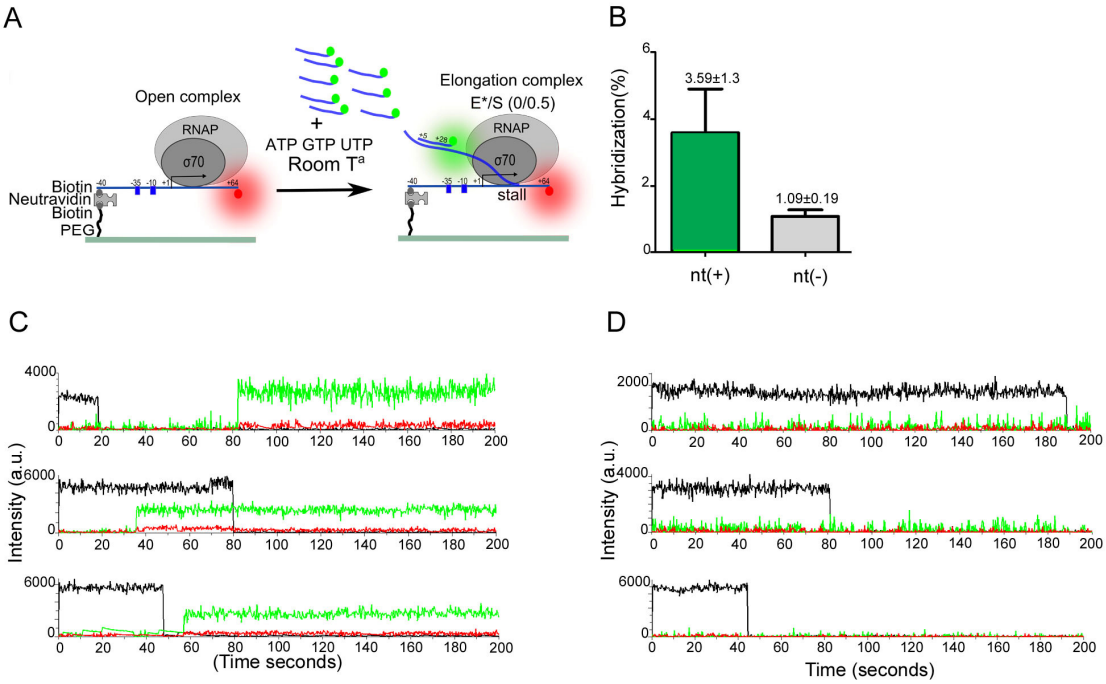


Fig. 6.6. Real time hybridisation of RDe complexes immobilised on TIRF slides. A schematic of the single-molecule transcription experiment. *A*) Schematic of real time single-molecule hybridisation on a TIRF slide. RPo complexes were added to a TIRF slide, the excess of RPo complexes removed by washing, and the immobilised RPos mixed with a transcription buffer containing G(+5/+28) probes (30 nM), ATP,CTP, UTP (2mM) at room temperature. Hybridisation was recorded in multiple field of views consecutively. *B*) Comparison of hybridisation percentage (mean ± SEM) in the presence of nucleotides (3.5 %) and the absence of nucleotides (1% control). A T-test comparison between hybridisation in the presence or absence of nucleotides does not show any significant differences (15 movies, 100 ms, 2000 frames; p-value >0.05). Hybridisation percentage of RDe per FOV was calculated as the percentage of LacUV5-50 immobilised on the slide which colocalised with G(+5/+28) probes. Hybridisation of probes to RDe per FOV was calculated out of the total number of LacUV5-50 promoter immobilised on the surface. *C*) Representative time traces from a single molecule, showing hybridisation to ssRNA in RDe bearing a Cy5 labelled LacUV50-5 (black) of a RDe with the G(+5/+28) probes (green) and no FRET (background red line). *D*) Representative time trace of negative controls performed in equal conditions as the transcription experiment but in the absence of nucleotides.

Due to the constraints and limitations of real time single-molecule experiments, another set of experiments were set up to obtain time scale information of the hybridisation of the G(+5/+28) probes to RDe. In this experiment RPo were mixed with nucleotides ATP, CTP, UTP and G(+5/+28) probes at 37ºC for 90 minutes. Aliquots were removed sequentially at 0, 5, 15, 30, 60, 90 min, added to TIRF

slides to measure the percentage of hybridisations of RDe per field of view. The results show hybridisation detection time of approximately 4 min (Fig. 6.7*A*.)

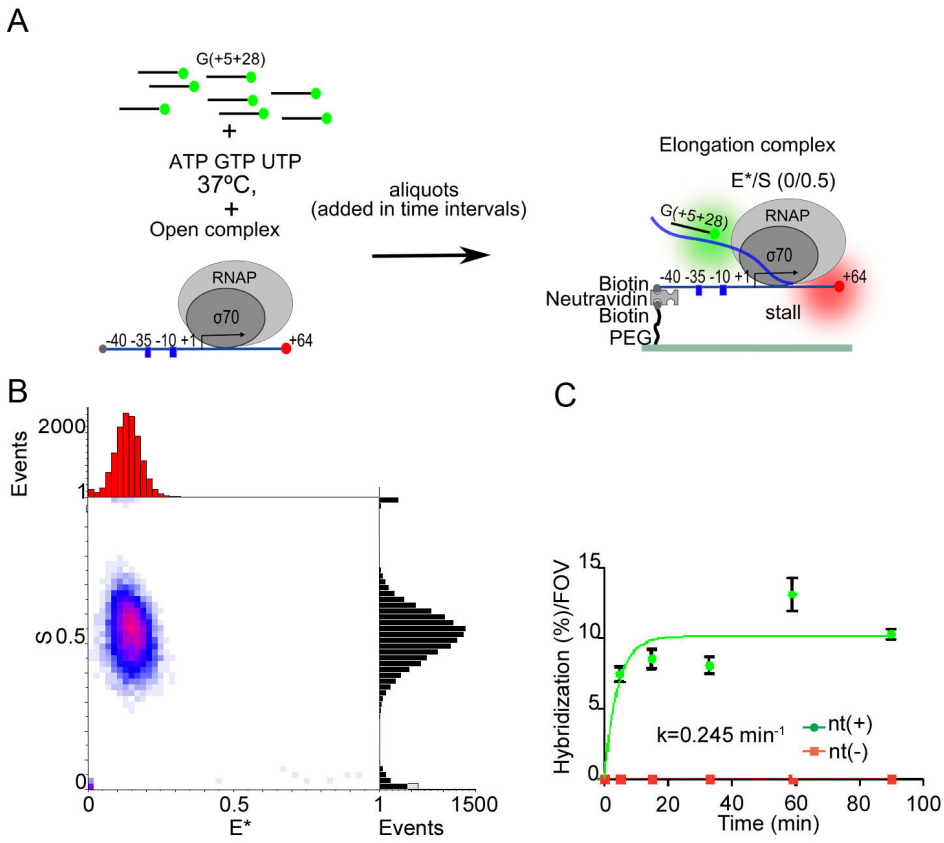


Fig. 6.7. Hybridisation kinetics of G(+5/+28) probes to stalled RDe complex. *A*) RPo were mixed with nucleotides ATP, CTP, UTP and 30 nM G(+5/+28) probes at 37ºC for 90 minutes. Aliquots were removed sequentially at 0, 5, 15, 30, 60, 90 min, added to a TIRF slides to measure the hybridisations per field of view of G(+5/+28) to RDe with LacUV5-50 promoter recorded. *B*) Example of stalled transcripts hybridised with G(+5/+28) probes to RDe complex with G(+5/+28) probes. The E*/S histogram (746 hybridisations) corresponds to the number of hybridisations detected after a 90 min incubation. The E*/S code of 0.15/0.5 matched with the presence of colocalisation of probes without any expected FRET (the value of 0.15 corresponds to leakage of green laser emission into the red channel and a S of 0.5 (a ratio of 1:1 of labelled probes).

C) Hybridisation of RDe to G(+5/+28) per FOV was calculated as the percentage of the total LacUV5-50 promoter particles immobilised on the surface. 26 movies were recorded per each point (100 ms, 40 frames). The data was fitted with a one phase exponential association function and show a hybridisation detection time of approximately 4.2 min.

After 15 min there was no further increase of hybridisation suggesting that the hybridisation of RDe in solution is saturated after 15 min. The E*/S histogram was also analysed per each data point incubation, an example of E*/S histogram corresponding with 746 particles observed after 90 min is shown (Fig. 6.7*B*.) The E*/S code of 0.15/0.5 agrees well with the expected code of one hybridisation of one probe per RDe and also with end point hybridisation assays shown before.

