JACOBS UNIVERSITY

Biochemical studies of the social amoeba *Dictyostelium discoideum*

by

JAN DIESEND

a thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
in Biochemistry

Approved Dissertation Committee

Prof. Dr. Christian Hammann
Ribogenetics, Jacobs University Bremen

Prof. Dr. Sebastian Springer
Biochemistry, Jacobs University Bremen

Prof. Dr. Stefanie Kaiser
Pharmaceutical Chemistry, Goethe University Frankfurt

Date of Defense: 11th of January 2022

Life Sciences & Chemistry

Statutory Declaration

Family Name, Given/First Name	Jan Diesend
Matriculationnumber	20331971
Type of thesis	PhD thesis

English: Declaration of authorship

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.

The Thesis has been written independently and has not been submitted at any other university for the conferral of a PhD degree; neither has the thesis been previously published in full.

German: Erklärung der Autorenschaft (Urheberschaft)

Ich erkläre hiermit, dass die vorliegende Arbeit ohne fremde Hilfe ausschließlich von mir erstellt und geschrieben worden ist. Jedwede verwendeten Quellen, direkter oder indirekter Art, sind als solche kenntlich gemacht worden. Mir ist die Tatsache bewusst, dass der Inhalt der Thesis in digitaler Form geprüft werden kann im Hinblick darauf, ob es sich ganz oder in Teilen um ein Plagiat handelt. Ich bin damit einverstanden, dass meine Arbeit in einer Datenbank eingegeben werden kann, um mit bereits bestehenden Quellen verglichen zu werden und dort auch verbleibt, um mit zukünftigen Arbeiten verglichen werden zu können. Dies berechtigt jedoch nicht zur Verwendung oder Vervielfältigung.

Diese Arbeit wurde in der vorliegenden Form weder einer anderen Prüfungsbehörde vorgelegt noch wurde das Gesamtdokument bisher veröffentlicht.

Bremen, Date

Jan Diesend

Acknowledgments

First and foremost, I would like to thank my academic supervisor Prof. Dr. Christian Hammann for his guidance and suggestions throughout these years. Thanks for your support in all pursued experiments and making this work possible in the first place.

Thanks to Prof. Dr. Sebastian Springer and Prof. Dr. Stefanie Kaiser for agreeing to be in my Dissertation committee and for the fruitful collaborations.

I would like to thank current and former members of the Ribogenetics Lab for their support and encouragement over the years. Special thanks to Monica Hagedorn for all the helpful discussions and support of the work.

In addition, I would like to thank my students: Gloria, Kshitij, Neela, and Nathalie. Separate thanks to Manuela for her continued support and making at times frustrating experiments more bearable.

The work for this thesis would not have been possible without the technical support of Tina Bidmon, Ursula Wellbrock, and Claudia Gerlich.

Finally, I would like to thank my friends and family for their continued support and of course special thanks to Lisa for her emotional support and patience throughout all these years.

Abstract

The amoeba *Dictyostelium discoideum* is a versatile and genetically tractable model organism used in research for many cellular processess. This work is focused on the biochemical study of two different processes in the amoeba.

The first part was the establishment of monitoring techniques for viral infections in the amoeba. For this purpose, two detection method should be established: a qPCR-based assay for the counting of the *Acanthamoeba polyphaga* mimivirus (APMV) and Tupanvirus (TPV) genomes and a high-throughput system for detection of giant viruses using flow cytometry. While qPCR-based genome counting was possible, flow cytometry was abandoned due to the high signal-to-noise ratios. In the qPCR-based monitoring of viral genome abundance it was observed that no net increase occurred in *D. discoideum* AX2 cells. Due to the better available toolbox, APMV was chosen to elucidate this phenomenon. APMV is taken up into the amoeba, however, the number of viral particles decreased until 6 hours post infection (hpi), indicating that APMV was not able to enter the replicative stage, possibly failing in the phagolysosomal pathway. Infection of phagolysosomal mutants indicated that lysosomal enzyme might be key in the defense against APMV. The RNAi machinery was not found to be involved.

The second part covers the potential ribosome heterogeneity during development of *D. discoideum* with a focus on the 2′-O-methylation (2′-O-Me) and pseudouridylation (Ψ). Both chemical modifications are introduced sequence-specifically by box C/D and box H/ACA small nucleolar RNPs (snoRNPs), respectively. Using bioinformatics and RNAseq analysis, 30 novel box C/D snoRNAs were identified in *D. discoideum*, many of which are differentially expressed during development. However, *in silico* approaches failed to deliver the responsible box H/ACA snoRNAs. Many box C/D snoRNAs were associated with a 2′-O-Me site. Modified positions were determined by the application of RiboMeth-seq for 2′-O-Me and HydraPsiSeq for Ψ. Several 2′-O-Me and Ψ sites were found to be substoichiometrically modified and some positions displayed dynamic modification levels during the development of the amoeba. All modified nucleotides were localized within the rRNA secondary structures predicted by homology-based modelling and many were found in the vicinity of integral structural elements. The U3 snoRNA, important for ribosome biogenesis, is featured in seven locis in the amoeba. Experiments and previous data implied a sudden generation of novel and potentially specialized ribosomes during development. The identification of U3 snoRNAs in representative members of the *Dictyostelia* allowed the establishment of a generalized secondary structure. Taken together, the presented data from *D. discoideum* is the first evidence for ribosome heterogeneity in the Amoebozoa supergroup, allowing to suggest that it is a common feature of all eukaryotes.

Zusammenfassung

Die Amöbe *Dictyostelium discoideum* ist ein vielseitiger und genetisch modifizierbarer Modellorganismus zur Untersuchung einer Vielzahl zellulärer Prozesse. Diese Arbeit beschäftigt sich mit zwei verschiedenen Aspekten in *Dictyostelium*.

Im ersten Teil der Arbeit wurden die Methoden zur Quantifizierung von Viren in *D. discoideum* etabliert. Zu diesem Zweck sollten zwei Methoden getestet werden: eine qPCR-basierte Zählung von *Acanthamoeba polyphaga* mimivirus (APMV)- und Tupanvirus (TPV)-Genomen und Durchflusszytometrie-basierte Detektion der Viren. Während eine qPCR-basierte Genomzählung möglich war, wurde die Durchflusszytometrie aufgrund des Hintergrundrauschens nicht weiterverfolgt. Die Experimente zeigten keinen Netto-Anstieg der Genomzahl an. APMV wurde aufgrund der besseren Werkzeug-Verfügbarkeit für die weitere Aufklärung dieses Phänomens gewählt. Während APMV in die Amöbe aufgenommen wurde, konnte bis 6 Stunden nach der Infektion eine Abnahme der Virenzahl beobachtet werden. Dies implizierte, dass APMV nicht in die replikative Phase eintrat und möglicherweise im Phagolysosom gescheitert war. Infektionen von phagolysosomalen Mutanten zeigten, dass potenziell die lysosomalen Enzyme eine Rolle in der APMV-Abwehr spielen.

Der zweite Teil dieser Arbeit befasste sich mit der potenziellen ribosomalen Heterogenität in der Entwicklung von *D. discoideum* mit Fokus auf die 2′-O-Methylierung (2′-O-Me) und die Pseudouridylierung (Ψ). Beide Modifikationen werden durch die box C/D und box H/ACA small nucleolar ribonucleoprotein (snoRNP)-Komplexe sequenzspezifisch durchgeführt. Durch bioinformatische und RNAseq-Analysen wurden 30 bisher unbekannte box C/D snoRNAs in *D. discoiduem* identifiziert und eine entwicklungsabhängige Expressionsänderung beobachtet, allerdings konnten box H/ACA snoRNA-Kandidation nicht gefunden werden. Die Modifikationsstellen wurden mit RiboMeth-seq für 2′-O-Me und HydraPsiSeq für Ψ identifiziert. Einige Positionen lagen substoichiometrisch modifiziert vor und einige zeigten eine entwicklungsspezifische Veränderung des Modifikationsgrades. Alle modifizierten Positionen wurden in den rRNA-Sekundärstrukturen lokalisiert und viele fanden sich in der Nähe wichtiger Strukturelemente. Die für die ribosomale Biogenese wichtige U3 snoRNA ist siebenmal im Genom von *D. discoiduem* kodiert. Die durchgeführten Experimente und bestehende Daten implizierten, dass die rRNA-Prozessierung im Entwicklungsprozess beschleunigt stattfinden könnte. Zudem gelang die Etablierung eines Sekundärstrukturmodells für die *Dictyostelia*. Zusammengefasst liefern die gesammelten Daten die ersten Beweise für die ribosomalen Heterogenität in den Amoebozoen. Dies erlaubt den Rückschluss, dass die Biogenese spezialisierter Ribosomen ein gemeinsames Merkmal aller Eukaryoten darstellt.

Publications and Manuscripts

Diesend, J., Kruse, J., Hagedorn, M., & Hammann, C. (2018). Amoebae, Giant Viruses, and Virophages Make Up a Complex, Multilayered Threesome. *Frontiers in Cellular and Infection Microbiology*, 7, 527. https://doi.org/10.3389/fcimb.2017.00527

Author contributions: All authors designed the mini review. JD: generated the figures and drafted the text; MH and CH: wrote the manuscript.

Borland, K., **Diesend, J.**, Ito-Kureha, T., Heissmeyer, V., Hammann, C., Buck, A. H., Michalakis, S., & Kellner, S. (2019). Production and Application of Stable Isotope-Labeled Internal Standards for RNA Modification Analysis. *Genes*, 10(1), 26. https://doi.org/10.3390/genes10010026

Author contributions: Conceptualization, K.B. and S.K.; methodology, K.B., S.K., T.I.-K., V.H. and J.D.; software, K.B. and S.K.; validation, K.B. and S.K.; formal analysis, K.B. and S.K.; investigation, K.B. and S.K.; resources, K.B., S.K., S.M., C.H., J.D. and A.H.B.; data curation, K.B. and S.K.; writing—original draft preparation, K.B. and S.K.; writing—review and editing, K.B., S.K., S.M., C.H., J.D., V.H.; visualization, K.B. and S.K.; supervision, S.K.; project administration, S.K.; funding acquisition, S.K., S.M., C.H., A.H.B.

Diesend, J., Birkedal, U., Kjellin, J., Zhang, J., Jablonski, K. P., Söderbom, F., Nielsen, H., & Hammann, C. (2022). Ribosome heterogeneity in Amoeboza: fractional 2′-O-Methylation in the ribosomal RNA of *Dictyostelium discoideum*. *Scientific Reports*, 12, 1952. https://doi.org/10.1038/s41598-022-05447-w

Author Contributions: JD designed and carried out most experiments; UB generated and analyzed the RMS data; JK analyzed RNAseq data; KPJ established tools for and performed bioinformatic analyses; JZ performed initial snoRNA analyses; FS, HN and CH supervised research. JD, HN and CH wrote the paper. All authors read and proved the final version of this manuscript.

Contents

Abbreviations

′	Prime
%	Percent
°C	Degree Celsius
2′-O-Me	Methylation of the 2′-hydroxyl of ribose
ACLaV	*Acanthamoeba castellani* lausannevirus
Agn	Argonaute proteins in *D. discoideum*
Amp	Ampicillin
AP	Alkaline phosphatase
APMV	*Acanthamoeba polyphaga* mimivirus
APMaV	*Acanthamoeba polyphaga* marseillevirus
APS	Ammonium persulfate
Atg1	Autophagy-related protein 1
ATP	Adenosine triphosphate
AX2	Axenic derivative of NC-4
BCIP	5-Brom-4-chlor-3-indolylphosphate
bp	Base pairs
Ca8V	Cannes 8 virus
cAMP	Cyclic adenosine monophosphate
CD RNA	Box C/D snoRNA of *D. discoideum* with target
CDMs	Cell-autonomous defense mechanisms
cDNA	Complementary DNA
CeV	Cedratvirus A11
CroV	*Cafeteria roenbergensis* virus
CS	Classifier score
Da	Dalton
DKC1	Human dyskerin
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
DrnB	Dicer-like protein B in *D. discoideum*
ds	Double stranded
DSE	Distal sequence element
DUSE	*D. discoideum* Upstream Sequence Element
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmatic reticulum
ES	Expansion segments
ETS	External transcribed spacer
FDR	False discovery rate
FL1	Fluorescent channel 1
FSC	Forward scatter
g	Gram
gDNA	Genomic DNA
GpdA	GAPDH

GTC	Guanidinium thiocyanate
h	Hour
hx	Helix in SSU rRNA with natural number x
Hx	Helix in LSU rRNA with natural number x
HerA	Helicase repair of Archaea
HPS	HydraPsiSeq
IP	Immunoprecipitation
ITS	Internal transcribed spacer
kb	Kilo base pairs
kDa	Kilo Dalton
Kil1	Sulfotransferase involved in bacterial killing
L	Liter
LB	Lysogeny broth
LRR	Leucine-rich repeats
LSM	Laser scanning microscope
LSU	Large subunit of the ribosome
M	Molarity [mol/L]
$m^1acp^3\Psi$	1-methyl-3-(3-amino-3-carboxypropyl)-Ψ
$m^{2,2,7}G$ cap	Trymethylguanosine cap
mA	Milliampere
Mavirus	Maverick-related virus
MBq	Mega Becquerel
MCP	Major capsid protein
MFE	Minimum free energy
mg	Milligram
min	Minute
mL	Milliliter
mM	Millimolar
mm	Millimeter
MOI	Multiplicity of infection
MOPS	3-(N-morpholino)propane sulfonic acid
MP	Maverick/Polinton retroelements
mRNA	Messenger RNA
NCLDV	Nucleocytoplasmatic large DNA viruses
ncRNA	Non-coding RNA
ng	Nanogram
nm	Nanometer
NMD	Nonsense-mediated decay
NNS termination	Nrd1-Nab3-Sen1 termination
Nramp1	Natural Resistance-Associated Macrophage Protein
NRO	Nuclear run-on transcription
nt	Nucleotide
NTP	Nucleoside triphosphate
OD	Optical density
OR	Box C/D snoRNA of *D. discoideum* without target

ORF	Open reading frame
P	Phosphate
p80	Putative copper transporter
PA	Polyacrylamide
PAGE	Polyacrylamide gel electrophoresis
PAMP	Pathogen-associated molecular pattern
PBS	Phosphate buffered saline
PCI	Phenol/chloroform/isoamyl alcohol
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PNK	Polynucleotide kinase
Pol II	RNA polymerase II
Pol III	RNA polymerase III
pre-mRNA	Precursor of mRNA
pre-rRNA	Precursor of rRNA
pre-snoRNA	Precursor of snoRNA
PSE	Proximal sequence element
Ψ	Pseudouridine
PTC	Peptidyl transferase center
qPCR	Quantitative PCR
R656	APMV-encoded amine oxidase
RacH	RAs-related-C3 botulinum toxin substrate H
RBP	RNA-binding protein
RdRP	RNA-dependent RNA polymerase
RMS	RiboMeth-seq
RNA	Ribonucleic acid
RNAi	RNA interference
rNTP	Ribonucleotide
RNV	Rio negro virophage
rpm	Revolutions per minute
rRNA	Ribosomal RNA
Rrp	RdRPs in *D. discoideum*
RT	Reverse transcriptase
SAM	S-adenosyl methionine
SDS	Sodium dodecyl sulfate
sec	Seconds
SNAREs	Soluble NSF attachment receptors
SNGV	Senegal virus
snRNA	Small nuclear RNA
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar ribonucleoprotein
SSC	Side scatter
SSU	Small subunit of the ribosome
TBE	Tris-Borat-EDTA
TEMED	Tetramethylethylenediamine

TE	Tris EDTA
TIR	Toll/interleukin 1 receptor
TLR	Toll-like receptor
TPV	Tupanvirus
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
tRNA	Transfer RNA
TSS	Transcription start site
u	Unit
UTP	Uridine triphosphate
UV	Ultraviolet
V	Volt
v/v	Volume/volume
VacB	Vacuolin B
VatA	V-ATPase A
w/v	Weight/volume
WASH	WASP and SCAR Homolog
WB	Western blot
x g	Gravity acceleration
α-32P	Alpha phosphate group with phosphor isotope ^{32}P
γ-32P	Gamma phosphate group with phosphor isotope ^{32}P
µg	Microgram
µL	Microliter

IUPAC nucleotide nomenclature:

A	adenine
C	cytosine
G	guanosine
T	thymine
U	uracil
R	A or G
Y	C or T
S	G or C
W	A or T
K	G or T
M	A or C
B	C or G or T
D	A or G or T
H	A or C or T
V	A or C or G
N	any nucleotide
. or -	gap

Introduction

Dictyostelium discoideum as a model organism

Dictyostelium discoideum is a well-established model organism of basic cellular processes discovered in 1935 [2, 3]. The AX2 strain has been established in 1970 and is able to grow in liquid medium, instead of being reliant on a bacterial food source [4]. A wide spectrum of experimental tools has been established for the amoeba [5], and these are frequently used the study of mechanisms governing cell motility, host-pathogen interactions, autophagy, social evolution (reviewed in [6]), mobile genetic elements [7], and their domestication by the RNA interference (RNAi) machinery [8-10]. *D. discoideum* cells usually propagate by mitotic division; upon starvation, however, a complex developmental process is initiated (Figure 1), in which about 100,000 cells aggregate to form within 24 h a fruiting body containing spores [11]. The organization of ribosomal DNA (rDNA) is unusual in *D. discoideum*. In most metazoans, the genes for the ribosomal RNAs (rRNAs) are organized in rDNA clusters, an arrangement that is thought to facilitate efficient rRNA transcription. Such rDNA clusters exist also in *D. discoideum*; however, they are not encoded in chromosomes but localized on extrachromosomal elements [12, 13]. Each nucleus contains about 100 copies of these elements of 88 kb, that each feature two rRNA transcription units organized as palindromes [14]. A first model for rRNA processing from the primary 37S transcript in the amoeba has been proposed, and sequences of the mature rRNAs in *D. discoideum* were determined experimentally [15]. In the last decades, *Dictyostelium* has been extensively used to study the molecular interactions and pathogenic bacteria like *Pseudomonas aeruginosa* [16, 17], *Legionella pneumophila* [18, 19], *Vibrio cholerae* [20], *Mycobacteria* [21-24], *Francisella noatunensis [25, 26]*, *Neisseria meningitidis* [27], *Burkholderia cenocepacia* [28], and *Salmonella* [29]. These bacteria are phagocytosed by the amoeba and subsequently avoid intracellular killing by modulating or escaping the phagosomal maturation pathway.

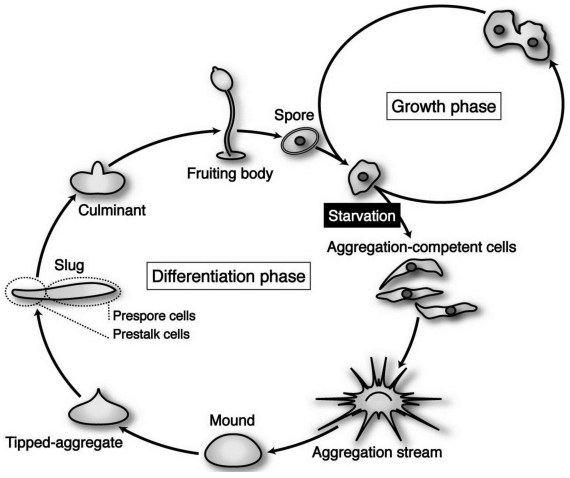


Figure 1. Development cycle of the amoeba *D. discoideum*. Shown are the vegetative growth phase and the development cycle induced upon starvation. The mobile slug stage is reached after 16 h and the fruiting body containing the spores is formed after 24 h. Taken from Maeda and Chida [1].

Introduction to giant viruses and their hosts

Giant viruses, their host, and their parasites: a multilayered interaction

This chapter is included in a manuscript previously published in "Frontiers in Cellular and Infection Microbiology":

Diesend, J., Kruse, J., Hagedorn, M., & Hammann, C. (2018). Amoebae, Giant Viruses, and Virophages Make Up a Complex, Multilayered Threesome. *Frontiers in Cellular and Infection Microbiology*, 7, 527. https://doi.org/10.3389/fcimb.2017.00527

Introduction to giant viruses

The discovery of giant viruses in the early 2000s led to a mind shift in the field of virology with respect to the potential origins of viruses [30, 31]. Originally, viruses were thought of as submicroscopic particles with a self-evident denial that viruses might exist, whose size would be large enough to be resolved with a simple light microscope [32, 33]. Due to this mindset, the large, gram-positive particles in an *Acanthamoeba polyphaga* population were at first erroneously classified as bacteria [31, 34, 35]. Only the absence of rDNA in the presumed bacterium, led to the discovery and definition of the *Acanthamoeba polyphaga* mimivirus (APMV) in 2003 [31]. The acronym mimivirus (for mimicking microbe) reflects the resemblance to bacteria upon gram staining. At the same time, the discovery of APMV was the first ever report of a virus infecting amoebae. Amongst other features that are detailed below, APMV is unusual as it contains a large genome of 1.14 Mbp, thereby even surpassing the genome size of some bacterial species [30]. APMV particles are characterized by an up to 700 nm large capsid (Figure 2A), which is well above the resolution of a simple light microscope. Once it was established that giant DNA viruses of amoeba exist, many more such viruses, belonging to the nucleocytoplasmatic large DNA viruses (NCLDV) were found in the environment, as well as within a wide range of host organisms from humans, monkeys, and oysters [36-38]. *Ex vivo* studies of human cell lines revealed that APMV is capable of infecting myeloid and mononuclear blood cells and interferes with the type I Interferon system [39]. In addition, a distantly APMV-related NCLDV family member has been shown to productively infect T-lymphocytes under laboratory conditions [40]. In 2008, a small viral particle was discovered in *A. polyphaga* and was identified to be a virus parasitizing upon viral factories of giant viruses called Sputnik 1 [41]. Due to the functional similarity to bacteriophages in

mediating lateral gene transfer, Sputnik was classified as a virophage [41]. Here, we will review the since 2008 expanding family of virophages and discuss the implications for giant virus reproduction inside amoebae.

The diverse families of giant viruses that infect amoebae

The discovery of APMV sparked the interest in giant viruses and spawned a contemporary research field of its own [31]. Up until today, two giant virus families belonging to the NCLDV have been described that primarily infect amoeba: the *Mimiviridae* and the *Marseilleviridae* (Figure 2B). The latter has the *A. polyphaga* marseillevirus (APMaV) as founding member, which was discovered in 2009 led to the classification of a second group of giant viruses termed *Marseilleviridae* [42, 43]. In the last decade, nine additional viruses have been associated with the *Marseilleviridae* group [44]. The *Acanthamoeba castellani* lausannevirus (ACLaV) was discovered by incubating water from the Seine river in France with *A. castellani*, a close relative of *A. polyphaga* [45]. ACLaV is the first known giant virus to encode histone-like proteins, which could point towards a DNA packaging mechanism similar to eukaryotes [45]. The Cannes 8 virus (Ca8V) [46] and the Senegal virus (SNGV) [47] have been isolated using similar co-culture methods and are grouped with the *Marseilleviridae*. The icosahedral capsid of the *Marseilleviridae* is between 190 and 250 nm in diameter [42]. Like the genome of the *Mimiviridae*, the 370,000 bp double-stranded DNA (dsDNA) genome is encased in a lipid bilayer and encodes about 450 proteins [43, 45-47]. Both, *Mimiviridae* and *Marseilleviridae*, share only nine core genes with all NCLDVs (Figure 2C) and 180 genes are shared with at least two of the NCLDV families [48, 49]. Based on the discovery of APMV and its complex genome, it was suggested to incorporate viruses into the tree of life by defining them as capsid-encoding organisms contrary to the ribosome-encoding organisms, which are represented by eukarya, bacteria, and archaea [50].

APMV – the best studied giant virus of amoeba

APMV was the first giant virus of amoeba to be discovered [41] and confronted the scientific community with features never observed in a virus before. Its capsid size and genetic complexity with many genes usually found in eukaryotic and prokaryotic cells challenged the Lwoff's characteristics of a virus [30, 50]. The AT-rich 1.14 Mbp APMV genome features an impressive number of 979 protein-encoding genes in a dense arrangement [30, 59]. Several of

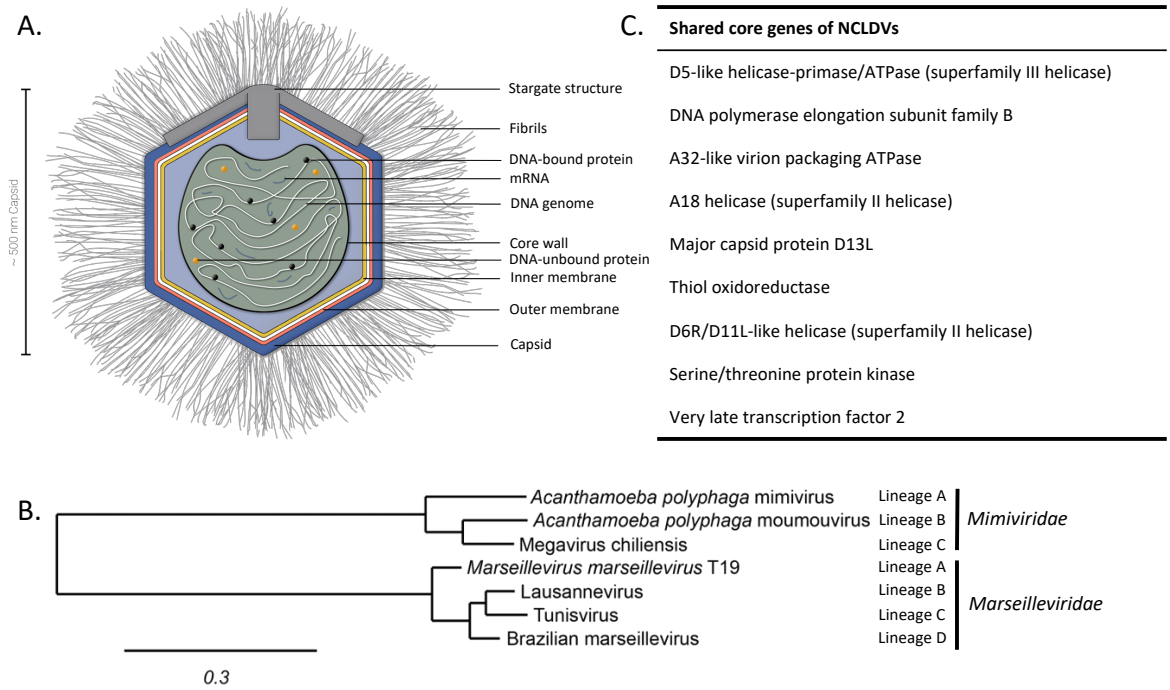


Figure 2. Structure of APMV and the core genes and relationship of giant viruses. (A) Viral particles of APMV feature a viral core with the genome, mRNAs and prefabricated proteins. This is surrounded by the indicated membranes and the capsid structure that contains a pentagonal, star-shaped structure termed 'stargate', which is involved in the release of the viral core into the host cell's cytosol upon phagocytosis. The capsid is decorated with a compact layer of fibrils. For details see main text. (B) Cladogram displaying the relationships of the different lineages of the Mimiviridae and Marseilleviridae. Since the discovery of APMV, over 100 new mimivirus strains have been characterized using samples of various origins in amoebal co-culture methods [51-53]. All Mimiviridae share a capsid size between 370 and 600 nm and a 1.02 to 1.26 Mb AT-rich genome which encodes about 1.000 putative proteins [44]. Based on sequence homology, the Mimiviridae can be divided into three distinct lineages (Figure 2B): lineage A with APMV as prototype and a total of 18 members, as reviewed recently [44], lineage B with the moumouvirus as prototype and four additional members [44, 54], and lineage C with Megavirus chiliensis as prototype and a total of 12 members [44, 55]. The tree was created using the sequences of the D13 major capsid proteins of the indicated prototype viruses using Phylogeny.fr, with the relative evolutionary distance indicated [56, 57]. (C) List of nine genes conserved throughout all NCLDV families. Figure taken from Diesend *et al.* [58].

its genes are only found in giant viruses of amoeba and code for virus-atypical proteins involved in DNA repair, protein folding, transfer RNA (tRNA) synthesis and translation, and more [30]. In addition, the APMV genome displays some plasticity and encodes self-splicing introns, inteins, and a specific set of mobile genetic elements called transpovirons [60]. Furthermore, the genome contains many genes likely acquired via horizontal gene transfer, paralogous genes, and so called ORFans, which are genes that encode proteins with unknown function [61-64]. Many of these genes are shared with the poxviruses, phycodnaviruses, and other NCLDVs [64]. ORFans represent roughly 50% of genes and about 40% of the APMV proteome, which results in a high number of factors with unknown functions that might act during viral replication and morphogenesis [65]. Alike "classical" viruses, APMV genes are partly under the control of early and late stage-specific promoters [30, 66].

The APMV particles possess remarkable structural features, separating them from the classical structures of viruses (Figure 2A). In its center, the viral DNA, messenger RNAs (mRNAs) and proteins are packed into the core compartment [67, 68] and enclosed by a lipid membrane. Among the pre-packed proteins are 12 enzymes involved in transcription, five in DNA repair, two in RNA modification, and five in protein modification [65]. The central compartment is surrounded by an approximately 340 nm-large lipid bilayer and a secondary bilayer directly underneath an icosahedral capsid. This is comprised of major capsid proteins (MCPs) and features a five-branch proteinaceous structure, the ‘stargate’, at one vortex [68]. The capsid itself is covered by a compact layer of about 120 to 140 nm long, heavily glycosylated fibrils, which potentially facilitate the attachment of APMV to its host cells [69].

Infection cycles of giant viruses in amoebae

Even though the replication cycle of most giant viruses differ in aspects like nuclear involvement, duration, assembly and release of the viral progeny, key steps in the infection appear to be conserved, as summarized recently [44]. For example, all known giant viruses enter the host cell by phagocytosis and release their DNA into the cytosol in a similar manner [70]. Furthermore, viral replication takes place in specialized endoplasmatic reticulum (ER)-derived compartments that are found in the cytosol and are called viral factories [67, 68, 71].

After uptake, the virus resides in a *de-novo* phagosome. Subsequently, the phagosomal and viral membranes fuse, which allows the release of the viral core, that contains the genome, proteins, and mRNAs into the cytosol [71, 72]. Alike the well-described poxvirus [73], the structural integrity of the viral core seems to be retained until viral factories arise [71, 74]. Intriguingly, recent experiments suggest that viral transcription might be initiated already before the release

of the viral core [75]. Once in the cytosol, replication of the viral genome begins immediately and the expression of early stage genes leads to the formation of early stage viral factories [75-77]. The replication cycle is confined to the cytosol, again a trait shared with the poxvirus [31, 74]. This also suggests that giant viruses (like the poxvirus) must carry transcription complexes to initiate transcription immediately after infection [74, 78]. In later stages of infection, these viral factories merge into one large cytosolic compartment for replication and capsid assembly [75, 76]. It should be noted that viral factories are not chaotic, but rather appear to feature distinct assembly lines for their progeny. The viral factory is made up of functional regions playing discrete roles in replication, capsid assembly, DNA packaging, and attachment of fibrils [75, 76]. In the outermost layer of the viral factory, the internal membrane layers of APMV are assembled from host-derived membrane vesicles, which are thought to rupture, thereby forming open single-layer membrane sheets [77]. Capsid assembly occurs around these membrane sheets and is scaffolded by the major capsid protein L425 [77]. Upon capsid formation, the genome is deposited into the empty viral particle through a transient interstice distal from the 'stargate' structure [72]. There is little evidence for a nuclear stage of giant viruses. However, the nuclei of *A. polyphaga* and *A. castellani* exhibit transient changes in their morphology during the early stages of infection with members of the *Marseilleviridae* family [79]. This indicates that nuclear host factors might play a role in the APMV replication, a notion that is supported by a two-fold decrease of the nuclear size in infected *A. polyphaga* cells [44]. This might be due to a substantial redistribution of nuclear factors for viral replication, transcription or other processes [44]. Albeit indirectly, this scenario is supported by data on the cytoplasmic replication of the Vaccinia virus (a poxvirus), to which mimivirus replication bears similarities [71] and for which the involvement of nuclear enzymes has been demonstrated [80].

Virophages as parasites of the Megavirales

The description of *Megavirales* infection of amoebae was followed by the discovery of the fascinating virophage Sputnik 1 in 2008 [41]. Sputnik 1 was found infecting the viral factories of the mamavirus, a close relative of APMV [41]. Replication of the Sputnik virophages inside APMV-infected *A. castellani* cells is deleterious to APMV replication and results in abortive DNA replication and disruption of capsid biogenesis [41]. Many published articles denote Sputnik and other virophages as satellite viruses [81-83]. Satellite viruses are characterized by their dependency on factors of a helper virus. However, the Sputnik genomes encodes factors involved in (DNA) replication [41]. Therefore, Sputnik is a full virus and not a defective viral particle or sub-viral agent. Furthermore, it was never observed that co-infection with a satellite

virus leads to the formation of defective helper virus particles [84]. Altogether, these observations suggest a virophage rather than a satellite virus [84, 85].

All known existing members of the dsDNA virophage family parasitizing on giant viruses are categorized into the large virus-dependent or -associated (*Lavida*-)*viridae* family divided into the *Sputnikvirus* and *Mavirus* genera [86]. At the species level, the *Sputnikvirus* genus can be differentiated into the APMV-dependent Sputnik virophage and the APMV-dependent Zamilon virophage, while *Mavirus* genus (not shown) contains only the *Cafeteria roenbergensis* virus (CroV)-dependent mavirus [86].

Sputnik replicates inside mamavirus-infected *A. castellani* cells within the viral factories, nonetheless, with different kinetics as the mamavirus and at multiple hot spots inside the factory [41]. Studies on amoebae infected with different mimiviruses revealed that the Sputnik virophages can parasitize on mimiviruses from all *Mimiviridae* lineages but not the *Marseilleviridae* lineages [87]. Productive infection of APMVs viral factories results in the emergence of newly generated Sputnik 1 particles six h post infection with a concomitant decrease of infective APMV particles [88]. The 18,343-kilobase circular dsDNA genome of the Sputnik virophages possesses 21 partly overlapping open reading frames (ORFs) encoding for several factors involved in DNA replication, e.g. a DNA polymerase or primase and a Tyr recombinase [41]. Furthermore, four of the 21 ORFs are strongly homologous to APMV-encoded genes [41, 87]. Since Sputnik virophages encode lambda-type integrases, the molecular tools for genome integration are present [41]. Until now, only the integration of the Sputnik 1 genome into the Lentillevirus genome – a relative of APMV - has been observed *in vitro* [60]. There is no indication of Sputnik genome integration into the host cell genome, yet.

The Zamilon virophage (belonging to the *Sputnikvirus* genus) was discovered together with the Mont1 virus which was recovered from soil samples from Tunisia [37, 89]. The 60 nm-wide, spherical virophage carries a 17,276 bp dsDNA genome encoding 20 genes. Although the Zamilon virophage shares 76% of its genomic sequence with the Sputnik virophages, Zamilon can only infect lineages B and C [89]. Furthermore, the tv_L8 protein, encoded in the transpovirons of the Monve virus (*Mimiviridae*, lineage B), shares significant homology with the ORF8-encoded protein of the Zamilon virophage [89]. This suggests that direct genetic interactions between the giant virus and the Zamilon virophage might occur within co-infected amoebae. However, no experimental data has confirmed such interactions so far. Recently, a new virophage, the Rio Negro virophage (RNV) has been discovered in Samba virus isolates from the Brazilian Amazon [90]. While the RNV genome is still incomplete, based on the

genomic sequence of its predicted major capsid proteins, it was associated with the *Sputnikvirus* genus [86, 90].

The Maverick-related virus (mavirus), belonging to the *Mavirus* genus, parasitizes upon the viral factories of CroV that infects the marine heterotrophic nanoflagellate *Cafeteria roenbergensis* [91, 92]. The 19.063 bp circular dsDNA mavirus genome possesses 20 ORFs including a retroviral rve-family integrase, a protein-primed DNA polymerase, a filamenting temperature-sensitive mutant K (FtsK)-helicase repair of Archaea (HerA)-type genome packaging ATPase, and a cysteine protease, and two major capsid proteins all also found conserved in Maverick/Polinton (MP) retroelements [86, 92, 93]. Additionally, the termini of the mavirus genome consist of long terminal repeats similar to those found in MP retroelements [86, 94]. Both findings suggest that these retroelements might have originated from mavirus genome integration events in mavirus co-infected cells (Fischer & Suttle, 2011, Krupovic et al., 2015). Nonetheless, this hypothesis for the origins of MP retroelements remains to be tested experimentally. Fischer and Hackl [95] succeeded to monitor complete mavirus genome integration into the *C. roenbergensis* genome by co-infection with a low multiplicity of infection (MOI) of CroV. Intriguingly, genes in the mavirus genome possess promoter sequences similar to the late stage promoter of CroV [95]. As a consequence, re-infection of *C. roenbergensis* carrying the integrated mavirus genome with CroV resulted in inhibition of viral DNA replication and increased host cell survival [95]. In contrast to Sputnik, the mavirus enters the amoebae independent of its viral host, likely passively via an endocytic pathway, and target the cytoplasmatic viral factory of CroV [92, 96].

Other virophages have been discovered by metagenomic analysis of water samples (e.g. the Organic Lake virophage [97], the Yellowstone Lake virophages 1-7 [98, 99]), however, the viral and cellular host platforms for these remain to be determined [86].

Known and potential defense mechanisms of *D. discoideum*

Phagocytosis and phagosomal maturation are basic cellular defense mechanisms and, accordingly, the molecular machinery is highly conserved between the amoeba and immune cells of higher eukaryotes. [100, 101]. These cellular defense mechanisms include the concerted action of phagocytosis/autophagy [102], which originally evolved from eukaryotic ancestors like *Dictyostelium* [100]. In addition, in the last decades evidence accumulated, that RNAi plays a role in the defense against intracellular pathogens in *D. melanogaster* and *C. elegans* [103-107]. This machinery is also highly conserved in *D. discoideum* [108], however, a role for the amoebal RNAi in cellular defense has not been established so far. In the following, the key

players of phagocytosis, autophagy, and RNAi in *Dictyostelium* and their potential in its defense mechanisms will be described.

Phagocytosis in *Dictyostelium*

Professional phagocytes like *D. discoideum* engulf and internalize particles larger than 200 nm with membrane protrusions called pseudopods and, subsequently, digest and/or kill the contents of the *de novo* freshly formed phagosome. In general, particles are recognized by surface receptors, triggering actin (de-)polymerization to move the membrane around the particle. The novel phagosomal compartment formed by this mechanism is transformed into a highly degradative environment through several well-defined fusion and fission steps, summarized as phagosomal maturation [109]. Phagocytosis in the amoeba evolved mainly as a feeding and defense mechanism, a feature that is conserved throughout evolution [110]. Therefore, many of the proteins and principles are highly conserved [100, 101].

Since *Dictyostelium* is a rather simple phagocyte, fewer and less specific phagocytic receptors of *D. discoideum* have been described (reviewed in Cosson and Soldati [101]). Mammalian phagocytic cells present on the cell surface several types of receptors able to initiate phagocytosis, including toll-like receptors (TLRs), integrins, scavenger receptors, and lectins [111]. Homologs of these receptors or at least proteins with functional domains associated with these specific receptors have been found in the amoeba [112]. *D. discoideum* also possesses five integrin-β-like Sib (SibA-E) family proteins partly involved in substrate/particle adhesion [113] and three scavenger proteins (LmpA-C) [114-116]. Of the latter, only LmpB is present on the cell surface [117], while LmpA and LmpC gradually accumulate on (post-)lysosomes [117, 118]. The exact molecular functions of the scavenger receptors in the amoeba, nonetheless, is still unclear. The lectin receptors in *D. discoideum* have only been passively confirmed by association of amoebal cells to carbohydrates immobilized to a polyacrylamide matrix [119, 120], however, the molecular characterization of these receptors involved is still pending.

Upon binding of a substrate to plasma membrane receptors, F-actin is re-arranged driven by the Arp2/3 complex, which is regulated by the SCAR/WAVE (WASH) complex [121-123]. The F-actin re-arrangement and its regulation by Rho GTPases are excellently and comprehensively reviewed in Bozzaro *et al.* [124] and Rivero and Xiong [125]. Nonetheless, phagosomal uptake and maturation is not only regulated by proteins but also by second messengers. For example, phosphatidylinositol phosphates (PIPs) regulate the phagosomal uptake and maturation [126,

127], e.g., the phosphatase Dd5P4 acts on PI(3,4,5)P_3 dephosphorylating it to PI(3,4)P_2, facilitating phagocytic cup closure [128].

Maturation of the phagosome is a tightly orchestrated sequence of events and begins during the closure of the phagocytic cup. The Rab GTPase Rab7 is recruited to the early phagosome as early as 1 min after the uptake and regulates the delivery of lysosomal enzymes into the phagosome [129, 130]. It usually acts in cooperation with Rab5 as one of the main regulators of phagosomal maturation [131, 132]. Another protein recruited upon closure of the phagocytic cup to the nascent phagosome is the H^+-vacuolar ATPase (V-ATPase) [133, 134]. Responsible for the acidification of the compartment [135], the V-ATPase decreases the pH inside the phagosome to 3.5-4 [136] after 10 to 30 min [137, 138]. This acidification is a prerequisite for the delivery and activity of lysosomal enzymes [139] and the killing and digestion of bacteria [109]. The V-ATPase and the lysosomal enzymes are retrieved by the WASH complex in waves of recycling [123, 140]. This results in a shift of the phagosomal pH back to neutral (pH ~7), marking the beginning of the maturation into a postlysosome [109, 141]. The postlysosomal compartment is characterized by an actin coat and the presence of various proteins like vacuolins (vacA and vacB), flotillins, and coronin [142-144]. The mature postlysosome fuses with the plasma membrane and ejects its undigested contents [143, 145]. Exocytosis is also regulated via local calcium concentrations, with mucolipin, a Ca^{2+} transporter, pumping Ca^{2+} ions into the compartment and inhibiting exocytosis [146]. While in *D. discoideum* exocytosis is a constitutive process, in mammalian immune cells the regulation by local Ca^{2+} concentrations is only observed secretory lysosomes [147].

Autophagy in *D. discoideum*

After the first description of autophagy, research of the process in the amoeba gained significant traction in the recent decades. In general, autophagy is a process of self-digestion for nutrient acquisition. The first description of autophagy in *Saccharomyces cerevisiae* [148, 149] and the growing interest in this process, led to the discovery of all relevant components in the amoeba [150-152]. First studies in *D. discoideum* focused on autophagy maintaining cell viability during nutrient deprivation, i.e., development of *Dictyostelium*. Cell viability is maintained by the gradual digestion and recycling of intracellular materials as a source of energy and essential metabolites for development [153]. There are, however, also nonrecycling functions of autophagy in aggregation and differentiation of the amoeba. For example, SDF-2, an essential factor for spore formation, and its precursor AcbA are secreted by an autophagy-dependent mechanism [154]. Until now, many forms of autophagy have been described and the lines

separating it from other processes like phagocytosis are blurring [155]. The following paragraphs will describe a general overview of the process.

The double-membraned, degradative vesicle formed around a target is a hallmark of autophagy [152]. This formation of the autophagosome can be divided into three stages, induction, elongation, and completion, that are facilitated by protein complexes consisting of autophagy-related (Atg) proteins [152]. The inductive phase is initiated by the Atg1 complex containing Atg1, Atg13, Atg101, and the scaffold protein FIP200 [156-158] by phosphorylation of Atg6 in the phosphoinositide-3-kinase (PtdIns3K) complex, triggering the production of phospholipid PtdIns3P [159]. This induction occurs in the so-called omegasome, a newly generated subdomain of the ER [160]. The PtdIns3P levels at this site are tightly regulated by factors like Vmp1 [161, 162], which is essential for the acquisition of Atg proteins involved in membrane trafficking during autophagosome elongation like Atg18, Atg2, and Atg9 [151, 163, 164]. In the following steps, the omegasome is expanded to form the isolation membrane (IM or phagophore), which requires the concerted action of eight conserved core Atg proteins [165]. This expansion is facilitated by a series of ubiquitin-like conjugations leading to the association of Atg8a and Atg8b to the lipid phosphatidylethanolamine (PE) in the inner and outer membranes of the IM, respectively [166, 167]. This process is excellently reviewed for *D. discoideum* in Fischer and Eichinger [158] and also displayed in Figure 3, which is taken from the same review. Following the conjugations of the inner and outer membranes, autophagosome biogenesis is completed with the closure of the double membrane [168, 169]. At some point, the outer membrane of the mature autophagosome fuses with a lysosome, with the lysosomal enzymes proceeding to degrade the inner autophagosomal membrane and its contents [170]. In *Dictyostelium*, the molecular actors involved in that process are not conclusively identified. However, the completed autophagosome is transported along the microtubules to the lysosome [171]. The subsequent fusion between autophagosome and lysosome in *D. discoideum* is dependent on soluble NSF attachment receptors (SNAREs), Rab GTPases, and membrane tethering factors [172].

In addition to its functions in development and nutrient recycling as outlined before, autophagy also constitutes a major first line of defense against pathogenic bacteria [109]. Upon the detection of bacterial invader trying to manipulate or escape the phagosome, *Dictyostelium* triggers, like its mammalian counterparts, a stringent pathway called xenophagy, leading to the engulfment and degradation of the damaged phagosomes and its contents [109, 173]. For this purpose, the bacterium and the phagosomal remnants are ubiquitinated, acting as a degradation signal [174, 175]. The association of the receptors to the ubiquitin moieties triggers the

recruitment of the nascent phagophore via the interaction of an LC3-interacting region on the receptor protein with the main autophagosomal marker Atg8 (called LC3 in higher eukaryotes) [174]. In *D. discoideum*, p62/SQSTM1 is one of such autophagy receptors [151], and it recognizes cytosolic *F. noatunensis* [26] and *M. marinum* [176, 177]. In summary, autophagy represents a major line of defense for the amoeba in addition to its function in nutrient recycling.

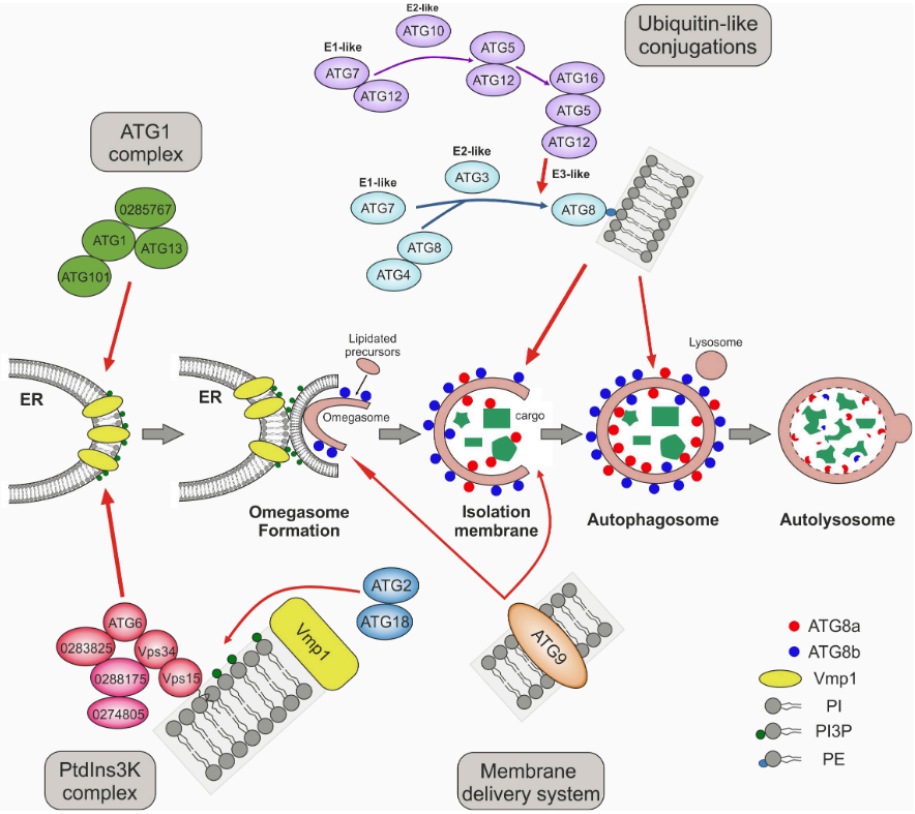


Figure 3. The molecular events occurring during autophagy. Induction of autophagosome formation is facilitated by the Atg1 and PtdIns3K complexes at the omegasome, a subdomain of the ER. Atg2, Atg9, and Atg18 are involved in the membrane trafficking promoting autophagosome elongation. A series of ubiquitin-like conjugations leads to the association of Atg8 to the phospholipid PE in the inner and outer membranes of the forming IM. After association of Atg8, autophagosome formation is completed upon closure of the double membrane. The figure was taken from the review by Fischer and Eichinger [158].

RNA interference (RNAi) as a defense mechanism

Downstream of these processes, RNAi constitutes another layer of defense against pathogenic bacteria or viruses in various organisms stretching from plants to invertebrates [178-180]. RNAi was originally observed by Fire *et al.* [181] when they injected complementary RNA into the gut of *C. elegans*, which resulted in efficient and inheritable gene silencing. For this discovery Andrew Mello and Craig Fire were awarded the Nobel prize in 2006. Shortly after this

discovery, small interfering RNAs (siRNAs) generated by the action of the RNase III protein Drosha were described to be effector molecule in RNAi [182, 183]. These siRNAs have a characteristic size of 21 or 22 nt with a 2 nt 3′-overhang on both strands and a 5′-monophosphate [184]. These characteristics and the process of the generation of siRNAs from a double-stranded RNA (dsRNA) by RNase III proteins called dicers is conserved throughout members of all eukaryotic supergroups [185]. Research in the last two decades uncovered more about the proteins and complexes involved, especially downstream of siRNA generation. The generated siRNAs form the RNA-induced silencing complex (RISC) by integrating one siRNA strand into an argonaute (Ago) family protein [186]. Selection of the integrated siRNA strand is purely dependent on the thermodynamic stability of the 3′-ends of the duplex [187, 188]. The siRNA-Ago complex targets complementary mRNA sequences while the slicer activity of Ago cleaves the bound substrate [189, 190]. The characterization of siRNA generation and RISC, however, did not explain the inheritability of the gene silencing occurring upon dsRNA formation. Fire *et al.* (1998) already speculated that this observation suggests the presence of some kind of amplification mechanism in the RNAi response. Indeed, the RNA-dependent RNA polymerase (RdRP) EGO-1 in *C. elegans* was found to be required for a robust RNAi response [191]. Since this discovery, RdRPs were identified in many other eukaryotes and the molecular mechanism of the RNAi response amplification was characterized [185]. In brief, a complementary strand is synthesized by an RdRP in a primer-dependent or primer-independent manner, which is subsequently processed into secondary siRNAs by dicer proteins [192-194]. Another mechanism has been described for *C. elegans*. In the guiding mode mechanism primarily 22-nt long RNA molecules, featuring a 5′-end trimethylated guanosine, are *de novo* generated by the RdRP RRF-1 [195-197].

Research in recent years revealed multiple functions of RNAi, i.e., constituting a layer of protection against viral infections in plants and invertebrates (as reviewed in Ding and Voinnet [178]). For example, dicer-like protein Dcr2 of *D. melanogaster* was involved in conferring immunity against four different insect viruses with a (+) single-stranded RNA (ssRNA) genome [103-105]. Similarly, RNAi plays a role in the antiviral defense mechanisms of *C. elegans* [106, 107]. In addition to these defense responses against RNA viruses, RNAi was also described to be involved in the repression of dsDNA virus replication, as demonstrated for the Cauliflower mosaic virus and the Cabbage leaf curl virus [198]. In case of the former, the 35S leader of the viral polycistronic transcripts forms an extensive secondary structure with long double-stranded stretches [199]. These double-stranded regions are targeted by dicer-like proteins in

Arabidopsis and processed into viral siRNAs, which then proceed to inhibit viral replication [178, 198]. Most of these RNAi-mediated antiviral defense mechanisms have been described for plants and higher eukaryotes and similar processes have not been described for basal eukaryotes, yet. However, the RNAi machinery seems to be a common feature of all eukaryotic supergroups including the Amoebozoa [108, 185]. Especially *D. discoideum* features a highly potent RNAi response, able to suppress the retrotransposition of transposable elements like DIRS-1, Skipper-1, and TRE5-A [200]. The first indications for a role of the amoebal RNAi machinery in pathogen defense come from transcriptome analyses of *Dictyostelium* cells infected with *M. marinum* and *L. pneumophila* [201]. Besides the classical responses to pathogens like the upregulation of phagosomal factors and autophagy, an induction of RNAi factors was observed independent of the identity of the bacterial invader [201]. So far, no viruses have been described to infect *D. discoideum*.

Introduction to (r)RNA modifications in *D. discoideum*

Parts of this chapter are included in the manuscript entitled "Ribosome heterogeneity in Amoebozoa: fractional 2′-O-Methylation in the ribosomal RNA of *Dictyostelium discoideum*", which is published in Scientific reports. All sections concerned with 2′-O-methylation in the amoeba are adapted from this manuscript.

The diverse world of RNA modifications

Chemical modifications of RNA

Chemical modifications of RNAs have a long-standing history. The first RNA modification to be discovered was the pseudouridine (Ψ) in 1957, only a few years after the first DNA modification, 5-methylcytosine, was described in the literature [202, 203]. Since then, the number of known modification of RNA bases has exploded to ~170 known RNA nucleosides [204]. The early discoveries of these modifications relied on chromatographic separation and characterization by organic chemistry [205]. RNA modifications can be found in many non-coding RNAs (ncRNAs) like tRNAs, rRNAs, small nuclear RNAs (snRNAs), and small nucleolar RNAs (snoRNAs) [204], but also in mRNAs [206]. A high number of RNA modifications affect molecular processes like transcription, splicing, mRNA stability, translation, RNA export, and RNA degradation [207-215], underlining the relevance of chemical modification of nucleotides.

As the most modified RNA molecule and the first small RNA to be discovered, tRNAs were the richest source for information on the high variety of chemical modifications (Figure 4A and B). Up to ~25% of their nucleotides are modified [204, 205, 216], ranging from simple methylations to highly complex, multistep attachment of chemical groups (Figure 4C; [217]), which contribute to folding, stability, and function of the RNA [218]. Nucleoside methylations in tRNAs play an important role in the maturation of these molecules. For example, N1-methyladenosine (m^1A) at position 9 of mitochondrial $tRNA^{Lys}$ stabilizes the canonical clover-leaf structure during tRNA biogenesis by hindering intra-stem base pairing, which otherwise forms a dysfunctional extended hairpin structure [219-221]. Another example is the family of methylguanosines (m^2G, m^2_2G, and m^2_2Gm) that are highly conserved at tRNA positions 10 and 26, where they disrupt early base pairing in tRNA maturation to allow for proper folding [222]. Additionally, the methylguanosines at tRNA positions 2, 3, and 10 are important for proper recognition during aminoacylation, as determined by substitution of these positions with inosine [223]. Other modifications like Ψ, dihydrouridine (D), and 5-methylcytosine (m^5C) are

more involved in the stabilization of the tRNA secondary structure [224-226]. Beyond the addition of single chemical groups, the position 34 and 37 in the anticodon loop are the most heavily modified nucleotides in tRNAs [227]. The position U34, the wobble position, is modified in the elongator-dependent pathway up to 5-methoxycarbonymethyl-2-thiouridine (mcm^5s^2U; Figure 4C) in a small subset of tRNAs and is important for the recognition of the cognate codon [228]. Position 37, located on the 3′ side of the anticodon, is usually hypermodified when an A or U is present at position 36. Modifications at this site include highly sophisticated structures like hydroxywybutosine (OHyW), N6-threonylcarbamoyladenosine (t^6A), etc. [204, 227, 229]. This diverse set of RNA modifications at position 37 serves to stabilize the A/U and U/A base pairs at position 36 in the codon-anticodon interaction and to prevent intraloop base pairing, thereby contributing to translational fidelity [230, 231].

With the rise of NGS techniques, RNA modifications were also found in less abundant RNA molecules and, especially, research of chemical modifications of mRNAs gained traction in the last decade [211, 232]. The most abundant modification in mRNAs, N6-methyladenosine (m^6A), is co-transcriptionally and reversibly introduced on the N6 of the adenosine base by the methyltransferase complex with methyltransferase-like 3 (METTL3) and METTL14 at its core [233-235]. m^6A affects many developmental processes and diseases including the maternal-to-zygotic transition in *Danio rerio* [236], the cortical neurogenesis [237], and acute myeloid leukemia [238]. Molecularly, this is achieved by the interaction of so-called reader proteins with m^6A, recruiting machineries such as translation, RNA decay, or localization [222].

The regulation, complexity, and diversity of RNA modifications found throughout all evolutionary supergroups is extensive [204, 206, 236]. Borland *et al.* [239] showed by mass spectrometry that *D. discoideum*, the model organism used in this work, also features a highly diverse set of chemical modifications. While many thought in the beginning that these modifications represent a steady state, scientific progress in the last decades revealed a dynamic landscape of RNA modifications. Especially tRNAs, mRNAs, and rRNAs have been found to be heterogenous in their set of chemical modifications in dependence of environmental and developmental processes [227, 233, 240]. The following chapters will mainly focus on the modifications introduced into rRNA.