6.4. Discussion

In this chapter, I used initially an ensemble experiment to identify optimal concentrations of reactans to form RPo complexes; it was a necessary preliminary step before setting up a single molecule hybridisation experiment.

However, the formation of RPo complexes in high saturating concentrations of σ70 and core never yield to a complete shift of LacUV5-50. One possibility is the presence of unfolded protein that could underestimate the concentration of native active σ70.

However, because the experiments were performed in bulk, it is not possible to discard the presence of unstable intermediates that would be undetected in the gel shift assays. For example, a single-molecule study in a non-homologous promoter σ54 from *S. typhimurium* showed that the formation of RPc is a process kinetically unstable where interactions of RNAP and σ54 comprised a minimal sequential kinetic scheme of two intermediate closed complexes, one short-lived closed complex and a stable long-lived closed complex (197).

The bulk experiments investigating RPo formation also showed that once the RPo is formed, a large fraction remains stable after 3 hours, which argues for the presence of several subpopulations one of them formed by stable long-lived RPo complexes.

Overall, the results showed that the hybridisation rate is not dramatically affected by the stability of the complex in the experiment.

Next, I verified that RPo were transcriptionally active and RNA synthesis from a biologically relevant model such as the *E. coli* can be detected at the single-molecule level with an end point assay. It was shown by mixing RPo with a buffer containing ATP, CTP, UTP but not GTP and G(+5/+28) which hybridised to newly synthesised transcripts stalled in the RDes. Hybridisation was detected by colocalisation spots of the Cy3B labelled probes, G(+5/+28) with Cy5 labelled promoter able to detect approximately 18% of the RPo complexes formed in solution which were mixed with transcription buffer. When the probes were incubated with DMSO the hybridisation increases up to 30% of the total of RPo formed in solution.

I attempted to modify the hybridisation buffer to improve the probe hybridisation. In this experiment I used DMSO, which has been used as an additive in PCR reactions in the synthesis and during the assembly of the reaction (198-200). The DMSO disrupts secondary structure formation and minimises high T_m oligonucleotides by disrupting inter and intra strand re-annealing; the results showed that a percentage of ~1% DMSO improved the specificity and quantity of product when added to the PCR reaction.

The reasons why a large fraction of the RPo were not colocalised with transcripts could be the result of the hybridisation conditions used in these experiments which are performed under low concentration, mild ionic strength conditions and at room temperature, these experimental conditions may have not broken the secondary structure, self association of the transcripts and probes respectively.

In a further attempt to investigate real time hybridisation at the single molecule level, an experiment was performed. Under the experimental conditions used, the results showed a very low number of hybridisation events that can be attributed to a combination of constraints during these measurements, the use of lower concentrations of fluorescent probes used in real time experiments and the fluctuations in temperature that RPo complexes undergo from the incubation conditions (37ºC) to the TIRF slide where measurements are taken at room temperature.

As a result the hybridisation rate was roughly investigated with a time course experiment. This measurement showed that 4.5 minutes is the minimum amount of time where transcripts were detected using a probe concentration of 30 nM. These results are longer than it would be expected when compared with the hybridisation rate for σ54 which was 4.5 seconds using probe concentrations of 20 nM (197). Hybridisation time is not measured using a microfluidic system so we cannot establish a precise initial time, hybridisation measurement was overlapped with several processes such as the synthesis of short and mature transcripts.

The experiment is designed to obtain kinetic information under the concentration of hybridisation probes that can be used in a TIRF experiment.

As a result, the results are only an approximate indication of the saturation of the hybridisation reaction and it would be necessary to decrease the concentration of probes to obtain more data points that provide an accurate saturation time.

The hybridisation rate in this experiment is constrained by the transcription buffer and the temperature used, which must ensure the stability of the RDe.

However, it has already been shown that under different buffer conditions, it would be possible to enhance the rate of hybridisation by varying, NaCl concentration, use

of solvent, and temperature. The effect of NaCl concentration in aqueous buffers has been investigated. The concentration of NaCl increases the hybridisation rate of DNA. The hybridisation rate is maximised at a concentration of 200 mM (201). Other buffer additives have also been used to optimise nucleic acid hybridisation including ethanol, methanol, formamide and DMSO (201).

For instance, a concentration of 56% of ethanol in the buffer gave a 70 fold increase in hybridisation rate (201). The mechanism by which hybridisation is increased is partially due to the observation of a concomitant decrease in Tm associated with the effect of hybridisation (201).

In summary, here I have reported a set of experiments that prove the validity of single-molecule hybridisation to detect RNA transcripts in biologically relevant transcription systems. These results show that under these buffer concentrations, RNA is detected in 4.5 minutes which point out that experiments in rotavirus transcription can be limited to fixed events rather than real time observations. The foregoing observations were limited by the low percentage of RPo complexes immobilised in TIRF slides which can be available for investigation. Further improvements in experimental design can benefit from new techniques developed recently such as single molecule immunoprecipitation (202, 203) that can specifically pull down native protein complexes concentrating the amount of RPo and limiting the amount of free uncoupled LacUV5-50 promoter that constituted the majority of the immobilised constructs in our single-molecule studies.

Chapter 7

General Discussion

7.1. General discussion

In this thesis, I have presented studies of the *Reoviridae* family which comprises a group of viruses of relevance for human health, such as rotavirus, one of the major causes of mortality and morbidity in infants, as well as zoonotic diseases. The family also includes viruses responsible for losses in agriculture such as blue tongue virus, which is now an emergent disease in Europe as a result of the propagation of the vector transmitting the disease to northern latitudes, possibly due to global climate change.

Common to all members of this family is the complex icosahedral capsid architecture consisting of concentric protein layers, within which the virus contains a segmented dsRNA genome and the transcription complexes. The virus synthesises (+)ssRNA which are extruded through 12 channels that communicate the internal space with the cytoplasm of the infected cell where the RNA is translated or packaged in newly assembled capsids.

Although the export of transcripts is a crucial step for viral replication, a detailed functional study on the type of transcripts extruded was still lacking. In particular, it was not known whether the release of a transcript from a specific segment was always coupled with the same assigned channel (in this thesis designated the "channel-specialisation" model) or instead, whether the channels were not specialised, and can participate in the release of a transcript from many different segments ("channel-generality" model). This posed an important question, which once fully addressed will provide an insight in how the segments are finally packaged in the virus, and how the release of (+)ssRNA transcripts is regulated, so intervention strategies can be developed to block either the packaging of the segments or the gates releasing newly synthesised transcripts.

To test the different models of transcript export, a novel capture and identification assay (CID) was developed in this thesis. This assay determined the copy number of specific transcripts during their extrusion from transcriptionally active DLPs. In this thesis, I showed that the CID assay using a E*/S code detects transcripts in nM concentrations at the single molecule level after ~15 minute incubation.

I used DLPs from RRV to study the mechanism of transcript extrusion for the *Reoviridae*. Unlike other members of the *Reoviridae* such as ORV, BTV and african horse sickness, the RV genome is not easily to be modified since this research model lacks reverse genetics, a type of technology that produces a defined virus progeny by using reverse transcription of ssRNAs into cDNA and subsequent re-transcription of engineered cDNA into ssRNA (204). RRV was nevertheless chosen because it is relevant to the study of the human RV, and DLPs from RV do not require restricted containment facilities in the lab (since DLPs are non infectious for humans or animals).

The CID assay detected unambiguously RNA molecules, because it used two observables E* and S and a transcript capture strategy; and only once the correct number and type of probes were hybridised to the right sequence, it produced the expected code. Therefore, the implementation of a E*/S code made the CID assay more specific than any other single molecule FISH assays which do not have a reliable way to test whether hybridisation is specific. The introduction of a capture probe increased the sensitivity because the capture probes pulled down on the slide all the hybridised probes, a development which was also advantageous with respect to previous E*/S approaches that used hybridised probes in solution using confocal microscopy (205).

7.1.2. RNA packaging and transcription in RV

The results from this thesis using the CID assay prove that the "channel-specialisation" model is correct in RV and excludes the alternative model in which an individual segment is transcribed by several complexes and extruded through multiple channels. The results from this thesis also suggest that ssRNA selective packaging must require ancillary proteins, since oligomerisation of 11 ssRNAs was not observed under the concentration and buffer conditions investigated. The specialisation of channels, together with the selective packaging, and the observation that partial packaging (DLPs with variable number of segments) do not occur in the infected cell, support a model in which the 11 segments occupy a highly ordered localisation in the capsid, although whether or not there is spatial conservation of the segments within the capsid and if so, what the mechanism for packaging is, remains unknown.