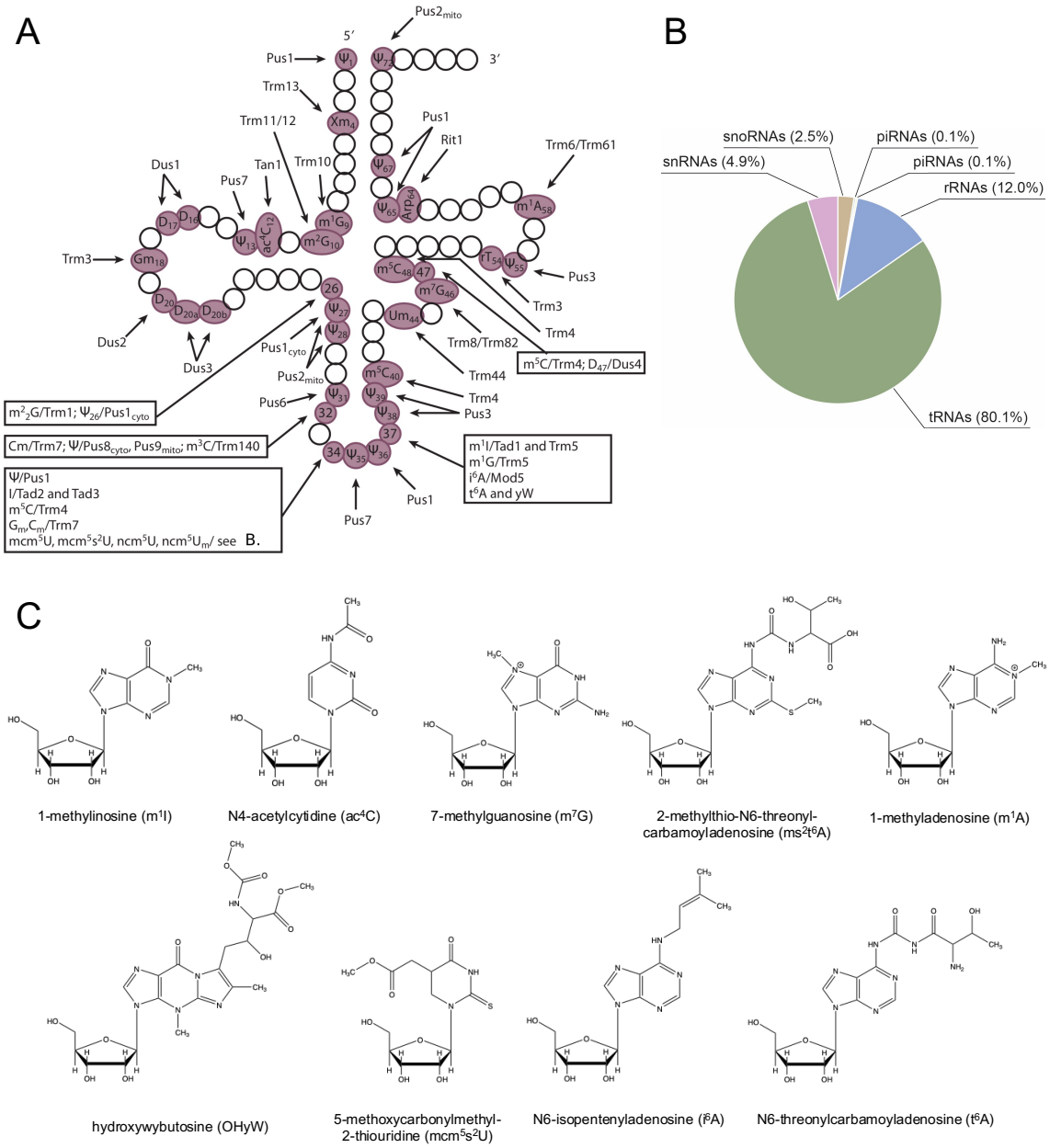


Figure 4. Transfer RNAs as treasure trove in the detection of chemical modifications of RNAs. (A) Schematic overview of RNA modifications in the secondary structure of a yeast tRNA. Modified sites are indicated in red, and the type of the RNA modification and the responsible enzymes are given. The figure is adapted from El Yacoubi *et al.* [229]. (B) Pie chart of non-coding RNA (ncRNA) species with known chemical modifications in the Modomics database [204]. (C) A small selection of the diverse RNA modifications found in eukaryotic tRNAs ranging from simple base methylations (e.g., m^7G or m^1A) to chemically complex modifications (e.g., mcm^5s^2U, i^6A, or t^6A). The structures were drawn with ChemDraw v. 19.1 based on information from the Modomics database [204].

Chemical modifications of ribosomal RNA

Early on, the peptidyltransferase reaction of the ribosome was shown to be resistant to protein degradative treatment [241]. This was the first indication that a non-proteinaceous factor is the catalytic entity in protein biosynthesis. Indeed, that rRNA is this catalytic entity, rather than proteins, was subsequently confirmed by ground-breaking and highly decorated crystallographic work [242-244]. This is also reflected in their biogenesis, as maturation of ribosomes is amongst the most complex cellular processes and requires about 200 facilitating proteins, as reviewed recently [245]. This process can be broken down into six crucial steps:

1. Synthesis of rRNA, snoRNAs, ribosomal proteins, and assembly factors
2. Cleavage and processing of the ribosomal primary transcript
3. Chemical modification of the nascent rRNA
4. Assembly of the ribonucleoprotein complex in the nucleus
5. Nuclear export of the complex
6. Quality control of synthesized ribosomes

Amongst many other processes, the introduction of post-transcriptional, covalent modifications (step 3) in rRNA is of utmost importance for ribosome biogenesis and function, as summarized by Sloan *et al.* [240]. Ribosomal RNA contains a high number of chemical modifications including acetylations, base methylations, or 3-amino-3-carboxypropylation (acp) [240]. The most prominent nucleotide modifications in rRNA are 2′-O-ribose methylation (2′-O-Me; Figure 5A) and Ψ (Figure 5B) that are introduced site-specifically. These modifications are thought to be important for RNA folding, ribosome stability and translational fidelity [246-248]. In recent years, a specialization of ribosomes in response to environmental changes and/or developmental processes has been suggested, with substoichiometric chemical modifications being implicated as a major source of ribosome heterogeneity [240, 249]. As such, examples for fractional rRNA modifications are found in various species, including *S. cerevisiae*, where 18 positions are modified in less than 85% of the ribosomal population [250], and also approximately a third of the 2′-O-Me positions in rRNA of *Homo sapiens* are found hypomodified [251]. Recently, altered 2′-O-Me levels were also discovered during the development of *D. rerio* [252]. Functionally, ribosome heterogeneity has been proposed to constitute a fine-tuning mechanism for translational activity of an unknown subset of mRNAs [253, 254].

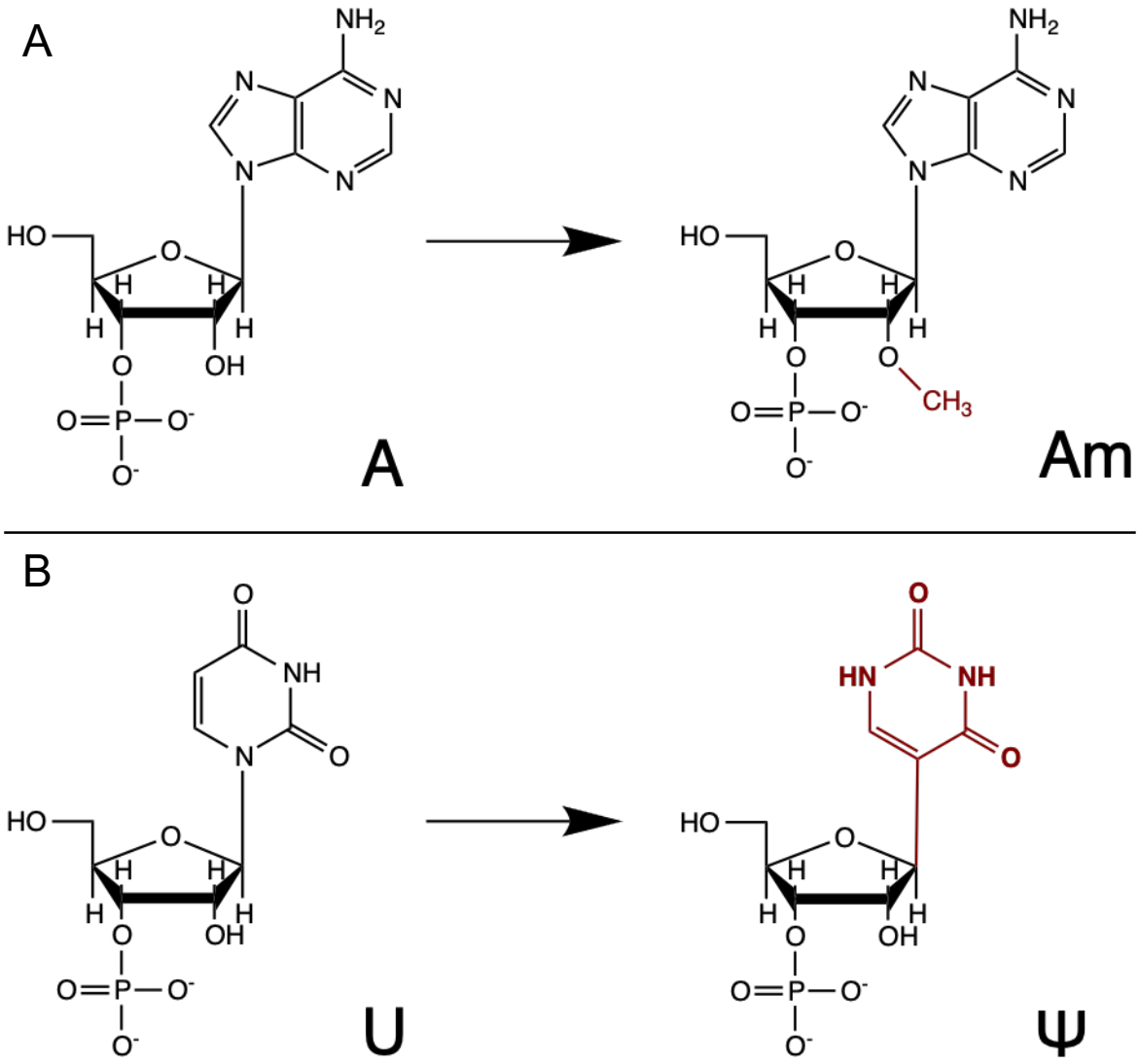


Figure 5. Chemical modification of ribonucleotides. Modification reactions were drawn using ChemDraw v. 19.1. Modified groups are marked in red. (A) Methylation of the 2′-hydroxyl group of the ribose moiety. (B) Isomerization of the uridine nucleotide by attachment of the uracil base via a C1-C5 glycosidic bond to the ribose.

2′-O-Me can be found on a high number of non-coding [216, 255, 256] and coding [257] RNA molecules. This modification is introduced either by stand-alone methyltransferases [258, 259] or by snoRNA-guided protein complexes containing the methyltransferase fibrillarin [260]. Independent of the mode of action, the 2′-hydroxyl in the ribose moiety of all four nucleotides can be modified (Figure 5A). The importance of ribosomal 2′-O-Me for translational fidelity is underlined by studies depleting cellular fibrillarin and, in consequence, the loss of translational activity [261]. Even though the complete loss of ribosomal 2′-O-Me has strong phenotypic effects, the functions of the vast majority of individual 2′-O-methylations are not clear, yet. Rather, they seem to contribute to the rRNA structure, as 2′-O-Me-modified nucleotides that are paired or intra-helically stacked exhibit altered conformational preferences, favoring the

C3′-*endo* over the C2′-*endo* sugar pucker confirmation (Figure 6; [262]). The stabilization of the C3′-*endo* form is caused by the steric repulsion among the 3′-phosphate, the 2′-O-methyl, and the 2-carbonyl groups in the C2′-*endo* conformation [263]. The transition to the C3′-endo form in ssRNA before folding, results in a stabilization of the helix upon duplex formation [264, 265]. The structural bias of 2′-O-methylated nucleotides is important for biologically relevant functions. For example, Polikanov *et al.* [247] described that the functionally relevant conformation of the G2553 base is maintained by the intercalation of the 2′-O-methyl group on U2552 in the A-loop of the *Thermus thermophilus'* 23S rRNA. G2553 is in direct contact with and helps accommodating the A site tRNA [247]. Similarly, the 2′-O-methylated G2251 contacts the C2065 ribose and U2449 base via hydrophobic interactions to stabilize the base pairing with the 3′-end of the P-site tRNA [247]. Both 2′-O-methylated positions are highly conserved in ribosomes in all kingdoms of life. In addition to these selected positions in rRNA, specific 2′-O-Me modified nucleotides have been implicated in spliceosomal assembly and disassembly [266, 267], in aiding exon ligation [268], and disrupting tRNA selection and proofreading [269]. A comprehensive overview on the (structural) impact of a whole 2′-O-Me landscape on RNA structures is largely missing and only beginning to emerge. It was shown recently, nonetheless, that 2′-O-Me stabilizes alternative secondary structures of the HIV-1 transactivation response element *in vitro*, which are unfavorable and exhibit shorter lifetimes in nuclear magnetic resonance (NMR) for unmodified RNA molecules [209].

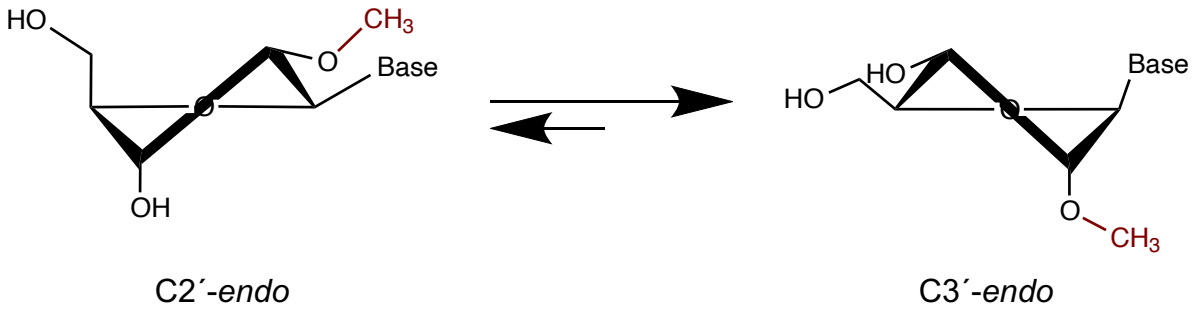


Figure 6. Bias of 2′-O-methylated towards C3′-*endo* sugar pucker conformation. The 'Base' can be any of the four RNA bases and the methylation of the 2′-hydroxyl is marked in red. The figure is drawn with ChemDraw v. 19.1 based on Abou Assi *et al.* [209] and information from Kawai *et al.* [263].

Upon its discovery in 1957, Ψ was originally mischaracterized as a 5th core nucleotide, due to its high abundance in cellular RNA molecules [202]. It was renamed, however, after thorough experimental characterization [270] indicated Ψ (5-ribosyluracil) was an isomer of uridine (1-ribosyluracil). The isomerization starts with the break of the N1-C1′-bond between the uracil

base and the ribose [271]. The new C5-C1′ bond is formed after a 180° rotation of the base along the N3-C6 axis (Figure 5B), yielding an additional hydrogen bond donor at the Hoogsteen edge [271]. The isomerization of uridine to Ψ is generally catalyzed by a highly conserved set of enzymes: the pseudouridine synthases. In bacteria, Ψ is introduced site-specifically by stand-alone proteins [272], while in eukaryotes and archaea only a fraction of modifications are occurring in this manner (reviewed in Rintala-Dempsey and Kothe [273]). Most pseudouridylations in the rRNA of eukaryotes are made by the snoRNA-guided box H/ACA small nucleolar ribonucleoprotein (snoRNP) complex containing dyskerin as the functional pseudouridine synthase [274]. Ψ is present on a wide range of cellular RNA molecules like snRNAs (reviewed in Guthrie and Patterson [275]), tRNA [216, 230], and rRNA (reviewed in Ofengand and Fournier [276]). Especially, rRNA pseudouridylation is essential for tRNA binding and translational fidelity [277]. This is compounded by the clustering of Ψ in functionally important sites like the peptidyl transferase center (PTC) [278], the decoding center [279, 280], the intersubunit bridges [248], and LSU's Helix 69 [281-284]. The exact role of ribosomal Ψ in the processes occurring in these functional regions, however, is still unclear. Ψ in snRNAs is equally important for snRNP function and biogenesis, with all tested Ψ in the snRNAs of Xenopus oocytes being essential for splicing [285, 286]. One example is the well-characterized function of Ψ35-containing U2 snRNA's involvement in mRNA splicing in yeast [287]. Crystal structures of the pre-mRNA/U2 duplex imply that the bulged branch-point nucleotide A is stabilized by Ψ35 in the first step of splicing [288]. Thereby, Ψ35 helps to expose the 2'-hydroxyl for a nucleophilic attack [288].

Since useless features are usually lost during evolution, the high abundance of Ψ supports a likely positive influence on the structural dynamics, stability, and conformation of modified cellular RNAs. Indeed, the Ψ-A base pairs formed exhibit similar geometry than U-A base pairs, but display a higher thermodynamic stability, thereby increasing RNA duplex formation [289, 290]. Ψ can form, in addition to A, stable base pairs with G and C, making it a flexible base pairing partner [291]. One stabilizing effect of Ψ first observed in crystal structures of tRNA(Gln) is the coordination of water molecules between its nucleobase (Ψn) and surrounding sugar-phosphate backbone (O1P n and O1P n-1), enhancing the rigidity of the RNA structure [292]. Other effects on RNA structure are, like 2′-O-Me, caused by the preference of the ribose for the C3′-*endo* sugar pucker conformation and the resulting augmented base stacking [289, 291]. In a biological context, it was described that the C3′-*endo* form may antagonize the binding of RNA molecules to their corresponding RNA-binding proteins (RBP), e.g., RNA binding to PUM2 [293] or U2AF2 [294], and *Xenopus*' snoRNP assembly [295]. Given these

experimental observations, RBPs may exhibit a C2′-*endo* conformation bias [296]. The mechanisms how RBP-binding is affected by Ψ either directly or indirectly by structural modulation of the RNA, however, remains to be elucidated.

Despite the well-characterized stabilizing nature of Ψ, the biological function of the modification is highly dependent on sequence context. A recent study by Hudson *et al.* [297] systematically characterized differences in the thermodynamic stability of RNA duplexes with and without Ψ in dependence of varying sequences context. Not surprisingly, duplexes containing NΨ/ΨN at their termini were found to be more stable than NU/UN-containing ones, while the stability of RNA duplex with internal NΨ/ΨN neighbors was further increased [297]. This context dependency is also observed in the base stacking of single-stranded RNA molecules [290]. Beyond the local sequence context, effects of Ψ depend on its position in the secondary structure of the RNA [289, 298]. The structural impact of Ψ in longer, cellular RNAs is uncharted at present.

RNAs guide RNA modifications: the small nucleolar RNAs (snoRNAs)

Box C/D and Box H/ACA snoRNAs

Ribose methylations (Figure 5A) and pseudouridylations (Figure 5B) are site-specifically introduced in rRNA by snoRNPs, as summarized recently [240, 249, 299]. They come in two flavors: H/ACA snoRNPs catalyze the conversion of uridine to Ψ, while box C/D snoRNPs introduce methyl groups at the 2′-hydroxyl of ribose residues [260, 300]. For each class of snoRNPs, a conserved and distinct set of four proteins form the catalytic complex, of which dyskerin in the H/ACA snoRNPs isomerizes uridine [301], while fibrillarin in box C/D snoRNPs acts as the methyltransferase on 2′-hydroxyl groups [302]. The rRNA target positions are defined by the individual snoRNA components of the snoRNPs. For both classes, specific base pairing patterns select the nucleotide to be modified.

Box C/D snoRNAs possess conserved box C (5′-RUGAUGA-3′) and box D (5′-CUGA-3′) motifs that are essential for their structure, function, and biogenesis (Figure 7) as well as less conserved box C′ and box D′ motifs [303-305]. Nucleotides of the box C and box D motifs interact with each other, forming a kink-turn, and similar, but weaker interactions may also occur between nucleotides of the box C′ and box D′ motifs. The kink-turn was first described in rRNA and typically consists of two stems separated by a short loop [306]. While the first

stem (Stem-I) is composed of standard Watson-Crick base pairs, the second stem begins with two *trans* Hoogsteen/sugar-edge A•G base pairs, resulting in the unpaired nucleotides in the loop inducing a ~120° kink in the helical axis [306]. Mutations in the nucleotides forming the kink-turn inhibits box C/D snoRNA localization to the nucleolus [307], suggesting that intramolecular base pairing of box C and box D are essential for snoRNA processing and snoRNP structure. Indeed, the k-turn structure is first recognized by 15.5K proteins (Snu13 in yeast and L7Ae in archaea) [307, 308]. Even though L7Ae is able to recognize the less conserved box C′/D′ motifs [309], the eukaryotic 15.5K protein only binds the highly conserved box C/D motifs [310]. There are indications, however, that 15.5K is recruited to the C′/D′ motifs in eukaryotes through protein-protein interactions with Nop56 and Nop58 [311]. Crytallographic observations imply that the 15.5K protein stabilizes an RNA structure favoring the recruitment of the remaining box C/D snoRNP core proteins [312, 313], thereby, constituting an initiation complex for snoRNP biogenesis. As a next step in eukaryotes, Nop56 binds to fibrillarin before assembly on the initiation complex occurs [314]. The archaeal Nop5 exhibits the same mechanism [315], however, the second eukaryotic Nop5-homolog Nop56 only associates with the pre-snoRNP after fibrillarin is present [314]. However, the interactions of Nop56 that are necessary for the recruitment are not yet described. In the fully assembled snoRNP, Nop56 and Nop58 bind to the box C and box C′, respectively, while one copy of fibrillarin each binds to the box D and the box D′ [316]. Immediately upstream of the D and/or D′ box are antisense elements that form 7 to 21 bp duplexes with the RNA target and direct the active site of fibrillarin to the 2′-hydroxyl of the nucleotide base paired to the 5th nucleotide upstream of the D or D′ [260, 302, 317]. Methylation of the target nucleotide occurs in an S-adenosylmethionine (SAM)-dependent reaction [302]. It has been shown that additional base pairing independent of the box C/D snoRNA antisense elements between the snoRNA and rRNA enhances methylation [318]. Since the D and D′ box can bind a copy of fibrillarin and both possess upstream antisense elements, box C/D snoRNAs can in principle guide two distinct 2′-O-methylations in one or multiple RNA molecules [260]. A reorientation of box C/D snoRNAs with non-canonical box C′/D′ motifs inside the snoRNP complex might lead to alternative base pairing, increasing the target number of a given snoRNA even further [319]. Structural analysis of archaeal box C/D sRNPs revealed that the substrate-binding channel of the complex accommodates only 10 base pairs of the snoRNA/rRNA duplex [320]. While shorter and stable duplexes are allowed, longer duplexes need to be unwound and therefore increase the residence time of the duplex within the complex [320]. It can be speculated that

this might constitute a control mechanism to ensure that important sites in the rRNA are modified, e.g., those interactions with long complementary stretches [320].

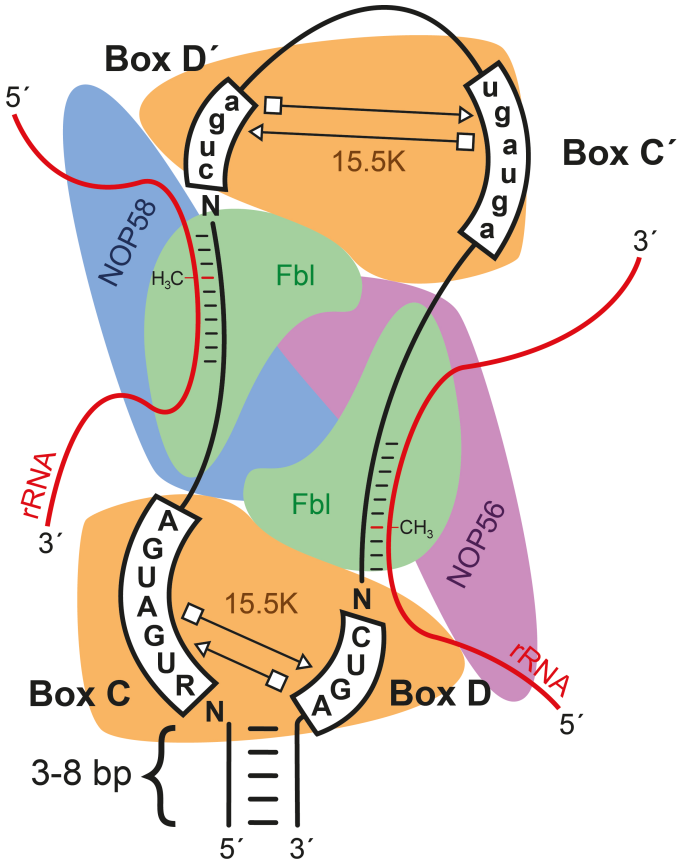


Figure 7. Features of box C/D snoRNAs. Conserved residues of boxes C and D are shown. They interact to form a functionally important k-turn by means of *trans* Hoogsteen/sugar-edge A•G base pairs, shown in conventional Leontis-Westhof symbols [321]. The core snoRNP proteins are shown schematically. The guiding sequences is upstream of the D box with methylation occurring in rRNA at the position pairing to the 5th nucleotide upstream of the D box (indicated with a methyl group). Base pairing with rRNA (red) is schematically shown. Boxes C´ and D´ are usually less well conserved (indicated by small lettering). The separate antisense sequence upstream of box D´, allows guidance to a further methylation site.

Box H/ACA snoRNAs are characterized by their eponymous box H (ANANNA) and box ACA, essential for nuclear localization and recruitment of the snoRNP complex proteins [300, 322-324]. Typically, these snoRNAs (Figure 8) consist of double hairpin loops separated by a hinge region with each of them possessing an internal pseudouridylation pocket where the target U is isomerized to Ψ [300]. The RNA target sites are identified by site-specific base pairing with the box H/ACA snoRNA [300]. The RNA component does not act alone in pseudouridylation, rather, they form a complex with four common core proteins: Cbf5 (dyskerin in humans), Nhp2,

Gar1, and Nop10 (Figure 8). These H/ACA snoRNP core proteins were found to be essential for the Ψ formation reaction and stability of the complex [325-328]. The molecular mechanism of pseudouridylation was a point of discussion for the last two decades with three main hypotheses being put forward: the Michael addition scheme, the acylal scheme, and the glycal scheme [329-332]. The first involves two nucleophilic attacks linking the uracil base to the Ψ synthase, followed by repositioning of the uracile C1′ close to the ribose C5 and formation of the C-C bond [329-331]. The acylal scheme mechanism would generate an acylal intermediate after uracil liberation upon the nucleophilic attack of the active site aspartate of the Ψ synthase with subsequent rotation of the base and re-attachment of C1′ to the ribose C5 [333]. However, recent studies rather support the glycal scheme of Ψ formation [332, 334]. In this model, the C2′ carbon deprotonation and detachment of the uracil base forms a glycal complex with double bonds between the C1′ and C2′ of the ribose moiety [334]. *In vitro* reconstitution and subsequent assays identified three pivotal sequence and structural features: an obligatory distance of 14-16 nt between the target U and either box H or ACA, the thermodynamic stability of the snoRNA/target RNA duplex, and the hairpin stability forming the pseudouridylation pocket [335]. The substrate RNA is not threaded through the pseudouridylation pocket during base pairing, forming a one-sided Ω-shaped interaction motif as observed by NMR studies [336]. Such an interaction would allow for a quick and consecutive target RNA base pairing, modification, and release from the H/ACA snoRNP complex [336]. Extensive crystallographic work throughout the last two decades yielded a detailed insight into H/ACA snoRNP function and structure [337-341].These studies have shown that Cbf5, Nop10, and archaeal L7Ae (ortholog of Nhp2) interact with the upper stem of the box H/ACA hairpin [337, 338]. Gar1, however, only binds to Cbf5 and does not directly interact with the box H/ACA snoRNA or the substrate RNA, likely regulating the release of the modified target RNA [339, 341]. The positioning of Cbf5's active site in proximity of the pseudouridylation pocket is facilitated by additional interactions with box H or ACA and the lower stem of the hairpin [337, 338]. Base paring of the substrate RNA triggers additional interaction between the snoRNP and the guide/substrate pair and ensures precise positioning of the target nucleotide in the catalytic site of Cbf5 [340, 341].

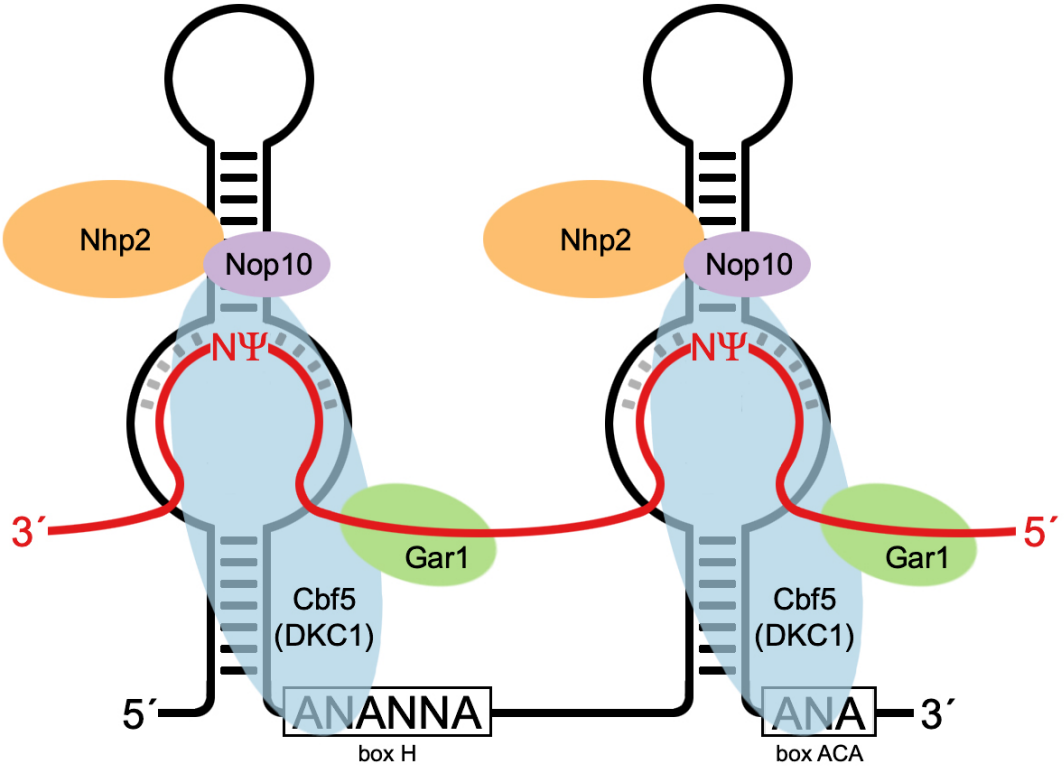


Figure 8. Features of the box H/ACA snoRNP in eukaryotes. The box H/ACA snoRNA (shown in black) is characterized by a double hairpin structure, with each being an individual functional unit. Box H and box ACA are shown with their conserved sequence motif, which is essential for the formation of the snoRNP complex [300]. The four core H/ACA snoRNP proteins, Cbf5 (DKC1 in humans; blue), Gar1 (green), Nop10 (purple), Nhp2 (orange) are shown. The substrate RNA and the nucleotides in the pseudouridylation pockets are drawn in red. The figure was adapted and is based on the information from Ge and Yu [342].

Transcription and processing of box C/D and box H/ACA snoRNAs

Throughout the evolutionary tree of life, snoRNA can mostly be divided into independently encoded and transcribed snoRNA genes and intronically encoded snoRNA genes (reviewed in Dieci *et al.* [343]). The majority of snoRNAs in plants [344] and yeast [345] are encoded in either monocistronic or polycistronic transcriptionally independent units. Polycistronic coding units can consist of homologous or heterologous snoRNA genes [346, 347]. While in *S. cerevisiae* and *S. pombe* only a minor fraction of snoRNAs can be found in introns [345, 348], the number of intronic snoRNA genes is increased in higher eukaryotes [349-351] with the exception of plants [344]. Independent of the mode of transcription or clustering, the same set of endo- and exoribonucleases is involved in the maturation of the pre-snoRNA [352, 353]. Pre-snoRNA transcription and processing will only be covered briefly in the next paragraphs

and for more detailed information we refer to the excellent reviews on transcription by Dieci *et al.* [343] and on snoRNA maturation by Kufel and Grzechnik [354].

Transcription of guide snoRNA from independent promoters is generally facilitated by RNA polymerase II (Pol II) [355-357]. The promoter elements driving transcription of guide snoRNAs, however, are largely unknown. Only snoRNA genes in yeast are described to feature TATA boxes and A/T-rich sequences in their promoters [358]. Also, snR52's promoter features A and B boxes which are typical for tRNA genes and recognized by RNA polymerase III (Pol III) [359], which is observed as well in a broad spectrum of eukaryotic organisms [356, 360, 361]. Transcription of intron-encoded guide snoRNA is coupled to its host gene and therefore its promoter elements [343]. As pre-mRNAs are synthesized by Pol II and the snoRNA residing in an intron is liberated by splicing [343, 362]. Interestingly, abundance of the mature snoRNA is not necessarily coupled to host gene expression. Some transcripts of human snoRNA-containing genes are targeted for degradation by nonsense-mediated decay (NMD) after splicing of the intronic snoRNA(s) [363]. In addition, alternative splicing of host genes containing snoRNA genes that are partially or fully located in one of its exons leads to NMD [363].

Processing of pre-snoRNAs into mature guide snoRNAs generally adheres to the same sequence of events independent of species and genomic organization. In yeast it begins with the 3′-end of the transcribed pre-snoRNA being processed upon Nrd1-Nab3-Sen1 (NNS) termination [364, 365] by the exosome and its exonuclease cofactor Rrp6 [366, 367]. In some of the pre-snoRNAs, the endoribonuclease Rnt1, an RNase III protein, creates the entry site for the exosome and Rrp6 by downstream cleavage of the snoRNA sequence [352]. Intronic snoRNAs are released by the splicing debranching enzyme Dbr1 and subsequent Rnt1 cleavage and are therefore not coupled to NNS termination [368, 369]. Since Pol II transcripts carry a m^7G cap [370], exonucleolytic processing can only occur after Rnt1 binds and cleaves a stem loop structure in the 5′ extension of the pre-snoRNA [369]. The resulting 5′-end can subsequently be processed by the 5′-3′ exonucleases Rat1 and Xrn1 into the mature snoRNA [371]. The mature 5′- and 3′-end of the snoRNA are determined by the snoRNP complex, as it physically blocks the exonucleolytic enzymes from proceeding [366, 371].

snoRNAs involved in rRNA biogenesis

Several different snoRNAs have been implicated in the processing of pre-rRNA in various species [372]. During evolution, a core set of snoRNAs or functional homologs established which fulfil the same steps in processing. This set includes the U3 snoRNA [373], the U14

snoRNA [374], snR10 [375], U17 snoRNA/snR30 [376], and U22 snoRNA [377]. Additionally, the RNase MRP complex has been shown to facilitate the processing of internal transcribed spacer I (ITSI) [378]. Due to preliminary data concerning its role in enhanced rRNA processing in development and its potential impact on ribosomal (Christian Hammann, personal communication), this work is focused on the U3 snoRNP.

The RNA component, the U3 snoRNA, carries a box C/C′ and a box D/D′ (B) motif, placing it in the family of box C/D snoRNAs [379]. In addition to the canonical box motifs, the U3 snoRNA contains three non-canonical highly conserved sequence motifs: box GAC, box A′, and box A [380]. Unlike other box C/D snoRNAs, the U3 snoRNP does not introduce 2′-O-Me but it is an essential part of the small subunit (SSU) processome, a complex involved in processing, assembly, and maturation of SSU [381], involved in the early processing (site A0, A1, and A2 in yeast) of the SSU rRNA [373, 381-383]. For its role rRNA processing, the 5′ domain of the U3 snoRNA (Figure 9) base pairs with four distinct stretches in the 5′-ETS and the SSU rRNA regions of the nascent rRNA transcript [384]. Base pairing of the U3 snoRNA and the rRNA precursor alone, however, is not sufficient for cleavage at A0, A1, or A2 in yeast (the A0 site is absent in *D. discoideum* [15]). The U3 snoRNA is generally transcribed as an independent transcript by Pol II [385, 386], however, Pol III transcription is described in some instances [387, 388]. Like many other Pol II transcripts, the U3 snoRNA carries a trimethylguanosine ($m^{2,2,7}G$) cap [389]. U3 snoRNA promoters in vertebrates and invertebrates typically contain the proximal sequence element (PSE) also driving the transcription of snRNAs [390]. In yeast, the promoter contains a TATA box in addition to some enhancing upstream promoter elements [391].

Even though the U3 snoRNA does not guide methylation, it still needs a complex of all members of a canonical box C/D snoRNP (Figure 9) for its activity [392, 393]. In the special case of the U3 snoRNP, an additional protein is found as part of the complex: Rrp9 in yeast or U3-55K in humans [394]. Rrp9 was first identified in purified U3 snoRNPs from Chinese hamster ovary cells [395] and quickly established as essential for rRNA processing [394]. Only recently, however, the interactions between Rrp9 and the U3 snoRNA and their impact on rRNA processing were characterized in detail by crosslinking and analysis of cDNA [396]. The study found that the arginine-289 is required for efficient cleavage at A1 and A2, with the combination of the mutations of R289A and helix VI of the U3 snoRNA-SSU rRNA being lethal [396]. The association of the U3 snoRNP to the SSU processome was found by co-precipitation of Sof1 [397] and Mpp10 [398]. Even though, the U3 snoRNA and its protein

complex are highly conserved throughout many evolutionary supergroups, a characterization in the Amoebozoa and its potential intricacies is lacking.

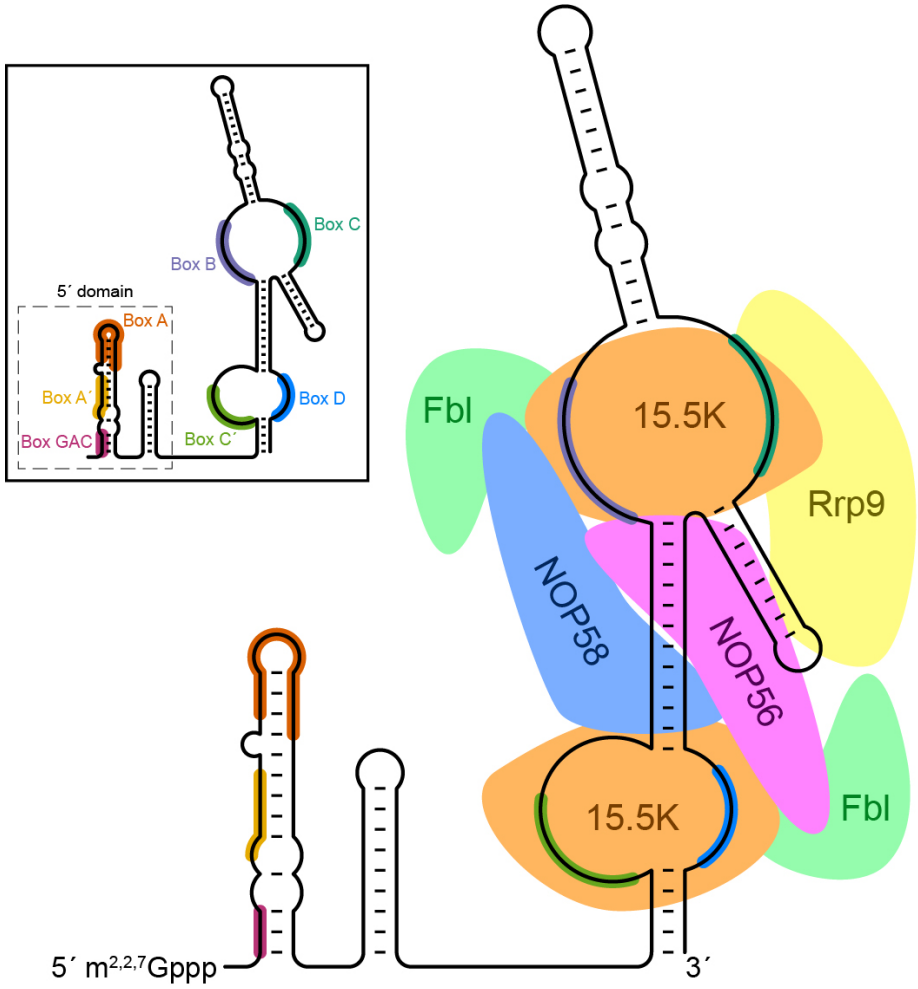


Figure 9. Conserved U3 snoRNA structure. Annotated are the conserved sequence motifs and the 5´ domain (inset of the figure). Binding of the core proteins is shown on the conserved U3 snoRNA structure. The figure was drawn on the basis of Ojha *et al.* [372].

Aims of the thesis

Establishment of monitoring techniques for viral infections of the amoeba

It was the aim of this part of the work to establish protocols and techniques to infect *D. discoideum* productively with giant viruses. For this purpose, a qPCR-based quantitative approach to count the viral genomes of APMV and Tupanvirus (TPV) during the infection was set up. This technique was chosen due to its ability to detect miniscule amounts of viral DNA in a large pool of amoebal DNA molecules. Furthermore, it was the aim to fluorescently label the virus and establish parameters needed for detection in flow cytometry. Setting up these techniques would allow the monitoring of the infection and its initial events in invasion of *D. discoideum*. Depending on the results of initial infections of the amoeba, different aims would be pursued:

Scenario 1: No productive infection of *D. discoideum* occurs

In case of no detectable viral replication, it would be the aim to investigate the cellular fate of the virus by fluorescence microscopy of APMV-infected cells. Furthermore, the possible pathways implicated in this way in the defense against the virus should be manipulated either by knockout of factors involved in the process or external manipulation. In case no effect could be observed, other giant viruses might be used in the infection of *D. discoideum*.

Scenario 2: Viral DNA replication occurs but viral morphogenesis is disrupted

There are many intracellular processes in *D. discoideum* which might contribute to the disruption of the viral morphogenesis of APMV and could potentially be manipulated. Therefore, it would be the aim to investigate the factors leading to these observations. One possible scenario might be the repression of viral gene expression. *Dictyostelium* might be able to detect and degrade viral transcripts using its RNAi machinery, as already observed for bacterial pathogens.

Scenario 3: APMV productively infects *D. discoideum*

In case of a productive infection cycle with detectable replication of the viral genome and generation of viable viral progeny, the aim would be to characterize the molecular mechanisms of the infection or parts of it. The factors involved in this process are not described on a molecular level, due to the natural host *A. polyphaga* not being genetically tractable. Being established as a model organism, infections of *D. discoideum* and the associated viral defense organisms would be more convenient to investigate the infection.

Elucidation of the landscape of chemical modification of rRNA in *D. discoideum*

Even though chemical modifications of rRNA are described across many evolutionary groups from archaea to higher eukaryotes, information on ribomal 2′-O-Me and Ψ in the Amoebozoa and its model organism *D. discoideum* is still missing. Earlier work has identified 17 box C/D and one box H/ACA snoRNA(s) in *D. discoideum*, and verified the function of the former in rRNA 2′-O-methylation [346]. The number of box C/D snoRNAs, however, was not sufficient to cover a potential 2′-O-Me landscape. In addition, ribosome heterogeneity in developmental processes has been described as a common theme in other evolutionary supergroups, nonetheless, it has not been addressed for the Amoebozoa, yet.

Owing to these observations, we set out here to elucidate the global pattern of chemical modifications in *D. discoideum* and its development by employing RiboMeth-seq (RMS) for 2′-O-Me [250] and HydraPsiSeq (HPS) for Ψ [399]. As of the time of this thesis, the rRNA secondary structure of *D. discoideum* and its differences predominantly in the expansion segments towards other species has not been described, yet. To locate the modifications in the rRNAs and create a comprehensive map, their secondary structure should be predicted by homology to other eukaryotic SSU and LSU rRNAs. Upon establishment of the global 2′-O-Me and Ψ, box C/D and box H/ACA snoRNAs of the amoeba should be identified using `snoScan`, a program developed to find potential box C/D snoRNA loci in the given genome [400]. Mapping to their corresponding target sites in the rRNAs should be performed using `RNAhybrid` [401]. A further description of the snoRNA/rRNA interaction would follow.

Previous work in the group established by using northern blots that the U3 snoRNA is developmentally regulated and that the abundance for the 37S primary rRNA transcript is decreased 8 and 16 h in the development. As a first step, the U3 snoRNA expression should be re-confirmed using RNAseq data sets from axenic growth and 16 h of development. In addition, it was of interest to determine the genomic organization and secondary structure of the U3 snoRNA in *D. discoideum* and its close relatives in the group of *Dictyostelia* using *in silico* approaches. Promoter analysis, furthermore, should reveal if any difference in the regulation of U3 snoRNA transcription appeared during the evolution of the dictyostelids.

Overall, this work aims to yield new insights in the chemical modifications and processing of *D. discoideum*'s rRNAs and reveal potential ribosome heterogeneity, also in its development.

Materials & Methods

Parts of the Materials & Methods section were taken and adapted from the PhD theses of Manfred Schäck [402], Janis Melanie Kruse [403], and Marek Malicki [404]. Parts of this chapter are included in the manuscript entitled "Ribosome heterogeneity in Amoeboza: fractional 2′-O-Methylation in the ribosomal RNA of *Dictyostelium discoideum*", which is published in Scientific Reports. All sections concerned with 2′-O-methylation in the amoeba are adapted from this manuscript.

Materials

Equipment

Autoclave	Systec, Linden
Balances:	
Finebalance	1413MP8; Sartorius, Göttingen
Binocular	Carl Zeiss Jena, Jena
Blotting equipment:	
Semidry blot	Fastblot; Biometra Göttingen
Centrifuges:	
Centrifuge 5417C	Eppendorf, Hamburg
Centrifuge 5418R	Eppendorf, Hamburg
Heraeus Multifuge 3 S-R	ThermoFisher Scientific, Waltham
Rotina 380R	Hettich, Tuttlingen
Cell counter	Casy TT Cell Counter; OLS, Bremen
Electrophoresis equipment:	
Agarose gels	Mupid-One; Advance, Tokyo
PA gels	University of Kassel
SDS PA gels	Invitrogen, Waltham
Geiger counter	MiniMonitor 900G; ThermoFisher Scientific, Waltham
Gel documentation	IX Imager 20; Intas, Göttingen
Incubators:	
Shake 'n' stack incubator	Thermo Hybaid, Heidelberg
Shaking incubator	Certomat BS-1, Sartorius, Göttingen
Incubator	ThermoFisher Scientific, Waltham

uniINCU 28 Cool incubator	LLG Labware, Meckenheim
Microscopes:	
Axio Vert.A1	Zeiss, Jena
Motic AE 2000	Motic Europe, Barcelona
LSM 510 /ConfoCor 2	Zeiss, Jena
pH-meter	Inolab pH7110; WTW Weilheim
Phosphoimager	FLA3000; Fujifilm, Düsseldorf
Power supplies:	
≤ 300 V	Enduro 300 V; Labnet, Edison
≥ 3,000 V	Consort EV232; Consort, Turnhout
Shaker for *Dictyostelium*	Laboshake; Gerhardt, Königswinter
Spectrophotometer:	
Nanodrop 2000	ThermoFisher Scientific, Waltham
Biophotometer plus	Eppendorf, Hamburg
Thermocyclers:	
Biometra TAdvanced	Analytik Jena, Jena
FlexCycler	Analytik Jena, Jena
Primus 25	Peqlab Hain Lifesciences, Nehren
realplex2 Mastercycler	Eppendorf, Hamburg
UV-table (365 nm)	ECX-F20.L; Vilber, Eberhardzell
Water bath	Memmert, Schwabach

Consumable items

24-well plates	Sarstedt, Nürnbrecht
Coverslips	ThermoFisher Scientific, Waltham
Cuvettes	Ratiolab, Dreieich
Falcon tubes (15 mL, 50 mL)	ThermoFisher Scientific, Waltham
Filter paper	Whatman, Dassel
Fluoresbrite® YG 4.5 µm carboxylate beads	Polysciences, Hirschberg a. d. Bergstraße
Nitrocellulose filter	Type HAWP 04700, Millipore, Eschborn
Nitrocellulose membrane	porablot™ NCP, Macherey-Nagel, Düren
PCR reaction tubes (0.2; 0.5 ml)	Sarstedt, Nürnbrecht
Petri dishes (10 ml)	Corning, New York
Reaction tubes (1.5 ml; 2 ml)	Sarstedt, Nürnbrecht

Syringes	Omnifix-F 1 ml B. Braun, Melsungen

Kits

GeneJet™ Gel extraction Kit	ThermoFisher Scientific, Waltham
GenElute Plasmid MiniPrep kit	Merck, Darmstadt
CloneJet PCR cloning Kit (pJET1)	ThermoFisher Scientific, Waltham

Chemicals

Unless indicated otherwise, chemicals were purchased from Roth, Karlsruhe and Merck, Darmstadt.

Antibiotics

Amphotericin B (250 µg/mL)	PAA, Cölbe
Ampicillin (50 µg/mL)	Roth, Karlsruhe
Penicillin (10000 u/mL)	PAA, Cölbe
Streptomycin (10 mg/mL)	PAA, Cölbe

Media for cell cultivation

Media for *E. coli*:

LB medium (1 L, pH 7.0)	1% (w/v)	Bacto-Tryptone
	0.5% (w/v)	Yeast extract
	85.5 mM	NaCl
LB-Amp medium	50 µg/mL	Ampicillin
	in LB medium	
LB-Agar	1.3% (w/v)	Agar-Agar
	in LB medium	
LB-Amp Agar	50 µg/mL	Ampicillin
	in LB medium	

Media for amoebae:

HL5 medium, pH 6.4	HL5 supplemented with glucose, (Formedium, 35.5 g/L)
G0 medium	50 µg/mL Ampicillin 0.25 µg/mL Amphotericin B 10 u/mL Penicillin 10 µg/mL Streptomycin in HL5 medium
BS10 medium	10 µg/mL Blasticidin in G0 medium
G10 medium	10 µg/mL Geniticin G418 in G0 medium
PYG medium	1.25 g Peptone 1.25 g Yeast extract 3 g Dextrose
Sørensen -Agar	13 g/L agar-agar in Sørensen phosphate buffer

Buffers and solutions

Alkaline phosphatase buffer (pH 9.5)	100 mM Tris 100 mM NaCl 5 mM $MgCl_2$
Blocking solution	5% (w/v) skimmed milk in 1X TBS-T

Church buffer (pH 7.2)	0.5 M	Na_2HPO_4
	1 mM	EDTA, pH 8.0
	7% (w/v)	SDS
	1% (w/v)	BSA
	0.34% (v/v)	H_3PO_4
DNA loading dye (6X)	20 mM	Tris/HCl, pH 8.0
	120 mM	EDTA pH 8.0
	50% (v/v)	Glycerol
	0.03% (w/v)	Bromphenol blue
	0.03% (w/v)	Xylene cyanol
Ethidium bromide solution	1 mg/mL	Ethidium bromide in ddH_2O
2X Laemmli buffer (pH 6.8)	0.5 M	Tris
	20% (v/v)	Glycerol
	10% (w/v)	SDS
	0.1% (w/v)	Bromophenolblue
	1 mM	DTT
NRO lysis buffer (without NP40)	50 mM	HEPES,
	40 mM	$MgCl_2$
	20 mM	KCl
	2 mM	DTT
	5% (w/v)	Sucrose
NRO reaction buffer (5X)	200 mM	Tris/HCl, pH 8.0
	250 mM	KCl
	50 mM	$MgCl_2$
	0.5 mM	DTT
	25% (v/v)	Glycerol

PA gel elution buffer	40% (v/v)	Formamide
	0.7% (w/v)	SDS
PA gel running buffer, pH 8.0	40 mM	MOPS, pH 7.0
	10 mM	Sodium acetate
	1 mM	EDTA, pH 8.0
PA gel stock solution (sterile-filtered)	7 M	Urea
	12% (v/v)	Rotiphorese® Gel40
	20 mM	MOPS, pH 7.0
Paraformaldehyde fixation solution	4% (w/v) paraformaldehyde in 20 mM phosphate buffer, pH 7.4	
10X PBS (pH 7.3)	1.37 M	NaCl
	27 mM	KCl
	14 mM	KH_2PO_4
	43 mM	Na_2HPO_4
Permeabilization solution	0.1% (v/v) Triton X-100 0.5% (v/v) FCS in 1X PBS	
PFA quenching solution	1 M glycine in 1X PBS	
Poly-L-lysine solution	25 mg/mL poly-L-lysine in 1X PBS	
Quenching solution (pH 4.4)	2 mg/mL trypan blue 5 mM sodium azide in 0.15 M NaCl / 0.15 M sodium citrate	

Ponceau S staining solution	0.5% (w/v)	Ponceau S
	5% (v/v)	Glacial acetic acid
RNA loading dye (2X)	95% (v/v)	Formamide
	17 mM	EDTA, pH 8.0
	0.025% (w/v)	Bromphenol blue
	0.025% (w/v)	Xylene cyanol
Sephadex G50 solution	5 g	Sephadex G50
	70 mL	1X TE buffer
Sørensen phosphate buffer (pH 6.0)	2 mM	Na_2HPO_4
	15 mM	KH_2PO_4
Solution I	25 mM	Tris/HCl, pH 7.4
	10 mM	EDTA
	15% (w/v)	Sucrose
Solution II	200 mM	NaOH
	1% (w/v)	SDS
Solution III	3 M	Sodium acetate
10X SDS running buffer (pH 8.3)	1.92 M	Glycine
	250 mM	Tris
	1% (w/v)	SDS
SSC (20X)	3 M	NaCl
	300 mM	Sodium citrate, pH 7.0
TBE buffer (5X)	445 mM	Tris/HCl, pH 8.2
	445 mM	Boric acid
	10 mM	EDTA, pH 8.0

10X TBS-T buffer (pH 7.6)	200 mM	Tris
	1.5 M	NaCl
	1% (v/v)	Tween-20
TE buffer (10X)	100 mM	Tris/HCl, pH 7.5
	10 mM	EDTA, pH 8.0
TNES-7U buffer	20 mM	Tris/HCl, pH 7.5
	125 mM	NaCl
	10 mM	EDTA, pH 8.0
	1% (w/v)	SDS
	7 M	Urea
Towbin transfer buffer (pH 8.3)	1.92 M	Glycine
	250 mM	Tris
	20% (v/v)	Methanol

Strains

E. coli:

XL10-Gold Ultracompetent Cells	Agilent, Santa Clara

D. discoideum:

AX2	Watts and Ashworth [4]
Δ*drnB*	Avesson *et al.* [405]
DH1-10	Cornillon *et al.* [406]
Δ*rrpA*	Wiegand and Hammann [407]
Δ*rrpB*	Wiegand and Hammann [407]
Δ*rrpC*	Wiegand and Hammann [407]
Δ*agnA*	Boesler *et al.* [408]
Δ*agnB*	Boesler *et al.* [408]
Δ*atg1*	King *et al.* [140]
Δ*wshA*	Park *et al.* [409]
Δ*vacB*	Jenne *et al.* [142]
Δ*nramp1*	Peracino *et al.* [23]
Δ*nrampB*	Peracino *et al.* [410]

Δ*racH*	Somesh *et al.* [411]
Δ*kil1*	Benghezal *et al.* [412]
Δ*alyA*	Müller *et al.* [413]
Δ*5P4*	Loovers *et al.* [414]
Δ*vsk1*	kindly provided by the Jin lab, NIH, USA
Δ*vsk2*	kindly provided by the Jin lab, NIH, USA
Δ*vsk3*	Fang *et al.* [415]

Viruses

Acanthamoeba polyphaga mimivirus	La Scola *et al.* [31]
Tupanvirus Deep Ocean	Abrahão *et al.* [416]
Cedratvirus A11	Andreani *et al.* [417]

Enzymes

Unless indicated otherwise, enzymes were purchased from ThermoFisher Scientific, Waltham or New England Biolabs, Ipswich.

Antibodies

Primary antibodies:

α-APMV	kindly provided by Bernard La Scola [31]
α-VatA	Neuhaus *et al.* [418]
α-VacB	Rauchenberger *et al.* [143]
α-p80	Ravanel *et al.* [419]
α-DKC1	Bethyl Laboratories, Montgomery

Secondary antibodies:

goat α-rabbit-Alexa488	ThermoFisher Scientific, Waltham
goat α-mouse-Alexa488	ThermoFisher Scientific, Waltham
goat α-rabbit-Alexa594	ThermoFisher Scientific, Waltham
Goat α-rabbit alkaline phosphatase conjugated	Dianova, Hamburg

Dyes

SYBR Green I	ThermoFisher Scientific, Waltham

DRAQ5™	biostatus, Leicestershire
Phalloidin-Alexa-468	ThermoFisher Scientific, Waltham
Phalloidin-Alexa-568	ThermoFisher Scientific, Waltham

Oligonucleotides

Desalted DNA oligonucleotides were purchased in a synthesis scale of 0.05 µmol from Merck, Darmstadt. Sequences are listed in Table S2 and Table S3.