An update for the concerted model, (shown in chapter 2) is shown (Fig. 7.1). In the updated model, selective packaging is linked to channel specialisation and spatial conservation of the segments within the capsid. In this stepwise, selective packaging model, the "conserved motifs" in the 3' and 5' regions which are common to each ssRNA recruit a TC; subsequently, the ssRNA-TC complex adopts a stable conformation that exposes "specific recognition motifs" unique to each ssRNA segment; the specific recognition motifs then recruit specific neighbouring ssRNA-TC complexes, resulting in a stepwise addition of ssRNA-TC complexes. The interactions ssRNA-ssRNA and ssRNA-TC maintain the complex inactive. The recruitment of ssRNAs based on specific recognition sequences allow a quality checkpoint in viral replication, because mutations either in the TC sequences or the ssRNAs will alter the complex conformation and will fail to recruit the subsequent

specific assembly complexes; this model accounts for the presence of replication intermediates and the absence of partially filled DLPs. A prediction from this model is that a single five fold TC channel is unoccupied. Specifically, VP2 capsid proteins are added by an interaction with the RNA charge and the TC. The complete and correct assembly of VP2 to the TC-ssRNAs causes a conformational change between the TC and the ssRNA and a rearrangement interaction of the ssRNA with the VP2 that leads to a loss of ssRNA-ssRNA interactions, both of these changes contributing to activate the TCs and make dsRNAs.

7.1.3. Prospects for the CID assay

The CID assay is a new development of the FISH assays, and due to its specificity and fast procedure can be used as general method to detect nucleic acids but also in cells. By using an E*/S code, and the capture effect of the biotinylated probe, the CID assay also has a potential use as a multiplexing diagnostic tool with high sensitivity. Several aspects of the transcription of the *Reoviridae* could be further investigated with this method or a slight variation of this method. This new research could investigate *in vitro* real-time studies of the polymerase; in the absence of capture probes in-cell experiments could be performed to monitor the localisation of transcripts extruded by the DLP and their fate during translation and packaging in the replication cycle. Although the CID assay has been implemented in the study of RV, the work presented here proves that it can also be used to investigate *E. coli* transcription and therefore can have broader applications. The CID assay can potentially circumvent some of the problems faced with standard FISH assays, widely utilised to investigate *in vitro* transcription models, detection of microbiological populations and pathogen identification in medical samples.

Specifically, FISH approaches are proven to be robust for identification of prokaryotic populations, but they lack single-molecule sensitivity and their use can be limited to the analysis of viral RNA in the cell. These assays include in situ PCR, RING-FISH or enzyme based amplification systems such as CARD-FISH, which uses probes coupled with Horseradish peroxidase (HRP) and tiramide signal amplification with a signal 3-4 fold higher than hybridisation on monolabelled multi probes, reviewed in ref. (206).

More specific single-molecule FISH (smFISH) approaches have also been used to investigate the cell localisation of ssRNA in viruses such as influenza (207) or in the study in eukaryotic cells of the transcriptional variability, localisation and splicing of mRNA in single eukaryotic cells (208).

However, these smFISH based approaches to detect a transcript target rely on the use of several probes which are labelled with multiple fluorophores (209) or alternatively, with just a single fluorophore moiety (210). Both approaches have several caveats; use of probes labelled with several fluorophores is more susceptible to self-quenching and present a heterogeneous signal due to the difficulty of fully labelling and purifying the DNA probes; the second approach (multiple probes singly labelled) depends on a large number of probes, and as a result, has a tendency to self- and non-specific hybridisation.

In summary, the CID assay will also benefit from the constant developments in FISH techniques (211). Refinement and optimisation in the current assay could include: careful choice in RNA target, improving RNA transcript stability, for example using RNA probes that render the hybridisations thermodynamically more stable than the ones using RNA-DNA, also improving the accessibility of the probes to the target, introducing adjacent helper oligos that unfold the secondary structure,

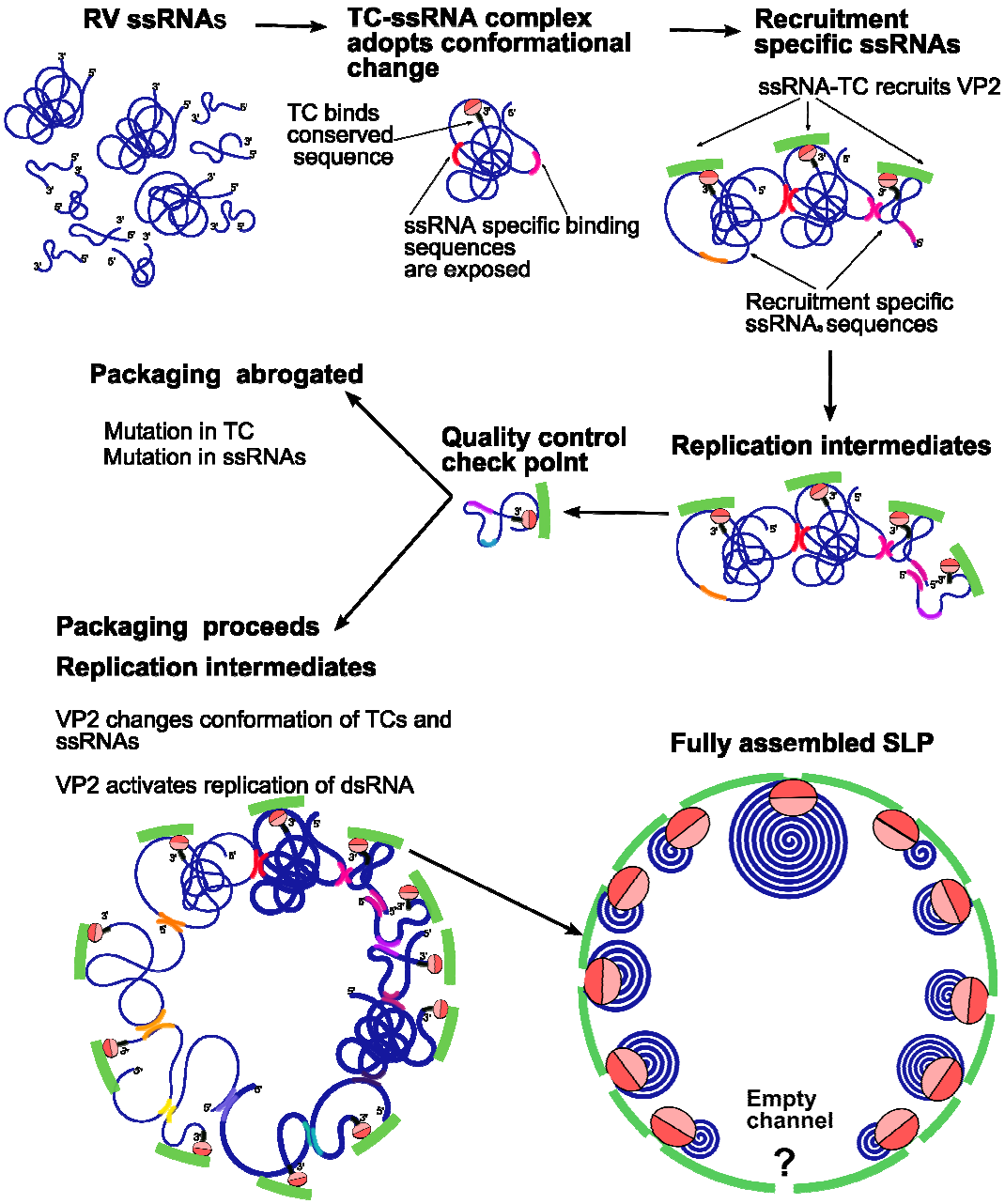


Fig. 7.1. Stepwise packaging model of ssRNAs in spatially conserved localisations within the RV capsid. In this model, free ssRNAs in solution adopt flexible and dynamic conformations. The ssRNAs have conserved motifs that bind the TC; the TC-ssRNA complex adopts a conformational change that exposes unique sequences that recruit specific ssRNAs, resulting in a conserved spatial localisation of the ssRNAs; the model predicts that only 11 VP1 molecules (in RV) are expected to be found within the DLP; segment recruitment follows a stepwise process, which results in the addition of new TC-ssRNA complexes resulting in the formation of replication intermediates; VP2 capsid protein are added by an interaction with the RNA charge and the TC. The interaction ssRNA-ssRNA and ssRNA-TC maintain the complex inactive. Mutations in the TC and or the ssRNAs can alter the complex conformation failing to recruit another TC-ssRNAs and blocking the assembly.

The correct assembly of VP2 monomers causes a conformational change between the VP1 and the ssRNA and a loss of ssRNA-ssRNA interactions that contributes to activate the TC.

looking at advances in cell permeability protocols (for in situ cell applications), designing labelled peptide nucleic acids (PNA) that hybridise with high affinity under low salt concentration and high temperature conditions or using methods for labelling probes such as click chemistry. The CID assay can be the basis for an identification test that targets important *Reoviridae* pathogens of veterinary or medical importance.