Vectors

pJET1/blunt	ThermoFisher Scientific, Waltham

Molecular weight standards

GeneRuler™ 100 bp DNA ladder Plus	ThermoFisher Scientific, Waltham
GeneRuler™ 1 kb DNA ladder	ThermoFisher Scientific, Waltham
GeneRuler™ Ultra-low Range DNA ladder	ThermoFisher Scientific, Waltham
PageRuler™ Prestained Protein Ladder Plus	ThermoFisher Scientific, Waltham
RiboRuler™ High Range RNA ladder	ThermoFisher Scientific, Waltham
RiboRuler™ Low Range RNA ladder	ThermoFisher Scientific, Waltham

Software

`BEDtools` v. 2.29.2	Quinlan and Hall [420]
`Bioconductor` v. 3.13	Huber *et al.* [421]
`BLAST+` v. 2.12.0	Altschul *et al.* [422]
`bowtie` v. 1.2.3	Langmead and Salzberg [423]
`cutadapt` v. 3.3	Martin [424]
`DataCombine` v. 0.2.21	Gandrud [425]
`EndNote X9` v. 9.3.3	The EndNote Team [426]
`fastqc` v. 0.11.9	Andrews [427]
`featureCounts` v. 2.0.0	Liao *et al.* [428]
`Fiji` v. 2.1.0	Schindelin *et al.* [429]
`FlowJo` v. 10.8.0	Becton Dickinson and Company [430]
`GraphPad Prism` v. 9.1.1	GraphPad Software [431]
`LinRegPCR` v. 11.0	Ramakers *et al.* [432], Ruijter *et al.* [433]

Infernal v. 1.1.1	Nawrocki and Eddy [434]
MAFFT v. 7.475	Katoh and Standley [435]
MEME v. 5.0.2	Bailey *et al.* [436]
miniconda v. 4.10.3	Anaconda Inc. [437]
miRTrace v. 1.0.1	Kang *et al.* [438]
MUSCLE v. 3.8.31	Madeira *et al.* [439]
snoScan v. 0.9.1	Lowe and Eddy [400]
SRA toolkit v. 2.11.1	www.ncbi.nlm.nih.gov/sra
R2R v. 1.0.6	Weinberg and Breaker [440]
R v. 4.1.0	R Core Team [441]
RStudio v. 1.3.959	RStudio Team [442]
R packages used:	
ComplexHeatmap v. 2.5.3	Gu *et al.* [443]
dendsort v. 0.3.3	Sakai [444]
DESeq2 v. 1.29.6	Love *et al.* [445]
GenomicRanges v. 1.41.5	Lawrence *et al.* [446]
GenonmeInfoDb v. 1.25.10	Arora [447]
ggplot2 v. 3.3.2	Wickham [448]
mdthemes v. 0.1.0	Neitmann [449]
RColorBrewer v. 1.1.2	Neuwirth [450]
stats v. 4.0.2	R Core Team [441]
tibble v. 3.0.3	Müller [451]
RNAhybrid v. 2.1.2	Rehmsmeier *et al.* [452]
RNAviz v. 2.0.3	De Rijk *et al.* [453]
SAMtools v. 1.10	Li *et al.* [454]
SnapGene v. 5.3	GSL Biotech LLC [455]
ViennaRNA v. 2.4.14	Lorenz *et al.* [401]
ViennaRNA packages used:	
RNAalifold v. 2.4.14	Lorenz *et al.* [401]
RNAduplex v. 2.4.14	Lorenz *et al.* [401]
RNAfold v. 2.4.14	Lorenz *et al.* [401]
Viral Recall v. 2.0	Aylward and Moniruzzaman [456]
WebLogo v. 3.7	Crooks *et al.* [457]

Cell culture techniques

Cultivation of *D. discoideum*

The *D. discoideum* strains used in this thesis were cultivated in G0 medium at 22°C in shaking suspension or adherent in a petri dish. Shaking cultures were cultivated at 140 rpm. Exchange of medium and splitting was performed at 70-100% confluence.

Generation of spores from *Dictyostelium*

For the preparation of spores of *D. discoideum* for long-term storage, 2 x 10^8 cells were pelleted for 3 min at 300 x *g* and washed twice with 25 mL Sørensen buffer. Subsequently, the pellet was resuspended in 1 mL Sørensen buffer and transferred on Sørensen agar plates. After three to four days, spores were harvested by banging the plates lid facing down. The spores were collected in 1 mL Sørensen buffer and stored at -80°C.

Filter development of *D. discoideum*

Filter development was performed using 5×10^8 of axenically grown *D. discoideum* cells pelleted for 5 min at 500 x *g* and washed three times with Sørensen buffer (2 mM Na_2HPO_4, 15 mM KH_2PO_4, (pH 6.7)). The pellet was resuspended in Sørensen buffer and transferred in a 6-cm dish containing two layers of Whatman® paper topped off with a nitrocellulose membrane. After 16 h, the slugs were harvested by washing the nitrocellulose membrane with Sørensen buffer and spun down by centrifugation at 500 x g for 5 min. RNA was isolated from the resulting pellet.

Cultivation of *A. polyphaga*

The natural host of many giant viruses, *A. polyphaga*, was cultivated in PYG medium supplemented with 7% fetal calf serum, 50 µg/mL ampicillin, 250 ng/µL amphotericin B, and 500 u/mL penicillin/streptomycin at 32°C. Cell culture was split every second day. The strain was kindly provided by Didier Raoult (Aix-Marseille Université).

Generation of giant viruses using *A. polyphaga*

For the generation of viral particles, *A. polyphaga* was infected at multiplicity of infection (MOI) 1 and kept at 32°C until amoebal lysis occurred. Depending on the virus the time until

complete lysis varied between 24 h and three to four days. However, purification for all viruses were performed as described here.

Upon complete amoebal lysis, cellular debris was removed by centrifugation for 5 min at 500 x *g* and all remaining particles were pelleted for 30 min at 4,000 x *g*. The supernatant was discarded and the pellet containing the viral particles resuspended in 500 µL Sørensen buffer. For the removal of remaining amoebal genomes, 60 µL 10X DNaseI reaction buffer (+$MgCl_2$) and 25 µL *DNase I* (1 u/µL; Thermo Fisher Scientific Inc., USA) was added and the reaction incubated for 60 min at 37°C. The DNase I was inactivated by addition of 25 µL 500 mM EDTA and incubation for 10 min at 65°C. Subsequently, 9.8 mL 1 M NaCl was added, and the mixture incubated for 60 min at 4°C. The solution containing the viral particles was layered upon a 50% sucrose cushion and centrifuged for 60 min at 4,000 x *g* without brake. Viral particles found in the pellet were resuspended in 2 mL Sørensen and frozen at -80°C in 250 µL aliquots.

Cultivation of *E. coli*

Cultivation of *E. coli* was performed either on LB-agar plates or in liquid LB medium at 200 rpm and 37°C in an incubator. Depending on the selectable markers, antibiotics were added to the medium. Glycerol stocks were created by mixing 1 mL of an exponentially growing culture with 250 µL 86% glycerol (sterile), incubating for 20 min at room temperature, and subsequent freezing at -80°C.

Generation of chemically competent *E. coli* cells using the $CaCl_2$ method

For the preparation of chemically competent cells, 5 mL LB medium was inoculated with *E. coli* XL-10 Gold and incubated overnight at 37°C and 200 rpm. Subsequently, 50 mL LB medium was inoculated with 2 mL of the overnight culture and grown to an OD_{600nm} = 0.3-0.4 at 37°C and 200 rpm. Upon reaching the desired OD, the cells were pelleted by centrifugation for 10 min at 3,200 x g, resuspended in 50 mL 50 mM $CaCl_2$ (sterile), and incubated for 30 min on ice. The cells were pelleted again for 10 min at 3,200 x g and resuspended in 20 mL 50 mM $CaCl_2$ containing 15% (v/v) glycerol (sterile). The suspension was frozen in 200 µL aliquots in liquid nitrogen and stored at -80°C.

Techniques used in infection of *D. discoideum*

Infection of *Dictyostelium* with giant viruses

For the infection of *Dictyostelium* with giant viruses, 2 x 10^7 AX2 cells were infected at MOI 1 at a temperature of 22°C, if not stated differently. Each infection was performed in independent, biological triplicates. At first, the appropriate number of viruses was spread in a 10-cm dish and the AX2 cells were spun down onto the viruses at 300 x *g* for 3 min to synchronize the infection. The time allowed for infection to take was place was dependent on the experiment and was indicated accordingly. Cells and viruses for qPCR were harvested after the infection by centrifugation for 20 min at 3,000 x g. Samples for immunofluorescence were harvested by centrifugation for 5 min at 500 x *g* onto poly-L-lysine coated coverslips and fixed using 4% paraformaldehyde (PFA). After 30 min of fixation, the remaining PFA was quenched using quenching solution and the coverslips were stored in Sørensen buffer at 4°C.

External manipulation of phagosomal conditions and acid treatment of APMV

The acid treatment of giant viruses was performed by centrifugation of the viral particle suspension in Sørensen buffer at 16,000 x *g* for 5 min and resuspension in 1 mL phosphate buffer at pH 2.0. After 60 min of incubation at room temperature, the viruses were pelleted by centrifugation at 16,000 x *g* for 5 min and washed three times in Sørensen buffer. Subsequently, the acid-treated viral particles were resuspended in Sørensen buffer and immediately used for infection. The external manipulation of phagosomal conditions were performed by supplementing the G0 medium used in infection with either 20 mM NH_4Cl or 0.1X protease inhibitor. The addition of NH_4Cl results in the increase of phagosomal pH by ~2. The protease inhibitor was added to inhibit lysosomal proteases. G0 medium without any additives was used as negative control. Infections were performed with either acid-treated or untreated virus at MOI 1 as described above.

Immunofluorescence of infected *Dictyostelium* cells

Immunostaining of infected AX2 cells was performed mostly according to Hagedorn *et al.* [458]. As a first step, permeabilization was carried out in permeabilization solution for 5 min. Subsequently, the coverslips were incubated in a 1X PBS solution containing the primary antibodies for at least 60 min. After the incubation time, the coverslips were dipped three times in 1X PBS for washing and incubated again for 5 min in permeabilization solution. The coverslips were incubated with the secondary antibodies in 1X PBS for at least 60 min in the

dark. DNA was stained in the same step by addition of DRAQ5™ in a 1:1,000 and RNase A in a 1:100 dilution. The actin cortex of the AX2 cells was stained using Alexa-conjugated phalloidin at a dilution of 1:700. Followingly, the coverslips were washed three times in 1X PBS and embedded in embedding medium. The coverslips were examined using the LSM 510 and the images were processed using Fiji [429].

SYBR Green I staining of viral particles

Giant virus particles were stained for flow cytometry using SYBR Green I. For this purpose, viral particles were pelleted briefly at 16,000 x *g* for 5 min and resuspended in Sørensen buffer contained 0.1X, 1X, or 10X SYBR Green I. This mixture was incubated for 48 hours on a rolling incubator at 4°C in the dark. To remove excess SYBR Green I dye, the viral particles were washed three times with Sørensen buffer and spun down at 16,000 x *g* for 5 min between each washing step. After the last wash, the fluorescent giant viruses were resuspended in Sørensen buffer and stored at -80°C until further use.

Flow cytometry of giant viruses

Flow cytometry was performed with SYBR Green I-stained Cedratvirus (CeV) particles and AX2 cells. An amount of 2 x 10^6 particles/live cells were used in each experiment, if not otherwise indicated. In general, cells and viral particles were fixed in 4% PFA and quenched with quenching solution before flow cytometry. Before injection, cells and/or viruses were mixed in 1 mL Sørensen buffer. The injection speed was set to 1 µl per second and, due to the mixture of particles, the measurement was automatically stopped after 60 µL. Subsequent data analysis was performed using `FlowJo` v. 10 [430].

For the establishment of the staining conditions, CeV was stained with 0.1X, 1X, and 10X SYBR Green I. Fluorescent beads were used as a positive control. The supplement of the staining solution with 1% (v/v) Triton X-100 was tested with a 10X SYBR Green I concentrations during staining. Unstained virus served as the negative control.

Monitoring of phagocytosis was performed by mixing AX2 and CeV particles. The mixture was either fixed and quenched immediately or incubated for 15 min to allow for phagocytosis. Unstained AX2 cells and stained CeV were used as negative controls and the samples immediately fixed served as negative phagocytosis control.

Quenching of the extracellular fluorescence was done by addition of a trypan blue-containing quenching solution before the measurement. Quenched and unquenched viral particles and AX2

cells served as negative control for this experiment. The monitoring of phagocytosis was repeated with and without the addition of quenching solution.

To reduce the surface binding of CeV particles, viral particles were pre-incubated with sugar solutions. Based on the study by Rodrigues *et al.* [69], the monosaccharides galactose and glucose were used at a concentration of 100 µg/mL. CeV particles were incubated in either a galactose or glucose solution on a rolling incubator for 60 min at 4°C. The virus was washed three times with Sørensen buffer to remove excess sugar molecules. CeV particles, AX2 cells, and untreated virus mixed with AX2 were used as negative controls.

Molecular biological techniques

Phenol/chloroform extraction of nucleic acids

Undesired proteins were removed from aqueous solutions containing nucleic acids by phenol/chloroform/isoamyl alcohol (PCI; 125:24:1) extractions. Equal volumes of sample (at least 200 µL) and PCI were mixed and separated by centrifugation at 16,000 x g for 2 min. The resulting aqueous phase was subjected to ethanol precipitation.

Ethanol precipitation of nucleic acids

Dissolved nucleic acids were precipitated in the presence 3 vol. (v/v) ethanol and 0.1 vol. (v/v) 3 M sodium acetate (pH 4.7). Nucleic acids were allowed to precipitate at -20°C for at least 30 min. Subsequently, nucleic acids were collected by centrifugation at 16,000 x *g* for 20 min and washed twice with 70% ethanol (DNA) or 75% ethanol (RNA) to remove residual salts. The air-dried nucleic acids were dissolved in nuclease-free ddH_2O.

Agarose gel electrophoresis

Quality and size were determined using agarose gel electrophoresis. Furthermore, desired nucleic acid fragments could be isolated from a mixture of heterologous-sized fragments ≥ 100 nt. The concentration of the agarose was chosen according to Table 1 adapted from Green & Sambrook (2012). Gels were prepared by boiling high-melting-temperature agarose in 1X TBE. Upon cooling to at least 65°C, the mixture was supplemented with 0.5 µg/mL (w/v) ethidium bromide, which intercalates in double-stranded DNA and allows visualization after excitation by UV light. Since intercalated ethidium bromide retards DNA migration, gels to determine accurate DNA or RNA sizes were run without ethidium bromide and stained for 20 min in 1X TBE supplemented with 0.5 µg/mL ethidium bromide. For gel loading, 6X DNA

loading dye was mixed with the sample to a 1X concentration. The analysis of RNA quality and size was performed as described for DNA. However, 10 mM guanidinium thiocyanate (GTC) was added to the agarose after cooling and RNA samples were denatured for 2-3 min at 95°C before loading. Instead of DNA loading dye, 2X RNA loading dye was mixed with the sample to achieve a 1X concentration. Nucleic acids were separated at 6 V/cm electrode distance for DNA and 2 V/cm electrode distance for RNA.

Table 1. Resolution of linear DNA fragments depending on agarose concentration. Adapted from Green and Sambrook [459].

Agarose concentration [%]	Range of separation [bp]
0.5	700 - 25,000
0.8	500 - 15,000
1.0	250 - 12,000
1.2	150 - 6,000
1.5	80 - 4,000
2.0	50 - 2,000

Polyacrylamide gel electrophoresis of nucleic acids

Small RNAs (< 1000 nt) were separated in polyacrylamide gels containing either 7 M Urea for denaturing or no additives for native polyacrylamide gels. Stock solutions including between 6 and 12% polyacrylamide were stored in the dark until use. For polymerization, 0.5% (v/v) APS and 0.1% (v/v) TEMED were added and immediately poured between two glass plates separated by 2 mm spacers. Gel loading occurred after denaturing at 95°C or immediately for native conditions and electrophoresis was conducted at a constant current of 25 mA.

Re-isolation of nucleic acids from agarose or polyacrylamide gels

Desired nucleic acid fragments were recovered from agarose gels using the GeneJET Gel Extraction Kit. Extraction of RNA from polyacrylamide gels was performed by addition of 300 µL PA gel elution buffer and incubation overnight while shaking. Nucleic acids were concentrated using PCI extraction and ethanol precipitation.

Polymerase chain reaction (PCR)

Desired DNA fragments were exponentially amplified by polymerase chain reaction (PCR) using a thermostable DNA polymerase and specific oligonucleotides for the DNA region of interest [460].

Generally, the *Taq* polymerase was used for analytic purposes. It can also be utilized for the generation of inserts used in TA cloning, as it adds a non-templated A at the 3′-end. The *Taq* polymerase, however, lacks a 3′-5′ proofreading activity, reducing its replication fidelity. Therefore, the more accurate Phusion polymerase, carrying a 3′-5′ exonuclease activity, was used in the generation of DNA inserts for molecular cloning. PCRs were performed as described in Table 2 for Taq polymerase and Table 3 for Phusion polymerase using the general PCR program indicated in Table 4. To accommodate for the AT-rich genome of *D. discoideum*, the elongation temperature was decreased to 68°C, effectively doubling the time of elongation.

Table 2. PCR setup for the *Taq* polymerase.

Component	Final concentration/amount	Volume
Template	100-250 ng	depends on template conc.
dNTP's (5 mM each nt)	250 µM each nt	2.5 µL
$MgCl_2$ (25 mM)	0.5-2.5 mM	1-5 µL
Forward primer (3 µM)	300 nM	5 µL
Reverse primer (3 µM)	300 nM	5 µL
Taq polymerase		1 µL
H_2O		ad 50 µL

Table 3. PCR setup for the *Phusion* polymerase.

Component	Final concentration/amount	Volume
Template	100-250 ng	depends on template conc.
dNTP's (5 mM each nt)	250 µM each nt	2.5 µL
$MgCl_2$ (50 mM)	2.5 mM	2.5 µL
Forward primer (3 µM)	300 nM	5 µL
Reverse primer (3 µM)	300 nM	5 µL
Phusion polymerase		0.5 µL
H_2O		ad 50 µL

Table 4. General PCR program utilized.

Step	Temperature	Duration	Number of cycles
Initial denaturation	95°C	2 min	1
Denaturation	95°C	20 sec	
Annealing	50-65°C	20 sec	30-40
Elongation	68°C	2 min/kb	
Final Elongation	68°C	2-10 min	1

Restriction digest of plasmids

Restriction enzymes were utilized to linearize, analyze plasmids, or facilitate directional insertion of the desired fragment into the plasmid of choice. Type II restriction enzymes usually cleave within a 6 nt-long, palindromic recognition site either generating 'sticky' ends with 3 nt-long overhangs on the 5' end or 'blunt' ends without any overhangs. The set-up of an analytical restriction digest is shown in Table 5, while the mix for preparative reactions used for linearization and cloning is shown in Table 6. Generally, the restriction digest was run for at least 2 h to a maximum of 16 h at 37°C. Double digests were prepared according to *DoubleDigest Calculator* (ThermoFisher Scientific Inc., Waltham).

Table 5. Set-up of analytical restriction digest.

Component	Volume/amount
Plasmid DNA	2 to 4 µg
Restriction enzyme (10 u/µL)	0.5 µL
10X reaction buffer	1 µL
Nuclease-free ddH_2O	ad 10 µL

Table 6. Set-up of preparative restriction digest.

Component	Volume/amount
Plasmid DNA	2 to 25 µg
Restriction enzyme (10 u/µL)	2 µL
10X reaction buffer	5 µL
Nuclease-free ddH_2O	ad 50 µL

Dephosphorylation of nucleic acids using alkaline phosphatase

Nucleic acids were dephosphorylated 1U FastAP per 1 µg nucleic acids in the supplied 1X FastAP buffer at 37°C for at least 30 min and PCI-extracted and precipitated in ethanol.

Ligation of DNA fragments

Cloning of PCR fragments into pJET1/blunt was performed according to the manufacturer's instructions.

Transformation of chemically competent *E. coli* XL-10 Gold cells

Transformation of chemically competent cells was performed with 10 ng plasmid for re-transformation and 50-100 ng plasmid for fresh ligations. Competent cells were thawed on ice.

Subsequently, the desired plasmid was added, and the tube was mixed with a few careful flicks. The mixture was incubated for 30 min on ice. Cells were transformed by a heat-shock for exactly 90 sec. Afterwards, the cells were stored for 5 min on ice. For plasmids carrying an ampicillin selectable marker, cells were immediately plated on LB-agar plates containing 50 µg/mL ampicillin. Cells carrying plasmids with different antibiotics were incubated with 900 µL LB medium (without antibiotic) for at least 1 h at 37°C. After incubations, cells were pelleted for 10 min at 1,700 x *g*, resuspended in 100 µL LB medium, and plated on a LB-agar plates containing the respective antibiotic.

Plasmid minipreparation

Plasmids were isolated from overnight culture of the desired *E. coli* grown at 200 rpm and 37°C. First, 2 mL culture were pelleted for 2 min at 16,000 x *g* and the step was repeated to increase yield. The pellet was resuspended in 200 µL solution I and mixed with 250 µL solution II by inverting the tube at least 10 times. After 5 min of incubation, 300 µL solution III was added and the sample carefully mixed by inverting. Subsequently, the cellular debris was pelleted by centrifugation for 15 min at 16,000 x *g*. The supernatant (approx. 750 µL) was transferred into a fresh tube containing 750 µL isopropanol and left at room temperature for at least 20 min. Precipitated plasmids were pelleted by centrifugation for 10 min at 16,000 x *g*. The pellet was washed twice with 70% ethanol and after drying dissolved in 50 µL nuclease-free ddH_2O.

In vitro transcription

RNA molecules used in nuclear run-on transcription were generated using *in vitro* transcription. PCR products with an added T7 promoter ligated into pJET1/blunt were used as template after linearization with XhoI (10 u/µL). The reaction was prepared as shown in Table 7 and incubated for 2 h at 37°C. Subsequently, 100 µL TE buffer (pH 7.0; RNase-free) was added and mixed with an equal volume phenol:chloroform:isoamyl alcohol (125:24:1; pH 4.3). Phases were separated by centrifugation for 2 min at 16,000 x *g* and 4°C. The aqueous phase was precipitated with ethanol and the pellet was resuspended in 30 µL nuclease-free ddH_2O or RNA loading dye.

Table 7. Set-up of in vitro transcription.

Component	Volume/amount
Linearized plasmid	1300 ng
20X transcription buffer	5 µL
NTPs (5 mM each)	16 µL
$MgCl_2$ (100 mM)	5 µL
RiboLock (40 u/µL)	2 µL
Inorganic pyrophosphatase (0.1 u/µL)	5 µL
T7 Polymerase (20 u/µL)	10 µL
Nuclease-free ddH_2O	ad 100 µL

Nuclear run-on transcription

Nuclear run-on transcription was used to chart the rate of transcription of the 37S primary rRNA transcript throughout *D. discoideum* development. For this purpose, cells were harvested at time points 0, 8, 16, and 24 h of development and their nuclei were isolated by gentle lysis of the cells in NRO lysis buffer. The isolated nuclei resumed transcription when provided with a reaction buffer containing ATP, CTP, GTP, and α-P32-UTP, which upon hybridization with complementary probes, allowed for the quantification of transcription of the 37S primary rRNA transcript.

Upon harvest by centrifugation for 5 min at 500 x *g*, the cells washed with 20 mL ice-cold Sørensen buffer and re-centrifuged. The resulting pellet was resuspended in 16 mL NRO lysis buffer and 2 mL Percoll and 2 mL 10% NP40 was added. Cell lysis occurred during 15 min incubation on a rolling incubator at 4°C and the nuclei were pelleted for 15 min at 3,200 x g and 4°C. The pelleted nuclei were resuspended in 20 µL NRO reaction buffer (5X) and stored at -80°C until use. For nuclear run-on transcription, the reaction was prepared as shown in Table 8.

Table 8. Set-up of reaction for nuclear run-on transcription.

Component	Volume [µL]
Nuclei (in NRO reaction buffer)	20
NTPs (ATP/CTP/GTP; 5 mM each)	5
RiboLock (40 u/µL)	1
α-P32-UTP (0.37 MBq/µL)	10
ddH_2O	74

Before the start of the reaction, the nitrocellulose membrane carrying the *in vitro*-transcribed probes complementary to the 5′ ETS and the housekeeping gene *gpdA* was pre-hybridized for

60 min with Church buffer in a hybridization chamber. The reaction was mixed carefully by pipetting and incubated for 20 min at room temperature. For the isolation of the nucleic acids, the volume of the reaction mix was increased to 200 µL with 1X TE, mixed 1:1 with phenol/chloroform, and the aqueous phase was run through a Sephadex G-50 column. The eluted fraction was denatured for 5 min at 95°C and added after pre-hybridization in fresh Church buffer to the membrane. Hybridization was carried out over night at 42°C. The following day, the membrane was washed two times for 10 min with 2X SSC and two times for 10 min with 0.5X SSC. For detection, a screen exposed to the membrane for three to four days was read with the phosphoimager. The signals of the 5′ ETS was normalized to the *gpdA* signal and plotted over time.

Radiolabeling of nucleic acids

DNA oligonucleotides were purchased from Merck and are listed in Table S2. For primer extension and northern blot analysis, 10 pmol oligonucleotide was 5′-end-labeled by incubation with 10 u T4 polynucleotide kinase for 30 min at 37°C in 50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 100 µM spermidine, and 0.37 MBq [γ-32P]-ATP. The reaction was stopped at 80°C for 5 min, the radiolabeled oligonucleotides were phenol/chloroform-extracted and purified using a Sephadex G50 (GE Healthcare) column. Commercial ladders were labelled after 5′-end dephosphorylation.

RNA extraction from *D. discoideum*

RNA was isolated from 2×10^7 axenically grown *D. discoideum* cells washed with pre-cooled Sørensen buffer. Cells were pelleted and resuspended in TRIzol reagent (Invitrogen) containing 10 mM EDTA (pH 8.0). RNA was extracted according to the manufacturer's instructions. RNA concentration was determined spectrophotometrically.

Primer extension

For primer extension, a box C/D snoRNA-specific 5′-radiolabeled oligonucleotide was annealed to 4 µg RNA at 65°C for 5 min and cooled for at least 1 min on ice. Upon annealing, 1X SuperScript IV buffer, 1 mM dNTP mix, 5 µM DTT, 40 u RiboLock RNase Inhibitor, 50 u SuperScript IV Reverse Transcriptase. The reaction was incubated at 55°C for 30 min and stopped at 85°C for 5 min. Products were phenol/chloroform-extracted, recovered by ethanol

precipitation and separated on a polyacrylamide gel (12% PAA, 20 mM MOPS, pH 7.0, 7 M Urea) for 3 h at 25 mA.

Extraction of genomic DNA of infected *Dictyostelium* cells

Genomic DNA of infection *Dictyostelium* cells was isolated using TRIzol and TNES-7U buffer. For that purpose, infected cells and extracellular viral particles were harvested by centrifugation for 10 min at 3,500 x *g* and resuspended in 1 mL TRIzol™ Reagent (ThermoFisher Scientific Inc., USA). Upon homogenization, the sample was incubated for 5 min at room temperature and, subsequently, 200 µL chloroform was added and the sample was mixed by vortexing. After 3 min of incubation at room temperature, the phases were separated by centrifugation for 20 min at 16,000 x *g* and 4°C. The resulting aqueous phase was discarded or was used for RNA isolation (according to the manufacturer's instructions). For the separation of the DNA and nucleoproteins contained in the interphase, 300 µL TNES-7U buffer was added, mixed properly by inverting, and incubated for 10 min at room temperature. Once more, the phases were separated by centrifugation for 20 min at 16,000 x *g* and 4°C The aqueous phase containing the genomic DNA was transferred into a fresh tube, while the remaining organic and interphase contains the proteins which could be isolated according to the manufacturer's instructions. The DNA was precipitated using ethanol at -20°C at least for 2 h and pelleted by centrifugation for 30 min at 16,000 x *g* and 4°C. Subsequently, the pellet was washed thrice with 70% ethanol, air-dried, and resuspend in nuclease-free ddH_2O. The yield of each DNA isolation was calculated with the spectrophotometrically measured DNA concentrations.

Determination of primer efficiency in quantitative PCR

To determine the quality of the designed oligonucleotides, the primer efficiency of all oligonucleotide pairs was performed on genomic DNA (gDNA) isolated from *D. discoideum* and each giant virus. Quantitative PCR was performed using the PowerUp™ SYBR™ Green Mastermix (2X; ThermoFisher Scientific Inc., USA). Oligonucleotides targeting the genes of the putative viral amine oxidase *R656* for APMV, the *MCP* for TPV and *gpdA* for *D. discoideum* were used. The primer efficiencies were determined by preparation of a ten-fold dilution series of viral and amoebal genomes, respectively. For each 10 µL reaction, 1 µL gene-specific primer (3 µM each; forward and reverse primers were pre-mixed), 5 µL PowerUp™ Mastermix, and 3 µL nuclease-free H_2O were added to 1 µL of each dilution. Each reaction was performed in triplicates and a negative control containing nuclease-free H_2O instead of template was included. The reactions were run with the settings shown in Table 9. The acquired

C_T values were plotted over the dilution factor and the slope was determined by a semi-logarithmic regression in GraphPad Prism 9 [431]. A doubling of the amplificate each cycle, i.e., 100% primer efficiency, corresponds to a shift of the C_T value by 3.323. Consequently, primer efficiency can be calculated with the following formula:

$$Primer\ effiency\ [\%] = \frac{Slope\ of\ dilution\ series}{3.323}$$

Table 9. Settings for quantitative PCR. Steps 3 and 4 were cycled 40 times and detection of SYBR Green I fluorescence was performed at the end of step 4. The melt curve was performed during steps 5 to 7 with an incremental increase in temperature by 0.01°C/sec with consecutive readings of the SYBR Green I fluorescence.

Step	Temperature [°C]	Time [sec]	
2	95	120	Initial denaturation
3	95	15	Denaturation
4	60	60	Annealing/Elongation
5	95	15	
6	60	60	Melt curve
7	95	15	

Quantitative PCR

After determination of the primer efficiency of the used oligonucleotides, qPCR was performed on genomic DNA isolated from infected *Dictyostelium* cells. For this purpose, 10 µL reactions were prepared containing 1 µL gene-specific primer (3 µM each; forward and reverse primers were pre-mixed), 5 µL PowerUp™ Mastermix, and 10 ng DNA. The reactions were run with the settings described above. The copy number normalized to the culture volume was calculated for the acquired data. For this purpose, the mean PCR efficiency for each oligonucleotide was calculated based on individual PCR reaction efficiencies using LinRegPCR v. 11.0 [432, 433]. This reduced between-reaction variation based on minor handling errors or other influences. The number of DNA molecules of each genome in the reaction was calculated:

$$Copy\ number = 10^{\left(\frac{C_{T\ mean} - y_{intercept}}{slope}\right)}$$

To normalize to the culture volume, the total yield of isolated DNA (representing 10 mL culture volume) was divided by the DNA amount in the reaction. The resulting dilution factor was multiplied by the copy number in the reaction, yielding the number of DNA molecules in 10 mL

culture volume. A division by ten resulted in the normalized viral genomes per mL. This calculation was performed for each biological replicate. Errors are displayed as standard error of the mean (SEM). Results were visualized using `GraphPad Prism 9` [431].

Sodium-dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed as described by Laemmli [461]. SDS was used for the denaturation of secondary and non-disulfide-bridged tertiary protein structures. The molecule possesses a polar sulfate group which results in a protein charge in proportion to their mass [461]. Therefore, proteins are separated exclusively by their electrophoretic mobility which depends on the molecular weight. The SDS-PAGE gel composition is shown in Table 10.

Table 10. Polyacrylamide gel composition.

Gel	Component	Volume [mL]
Separating gel (12% polyacrylamide)	1.5 M Tris/HCl, 0.4% (w/v) SDS, pH 8.8	2.5
	37.5% (v/v) acrylamide/1% (v/v) bisacrylamide	3
	Millipore H_2O	4.4
	TEMED	0.004
	10% (w/v) APS	0.05
Resolving gel (4% polyacrylamide)	0.5 M Tris/HCl, 0.4% (w/v) SDS, pH 6.8	1
	37.5% (v/v) acrylamide/1% (v/v) bisacrylamide	0.5
	Millipore H_2O	2.42
	TEMED	0.005
	10% (w/v) APS	0.02

Samples were denatured in 2X Laemmli buffer at 60°C or 70°C for 10 min in thermoblock. The polyacrylamide gel (1.0 mm) was loaded with 6 µL PageRuler Prestained protein ladder (10 to 170 kDa) and up to 30 µL of the desired samples (~1 x 10^6 *D. discoideum* cells). Gel electrophoresis was conducted at a constant current of 25 mA and a maximum of 200 V in 1X SDS-running buffer. A western blot was immediately performed after the electrophoresis.

Western blot

For the detection of the dyskerin-ortholog nola4, proteins separated by SDS-PAGE were transferred to a nitrocellulose membrane by semi-dry electroblotting in Towbin transfer buffer at a constant current of 75 mA for 90 min. Transfer efficiency was controlled by Ponceau S staining for 5 min. After destaining with ddH_2O, the blotted membrane was shaken in the blocking solution for at least 30 min before the blotted membrane was incubated with the

primary antibody α-DKC1 at a 1:2,000 dilution overnight at 4°C on a rolling incubator. On the next day, three washing steps with blocking solution were performed for 10 min each. A secondary alkaline phosphatase-conjugated antibody was subsequently added in a 1:10,000 dilution and allowed to bind for at least 60 min. After washing again with blocking solution three times 10 min each, the membrane was washed once in AP staining solution. Staining was performed in 40 mL AP staining solution containing 50 µg/mL NBT and 25 µg/mL BCIP for 15 to 30 min in the dark. As soon as the protein bands were visible, the reaction was stopped by rinsing the membrane with ddH_2O.

Bioinformatics

Resources for RNA-seq datasets

RNA-seq datasets of AX2 and Δ*drnB* in axenic growth and slug stage of development were acquired from the sequence read archive (https://www.ncbi.nlm.nih.gov/sra) and used for RNA-seq validation of box C/D snoRNA candidates and expression analysis. Accession numbers of the utilized data sets can be found in Table 11. Sample preparation and sequencing was described in Liao *et al.* [462].

Table 11. RNA-seq data sets retrieved from SRA[a] [462].

Strain	Stage	Replicate	Accession
AX2	axenic	1	SRX3776204
AX2	axenic	2	SRX3776205
AX2	slug	1	SRX3776206
AX2	slug	2	SRX3776207
Δ*drnB*	axenic	1	SRX3776208
Δ*drnB*	axenic	2	SRX3776209
Δ*drnB*	slug	1	SRX3776210
Δ*drnB*	slug	2	SRX3776211

[a]https://www.ncbi.nlm.nih.gov/sra

In silico identification and validation of box C/D snoRNA candidates

The genomic sequences were retrieved from dictyBase (www.dictybase.org) and the sequences of the 17S rRNA and 26S rRNA [15] were retrieved from GenBank (www.ncbi.nlm.nih.gov/genbank/). The identification of box C/D snoRNA candidates in *D. discoideum* was performed using `snoScan` v. 0.9.1 [400] with threshold settings (`-C 0 -D 0 -X 0`) disabled. Candidates with a combined box C and box D score higher than 9 and a box

C-D distance between 50 and 100 nt were selected for RNA-seq validation. Sequencing reads from the axenic AX2 dataset were aligned to the genomic coordinates +/- 150 bp using `bowtie` v. 1.2.3 [463] allowing for one mismatch. Box C/D snoRNAs were considered validated, if reads specifically matched the predicted loci and read coverage calculated with `BEDTools coverage` v. 2.29.2 [420] indicated a distinct 5′ end, yielding an expression score of 15. Box C/D snoRNA candidates lacking expression or a distinct 5′ end received a penalty of -15. All scores were combined into a classifier score containing C/D box scores, terminal stem score, Box C-D distance score, and the expression score (Figure 10). If a total classifier score of 29 or higher was achieved, the candidate was considered to be an expressed *bona fide* box C/D snoRNA and kept for further analyses and assignment to the predicted ribosomal 2'-O-Me pattern.

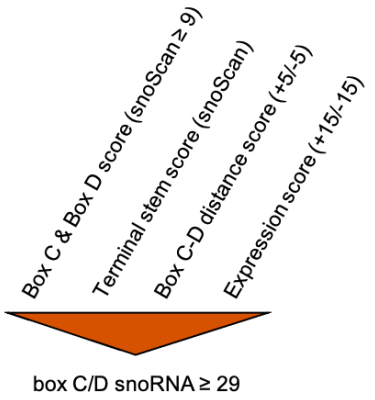


Figure 10. Selection criteria for box C/D snoRNAs in *D. discoideum*. Classification as box C/D snoRNAs based on snoScan scores for box C and box D, Box C-D distance and terminal stem score, augmented by score or penalty for RNA expression..

RNA-seq analysis of box C/D snoRNAs and U3 snoRNA in development

Reads were aligned using `bowtie` v. 1.2.3 [463] allowing for one mismatch and counted with `featureCounts` v. 2.0.0 [428]. Between-sample normalization was done by `DEseq2` v. 1.29.6 [445]. P-values were adjusted using the false discovery rate (FDR) method. Principal component analysis was performed on DESeq2-normalized reads using the `R-stats` v. 4.1.0 package and visualized with `R-ggplot2` v. 3.3.2. The heatmap of $\log_2$ fold-change of box C/D snoRNAs was generated using `ComplexHeatmap` v. 2.5.3 [443].

RiboMeth-seq

The RiboMeth-seq analysis was performed in triplicates with barcoded adapters according to previously described protocols [250, 464]. In brief, 10 µg RNA from each sample was degraded

by alkaline for 6 minutes at 90°C and the 20–40 nt fraction was excised and purified from a 10% urea polyacrylamide gel. A modified *Arabidopsis* tRNA ligase was used to ligate adaptors to the library fragments, and sequencing was carried out on the Ion Proton sequencing platform. The reads were mapped to 17S and 26S from *D. discoideum* (GenBank: 17S: FR733593.1; 26S: FR733594.1) using `bowtie2` [423] and scored for read-end counts. RMS scores representing "fraction methylated" were calculated as described previously ("score C") in [250] and barcode correction was applied when necessary [255]. The commercial RNA oligonucleotides used as 3′adaptors were found to be slightly heterogeneous in length, which can cause a fractional shift in the 3′-read-end count, if the 3′-library fragment nucleotide is identical to the expected 5′-end of the oligonucleotide. As the experiments were made in triplicate with barcodes carrying different 5′-ends, such errors were easily detected, and a manual correction was made at a few sites to counter the effect by excluding the 3′-read-end counts from the analysis. The experiments were performed in cooperation with Ulf Birkedal and Henrik Nielsen (University of Copenhagen, Denmark).

HydraPsiSeq

The HydraPsiSeq analysis was performed in independent triplicates according to the published protocol [399]. In brief, 50 ng total RNA was treated with 50% (w/v) hydrazine for 60 min on ice. After precipitation, the resuspended RNA was incubated for 15 min in 1 M aniline (pH 4.5) at 60°C in the dark. The library was constructed after 3′-end dephosphorylation using the NEBNext Small RNA Library kit and sequencing was carried out using an Illumina HiSeq 1000 with 50 bp single-end reads. The reads were mapped to the rRNA sequences of *D. discoideum* (Genbank: 5S = FR733597.1; 5.8S = FR733595.1, 17S: FR733593.1; 26S: FR733594.1) `bowtie2` in end-to-end mode. Normalization of 5′-end read counts was performed in a 10 nt rolling window and the U cleavage profiles were determined. The values for scoreA, scoreB and scoreC (Ψ score) were calculated based on these profiles [399, 465]. ScoreA is based on the average and standard deviation of positions neighboring methylated positions, while scoreB weighted average of these positions used for manual inspection of the data [250]. Normalization of scoreB yields scoreC and represents the fraction of Ψ at a given position [250]. The experiments were performed in cooperation with Virginie Marchand and Yuri Motorin (Université de Lorraine, France).

Prediction of rRNA secondary structure

To locate the predicted 2′-O-Me sites in the mature rRNA, we predicted the secondary structure and the nucleotides in the A, P, and E sites by comparative analysis with the LSU and SSU rRNAs of *A. thaliana*, *C. elegans*, *H. sapiens*, and *D. melanogaster*. For that purpose, we retrieved the corresponding SSU and LSU rRNA sequences for these organisms from GenBank (Table 12). We aligned the sequences to the 17S and 26S rRNA of *D. discoideum* using `MUSCLE` v. 3.8.31 [439] in the ClustalW output format and inferred the secondary structure by homology manually. The resulting secondary structure diagrams were drawn using `RNAviz` v. 2.0.3 [453].

Table 12. Accession numbers of rRNA sequences used in this study.

Species	Subunit	Accession	length [nt]
A. thaliana	18S	X16077.1	1809
A. thaliana	25S	X52320.1	3539
C. elegans	18S	NR_000054.1	1754
C. elegans	26S	NR_000055	3509
D. discoideum	17S	FR733593.1	1871
D. discoideum	26S	FR733594.1	3741
D. melanogaster	18S	NR_133559.1	1995
D. melanogaster	28S	NR_133562.1	3970
H. sapiens	18S	NR_146146.1	1869
H. sapiens	28S	NR_003287.4	5070

Mapping of predicted snoRNA candidates to the rRNA 2′-O-Me pattern

Mapping of box C/D snoRNAs to the predicted 2′-O-Me sites was performed using `RNAhybrid` v. 2.1.2 [452]. 10 nt upstream and downstream of the 2′-O-Me sites were used as target sites against the full-length sequences of the box C/D snoRNAs. Selection of the likely correct duplex was achieved using the following criteria: (I) 2′-O-Me site is located at the 5th base paired nucleotide upstream of a D or D′ box and (II) a box C/D snoRNA/rRNA duplex length of minimum 7 bp with (III) a maximum of 1 mismatch. Conservation of box C and box D motifs was visualized using `WebLogo` v. 3.7 [457]. Calculation of the predicted duplex' minimum free energy (MFE) in kcal/mol was performed using `RNAduplex` v. 2.4.14 [401]. Box C/D snoRNAs that were not mapped to any predicted 2′-O-Me sites but were validated by RNA-seq, were classified as orphans.

Annotation of genomic copies and promoter analysis of the U3 snoRNA

The identification of the genomic copies of the U3 snoRNA was performed using the `cmsearch` function part of `Infernal` v. 1.1.1 [434] on the genomes of representatives of the major groups of the Dictyostelia (Table 13). The covariance model (CM) of the U3 snoRNA family (www.rfam.xfam.org; Rfam ID: RF00012) deposited in the Rfam database [466] was used for this purpose. Since the U3 snoRNA is structurally conserved, `Infernal` hits with a score higher than 25 were considered *bona fide* U3 snoRNA loci.

Table 13. Utilized genomes of representatives of all major groups of the Dictyostelia.

Name	Group	GenBank accession	Assembly level
Dictyostelium discoideum	4	www.dictybase.org	chromosome
Dictyostelium purpureum	4	www.dictybase.org	scaffold
Dictyostelium lacteum	3	GCA_001606155.1	contig
Polysphondylium pallidum	2A	GCA_000004825.1	scaffold
Acytostelium subglobosum	2B	GCA_000787575.2	scaffold
Dictyostelium fasciculatum	1	GCA_000203815.1	scaffold

In order to identify upstream sequence elements occurring in the potential promoter regions of the U3 snoRNA, 150 bp upstream were analyzed using `MEME` v. 5.0.2 [436]. The TATA-like motifs were manually curated, due to the AT richness of many dictyostelid genomes. Logos of identified upstream sequence elements were created using `WebLogo` v. 3.7 [457]. Heterogeneity in the U3 snoRNA sequences in Dictyostelia with two or more U3 snoRNA genes was determined by alignment using `MAFFT` v. 7.475 [435].

Prediction of the (consensus) secondary structure of the U3 snoRNA

The secondary structure of the U3 snoRNA in D. discoideum was predicted using the `RNAfold` function part of the `ViennaRNA` package v. 2.4.14 [401] at 22°C (`-T 22`) and the diagrams were drawn using `RNAviz` v. 2.0.3 [453]. Annotations are based on the generalized U3 snoRNA model by Marz and Stadler [467]. The consensus secondary structure was generated by alignment of all dictyostelid U3 snoRNA sequences using `MAFFT` v. 7.475 [435], secondary structure prediction by `RNAalifold` v. 2.4.14 [401], and subsequent drawing using `R2R` v. 1.0.6 [440].

Viral Recall analysis of the genome of *D. discoideum*

The genome of *D. discoideum* was analyzed with the tool `Viral Recall` v. 2.0 [456]. The program predicts the proteins encoded in the genomes and analyzes the protein sequences in a rolling window of 13-16 ORFs for homology with NCLDV proteins. Positive scores indicate regions with NCLDV signatures. To identify the viruses corresponding to the predicted NCLDV regions, the resulting sequences were aligned against the NCLDV database at NCBI using `BLAST+` v. 2.12.0 [422]. Ten viruses with the highest number of homologous protein sequences were listed.

Results

Detection and quantification of giant viruses in *D. discoideum*

Establishment of qPCR-based giant virus quantification

Assessment of primer efficiency and viral genome quantification

As the first step of this project, a qPCR-based technique of detecting and quantifying giant viruses was established. This approach did not only allow for the detection of small amounts of viruses but also the exact quantification of viral genomes. For this purpose, oligonucleotides specific to the virus of interest were utilized. In case of APMV, oligonucleotides were chosen that were specific for the viral amine oxidase *R656*. For the TPV and CeV, oligonucleotides were designed which bind to the coding region of the *MCP* gene of the individual viruses. Oligonucleotides targeting *gpdA* have been widely established and genome counting of *D. discoideum* yielded an indication of amoebal growth [468-470].

As the next step, primer pairs for APMV (Figure 11A) and TPV (Figure 11B) were tested on defined amounts of viral gDNA. This experiment served to first demonstrate the primer

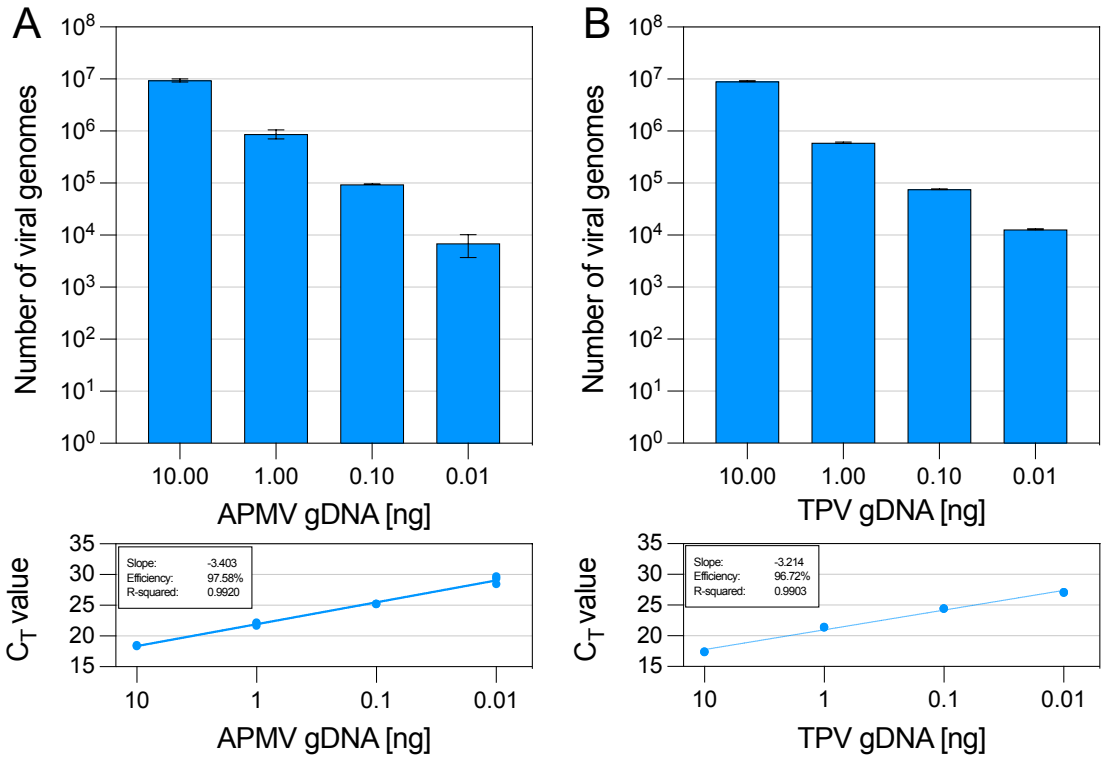


Figure 11. Test of oligonucleotides and determination of their primer efficiencies. The threshold cycle (C_T) values indicate the number of cycles required for the signal to cross the fluorescence background signal. (A) Test of oligonucleotides for R656 in APMV (n = 3). (B) Test of oligonucleotides for MCP in TPV (n = 3).

efficiencies of the oligonucleotide used in qPCR and, secondly, whether the DNA copy numbers (i.e., number of viral genomes) calculated *in silico* and measured by qPCR correlate. The latter calculations determined that 10 ng APMV gDNA correlated to ~7.84 x 10^6 molecules and 10 ng TPV gDNA correlated to ~6.11 x 10^6 molecules. For APMV and the primer pair targeting R656, the primer efficiency was ~98%, with the number of viral genomes almost exactly matching the *in silico* calculation (Figure 11A). The *MCP* oligonucleotides for TPV exhibited a similar primer efficiency at ~97%, however, the abundance of viral genomes did not completely correlate with the calculated number (Figure 11B). For detecting and quantifying viral replication in *D. discoideum*, properties of the primers were considered sufficient. Next, it was tested whether the detection works equally well in a mixture of amoebal and viral gDNAs. For this purpose, dilutions of APMV gDNA with 100 ng AX2 gDNA were mixed and qPCR performed. Comparison of the results showed no differences between samples with or without AX2 gDNA (Figure 12).

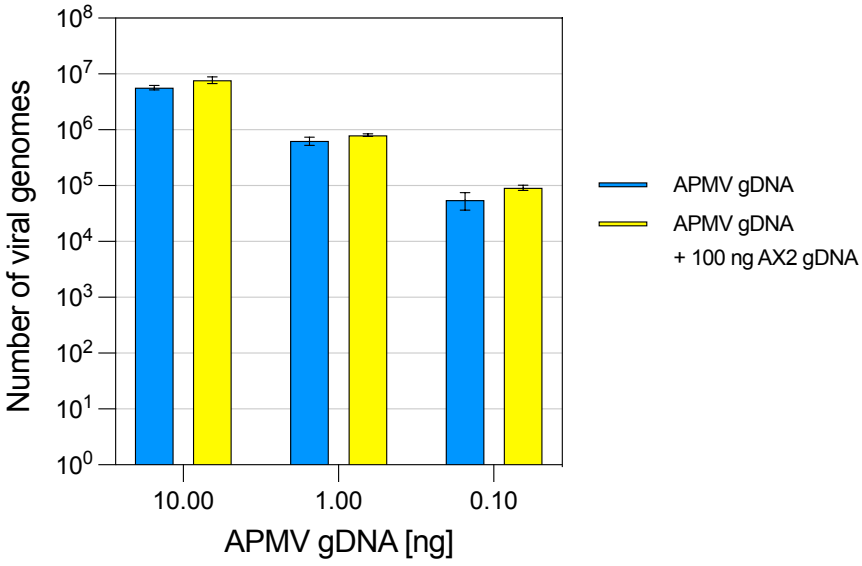


Figure 12. qPCR quantification of APMV genomes in mixture with AX2 genomes. The experiment was performed in triplicate. Quantification of isolated APMV gDNA (blue), and upon addition of 100 ng AX2 gDNA (yellow) to a dilution series of APMV gDNA.

As the last step, qPCR was performed on a mixture of virus and AX2 cells, to investigate whether the proper amount of virus was detected. For this purpose, either 2 x 10^6 (i.e., MOI 0.1), 2 x 10^7 (i.e., MOI 1), or 1 x 10^8 (i.e., MOI 5) APMV or TPV particles were added to 2 x 10^7 cells and the gDNA was immediately isolated. For both viruses, APMV (Figure 13A) and TPV (Figure 13B), the amount of virus per cell was accurately quantified, reflecting the

correct number of viral particles per cell. The number of cells was determined by qPCR targeting *gpdA* and displayed similar values as the number of cells used for this experiment.

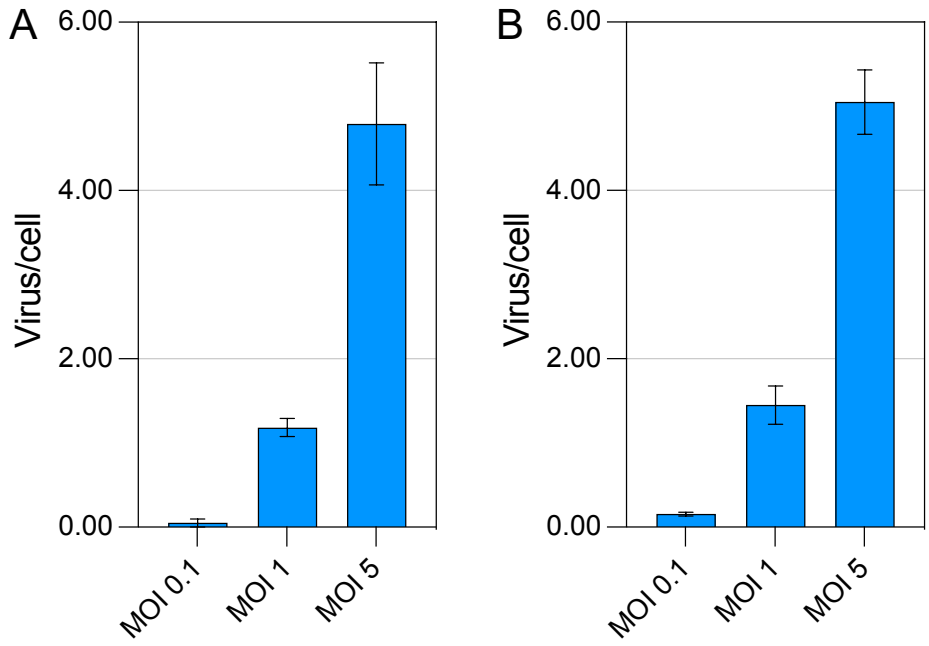


Figure 13. Test of quantifying different MOIs in gDNA isolated from an infection at 0 hpi. This experiment was performed for (A) APMV and (B) TPV in independent triplicates (n = 3). Number of cells was determined via qPCR targeting *gpdA*.

The herewith established method for the quantification of giant virus in *Dictyostelium* cultures allowed for the detection of changes of the viral genome abundance over time, i.e., in a time course during infection. Besides the molecular biological detection of giant viruses, a system for the fluorescent staining of giant viruses and their detection by flow cytometry is already established and in use as described by Khalil *et al.* [52]. Flow cytometry allows for the high-throughput quantification of (fluorescent) particles and the next step was the establishment of this technique for giant viruses and *D. discoideum*.

Establishment of flow cytometry for the detection of giant viruses

Staining of Cedratvirus (CeV) using SYBR Green I

To investigate the number of cells that internalized a viral particle, a flow cytometry-based method for their detection in samples containing *Dictyostelium* cells should be established. Earlier work provided the protocols for the staining of giant viruses with fluorescent DNA dyes like SYBR Green I and sorting using fluorescence-assisted cell sorting (FACS) [52]. For this purpose, the CeV was selected, however, DNA staining with other giant viruses worked equally well.

As a first step, various concentrations of SYBR Green I were tested, with fluorescent beads (Fluoresbrite YG 4.5 µm) being used as an internal control. SYBR Green I is a sensitive DNA dye, which works by binding DNA and the resulting DNA-dye-complex absorbs at 497 nm and emits at 520 nm. In flow cytometry, a low but detectable fluorescence of SYBR Green I-stained CeV was detected (Figure 14A). The observed non-linearity between the SYBR Green I concentration and the mean fluorescence might be explained by a saturation of viral DNA even at low dye concentrations. The staining of the virus with SYBR Green I was confirmed by fluorescence microscopy (Figure 14B). As it resulted in the strongest signal, a concentration of 10X SYBR Green I was used for the subsequent experiments.

In common immunofluorescence and staining protocols, cells are permeabilized by Triton X-100 or a similar non-ionic detergent. This treatment allows for an increased accessibility for antibodies and dyes and thus improve staining. Therefore, it was investigated whether detection of CeV particles could be enhanced by the addition of 1% Triton X-100. This treatment, however, did not result in a significant increase in the number of stained particles or the mean fluorescence of the CeV population (Figure 15). Thus, it was decided to continue staining of CeV particles without the addition of any detergent.

Monitoring endocytosis of stained CeV using flow cytometry

Upon establishment of the staining conditions and the high-throughput quantification of the viral particles, the question was posed whether the stained CeV particles were detectable when mixed with *Dictyostelium* cells. For this purpose, 2×10^6 cells were mixed with the same number of viral particles. Half of the samples was fixed immediately in 4% PFA, while the other half was allowed to endocytose CeV for 15 min before fixing. In the samples not allowed to internalize the virus, a shift of the AX2 population in fluorescence was observed, suggesting that CeV immediately attached to its cell surface (Figure 16). The AX2 population, however,

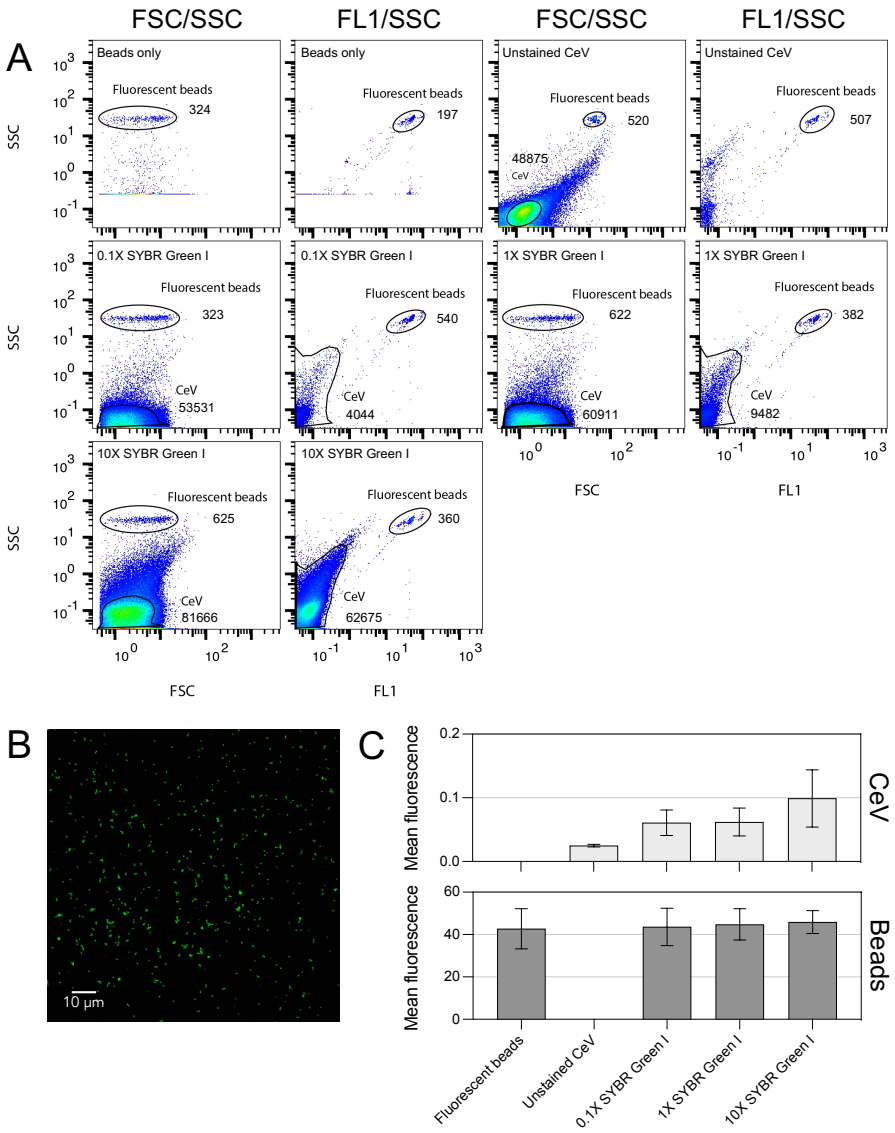


Figure 14. SYBR Green I staining of CeV particles. (A) Flow cytometric detection of fluorescent CeV. The FSC channel indicates particle size, the SSC channel particle granularity, and the FL1 channel represents SYBR Green I upon DNA binding. Fluorescent beads were used as a positive control and unstained CeV particles as negative control. CeV was stained with 0.1X, 1X, and 10X concentrated SYBR Green I. Number of particles is indicated for each gate. (B) Fluorescence microscopy of SYBR Green I-stained CeV. (C) Quantification of particle counts and mean fluorescence.

split into two subpopulations with a different mean fluorescence, which did not correspond to the AX2 control. This might be the result of varying amounts of stained CeV particles attaching to the cells. The incubation time of 15 min did not increase this effect (Figure 16), however, attached, or internalized CeV could not be discerned. Fluorescence microscopy of an aliquot of these samples revealed that a large fraction of CeV particles was stuck to the cell surface at 0 and 15 min. Phagocytosis was observed after 15 min of incubation, however, at low level (data not shown).