The virus of choice should be an exotic virus, an emergent pathogen or a novel virus recently discovered for which diagnostic tests are not available, or are not routinely performed, but nevertheless rapid tests are needed to inform the decision-making process leading to the establishment of rapid outbreak control measurements. The test should include multiplexing, so it is possible to establish a differential diagnostic, identify virulence and overlapping strains, and provide epidemiological information about the vector strains transmitting the disease. With this view in mind an interesting target in the *Reoviridae* family can be BTV, and the zoonotic African horse sickness.

7.2. Further research

The output of this thesis has left still many opened questions but has also pointed out new research avenues that now can be the focus of further investigation.

a) A direct question raised by the research in this thesis, is whether there is a spatial conservation in the position that the segments occupy within the capsid. In other words, is each segment, which is packaged following an ordered interaction keeping always the same spatial relationship with a set of specific transcripts? Selective packaging of exactly 11 segments and channel specialisation are indeed pointing to

a packaging model of spatial conservation of segments. To test a model of conserved spatial localisation of the segments within the *Reoviridae,* two complementary approaches are possible.

1) Super-resolution microscopy of a pair of labelled segments in a core of BTV reconstituted *in vitro* (Fig. 7.2). Formation of singly or multiply labelled constructs can be done using azido-derivatised dinucleotides (212) to be introduced at the 5'-end of synthetic ssRNAs; these modified dinucleotides can be later labelled with alkyne-fluorophores using click chemistry (212). Alternatively, other methods such as labelled nucleotides or short labelled probes can be used to decorate the transcripts. The transcripts can be then incorporated *in vitro* in the core following a recently published protocol (167). Next, the cores are attached to the TIRF slide and pairwise distances (as 2D projections) can be measured between labelled segments using a super-resolution microscopy based method such as direct stochastic optical reconstruction microscopy (dSTORM) (213) or two-colour co-localization of fluorophores using precise registration in imaging channels (214). With dSTORM, the positions of ssRNAs can be resolved with 20 nm accuracy. These measurements can provide information about the shape, the position and the distance between labelled transcripts. To further improve the localisation, we can also use three-dimensional positioning of these segments by implementing 3D-STORM (215), or it could be possible to improve the localisation accuracy below the 20 nm using interferometric photoactivated localisation microscopy (iPALM) (216).

Even if only the labelling of the core is achieved, it would be an invaluable reagent for the research in *Reoviridae*, as the core could be tracked in the whole process of virus replication from invasion, replication in the cell to the budding, and BTV release and cell lysis.

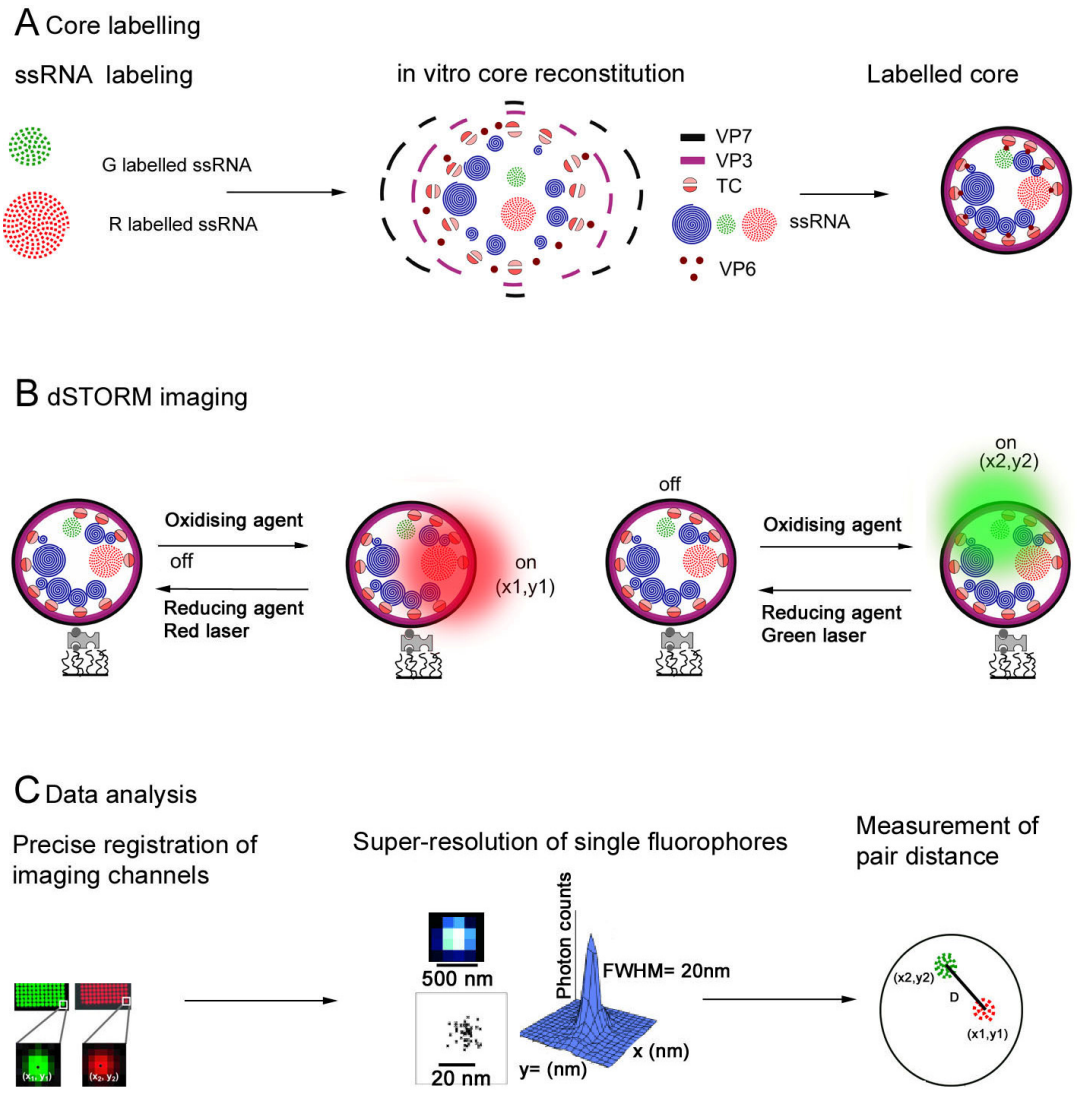


Fig. 7.2. Study of spatial conservation of segments within the BTV viral capsid. *A*) Synthetic ssRNAs from BTV are labelled with photoswitchable dyes, using an array of approaches including azido-derivatised dinucleotides, labelled nucleotides or short labelled complementary probes. BTV is assembled with a cell-free method. *B*) dSTORM is performed with buffers that photo-switch green and red fluorophores. The two-colour super-resolution of fluorophores is achieved by exploiting the differences in spectral emission, and stochastic fluorophore emission caused by introducing reducing agents in the buffer. *C*) Data analysis. Localisation of a fluorophore requires a precise registration with nm accuracy of imaging channels. Then, dSTORM imaging is performed; each single fluorophore emission is time-resolved and recorded multiple times generating point spread functions (PSF) of ~500 nm in lateral resolution. Then, the centroid coordinates of each PSF are calculated

with the FWHM (full width height maximum) by fitting a Gaussian function, which allows fluorophore localisation with ~20 nm accuracy (depending on the number of photons collected). Distance measurements between mapped positions are performed per each core generating a distance distribution per each pair. An investigation of pair distances will allow me to establish a map of the segments within the capsid.

2) A complementary approach can also be attempted using DLPs from RV. The implementation of an independent method in a different study model will add several advantages. It will provide robustness to our results from BTV; experiments with DLPs from RRV are easy to perform in a biological physics lab, since they do not have the biosecurity restrictions of BTV and it will allow me to test whether this is a general mechanism along the *Reoviridae*. In this approach, spatial localisation of extruded ssRNAs from RV DLPs is investigated. DLPs are captured on the TIRF slide using the CID assay developed in this thesis (Fig.7.3). Extruded transcripts will be hybridised using one or multiple singly labelled probes for each target. The implementation of multiply labelled probes can potentially improve the signal and the precision in the transcript localisation. The design of probes, could be optimised by using PNA probes that have better access to the target. An advantage of using DLPs, is the possibility of fixing the attachment position to the surface, limiting the number of possible projections to be found and simplifying the data analysis. Nevertheless, there are also major drawbacks in this approach: DLP capture with CID assay is an inefficient process, and the CID assay cannot control the length of the extruded transcript which will generate multiple distances; as a result, interpretation will be difficult, since there will be an added degree of variability between particles and experiments.