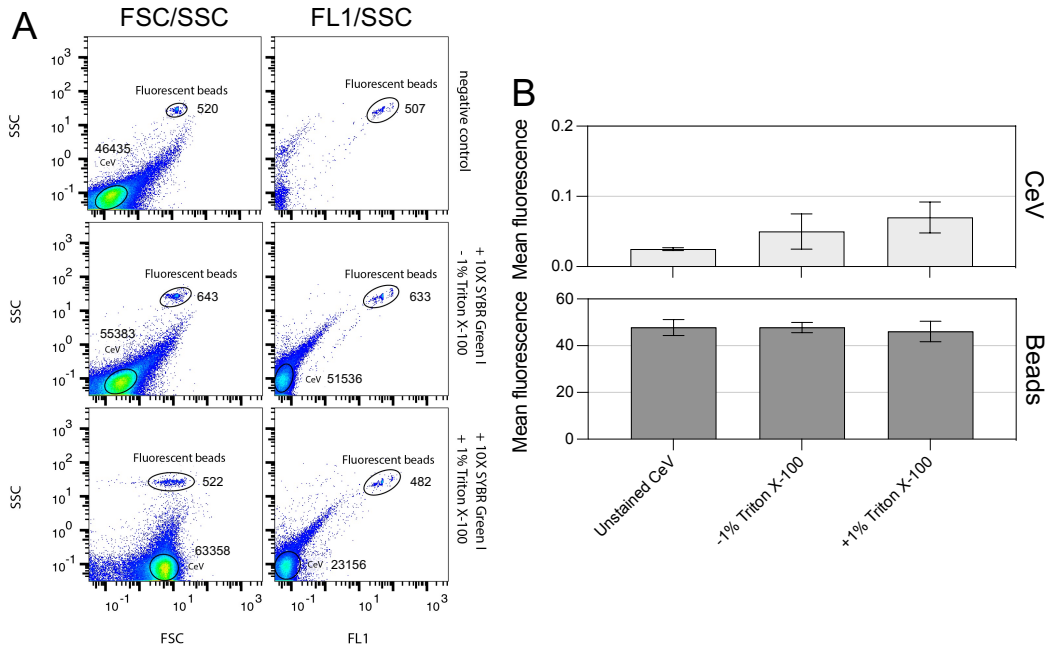


Figure 15. Flow cytometry of stained CeV particles (un-)treated with Triton X-100. (A) Flow cytometric detection of fluorescent CeV. Channel legend as described in Figure 14. FSC/SSC plots are shown in the first column and FL1/SSC plots in the second. Fluorescent beads were used as a positive control and unstained CeV particles as negative control (first row). SYBR Green I staining at concentration of 10X was performed without (second row) and with (third row) 1% Triton X-100. Number of particles is indicated for each gate. (B) Quantification of particle counts and mean fluorescence.

Differentiating extracellular and internalized viruses by quenching of fluorescence

To distinguish between adhered and engulfed virus particles, protocols to quench the extracellular fluorescence were tested. For this purpose, trypan blue, a fluorescent dye that shifts the green fluorescence from SYBR Green I into the red spectrum via resonance energy transfer, was utilized. Quenching is dependent on the proximity of both fluorophores, therefore, fluorescence of viral particles internalized by *Dictyostelium* would not expected to be quenched. Trypan blue treatment should lead to a shift in the FACS pattern with a population of AX2 cells similar to the negative control and a small population of AX2 with phagocytosed, fluorescent CeV particles. However, this was not observed, as trypan blue treatment had almost no effect on the mean fluorescence of the AX2+CeV population at 0 and 15 min (Figure 17). This suggested that trypan blue was not able to quench the fluorescence of extracellular CeV.

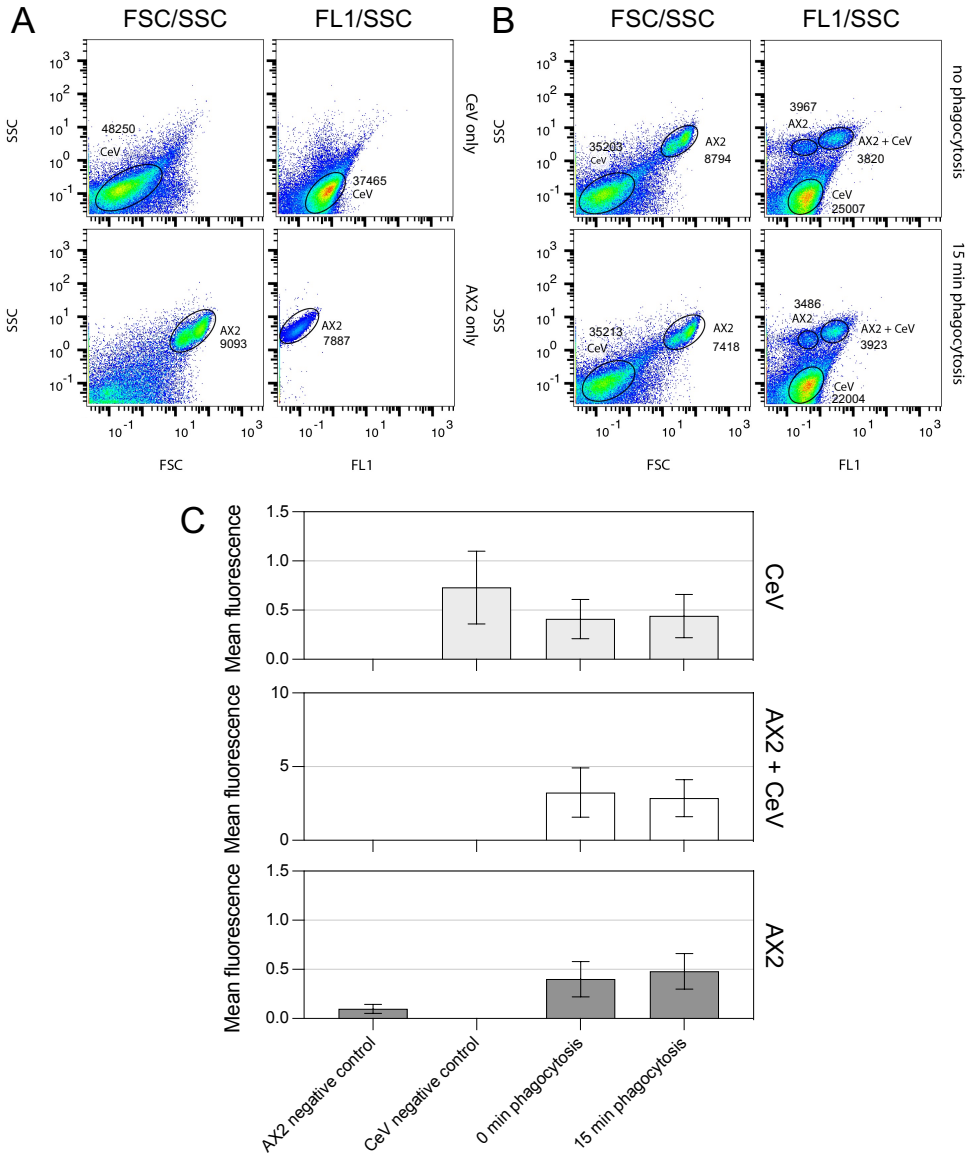


Figure 16. Monitoring phagocytosis events of CeV particles using flow cytometry. Channel legend as described in Figure 14. FSC/SSC plots were displayed in the first and third column and FL1/SSC plots were shown in the second and fourth column. (A) Negative controls for the detection of uptake. Flow cytometry was performed with SYBR Green I-stained CeV particles (first row) and AX2 cells (second row). Number of particles is indicated for each gate. (B) Detection of endocytosis of fluorescent CeV particles. The viral particles and AX2 cells were mixed at a ratio of 1:1 and were either immediately fixed (first row) or incubated for 15 min to allow for uptake (second row). (C) Quantification of particle counts and mean fluorescence.

Addition of glucose or galactose reduces extracellular attachment

Instead of quenching of extracellular fluorescence of viral particles adhering to the cell surface, another approach employed was to reduce the attachment of CeV to the surface of AX2 cells. For APMV, the attachment to glycans presented on the cell surface is mediated by its fibrils

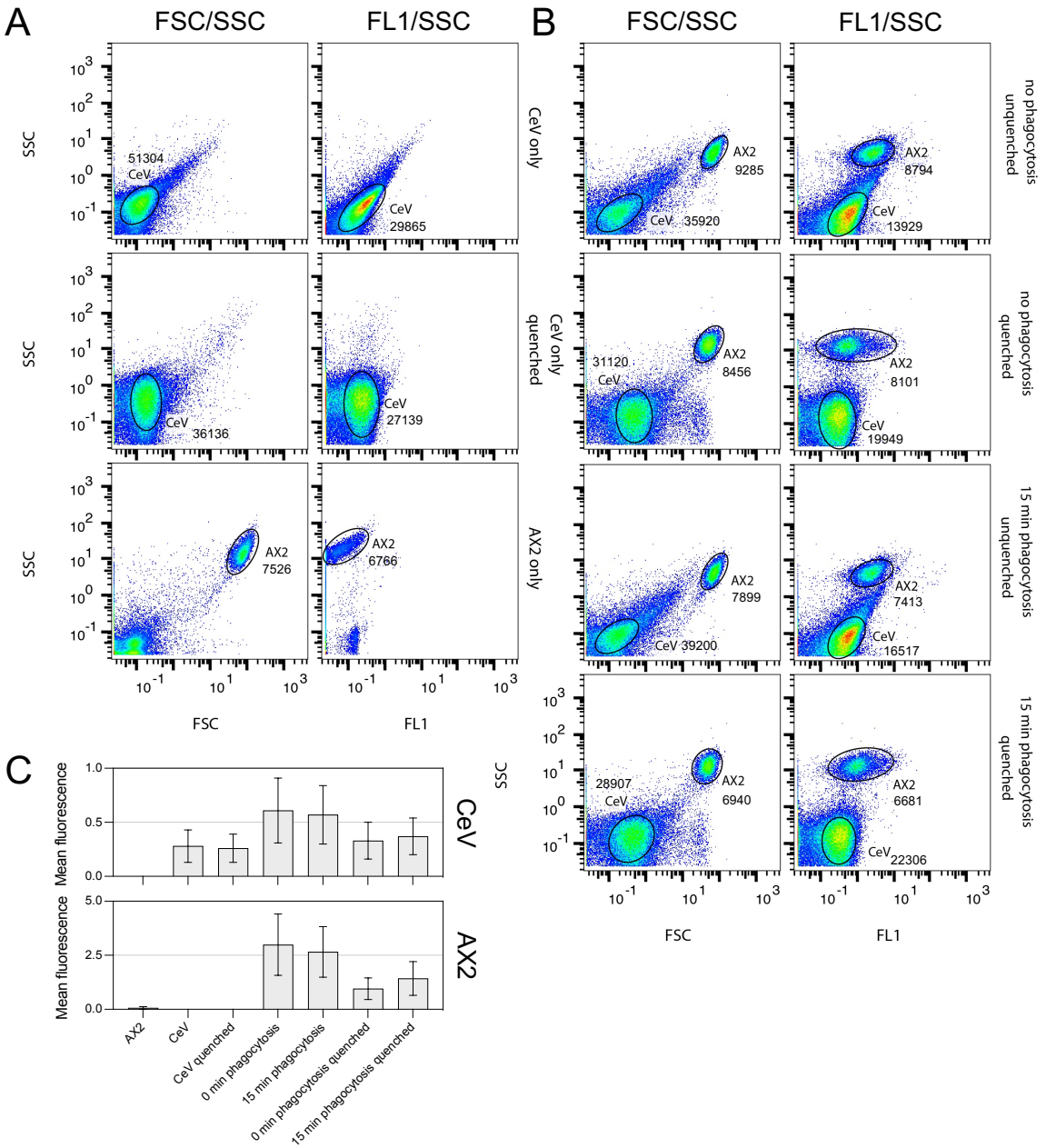


Figure 17. Quenching of the extracellular fluorescence of stained CeV. Channel legend as described in Figure 14. FSC/SSC plots were displayed in the first and third column and FL1/SSC plots were shown in the second and fourth column. (A) Negative controls for the detection of uptake and quenching. Flow cytometry was performed with SYBR Green I-stained untreated CeV particles (first row), CeV particles treated with quenching solution, and AX2 cells (third row). Number of particles is indicated for each gate. (B) Detection of endocytosis of fluorescent CeV particles without and with quenching of extracellular fluorescence. As before, the viral particles and AX2 cells were mixed at a ratio of 1:1. The samples were either not quenched after 0 or 15 min of incubation (first and second row) or quenched with trypan blue (third and fourth row). (C) Quantification of particle counts and mean fluorescence.

[69]. In this study, a reduction of viral attachment after saturation of the fibrils with sugars has

been described. Consequently, it was attempted to saturate stained CeV particles with glucose and galactose, to reduce the extracellular attachment of the virus. Indeed, a shift in the population of AX2 and CeV upon saturation of the viral particles with either glucose or galactose was observed (Figure 18), suggesting that the number of attached viral particles did reduce.

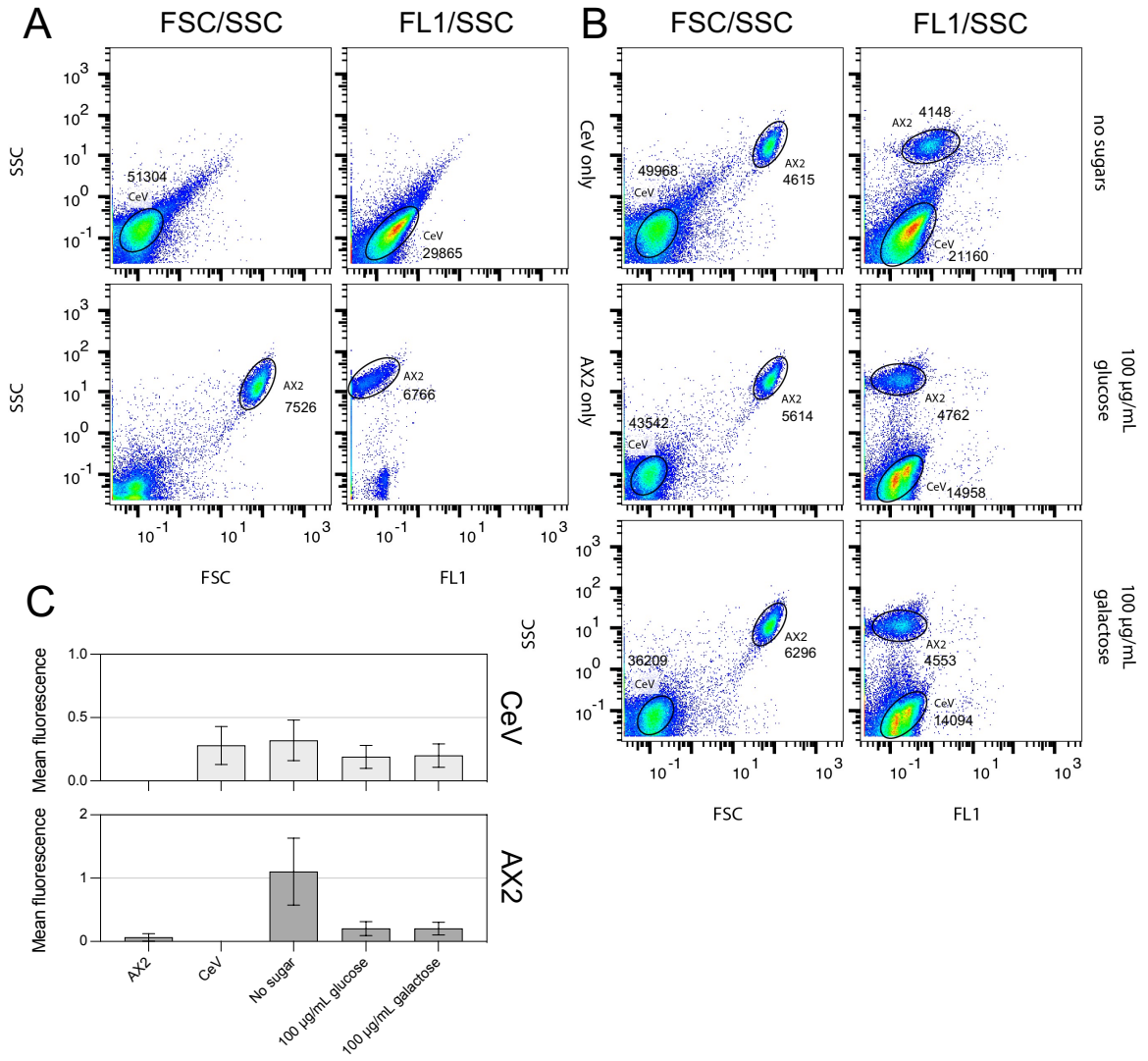


Figure 18. Effect of sugar saturation of CeV particles on cell surface attachment. Channel legend as described in Figure 14. FSC/SSC plots were displayed in the first and third column and FL1/SSC plots were shown in the second and fourth column. (A) Negative controls for the detection of uptake. Flow cytometry was performed with SYBR Green I-stained CeV particles (first row) and AX2 cells (second row). Number of particles was given for each gate. (B) Flow cytometry of untreated or sugar saturated CeV particles mixed with *Dictyostelium* cells. Untreated CeV particles were visualized in the first row. The treatment with glucose was displayed in the second row and galactose treatment was shown in the third row. (C) Quantification of particle counts and mean fluorescence.

Internalization of APMV particles and their cellular fate in the amoeba

Optimization of the multiplicity of infection (MOI)

In the next experiments, different MOIs were tested in the infection of AX2. In brief, the MOI is an indicator in infection experiments for the number of pathogens per cell. The MOIs 0.1, 1, and 5 were selected and these infections were performed in independent triplicates. To control the MOIs, the number of viruses normalized to the number of cells was analyzed (Figure 19A). The mean of the replicates at 0 hpi reflected the MOIs, suggesting that the correct amount of virus was added to each infection setup. However, abundance of the viral genome diminished at 24 hpi in all tested MOIs (Figure 19A), indicating that the number of genomes decreased. At this point, it could not be excluded that this is an effect of cell growth, as *Dictyostelium* divides every 8 to 12 h. This would result in the presence of more amoebal cells than viruses. To examine this possibility, the number of viruses was normalized to the volume of the culture medium, as it is unchanged throughout the whole experiment. Even though this approach reflects the number of viruses in infection better (Figure 19B), growth of *D. discoideum* (Figure 19C) did not mask any significant virus replication. It is unclear, however, at which stage of the infection the replication or invasion is halted.

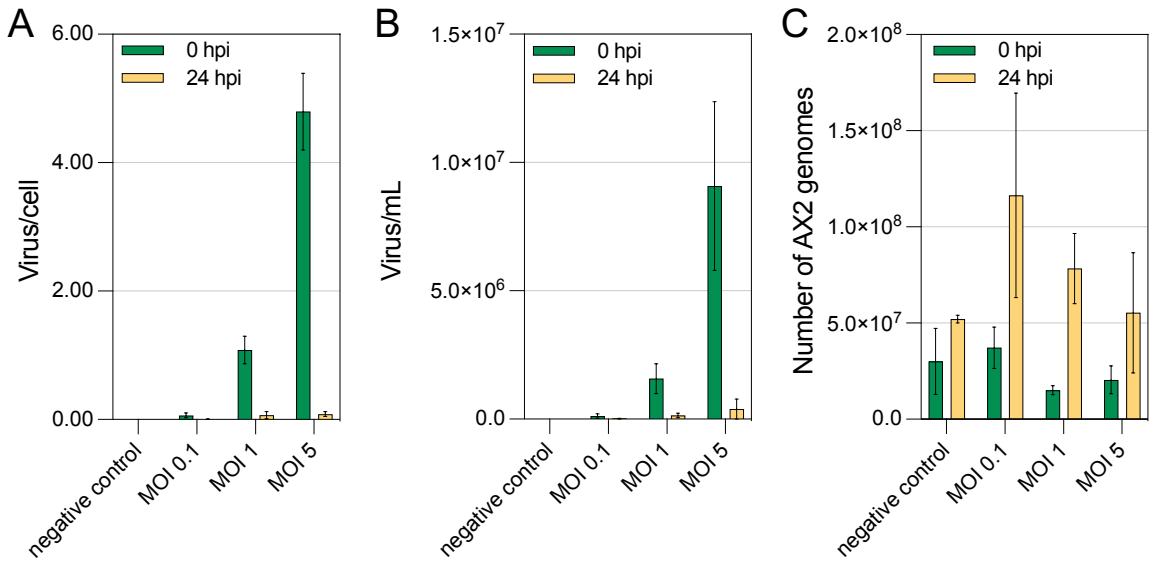


Figure 19. MOI optimization of APMV infection of *D. discoideum*. The experiment was performed in independent triplicates and the SEM was calculated. The cells were infected at the MOIs 0.1, 1, and 5 and harvested at 24 hpi for DNA isolation and qPCR analysis (A) Calculation of the number of APMV genomes per cell and (B) number of APMV genomes per mL culture medium at 0 and 24 hpi. (C) Counting of AX2 genomes at 0 and 24 hpi.

APMV enters *D. discoideum*

For the investigation the fate of APMV, uptake of the virus by *Dictyostelium* was examined. To ensure synchronous and efficient uptake, AMPV was spread evenly in a cell culture dish and AX2 cells were centrifuged onto the viral particles. Viral phagocytosis was examined using fluorescence microscopy of fixed cells at 30 mpi and it was observed that APMV is taken up by AX2 cells (Figure 20). The viral load per cell varied with some amoeba showing only a single virus and others with multiple viruses taken up (examples for both in Figure 20).

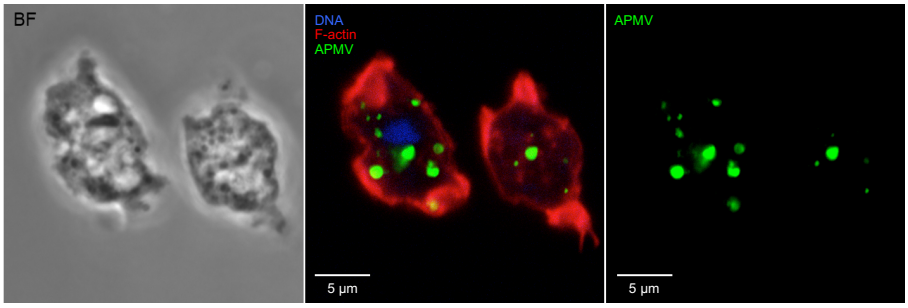


Figure 20. Immunofluorescence of APMV uptake in *D. discoideum* 30 mpi. The APMV-infected AX2 cells were fixed in 4% PFA. The brightfield (BF) images can be found in the left panel. APMV is visualized in green. DNA was stained blue using DRAQ5™ and phalloidin was used to stain F-actin in the cell cortex.

In *A. polyphaga*, APMV stays in the phagosome until acidification, which triggers the opening of the stargate structure and the subsequent escape of the viral core into the cytosol [67, 471]. As a next step, it was investigated whether the observed internalized APMV particles in *Dictyostelium* is present in phagosomal structures. For this purpose, the phagosomal markers p80 and the V-ATPase subunit VatA were immunostained and observed by fluorescence microscopy. The V-ATPase subunit VatA is recruited within minutes to the early phagosome [135], while the putative copper transporter p80 is a late endosomal marker, commonly utilized in the identification of post-lysosomal compartments [419]. At 30 mpi, APMV was mainly observed in p80 or VatA-positive compartments (Figure 21A and B). Very few particles, however, seemed to be neither colocalized with p80 or VatA. This indicates that APMV still resided in the phagosome of *Dictyostelium* in the late stages of phagosomal maturation prior to exocytosis.

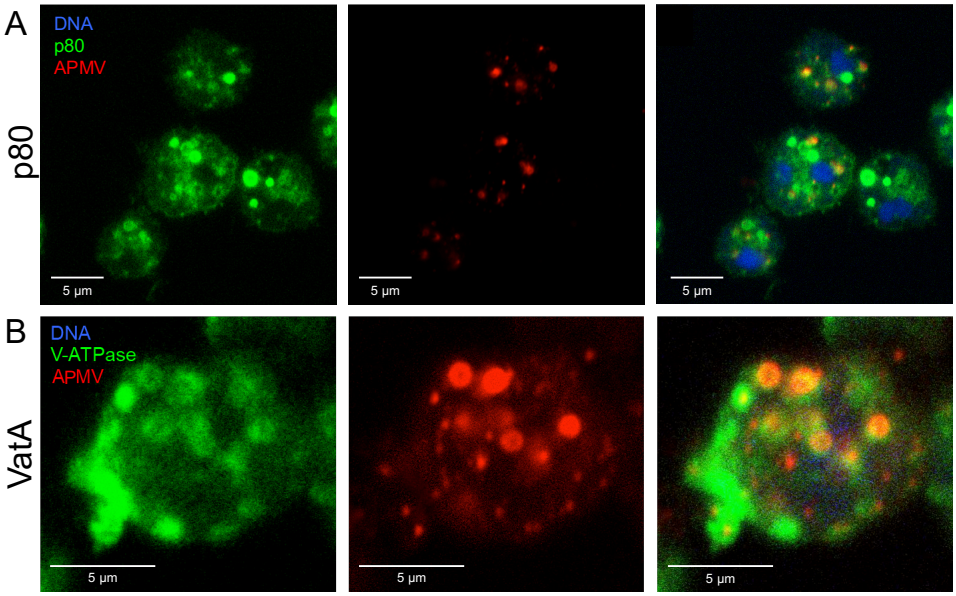


Figure 21. Surrounding of internalized APMV particles with phagosomal markers at 30 mpi. The APMV-infected AX2 cells (MOI 1) were fixed in 4% PFA. Colocalization of (A) p80 (green) and (B) VatA (green) with APMV (red). DNA was stained using DRAQ5™ (blue).

Viral factories are not formed in the amoeba

After the finding that APMV entered *D. discoideum* via the phagosomal pathway, the question was posed whether viral escape from the phagosome and the subsequent formation of a viral factory could be observed. Since the early viral factories start to emerge at 4 hpi in *Acanthamoeba* [76], the APMV-infected cells (MOI 1) at 1 hpi, 3 hpi, and 6 hpi were examined using fluorescence microscopy. In the first hour of the infection, a high number of extracellular APMV particles next to cells carrying one or multiple virus was observed (Figure 22). At 3 hpi almost all viruses have been internalized by the amoeba and, subsequently, at 6 hpi only a fraction of viral particles was still detectable using fluorescence microscopy (Figure 22). This is similar to the observations made by qPCR (Figure 19), with viral genome abundance diminishing between 0 hpi and 24 hpi. In addition, structures resembling a viral factory at the latter time point could not be observed (compared to 24 hpi of APMV-infected *A. polyphaga* cells in Figure S1). This suggested that no replication of APMV occurred. However, it cannot be excluded that replication still occurs in small fraction of *Dictyostelium* cells, due this approach being a single cell analysis method rather than a one averaging the whole population like qPCR.

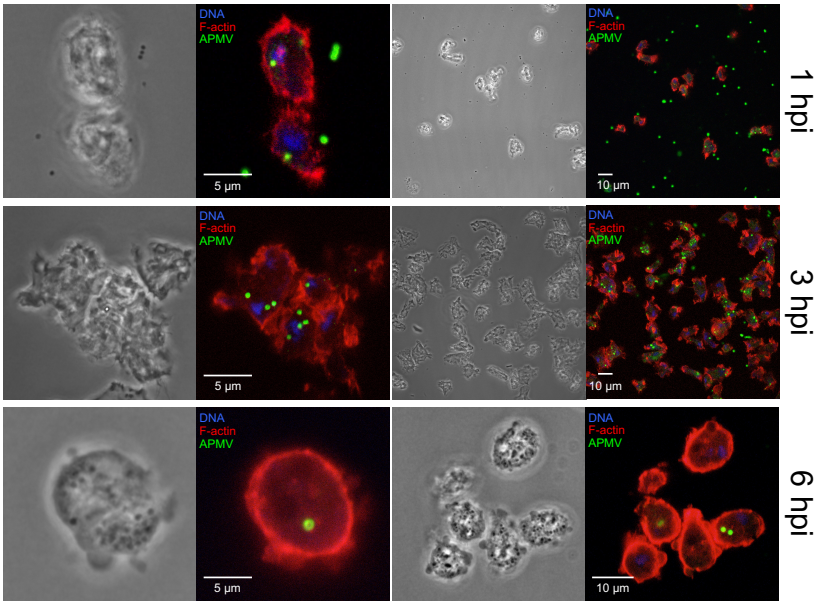


Figure 22. Fluorescence microscopy of APMV-infected AX2 cells at early time points. AX2 was infected with APMV at MOI 1 and fixed with 4% PFA. Viruses were detected with a polyclonal antibody and visualized in green. The F-actin in the cell cortex was stained using phalloidin-568 (red). DNA was stained using DRAQ5™ (blue).

Cellular fate of APMV particles in AX2

Since APMV particles were observed colocalized with the late phagosomal marker p80, it was investigated whether these compartments also acquire vacuolin, an endocytic marker recruited immediately before exocytosis [143]. In experiment with latex beads, colocalization of vacuolin and subsequent exocytosis occurs after 3 to 5 hours [472]. At 1 hpi no APMV-containing phagosomes staining positive for vacuolin were recognized (Figure 23A). However, at 6 hpi vacuolin was found localized to APMV-containing endosomes in a fraction of infected cells (Figure 23B), indicating that at least some viral particles are exocytosed after a 6-7 hpi.

Concomitantly, a decrease of viral abundance at 6 hpi could be observed, which suggests that either APMV was not stable in the utilized medium or that *D. discoideum* gradually degrades the viral particles. To exclude the former possibility, the persistence of APMV in HL5 medium was tested by incubation for 24 h and quantification of viruses using qPCR. No significant difference between the control at 0 h and 24 h incubation in HL5 medium could be seen (Figure 24). This suggests that *D. discoideum* phagocytosed and digested APMV particles during infection, leading to the observed decrease of viral abundance.

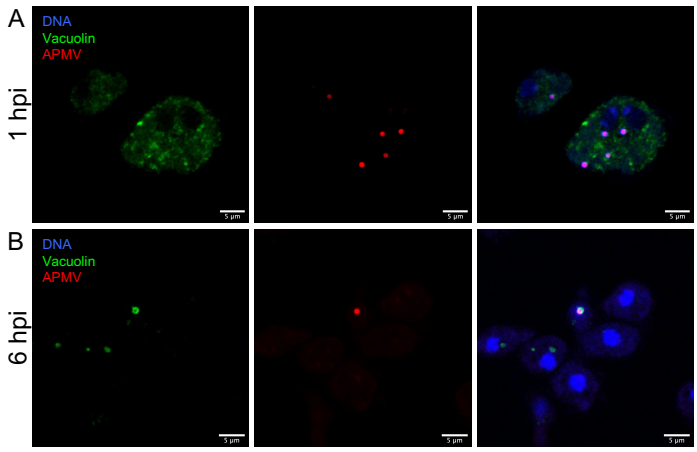


Figure 23. Colocalization of vacuolins with APMV-containing endosomes. AX2 was infected with APMV at MOI 1 and fixed at (A) 1 hpi and (B) 6 hpi with 4% PFA. Viruses were detected with a polyclonal antibody and visualized in red. Vacuolins were detected by a specific antibody and are visualized in green. DNA was stained using DRAQ5™ (blue).

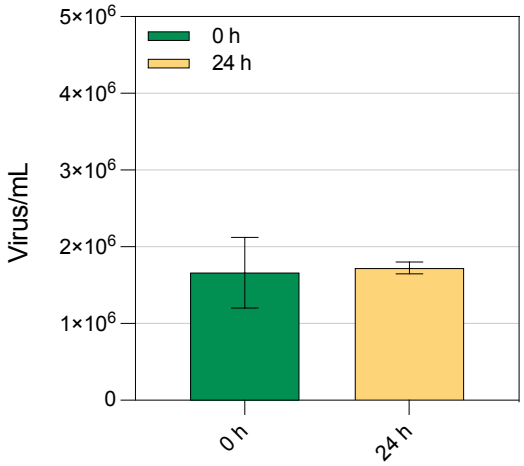


Figure 24. APMV persistence in HL5 medium. The experiment was performed in triplicate and the viral abundance was quantified by qPCR at 0 and 24 hpi. The initial viral load was 2 x 10^6 viruses /mL culture medium.

External manipulation of endosomal conditions

The indications that APMV might be degraded by *D. discoideum*, raised the question whether the manipulation of phagosomal conditions might alter APMV persistence. After initial phagocytosis, several maturation steps turn the compartment into a highly degradative, acidic, and oxidative environment/conditions, facilitating the intracellular killing and digestion of food bacteria [109]. Therefore, as a next step, it was investigated whether manipulation of this environment might change the viral abundance after 24 h. For this purpose, the medium used in infection was supplemented with either NH_4Cl to buffer phagosomal pH or protease inhibitor

to lower the activity of the lysosomal enzymes. Compared to the unbuffered conditions after 24 h, buffering of the phagosomal pH did not have a significant effect on the viral abundance after the same time (Figure 25A). The addition of protease inhibitor, however, seemed to diminish the APMV amount even further, suggesting that the lysosomal enzymes might be necessary for release of the viral core. This is supported by a study observing the release of the viral core into the cytosol after the delivery of the lysosomal enzymes to the APMV-containing phagosome [72].

In a recent study, *in vitro* loosening of the stargate structure was achieved in ~60% of giant virus particles at an acidic pH 3.0 or lower [471]. It could not be excluded that the phagosomal pH of 3.5-4 in *D. discoideum* were not ideal for viral core release. To prime the viral particles for core release, the virus was incubated in a phosphate buffer at pH 2.0 before the infection. Subsequently, the infection was performed in medium supplemented with protease inhibitor or NH_4Cl with the acid-treated virus. A significant difference in viral abundance after 24 h compared to untreated virus after the same time, nonetheless, was not observed in any of the treatments (Figure 25B). Therefore, it is tempting to speculate, that the proper release of the viral core might not only dependent on the phagosomal conditions but also on the timing of

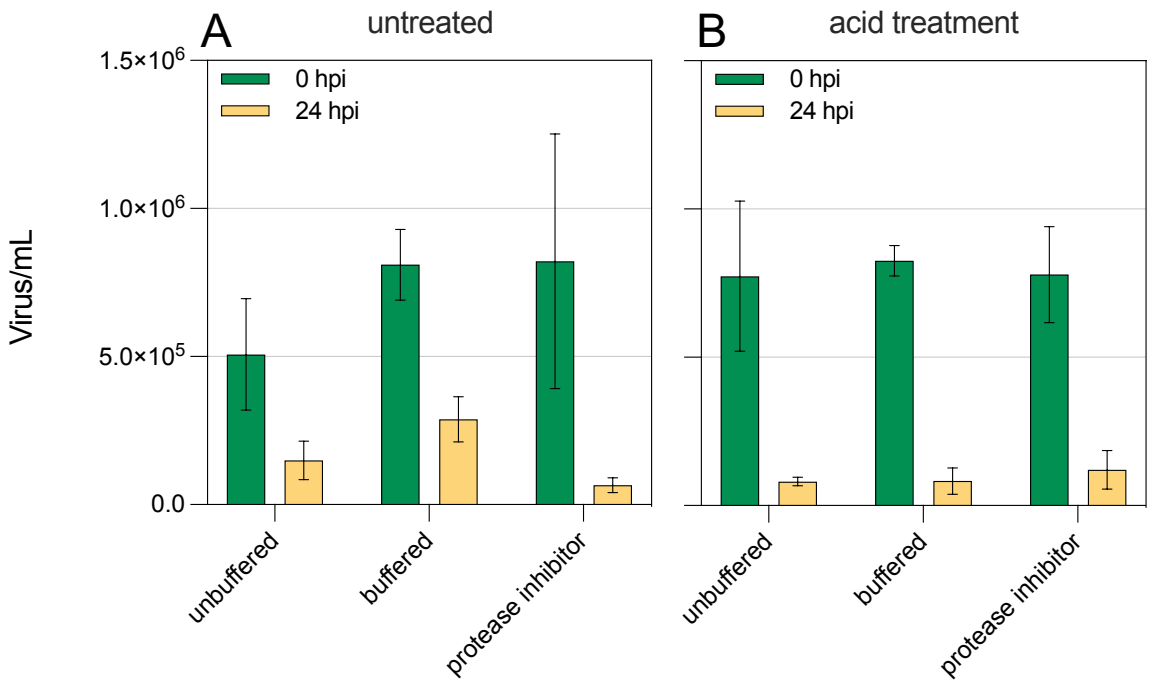


Figure 25. External manipulation of phagosomal conditions. Infection of AX2 was performed either with acid-treated or untreated viruses in medium supplemented with NH_4Cl or protease inhibitor. Untreated virus and medium without supplement constituted the negative control. The experiments were performed in triplicates with (A) untreated APMV particles and (B) acid-treated viruses. Quantification of virus was performed via qPCR and the copy numbers normalized to the volume of the culture medium.

capsid opening. An early opening, as simulated with the pre-treatment, might result in the degradation of the virus before reaching the appropriate point of phagosomal maturation.

Screening for APMV-permissive mutants of *D. discoideum*

Infection of mutants affected in phagosomal maturation and autophagy

Since no productive infection occurred, the question was posed what factors led to these observations and what might have conferred immunity to the amoeba. Fluorescence microscopy revealed that giant viruses can rarely be found outside of the phagosome, suggesting that the viral particles might degraded in the compartment. Therefore, a library of *Dictyostelium* mutant strains that show defects in phagosomal maturation was examined either by qPCR or fluorescence microscopy.

Acidification of the phagosomal compartment is a hallmark of phagosomal maturation and delivery and recycling of the responsible V-ATPase complex is a highly orchestrated process [109, 135]. A key regulator in this process is the WASH protein complex, which facilitates the retrieval and recycling of phagosomal components like lysosomal enzymes, Rab GTPases, and the aforementioned V-ATPase, etc. [123, 140]. Knockout of WshA, an integral part of the WASH complex, leads to a reduction of proteolytic activity and exocytosis extending the residence time of particles and exposing the phagosomal contents to less degradative conditions [140]. Thus, it was expected that APMV would meet more favorable phagosomal conditions for infection. Indeed, while an ~11-fold reduction in viral genome abundance was detected for the AX2 strain after 24 h (Figure 26), the number of APMV genomes in Δ*wshA* did only decrease ~1.5-fold in the same time frame, staying within the margin of error (Figure 26). The observation that the number of APMV particles did not decrease significantly in the Δ*wshA* strain, might be explained by the reduced degradative capabilities of the endosomal environment rather than increased APMV replication.

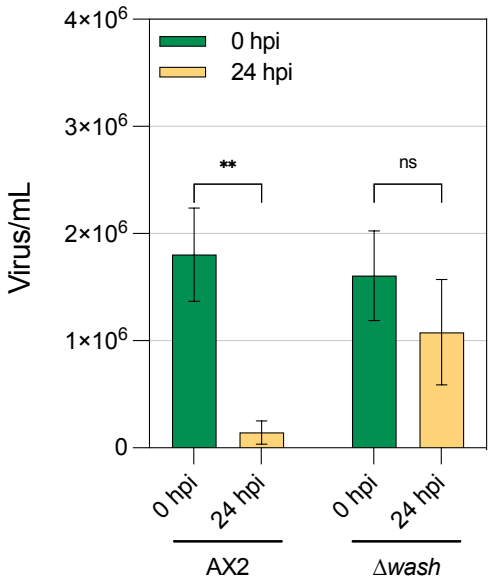


Figure 26. APMV infection of the *ΔwshA* strain. The experiment was performed with AX2 and *Δwash* cells infected at MOI 1 in independent triplicates. DNA for the qPCR quantification of viral abundance at was isolated at 0 and 24 hpi. The copy number was normalized to the volume of the culture medium. The asterisks indicate significant differences as determined by two-way ANOVA (n = 3; ** = $p < 0.005$; ns = no significance).

The knockout of VacB was screened next. During the WASH-mediated recycling of phagosomal components, the phagosome acquires postlysosomal markers VacA and VacB, which share a domain structure and lipid raft-association similar to mammalian flotillins [109, 142-144]. Both vacuolins are inhibitors of the fusion of the late endosomes [142, 473] and a study by Bosmani *et al.* [474] described that all vacuolins are necessary for efficient uptake of beads or bacteria and proper phagosome acidification and exocytosis [474]. Knockout of VacB leads to a reduction of phagocytic uptake of up to 90% [142, 474]. Thus, it was expected APMV would be degraded as in AX2 but to a lesser extent, due to the reduced uptake. This, nonetheless, did not manifest in the infection. The ~5.2-fold reduction of viral genome abundance in Δ*vacB* was comparable to APMV-infected AX2 cells (~7.2-fold) at 24 hpi (Figure 27). This might be explained by the fact that the intracellular killing of bacteria in vacuolin mutants is not reduced [474] and, therefore, arguing that the degradative conditions of the phagosomal compartment are comparable to the wildtype.

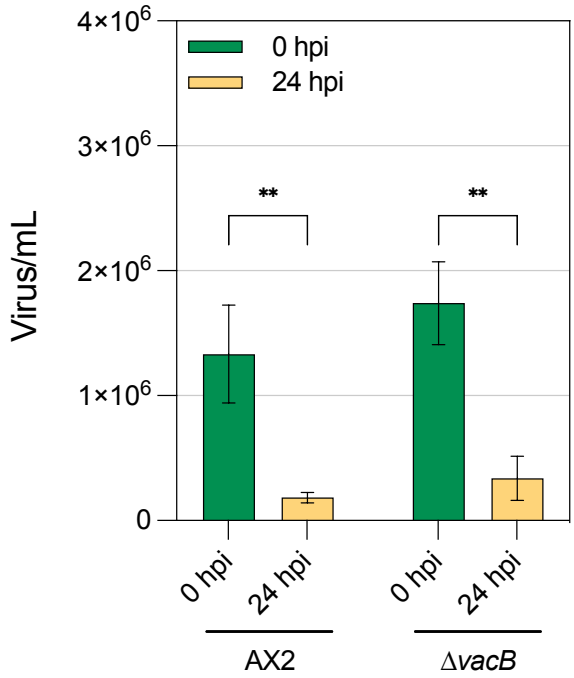


Figure 27. APMV infection of the *ΔvacB* strain. The experiment was performed with AX2 and *ΔvacB* cells infected at MOI 1 in independent triplicates. DNA for the qPCR quantification of viral abundance at was isolated at 0 and 24 hpi. The copy number was normalized to the volume of the culture medium. The asterisks indicate significant differences as determined by two-way ANOVA (n = 3; ** = p < 0.005; ns = no significance).

The next strain, Δ*racH*, is a knockout of a Rho GTPase involved in the actin-based vesicle trafficking [411]. Due to the defects in vesicle trafficking, Δ*racH* also exhibits defects in exocytosis and early phagosome acidification, which results in less harsh phagosomal conditions and longer residence time of the phagosomal cargo in the compartment [411]. Furthermore, RacH has been implied in the resistance to *M. marinum* [22] and *L. pneumophila* [475] infection, as a RacH knockout leads to increased growth of these pathogens in *Dictyostelium*. Therefore, Δ*racH* was infected with APMV. The ~2.3-fold reduction of viral abundance at 24 hpi was less compared to AX2's ~21-fold reduction (Figure 28A). At 24 hpi, viral factories and a high number of APMV particles can usually be observed in its natural host *A. polyphaga* (Figure S1). This was not seen in fluorescence microscopy of Δ*racH* cells at 24 hpi, suggesting that in summary APMV was more persistent but not necessarily productive in its infection (Figure 28B). This effect might rather be explained by the less degradative conditions in the Δ*racH* phagosomes compared to wildtype ones.

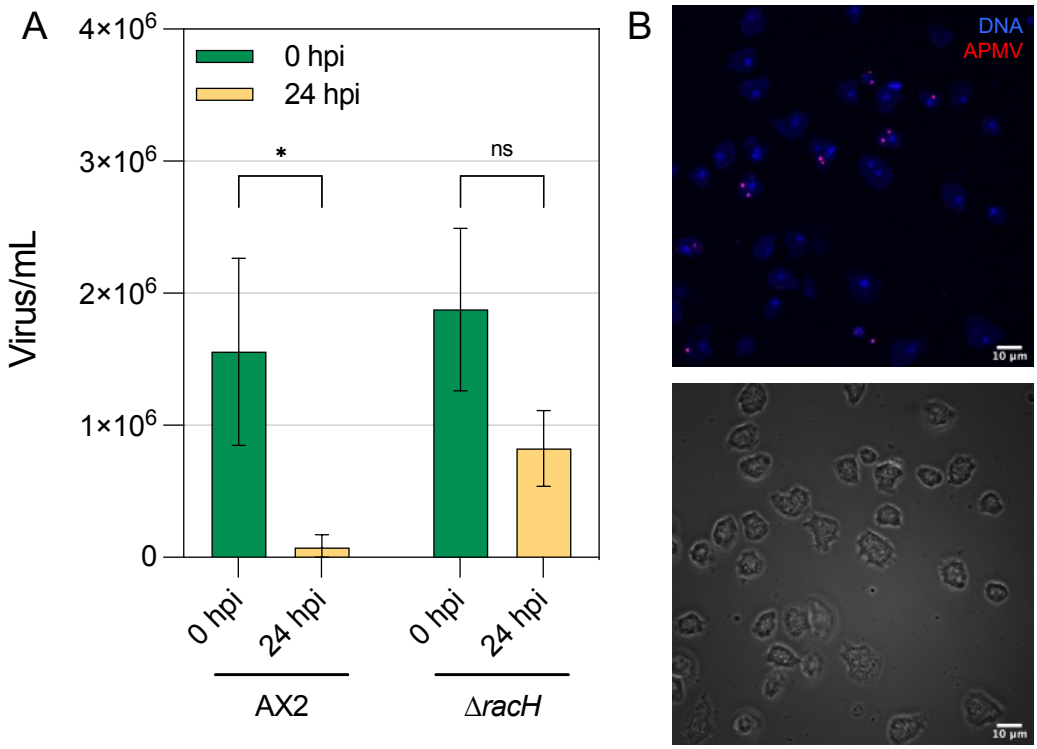


Figure 28. APMV infection of the *ΔracH* strain. The experiment was performed with AX2 and *ΔracH* cells infected at MOI 1 in independent triplicates. (A) DNA for the qPCR quantification of viral abundance at was isolated at 0 and 24 hpi. The copy number was normalized to the volume of the culture medium. The asterisks indicate significant differences as determined by two-way ANOVA (n = 3; * = $p < 0.05$; ns = no significance). (B) Fluorescence microscopy of APMV-infected *ΔracH* cells at 24 hpi.

Eight additional phagolysosomal and autophagy mutants were screened for a productive infection with APMV. However, no viral replication or any visible hallmarks of the infection up until 24 hpi were observed (Figure 29 and Figure 30) except for Δ*vsk1* and Δ*vsk2*. Both mutant strains exhibited an increased signal intensity upon immunodetection with an APMV-specific antibody. However, extracellular viral particles were not observed for Δ*vsk1* and Δ*vsk2* at 24 hpi, as seen with APMV-infected *A. polyphaga* cells. An overview over all tested phagolysosomal mutant strains, their phenotypes, and their experimental outcome in APMV infection was compiled in Table 14.

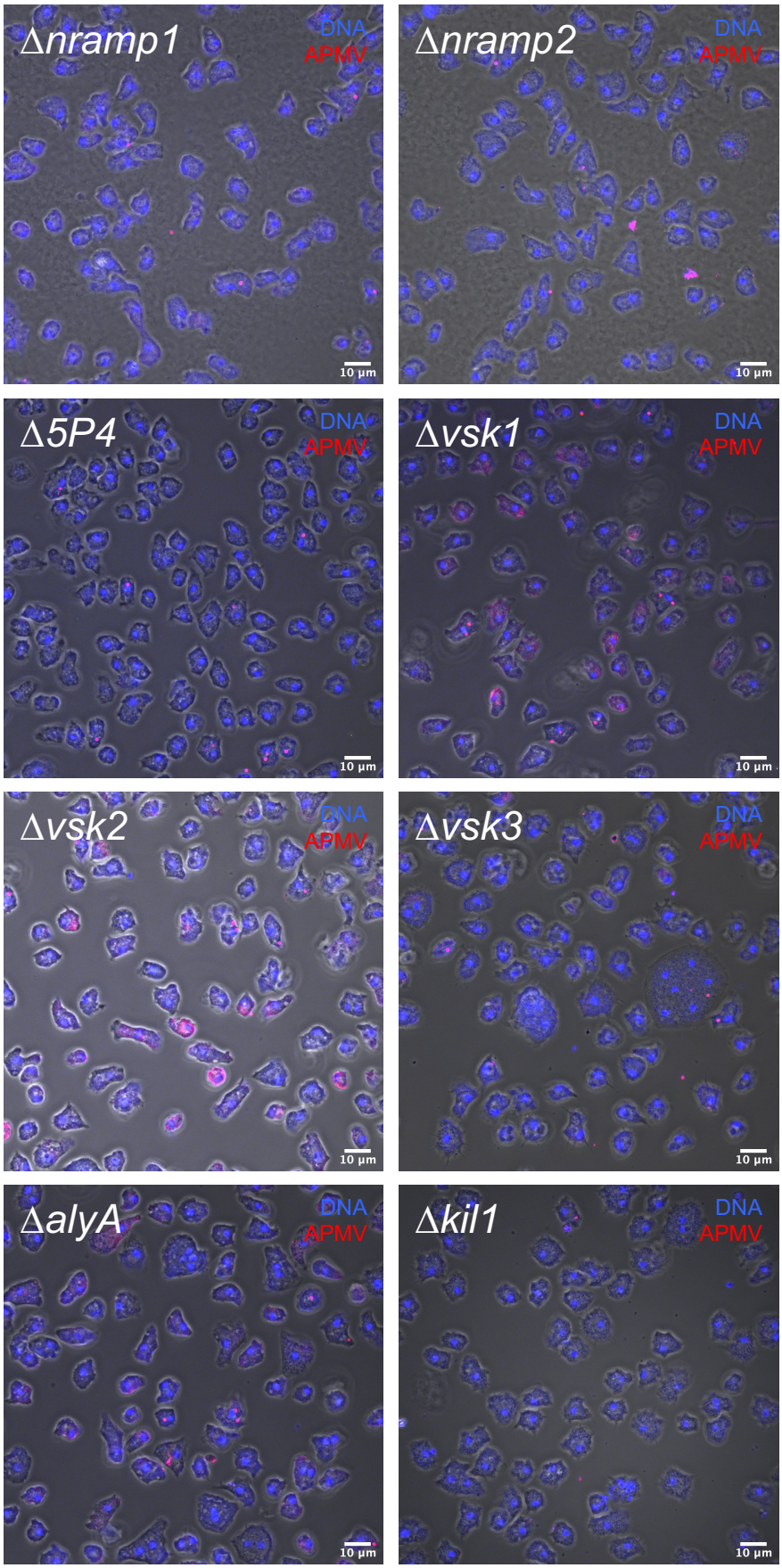


Figure 29. Screening of mutants defective in the phagolysosomal pathway at 24 hpi. APMV was stained with a specific antibody (red) and DNA was stained with DRAQ5™ (blue). No mutant exhibited any observable virus replication.

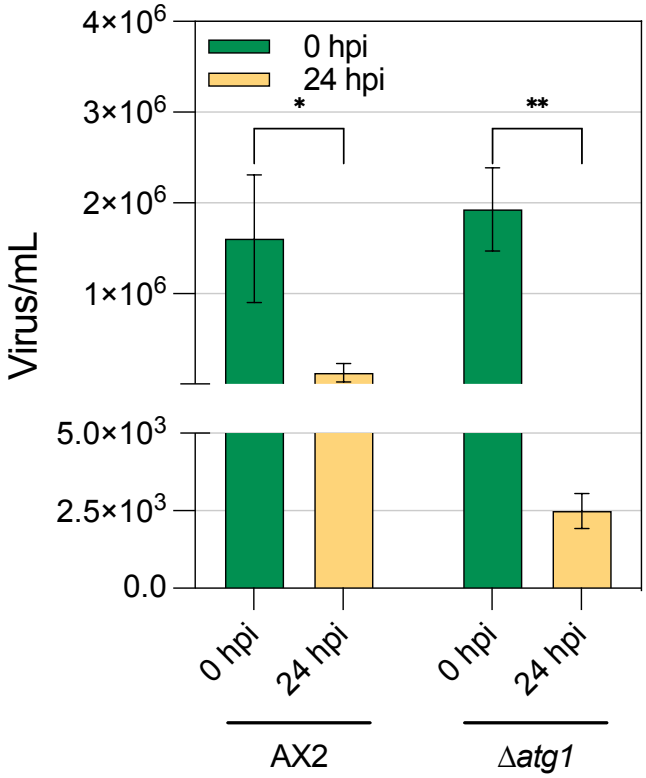


Figure 30. APMV infection of the *Δatg1* strain. The experiment was performed with AX2 and *Δatg1* cells infected at MOI 1 in independent triplicates. DNA for the qPCR quantification of viral abundance at was isolated at 0 and 24 hpi. The copy number was normalized to the volume of the culture medium. The asterisks indicate significant differences as determined by two-way ANOVA (n = 3; * = $p < 0.05$; ** = $p < 0.005$).

Table 14. Overview of (auto-)phagosomal mutant screening. All tested mutants with their experimental outcome by qPCR (Q) or microscopy (M) were listed with their respective phenotypes. Strains indicated with an asterisk were kindly provided by Tian Jin, USA.

Mutant	Viral Persistence	Mutant phenotype	Reference
ΔwshA	Yes (Q/M)	Defects in reneutralization, digestion, VatA recycling	[123, 140]
Δnramp1	No (M)	Defects in metal ion transport, less resistent to pathogens	[23, 25]
ΔnrampB	No (M)	Defects in metal ion transport, less resistent to pathogens	[25, 410, 476]
Δkil1	No (M)	Defects in intracellular killing of bacteria	[412]
ΔracH	Yes (Q/M)	Defects in endosomal pathway, less resistant to pathogens	[22, 411, 475]
Δvsk1	Possibly (M)	No phenotype described	*
Δvsk2	Possibly (M)	No phenotype described	*
Δvsk3	No (M)	Reduced phagocytosis and fusion with lysosome	[415]
Δ5P4	No (M)	Reduced phagocytosis, less resistant to pathogens	[414, 477]
ΔvacB	No (Q)	Impaired particle recognition, more resistant to pathogens	[22, 474]
ΔalyA	No (M)	Reduced lysozyme activity, increased phagocytosis	[413]
Δatg1	No (Q)	Aberrant macroautophagy	[478]

Infection mutants with defects in RNA interference (RNAi)

A common and widely described antiviral mechanism is constituted by the components of the RNAi machinery. It is already known that this pathway is able to facilitate a defense against a wide variety of viruses with diverse genomic organizations (reviewed in [178]). The RNAi machinery and the mechanisms are conserved in *D. discoideum* [108, 407, 479]. Even though, it was observed that the genome abundance for APMV was decreasing after 24 h of infection of *Dictyostelium*, it could not be excluded that a minute fraction of the viruses was able to escape the phagosome, but replication was stopped by downstream defense mechanisms like RNAi. Thus, mutants missing RNAi components were infected with APMV.

For this purpose, strains with knockouts of RdRPs and argonaute proteins were selected: Δ*rrpA*, Δ*rrpB*, Δ*rrpC*, Δ*agnA*, and Δ*agnB*. The selected mutants were infected with APMV, and the viral genome abundance was quantified at 0 and 24 hpi. While the APMV-infected AX2 cells exhibited an ~7.3-9.5-fold reduction in the number of viral genomes, the infected RdRP mutants showed a ~1.7-3.8-fold decrease and the Agn mutants revealed a ~21.1-30.0-fold decrease in viral genome abundance (31A and B). Especially in Δ*rrpB* (~1.7-fold reduction), the number of viral genomes at 24 hpi seemed to suggest viral persistence. In contrast, the Agn mutants exhibited an approximately three times stronger decrease in the viral genome abundance at 24 h. Even though the RNAi mutant strains did not exhibit a productive infection, the absence of RNAi factors seems to influence viral genome abundance in some unknown manner and could be of interest once more factors involved in *Dictyostelium* anti-viral defenses are known.