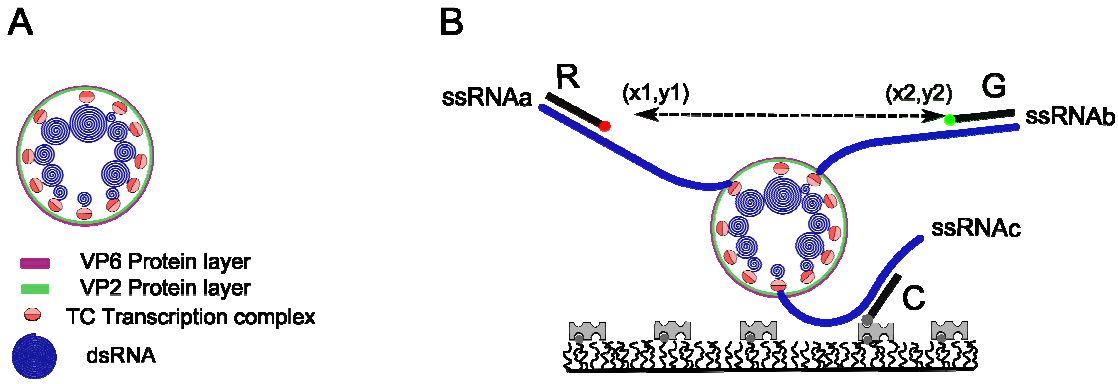


Fig. 7.3. Study of spatial conservation of segments within the viral capsid using a CID assay. *A*) Schematic of the RV DLP structure. *B*) Schematic of a modified CID assay to investigate spatial conservation of segments within the viral capsid. DLP transcription is blocked after 2 min. DLPs are then incubated with identification probes labelled with green (G) and red (R) fluorophores and a capture probe C) which are complementary to three different extruded ssRNAs (a, b and c). Orientation between extruded labelled ssRNAs is fixed with the capture probe. Distance between fluorophores is investigated using super-resolution microscopy that relies on precise registration of molecules in both channels, and the different spectral emission of the fluorophores.

b) How, where and when the virus decides whether the transcripts are translated or co-packaged in a new capsid. Single-molecule hybridisation combined with super-resolution switchable FRET (165) could be used to study spatiotemporal changes of specific ssRNA transcripts in the infected cell. This method will allow me to identify multiple ssRNAs using pairs with a defined E*, but by using switchable dyes it will also be possible to find the precise localisation in the cell and an approximate amount of transcripts in these positions.

References

1. Mertens, P. (2004) *Virus Res* **101,** 3-13.

2. Mertens, P., Attoui, H., Duncan, R. & Dermody, T. S. (2005) *In: C.M. Fauquet, M.A. Mayo, J.Maniloff, U. Desselberger, and L.A. Ball (eds.), Virus Taxonomy. Eighth Report of the International Commitee on Taxonomy of Viruses. London: Elsevier/Academic Press***,** 447-54.
3. Lawton, J. A., Estes, M. K. & Prasad, B. V. (1997) *Nat Struct Biol* **4,** 118-21.
4. Attoui, H., Fang, Q., Mohd Jaafar, F., Cantaloube, J. F., Biagini, P., de Micco, P. & de Lamballerie, X. (2002) *J Gen Virol* **83,** 1941-51.
5. Attoui, H., Maan, S., Simon, J. A. & Mertens, P. P. C. (2009) *In: P.S. Mellor, M. Baylis, P.P.C. Mertens (eds.), Bluetongue.London: Elsevier/Academic Press* **3,** 23-52.
6. Tate, J. E., Burton, A. H., Boschi-Pinto, C., Steele, A. D., Duque, J. & Parashar, U. D. (2012) *Lancet Infect Dis* **12,** 136-41.
7. Estes, M. K. & Kapikian, A. Z. (2007) *in: Editors: Knipe, David M., Howley, Peter M. Fields Virology (5th Edition). Publisher: Lippincott Williams & Wilkins.* **II,** 1918-1974.
8. Bishop, R. F., Davidson, G. P., Holmes, I. H. & Ruck, B. J. (1973) *Lancet* **2,** 1281-3.
9. Flewett, T. H., Bryden, A. S., Davies, H., Woode, G. N., Bridger, J. C. & Derrick, J. M. (1974) *Lancet* **2,** 61-3.
10. Yolken, R., Arango-Jaramillo, S., Eiden, J. & Vonderfecht, S. (1988) *J Infect Dis* **158,** 1120-3.
11. Ogden, K. M., Johne, R. & Patton, J. T. (2012) *Virology* **431,** 50-7.
12. Hoshino, Y. & Kapikian, A. Z. (1996) *Arch Virol Suppl* **12,** 99-111.
13. Estrozi, L. F., Settembre, E. C., Goret, G., McClain, B., Zhang, X., Chen, J. Z., Grigorieff, N. & Harrison, S. C. (2013) *J Mol Biol* **425,** 124-32.
14. Trask, S. D., Ogden, K. M. & Patton, J. T. (2012) *Curr Opin Virol* **2,** 373-79.
15. Pesavento, J. B., Lawton, J. A., Estes, M. E. & Venkataram Prasad, B. V. (2001) *Proc Natl Acad Sci U S A* **98,** 1381-6.
16. Kapikian, A. Z., Kalica, A. R., Shih, J. W., Cline, W. L., Thornhill, T. S., Wyatt, R. G., Chanock, R. M., Kim, H. W. & Gerin, J. L. (1976) *Virology* **70,** 564-9.
17. Prasad, B. V., Wang, G. J., Clerx, J. P. & Chiu, W. (1988) *J Mol Biol* **199,** 269-75.
18. Yeager, M., Dryden, K. A., Olson, N. H., Greenberg, H. B. & Baker, T. S. (1990) *J Cell Biol* **110,** 2133-44.
19. Xiaofang, J., Crawford, S., Estes, B. & Prasad, V. (2008) *In:Patton J.ed. Segmented Double-stranded RNA Viruses. Caister Academic Press***,** 45-61.
20. Mertens, P., Attoui, H. & Bamford, D. H. (2003) *http://www.reoviridae.org/dsRNA_virus_proteins/Rotavirus.htm*.
21. Mathieu, M., Petitpas, I., Navaza, J., Lepault, J., Kohli, E., Pothier, P., Prasad, B. V., Cohen, J. & Rey, F. A. (2001) *Embo J* **20,** 1485-97.
22. Uzri, D. & Greenberg, H. B. (2013) *PLoS One* **8,** e69825.

23. Diprose, J. M., Burroughs, J. N., Sutton, G. C., Goldsmith, A., Gouet, P., Malby, R., Overton, I., Zientara, S., Mertens, P. P., Stuart, D. I. & Grimes, J. M. (2001) *Embo J* **20,** 7229-39.
24. Prasad, B. V., Rothnagel, R., Zeng, C. Q., Jakana, J., Lawton, J. A., Chiu, W. & Estes, M. K. (1996) *Nature* **382,** 471-3.
25. Bamford, D. H. (2002) *EMBO Rep* **3,** 317-8.
26. McClain, B., Settembre, E., Temple, B. R., Bellamy, A. R. & Harrison, S. C. (2010) *J Mol Biol* **397,** 587-99.
27. Chen, D., Luongo, C. L., Nibert, M. L. & Patton, J. T. (1999) *Virology* **265,** 120-30.
28. Lawton, J. A., Estes, M. K. & Prasad, B. V. (2001) *J Virol* **75,** 1632-42.
29. McDonald, S. M. & Patton, J. T. (2011) *J Virol* **85,** 3095-105.
30. Lu, X., McDonald, S. M., Tortorici, M. A., Tao, Y. J., Vasquez-Del Carpio, R., Nibert, M. L., Patton, J. T. & Harrison, S. C. (2008) *Structure* **16,** 1678-88.
31. Settembre, E. C., Chen, J. Z., Dormitzer, P. R., Grigorieff, N. & Harrison, S. C. (2011) *Embo J* **30,** 408-16.
32. Poruchynsky, M. S. & Atkinson, P. H. (1991) *J Virol* **65,** 4720-7.
33. Crawford, S. E., Mukherjee, S. K., Estes, M. K., Lawton, J. A., Shaw, A. L., Ramig, R. F. & Prasad, B. V. (2001) *J Virol* **75,** 6052-61.
34. Dormitzer, P. R., Nason, E. B., Prasad, B. V. & Harrison, S. C. (2004) *Nature* **430,** 1053-8.
35. Dormitzer, P. R., Sun, Z. Y., Wagner, G. & Harrison, S. C. (2002) *Embo J* **21,** 885-97.
36. Aoki, S. T., Settembre, E. C., Trask, S. D., Greenberg, H. B., Harrison, S. C. & Dormitzer, P. R. (2009) *Science* **324,** 1444-7.
37. Trask, S. D. & Dormitzer, P. R. (2006) *J Virol* **80,** 11293-304.
38. Wolf, M., Vo, P. T. & Greenberg, H. B. (2011) *J Virol* **85,** 2492-503.
39. Li, W., Manktelow, E., von Kirchbach, J. C., Gog, J. R., Desselberger, U. & Lever, A. M. (2010) *Nucleic Acids Res* **38,** 7718-35.
40. McIntyre, M., Rosenbaum, V., Rappold, W., Desselberger, M., Wood, D. & Desselberger, U. (1987) *J Gen Virol* **68 (Pt 11),** 2961-6.
41. Estes, M. K. & Cohen, J. (1989) *Microbiol Rev* **53,** 410-49.
42. Imai, M., Akatani, K., Ikegami, N. & Furuichi, Y. (1983) *J Virol* **47,** 125-36.
43. Mattion, N. M., Mitchell, D. B., Both, G. W. & Estes, M. K. (1991) *Virology* **181,** 295-304.
44. Ghosh, S., Kobayashi, N., Nagashima, S., Chawla-Sarkar, M., Krishnan, T., Ganesh, B. & Naik, T. N. (2010) *Arch Virol* **155,** 159-67.
45. McCrae, M. A. & Faulkner-Valle, G. P. (1981) *J Virol* **39,** 490-6.
46. Jourdan, N., Maurice, M., Delautier, D., Quero, A. M., Servin, A. L. & Trugnan, G. (1997) *J Virol* **71,** 8268-78.
47. Dormitzer, P. R. (2008) *In: J. Patton (eds), Segmented dsRNA viruses. Structure and molecular biology. Caister Academic Press.* **12,** 189-214.
48. Trask, S. D., McDonald, S. M. & Patton, J. T. (2012) *Nat Rev Microbiol* **10,** 165-77.
49. Arias, C. F., Romero, P., Alvarez, V. & Lopez, S. (1996) *J Virol* **70,** 5832-9.
50. Trask, S. D., Kim, I. S., Harrison, S. C. & Dormitzer, P. R. (2010) *J Virol* **84,** 1764-70.
51. Yoder, J. D. & Dormitzer, P. R. (2006) *Embo J* **25,** 1559-68.

52. Pesavento, J. B., Crawford, S. E., Roberts, E., Estes, M. K. & Prasad, B. V. (2005) *J Virol* **79,** 8572-80.
53. Cohen, J., Laporte, J., Charpilienne, A. & Scherrer, R. (1979) *Arch Virol* **60,** 177-86.
54. Dormitzer, P. R., Sun, Z. Y., Blixt, O., Paulson, J. C., Wagner, G. & Harrison, S. C. (2002) *J Virol* **76,** 10512-7.
55. Bass, D. M., Mackow, E. R. & Greenberg, H. B. (1991) *Virology* **183,** 602-10.
56. Rolsma, M. D., Kuhlenschmidt, T. B., Gelberg, H. B. & Kuhlenschmidt, M. S. (1998) *J Virol* **72,** 9079-91.
57. Ciarlet, M., Crawford, S. E. & Estes, M. K. (2001) *J Virol* **75,** 11834-50.
58. Ciarlet, M. & Estes, M. K. (1999) *J Gen Virol* **80 (Pt 4),** 943-8.
59. Graham, K. L., Fleming, F. E., Halasz, P., Hewish, M. J., Nagesha, H. S., Holmes, I. H., Takada, Y. & Coulson, B. S. (2005) *J Gen Virol* **86,** 3397-408.
60. Graham, K. L., Halasz, P., Tan, Y., Hewish, M. J., Takada, Y., Mackow, E. R., Robinson, M. K. & Coulson, B. S. (2003) *J Virol* **77,** 9969-78.
61. Beaulieu, J. F. (1992) *J Cell Sci* **102 (Pt 3),** 427-36.
62. Coulson, B. S., Londrigan, S. L. & Lee, D. J. (1997) *Proc Natl Acad Sci U S A* **94,** 5389-94.
63. Hewish, M. J., Takada, Y. & Coulson, B. S. (2000) *J Virol* **74,** 228-36.
64. Zarate, S., Romero, P., Espinosa, R., Arias, C. F. & Lopez, S. (2004) *J Virol* **78,** 10839-47.
65. Zarate, S., Cuadras, M. A., Espinosa, R., Romero, P., Juarez, K. O., Camacho-Nuez, M., Arias, C. F. & Lopez, S. (2003) *J Virol* **77,** 7254-60.
66. Kaljot, K. T., Shaw, R. D., Rubin, D. H. & Greenberg, H. B. (1988) *J Virol* **62,** 1136-44.
67. Sanchez-San Martin, C., Lopez, T., Arias, C. F. & Lopez, S. (2004) *J Virol* **78,** 2310-8.
68. Dowling, W., Denisova, E., LaMonica, R. & Mackow, E. R. (2000) *J Virol* **74,** 6368-76.
69. Dormitzer, P. R., Greenberg, H. B. & Harrison, S. C. (2000) *Virology* **277,** 420-8.
70. Keljo, D. J. & Smith, A. K. (1988) *J Pediatr Gastroenterol Nutr* **7,** 249-56.
71. Keljo, D. J., Kuhn, M. & Smith, A. (1988) *J Pediatr Gastroenterol Nutr* **7,** 257-63.
72. Tao, Y., Farsetta, D. L., Nibert, M. L. & Harrison, S. C. (2002) *Cell* **111,** 733-45.
73. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H. & Stuart, D. I. (2001) *Nature* **410,** 235-40.
74. Boyce, M., Wehrfritz, J., Noad, R. & Roy, P. (2004) *J Virol* **78,** 3994-4002.
75. McDonald, S. M., Tao, Y. J. & Patton, J. T. (2009) *Curr Opin Struct Biol* **19,** 775-82.
76. Tortorici, M. A., Broering, T. J., Nibert, M. L. & Patton, J. T. (2003) *J Biol Chem* **278,** 32673-82.
77. Patton, J. T., Jones, M. T., Kalbach, A. N., He, Y. W. & Xiaobo, J. (1997) *J Virol* **71,** 9618-26.
78. Chen, J. Z., Settembre, E. C., Aoki, S. T., Zhang, X., Bellamy, A. R., Dormitzer, P. R., Harrison, S. C. & Grigorieff, N. (2009) *Proc Natl Acad Sci U S A* **106,** 10644-8.