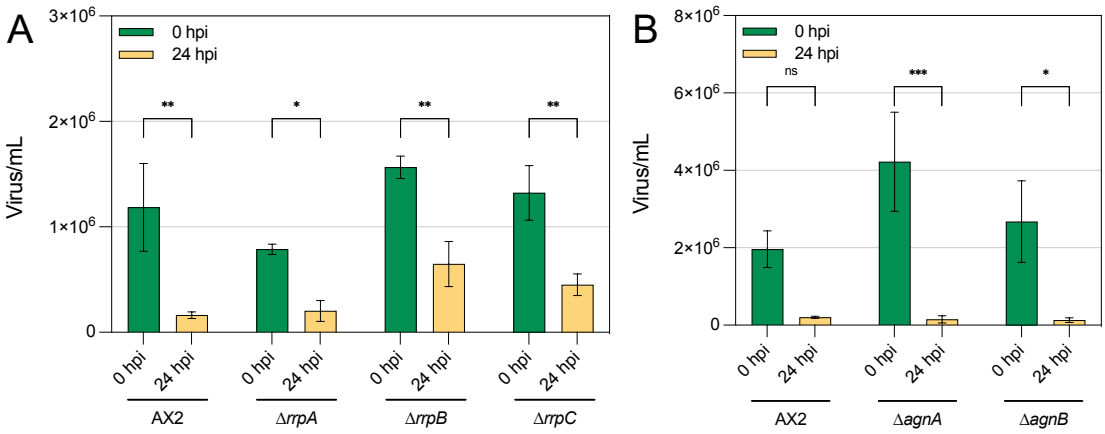


Figure 31. Infection of RNAi mutants with APMV. The experiment was performed with AX2, the RdRP mutant (A) and the Agn mutant (B) strains infected at MOI 1 in independent triplicates. DNA for the qPCR quantification of viral abundance at was isolated at 0 and 24 hpi. The copy number was normalized to the volume of the culture medium. The asterisks indicate significant differences as determined by two-way ANOVA (n = 3; ns = no significance; * = $p < 0.05$; ** = $p < 0.005$).

Tupanvirus - a similar fate as APMV

To investigate the optimal MOI for TPV, infection of the AX2 strain with TPV were performed at the MOIs 0.1, 1, and 5 in independent triplicates. As with APMV, a decrease in viral genome abundance was observed for all MOIs between 0 hpi and 24 hpi. The reductions at MOIs 1 and 5, however, were not as strong compared to APMV, arguing that TPV was persisting to some degree (Figure 32A). Normalization of viral abundance to the volume of culture medium reduced this apparent persistence (Figure 32B), suggesting that cell division was affected. Indeed, the number of cells was slightly reduced at 24 hpi for MOI 1 and unchanged for MOIs 0.1 and 5 (Figure 32C). It cannot be ascertained by this method, however, if this increased

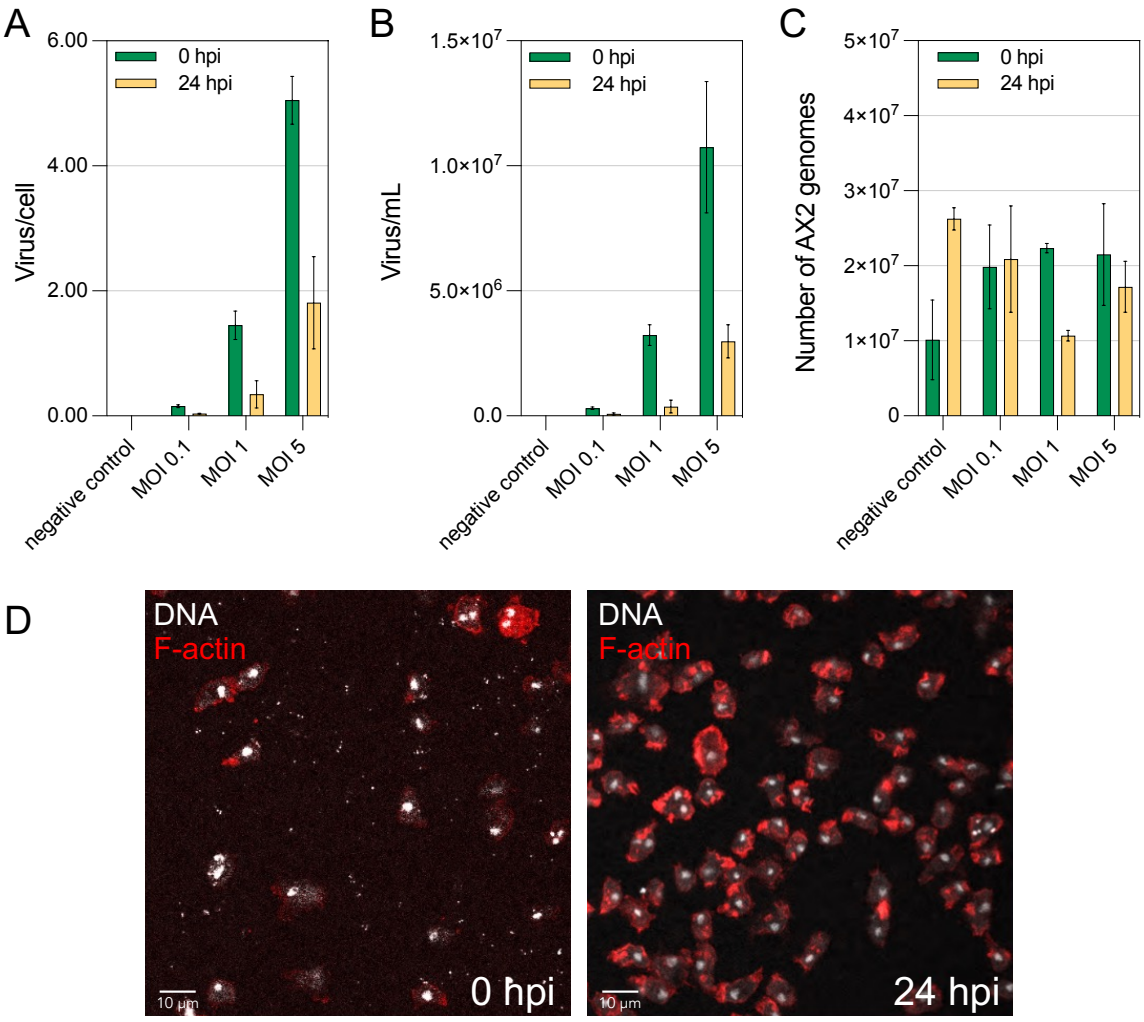


Figure 32. MOI optimization of TPV infection of *D. discoideum*. (A) Calculation of the number of TPV genomes per cell and (B) number of TPV genomes per mL culture medium at 0 and 24 hpi. (C) Cell growth was quantified using oligonucleotides specifically targeting *GpdA*. (D) Fluorescence microscopy of TPV-infected cells at 3 and 24 hpi did not confirm the observation from qPCR. To increase the visibility of the DNA staining, the color was converted to white.

persistence and the effect of cell growth can be attributed to viral replication or higher resistance to degradation of the viral particles. To investigate this possibility, fluorescence microscopy was employed on AX2 cells infected with TPV at MOI 5. Since no specific antibodies against TPV exist, DRAQ5^{TM} staining was used to stain the nucleic acids of the virus. This approach revealed that, like it was observed to be the case for APMV, after 24 h extracellular viral abundance seemed to be drastically reduced (exemplified in Figure 32D).

Chemical modifications in rRNA & ribosomal biogenesis

Parts of this chapter concerned with 2′-O-methylation are included in the manuscript entitled "Ribosome heterogeneity in Amoebozoa: fractional 2′-O-Methylation in the ribosomal RNA of *Dictyostelium discoideum*", which is published in Scientific Reports.

2'-O-Me and associated box C/D snoRNAs in *D. discoideum*

Identification and validation of 30 novel box C/D snoRNAs in the genome of *D. discoideum*

The number of 17 box C/D snoRNAs identified in *D. discoideum* prior to this study is relatively small for normally-sized rRNA sequences [15] compared to orthologous RNAs found in other species [343]. Therefore, we set out here to search for additional box C/D snoRNAs in the amoeba. To this end, we employed an *in silico*-approach for the identification of novel box C/D snoRNAs by using the probabilistic model-dependent search tool snoScan [400], which we combined with RNA-seq analyses. The sizes of previously described box C/D snoRNAs of *D. discoideum* range between 66 and 113 nt, with box C-D distances between 50 and 97 nt [346]. We searched accordingly first with snoScan in the genome of *D. discoideum* (available at www.dictybase.org) for sequences containing box C and box D motifs with a box C-D distance between 50 to 100 nt. Since inverted repeats at the 5′ and 3′ ends were not observed before [346], we did not pre-require the presence of a terminal stem structure for a classification as a *bona fide* box C/D snoRNA. Using these settings, we identified 577 box C/D snoRNA candidates in the genome of *D. discoideum* (data not shown), including the set described before [346]. In order to refine our search, we next addressed the expression of these candidates in publicly available RNA-seq data of the axenic AX2 wild type strain, deposited in duplicate [462] at the sequence read archive (https://www.ncbi.nlm.nih.gov/sra). Specifically, we mapped reads to the genomic loci of the candidates and selected only those sequences that exceeded a read count of 100 and were not part of a longer transcript, as indicated by a distinct 5′ end. Both, the lack of specific RNA-seq reads or a distinct 5′ end, were penalized ('expression score', Figure 10). Sequences scoring 29 or higher in the classifier score (Table 15) were classified as *bona fide* box C/D snoRNAs. This routine allowed us to identify 47 box C/D snoRNAs in *D. discoideum*, of which 30 are novel [346]. In addition, we carried out primer extension experiments on RNA isolated from axenically grown or developed AX2 and Δ*drnB* cells. This resulted for the majority of the snoRNAs in a single signal at the predicted size (Figure 33), indicating that they have homogeneous 5′-ends. Their genomic locations are listed in Table S3, allowing to characterize the properties of box C/D snoRNA genes in the amoeba.

Table 15. Parameters for snoScan and Classifier Scores.

	Parameters and values from Scan								Parameters and values additionally used for Classifier Score (CS)			
RNA[a]	C box	C score	D box	D score	TS[b] length (bp)	TS[b] likely formed	TS[b] score	Scan Score[c]	RNAseq expression	C-D gap (nt)	Gap score	CS[d]
CD1	AUGAUGA	12.73	CUGA	8.05	6	Yes	0.27	21.05	15	50	5	41.05
CD2	AUGAUGA	12.73	CUGA	8.05	2	No	-1.65	19.13	15	68	5	39.13
CD3	AUGAUGA	12.73	CUGA	8.05	1	No	-1.90	18.88	15	68	5	38.88
CD4	GUGAUGA	10.76	CUGA	8.05	4	No	-0.93	17.88	15	90	5	37.88
CD5	AUGAUGA	12.73	CUGA	8.05	4	No	-0.93	19.85	15	65	5	39.85
CD6	AUGAUGA	12.73	CUGA	8.05	5	Yes	2.88	23.66	15	70	5	43.66
CD7	AUGAUGA	12.73	CUGA	8.05	6	Yes	2.61	23.39	15	81	5	43.39
CD8	GUGAUGA	10.76	CUGA	8.05	4	No	-0.91	17.90	15	73	5	37.90
CD9	AUGAUGA	12.73	CUGA	8.05	2	No	-0.95	19.83	15	65	5	39.83
CD10	AUGAUGA	12.73	CUGA	8.05	6	Yes	2.42	23.20	15	65	5	43.20
CD11b	AUGAUGU	7.48	CUGA	8.05	7	Yes	3.88	19.41	15	60	5	39.41
CD11a	AUGAUGU	7.48	CUGA	8.05	2	No	-1.65	13.88	15	60	5	33.88
CD12	GUGAUGA	10.76	CUGA	8.05	5	Yes	2.36	21.17	15	59	5	41.17
CD13	AUGAUGA	12.73	CUGA	8.05	3	Yes	0.52	21.30	15	66	5	41.30
CD14	AUGAUGA	12.73	CUGA	8.05	5	Yes	3.93	24.71	15	71	5	44.71
CD15	CUGAUGA	8.44	CUGA	8.05	1	No	-1.90	14.59	15	97	5	34.59
CD16	AUGAUGA	12.73	CUGA	8.05	3	Yes	0.14	20.92	15	62	5	40.92
CD17	AUGAUGA	12.73	CUGA	8.05	2	No	-0.62	20.16	15	54	5	40.16
CD18	AUGAUGA	12.73	CUGA	8.05	8	Yes	3.94	24.72	15	54	5	44.72
CD19	GUGAUGA	10.76	CUGA	8.05	5	No	-1.48	17.33	15	57	5	37.33
CD20	GUGAUGA	10.76	AUGA	3.77	1	No	-2.41	12.12	15	65	5	32.12
CD21	GUGAUGA	10.76	CUGA	8.05	5	Yes	0.77	19.58	15	59	5	39.58
CD22	GUGAUGA	10.76	CUGA	8.05	5	No	-1.56	17.25	15	61	5	37.25
CD23	AUGAUGA	12.73	CUGA	8.05	4	Yes	0.90	21.68	15	61	5	41.68
CD24	AUGAUGA	12.73	CUGA	8.05	4	No	-0.93	19.85	15	59	5	39.85
CD25	AUGAUGA	12.73	CUGA	8.05	7	Yes	1.98	22.76	15	62	5	42.76
CD26	AUGAUGA	12.73	CUGA	8.05	5	Yes	1.14	21.92	15	62	5	41.92
CD27	AUGAUGA	12.73	AUGA	8.05	6	Yes	0.18	20.96	15	63	5	40.96
CD28	AUGAUGU	7.48	CUGA	8.05	7	Yes	4.08	19.61	15	65	5	39.61
CD29	AUGAUUG	3.27	CUGA	8.05	7	Yes	1.78	13.10	15	66	5	33.10
CD30	AUGAUUA	7.81	AUGA	3.77	2	No	-0.62	10.96	15	74	5	30.96
CD31	AUGAUGA	12.73	CUGA	8.05	2	No	-0.62	20.16	15	61	5	40.16
CD32	AUGAUGA	12.73	CUGA	8.05	7	Yes	0.93	21.71	15	55	5	41.71
CD33	AUGAUUA	7.81	AUGA	3.77	2	No	-0.62	10.96	15	74	5	30.96
CD34	AUGAUUA	7.81	AUGA	3.77	3	No	-1.69	9.89	15	61	5	29.89
CD35	AUGAUGA	12.73	CUGA	8.05	4	No	-0.32	20.46	15	75	5	40.46
CD36	AUGAUGA	12.73	CUGA	8.05	4	Yes	0.08	20.86	15	74	5	40.86
CD37	AUGAUUA	7.81	AUGA	3.77	3	Yes	0.33	21.05	15	66	5	31.91
CD38	AUGAUGA	12.73	CUGA	8.05	4	Yes	0.81	17.55	15	74	5	41.59
OR1	AUGAUGA	12.73	CUGA	8.05	1	No	-1.90	21.09	15	71	5	38.88
OR2	GUGAUGA	10.76	CUGA	8.05	6	Yes	2.28	18.88	15	61	5	41.09
OR3	AUGAUGA	12.73	CUGA	8.05	1	No	-1.90	18.88	15	71	5	38.88
OR4	AUGAUGA	12.73	CUGA	8.05	1	No	-1.90	13.30	15	72	5	38.88
OR5	AUGCUGA	8.15	CUGA	8.05	2	No	-1.66	11.91	15	71	5	34.54
OR6	AUGAUUA	7.81	AUGA	3.77	6	Yes	1.72	21.59	15	70	5	33.30
OR7	AUGAAGA	7.52	CUGA	8.05	4	Yes	1.98	16.56	15	53	5	37.55
OR8	AUGAGGA	8.23	CUGA	8.05	5	Yes	0.28	15.30	15	60	5	36.56
OR9	CUGAUGA	8.44	CUGA	8.05	1	No	-1.19	14.54	15	83	5	35.30

[a] Sequences with predicted methylation sites in rRNA are named CDx. and those without ORx for ORphan (x: natural number)
[b] TS: terminal stem
[c] snoScan Score = C score + D score + TS score
[d] Classifier Score (CS) = snoScan Score + RNAseq expression + Gap score

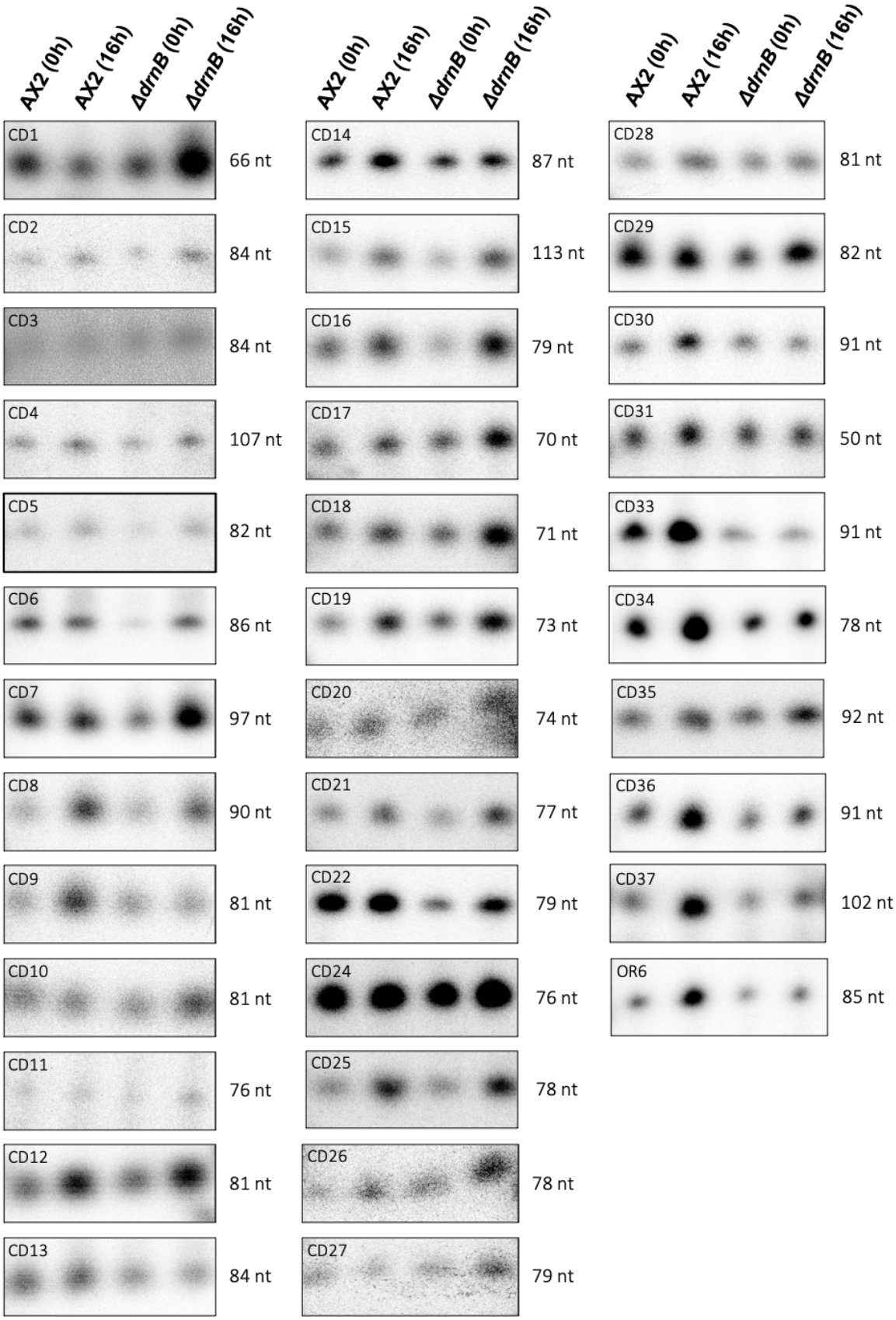


Figure 33. Size estimation of box C/D snoRNAs in *D. discoideum* using primer extension. Shown are the primer extension products for the indicated box C/D snoRNAs in the Ax2 and *ΔdrnB* strains in axenic growth and in the slug stage of development. Inferred RNA sizes are shown to the right.

The box C/D snoRNA genes in *D. discoideum*

Usually, box C/D snoRNAs are encoded in intergenic regions or as part of introns in protein-coding genes, and in either set-up, they can be generated as mono- or poly-cistronic transcriptional units [343]. Aspegren et al. [346] predicted four bi-cistronic transcriptional units of snoRNAs in *D. discoideum* and confirmed expression for several of them using RT-PCR. An analysis of the genomic location of the genes for our set of 47 box C/D snoRNAs revealed five additional clusters containing two box C/D snoRNAs and two clusters comprised of three box C/D snoRNAs (Figure 34). The genes for these box C/D snoRNAs appear equally spaced in the clusters. All box C/D snoRNA genes, in clusters or not, were found in intergenic regions, except CD38, which is encoded in an intron (Table S3). The CD RNAs are encoded on all chromosomes without a noticeable pattern, but we observed that the majority of OR RNAs are encoded on chromosome 4. The biological significance of this, if any, remains to be elucidated, and we cannot exclude that it is a random localization. Next, we set out next to investigate the 2′-O-Me patterns in *D. discoideum*'s rRNAs, that would be guided by the encoded box C/D snoRNAs.

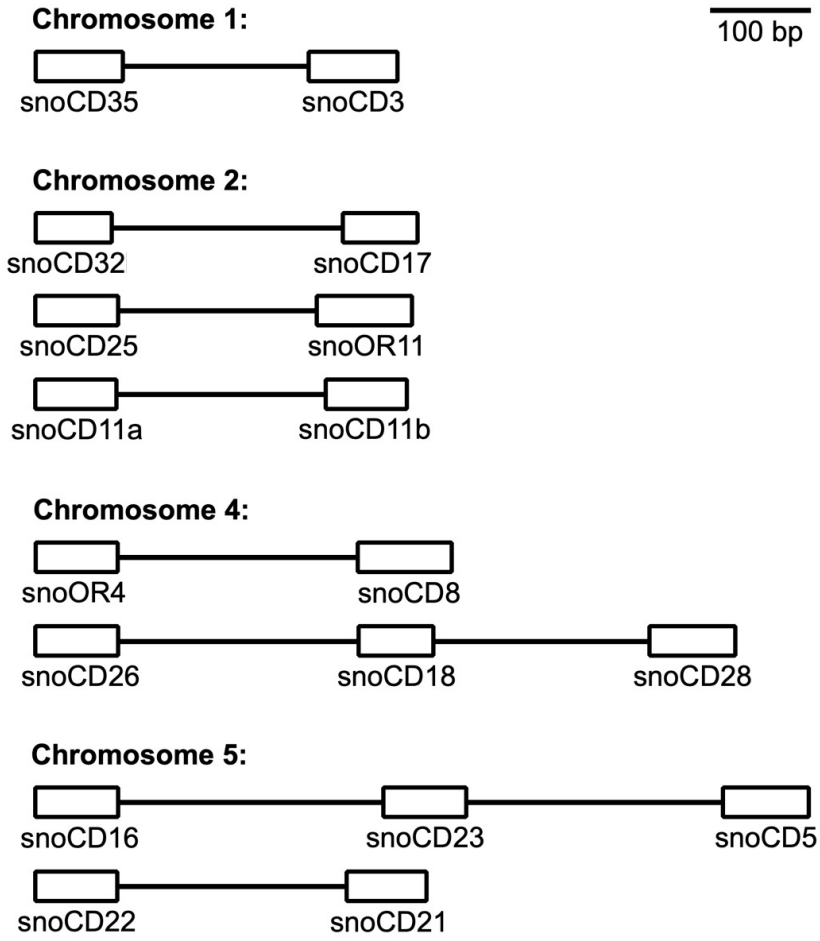


Figure 34. Novel genomic clusters of box C/D snoRNAs in *D. discoideum*. A scale is indicated on the upper right.

D. discoideum 17S and 26S rRNAs have 49 high-confidence 2′-O-Me sites

To address 2′-O-Me in the 17S and 26S rRNA of *D. discoideum*, we employed RMS, a method introduced on yeast rRNA [250], and subsequently used in a several other organisms [251, 252, 480]. In brief, RMS is a next-gen sequencing-based method that relies on the cleavage-resistance of 2′-O-methylated nucleotides under alkaline conditions, resulting in an underrepresentation of read ends in fragmented RNA. The results are expressed as RMS scores, which represent the fraction of modified molecules at a given position. The method yields methylation stoichiometry comparable to RP-HPLC [481]. We generally considered sites with an RMS score >0.75 as high-confidence 2′-O-Me sites.

To investigate the global 2′-O-Me landscape in wild type *Dictyostelium*, we initially determined the RMS scores of rRNA isolated from axenic AX2 cells. During these experiments, we realized that one nucleotide (C784) was missing in the 17S reference sequence [15], and its presence was independently confirmed by sequencing of a PCR product on total DNA. Using the criteria outlined above, we determined in total 17 and 32 positions with a 2′-O-Me moiety on the 17S rRNA and the 26S rRNA, respectively (Figure 35A). Of these high-confidence sites,

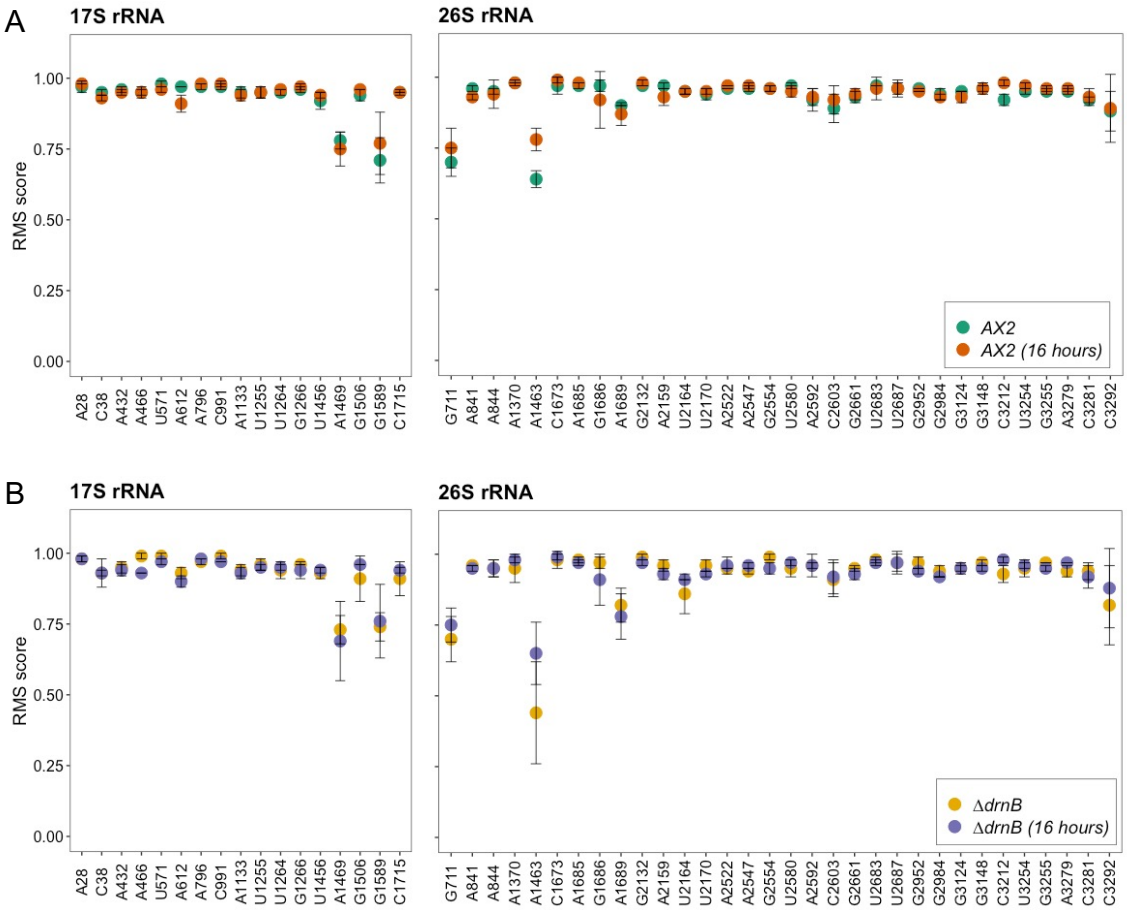


Figure 35. RiboMeth-seq analysis of the 17S and 26S in *D. discoideum*. RMS scores at 2′-O-Me sites on the 17S and 26S rRNA in axenic growth and development of AX2 (A) and *ΔdrnB* (B) cells (n=3).

the majority appeared to be fully methylated. In axenically-grown AX2 cells, we identified 2 hypomethylated positions each in the 17S and 26S rRNAs. This indicates, to our knowledge for the first time, heterogeneity of the ribosome population in *D. discoideum*. Heterogeneity in rRNA modifications had been, however, reported previously for mouse, human, thale cress, and zebrafish [251, 252, 261, 480, 482, 483]. In these studies, differences in the ribosome 2′-O-Me patterns between cultured cells and differentiated tissues, or during development have been described. Since *D. discoideum* undergoes development upon starvation, we set out next to elucidate any changes of the 2′-O-Me pattern in rRNAs of the slug stage of development in the AX2 wild type. The fractionally methylated positions in axenically-grown wild type cells were also substoichiometrically methylated during development, while the RMS score of most 2′-O-Me sites remained unchanged (Figure 35A).

For *Dictyostelium*, box C/D snoRNA gene clusters have been described [346]. Primary transcripts of such clusters are often processed by an RNase III before exonucleolytic processing can occur [352, 353, 358]. We therefore included the knock-out strain of the nucleolar RNase III DrnB [405, 484, 485] in our experiments, reasoning that box C/D snoRNA availability might be altered in that strain. At large, the 2′-O-Me pattern of the AX2 strain, however, was also observed for axenic growth and development of the Δ*drnB* strain (Figure 35B and Figure36). Only one position, 26S-Am1463, exhibited a noticeable difference between

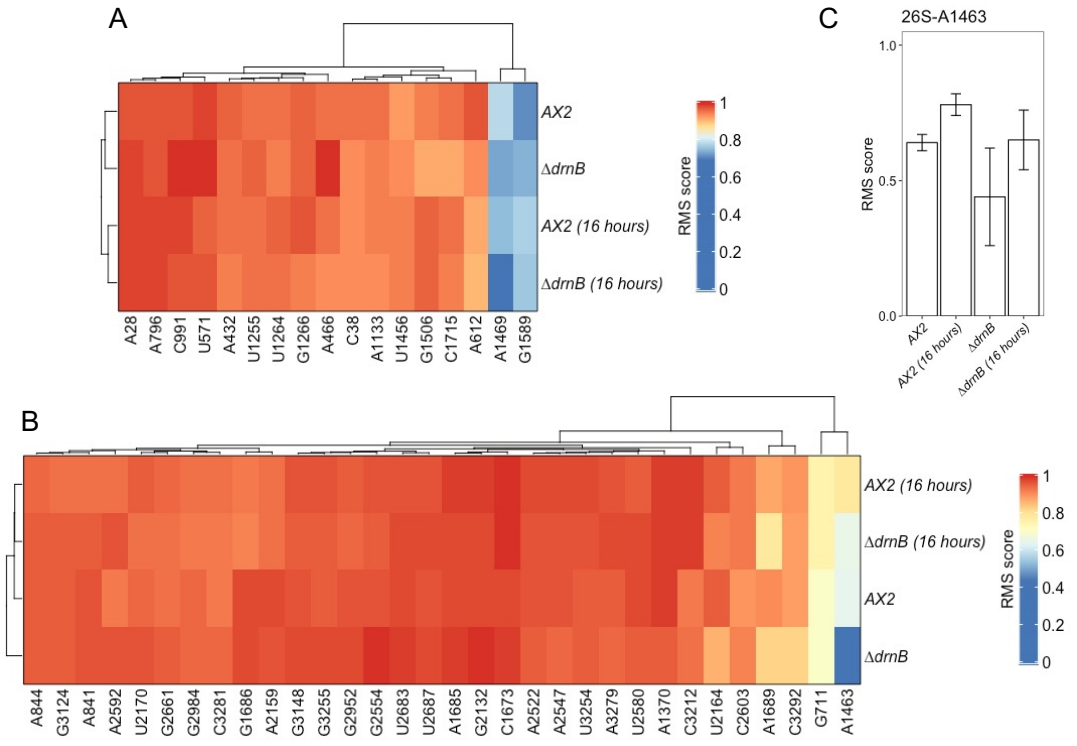


Figure 36. Comparison of RMS scores in axenic growth and the development of AX2 and *ΔdrnB*. Heatmap of RMS scores at all methylated positions on the 17S (A) and 26S rRNA (B).

Table 16. Sites of 2´-O-Me in rRNA with guiding box C/D snoRNAs in *D. discoideum* and further species.

D. discoideum[a]		*S. cerevisiae*[b]		*H. sapiens*[c]		*A. thaliana*[d]	
Position	**guided by**	**Position**	**guided by**	**Position**	**guided by**	**Position**	**guided by**
				SSU[e]			
Am28	**CD18**	Am28	snR74	Am27	U27	Am28	AtU27
Cm38	**CD35**	-		-		-	
Am432	**CD8**	Am436	snR87	Am484	U16	Am438	AtU16
Am466	**CD8**	-		-		-	
Um571	**CD21**	Um578	snR77	Um627	HBII-135	Um580	AtsnoR77Y
Am612	**CD25**	Am619	snR47	Am668	U36A/B	Am621	AtU36
Am796	**CD19**	-		-		-	
Cm991	**CD7**	-		-		-	
Am1133	**CD10**	-		-		-	
Um1255	**CD20**	-		-		-	
Um1264	**CD29**	Um1269	snR55	Um1326	U33	Um1270	AtsnoR34
Gm1266	**CD37**	Gm1271	snR40	Gm1328	U232A	Gm1272	AtsnoR21
Um1456	**CD1**	-		Um1442	U61	Um1281	AtU61
Am1469		-		-		-	
Gm1506	**CD1**	Gm1428	snR56	Gm1490	U25	Gm1431	AtsnoR19
Gm1588	**CD16**	-		-		-	
Cm1715	**CD28**	Cm1639	snR70	Cm1703	U43	Cm1641	AtU43
				LSU[e]			
Gm711	**CD7**	-		-		-	
Am841	**CD12**	-		-		-	
Am844	**CD24**	Am649	U18	Am1313	U18A/B/C	Am647	AtU18
Am1370	**CD9/13**	Am1133	snR61	Am1858	U38A/B	Am1140	AtU38
Am1463	**CD9/13**	-		-		-	
Cm1673	**CD19**	Cm1437	Um24	Cm2338	U24	Cm1439	AtU24
Am1685	**CD19**	Am1449	Um24	Am2350	U76	Am1451	AtU24
Gm1686	**CD16**	Gm1450	Um24	Gm2351	U24	Gm1452	
Am1689	**CD33**	-		-		-	
Gm2132	**CD1**	-		-		-	
Am2159	**CD27**	-		-		-	
Um2164	**CD31**	-		-		-	
Um2170	**CD17/32**	Um1888	snR62	Um2824	U34	Um1882	AtU34
Am2522	**CD4**	Am2256	snR63	Am3739	U46	-	
Am2547		Am2281	snR13	Am3764	U15A/B	Am2271	AtU15
Gm2554	**CD14**	Gm2288	snR75	-		Gm2278	AtU15
Um2580	**CD12/34**	-		-		-	
Am2592	**CD34**	-		-		-	
Cm2603	**CD26**	Cm2337	snR64	Cm3820	U74	-	
Gm2661	**CD2/3**	-		-		-	
Um2683	**CD30**	Um2417	snR66	-		-	
Um2687	**CD22**	Um2421	snR78	Um3904	U52	Um2411	AtsnoR37
Gm2952	**CD13**	Gm2619	snR67	Gm4166	U31	Gm2610	AtsnoR35
Gm2984	**CD5**	-		-		-	
Gm3124	**CD36**	-		-		-	
Gm3148	**CD6**	Gm2815	snR38	Gm4362	snR38A/B/C	Gm2805	AtsnoR38Y
Cm3212	**CD38**	-		-		-	
Um3254	**CD25**	Um2921	snR52	Um4468		-	
Gm3255		Gm2922	Spb1	Gm4469		-	
Am3279	**CD11**	Am2946	snR71	Am4493	U29	Am2936	AtU29
Cm3281	**CD15**	Cm2948	snR69	-		-	
Cm3292	**CD23**	Cm2959	snR73	Cm4506	U35A/B	Cm2949	AtU35

[a] This study
[b] https://people.biochem.umass.edu/fournierlab/snornadb/mastertable.php
[c] https://www-snorna.biotoul.fr/human_yeast/
[d] https://ics.hutton.ac.uk/cgi-bin/plant_snorna/home
[e] SSU: small subunit; LSU: large subunit.

the axenically-grown AX2 and Δ*drnB* strains (Figure 36C). This indicates that any effect that DrnB might have on the processing of box C/D snoRNA precursors does not manifest

substantially in altered 2′-O-Me patterns. Similarly, position 26S-Am1463 displayed different RMS scores were observed between axenic growth and the slug stage in both, the AX2 and Δ*drnB* strains (Figure 35C, Figure 36). The four 2′-O-methylated residues that we found either fractionally modified or changed in development had no orthologous modified sites in *S. cerevisiae*, *H. sapiens*, and *A. thaliana* (Table 16, and see below).

Secondary structure models for the small and large ribosomal subunits in Amoebae

As methylated rRNA positions are required for folding and structural stabilization of rRNAs, thereby contributing to ribosome function [247], it was of interest to localize the 2′-O-methylated positions in the context of the rRNA structure of *D. discoideum*. A partial structure of the large ribosomal subunit of *D. discoideum* has been published recently [486], but no high-resolution structural data is available for complete ribosomes from any species of the Amoebozoa. To obtain a model for the rRNA secondary structures, we therefore employed homology modelling using sequences of species from the evolutionary supergroups of Opisthokonta and Archaeplastida [2]. In brief, we aligned the rRNAs from the amoeba with the corresponding small and large subunits' (SSU and LSU, respectively) rRNA sequences from *A. thaliana*, *Caenorhabditis elegans*, *Drosophila melanogaster*, and *H. sapiens* (Table 16). The inferred secondary structure models of the 17S and 26S (with the 5.8S) rRNAs of *D. discoideum* are shown in Figure 37 and Figure 38, respectively, and include the 2′-O-methylated positions. Central parts of ribosomes from different species are structurally highly conserved and variation appears restricted to peripheral regions and the so-called expansion segments (ES) [487], which often harbor species-specific sequences. This is exactly what the models for the amoebal rRNA structures display (Figure 37 and Figure 38). This holds particularly true for the conserved regions involved in the formation of A, P and E sites. Not surprisingly, the ES of *D. discoideum*, which are not covered in the aforementioned structure [486], exhibited significant differences as compared to the ES in other species (exemplified for *H. sapiens*; Table 17).

About half of the 2′-O methylated positions were found in the vicinity of nucleotides residing in the A, P and E sites, and the other half in other regions of the rRNAs (Figure 37 and Figure 38). These latter positions localized frequently to formally single stranded regions, or to nucleotides at the very beginning of helical stems. When comparing the 2′-O-Me patterns in wild type *D. discoideum* to those in *S. cerevisiae*, *H. sapiens*, and *A. thaliana*, we found 28 of the 2′-O-Me sites conserved in at least one of these organisms, and therefore, the other 21 sites are specific to *D. discoideum* (Table 16). Only one of these positions, Gm711 in the 26S rRNA, was found in an ES (Figure 38), indicating that 2′-O-Me is largely restricted to the core of the

Table 17. Main differences of ribosomal expansion segments in *D. discoideum* compared to *H. sapiens*.

Subunit	Expansion segment	Difference
SSU	h9ES3	h9ES3a completely missing; h9ES3b conserved in structure; h9ES3c prolonged
	h21ES6	h21ES6a conserved. contains the *D. discoideum*-specific methylation site Am796; h21ES6b extended; h21ES6c missing; h21ES6d conserved
	h26ES	completely missing
	h39ES9	almost 70 nt longer
LSU	H9ES3	constituted by the loop of H9. 6 nt smaller
	H25ES7	H25ES7a: ca. 200 nt shorter; H25ES7b: ca. 160 nt shorter; H25ES7c: 6 bp shorter; H25ES7d: completely missing; H25ES7e: ca. 40 nt shorter; H25ES7f: ca. 70 nt shorter; H25ES7g: completely missing; H25ES7h: similar length. different structure; three additional helices (now called H25ES7i-k)
	H28ES8	stem loop 3 nt shorter
	H30ES9	similar length. different structure
	H31ES9	stem loop. 4 nt longer
	H38ES10	completely absent
	H38ES12	H38ES12a corresponds to H38ES12 in human LSU. H38ES12a 2 bp shorter with additional bulge near helix terminus; Additional helix (now called H38ES12b).
	H45ES15	completely absent
	H52ES19	H52ES19a corresponds to H52ES19 in human LSU. H52ES19a is similar length. different structure; Additional helix (now called H52ES19b).
	H54ES20	H54ES20a 5 bp shorter in Dicty; H54ES20b completely missing
	H63ES27	Overall almost 570 nt shorter: H63ES27a is 280 nt shorter. H63ES27b is almost completely missing.
	H78ES30	completely absent.
	H79ES31	Size is similar. but less structured with H79ES31a being structurally conserved and H79ES31b not formed
	H98ES39	Overall 110 nt shorter and less structured with H98ES39a being 95 nt shorter and H98ES39b being 15 nt shorter

ribosome in *D. discoideum*. Noteworthy, five of the 13 specific 2´-O-Me sites on the 26S rRNA were locating in domain 0, which has been shown in other species to coordinate folding of all other domains of the LSU rRNA, including the PTC [488].

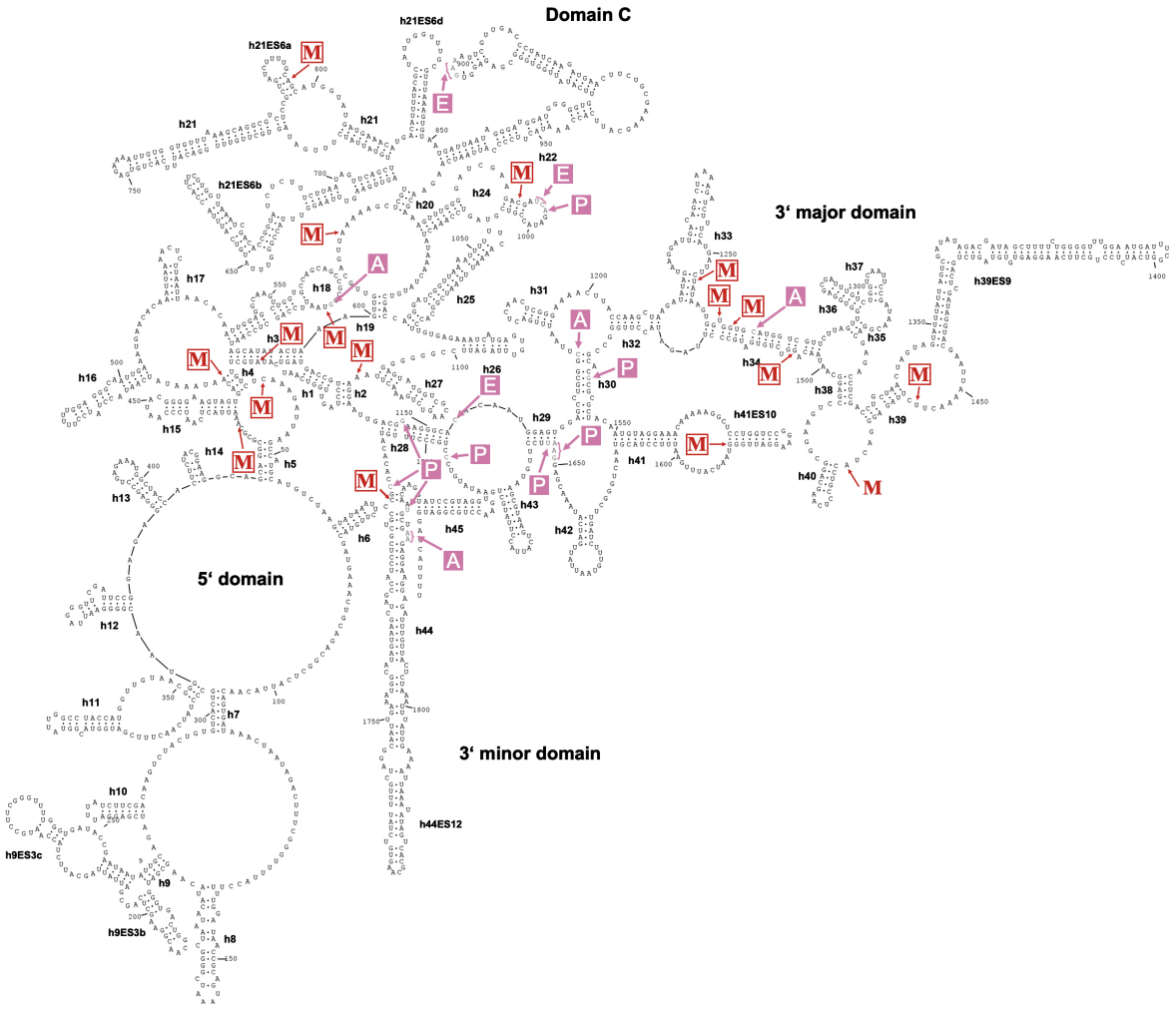


Figure 37. Secondary structure of the 17S rRNA of *D. discoideum* with 2´-O-Me sites. The secondary structure of the 17S rRNA was inferred by homology and drawn using RNAviz (v. 2.0.3). The 2´-O-methylated nucleotides as identified by RiboMeth-seq are marked with an arrow and ‘m’ (red). Nucleotides located in the A, P, and E sites of the ribosome are indicated in pink. Helices (hx) are named to convention and expansion segments (ESx) are labeled with x: natural number.

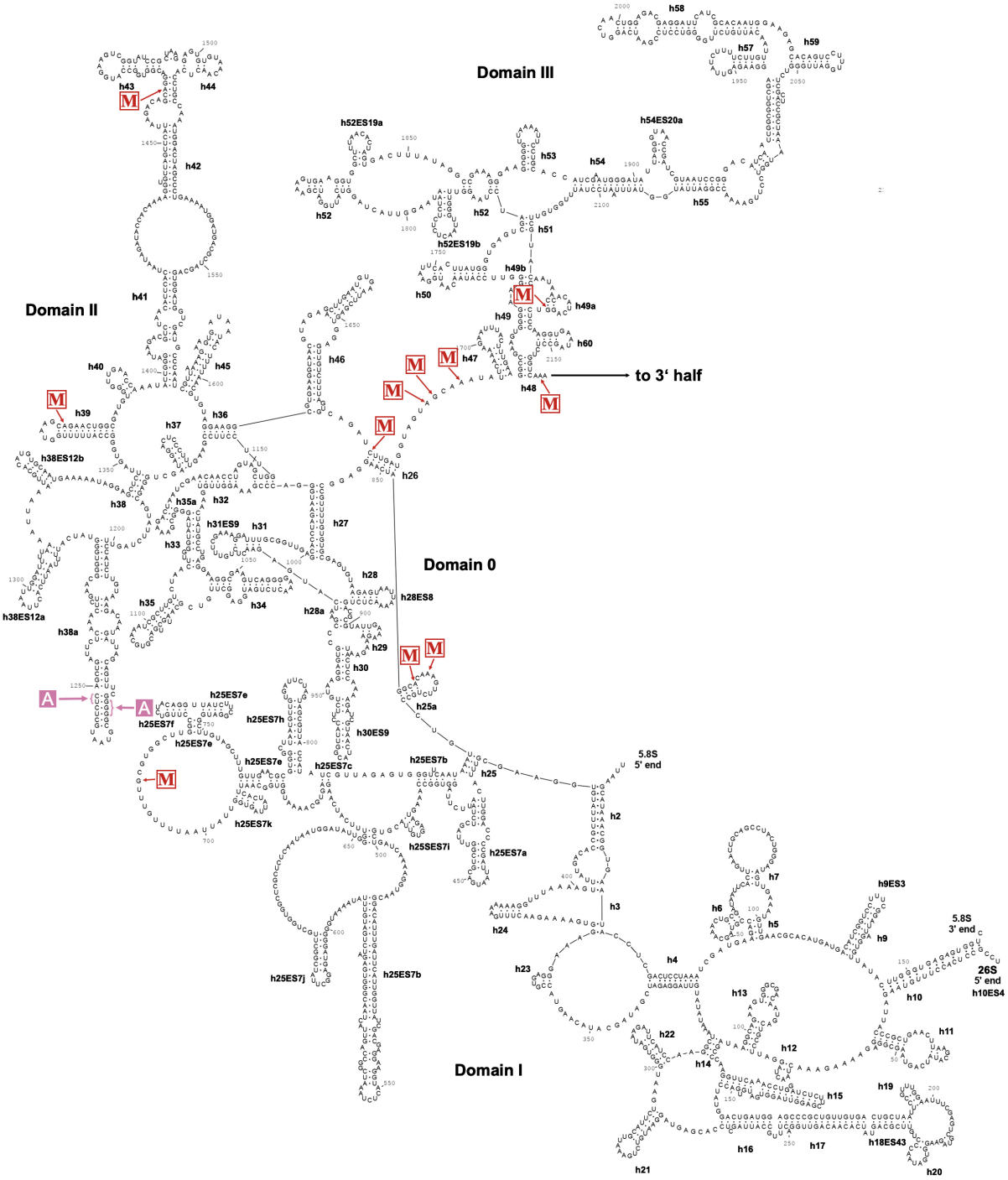


Figure 38. Secondary structure of the 26S rRNA of *D. discoideum* with 2´-O-Me sites. The secondary structure of the 26S rRNA was inferred by homology and drawn using RNAviz (v. 2.0.3). The 2´-O-methylated nucleotides as identified by RiboMeth-seq are marked with an arrow and 'm' (red). Nucleotides located in the A, P, and E sites of the ribosome are indicated in pink. Due to the size of the 26S rRNA, the figure is split into the 5´ half (this) and 3´ half (next page). The predicted interaction with the 5.8S rRNA (previous page) is shown at the 5´end. Helices (Hx) are named to convention and expansion segments (ESx) are labeled.

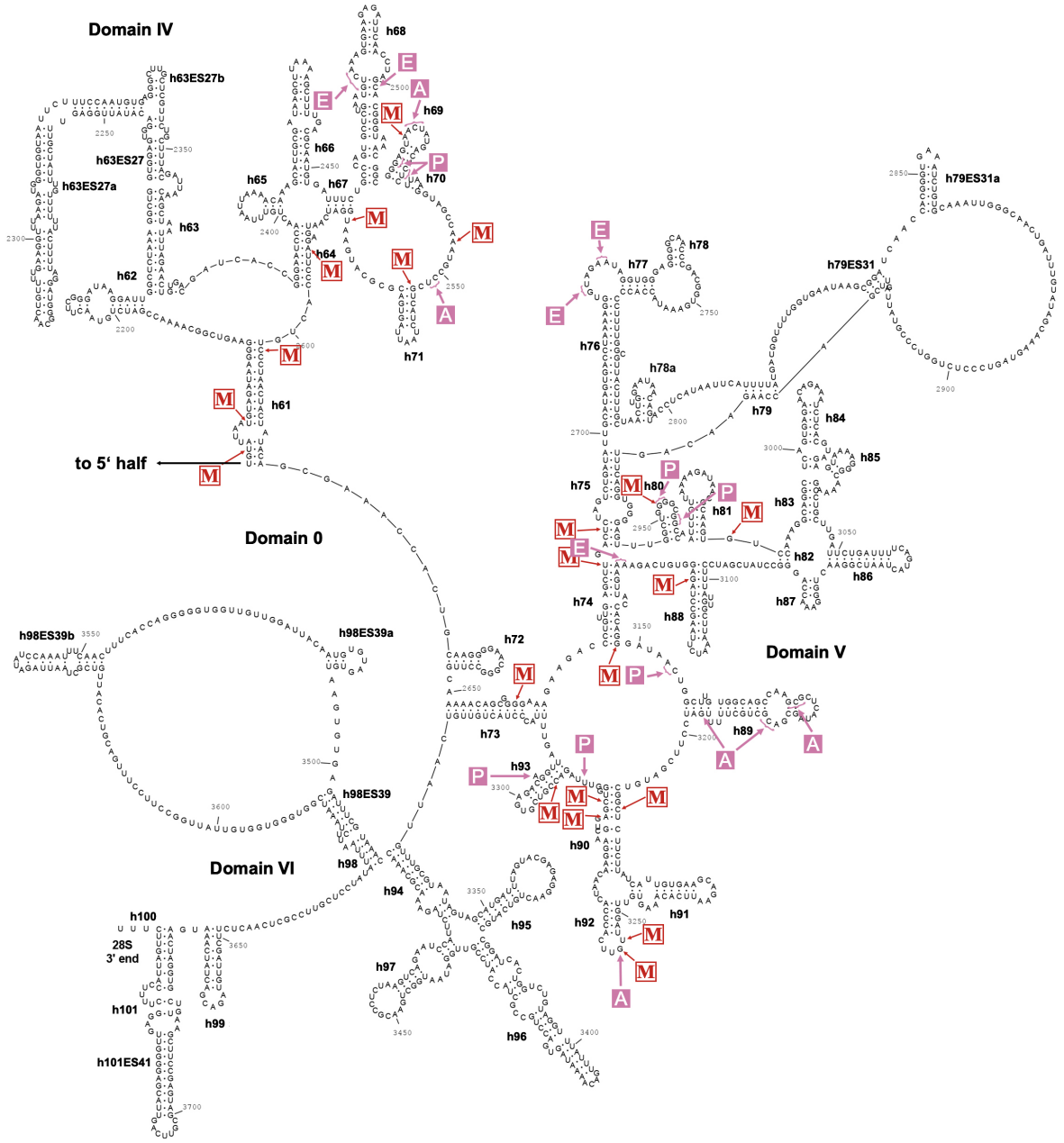


Figure 38. Secondary structure of the 26S rRNA of D. discoideum with 2´-O-Me sites. ctd.

The majority of 2′-O-Me sites in *D. discoideum* can be associated to box C/D snoRNAs

To identify snoRNA guides for the 2′-O-methylated sites, we employed next RNAhybrid, since snoScan alone was not able to predict all targets for our set of box C/D snoRNAs. This resulted in the prediction of 46/49 2′-O-Me sites with at least one, occasionally two box C/D snoRNA guides (Table 16). The snoRNAs guiding 2′-O methylation at these rRNA sites were named CDx (x = natural numbers; Table S3). For the remaining 9 box C/D snoRNAs, we could not assign a 2'-O-Me site in either rRNA, and therefore we classified these sequences as orphans, and named them accordingly ORx (Table S3). Seven of the CD RNAs can make use of both their D and D′ boxes to guide 2′-O-Me in one or both rRNAs (Table 16 and Table 18). For most positions targeted by these CD RNAs, no alternative guides were found. Rather, CD1 and CD19 have two targets each for their D′ boxes, additional to the targets of their D boxes (Table 18). The majority of CD RNAs, however, is predicted to employ either its D or D′ box. Figure 41A displays examples for single and double usage of D boxes, shown exemplarily for one case each in the 17S and 26S rRNA. The predicted bimolecular interactions of the CD RNAs with their rRNA targets are shown in Figure 39 and Figure 40 for D and D′ box guides, respectively.

Table 18. Targets of CD RNAs which utilize box D and box D′.

CD RNA	box D target	Domain	box D′ target	Domain	Distance [nt]
CD1	17S-U1456	3′ major	17S-G1506	3′ major	50
CD1	17S-U1456	3′ major	26S-A2132	III	-
CD7	26S-G711	I	17S-C991	C	-
CD8	17S-A466	5′	17S-A432	5′	34
CD13	26S-A1370	II	26S-G2952	V	1581
CD15	26S-C3292	V	26S-C3281	V	11
CD19	26S-C1673	II	26S-C1685	II	12
CD19	26S-C1673	II	17S-A796	C	-
CD25	17S-A612	5′	26S-U3254	V	-

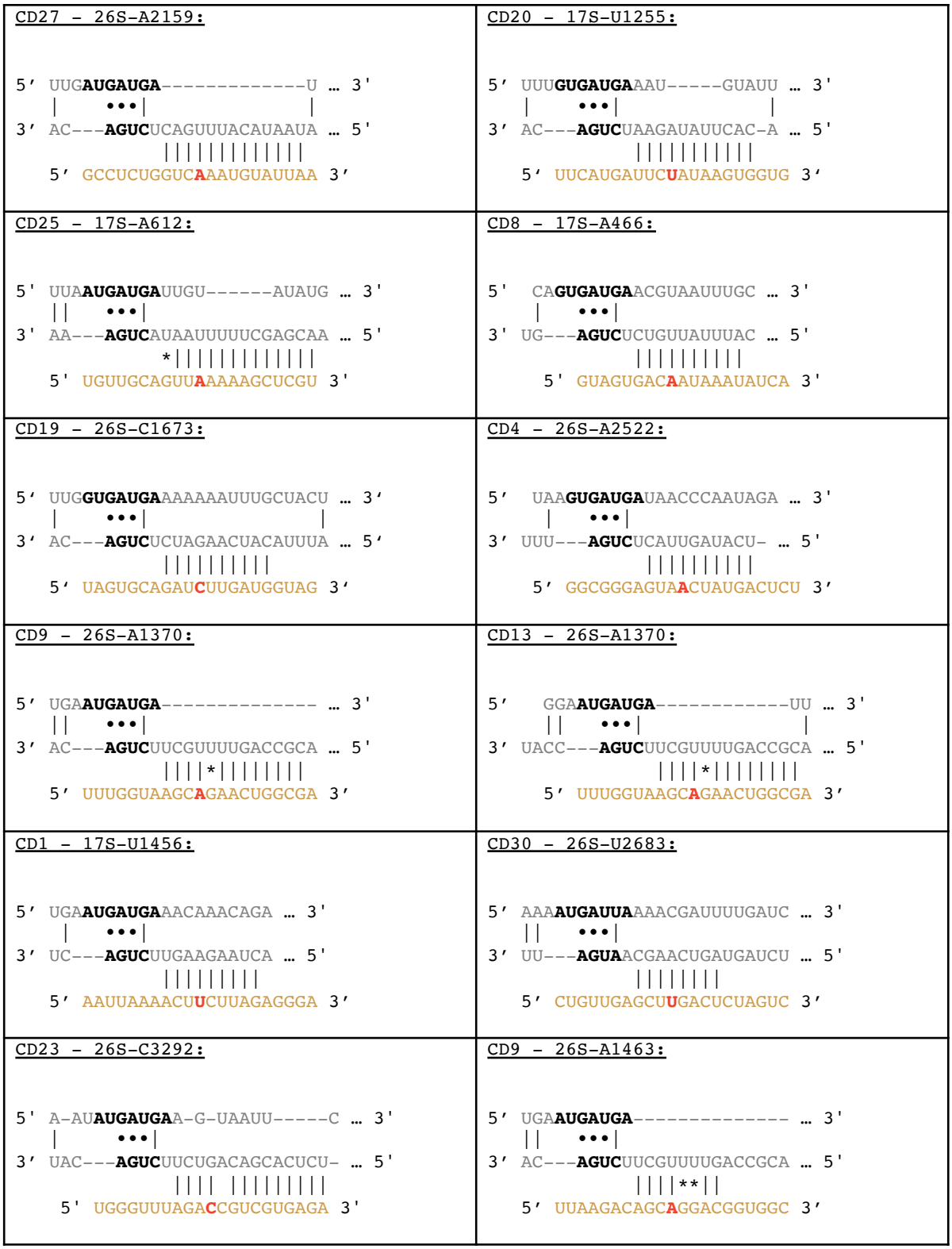


Figure 39. Predicted base pairing between rRNA and CD RNAs utilizing the D box motif. Shown are part of the rRNA (beige) with the methylated residue (red), and base pairs with relevant sequences of the guiding snoRNAs (grey) and their predicted C and D boxes (bold). Watson-Crick (|) and wobble base pairing (*) between rRNA and snoRNA are indicated. The predicted k-turns formed via non-Watson-Crick base pairing between C/D boxes are designated by (•). Except for the CD29 - 17S-U1264 pair, all k-turns follow the consensus. Pairs are sorted by decreasing stability of the duplex.