79. Jayaram, H., Estes, M. K. & Prasad, B. V. (2004) *Virus Res* **101,** 67-81.
80. Deo, R. C., Groft, C. M., Rajashankar, K. R. & Burley, S. K. (2002) *Cell* **108,** 71-81.
81. Harb, M., Becker, M. M., Vitour, D., Baron, C. H., Vende, P., Brown, S. C., Bolte, S., Arold, S. T. & Poncet, D. (2008) *J Virol* **82,** 11283-93.
82. Piron, M., Vende, P., Cohen, J. & Poncet, D. (1998) *Embo J* **17,** 5811-21.
83. Poncet, D., Aponte, C. & Cohen, J. (1993) *J Virol* **67,** 3159-65.
84. Gonzalez, R. A., Torres-Vega, M. A., Lopez, S. & Arias, C. F. (1998) *Arch Virol* **143,** 981-96.
85. Gonzalez, R. A., Espinosa, R., Romero, P., Lopez, S. & Arias, C. F. (2000) *Arch Virol* **145,** 1963-73.
86. Hu, L., Crawford, S. E., Hyser, J. M., Estes, M. K. & Prasad, B. V. (2012) *Curr Opin Virol* **2,** 380-8.
87. Stirzaker, S. C. & Both, G. W. (1989) *Cell* **56,** 741-7.
88. Xu, A., Bellamy, A. R. & Taylor, J. A. (2000) *Embo J* **19,** 6465-74.
89. Storey, S. M., Gibbons, T. F., Williams, C. V., Parr, R. D., Schroeder, F. & Ball, J. M. (2007) *J Virol* **81,** 5472-83.
90. Bergmann, C. C., Maass, D., Poruchynsky, M. S., Atkinson, P. H. & Bellamy, A. R. (1989) *Embo J* **8,** 1695-703.
91. Mirazimi, A., Nilsson, M. & Svensson, L. (1998) *J Virol* **72,** 8705-9.
92. Hyser, J. M., Collinson-Pautz, M. R., Utama, B. & Estes, M. K. (2010) *MBio* **1**. e00265-10. doi: 10.1128/mBio.00265-10
93. Parr, R. D., Storey, S. M., Mitchell, D. M., McIntosh, A. L., Zhou, M., Mir, K. D. & Ball, J. M. (2006) *J Virol* **80,** 2842-54.
94. Musalem, C. & Espejo, R. T. (1985) *J Gen Virol* **66 ,** 2715-24.
95. Luque, D., Rivas, G., Alfonso, C., Carrascosa, J. L., Rodriguez, J. F. & Caston, J. R. (2009) *Proc Natl Acad Sci U S A* **106,** 2148-52.
96. Bahar, M. W., Sarin, L. P., Graham, S. C., Pang, J., Bamford, D. H., Stuart, D. I. & Grimes, J. M. (2013) *J Virol* **87,** 3229-36.
97. McDonald, S. M. & Patton, J. T. (2011) *Trends Microbiol* **19,** 136-44.
98. Borodavka, A., Tuma, R. & Stockley, P. G. (2013) *RNA Biol* **10,** 481-9.
99. Silvestri, L. S., Taraporewala, Z. F. & Patton, J. T. (2004) *J Virol* **78,** 7763-74.
100. Gallegos, C. O. & Patton, J. T. (1989) *Virology* **172,** 616-27.
101. Taraporewala, Z. F. & Patton, J. T. (2004) *Virus Res* **101,** 57-66.
102. Taraporewala, Z. F. & Patton, J. T. (2001) *J Virol* **75,** 4519-27.
103. Patton, J. T. (2008) *Caister Academic Press. Norfolk, UK***,** 1-373.
104. Taraporewala, Z., Chen, D. & Patton, J. T. (1999) *J Virol* **73,** 9934-43.
105. Vasquez-Del Carpio, R., Gonzalez-Nilo, F. D., Riadi, G., Taraporewala, Z. F. & Patton, J. T. (2006) *J Mol Biol* **362,** 539-54.
106. Arnoldi, F., Campagna, M., Eichwald, C., Desselberger, U. & Burrone, O. R. (2007) *J Virol* **81,** 2128-37.
107. Vende, P., Taraporewala, Z. F. & Patton, J. T. (2002) *J Virol* **76,** 5291-9.
108. Zeng, C. Q., Estes, M. K., Charpilienne, A. & Cohen, J. (1998) *J Virol* **72,** 201-8.
109. Desselberger, U. (1996) *Adv Virus Res* **46,** 69-95.
110. Roner, M. R. & Steele, B. G. (2007) *Virology* **358,** 89-97.
111. Noda, T., Sagara, H., Yen, A., Takada, A., Kida, H., Cheng, R. H. & Kawaoka, Y. (2006) *Nature* **439,** 490-2.

112. Hutchinson, E. C., von Kirchbach, J. C., Gog, J. R. & Digard, P. (2010) *J Gen Virol* **91,** 313-28.
113. Patton, J. T. & Gallegos, C. O. (1990) *J Gen Virol* **71,** 1087-94.
114. Mindich, L. (2004) *Virus Res* **101,** 83-92.
115. Gottlieb, P., Qiao, X., Strassman, J., Frilander, M. & Mindich, L. (1994) *Virology* **200,** 42-7.
116. Mindich, L. (1999) *Microbiol Mol Biol Rev* **63,** 149-60.
117. Lakowicz, J. R. (2006) *Principles of Fluorescence Microscopy.*Third edition. Springer, 1-25.
118. Turro, N. J., Ramamurthy, V. & Scaiano, J. (2010) *Modern Molecular Photochemistry of Organic Molecules.* University Science Books, 1-16.
119. Benesch, R. E. & Benesch, R. (1953) *Science* **118,** 447-8.
120. Cordes, T., Vogelsang, J. & Tinnefeld, P. (2009) *J Am Chem Soc* **131,** 5018-9.
121. Rasnik, I., McKinney, S. A. & Ha, T. (2006) *Nat Methods* **3,** 891-3.
122. Le Reste, L. (2010) in *Biological Physics*, ed. Oxford, U. o. (University of Oxford Oxford), DPhil, pp. 1-174.
123. Weiss, S. (1999) *Science* **283,** 1676-83.
124. Stryer, L. & Haugland, R. P. (1967) *Proc Natl Acad Sci U S A* **58,** 719-26.
125. Clegg, R. M. (1992) *Methods Enzymol* **211,** 353-88.
126. Iqbal, A., Arslan, S., Okumus, B., Wilson, T. J., Giraud, G., Norman, D. G., Ha, T. & Lilley, D. M. (2008) *Proc Natl Acad Sci U S A* **105,** 11176-81.
127. Dale, R. E., Eisinger, J. & Blumberg, W. E. (1979) *Biophys J* **26,** 161-93.
128. Axelrod, D. (2001) *Traffic* **2,** 764-74.
129. Mattheyses, A. L., Simon, S. M. & Rappoport, J. Z. (2010) *J Cell Sci* **123,** 3621-8.
130. Lee, N. K., Kapanidis, A. N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R. H. & Weiss, S. (2005) *Biophys J* **88,** 2939-53.
131. Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E. & Weiss, S. (2004) *Proc Natl Acad Sci U S A* **101,** 8936-41.
132. Kapanidis, A. N., Laurence, T. A., Lee, N. K., Margeat, E., Kong, X. & Weiss, S. (2005) *Acc Chem Res* **38,** 523-33.
133. Hohlbein, J., Craggs, T. D. & Cordes, T. (2013) *Chem Soc Rev***,** DOI: 10.1039/c3cs60233h (Epub ahead of print).
134. Holden, S. J., Uphoff, S., Hohlbein, J., Yadin, D., Le Reste, L., Britton, O. J. & Kapanidis, A. N. (2010) *Biophys J* **99,** 3102-11.
135. Crawford, R. (2011) in *Biological Physics*, ed. Oxford, U. o. (University of Oxford Oxford), DPhil, pp. 1-150.
136. Selvin P, H. T. (2008) *Single Molecule Techniques: A Laboratory Manual* (Cold Spring Harbor laboratory Press
137. Muschielok, A., Andrecka, J., Jawhari, A., Bruckner, F., Cramer, P. & Michaelis, J. (2008) *Nat Methods* **5,** 965-71.
138. Guglielmi, K. M., McDonald, S. M. & Patton, J. T. (2010) *J Biol Chem* **285,** 18123-8.
139. Kapanidis, A. N., Margeat, E., Laurence, T. A., Doose, S., Ho, S. O., Mukhopadhyay, J., Kortkhonjia, E., Mekler, V., Ebright, R. H. & Weiss, S. (2005) *Mol Cell* **20,** 347-56.
140. Cooper, M., Ebner, A., Briggs, M., Burrows, M., Gardner, N., Richardson, R. & West, R. (2004) *J Fluoresc* **14,** 145-50.

141. ATTO-TECH (2009) *ATTO647N product information.*
142. Gell, C., Brockwell, D. & Smith, A. (2006) *Handbook of single molecule fluorescence spectroscopy. OUP.Oxford***,** 159-96.
143. Boyce, M., Celma, C. C. & Roy, P. (2008) *J Virol* **82,** 8339-48.
144. Maan, S., Rao, S., Maan, N. S., Anthony, S. J., Attoui, H., Samuel, A. R. & Mertens, P. P. (2007) *J Virol Methods* **143,** 132-9.
145. Patton, J., Chizhikov, V., Taraporewala, Z. & Chen, D. (2000) *In:Gray, J & Desselberger U.eds. Methods and Protocols: Virus Replication. Humana Press* **34,** 33-66.
146. Fadouloglou, V. E., Kokkinidis, M. & Glykos, N. M. (2008) *Anal Biochem* **373,** 404-6.
147. Choi, H. M., Chang, J. Y., Trinh le, A., Padilla, J. E., Fraser, S. E. & Pierce, N. A. (2010) *Nat Biotechnol* **28,** 1208-12.
148. Chou, Y. Y., Vafabakhsh, R., Doganay, S., Gao, Q., Ha, T. & Palese, P. (2012) *Proc Natl Acad Sci U S A* **109,** 9101-6.
149. Stoughton, R. B. (2005) *Annu Rev Biochem* **74,** 53-82.
150. Skinner, S. O., Sepulveda, L. A., Xu, H. & Golding, I. *Nat Protoc* **8,** 1100-13.
151. Tyagi, S. & Kramer, F. R. (2012) *F1000 Med Rep* **4,** 10.
152. Tyagi, S. & Kramer, F. R. (1996) *Nat Biotechnol* **14,** 303-8.
153. Santoso, Y., Hwang, L. C., Le Reste, L. & Kapanidis, A. N. (2008) *Biochem Soc Trans* **36,** 738-44.
154. Roy, R., Hohng, S. & Ha, T. (2008) *Nat Methods* **5,** 507-16.
155. Mason, B. B., Graham, D. Y. & Estes, M. K. (1980) *J Virol* **33,** 1111-21.
156. Kebbekus, P., Draper, D. E. & Hagerman, P. (1995) *Biochemistry* **34,** 4354-7.
157. Porschke, D. & Eigen, M. (1971) *J Mol Biol* **62,** 361-81.
158. Amann, R. & Fuchs, B. M. (2008) *Nat Rev Microbiol* **6,** 339-48.
159. Levsky, J. M. & Singer, R. H. (2003) *J Cell Sci* **116,** 2833-8.
160. Clamme, J. P. & Deniz, A. A. (2005) *Chemphyschem* **6,** 74-7.
161. Johnson, M. A. & McCrae, M. A. (1989) *J Virol* **63,** 2048-55.
162. Lawton, J. A., Zeng, C. Q., Mukherjee, S. K., Cohen, J., Estes, M. K. & Prasad, B. V. (1997) *J Virol* **71,** 7353-60.
163. Lawton, J. A., Estes, M. K. & Prasad, B. V. (1999) *Proc Natl Acad Sci U S A* **96,** 5428-33.
164. Thouvenin, E., Schoehn, G., Rey, F., Petitpas, I., Mathieu, M., Vaney, M. C., Cohen, J., Kohli, E., Pothier, P. & Hewat, E. (2001) *J Mol Biol* **307,** 161-72.
165. Uphoff, S., Holden, S. J., Le Reste, L., Periz, J., van de Linde, S., Heilemann, M. & Kapanidis, A. N. (2010) *Nat Methods* **7,** 831-6.
166. Gouet, P., Diprose, J. M., Grimes, J. M., Malby, R., Burroughs, J. N., Zientara, S., Stuart, D. I. & Mertens, P. P. (1999) *Cell* **97,** 481-90.
167. Lourenco, S. & Roy, P. (2011) *Proc Natl Acad Sci U S A* **108,** 13746-51.
168. Robb, N. C., Cordes, T., Hwang, L. C., Gryte, K., Duchi, D., Craggs, T. D., Santoso, Y., Weiss, S., Ebright, R. H. & Kapanidis, A. N. (2013) *J Mol Biol* **425,** 875-85.
169. Kapanidis, A. N., Margeat, E., Ho, S. O., Kortkhonjia, E., Weiss, S. & Ebright, R. H. (2006) *Science* **314,** 1144-7.

170. Cordes, T., Santoso, Y., Tomescu, A. I., Gryte, K., Hwang, L. C., Camara, B., Wigneshweraraj, S. & Kapanidis, A. N. (2010) *Biochemistry* **49,** 9171-80.
171. Margeat, E., Kapanidis, A. N., Tinnefeld, P., Wang, Y., Mukhopadhyay, J., Ebright, R. H. & Weiss, S. (2006) *Biophys J* **90,** 1419-31.
172. Borukhov, S. & Nudler, E. (2008) *Trends Microbiol* **16,** 126-34.
173. Busby, S. & Ebright, R. H. (1994) *Cell* **79,** 743-6.
174. Burgess, R. R. & Anthony, L. (2001) *Curr Opin Microbiol* **4,** 126-31.
175. Gross, C. A., Lonetto, M. & Losick, R. (1998) *In: Yamamoto,K and McKnight (eds).Transcriptional Regulation. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.***,** 447-54.
176. Paget, M. S. & Helmann, J. D. (2003) *Genome Biol* **4,** 203.
177. deHaseth, P. L. & Helmann, J. D. (1995) *Mol Microbiol* **16,** 817-24.
178. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. (2002) *Science* **296,** 1285-90.
179. Murakami, K. S., Masuda, S. & Darst, S. A. (2002) *Science* **296,** 1280-4.
180. deHaseth, P. L., Zupancic, M. L. & Record, M. T., Jr. (1998) *J Bacteriol* **180,** 3019-25.
181. Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. (2006) *Cell* **127,** 941-54.
182. Revyakin, A., Liu, C., Ebright, R. H. & Strick, T. R. (2006) *Science* **314,** 1139-43.
183. Hsu, L. M., Vo, N. V., Kane, C. M. & Chamberlin, M. J. (2003) *Biochemistry* **42,** 3777-86.
184. Nudler, E., Mustaev, A., Lukhtanov, E. & Goldfarb, A. (1997) *Cell* **89,** 33-41.
185. Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A. & Darst, S. A. (2000) *Science* **289,** 619-25.
186. Kireeva, M. L., Komissarova, N., Waugh, D. S. & Kashlev, M. (2000) *J Biol Chem* **275,** 6530-6.
187. von Hippel, P. H. & Pasman, Z. (2002) *Biophys Chem* **101-102,** 401-23.
188. Nudler, E. *Cell* **149,** 1438-45.
189. Straney, D. C. & Crothers, D. M. (1985) *Cell* **43,** 449-59.
190. Metzger, W., Schickor, P., Meier, T., Werel, W. & Heumann, H. (1993) *J Mol Biol* **232,** 35-49.
191. Mukhopadhyay, J., Kapanidis, A. N., Mekler, V., Kortkhonjia, E., Ebright, Y. W. & Ebright, R. H. (2001) *Cell* **106,** 453-63.
192. Bar-Nahum, G. & Nudler, E. (2001) *Cell* **106,** 443-51.
193. Callaci, S., Heyduk, E. & Heyduk, T. (1998) *J Biol Chem* **273,** 32995-3001.
194. Mukhopadhyay, J., Mekler, V., Kortkhonjia, E., Kapanidis, A. N., Ebright, Y. W. & Ebright, R. H. (2003) *Methods Enzymol* **371,** 144-59.
195. Kapanidis, A. N. & Strick, T. (2009) *Trends Biochem Sci* **34,** 234-43.
196. Kapanidis, A. N., Ebright, Y. W., Ludescher, R. D., Chan, S. & Ebright, R. H. (2001) *J Mol Biol* **312,** 453-68.
197. Friedman, L. J. & Gelles, J. (2012) *Cell* **148,** 679-89.
198. Jensen, M. A., Fukushima, M. & Davis, R. W. (2010) *PLoS One* **5,** e11024.
199. Varadaraj, K. & Skinner, D. M. (1994) *Gene* **140,** 1-5.
200. Winship, P. R. (1989) *Nucleic Acids Res* **17,** 1266.
201. Dave, N. & Liu, J. (2010) *J Phys Chem B* **114,** 15694-9.

202. Jain, A., Liu, R., Ramani, B., Arauz, E., Ishitsuka, Y., Ragunathan, K., Park, J., Chen, J., Xiang, Y. K. & Ha, T. (2011) *Nature* **473,** 484-8.
203. Jain, A., Liu, R., Xiang, Y. K. & Ha, T. (2012) *Nat Protoc* **7,** 445-52.
204. Richards, J. E., Desselberger, U. & Lever, A. M. (2013) *PLoS One* **8,** e74328.
205. Lymperopoulos, K., Wirblich, C., Brierley, I. & Roy, P. (2003) *J Biol Chem* **278,** 31722-30.
206. Zwirglmaier, K. (2005) *FEMS Microbiol Lett* **246,** 151-8.
207. Chou, Y. Y., Heaton, N. S., Gao, Q., Palese, P., Singer, R. & Lionnet, T. (2013) *PLoS Pathog* **9,** e1003358.
208. Zenklusen, D. & Singer, R. H. (2010) *Methods Enzymol* **470,** 641-59.
209. Zenklusen, D., Larson, D. R. & Singer, R. H. (2008) *Nat Struct Mol Biol* **15,** 1263-71.
210. Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S. (2008) *Nat Methods* **5,** 877-9.
211. Wagner, M., Horn, M. & Daims, H. (2003) *Curr Opin Microbiol* **6,** 302-9.
212. Winz, M. L., Samanta, A., Benzinger, D. & Jaschke, A. (2012) *Nucleic Acids Res* **40,** e78 0 online method only.
213. van de Linde, S., Loschberger, A., Klein, T., Heidbreder, M., Wolter, S., Heilemann, M. & Sauer, M. (2011) *Nat Protoc* **6,** 991-1009.
214. Churchman, L. S., Okten, Z., Rock, R. S., Dawson, J. F. & Spudich, J. A. (2005) *Proc Natl Acad Sci U S A* **102,** 1419-23.
215. Huang, B., Wang, W., Bates, M. & Zhuang, X. (2008) *Science* **319,** 810-3.
216. Shtengel, G., Galbraith, J. A., Galbraith, C. G., Lippincott-Schwartz, J., Gillette, J. M., Manley, S., Sougrat, R., Waterman, C. M., Kanchanawong, P., Davidson, M. W., Fetter, R. D. & Hess, H. F. (2009) *Proc Natl Acad Sci U S A* **106,** 3125-30.