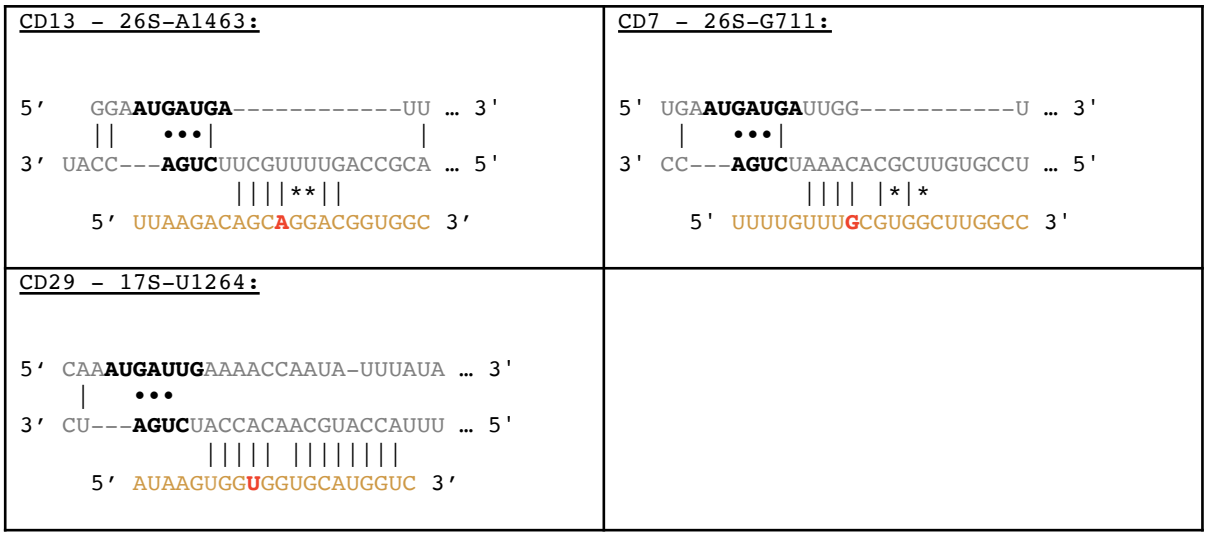


Figure 39. Predicted base pairing between rRNA and CD RNAs utilizing the D box motif. ctd.

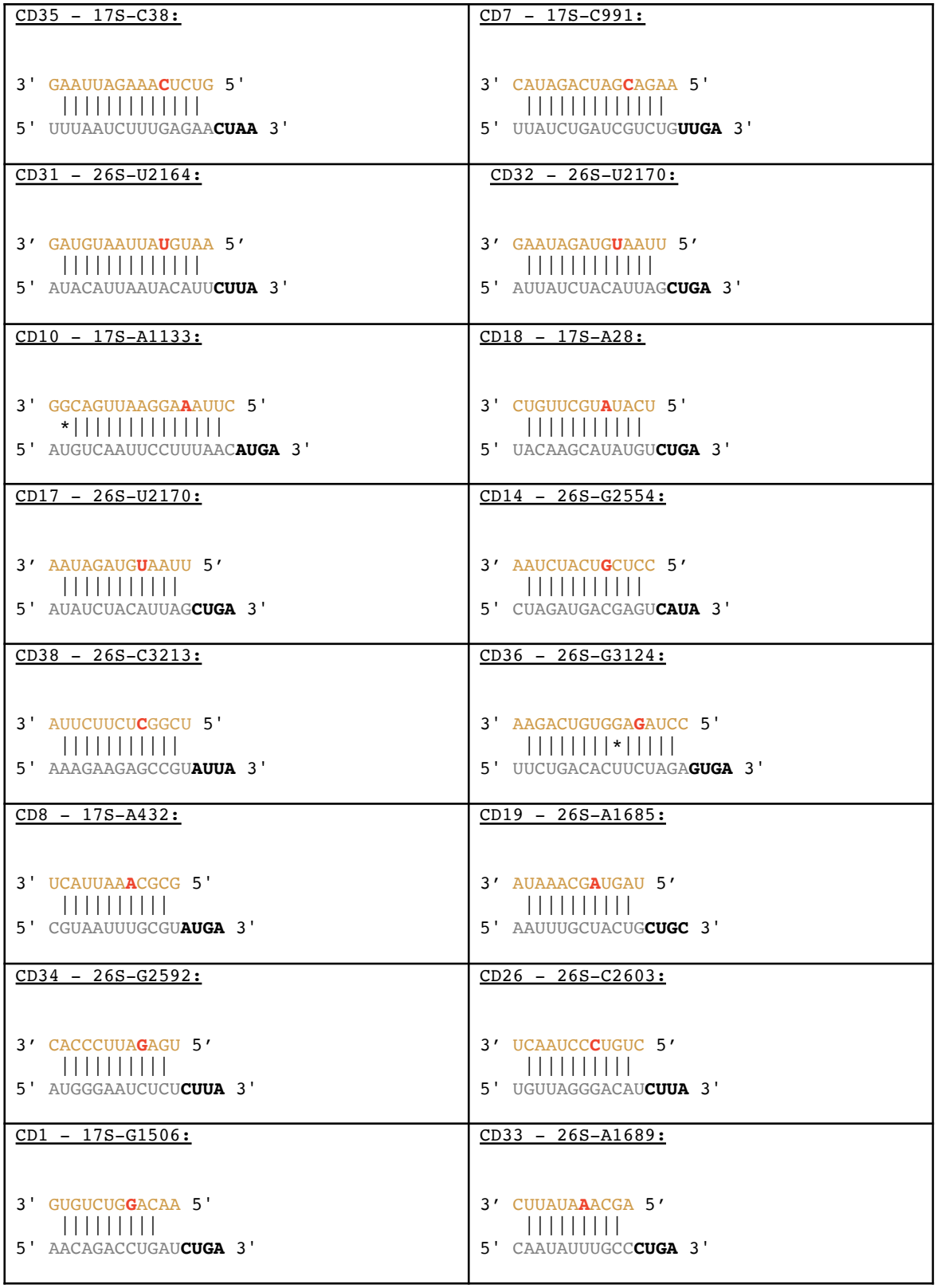


Figure 40. Predicted base pairing between rRNA and CD RNAs utilizing the D´ box motif. Shown are part of the rRNA (beige) with the methylated residue (red), and base pairs with relevant sequences of the guiding snoRNAs (grey) and their predicted D box (bold). Watson-Crick (|) and wobble base pairing (*) between rRNA and snoRNA are indicated. As the k-loop is less well conserved, the sequences of the C´ boxes are not shown. Pairs are sorted by decreasing stability of the duplex

Figure 40. Predicted base pairing between rRNA and CD RNAs utilizing the D´ box motif. ctd.

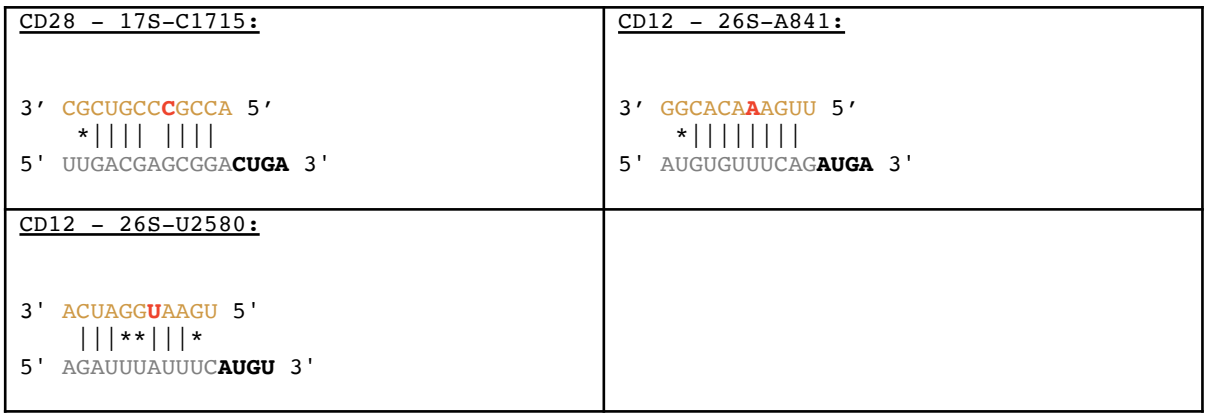


Figure 40. Predicted base pairing between rRNA and CD RNAs utilizing the D´ box motif. ctd.

Features of box C/D snoRNAs and their interactions with rRNA

The box C/D snoRNAs in *Dictyostelium* are between 66 nt and 113 nt in length, with an average GC content of 32.2% and box C-D distances between 50 nt and 97 nt (Table 16). The terminal stem often found in box C/D snoRNAs in other species (Figure 7), is predicted by snoScan only in 25 of the 47 box C/D snoRNAs of *D. discoideum* (indicated with a positive TS score in Table 16). In contrast, the box C and box D sequences forming the k-turn motif are highly conserved (Figure 41B); in particular, the GA dinucleotides forming *trans* hoogsteen/sugar-edge A•G base pairs are present in all CD RNAs selected by the described criteria (but not in all OR RNAs, see Table 16). Furthermore, we found that almost all CD RNAs abide to the box D consensus sequence CUGA, with a small fraction of snoRNAs featuring an AUGA instead (Figure 41B and Table 16). Compared to these motifs, the box C´ and box D´ sequences show considerably more variation in *Dictyostelium* (Figure 41B). Despite this, the majority of methylated positions is predicted to be guided by the D´ boxes of individual CD RNAs (Figure 41C), similar to observations made for the human box C/D snoRNAs [251]. The lengths of the CD RNA/rRNA duplexes distributed around 11 bp within a range of 7 - 15 bp, with average minimal free energies (MFE) of -13.9 kcal/mol (Figure 41D, E). In these predicted CD RNA/rRNA interactions, we observe the frequent occurrence of G*U base pairs [489], occasionally A/C base pairs [321], and a single G/A mismatch (Figure 39 and Figure 40). Only for the CD16/17S-G1589 duplex, we noticed that the +6 position is methylated, rather than the consensus +5 position, as had also been observed before in other species, e.g. *D. rerio* [252].

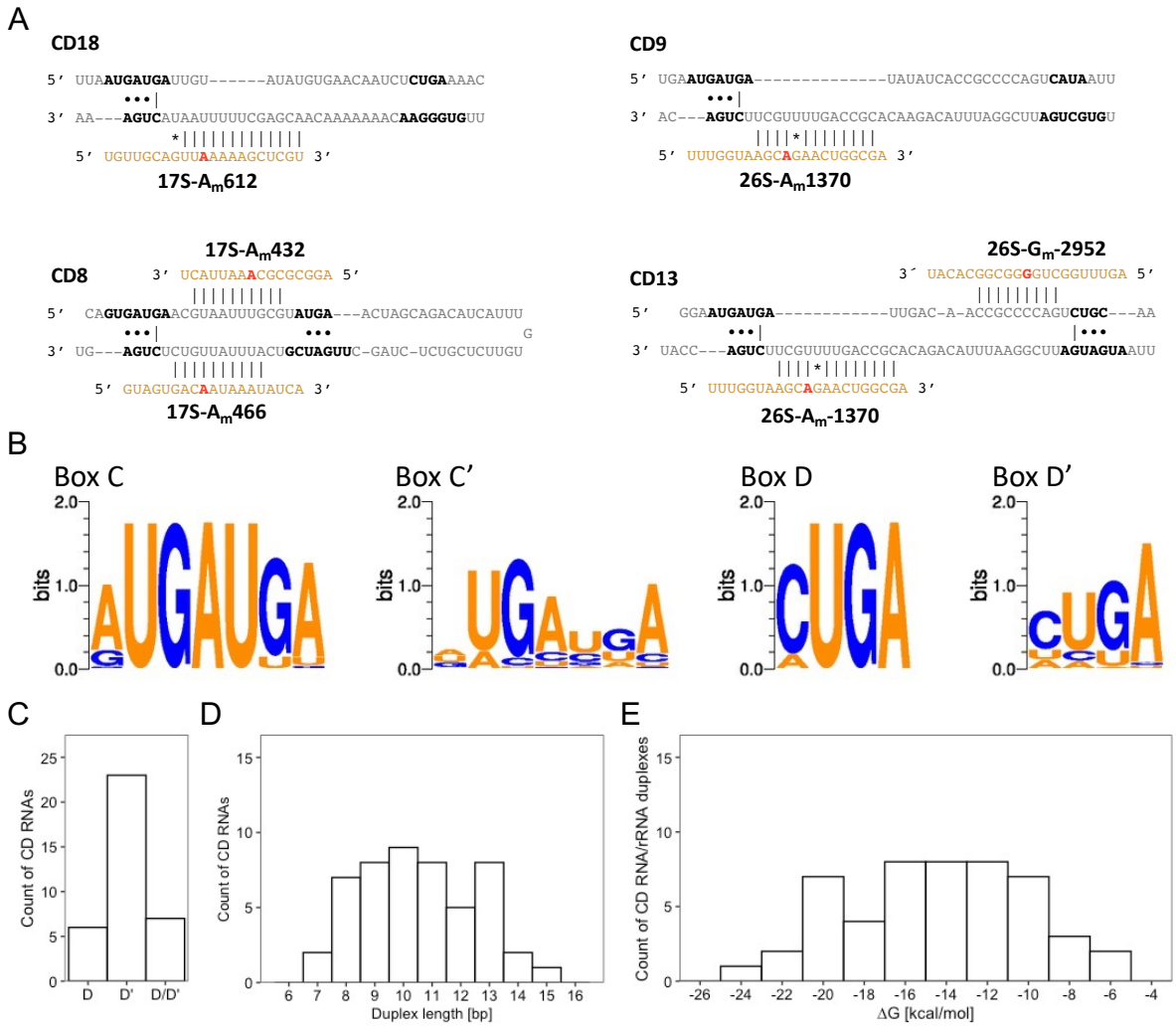


Figure 41. Features of CD RNAs in *D. discoideum*. (A) Examples of CD RNAs guiding 2´-O-Me at one or two rRNA positions. Single (top) and double (bottom) usage of D boxes of selected CD RNAs guiding positions in the 17S (left) and the 26S rRNA (right). Shown are CD RNA sequences (grey) with nucleotides involved in the formation of the k-turn (black). The guided part of the rRNA is shown in orange with the methylated residue highlighted in red. Intra- and intermolecular interactions are denoted for Watson-Crick (|) and G/U base pairing (*), as are the A/G and U/U base pairs (•) involved in the formation of the k-turn. (B) Conservation of C, C´, D and D´ box sequences shown with WebLogo [457]. (C) Distribution of CD RNAs using box D, D´ or both. Duplex lengths (in bp; D) and minimal free energies ΔG (in kcal/mol; E) of the interaction between CD RNA and the guided rRNA position.

Box C/D snoRNAs accumulate differentially during development of *D. discoideum*

Our primer extension experiments (Figure 33) indicated no 5´-end size heterogeneity of box C/D snoRNAs in *D. discoideum*. In absence of an internal control, a correlation between band intensity and expression levels is difficult. Furthermore, we could not obtain a product for several snoRNAs, despite the use of several distinct primers in these experiments. Therefore, to obtain a more complete view on box C/D snoRNA accumulation, we retrieved RNA-seq

datasets for AX2 and Δ*drnB* in axenic growth and in the slug stage of development from NCBI, which were originally deposited by Liao *et al.* [462]. As a first step, we performed a principal component analysis (PCA) of box C/D snoRNA expression on two biological replicates for each time point per strain. The analysis revealed global changes of box C/D snoRNA abundance in the development of the AX2 and Δ*drnB* strains (Figure 42A), however, not between AX2 and Δ*drnB*. This is corroborated by comparative 2D plots of DESeq2-normalized reads of individual box C/D snoRNAs in the two strains and under the two growth conditions (Figure 43A). In a subsequent analysis of individual box C/D snoRNAs, we considered changes significant if an adjusted p-value < 0.05 and an at least 0.5-log$_2$fold-change in RNA quantity was observed. Using these criteria, 22 box C/D snoRNAs were significantly up- or downregulated in the slug stage of development of AX2 (Figure 42B, C and Figure 43B). In contrast to this and as seen before (Figure 42A), we did not observe significant differences in the box C/D snoRNA between AX2 and Δ*drnB* except for OR9 and CD37, which were upregulated in the slug stage in Δ*drnB*, but not in AX2 (Figure 42B, C and Figure 43B). We wondered whether the changes that we observe in the 2′-O-Me patterns (Figure 35) can be explained by differences in the accumulation of the guiding CD RNAs. This is clearly not the case, as a 2D plot of the DESeq2-normalized reads of CD RNAs against the RMS scores at all

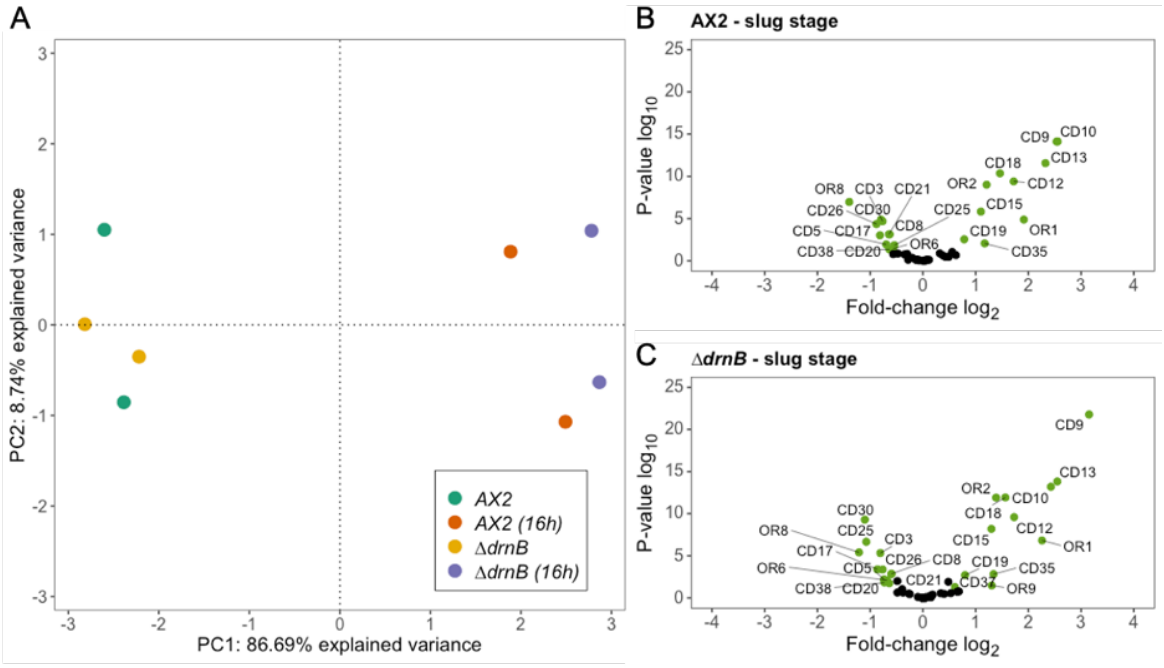


Figure 42. Analysis of box C/D snoRNA accumulation in axenic growth and development of the AX2 and *ΔdrnB* strains. (A) Principal component analysis (PCA) of data from RNA-seq on the indicated strains and conditions. Volcano plots of box C/D snoRNA abundance changes in the slug stage of AX2 (C) and *ΔdrnB* (C). Significantly up- or downregulated box C/D snoRNAs are labelled and colored green.

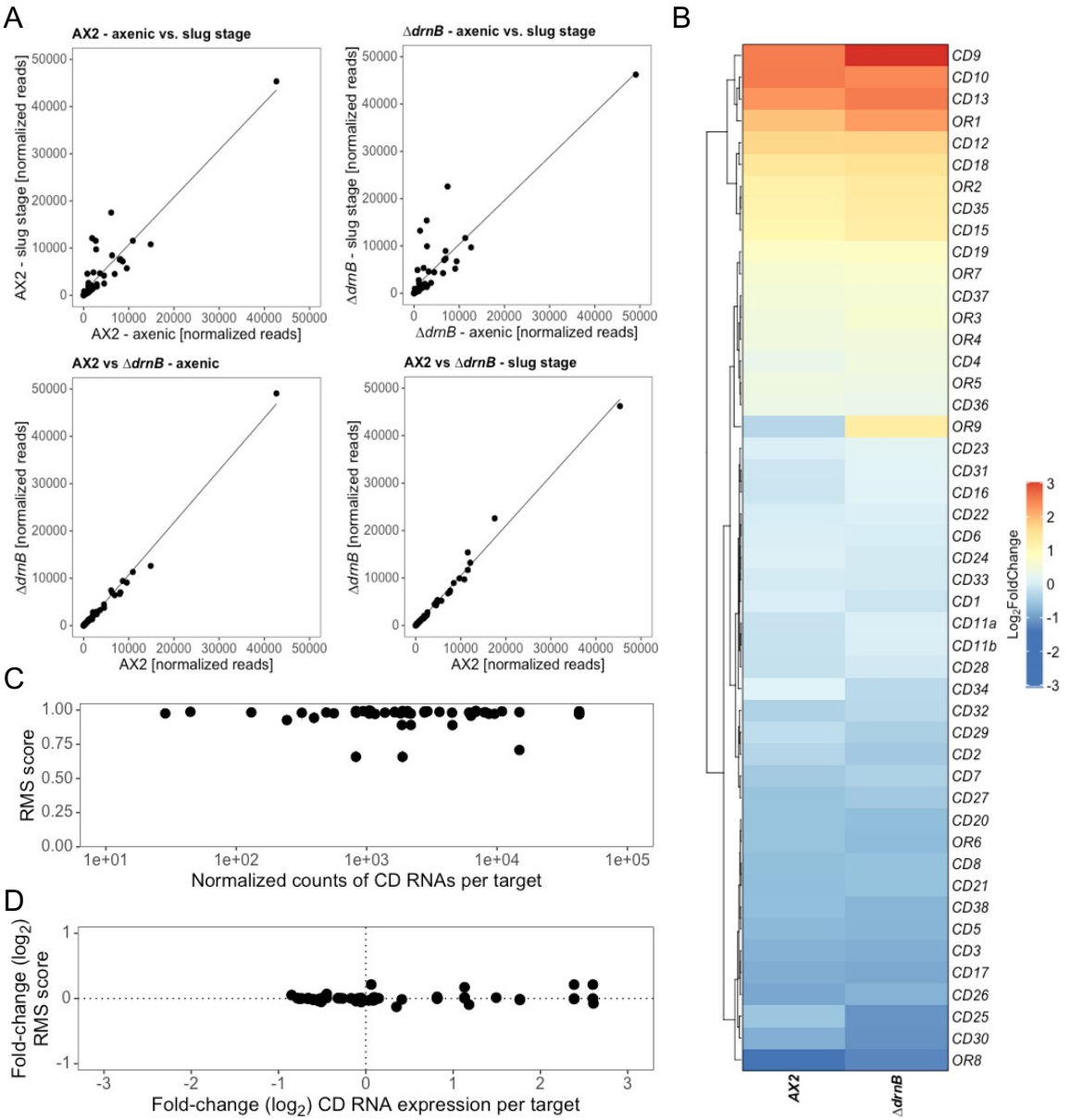


Figure 43. Analysis of box C/D snoRNA accumulation in axenic growth and development of the AX2 and *ΔdrnB* strains, and the relationship to 2´-O-Me in rRNA. (A) 2D plots of DESeq2-normalized reads in AX2 and *ΔdrnB* in axenic growth vs. slug stage (top) and in axenic growth or the slug stage of development of AX2 vs. *ΔdrnB* (bottom). (B) Hea™ap of log_2 fold-change of box C/D snoRNA levels in the slug stage of development in the indicated strains. (C) 2D plot of log_2 fold-change of box C/D snoRNA abundance per target in development and log_2 fold-change of the RMS score at all predicted methylation sites. (D) 2D plot of DESeq2-normalized counts of box C/D snoRNAs and RMS scores.

methylated sites revealed no correlation in axenic growth; rather, full and fractional methylation is observed independent of the CD RNAs abundance (Figure 43C). Furthermore, the log_2 fold-change of the RMS score and the CD RNA accumulation in the slug stage (Figure 43D) showed no differences. Thus, changes in the 2´-O-Me patterns can in general not be attributed to altered CD RNA amounts in the development of *D. discoideum*.

Pseudouridylation in the rRNAs of *D. discoideum*

D. discoideum's rRNAs have 66 predicted Ψ sites

After the description of the 2′-O-Me landscape in *D. discoideum*, the global ribosomal Ψ pattern in the amoeba should be elucidated. For this purpose, the recently established method HPS was used [399]. Beforehand, mapping of Ψ usually involved its derivatization using soluble carbodiimide [490], yielding noisy results due to harsh conditions during the preparation of RNA for deep sequencing [399]. In brief, HPS is based on the resistance of Ψ (and A, C, G) to treatment with hydrazine and aniline, which results in cleavage at all unmodified U residues [399]. Since sequencing reads start from the n + 1 position relative to an unmodified U, counting of the 5′-ends of the fragments therefore yields an underrepresentation of pseudouridylated sites. As RMS, Ψ scores immediately reflect the fraction of modified nucleotides at a given site [399]. For the analysis of the Ψ landscape in *D. discoideum*, nucleotides were considered pseudouridylated when the ScoreMEAN was higher than 0.92 and ScoreA was higher than 0.5. Besides AX2 in axenic growth, developing cells after 8, 16, and 24 h were included in the experiment.

At first, the pseudouridylated sites for axenically-growing amoebae were established. Using medium stringency as outlined above, 20 and 45 positions with an isomerized U were predicted on the 17S and 26S rRNA, respectively (Figure 44). One sample was excluded due to technical errors in sample preparation. In addition, a putative Ψ was identified at position 70 of the 5S rRNA (Figure 44). Most of these modified sites appeared fully modified, nonetheless, substoichiometric Ψ levels (Ψ score > 0.85) were observed in 5 positions on the 17S and 26S rRNA each. The Ψ score at position 26S-U2048, however, exhibits a high standard deviation and therefore might be fully pseudouridylated rather than fractionally modified (Figure 44). In

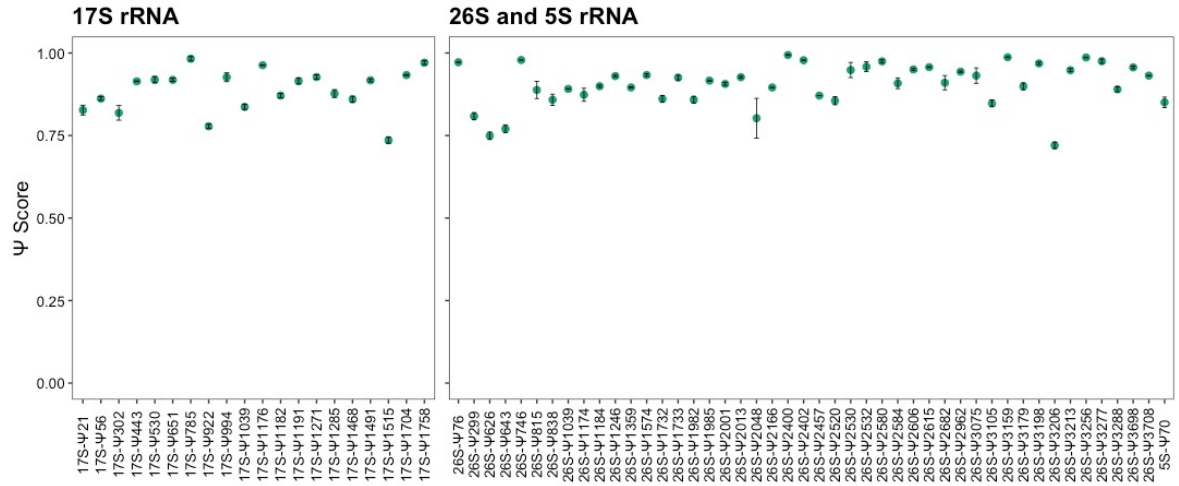


Figure 44. HydraPsiSeq analysis of the rRNAs in *D. discoideum*. Ψ scores at the predicted Ψ sites on the 5S, 17S, and 26S rRNA in axenic growth of the amoeba (n = 2).

addition to fractional 2′-O-Me, fractional pseudouridylation consolidates the presence of ribosome heterogeneity in *D. discoideum*.

Localization of Ψ sites in the secondary structure of *D. discoideum*'s rRNA

As a next step and to identify Ψ in the functional sites of the ribosome, the predicted Ψ sites were located in the secondary structure established previously. Almost two thirds (40/66) of pseudouridylated nucleotides reside in the vicinity of functionally important helices, with some of these modifications directly at the or immediately adjacent to the nucleotide in regions like the tRNA binding sites and the PTC (Figure 45 and Figure 46). Many of these positions are conserved between *D. discoideum*, *S. cerevisiae*, *H. sapiens*, and *A. thaliana* and are predominantly found in the 3′-major domain on the 17S rRNA and domains IV and V on the 26S rRNA (Table 19 and Table 20). Of the positions linked to important regions in the amoeba, 18 Ψ are part of the helices forming important structural elements like the exit tunnels, with the helices H65, H67, and H69 in domain IV of the 26S rRNA being heavily pseudouridylated. These helices form an integral part of the subunit interface [492], thus, the Ψ in these helices might serve a stabilizing function of the LSU-SSU interaction by rigidifying the RNA

Table 19. Ψ sites in the 17S rRNA of *D. discoideum* and orthologous sites in other organisms with their guiding snoRNAs.

Domain	*D. discoideum*	*S. cerevisiae*[B]		*H. sapiens*[C]		*A. thaliana*[DE]	
	Position	Position	snoRNA	Position	snoRNA	Position	snoRNA
5'	Ψ21	-	-	-	-	-	-
	Ψ56	-	-	-	-	-	-
	Ψ302	-	-	-	-	-	-
	Ψ443	-	-	-	-	-	-
	Ψ530	-	-	-	-	-	-
C	Ψ651	-	-	-	-	-	-
	Ψ785	-	-	-	-	-	-
	Ψ922	-	-	-	-	-	-
	Ψ994	Ψ999	snR31	Ψ1056	ACA8	Ψ1000	snoR72
	Ψ1039	-	-	-	-	-	-
3' major	Ψ1176	Ψ1181	snR85	Ψ1238	ACA5	-	-
	Ψ1182	Ψ1187	snR36	Ψ1244	ACA36	-	-
	Ψ1191	-	-	-	-	-	-
	Ψ1271	-	-	-	-	-	-
	Ψ1285	Ψ1290	snR83	Ψ1347	ACA4	-	-
	Ψ1468	-	-	-	-	-	-
	Ψ1491	-	-	-	-	-	-
	Ψ1515	-	-	-	-	-	-
	Ψ1704	-	-	Ψ1692	U70	Ψ1634	snoR134
3' minor	Ψ1758	-	-	-	-	-	-

[A] This study

[B] https://people.biochem.umass.edu/fournierlab/snornadb/mastertable.php

[C] https://www-snorna.biotoul.fr/human_yeast/

[D] https://ics.hutton.ac.uk/cgi-bin/plant_snorna/home

[E] [53, 491]

Table 20. Ψ sites in the 5S and 26S rRNA of *D. discoideum* and orthologous sites in other organisms with their guiding snoRNAs.

Domain	*D. discoideum*	*S. cerevisiae*[B]		*H. sapiens*[C]		*A. thaliana*[DE]	
	Position	Position	snoRNA	Position	snoRNA	Position	snoRNA
I	Ψ76	-	-	-	-	-	-
	Ψ299	-	-	-	-	-	-
	Ψ626	-	-	-	-	-	-
	Ψ643	-	-	-	-	-	-
	Ψ746	-	-	-	-	-	-
	Ψ815	-	-	-	-	-	-
0	Ψ838	-	-	-	-	-	-
II	Ψ1039	-	-	-	-	-	-
	Ψ1174	Ψ960	snR8	Ψ1664	ACA56	Ψ967	-
	Ψ1184	-	-	-	-	-	-
	Ψ1246	-	-	-	-	-	-
	Ψ1359	Ψ1124	snR5	Ψ1847	ACA32	Ψ1130	snoR80
	Ψ1574	-	-	-	-	-	-
III	Ψ1732	-	-	-	-	-	-
	Ψ1733	-	-	-	-	-	-
	Ψ1982	-	-	-	-	-	-
	Ψ1985	-	-	-	-	-	-
	Ψ2001	-	-	-	-	-	-
	Ψ2013	-	-	-	-	-	-
	Ψ2048	-	-	-	-	-	-
IV	Ψ2166	-	-	-	-	-	-
	Ψ2400	Ψ2133	snR3	Ψ3616	ACA6	Ψ2124	-
	Ψ2402	-	-	Ψ3618	ACA19	Ψ2126	snoR87
	Ψ2457	Ψ2191	snR32	Ψ3674	-	Ψ2181	-
	Ψ2520	-	-	Ψ3737	ACA23	Ψ2244	snoR92
	Ψ2530	Ψ2264	snR3	Ψ3747	-	Ψ2248	snoR79
	Ψ2532	Ψ2266	snR84	Ψ3749	-	Ψ2250	-
	Ψ2580	Ψ2314	snR86	Ψ3797	ACA48	Ψ2304	-
	Ψ2584	-	-	Ψ3801	ACA54	-	-
	Ψ2606	Ψ2340	snR9	Ψ3823	ACA58	-	-
	Ψ2615	Ψ2349	snR82	Ψ3832	ACA8	Ψ2339	snoR83
V	Ψ2682	Ψ2416	snR11	Ψ3899	ACA3	Ψ2406	-
	Ψ2962	-	-	-	-	Ψ2620	snoR78
	Ψ3075	-	-	-	-	-	-
	Ψ3105	-	-	-	-	-	-
	Ψ3159	Ψ2826	snR34	Ψ4373	U65	Ψ2816	U65
	Ψ3179	-	-	Ψ4393	U68	-	-
	Ψ3198	Ψ2865	snR46	Ψ4412	ACA16	Ψ2855	-
	Ψ3206	-	-	-	-	-	-
	Ψ3213	Ψ2880	snR34	Ψ4427	U65	Ψ2870	U65
	Ψ3256	Ψ2923	snR10	Ψ4470	ACA21	Ψ2913	snoR74
	Ψ3277	Ψ2944	snR37	Ψ4491	ACA10	-	-
	Ψ3288	-	-	-	-	Ψ2945	-
VI	Ψ3698	-	-	-	-	-	-
	Ψ3708	-	-	-	-	-	-
5S rRNA	Ψ70	-	-	-	-	-	-

[ABCDE] as in Table 19

phosphate-sugar backbone. Three Ψs are found on uridines directly located in the A, P, and E sites: 17S-Ψ994 on h24 residing in the E site, 26S-Ψ2532 on H69 residing in the P site, and 26S-Ψ3256 being part of the A loop in H92. Since the isomerization of uridine results in conformational change in the nucleotide, these modifications might be essential for tRNA binding in addition of their stabilizing function. For example, a loss of modifications on H69

of the 25S rRNA in yeast decreases translational fidelity drastically [248] and loss of 25S-Ψ2923 (orthologous to 26S-Ψ3256 in *D. discoideum*) increased the misreading of near- and non-cognate codons [493]. Even though we found five 2′-O-Me in domain 0, only Ψ838 was predicted in that region.

Overall, pseudouridylated nucleotides cannot be found in the ES of the amoeba, except for 4 Ψ located to ES7 on H25 of the 26S rRNA. This might be of functional importance as will be discussed later.

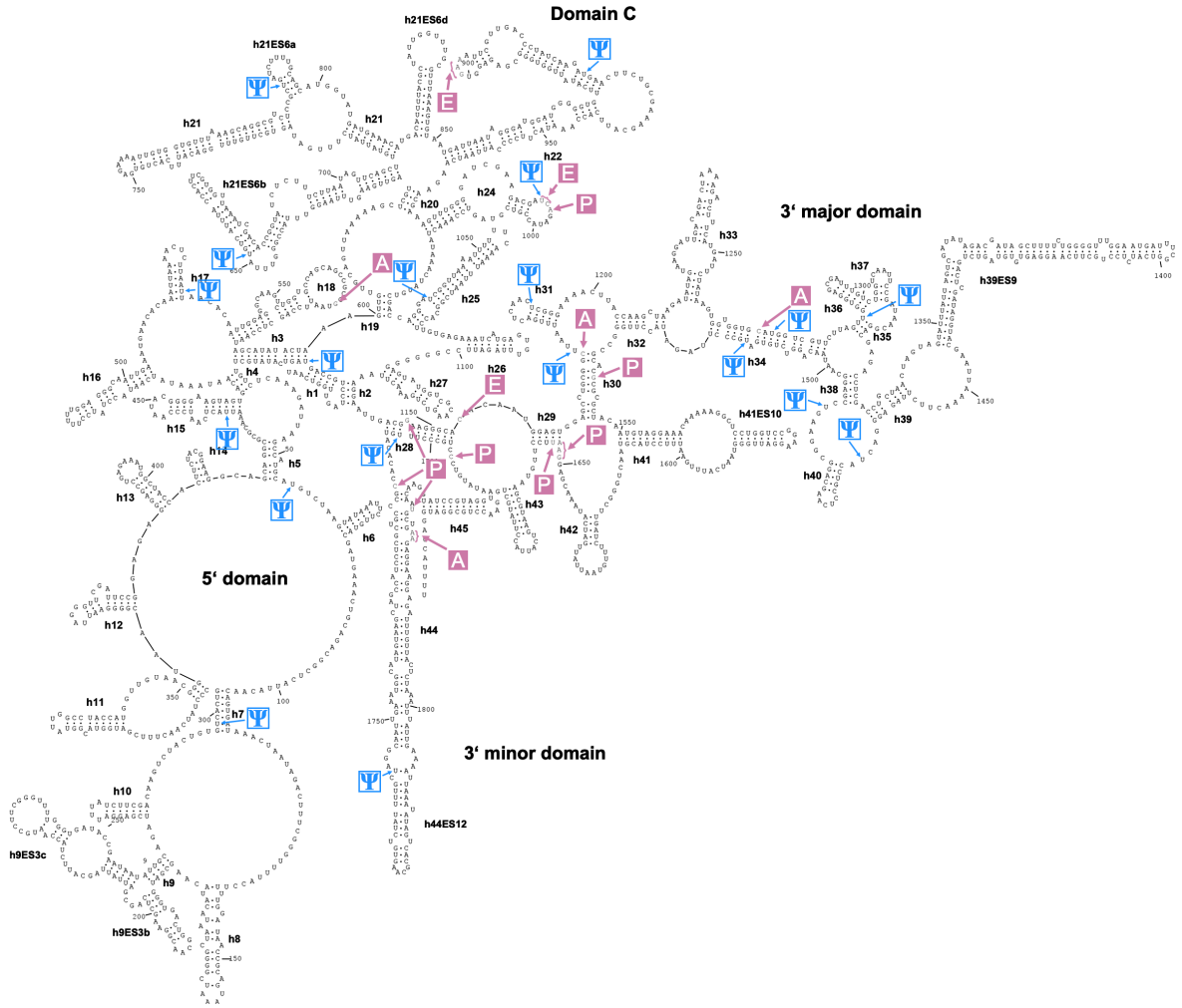


Figure 45. Ψ sites in the 17S rRNA of *D. discoideum* as identified by HydraPsiSeq. The secondary structure of the 17S rRNA was inferred by homology and drawn using RNAviz (v. 2.0.3). Ψ sites are marked in blue. Helices (hx) are named to convention and expansion segments (ESx) are labeled with x: natural number.

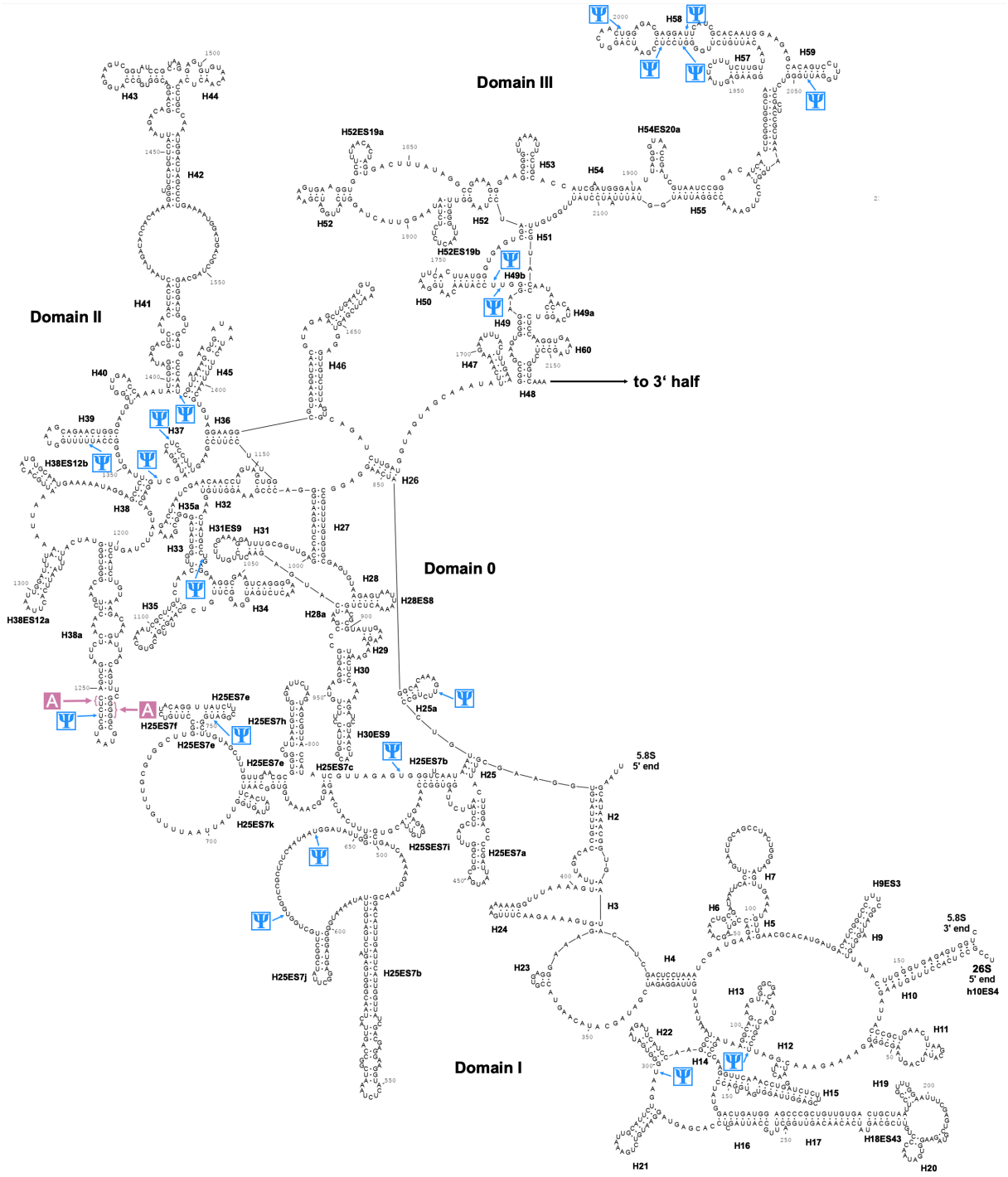


Figure 46. Ψ sites in the 26S rRNA and 5S rRNA of *D. discoideum* as identified by HydraPsiSeq. The secondary structure of the 26S rRNA and 5S rRNA was inferred by homology and drawn using RNAviz (v. 2.0.3). The 5S rRNA is displayed in the inset of the 5’ half of the 26S rRNA. Ψ sites are marked in blue. The predicted interaction with the 5.8S rRNA (previous page) is shown at the 5’end. Helices (Hx) are named to convention and expansion segments (ESx) are labeled with x: natural number.

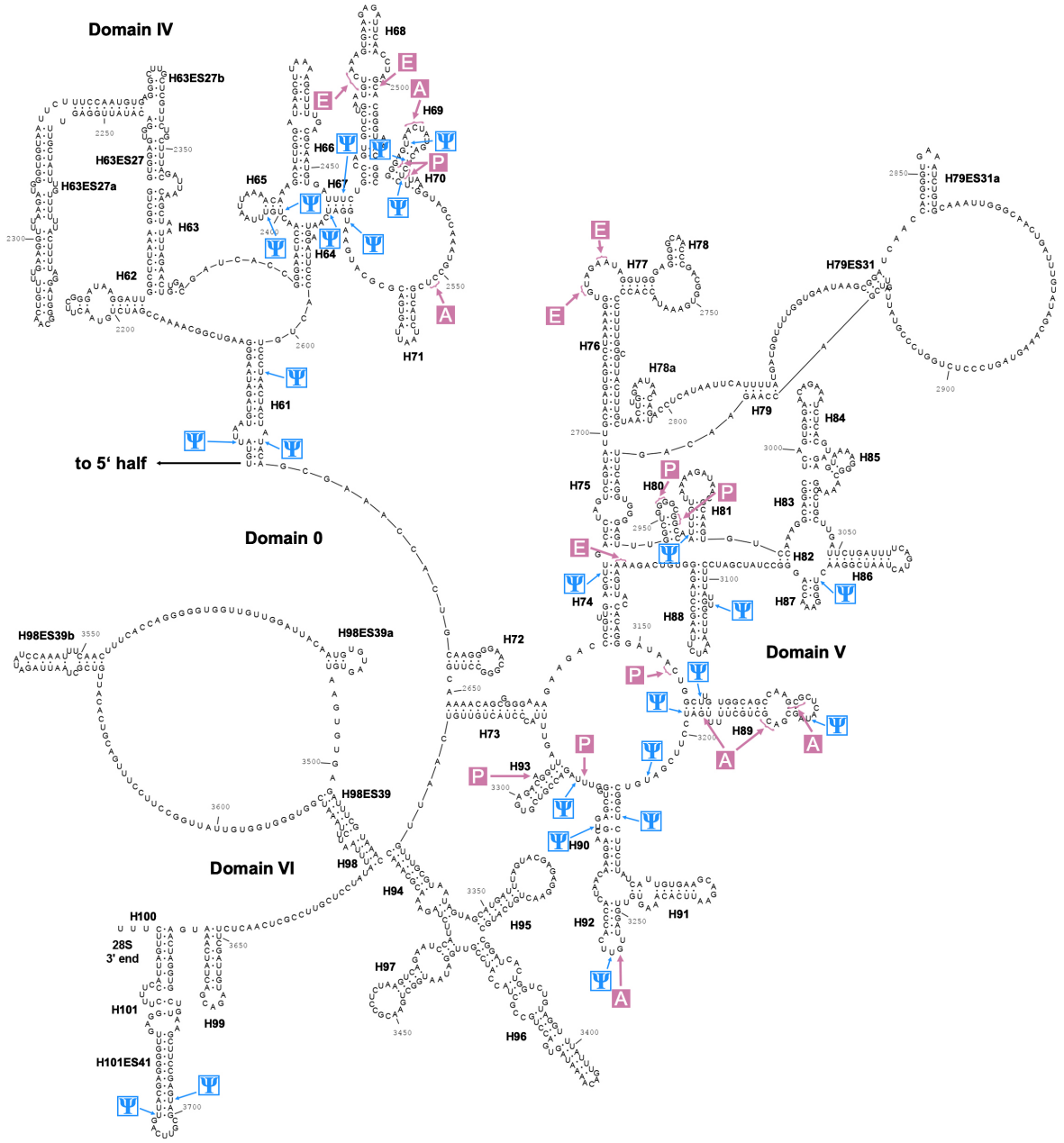


Figure 46. Ψ sites in the 26S rRNA and 5S rRNA of *D. discoideum* as identified by HydraPsiSeq. ctd.

Pseudouridylation in the development of *D. discoideum*

Upon establishment of the Ψ landscape in the amoeba, the impact of development of pseudouridylation levels was investigated. For this purpose, HPS was performed in triplicate for three developmental time points: 8 h, 16 h, and 24 h. Only potential changes the Ψ sites were considered, as the application of medium stringency filters (as outlined above) yielded no additional Ψ candidate sites (data not shown).

As a first step, exploratory data analysis was performed on the Ψ scores of the triplicates of each developmental time point including axenic growth using PCA. As with the axenic cells, one sample from the 16 h development was excluded due to technical errors in sample preparation. The analysis revealed that Ψ scores generally differed most between 8 h of development versus axenic growth, 16 h and 24 h of development (Figure 47A). In general, Ψ levels were significantly decreased after 8 h of development (Figure 47B), as confirmed by a one-way ANOVA. At 16 h of development, pseudouridylation levels recovered to the levels observed in axenic growth and were not further affected by development, except for some individual sites. In the context of *Dictyostelium*'s development, at 4 to 6 h of development the amoeba commits to development and starts aggregating after 8 h [494]. This also might influence pseudouridylation of rRNAs downstream.

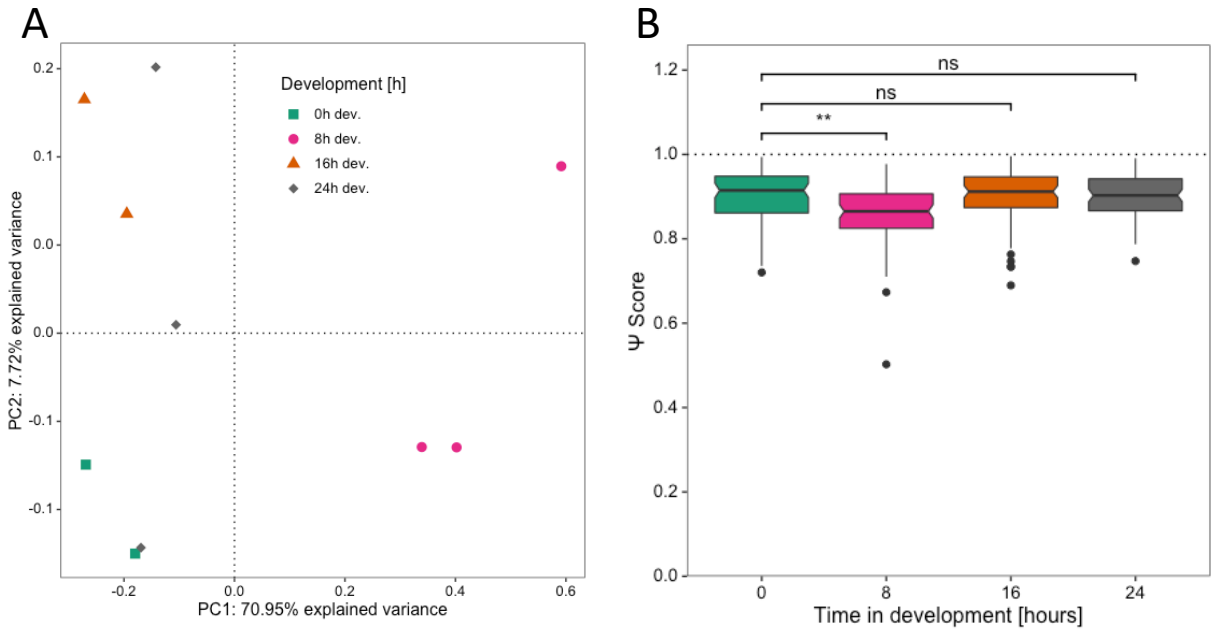


Figure 47. HydraPsiSeq analysis of the ribosomal RNA in *D. discoideum*. (A) Principal component analysis of Ψ scores in axenic growth (n = 2) and development (8 h (n = 3), 16 h (n = 2), and 24 h(n = 3)) of AX2. (B) Comparison of total ribosomal Ψ levels in axenic growth against development of AX2. The asterisks indicate significant differences (** ≙ $p < 0.005$) from axenic growth determined by one-way ANOVA and the notch indicates the 95% confidence interval.

At the level of the individual Ψ sites, Ψ levels were considered changed, if the Ψ score deviated more than 0.05 from axenic growth (Yuri Motorin, personal communication). Using this approach, 27 positions were found hypomodified and 4 positions hypermodified after 8 h development, 3 hypomodified and 1 hypermodified after 16 h development, and 3

hypomodified and 4 hypermodified after 24 h development. Since the most significant changes occur at 8 h of development, focus shifted to the changes in Ψ score observed after 8 h development compared to axenic growth in the following paragraphs.

The highest change in Ψ score was found at position U1174 in the loop of H37 in the 26S rRNA (Figure 46) which changed from fully modified in axenic growth to being modified in approximately 50% of the ribosomal population in 8 h development (Figure 48A). Even though a clear function has not been assigned to H37, there are indications that its loop section partakes in tRNA binding at the A site [495] as well as in the interaction with the mRNA located in the mRNA channel of the ribosome [496, 497]. The altered loop geometry induced by the introduction of a Ψ might be necessary in these interactions, as substantiated by the strong

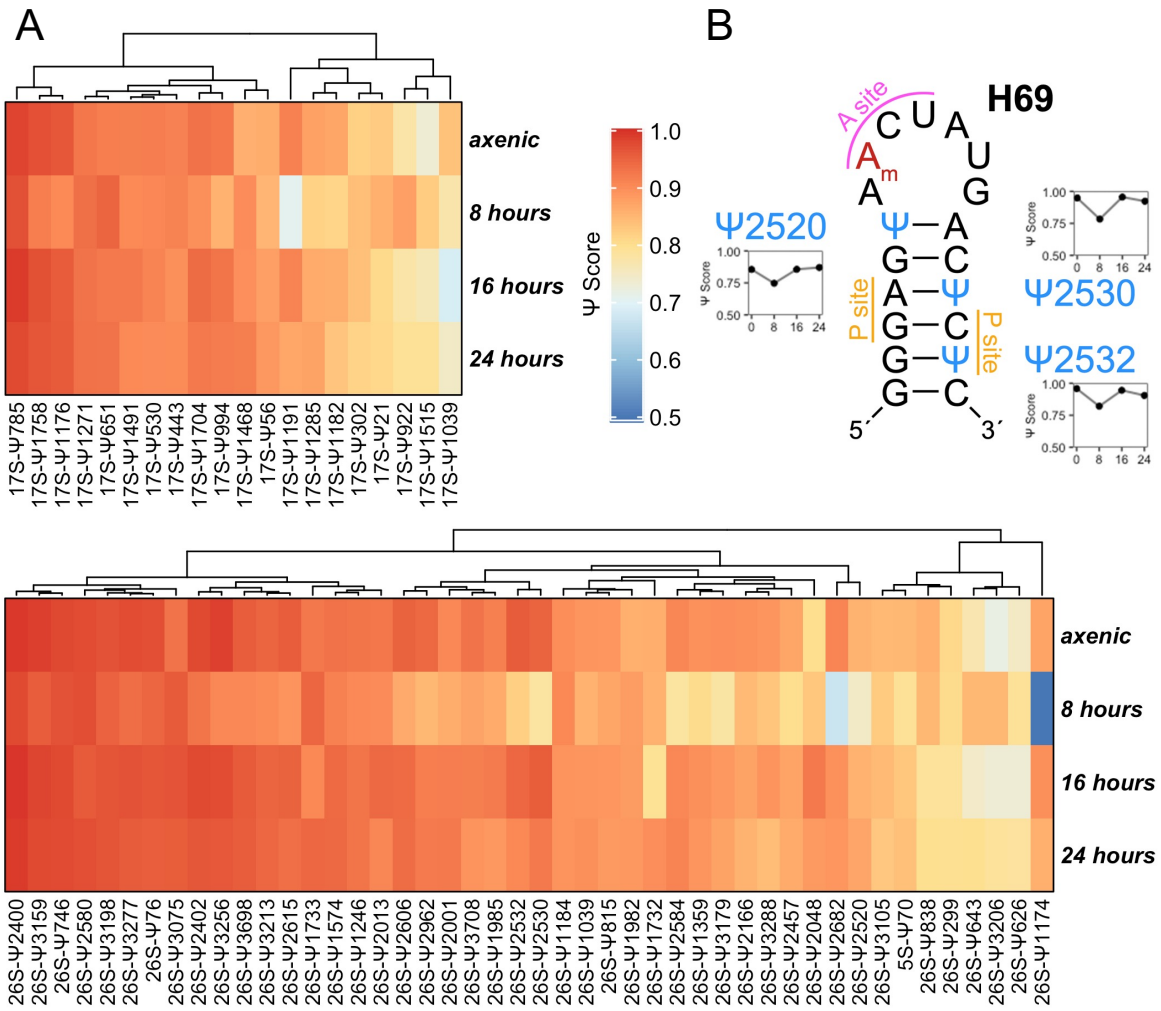


Figure 48. Heatmap of Ψ scores during development on the 5S, 17S, and 26S rRNAs in *D. discoideum*. (A) Visualized are the mean Ψ scores (n = 3) in the development of AX2 at each individual modified position at the indicated timepoints. (B) Pseudouridylation in the development of AX2 on H69 of the 26S rRNA. The nucleotides residing in the A (pink) and P (yellow) sites are indicated and the 2´-O-methylated A in the loop of H69 is shown in red. Ψ sites are shown in blue and the change of Ψ scores during development is displayed.

conservation across different evolutionary supergroups (Table 20). The position U1191, located at the base of the loop of h31 in the 17S rRNA (Figure 45), exhibited a Ψ score decrease from ~0.91 to ~0.71, again possibly altering the loop conformation in development. Close by at position U1186 in the loop of h31, the hypermodified 1-methyl-3-(3-amino-3-carboxypropyl)-Ψ ($m^1acp^3\Psi$) was found to be essential for rRNA biogenesis [498] and tRNA binding in the P site [499]. Thus, a conformational change of the h31 loop might modulate the tRNA interaction with the mRNA. Three of the strongest reductions in pseudouridylation levels clustered in the stem of H69 of the 26S rRNA (Figure 48B), with the decreases ranging from 0.11 to 0.17 at 8 h of development. As mentioned above, the set of modifications on H69 are essential for translational fidelity [248], suggesting substoichiometric modifications might constitute a fine-tuning mechanism for tRNA binding. It might also alter long-range interactions in the ribsome, as H69 is part of the intersubunit bridges [492]. Without structural data, however, this remains speculation. Overall, differential pseudouridylation in development appeared to occur mainly in the domains IV and V of the 26S rRNA (Figure 46 and Figure 48A) in the amoeba, which indicated that these changes might fulfill a functional role and are not random side effects of the developmental process.

Identification of small RNAs involved in rRNA processing of Dictyostelia

Prediction of U3 snoRNA loci in the genome of *D. discoideum*

In the RNAseq analysis, differential accumulation of a high number of box C/D snoRNAs was observed during the development of *D. discoideum*. The levels of 2´-O-Me, however, were not affected by the change in box C/D snoRNA abundance, indicating distinct roles for these molecules in the development of the amoeba. In several other (eukaryotic) organisms, snoRNAs are implied in the processing and folding of rRNA during ribosomal biogenesis besides guiding chemical modifications (reviewed in Ojha *et al.* [372]). Searching the genome of *D. discoideum* for the U14, U17, U22 snoRNA and snR10 using Infernal during this work did not yield any results, however, a previous study by Wise and Weiner [500] described five genomic loci of the U3 snoRNA in the *D. discoideum* AX3 strain. The search of the genomic copies of the U3 snoRNA in AX2 was performed using `Infernal` [434] and seven distinct loci in the genome of the amoeba were identified (Table 21). The discrepancy to the data by Wise and Weiner [500] might be attributed to strain-specific differences as AX3 was derived directly from the wildtype NC-4 [501], while AX2 was isolated from a subculture of AX1 which originated as well from NC-4 [4, 502]. Four of the seven U3 loci are located on chromosome 4 in relative proximity between positions 2,904,295 and 2,931,856 (Table 21). Similarly, the two copies on chromosome 6 are encoded in vicinity of each other (Table 21).

The sequences of all U3 snoRNA copies were 209 nt long and 96.2% identical to each other, however, they exhibit some minor sequence heterogeneity, of which two (C/U heterogeneity at pos. 91 and G/A heterogeneity at pos. 190) were already described by Wise and Weiner [500]. In addition to those positions, six variable positions in the sequence of the U3 snoRNA genes were identified (Table 21). Based on the pattern of variable nucleotides, the copies on chromosome 6 diverge in 7/8 positions (Table 21) and suggest that the duplication events might not have occurred recently in the evolution of the amoeba.

Table 21. Locations of the U3 snoRNA genes and positions with sequence heterogeneity.

Name	Chr.	Start	End	Strand	Positions with sequence heterogeneity							
					91	102[A]	128	137	144	145	153	190[A]
U3-1	3	3806685	3806893	+	U	C	G	A	U	U	C	G
U3-2	4	2904295	2904503	+	U	C	G	A	C	U	C	G
U3-3	4	2927177	2927385	+	C	C	G	A	C	U	C	G
U3-4	4	2931648	2931856	+	C	C	G	A	C	U	U	G
U3-5	4	2936427	2936635	+	U	C	G	A	C	U	U	G
U3-6	6	2268599	2268807	+	C	U	A	G	U	G	C	A
U3-7	6	2270802	2271010	+	C	U	A	A	U	G	C	A

[A] Positions already described in Wise and Weiner [500]

As the sequence heterogeneities described above could potentially influence the base pairing in the structure of the U3 snoRNA, the secondary structure of the RNA was predicted in the amoeba using `RNAfold`. The inferred secondary structure is shown Figure 49A and is similar to U3 snoRNA homologs in other eukaryotes like the clawed frog [503] and yeast [504]. In order to increase comparability across species, the helix numbering scheme suggested by Marz and Stadler [467] was adopted. This publication also features a generalized model for the U3 snoRNA secondary structure in eukaryotes ranging from Kinetoplastids to Metazoans, which was used for comparison [467]. Helices M1, M2, M3, and M4 were found structurally conserved in *D. discoideum* (Figure 49). The loop between helices M6 and M8 seemed to be degenerated to one (U3F and U3G) or two (U3A to U3E) nucleotides (Figure 49). It is possible, however, that M6 and M8 rather fused into one helix during the evolution of the amoeba. Since helix M7 is only present in fungi [467], it is not surprising that this helix was found absent in *D. discoideum*.

Next, the outlined set of variable positions found across all U3 snoRNA genes were mapped in the secondary structure of the RNA (Figure 49A and B). As mentioned above, Wise and Weiner [500] described the sequence variation at positions 102 and 190 before. Both nucleotides base pair at position 8 of the central helix M4 and variation at one position inadvertently leads to a compensatory change at the base paired nucleotide. A similar phenomenon was observed for the base pair of the variable positions 128 and 144 in helix M8. Next to the divergent base pair 128-144, position 145 exhibited a U/G heterogeneity with U3 isoforms on chromosome 3 and 4

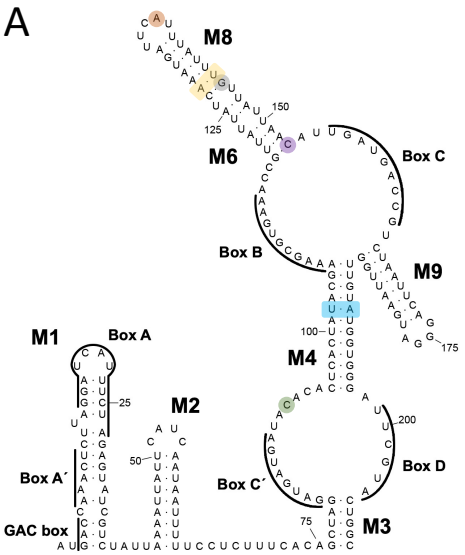


Figure 49. Secondary structure of the U3 snoRNA of *D. discoideum*. (A) Secondary structure of the U3 snoRNA. The structure was predicted using RNAfold (based on U3-7) and drawn using RNAviz (v. 2.0.3). Annotations are based on the generalized U3 snoRNA model by Marz and Stadler [467]. (B) Heterogeneric base pairs in the hairpins and stems of the U3 snoRNA.

(U3-1-U3-5) featuring a U and isoforms encoded on chromosome 6 (U3-6 and U3-7) featuring a G. Thus, the U/U bulge positioned between helices M6 and M8 was found extended by one nucleotide in U3-1-U3-5.

rDNA transcription and U3 snoRNA abundance in development of the amoeba

Previous work in the group established that rRNA precursors are less abundant after 8 and 16 h of development compared to axenic growth, with an approximate two-fold accumulation at 24 h (Christian Hammann, personal communication). To investigate whether this reduction at 8 and 16 h of development is caused by an increase in rRNA processing or by a decrease in rDNA transcription, nuclear run-on transcription was employed. In these experiments, isolated nuclei are allowed to transcribe and incorporate radioactively labelled UTP into the nascent transcripts. Specific detection of the precursor rRNAs is achieved by hybridization to a 5′-ETS *in vitro* transcript immobilized to a membrane. This approach allowed for the quantification of the transcription of specific genes by normalization to a housekeeping gene with stable expression pattern (*gpdA* in this case). By applying this technique to nuclei isolated from axenic cells and different developmental time points, no change of transcription of the rDNA genes in the amoeba was observed (Figure 50A). This suggests that the reduced abundance of rRNA precursor at 8 and 16 h of development was likely caused by an acceleration of the rRNA processing. Rather surprisingly, the rDNA transcription was not altered during development, even though approximately a third of the cells mainly localized to the stalk die in the process and transcription of ribosomal genes is repressed [505]. However, due to the chosen approach, it cannot be excluded that the observed transcription mainly occurs in the spores. The accumulation of the rRNA precursors at 24 h of development was most likely caused by a decreased rRNA processing speed and independent of rDNA transcription.

Despite the absence of the A0 site in the 37S primary transcript in the amoeba [15], A1 and A2 (*D. discoideum*: c4 and c2) sites are still present and processing at these sites is usually facilitated by base pairing with the U3 snoRNA [373]. To investigate the possibility that the increased rRNA processing is (partly) a result of differential U3 snoRNA accumulation in development, the abundance of the U3 snoRNA was determined using RNAseq data. Axenic cells and the 16 h time point of development of the AX2 and Δ*drnB* strains were included in the analysis, as the data was already deposited and readily available. In doing so, an accumulation of the U3 snoRNA at the 16 h time point by 1.5-2-fold was observed dependent of the analyzed strain (Figure 50B). There was, however, no significant difference between AX2 and Δ*drnB* (Figure 50A).

Together with the results from the nuclear run-on transcription, the increased abundance of the U3 snoRNA in developed *D. discoideum* cells strongly suggested that rRNA processing is enhanced in development. Considering the emerging theme of ribosome heterogeneity in dependence of developmental and/or environmental cues, it is tempting to speculate that the enhanced rRNA processing is due to the need of specialized ribosomes during *Dictyostelium*'s development.

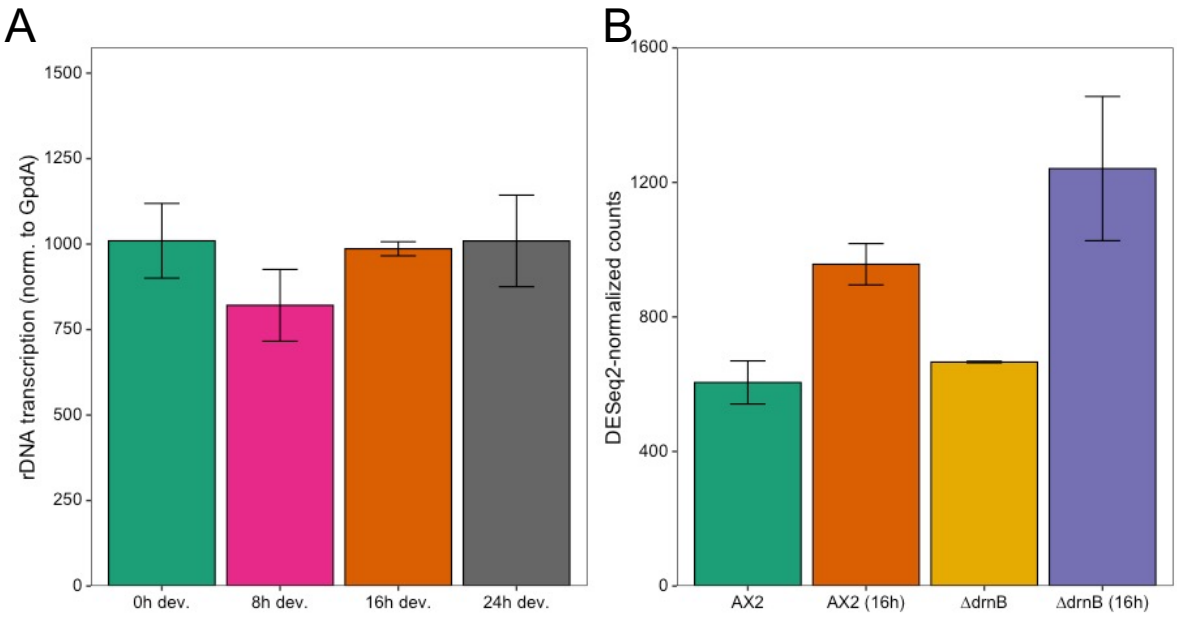


Figure 50. The rDNA transcription rate and abundance of the U3 snoRNA in the development of *D. discoideum*. (A) Nuclear run-on transcription of rDNA in the nuclei of developing AX2 cells (n = 3). (B) U3 snoRNA levels in the development of AX2 and *ΔdrnB* as determined by RNAseq (n = 2, data from Liao *et al.* [462]).

Identification of U3 snoRNA loci in all major groups of Dictyostelia

Since *D. discoideum*'s genome codes for an unusual high number of genomic copies of the U3 snoRNA for a basal eukaryote [467], the question was posed whether this is a common theme in all major groups of dictyostelids. For this purpose, the complete and well-annotated genomes of five representatives of each group of the Dictyostelia besides *D. discoideum* (group 4) were used, i.e., *Dictyostelium purpureum* (group 4) [506], *Dictyostelium lacteum* (group 3) [507], *Polysphondylium pallidum* (group 2A) [508], *Acytostelium subglosum* (group 2B) [509], and *Dictyostelium fasciculatum* (group 1) [508]. As before with *D. discoideum*, a CM of the U3 snoRNA to search the dictyostelid genomes was utilized with the program `Infernal` [434]. Due to the high conservation in structure and/or sequence, hits scoring 25 or higher in `Infernal` were assumed to be *bona fide* U3 snoRNA loci. This approach indicated the presence of at least

one U3 snoRNA locus in each of the representative dictyostelids (Table 22). Besides the one predicted copy in *D. fasciculatum* (Table 22), another sequence that resembles a U3 snoRNA was found. That putative RNA molecule, however, seems to be truncated at the 3′-end, missing the terminal stem (helix M3) and box D. Both of these are a likely prerequisites for a functional RNA molecule, as they are present in every eukaryotic U3 snoRNA locus [467]. Therefore, this sequence was not considered as a *bona fide* U3 snoRNA. With the exclusion of that truncated version in *D. fasciculatum*, the representatives of group 1 and group 2A/B only code for one U3 snoRNA gene each, while the genomes of the group 3 and group 4 representatives feature multiple copies. *D. lacteum*, a group 3 member, possesses three distinct U3 snoRNA loci, while both group 4 dictyostelids, *D. purpureum* and *D. discoideum*, code for six and seven copies, respectively. This indicated that at least in the subset of investigated dictyostelids, U3 snoRNA gene duplication is restricted to the group 3 and group 4 Dictyostelia.

Table 22. Genomic locations and conserved box motifs of dictyostelid U3 snoRNAs. Nucleotides that are conserved in all tested dictyostelids are marked in red.

Species	Name	Sequence	Start	End		Conserved motifs						
						GAC	Box A′	Box A	Box C′	Box B	Box C	Box D
D. discoideum	U3-1	DDB0232430	3806685	3806893	+	GAC	AACTC	GGATCATTTCTA	GATGATGA	AAGCGTGAA	TGATGACCG	TTCGTA
	U3-2	DDB0232431	2904295	2904503	+	GAC	AACTC	GGATCATTTCTA	GATGATGA	AAGCGTGAA	TGATGACCG	TTCGTA
	U3-3	DDB0232431	2927177	2927385	+	GAC	AACTC	GGATCATTTCTA	GATGATGA	AAGCGTGAA	TGATGACCG	TTCGTA
	U3-4	DDB0232431	2931648	2931856	+	GAC	AACTC	GGATCATTTCTA	GATGATGA	AAGCGTGAA	TGATGACCG	TTCGTA
	U3-5	DDB0232431	2936427	2936635	+	GAC	AACTC	GGATCATTTCTA	GATGATGA	AAGCGTGAA	TGATGACCG	TTCGTA
	U3-6	DDB0232433	2268599	2268807	+	GAC	AACTC	GGATCATTTCTA	GATGATGA	AAGCGTGAA	TGATGACCG	TTCGTA
	U3-7	DDB0232433	2270802	2271010	+	GAC	AACTC	GGATCATTTCTA	GATGATGA	AAGCGTGAA	TGATGACCG	TTCGTA
D. purpureum	U3-1	scaffold_688	3987	4195	-	GAC	TACTC	GGATCATTTCTA	GATGAAGA	AAGCGGGAA	TGATGATCG	TTCTTA
	U3-2	scaffold_140	37498	37706	-	GAC	TACTC	GGATCATTTCTA	GATGAAGA	AAGCGGGAA	TGATGATCG	TTCTTA
	U3-3	scaffold_73	89977	90185	+	GAC	TACTC	GGATCATTTCTA	GATGAAGA	AAGCGGGAA	TGATGATCG	TTCTTA
	U3-4	scaffold_101	57185	57393	-	GAC	TACTC	GGATCATTTCTA	GATGAAGA	AAGCGGGAA	TGATGATCG	TTCTTA
	U3-5	scaffold_202	34893	35101	+	GAC	TACTC	GGATCATTTCTA	GATGAAGA	AAGCGGGAA	TGATGATCG	TTCTTA
	U3-6	scaffold_688	3598	3806	-	GAC	TACTC	GGATCATTTCTA	GATGAAGA	AAGCGGGAA	TGATGATCG	TTCTTA
D. lacteum	U3-1	LODT01000037.1	192486	192702	-	GAC	AACTC	GGATTCATTTCA	GATGTTGA	TAGCGGGAA	TGATGATCG	TTCTGA
	U3-2	LODT01000037.1	423377	423592	+	GAC	AACTC	GGATTCATTTCA	GATGTTGA	TAGCGGGAA	TGATGATCG	TTCTGA
	U3-3	LODT01000035.1	825082	825298	-	GAC	AACTC	GGATTCATTTCA	GATGTTGA	TAGCGGGAA	TGATGATCG	TTCTGA
P. pallidum	U3	GL290983.1	372856	373065	-	GAC	TATTC	GGATCATTTCTT	GATGAAGA	GAGCGCGAT	AGATGACCG	CTCTAA
A. subglosum	U3	DF837589.1	656447	656660	-	GAC	TATTC	GGATCATTTCTA	GATGATGA	GAGCGTGAT	AGATGACCG	GTCGAA
D. fasciculatum	U3	GL883010.1	1346903	1347114	-	GAC	TACTC	GGATCTCTTAAC	GATGATGA	GAGCGTGAA	TGATGATCG	GGCAAA

In terms of the characteristic sequence motifs, the dictyostelid U3 snoRNAs mostly adhered to the consensus. Box A′ with the consensus 5′-TACTY-3′ was highly conserved in the group 1, 3, and 4 Dictyostelia, while the representatives of groups 2A and B displayed a C/T exchange at position 3 of the motif. The same pattern of conservation can be observed in the core of box A with the consensus 5′-GYATCW-3’. As in other box C/D snoRNAs, the core of the box C and box C′ motifs mostly adhered to the 5′-RTGATGA-3′ consensus, with box C′ showing

more variation between species. In contrast, the dictyostelid box B, corresponding to box D′, featured 5′-CGBGA-3′, exchanging the first A compared to the usual 5′-AGYGA-3′ consensus. Surprisingly, the box D appeared degenerated in most dictyostelids. Only the U3 snoRNA of *D. lacteum* featured a box D according to the 5′-CTGA-3′ consensus. A graphical overview of the motif conservation in the dictyostelids can be found in Figure 51.

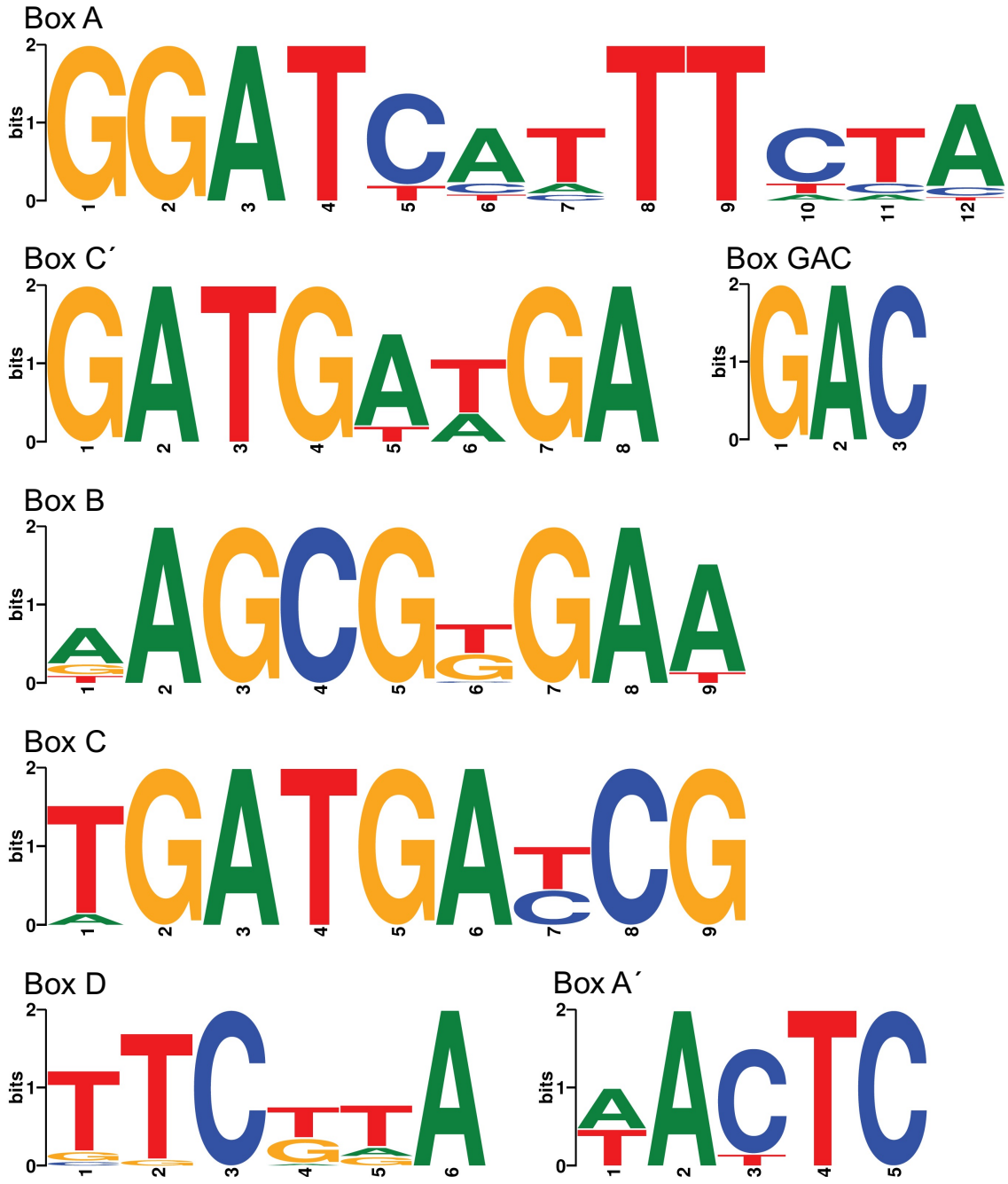


Figure 51. Conserved sequence elements in the U3 snoRNAs of the Dictyostelia. The logos were generated with WebLogo [457] and represent the box A, A′, B, GAC, C, C′, and D motifs. Most sequence motifs of the dictyostelid adhere to the consensus, with the exception of the box D motif.

As in *D. discoideum*, minor sequence heterogeneities can be observed in all Dictyostelia possessing more than one U3 snoRNA gene (Figure 52A). Even though *D. purpureum* features a similar number of U3 loci, the number of variable positions was found to be lower compared to *D. discoideum*. This might be explained by more recent duplication events in the former species. In the Dictyostelia investigated here, there seems to be no distinct pattern to the variable nucleotides in the U3 snoRNA sequences, suggesting that these variations occurred after their evolutionary split.

Since no data for the U3 snoRNA of the Dictyostelia was available, the sequnces were compared (Figure 52A) and the consensus secondary structure encompassing all representatives of the major dictyostelid groups was predicted. The alignment of the dictyostelid U3 snoRNAs yielded a phylogenetic tree that matched the multi-gene phylogeny by Schilde *et al.* [510]. The lengths the U3 snoRNA varied between 209 nt and 217 nt. For the prediction of the secondary structure, sequence alignments of all dictyostelid U3 snoRNAs were produced and utilized in R2R [440] to draw the consensus secondary structure diagram (Figure 52B). Like seen for

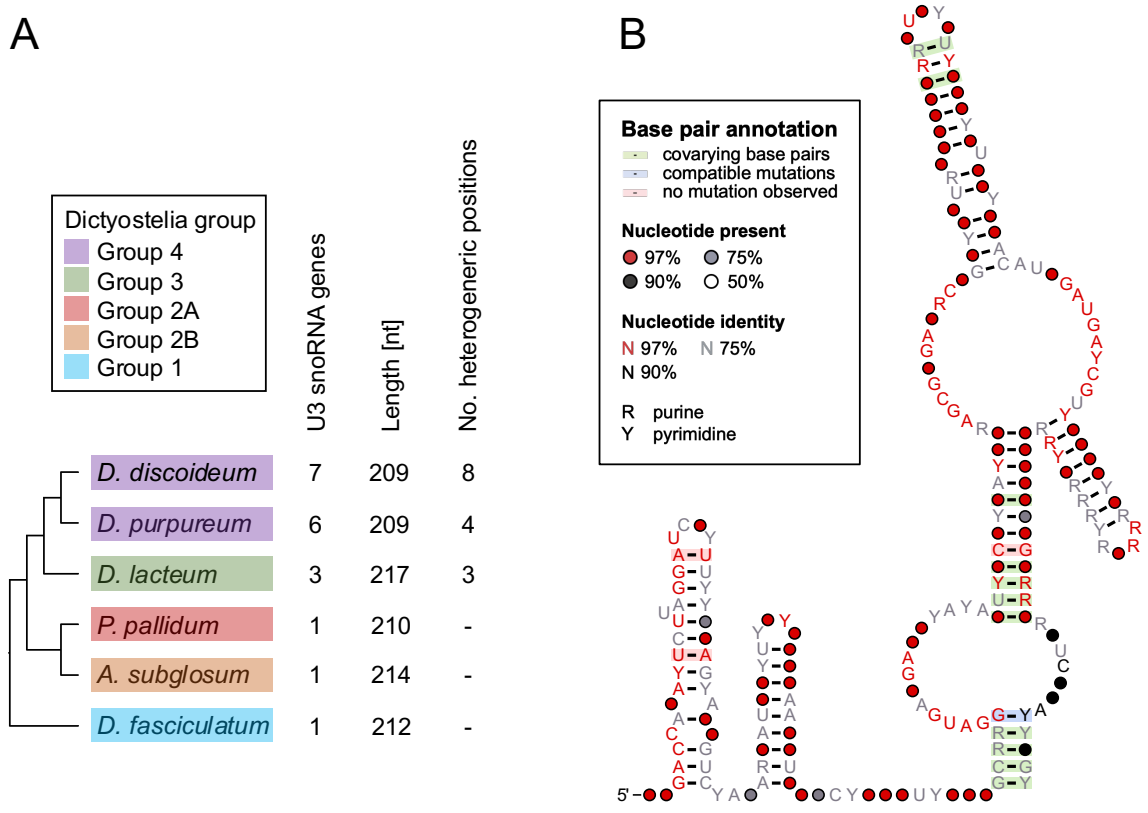
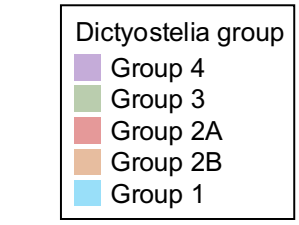


Figure 52. Number and characteristics of U3 snoRNA genes and the consensus secondary structure of the U3 snoRNA in representatives of all major groups of the Dictyostelia. (A) Shown are the number of predicted U3 snoRNA genes, their length, and the number of variable positions. Each major dictyostelid group is represented by at least one member (boxed legend). The phylogenetic tree is based on an alignment of the U3 snoRNA genes. (B) Consensus secondary structure of the U3 snoRNA of the Dictyostelia.

D. discoideum, helices M1, M2, M3, and M4 were structurally conserved, while the fungi-specific helix M7 was absent. The loop between helices M6 and M8 seemed to be degenerated with the possibility that M6 and M8 fused into one helix. Besides their conserved secondary structure, most helices are also conserved in sequence in the dictyostelids, with covariation mainly occurring in helix M3 and the lower part of helix M4 (Figure 52B).

Analysis of upstream sequence elements of the dictyostelid U3 snoRNA genes

Even though the U3 snoRNA carries a $m^{2,2,7}G$ cap, a hallmark of Pol II transcription, U3 snoRNA genes across the eukaryotes have no uniform mode of transcription, as Pol II [385, 386] and Pol III [387, 388] products have been described. To gather indications about the responsible RNA polymerase in the Dictyostelia, the immediate upstream sequence elements of all U3 snoRNA genes was analyzed using `MEME` [436]. Using this approach, the highly conserved *D. discoideum* Upstream Sequence Element (DUSE; Figure 53) was found 55 to 57 nt upstream of the predicted transcription start site (TSS). The DUSE motif was discovered earlier in the promoter regions of many (predicted) ncRNAs of *D. discoideum* [346, 511, 512]. The recent bioinformatic classification of Class I RNAs revealed that DUSE is a conserved upstream sequence element across all major group of dictyostelids with the consensus sequence 5′-WMCCAYAA-3′ [513]. The DUSE located to the likely promoter region of the dictyostelid



	DUSE		TATA	
	Consensus	Position	Consensus	Position
D. discoideum	TCCCATAA	-55/-56	TWWAAATA*	-30
D. purpureum	WCCCATAA	-53	TATAWWCA*	-21
D. lacteum	WWCCATAA	-55	missing	
P. pallidum	TCCCATAA	-57	missing	
A. subglosum	TACCATAA	-57	missing	
D. fasciculatum	ACCCATAA	-57	missing	

* *D. discoideum*: 4/7 loci; *D. purpureum*: 5/6 loci

Figure 53. Upstream sequence element found in the dictyostelid U3 snoRNA loci. The phylogenetic tree was generated based on an alignment of the U3 snoRNAs of the different Dictyostelia. Indicated are the consensus sequences of the DUSE and TATA box found in these loci, as well as their position upstream from the start of the gene.

U3 snoRNA genes adhered to this consensus (Dictyostelia: 5´-WHCCATAA-3´). The presence of the DUSE alone, however, did not allow for prediction of the RNA polymerase involved, as it was found in front of *D. discoideum* genes known for Pol II-dependent and Pol III-dependent transcription [514]. Thus, the identification of further upstream sequence elements would be needed for the determination of the RNA polymerase, however, MEME did not uncover any additional common motif. As additional motifs (e.g., TATA boxes) might have been hard to find in the AT-rich intergenic space of the dictyostelids, manual inspection of the 150 nt upstream of the predicted TSS was performed. By this approach, TATA-like sequences were found exclusively in both group 4 dictyostelids, *D. discoideum* and *D. purpureum*, in 5/7 and 5/6 loci, respectively. This sequence was observed at a fixed position independent of the analyzed putative promoter sequence of -30 for *D. discoideum* and -21 for *D. purpureum* (Figure 53 and Figure 54). With these observations, it can be speculated that group 4 dictyostelids rather use Pol II transcription for the U3 snoRNA genes, while other dictyostelids might utilize Pol III transcription. However, all U3 snoRNA loci in this analysis featured a poly(T) stretch ≥ 4 downstream of their presumed end of transcription, which is a common Pol III terminator [515].

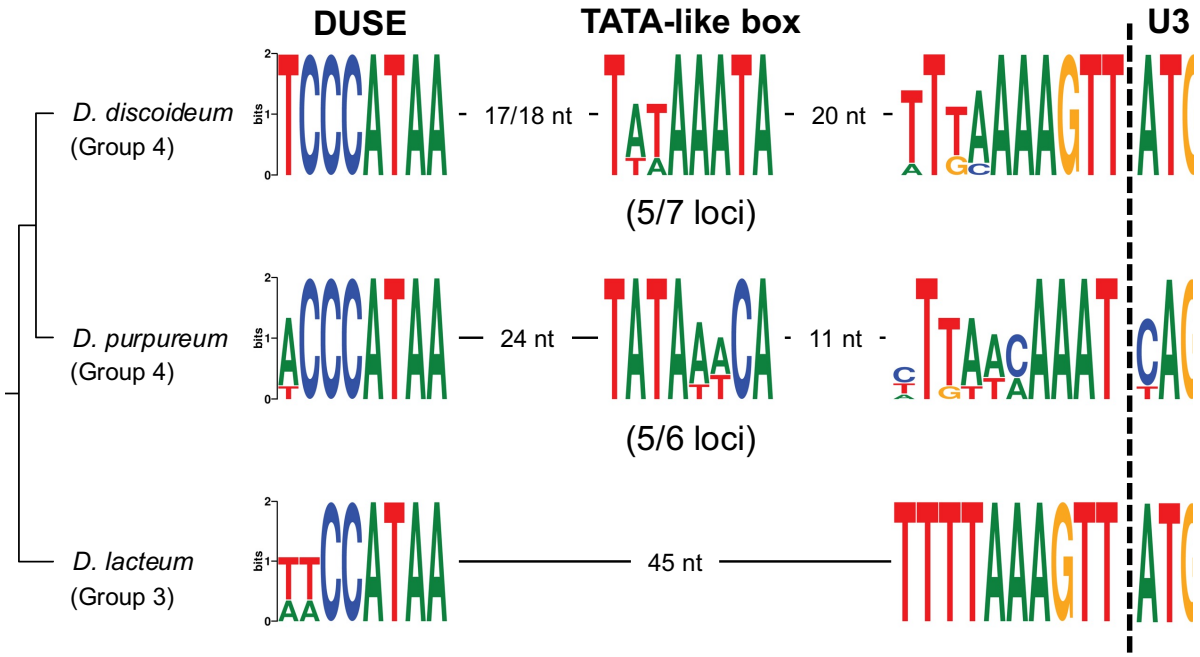


Figure 54. Comparison of upstream sequence elements of dictyostelids possessing more than one U3 snoRNA gene. Shown are the predicted DUSE promoter element, a TATA box-like sequence (if applicable), and ten nucleotides before the start of the U3 snoRNA gene (generated using WebLogo). The dashed line indicates the start of the U3 snoRNA genes.

Discussion

What confers immunity to the amoeba *D. discoideum*?

APMV is degraded in the phagolysosomal pathway

In this study, the question was posed, if *D. discoideum* is infected with amoebal viruses. For this purpose, a protocol for infection was established. To monitor the infection and viral load over time, both, qualitative and quantitative approaches were set up. For the quantitative techniques, a qPCR-based detection system for APMV and TPV was established. Qualitative surveillance of infection was achieved by the optimization of immunostaining of virus-infected *Dictyostelium* cells. Subsequently, the question was posed whether mutant strain could be identified, which are permissive for viral replication or persistence. For this purpose, APMV was selected for initial experiments, and it was observed that the overall number of viral genomes did not increase over time in the AX2 strain (Figure 19). The lack overall genome replication could be explained by a successful defense strategy by the amoeba or, since *D. discoideum* is not the natural host of the virus, some factor essential for the infection was missing. The microscopic investigation APMV uptake into the *Dictyostelium* cells revealed that the virus was surrounded by phagolysosomal markers most of the time (Figure 21). In consequence, viral particles were rarely observed not associated with these markers, suggesting that only few viruses were able to escape into the cytosol. Based on these observations, it was assumed that APMV was likely not able to escape the phagosomal compartment. The decrease in the abundance of viral genomes after 24 h, additionally suggested that the viral particles might be degraded in the phagolysosomal pathway. The phagolysosomal pathway in APMV's natural host *A. polyphaga* is not well described but the same mechanisms of internalization, acidification, and exocytosis of undigested materials seem to be conserved [516, 517]. However, due to the lack of details on *A. polyphaga*'s phagolysosomal pathway like phagosomal pH, the composition of the cocktail of lysosomal enzymes, etc., factors that might be leading to the putative degradation could not be inferred.

External manipulation of phagosomal degradative conditions

Phagocytosis not only provides a feeding mecanism but also represents the first line of defense against pathogens [6, 109, 112]. Also, (amoebal) phagocytosis represents an entry portal for many known bacterial pathogens and many phagosomal proteins have been implied in the defense against these invaders (as reviewed in Dunn *et al.* [109]). To potentially improve the

chances of APMV escaping into the cytosol, the phagolysosomal pathway of *D. discoideum* was manipulated by additives (Figure 25). A recent study described *in vitro* opening of the stargate structure [471], thus, viral particles were primed for core release incubation at pH 2.0 before the experiment. During the former experiment, the pre-treatment of the virus, buffering of the phagosomal pH, or the inhibition of lysosomal proteases were observed to rather enhance the decrease of viral genomes. This might be an indication that the "activation" of the virus and phagosomal escape might be highly orchestrated processes dependent on the order of phagosomal maturation [518]. However, no pH for the acidification of *A. polyphaga*'s phagosome is reported and, thus, this argument remains speculation. It is also conceivable that by enforcing the opening of the stargate, the core is made more accessible for the degradative components of the phagolysosomal pathway. In addition, Schrad *et al.* [471] described that the acid treatment alone did not suffice to completely open the viral capsid. In their *in vitro* approach they found that only the combination of high temperatures and a low pH triggers the full opening of the stargate [471]. Another observation was that the decrease in the number of viral genomes observed upon the addition of protease inhibitors independent of the pre-treatment was higher compared to the untreated samples. This suggested that lysosomal proteases might be somehow involved in the process of virus activation. Concrete experimental evidence and elucidation of the underlying molecular mechanisms, however, is lacking. This could potentially be addressed by screening a mutant library of lysosomal enzymes, yielding the answer to two questions: 1. Which lysosomal enzyme might be involved in virus activation and 2. Which lysosomal enzyme(s) are potentially degrading the virus.

Factors involved in bacterial infections have no effect on the APMV infection

As a model organism for a vast array of cellular processes and host-pathogen interactions, a high number of methods and a complete and annotated genome have been established for *D. discoideum*. This contrasts with any *Acanthamoeba* species, which lack most convenient research techniques, e.g., easy genetic manipulation. In addition to the user-submitted mutant strains at dictyBase (www.dictybase.org), an AX4 mutant library has been made available through dictyBase recently (www.remi-seq.org; [519]). This library was generated by restriction enzyme-mediated integration (REMI) insertions and contains almost every viable knockout mutant possible [519]. In case of the known bacterial pathogen-host interactions, many phagolysosomal factors are involved in the amoebal defense [109] and mutants for these are also readily available. For example, PIPs are important for phagosomal maturation and are shown to be used by bacterial invaders to manipulate this process. The pathogen

L. pneumophila secretes effector proteins that manipulate the identity of the compartment by acquiring PI(4)P, which is usually a marker for membranes of the ER and the *trans*-Golgi apparatus [520]. Proteins involved in the altering of the PIP dynamics like the inositol 5-phosphatase Dd5P4 have been shown to restrict *L. pneumophila* growth [477]. This effect can be alleviated, however, if Dd5P4 was knocked out [477]. The loss of Dd5P4, however, had no impact on the infection of *Dictyostelium* with APMV (Figure 29), suggesting that PIP dynamics might play only a minor role if any.

Metal poisoning or deprivation of essential metal ions is another mechanism in the defense against bacterial pathogens [109]. Divalent metal ions are pumped out of the phagosomal compartment by Nramp1, starving the bacteria in the compartment of essential metals [23]. This process is manipulated by some pathogenic bacteria [109]. For example, the divalent metal transporter Nramp1 is retained at the phagosomal membrane in the absence of the *F. noatunensis* virulence factor IglC [25]. Knockout or displacement of Nramp1 results in steady iron levels in the phagosome, allowing for efficient *F. noatunensis* replication [25]. The second Nramp isoform in *D. discoideum* is NrampB and it is located at the contractile vacuole, an osmoregulatory organelle of the amoeba, and plays a role in the cellular iron homeostasis [410]. Furthermore, loss of NrampB renders *Dictyostelium* more susceptible to *Francisella* infection, however, the underlying mechanisms are not known, yet [25]. In case of APMV infections, no increased persistence or amplification of virus particles was observed (Figure 29). This suggests that either the ions transported by the Nramp transporters did not play a role in APMV infection or that different ions and, therefore, the knockout other ion transporters might increase viral persistence or replication. This includes zinc transporters like zplA-G and zntA-D and copper transporters like p80 and three putative copper-translocating P-type ATPases [109]. In addition to expand the chances of viral persistence or replication, knockouts of metal ion transporters could be combined with external buffering of phagosomal pH with NH_4Cl. Also, PIP signaling mutants could be combined with divalent metal chelators and the NH_4Cl buffering.

In general, the family of receptor tyrosine kinases (RTKs) are involved in signal transduction and regulation in eukaryotic cells [521]. In mammalian phagocytes, phosphorylation of tyrosines is involved in the signaling in the early and later steps of phagocytic uptake [522-524]. RTKs in *D. discoideum* have only been speculated to be connected to the phagosomal manipulation occurring during bacterial infection [415]. At least three RTKs have been identified in the amoeba, however, only vesicle-associated kinase 3 (Vsk3) has been characterized on a molecular level. This protein was found to be located on the phagosomal

membrane with a N-terminal immunoglobulin-like fold domain extending into the lumen of the phagosome [415]. Strains lacking Vsk3 exhibited disturbed fusion of the early phagosome with the lysosome and a reduced phagocytosis rate but phagosomal acidification remained unaffected, suggesting a regulatory role in phagosome maturation [415]. Even though experimental data is absent, Fang *et al.* [415] speculated that Vsk3 or one of the other RTKs might be involved in the process of phagosomal manipulation during bacterial infections. At least for the APMV infection, however, the knockout of the Vsk3 did not have any noticeable effect (Figure 29). A signal in Δ*vsk1* and Δ*vsk2* cells, however, was detectable with an APMV-specific antibody at 24 hpi, while extracellular viruses could not be observed. This could potentially be explained by a delay in the infection cycle. To address whether the viral genome was replicated in the infection, the next step for the Δ*vsk1* and Δ*vsk2* strains would be the quantification of APMV genomes at 0 and 24 hpi using the established qPCR system.

APMV is likely degraded by lysosomal enzymes in *Dictyostelium* wildtype cells

An increase in the persistence of viral genome abundance in the *Dictyostelium* mutants for WshA (Figure 26) and RacH (Figure 28) was observed. Both strains feature strong phenotypes with a vast array of defects [123, 411]. For example, while Δ*wshA* exhibits prolonged phagosomal acidification [123], reneutralization in Δ*racH* occurs faster [411]. These observations suggest that exposure time to acidic conditions did not seem to be the sole factor promoting APMV inactivation. The common defect between both mutants, however, is the strong reduction of proper delivery of the lysosome to and lysosomal fusion with the early phagosome. *Dictyostelium*, as a bacterial predator, possesses a set of highly diverse lysosomal enzymes might also play a role in the inactivation of APMV [525, 526]. Since the amoeba needs to digest its bacterial prey, this cocktail of lysosomal enzymes is also able to break down peptidoglycans [525], which is usually an integral part of the cell wall of Gram-positive bacteria. These peptidoglycans are heteropolymers of alternating β-1,4-linked N-acetylglucosamine (NAG) and N-acetyl muramic acid (NAM) crosslinked by small peptides [527]. Strikingly, the surface layer fibrils of APMV consist of four distinct proteins protected by a layer of peptidoglycans likely making it Gram-positive [31, 34, 67], also explaining the early misconception of this virus as a bacterium. The fibril structure of APMV suggests that the amoeba is equipped to degrade the fibrils and, subsequently, the virus. This might explain the observations made for the Δ*wshA* and Δ*racH* strains. Additionally, it can be speculated that the surface fibrils somehow play a role in phagosomal escape and the productive infection of the host cell. Interestingly, the APMV particles lose their fibers completely and reduce the size of

their genomes by ~16% upon 150 consecutive passages in *A. polyphaga* [528]. This might suggest that the surface fibers are a necessity for the infection of a broad host range and therefore maybe for *D. discoideum*. However, experimental data supporting this notion is lacking. Interestingly, the knockout of AlyA, a lysozyme involved in this process [413, 525], did not have an impact on the reduced number of APMV genomes (Figure 29). This might be explained, however, by the high redundancy of aly proteins in *D. discoideum* [525]. Besides AlyA, the amoeba encodes three other glycosidases involved in the breakdown of peptidoglycans: AlyB, AlyC, and AlyD [525]. Either the knockout of AlyA alone is not sufficient to increase viral persistence or another aly family enzyme is the factor involved in the defense against APMV. This could be addressed in future experiments by multiple knockouts of the Aly family proteins and subsequent infection with APMV. The effect of each mutant strains of the phagolysosomal pathway on APMV abundance is shown in Figure 55.

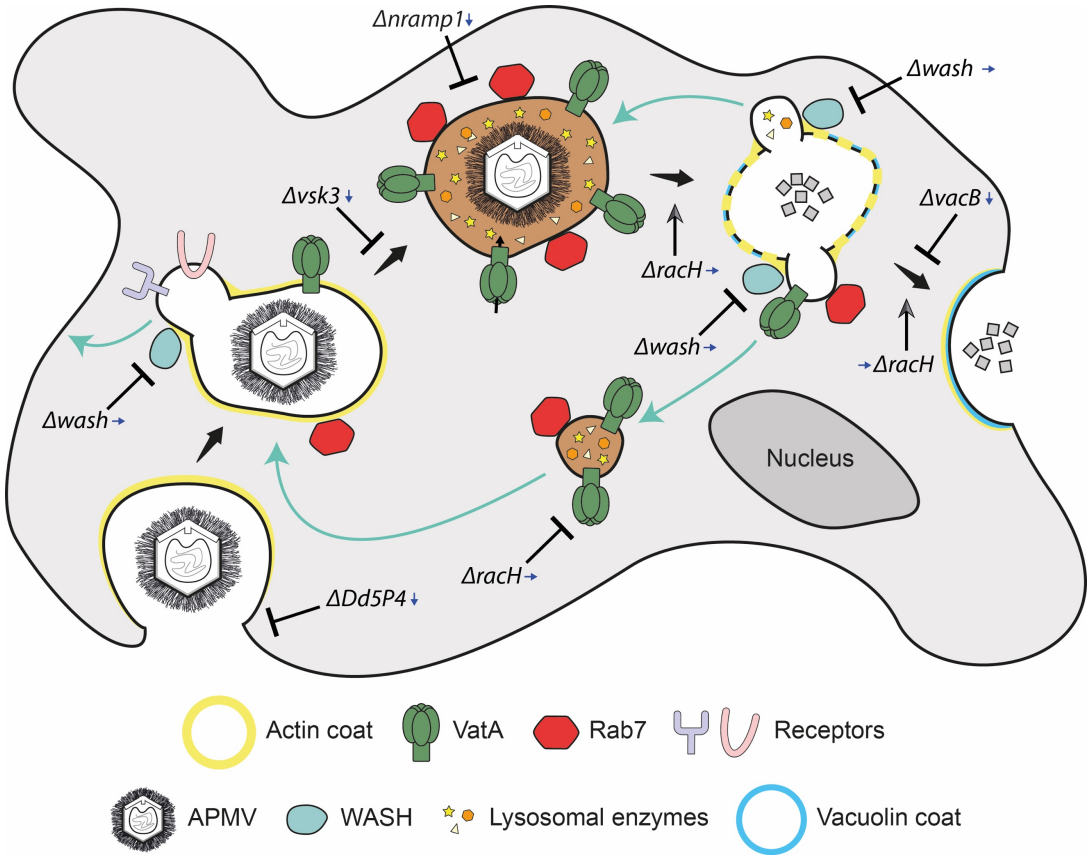


Figure 55. Mutants of the phagolysosomal pathway and their impact on APMV genome abundance. Shown is the phagosomal compartment during phagosome maturation in a model of a *D. discoideum* cell. The phagosomal maturation pathway is indicated by black arrows and recycling of phagosomal components by the WASH retromer is shown with cyan arrows. Acidification of the phagosomal compartment is displayed in brown. Mutants were annotated at the site of their described defect. The nucleus was labelled and shown in dark grey. The direction of the blue arrows indicates viral degradation (↓) or persistence (→). Figure was derived and modified from Dunn *et al.* [109].

Viral infections and defenses of *Dictyostelium*: an outlook

The screening and exploration of the interactions between *D. discoideum* and APMV gave indications, which mechanisms might play a role in the amoebal protection against viral infections. However, the possibility remains that *Dictyostelium* might be susceptible to other viruses. The next paragraphs will be concerned with an outlook on experiments that might elucidate which factors might confer immunity to APMV infection and what steps can be taken to possibly identify a virus that productively infects *Dictyostelium*.

Chasing APMV immunity: investigation of *Dictyostelium*'s defense mechanisms

This work established detection systems to investigate the amoebal resistance against APMV in much more details. The observations made for Δ*wshA* and Δ*racH* indicated that lysosomal enzymes might be a key player involved in the defense against the virus. To investigate which of the lysosomal enzymes might play a part in defense against APMV, the next step would be the screening of the knockouts of the lysosomal enzymes. For the mutant strain for AlyA, an increase the APMV genome abundance during infection was not observed, suggesting that it either might not act alone or that another Aly family enzyme might be involved. It is conceivable, however, that other lysosomal enzymes might play a role in the observed decrease in viral genome abundance, as *Dictyostelium* possesses 22 potential genes for lysosomal enzymes [525]. That question would then be addressed with by screening of their mutants available through the REMI-seq database on dictyBase. In parallel, combinations of external manipulation of phagosomal conditions with mutants of the phagolysosomal pathway could be performed to investigate whether the viral genome persistence is increased.

On a different note, it would also be of interest to establish a high-throughput screening of the APMV infection of *Dictyostelium* mutants, which are available through the REMI-seq database at dictyBase. Since the protocol for the screening infections *D. discoideum* in flat-bottomed 96-well plates was already established by Koller *et al.* [529] and the parameters for APMV infection of the amoeba were optimized in this work, the method could be adapted for the amoebal infection with APMV. Since a viral genome persistence is expected rather than observing cytopathic effects on a large scale like in APMV-infected *Acanthamoeba*, live-dead staining with trypan blue followed by flow cytometry might not be suitable for detection of permissive mutants. This system would require the introduction of a convenient method for DNA isolation from many wells simultaneously to allow for a qPCR-based detection of the

viral genome numbers. The establishment of such a high-throughput platform for detection of viral persistence or replication would help to elucidate the key factors involved in the viral resistance of *D. discoideum*.

Searching genomic sequencing data from environmental samples for NCLDV signatures

Most giant viruses were isolated in waterborne environments, including APMV, TPV, and CeV [31, 58, 416, 417]. Nonetheless, *D. discoideum* is a soil-dwelling eukaryote and might therefore be infected by different giant viruses isolated from the same environment. Indications for this possibility can be inferred by bioinformatic approaches using mainly the program Viral Recall v. 2.0, which constitutes an efficient tool to identify signatures of NCLDV in any kind of -omic data [456]. In brief, Viral Recall identifies NCLDV signatures in four steps: 1.) Prediction of proteins in the given sequence, 2.) Comparison of these predicted proteins to the Pfam and giant virus orthologous group (GVOG) databases based on an HMM search, 3.) Normalization of discrete final scores for each protein, and 4.) Generation of average scores for rolling windows of 15 ORFs [456]. The resulting regions are further analyzed using alignment tools. The two following paragraphs will briefly discuss two possible approaches to identify giant viruses that might infect *Dictyostelium* in its natural environment.

With the rise of modern techniques, sequencing was adopted as a common technique to assess the composition of environmental samples. This common practice includes the publication of the whole raw sequencing data on databases like the Sequence Read Archive, allowing for easy access of large amounts of environmental data sets. For *D. discoideum*, such data sets stem mainly from studies of cooperation and competition of the amoeba like Ostrowski *et al.* [530]. These sequence reads, however, also contain an average of 25% reads that do not account for any known sequence [530]. While these reads might be erroneous or part of genomes of not yet identified species, they possibly also contain NCLDV signatures, which might be detectable by analysis with Viral Recall. The presence of such signatures would then allow for more detailed analyses or even the isolation of one or more novel giant viruses from the original samples with potentially *D. discoideum* as their natural host.

Investigating the amoebal genome for NCLDV signatures

One of the key features of giant viruses are their large genomes with a high abundance of ORFs that are likely acquired by horizontal gene transfer [61, 62, 64, 531]. Due to this process, the acquired ORFs could potentially traced back to their origin, as already demonstrated APMV and protists like *Heterolobosea* and *Kinetoplastida* [62], the phycodnavirus EhV-86 and its

microalgal host *Emiliania huxleyi* [532] and other ORFs of giant viruses. Therefore, it is possible that due to this horizontal gene transfer from host to virus, a search with Viral Recall v. 2.0 might identify regions that at least resemble ORFs transferred from the searched organism to any NCLDV. Even though this does not represent evidence that the transfer occurred in this specific organism or even that an infection took place, but it provides some indication what kind of virus might infect it. Application of this approach to the genome of *D. discoideum* yielded hits to 37 distinct virus and the top ten hits are shown in Table 23. Interestingly, 5/10 viruses (marked in bold in Table 23) were found in metagenomic analyses of environmental forest soil samples [533]. In the same study, the authors found 16 giant viruses in these samples, however, they also indicated already that the bulk metagenomic data contained signatures of many more MCPs, suggesting that a high diversity of giant viruses is yet to discovered [533]. However, all the viruses identified in this study are not associated to a specific host organism and, therefore, cultivation in a lab setting is not possible, yet. Nonetheless, since *D. discoideum* was originally found in the forest soil [3], it still is a possibility that a giant virus infects the amoeba.

Table 23. Analysis of the genome of *D. discoideum* using Viral Recall. Viruses that were found in metagenomic analysis of soil environmental samples are marked in bold. The GenBank accession number for the genomes of the viruses were indicated. The number of homologous protein sequences between the viruses and *Dictyostelium* is shown.

Virus	GenBank Accession	No. of homologous proteins
Klosneuvirus	KY684108.1	12
Catovirus	KY684084.1	10
Terrestrivirus	MK071981.1	9
Homavirus	MK072333.1	9
Tupanvirus Deep Ocean	MF405918.2	9
Tupanvirus Soda Lake	KY523104.2	9
Hokovirus	KY684105.1	8
Hyperionvirus	MK072384.1	7
Edafosvirus	MK072073.1	7
Harvfovirus	MK072249.1	7

The intriguing possibility of viral immunity of *D. discoideum*

The amoeba *D. discoideum* is used as a model organism in the lab setting for more than five decades with the first description of an axenic strain dating back to Sussman and Sussman [502] in 1967. Despite frequent environmental probing by different labs, e.g., the Queller/Strassmann

lab in St. Louis, USA (collecting 1,422 distinct clones from environmental samples), not a single virus has been described to infect the amoeba in the literature in all this time. Even though the discovery of APMV led to the subsequent explosion in the number of detected giant viruses infecting amoeba (reviewed in Diesend *et al.* [58]), still none of the giant virus infecting its distant cousins has been shown to productively infect *Dictyostelium*. This could potentially be explained by the fact that most giant viruses were isolated from a waterborne environment. However, recent studies as Schulz *et al.* [533] suggested that forest soils harbor a diverse array of giant viruses yet to be discovered and characterized, making it unlikely that *D. discoideum* never met a virus. This was substantiated by the Viral Recall analysis performed for *Dictyostelium*'s genome above, nonetheless, this data is implying a productive infection of these viruses of the amoeba. More intriguing is the possibility that *Dictyostelium* might be able to defend itself against this high diversity of giant viruses with potent cell autonomous defense mechanisms. The work performed during this thesis seems to support that notion at least for APMV and the other viruses screened, opening up new possibilities for future experiments dedicated to identifying key players that confer viral immunity in the amoeba.

Ribosome heterogeneity - a common theme among eukaryotes?

Parts of this chapter concerned with 2′-O-methylation are included in the manuscript entitled "Ribosome heterogeneity in Amoebozoa: fractional 2′-O-Methylation in the ribosomal RNA of *Dictyostelium discoideum*", which is published in Scientific Reports.

Ribosome heterogeneity in Amoebozoa

In this study, it was investigated the 2′-O-Me and Ψ landscape of *D. discoideum*'s rRNAs and associated box C/D snoRNAs. To our knowledge, this is the first comprehensive report on this topic for any species from the Amoebozoa, one of five eukaryotic evolutionary supergroups [2]. Using RMS [250], 45 positions that are fully methylated were identified in the rRNAs of the amoeba, and additionally 4 positions that exhibit a sub-stoichiometric 2′-O-Me (Figure 35 and Figure 36). For pseudouridylation, 10 substoichiometrical Ψ sites and 56 fully modified were found using HPS [399]. This indicates that ribosome heterogeneity exists in Amoebozoa. Such variations have been reported for the 2′-O-methylation of nucleic acids making up the translation apparatus in organisms from other evolutionary supergroups, in particular Opisthokonta [250-252, 254, 480], but also in Archaeplastida [483]. Such data is missing for Ψ due to the only recent rise of HPS and the strong limitations of previous techniques in accurately quantifying this modification [534]. Based on the data for 2′-O-Me from a third evolutionary supergroup, the Amoebozoa, and the first reliable data set showing variation in ribosomal Ψ, it can be suggested that ribosome heterogeneity represents a trait common to all eukaryotes.

Ribose methylation and pseudouridylation is thought to occur largely co-transcriptionally [250, 535]. Thus, variation in the levels of this modification could be related to the rDNA organization. In *D. discoideum*, rRNAs are transcribed [15] from extrachromosomal, palindromic elements [12, 13], which is rare but has been described also for, e.g., *D. rerio* [253]. In the amoeba, clusters of the rDNA palindromes can condense into chromosome-like bodies [14] and in these, chemical modifications might be affected by limited accessibility of the nascent transcript to the snoRNPs. Therefore, it is worth noting that the 2′-O-Me and Ψ modification can be introduced equally well on rRNAs transcribed from extrachromosomal rDNA.

A single 2′-O-methylated positions, 26S-A1463, displayed altered RMS scores in the development of the amoeba and between the investigated strains (Figure 35 and Figure 36).

Such changes were also observed in the development of mouse [480] and zebrafish [252]. Further, fractionally methylated sites in rRNA residues in cultured human cells became (close to) fully modified in differentiated tissues [482]. All the aforementioned studies also used RMS, as the preferred high-throughput analysis method of 2′-O-Me patterns, allowing for single nucleotide analysis in a quantitative manner, unlike alternative approaches. This was concluded from a comparative study on rRNA from *Trypanosoma brucei* that further revealed 2′-O-Me patterns, which depended on the living conditions of the parasite [536]. Similar methodological advantages to RMS are also realized by the recently introduced and validated RiboMethSeq tool [465, 537] and the methylated positions reported here for the AX2 strain were at large confirmed independently using this method (Virginie Marchard and Yuri Motorin, personal communication).

For ribosomal Ψ, the modification levels changed more drastically during development of *D. discoideum* than 2′-O-Me patterns do, with the strongest effects seen at 8 h into development (Figure 48). At this time point, 31 of the 66 predicted Ψ sites were differentially modified compared to axenic growth, with most of them showing reduced Ψ levels. Besides the clear notion of a dedicated specialization of ribosomes in development, other possibilities are conceivable and need to be addressed. For example, the expression of ribosomal genes is silenced in the first few hours of development of *D. discoideum* [501]. It is conceivable, that this creates a "lagging" effect of ribosomal modifications, i.e., the sudden need for new ribosomes enhances rRNA processing and the modification machinery cannot keep up. Indeed, previous northern blot analysis of rRNA precursors (data by Sandeep Ojha), the nuclear run-on transcriptions, and U3 snoRNA expression (Figure 50) suggested that rRNA processing might be increased at the 8 and 16 h time points of development compared to axenic cells. As described above for 2′-O-Me, however, a dedicated specialization of ribosome for the developmental cycle is more likely.

Expansion segments in *D. discoideum* had no predicted Ψ sites, with the exception of ES7 in the 26S rRNA (Figure 46). In *S. cerevisiae*, ES7 is involved in rRNA processing [538] and constitutes a binding hub for different classes of proteins, like ribosomal biogenesis factors and aminoacyl tRNA synthetases [539]. Whereas recently, it was observed for human expansion segments, including ES7, to feature complementary sequences to and interactions with mRNAs, a phenomenon usually observed in the positioning of an IRES by the 18S rRNA and the SSU proteins [540]. Since we found a single 2′-O-Me in the same expansion segment (see above), it is tempting to speculate that ES7 fulfills similar functions in the amoeba and that these

modifications stabilize the RNA structure necessary for protein binding. A clear indication for that scenario, however, is missing.

Even though a comparable and reliable quantification of the ribosomal Ψ landscape in other organisms in the context of development and/or environmental factors is still missing, functional assays already described a connection between low pseudouridylation levels and defects and translational fidelity [277, 541-543]. In addition, changes in the Ψ levels of individual positions have been characterized in the past. For example in *S. cerevisiae*, a diauxic shift causes an approximately 2-fold change in levels of 25S-Ψ2314 [544] and the U2 snRNA is inducibly pseudouridylated at positions 56 and 93 upon nutrient deprivation [545]. The recent advent of HPS as a reliable next-generation sequencing approach, will most likely reveal more about (substoichiometric) Ψ and its role in ribosome heterogeneity.

An additional source of ribosome heterogeneity might be the rapid generation of new ribosomes in the first hours of development before aggregation starts. As discussed above, the experiments determining the rDNA transcription, the rRNA precursor accumulation, and the U3 snoRNA expression point towards an increase in rRNA processing between 8 and 16 h (Figure 50). This data is supported by the observation that approximately 75% of the ribosome population is replaced during development [546]. As the expression of ribosomal genes is repressed in the early hours of development [505], it can be speculated that the transcripts for the needed ribosomal proteins are accumulating before their transcription stops. Indeed, previous data showed that at least the five tested ribosomal protein transcripts are accumulated at 8 h (data by Sandeep Ojha), making their translation for the biogenesis of novel ribosomes at least feasible. The new ribosome generation could potentially feature an altered protein composition or - as we observed in RMS and HPS - a different degree and/or set of chemical modifications. In summary, the available data suggests that a specialization of ribosomes occurs in the first few hours of the development of *D. discoideum*.

The snoRNAs of *D. discoideum*

Identification of box C/D and box H/ACA snoRNAs: a cautionary tale

For the majority of 2′-O-methylated rRNA positions, suitable CD RNAs were bioinformatically identified (Figure 29 and Table 16). A subset of 17 such molecules had been reported earlier [346], and here additional 21 novel box C/D snoRNAs with a target in rRNAs were added, plus nine without. Previously, small non-coding RNAs in the amoeba were all called DdR-x (x =

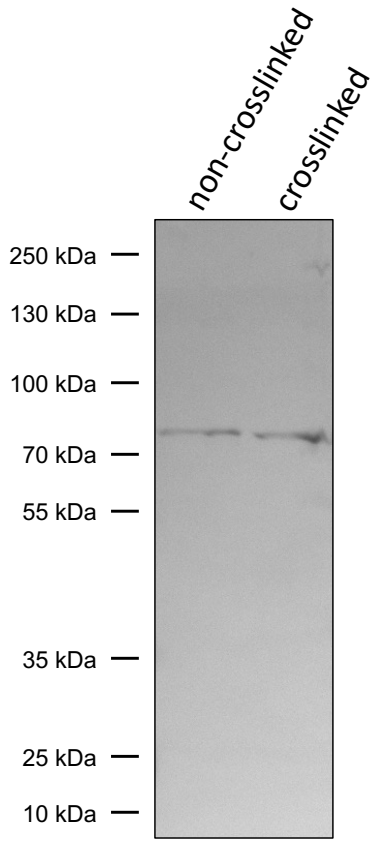


Figure 56. Western blot of (non-)crosslinked AX2 nuclei. Total protein of AX2 was isolated and separated on a 12% PA gel after either UV-crosslinking or no treatment. Marker was the PageRuler Prestained Plus. Detection was performed with α-DKC1 and a secondary AP-conjugated antibody. Nola4 has an expected size of ca. 66 kDa. The size observed was higher, however, this might be due to Nola4 carrying an RNA.

natural number), for ***D**ictyostelium **d**iscoideum* **R**NA [346]. With a functional association, the box C/D snoRNAs with an rRNA target were renamed to CDx (x = natural number), and those without to ORx RNA (for **or**phan).

Unlike for the 2′-O-methylations, no novel box H/ACA snoRNAs besides the single one characterized in an earlier study by Aspegren *et al.* [346] (annotated as sno18) was identified. In cooperation with Christine Gaspin (INRAE institute, Occitanie-Toulouse) the bioinformatic tool snoGPS was employed in tandem with RNAseq analysis, an approach similar to the one which successfully identified the box C/D snoRNAs. However, all these approaches either yielded no candidates or the candidates were not found to be expressed in the RNAseq datasets:

- Searching for canonical box H/ACA snoRNAs based on the predicted targets.
- Searching for individual stems with less stringent conservation of box ACA.
- Searching for less canonical box H/ACA snoRNAs based on the archaeal model.
- Searching for sequences similar to box H/ACA snoRNAs with an AT-rich bias.
- Searching in the RNAseq data for expressed canonical box H/ACA snoRNAs.

It can be excluded that the selected RNAseq datasets are not suitable for this analysis, since sno18 was consistently found with this method. However, it is unclear why all of these approaches failed, since sno18 is a canonical box H/ACA snoRNA [346]. As a next step and to identify the RNAs associated with the snoRNP complex, UV-crosslinked snoRNPs will be isolated, followed by deep sequencing of the contained RNA molecules. Initial experiments with a monoclonal antibody raised against human dyskerin was able to specifically detect the dyskerin-ortholog of the amoeba nola4

(Figure 56). This approach should be able to identify the illusive box H/ACA snoRNAs in *D. discoideum*.

A secondary structure model for the ribosomal RNA in *D. discoideum*

For the localization of the 2′-O-methylated positions a complete model for the secondary structure of the large rRNAs in the amoeba was proposed (Figure 37 and Figure 38), additionally to the partial Cryo-EM structure of the nascent ribosome [486]. This model is based on a homology alignment of rRNA sequences from organisms of two evolutionary supergroups, the Opisthokonta and Archaeplastida [2]. In the rRNA models for the Amoebozoan *D. discoideum*, about half of the 2′-O-methylated nucleotides are found close to the A, P and E sites of the ribosome. The remainder localize either in formally single stranded regions or at the very beginning of helical stems where they presumably fulfil a stabilizing function or support rRNA folding. Our models of the *D. discoideum* rRNAs are greatly supported by the previously introduced Cryo-EM structure of the nascent 60S subunit of *Dictyostelium* [486], that features parts of the proposed structural elements of the 26S rRNA (Figure 38), while the ESs are not covered in this structure.

In *D. discoideum*, the 2′-O-methylated positions U3254 and G3255 on the 26S rRNA are orthologous to the methylated sites U2921 and G2922 in *S. cerevisiae* (Table 16). In yeast, G_m2922 is highly important for the docking of transfer RNAs (tRNA) in the A-site via base pairing with C_{75} in their CCA-tail [259]. This suggests that G_m3255 might fulfill the same function in *Dictyostelium*. U3254 is likely modified by the CD25 RNP (see also below), however, a guide for G_m3255 is missing (Table 16). Intriguingly, position G2922 in *S. cerevisiae* is modified by the SAM-dependent methyltransferase Spb1, independent of a box C/D snoRNA guide [259]. *Dictyostelium*'s genome encodes the homologous *fsjC* gene (http://dictybase.org/gene/DDB_G0284945), and by analogy we hypothesize that its gene product might fulfil the same function as Spb1 in yeast.

The box C/D snoRNA genes

Box C/D snoRNAs in *D. discoideum* are encoded in intergenic regions or as part of introns of protein-coding genes, and in either set-up, they can be generated from mono- or poly-cistronic transcriptional units [343]. The selected set of 38 CD RNAs and their encoding genes display overall features similar to those seen in the original 17 sequences [346]. We found all box C/D snoRNAs in intergenic regions except for CD38, which is encoded in an intron of *DDB_G0283293* (Table S3).

Aspegren *et al.* [346] had reported three bi-cistronic transcriptional units of snoRNAs being expressed in *D. discoideum*. We identified seven additional clusters with two or three box C/D snoRNA genes (Figure 34). One of the tri-cistronic clusters (on chromosome 5; Figure 34), had been reported to contain CD16 and CD5, but the central CD23 gene had not been noticed at the time [346]. A primary transcript of that cluster was not observed, but for the other three originally reported bi-cistrons, primary transcripts had been shown [346]. The former observation might be explicable if the CD16-CD23-CD5 tri-cistron consists of independent mono- or bicistronic transcription units. In summary, box C/D snoRNAs in *D. discoideum* appear predominantly encoded in intergenic regions, half each as mono- and poly-cistrons.

Not only in *D. discoideum*, but also in other species with three-digit intron sizes, like A. thaliana, *S. cerevisiae* or *Schizosaccharomyces pombe* are box C/D snoRNAs largely encoded by independent genes (Table 24). In contrast, in eukaryotes with larger introns such as *D. melanogaster* or *H. sapiens*, snoRNAs are more frequently encoded in the intervening sequences of protein-coding genes [354]. Neither the global abundance of introns in protein-coding genes, nor their frequency/gene appear to be correlated with an "intronization" of the box C/D snoRNA genes (Table 24). Instead, their number appears increased in the analyzed multicellular organisms compared to those that can exist as unicellular species. In the evolutionary tree, the Amoebozoa with *D. discoideum* branched off after the split of the Archaeplastida (*A. thaliana*) and before the separation of the Opisthokonta encompassing as diverse organisms as *D. melanogaster*, *H. sapiens*, *S. cerevisiae*, or *S. pombe* [2]. This current observation might be explained by snoRNA numbers and their intronization having evolved after the split of the individual supergroups to meet the needs of the individual organism.

Table 24. Intronization of box C/D snoRNAs in selected eukaryotic organisms.

Organism	Fraction of genes with introns	Average intron size	Introns per gene	Number of C/D snoRNA genes
D. discoideum	65.2% [a]	129 bp [a]	1.85 [a]	47 [b]
S. pombe	43.0% [547]	107 bp [548]	0.9 [549]	32 [343]
S. cerevisiae	5.0% [550]	256 bp [548]	0.05 [549]	46 [343]
D. melanogaster	80.0% [551]	1639 bp [552]	4.67 [553]	111 [343]
H. sapiens	97.0% [554]	3365 bp [555]	7.8 [556]	275 [343]
A. thaliana	79.1% [557]	168 bp [558]	4.8 [559]	185 [343]

[a] www.dictybase.org

[b] this study

The non-canonical box C/D snoRNA U3 of *D. discoideum* are encoded in seven distinct copies carrying several variable positions (Table 21 and Figure 49). Multiple genomic copies of the U3 snoRNA are usually only observed in higher eukaryotes [467]. Group 3 and group 4 (but not group 1 and group 2A/B) dictyostelids, however, were observed to code for multiple copies of that RNA. The consensus secondary structure of the dictyostelid U3 snoRNA appears largely in line with eukaryotic U3 snoRNA, however, with some degenerated helices. This observation represented the first description of a consensus secondary structure of the evolutionary supergroup of Amoebozoa.

In the analysis of the promoter and terminator regions of all U3 snoRNA loci, the presence of TATA-like sequences in the promoters of the group 4 dictyostelids (Figure 53) was observed but poly(T) stretches were also found in all terminators of the Dictyostelia. The poly(T) termination is a hallmark of eukaryotic Pol III transcription [515], suggesting that Pol III instead of Pol II is utilized in the transcription of the dictyostelid U3 snoRNAs. Adding to this notion, TATA-like boxes were implied in type 3 Pol III transcription [560, 561] and the core of the SphI postoctamer homology (SPH) element (consensus: 5′-ATT**ACCCATAA**TGCATYGCG-3′; [562]) resembles the DUSE consensus 5′-WCCCAYAA-3′. The SPH element is commonly found in the distal sequence elements of Pol III promoters [561], however, it has salso been found closer to the TSS [562]. Altogether, it seems very likely that Pol III transcribes the U3 snoRNAs in the chosen representative of the Dictyostelia, however, experimental verification is still lacking. Nevertheless, the data presented here indicates that an ancestor of the U3 snoRNA gene was present already in the last common ancestor of the dictyostelids.

Interactions of CD RNAs with rRNAs in *D. discoideum*

A productive interaction between a box C/D snoRNA and its target has been suggested to require 7-20 base pairs, thereby allowing for G-U pairs and a few mismatches but excluding bulges [563]. However, only 10 base pairs fit in the substrate binding channel, as observed for an archaeal box C/D sRNP [320]. Overall, the interactions that we are proposing for the CD RNA/rRNA pairs adhere to these rules (Figure 39 and Figure 40). The minimum free energy for the formation of the duplexes (Figure 41) is, however, considerably higher compared to *H. sapiens* [251]. At the same time, the lengths of the interactions do not differ as much. This discrepancy can be attributed to the frequent occurrence of G*U base pairs, the occasional presence of A/C base pairs, and a single G/A mismatch (see below) that are predicted in individual interaction pairs. G*U base pairs have been observed also in analogous pairs of other species [251, 252], and they can be isosteric to Watson-Crick base pairs [489]. However, their

occurrence appears more frequent in the amoeba, and in the extreme case of the CD12/26S-U2580 interaction (Figure 40), 3/9 base pairs are G*U. Also unusual is the G*U interaction exactly at the 2′-O-Me site seen for CD16/17S-G1588 (Figure 40). In three predicted duplexes, we noted an A/C base pair that appeared to be confined to the 6th position upstream of the D box (CD7/26S-G711 and CD23/16S-C3292; Figure 39) or D′ box (CD28/17S-C1715; Figure 40). An A/C interaction can also substitute for a canonical Watson-Crick base pair, if the adenosine is protonated, i.e. A(+)/C [321]. Distinct from these is the single G/A mismatch seen in the CD29/U1264 pair (Figure 39) that is likely to cause structural perturbations in the interaction, which possibly is counteracted by the overall 13 base pairs surrounding the mismatch. As had been observed before [252], the methylated position 17S-G1589 appears to be guided by the +6 position of CD16 (Figure 40). We noted that non-Watson-Crick interactions occur in all predicted pairs that result in a fractional methylation (Table 16, Figure 39 and Figure 40). However, the overall strength (or weakness) of the CD RNA/rRNA interaction in *D. discoideum* does not appear to correlate with the RMS score (Figure 56), similar to observations made in human cells [251]. The lower free energies observed for the resulting duplexes (Figure 41E) might rather be explained by the lower optimal growth temperature of 21°C of *D. discoideum* [5], compared to yeast or humans. At this temperature, the inferred stabilities apparently warrant appropriate 2′-O-Me levels in the rRNAs in the amoeba (Figure 35).

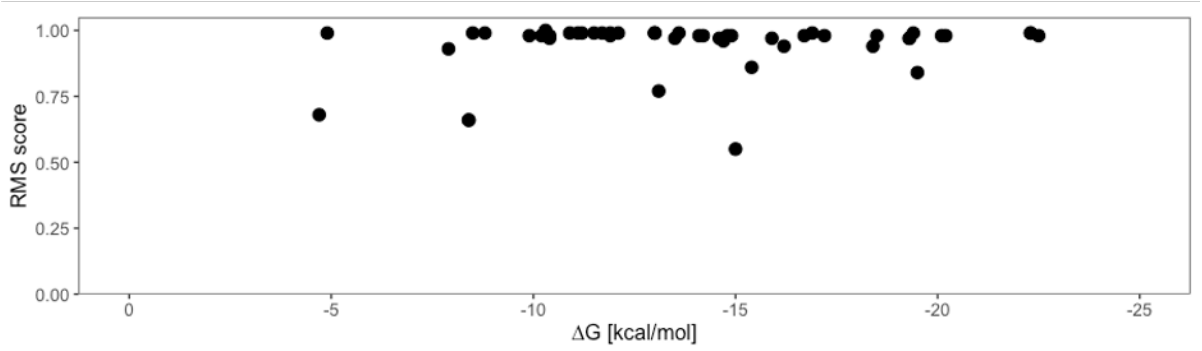


Figure 57. 2D plot of RMS score against MFE. Shown is the minimal free energy (ΔG in kcal/mol) and the RMS score of each CD RNA/rRNA interaction in *D. discoideum*.

Features of the box C/D snoRNAs

The mature box C/D snoRNAs in *D. discoideum* exhibit generally established characteristics of this class of ncRNAs (Figure 29). A stable terminal stem, however, is absent in about half of the mature box C/D snoRNAs (Table 16). Such stems are considered important for the recognition by the box C/D snoRNA processing machinery [303, 304, 358, 564, 565]. In

H. sapiens or *Xenopus laevis*, a lack of the terminal stem in mature snoRNAs appears to be compensated by self-complementary sequences in their precursors [566, 567]. This allows for productive interactions with the processing machinery, upon which these sequences are thought to be removed [303, 304, 564]. Also in *D. discoideum*, complementary stretches can be found up- and downstream of some box C/D snoRNAs without a terminal stem (data not shown). Therefore, we speculate that these sequences might be present in presumed precursor molecules.

D. discoideum CD RNAs are predicted to use the antisense elements associated with the weakly conserved D′ box sequences more frequently than those with the highly conserved D boxes (Figure 41B, C). The latter form, together with the in *D. discoideum* equally conserved C boxes, the terminal k-turn structure (Figure 29), which is essential for maturation and assembly of the box C/D snoRNP complexes [310, 568]. To some extent similar, a preferred usage of the D′ boxes in guiding 2′-O-Me to rRNA targets has also been reported for *H. sapiens* and *D. rerio* [251, 252]. These studies revealed that in humans, the box C′ and D′ sequences displayed a considerably stronger conservation than seen for the amoeba, while in zebrafish box D′ was also less conserved and box C′ appeared degenerated.

Seven CD RNAs of *D. discoideum* are predicted to utilize both antisense elements (Table 18), with no paralogs or other box C/D snoRNAs known to be able to target the associated rRNA positions. At present, it is unknown, whether an interaction of both antisense elements with the target RNA(s) takes place simultaneously or sequentially. For *S. cerevisiae*, a simultaneous usage of both the antisense elements upstream the D and D′ boxes has been proposed, which might bring distant parts of the rRNA structure into proximity, thereby facilitating ribosomal maturation [569, 570]. We wondered whether a similar situation might exist for “dual-use” CD RNAs in the amoeba. Since only a partial structure is available for the nascent 60S ribosomal subunit of *Dictyostelium* [486], we inferred positions not included in that structure by homology to the human ribosome (PDB accession: 4UG0) [571]. Positions targeted by CD1, CD7 and CD19 (Table 18) were not considered, as no orthologous methylated sites were found in other species (Table 18). CD25 of *D. discoideum* targets 17S-A612 and 26S-U3254 and the orthologous positions 18S-A668 and 28S-U4468 in the *H. sapiens* ribosome are around 100 Å apart, indicating sequential modification. Despite being distant in sequence, A1370 in helix H39 and G2952 in helix H80, which are both predicted targets of CD13 (Figure 58A), lie only 16.7 Å apart in the available structure [486] of the *D. discoideum* 60S subunit (Figure 58B, C). That structure describes the large subunit at a late stage of maturation. It contains already helices H39 and H80, suggesting that the 2′-O-Me (not featured in the structure) must have taken place,

as it requires the accessibility of the target sequences. We also cannot exclude that CD13 binds its targets after they reach proximity (Figure 57). It is tempting to speculate, however, that the CD RNA actually might first spatially orient the target positions, then trigger their methylation, before the helices finally form. This would be supported by similar reports from *S. cerevisiae* [569, 570]. Notably, in other species [242], the orthologous nucleotides are part of the PTC, with G2952 being directly involved in the interaction with the CCA-tail of the tRNA residing

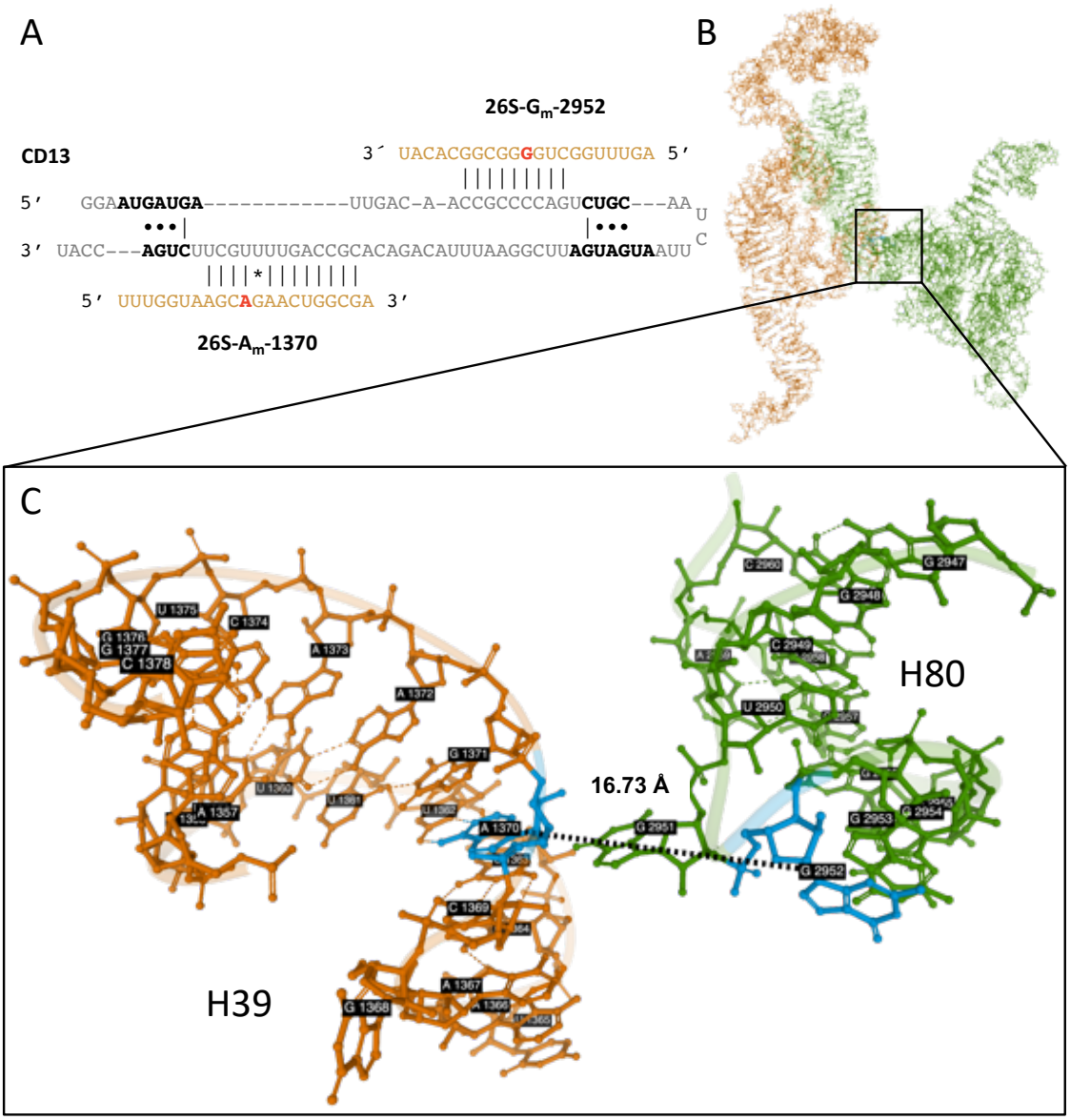


Figure 58. A model on the function of CD13 in guiding 2´-O-Me at two positions in the 26S rRNA. (A) Binary secondary structure of CD13 bound to positions A1370 and G2952 in the 26S rRNA of *D. discoideum*. (B) Scheme of relevant structure parts of the nascent 60S ribosomal subunit of *D. discoideum* (PDB accession: 5AN9) determined at 3.3 Å resolution via cryo-EM [486]. Domain II is displayed in orange and domain V in green (cf. Figure 4). (C) Close vicinity (16.7 Å) of nucleotide A1370 in helix H39 and nucleotide G2952 in helix H80 (both positions colored in blue).

in the ribosomal P site. The two predicted 26S rRNA targets of CD15 and CD19 (Table 18) are so close that a simultaneous occupation of both positions would appear sterically challenging, if not impossible. On the other hand, it seems feasible that CD1 and CD8 might interact with their respective two predicted 17S positions (Table 18) given their spacing. Thus, a simultaneous interaction with the two target sites appears unlikely for some of the "dual use" CD RNAs, but conceivable for others (CD1, CD8 or CD13).

Alternative functions of *D. discoideum* box C/D snoRNAs?

We noted that a substantial set of 22 box C/D snoRNAs are differentially accumulated in the development of the amoeba compared to axenic growth, however, without manifesting in altered 2′-O-Me levels at the targeted positions (Figure 42). This indicates that the amounts of CD RNAs are under either condition sufficient to warrant the appropriate 2′-O-Me levels (Figure 43C). Changes in the level of individual CD RNAs during development of the amoeba had already been observed in northern blots, e.g. for CD9, CD13 or CD15 [346]. This is similar to data from *D. melanogaster* [572] and *D. rerio* [252]. In the absence of an influence on 2′-O-Me levels in the amoeba (Figure 43D), developmental changes of many box C/D snoRNAs might instead point towards other physiological roles. Established is an alternative function as small Cajal Body RNAs (scaRNAs), which are structurally similar to box C/D snoRNAs, carrying an additional CAB box motif, but guide the sequence-specific methylation of small nuclear RNAs (reviewed for example in [299, 573]). Also, some box C/D snoRNAs are involved in the processing of precursor rRNA molecules in a variety of organisms (summarized in [568]). While 2′-O-Me in tRNA is usually introduced by specialized stand-alone methyltransferases, e.g. [574], certain positions are also guided by specific box C/D snoRNAs (reviewed in [299]), either alone or together with a dedicated box C/D scaRNA, like in the case of the wobble cytidine 34 of human $tRNA^{Met}$ [575]. Further functions that are conceivable also for *D. discoideum* box C/D snoRNAs encompass rRNA acetylation [569, 576], regulation of 3′ pre-mRNA processing [577, 578] or even the generation of small, sno-derived RNAs that might have regulatory functions, as described for other organisms [579-581]. Future work will show whether these possible functions are realized in *D. discoideum* by any of the OR RNAs or those CD RNAs, in which one antisense sequence lacks an identified rRNA target.

Bibliography

1. Maeda, Y. and Chida, J. (2013) Control of Cell Differentiation by Mitochondria, Typically Evidenced in Dictyostelium Development. *Biomolecules*, **3**.
2. Adl, S.M., Simpson, A.G., Lane, C.E., Lukes, J., Bass, D., Bowser, S.S., Brown, M.W., Burki, F., Dunthorn, M., Hampl, V. *et al.* (2012) The revised classification of eukaryotes. *J Eukaryot Microbiol*, **59**, 429-493.
3. Raper, K.B. (1935) Dictyostelium discoideum, a new species of slime mold from decaying forest leaves. *J. agri.res. (Wash. D.C., Dept. of Agriculture, G.P.O. [distributor], 1935.*, **50**, 135-147.
4. Watts, D.J. and Ashworth, J.M. (1970) Growth of myxameobae of the cellular slime mould Dictyostelium discoideum in axenic culture. *The Biochemical journal*, **119**, 171-174.
5. Fey, P., Kowal, A.S., Gaudet, P., Pilcher, K.E. and Chisholm, R.L. (2007) Protocols for growth and development of Dictyostelium discoideum. *Nat Protoc*, **2**, 1307-1316.
6. Bozzaro, S. (2013) The model organism Dictyostelium discoideum. *Methods in molecular biology (Clifton, N.J.)*, **983**, 17–37.
7. Glöckner, G., Szafranski, K., Winckler, T., Dingermann, T., Quail, M.A., Cox, E., Eichinger, L., Noegel, A.A. and Rosenthal, A. (2001) The complex repeats of Dictyostelium discoideum. *Genome Res*, **11**, 585-594.
8. Wiegand, S., Meier, D., Seehafer, C., Malicki, M., Hofmann, P., Schmith, A., Winckler, T., Földesi, B., Boesler, B., Nellen, W. *et al.* (2014) The *Dictyostelium discoideum* RNA-dependent RNA polymerase RrpC silences the centromeric retrotransposon DIRS-1 post-transcriptionally and is required for the spreading of RNA silencing signals. *Nucleic Acids Research*, **42**, 3330–3345.
9. Schmith, A., Spaller, T., Gaube, F., Fransson, Å., Boesler, B., Ojha, S., Nellen, W., Hammann, C., Söderbom, F. and Winckler, T. (2015) A host factor supports retrotransposition of the TRE5-A population in *Dictyostelium* cells by suppressing an Argonaute protein. *Mobile DNA*, **6**, 14.
10. Malicki, M., Spaller, T., Winckler, T. and Hammann, C. (2020) DIRS retrotransposons amplify via linear, single-stranded cDNA intermediates. *Nucleic Acids Research*, **48**, 4230-4243.
11. Kessin, R.H. (2001) *Dictyostelium – Evolution, Cell Biology and the Development of Multicellularity*. Cambridge University Press, Cambridge.
12. Cockburn, A.F., Newkirk, M.J. and Firtel, R.A. (1976) Organization of the ribosomal RNA genes of Dictyostelium discoideum: mapping of the nontranscribed spacer regions. *Cell*, **9**, 605-613.
13. Maizels, N. (1976) Dictyostelium 17S, 25S, and 5S rDNAs lie within a 38,000 base pair repeated unit. *Cell*, **9**, 431-438.
14. Sucgang, R., Chen, G., Liu, W., Lindsay, R., Lu, J., Muzny, D., Shaulsky, G., Loomis, W., Gibbs, R. and Kuspa, A. (2003) Sequence and structure of the extrachromosomal palindrome encoding the ribosomal RNA genes in Dictyostelium. *Nucleic Acids Res*, **31**, 2361-2368.
15. Boesler, C., Kruse, J., Soderbom, F. and Hammann, C. (2011) Sequence and generation of mature ribosomal RNA transcripts in Dictyostelium discoideum. *The Journal of biological chemistry*, **286**, 17693–17703.
16. Cosson, P., Zulianello, L., Join-Lambert, O., Faurisson, F., Gebbie, L., Benghezal, M., Van Delden, C., Curty, L.K. and Köhler, T. (2002) Pseudomonas aeruginosa virulence analyzed in a Dictyostelium discoideum host system. *J Bacteriol*, **184**, 3027-3033.
17. Pukatzki, S., Kessin, R.H. and Mekalanos, J.J. (2002) The human pathogen Pseudomonas aeruginosa utilizes conserved virulence pathways to infect the social amoeba Dictyostelium discoideum. *Proc Natl Acad Sci U S A*, **99**, 3159-3164.
18. Hägele, S., Köhler, R., Merkert, H., Schleicher, M., Hacker, J. and Steinert, M. (2000) Dictyostelium discoideum: a new host model system for intracellular pathogens of the genus Legionella. *Cell Microbiol*, **2**, 165-171.
19. Solomon, J.M., Rupper, A., Cardelli, J.A. and Isberg, R.R. (2000) Intracellular growth of Legionella pneumophila in Dictyostelium discoideum, a system for genetic analysis of host-pathogen interactions. *Infect Immun*, **68**, 2939-2947.
20. Pukatzki, S., Ma, A.T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W.C., Heidelberg, J.F. and Mekalanos, J.J. (2006) Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. *Proc Natl Acad Sci U S A*, **103**, 1528-1533.
21. Hagedorn, M., Rohde, K.H., Russell, D.G. and Soldati, T. (2009) Infection by tubercular mycobacteria is spread by nonlytic ejection from their amoeba hosts. *Science*, **323**, 1729-1733.
22. Hagedorn, M. and Soldati, T. (2007) Flotillin and RacH modulate the intracellular immunity of Dictyostelium to Mycobacterium marinum infection. *Cell Microbiol*, **9**, 2716-2733.

23. Peracino, B., Wagner, C., Balest, A., Balbo, A., Pergolizzi, B., Noegel, A.A., Steinert, M. and Bozzaro, S. (2006) Function and mechanism of action of Dictyostelium Nramp1 (Slc11a1) in bacterial infection. *Traffic*, **7**, 22-38.
24. Solomon, J.M., Leung, G.S. and Isberg, R.R. (2003) Intracellular replication of Mycobacterium marinum within Dictyostelium discoideum: efficient replication in the absence of host coronin. *Infect Immun*, **71**, 3578-3586.
25. Brenz, Y., Ohnezeit, D., Winther-Larsen, H.C. and Hagedorn, M. (2017) Nramp1 and NrampB Contribute to Resistance against Francisella in Dictyostelium. *Front Cell Infect Microbiol*, **7**, 282.
26. Lampe, E.O., Brenz, Y., Herrmann, L., Repnik, U., Griffiths, G., Zingmark, C., Sjöstedt, A., Winther-Larsen, H.C. and Hagedorn, M. (2015) Dissection of Francisella-Host Cell Interactions in Dictyostelium discoideum. *Appl Environ Microbiol*, **82**, 1586-1598.
27. Colucci, A.M., Peracino, B., Tala, A., Bozzaro, S., Alifano, P. and Bucci, C. (2008) Dictyostelium discoideum as a model host for meningococcal pathogenesis. *Med Sci Monit*, **14**, Br134-140.
28. Aubert, D.F., Flannagan, R.S. and Valvano, M.A. (2008) A novel sensor kinase-response regulator hybrid controls biofilm formation and type VI secretion system activity in Burkholderia cenocepacia. *Infect Immun*, **76**, 1979-1991.
29. Jia, K., Thomas, C., Akbar, M., Sun, Q., Adams-Huet, B., Gilpin, C. and Levine, B. (2009) Autophagy genes protect against Salmonella typhimurium infection and mediate insulin signaling-regulated pathogen resistance. *Proc Natl Acad Sci U S A*, **106**, 14564-14569.
30. Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M. and Claverie, J.-M. (2004) The 1.2-megabase genome sequence of Mimivirus. *Science (New York, N.Y.)*, **306**, 1344–1350.
31. La Scola, B., Audic, S., Robert, C., Jungang, L., Lamballerie, X.d., Drancourt, M., Birtles, R., Claverie, J.-M. and Raoult, D. (2003) A giant virus in amoebae. *Science (New York, N.Y.)*, **299**, 2033.
32. Raoult, D. (2013) TRUC or the Need for a New Microbial Classification. *Intervirology*, **56**, 349–353.
33. Lwoff, A. (1957) The concept of virus. *J Gen Microbiol*, **17**, 239–253.
34. Raoult, D., La Scola, B. and Birtles, R. (2007) The discovery and characterization of Mimivirus, the largest known virus and putative pneumonia agent. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*, **45**, 95–102.
35. Birtles, R.J., Rowbotham, T.J., Storey, C., Marrie, T.J. and Raoult, D. (1997) Chlamydia-like obligate parasite of free-living amoebae. *The Lancet*, **349**, 925–926.
36. Dornas, F.P., Rodrigues, F.P., Boratto, P.V.M., Silva, L.C.F., Ferreira, P.C.P., Bonjardim, C.A., Trindade, G.S., Kroon, E.G., La Scola, B. and Abrahão, J.S. (2014) Mimivirus circulation among wild and domestic mammals, Amazon Region, Brazil. *Emerg Infect Dis*, **20**, 469–472.
37. Boughalmi, M., Pagnier, I., Aherfi, S., Colson, P., Raoult, D. and La Scola, B. (2013) First isolation of a giant virus from wild Hirudo medicinalis leech: Mimiviridae isolation in Hirudo medicinalis. *Viruses*, **5**, 2920–2930.
38. Andrade, K.R., Boratto, P.P.V.M., Rodrigues, F.P., Silva, L.C.F., Dornas, F.P., Pilotto, M.R., La Scola, B., Almeida, G.M.F., Kroon, E.G. and Abrahão, J.S. (2015) Oysters as hot spots for mimivirus isolation. *Archives of virology*, **160**, 477–482.
39. Silva, L.C.F., Almeida, G.M.F., Oliveira, D.B., Dornas, F.P., Campos, R.K., La Scola, B., Ferreira, P.C.P., Kroon, E.G. and Abrahão, J.S. (2014) A resourceful giant: APMV is able to interfere with the human type I interferon system. *Microbes and infection*, **16**, 187–195.
40. Popgeorgiev, N., Boyer, M., Fancello, L., Monteil, S., Robert, C., Rivet, R., Nappez, C., Azza, S., Chiaroni, J., Raoult, D. *et al.* (2013) Marseillevirus-like virus recovered from blood donated by asymptomatic humans. *The Journal of infectious diseases*, **208**, 1042–1050.
41. La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., Merchat, M., Suzan-Monti, M., Forterre, P., Koonin, E. *et al.* (2008) The virophage as a unique parasite of the giant mimivirus. *Nature*, **455**, 100–104.
42. Colson, P., Pagnier, I., Yoosuf, N., Fournous, G., La Scola, B. and Raoult, D. (2013) "Marseilleviridae", a new family of giant viruses infecting amoebae. *Archives of virology*, **158**, 915–920.
43. Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., Robert, C., Azza, S., Sun, S., Rossmann, M.G. *et al.* (2009) Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 21848–21853.
44. Colson, P., La Scola, B., Levasseur, A., Caetano-Anollés, G. and Raoult, D. (2017) Mimivirus: leading the way in the discovery of giant viruses of amoebae. *Nature reviews. Microbiology*, **15**, 243–254.
45. Thomas, V., Bertelli, C., Collyn, F., Casson, N., Telenti, A., Goesmann, A., Croxatto, A. and Greub, G. (2011) Lausannevirus, a giant amoebal virus encoding histone doublets. *Environmental microbiology*, **13**, 1454–1466.

46. La Scola, B., Campocasso, A., N'Dong, R., Fournous, G., Barrassi, L., Flaudrops, C. and Raoult, D. (2010) Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. *Intervirology*, **53**, 344–353.
47. Lagier, J.-C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., Bittar, F., Fournous, G., Gimenez, G., Maraninchi, M. *et al.* (2012) Microbial culturomics: paradigm shift in the human gut microbiome study. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, **18**, 1185–1193.
48. Yutin, N., Wolf, Y.I., Raoult, D. and Koonin, E.V. (2009) Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virol J*, **6**, 223.
49. Yutin, N. and Koonin, E.V. (2012) Hidden evolutionary complexity of Nucleo-Cytoplasmic Large DNA viruses of eukaryotes. *Virol J*, **9**, 161.
50. Raoult, D. and Forterre, P. (2008) Redefining viruses: lessons from Mimivirus. *Nature reviews. Microbiology*, **6**, 315–319.
51. Khalil, J.Y.B., Andreani, J. and La Scola, B. (2016) Updating strategies for isolating and discovering giant viruses. *Current opinion in microbiology*, **31**, 80–87.
52. Khalil, J.Y.B., Robert, S., Reteno, D.G., Andreani, J., Raoult, D. and La Scola, B. (2016) High-Throughput Isolation of Giant Viruses in Liquid Medium Using Automated Flow Cytometry and Fluorescence Staining. *Frontiers in microbiology*, **7**, 26.
53. Pagnier, I., Reteno, D.-G.I., Saadi, H., Boughalmi, M., Gaia, M., Slimani, M., Ngounga, T., Bekliz, M., Colson, P., Raoult, D. *et al.* (2013) A decade of improvements in Mimiviridae and Marseilleviridae isolation from amoeba. *Intervirology*, **56**, 354–363.
54. Yoosuf, N., Yutin, N., Colson, P., Shabalina, S.A., Pagnier, I., Robert, C., Azza, S., Klose, T., Wong, J., Rossmann, M.G. *et al.* (2012) Related giant viruses in distant locations and different habitats: Acanthamoeba polyphaga moumouvirus represents a third lineage of the Mimiviridae that is close to the megavirus lineage. *Genome biology and evolution*, **4**, 1324–1330.
55. Arslan, D., Legendre, M., Seltzer, V., Abergel, C. and Claverie, J.-M. (2011) Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 17486–17491.
56. Dereeper, A., Audic, S., Claverie, J.-M. and Blanc, G. (2010) BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC evolutionary biology*, **10**, 8.
57. Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-F., Guindon, S., Lefort, V., Lescot, M. *et al.* (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research*, **36**, W465-469.
58. Diesend, J., Kruse, J., Hagedorn, M. and Hammann, C. (2018) Amoebae, Giant Viruses, and Virophages Make Up a Complex, Multilayered Threesome. *Frontiers in Cellular and Infection Microbiology*, **7**.
59. Legendre, M., Santini, S., Rico, A., Abergel, C. and Claverie, J.-M. (2011) Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. *Virol J*, **8**, 99.
60. Desnues, C., La Scola, B., Yutin, N., Fournous, G., Robert, C., Azza, S., Jardot, P., Monteil, S., Campocasso, A., Koonin, E.V. *et al.* (2012) Provirophages and transpovirons as the diverse mobilome of giant viruses. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 18078–18083.
61. Suhre, K. (2005) Gene and genome duplication in Acanthamoeba polyphaga Mimivirus. *Journal of virology*, **79**, 14095–14101.
62. Moreira, D. and Brochier-Armanet, C. (2008) Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes. *BMC evolutionary biology*, **8**, 12.
63. Forterre, P. (2010) Giant viruses: conflicts in revisiting the virus concept. *Intervirology*, **53**, 362–378.
64. Filée, J., Siguier, P. and Chandler, M. (2007) I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses. *Trends in genetics : TIG*, **23**, 10–15.
65. Renesto, P., Abergel, C., Decloquement, P., Moinier, D., Azza, S., Ogata, H., Fourquet, P., Gorvel, J.-P. and Claverie, J.-M. (2006) Mimivirus giant particles incorporate a large fraction of anonymous and unique gene products. *Journal of virology*, **80**, 11678–11685.
66. Suhre, K., Audic, S. and Claverie, J.-M. (2005) Mimivirus gene promoters exhibit an unprecedented conservation among all eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 14689–14693.
67. Xiao, C., Kuznetsov, Y.G., Sun, S., Hafenstein, S.L., Kostyuchenko, V.A., Chipman, P.R., Suzan-Monti, M., Raoult, D., McPherson, A. and Rossmann, M.G. (2009) Structural studies of the giant mimivirus. *PLoS Biol*, **7**, e92.
68. Kuznetsov, Y.G., Klose, T., Rossmann, M. and McPherson, A. (2013) Morphogenesis of mimivirus and its viral factories: an atomic force microscopy study of infected cells. *Journal of virology*, **87**, 11200–11213.
69. Rodrigues, R.A.L., dos Santos Silva, L.K., Dornas, F.P., Oliveira, D.B.d., Magalhães, T.F.F., Santos, D.A., Costa, A.O., Macêdo Farias, L.d., Magalhães, P.P., Bonjardim, C.A. *et al.* (2015) Mimivirus Fibrils

Are Important for Viral Attachment to the Microbial World by a Diverse Glycoside Interaction Repertoire. *Journal of virology*, **89**, 11812–11819.
70. Ghigo, E., Kartenbeck, J., Lien, P., Pelkmans, L., Capo, C., Mege, J.-L. and Raoult, D. (2008) Ameobal pathogen mimivirus infects macrophages through phagocytosis. *PLoS pathogens*, **4**, e1000087.
71. Mutsafi, Y., Zauberman, N., Sabanay, I. and Minsky, A. (2010) Vaccinia-like cytoplasmic replication of the giant Mimivirus. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 5978–5982.
72. Zauberman, N., Mutsafi, Y., Halevy, D.B., Shimoni, E., Klein, E., Xiao, C., Sun, S. and Minsky, A. (2008) Distinct DNA exit and packaging portals in the virus Acanthamoeba polyphaga mimivirus. *PLoS Biol*, **6**, e114.
73. Broyles, S.S. (2003) Vaccinia virus transcription. *Journal of General Virology*, **84**, 2293-2303.
74. Claverie, J.-M., Grzela, R., Lartigue, A., Bernadac, A., Nitsche, S., Vacelet, J., Ogata, H. and Abergel, C. (2009) Mimivirus and Mimiviridae: Giant viruses with an increasing number of potential hosts, including corals and sponges. *Journal of Invertebrate Pathology*, **101**, 172-180.
75. Mutsafi, Y., Fridmann-Sirkis, Y., Milrot, E., Hevroni, L. and Minsky, A. (2014) Infection cycles of large DNA viruses: emerging themes and underlying questions. *Virology*, **466-467**, 3–14.
76. Suzan-Monti, M., La Scola, B., Barrassi, L., Espinosa, L. and Raoult, D. (2007) Ultrastructural characterization of the giant volcano-like virus factory of Acanthamoeba polyphaga Mimivirus. *PLoS One*, **2**, e328.
77. Mutsafi, Y., Shimoni, E., Shimon, A. and Minsky, A. (2013) Membrane assembly during the infection cycle of the giant Mimivirus. *PLoS pathogens*, **9**, e1003367.
78. Resch, W., Hixson, K.K., Moore, R.J., Lipton, M.S. and Moss, B. (2007) Protein composition of the vaccinia virus mature virion. *Virology*, **358**, 233-247.
79. Arantes, T.S., Rodrigues, R.A.L., dos Santos Silva, L.K., Oliveira, G.P., Souza, H.L.d., Khalil, J.Y.B., Oliveira, D.B.d., Torres, A.A., da Silva, L.L., Colson, P. *et al.* (2016) The Large Marseillevirus Explores Different Entry Pathways by Forming Giant Infectious Vesicles. *Journal of virology*, **90**, 5246–5255.
80. Oh, J. and Broyles, S.S. (2005) Host cell nuclear proteins are recruited to cytoplasmic vaccinia virus replication complexes. *Journal of virology*, **79**, 12852-12860.
81. Krupovic, M. and Cvirkaite-Krupovic, V. (2011) Virophages or satellite viruses? *Nature Reviews Microbiology*, **9**, 762.
82. Koonin, E.V. and Krupovic, M. (2017) Polintons, virophages and transpovirons: a tangled web linking viruses, transposons and immunity. *Current Opinion in Virology*, **25**, 7-15.
83. Blanc, G., Gallot-Lavallée, L. and Maumus, F. (2015) Provirophages in the Bigelowiella genome bear testimony to past encounters with giant viruses. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, E5318-E5326.
84. Desnues, C. and Raoult, D. (2012) Virophages question the existence of satellites. *Nature Reviews Microbiology*, **10**, 234.
85. Fischer, M.G. (2011) Sputnik and Mavirus: more than just satellite viruses. *Nature Reviews Microbiology*, **10**, 78.
86. Krupovic, M., Kuhn, J.H. and Fischer, M.G. (2016) A classification system for virophages and satellite viruses. *Archives of Virology*, **161**, 233-247.
87. Gaia, M., Pagnier, I., Campocasso, A., Fournous, G., Raoult, D. and La Scola, B. (2013) Broad Spectrum of Mimiviridae Virophage Allows Its Isolation Using a Mimivirus Reporter. *PLOS ONE*, **8**, e61912.
88. Ogata, H. and Claverie, J.-M. (2008) How to Infect a Mimivirus. *Science*, **321**, 1305.
89. Gaia, M., Benamar, S., Boughalmi, M., Pagnier, I., Croce, O., Colson, P., Raoult, D. and La Scola, B. (2014) Zamilon, a novel virophage with Mimiviridae host specificity. *PLoS One*, **9**, e94923.
90. Campos, R.K., Boratto, P.V., Assis, F.L., Aguiar, E.R.G.R., Silva, L.C.F., Albarnaz, J.D., Dornas, F.P., Trindade, G.S., Ferreira, P.P., Marques, J.T. *et al.* (2014) Samba virus: a novel mimivirus from a giant rain forest, the Brazilian Amazon. *Virology Journal*, **11**, 95.
91. Fischer, M.G., Allen, M.J., Wilson, W.H. and Suttle, C.A. (2010) Giant virus with a remarkable complement of genes infects marine zooplankton. *Proceedings of the National Academy of Sciences*, **107**, 19508-19513.
92. Fischer, M.G. and Suttle, C.A. (2011) A Virophage at the Origin of Large DNA Transposons. *Science*, **332**, 231.
93. Krupovic, M., Bamford, D.H. and Koonin, E.V. (2014) Conservation of major and minor jelly-roll capsid proteins in Polinton (Maverick) transposons suggests that they are bona fide viruses. *Biology Direct*, **9**, 6.
94. Yutin, N., Raoult, D. and Koonin, E.V. (2013) Virophages, polintons, and transpovirons: a complex evolutionary network of diverse selfish genetic elements with different reproduction strategies. *Virology Journal*, **10**, 158.
95. Fischer, M.G. and Hackl, T. (2016) Host genome integration and giant virus-induced reactivation of the virophage mavirus. *Nature*, **540**, 288.

96. Taylor, B.P., Cortez, M.H. and Weitz, J.S. (2014) The virus of my virus is my friend: Ecological effects of virophage with alternative modes of coinfection. *Journal of Theoretical Biology*, **354**, 124-136.
97. Yau, S., Lauro, F.M., DeMaere, M.Z., Brown, M.V., Thomas, T., Raftery, M.J., Andrews-Pfannkoch, C., Lewis, M., Hoffman, J.M., Gibson, J.A. *et al.* (2011) Virophage control of antarctic algal host–virus dynamics. *Proceedings of the National Academy of Sciences*, **108**, 6163-6168.
98. Zhou, J., Sun, D., Childers, A., McDermott, T.R., Wang, Y. and Liles, M.R. (2015) Three Novel Virophage Genomes Discovered from Yellowstone Lake Metagenomes. *Journal of Virology*, **89**, 1278-1285.
99. Zhou, J., Zhang, W., Yan, S., Xiao, J., Zhang, Y., Li, B., Pan, Y. and Wang, Y. (2013) Diversity of Virophages in Metagenomic Data Sets. *Journal of Virology*, **87**, 4225-4236.
100. Gaudet, R.G., Bradfield, C.J. and MacMicking, J.D. (2016) Evolution of Cell-Autonomous Effector Mechanisms in Macrophages versus Non-Immune Cells. *Microbiol Spectr*, **4**.
101. Cosson, P. and Soldati, T. (2008) Eat, kill or die: when amoeba meets bacteria. *Curr Opin Microbiol*, **11**, 271-276.
102. Randow, F., MacMicking John, D. and James Leo, C. (2013) Cellular Self-Defense: How Cell-Autonomous Immunity Protects Against Pathogens. *Science*, **340**, 701-706.
103. Galiana-Arnoux, D., Dostert, C., Schneemann, A., Hoffmann, J.A. and Imler, J.L. (2006) Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila. *Nat Immunol*, **7**, 590-597.
104. van Rij, R.P., Saleh, M.C., Berry, B., Foo, C., Houk, A., Antoniewski, C. and Andino, R. (2006) The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. *Genes Dev*, **20**, 2985-2995.
105. Wang, X.H., Aliyari, R., Li, W.X., Li, H.W., Kim, K., Carthew, R., Atkinson, P. and Ding, S.W. (2006) RNA interference directs innate immunity against viruses in adult Drosophila. *Science*, **312**, 452-454.
106. Wilkins, C., Dishongh, R., Moore, S.C., Whitt, M.A., Chow, M. and Machaca, K. (2005) RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. *Nature*, **436**, 1044-1047.
107. Schott, D.H., Cureton, D.K., Whelan, S.P. and Hunter, C.P. (2005) An antiviral role for the RNA interference machinery in Caenorhabditis elegans. *Proc Natl Acad Sci U S A*, **102**, 18420-18424.
108. Martens, H., Novotny, J., Oberstrass, J., Steck, T.L., Postlethwait, P. and Nellen, W. (2002) RNAi in Dictyostelium: the role of RNA-directed RNA polymerases and double-stranded RNase. *Mol Biol Cell*, **13**, 445-453.
109. Dunn, J.D., Bosmani, C., Barisch, C., Raykov, L., Lefrançois, L.H., Cardenal-Muñoz, E., López-Jiménez, A.T. and Soldati, T. (2018) Eat Prey, Live: Dictyostelium discoideum As a Model for Cell-Autonomous Defenses. *Front Immunol*, **8**, 1906-1906.
110. Boulais, J., Trost, M., Landry, C.R., Dieckmann, R., Levy, E.D., Soldati, T., Michnick, S.W., Thibault, P. and Desjardins, M. (2010) Molecular characterization of the evolution of phagosomes. *Mol Syst Biol*, **6**, 423.
111. Freeman, S.A. and Grinstein, S. (2014) Phagocytosis: receptors, signal integration, and the cytoskeleton. *Immunol Rev*, **262**, 193-215.
112. Chen, G., Zhuchenko, O. and Kuspa, A. (2007) Immune-like phagocyte activity in the social amoeba. *Science*, **317**, 678-681.
113. Cornillon, S., Gebbie, L., Benghezal, M., Nair, P., Keller, S., Wehrle-Haller, B., Charette, S.J., Brückert, F., Letourneur, F. and Cosson, P. (2006) An adhesion molecule in free-living Dictyostelium amoebae with integrin β features. *EMBO reports*, **7**, 617-621.
114. Gotthardt, D., Warnatz, H.J., Henschel, O., Brückert, F., Schleicher, M. and Soldati, T. (2002) High-Resolution Dissection of Phagosome Maturation Reveals Distinct Membrane Trafficking Phases. *Molecular Biology of the Cell*, **13**, 3508-3520.
115. Janssen, K.P., Rost, R., Eichinger, L. and Schleicher, M. (2001) Characterization of CD36/LIMPII homologues in Dictyostelium discoideum. *The Journal of biological chemistry*, **276**, 38899-38910.
116. Harris, T.J., Ravandi, A. and Siu, C.H. (2001) Assembly of glycoprotein-80 adhesion complexes in Dictyostelium. Receptor compartmentalization and oligomerization in membrane rafts. *The Journal of biological chemistry*, **276**, 48764-48774.
117. Sattler, N., Bosmani, C., Barisch, C., Guého, A., Gopaldass, N., Dias, M., Leuba, F., Bruckert, F., Cosson, P. and Soldati, T. (2018) Functions of the Dictyostelium LIMP-2 and CD36 homologues in bacteria uptake, phagolysosome biogenesis and host cell defence. *Journal of Cell Science*, **131**.
118. Gotthardt, D., Blancheteau, V., Bosserhoff, A., Ruppert, T., Delorenzi, M. and Soldati, T. (2006) Proteomics Fingerprinting of Phagosome Maturation and Evidence for the Role of a Gα during Uptake*S. *Molecular & Cellular Proteomics*, **5**, 2228-2243.
119. Bozzaro, S. and Roseman, S. (1983) Adhesion of Dictyostelium discoideum cells to carbohydrates immobilized in polyacrylamide gels. I. Evidence for three sugar-specific cell surface receptors. *The Journal of biological chemistry*, **258**, 13882-13889.

120. Vogel, G., Thilo, L., Schwarz, H. and Steinhart, R. (1980) Mechanism of phagocytosis in Dictyostelium discoideum: phagocytosis is mediated by different recognition sites as disclosed by mutants with altered phagocytotic properties. *J Cell Biol*, **86**, 456-465.
121. Seastone, D.J., Harris, E., Temesvari, L.A., Bear, J.E., Saxe, C.L. and Cardelli, J. (2001) The WASp-like protein scar regulates macropinocytosis, phagocytosis and endosomal membrane flow in Dictyostelium. *J Cell Sci*, **114**, 2673-2683.
122. Insall, R., Müller-Taubenberger, A., Machesky, L., Köhler, J., Simmeth, E., Atkinson, S.J., Weber, I. and Gerisch, G. (2001) Dynamics of the Dictyostelium Arp2/3 complex in endocytosis, cytokinesis, and chemotaxis. *Cell Motility*, **50**, 115-128.
123. Carnell, M., Zech, T., Calaminus, S.D., Ura, S., Hagedorn, M., Johnston, S.A., May, R.C., Soldati, T., Machesky, L.M. and Insall, R.H. (2011) Actin polymerization driven by WASH causes V-ATPase retrieval and vesicle neutralization before exocytosis. *J Cell Biol*, **193**, 831-839.
124. Bozzaro, S., Bucci, C. and Steinert, M. (2008), *International Review of Cell and Molecular Biology*. Academic Press, Vol. 271, pp. 253-300.
125. Rivero, F. and Xiong, H. (2016) In Jeon, K. W. (ed.), *International Review of Cell and Molecular Biology*. Academic Press, Vol. 322, pp. 61-181.
126. Blanc, C., Charette, S., Cherix, N., Lefkir, Y., Cosson, P. and Letourneur, F. (2005) A novel phosphatidylinositol 4,5-bisphosphate-binding domain targeting the Phg2 kinase to the membrane in Dictyostelium cells. *European Journal of Cell Biology*, **84**, 951-960.
127. Dormann, D., Weijer, G., Dowler, S. and Weijer, C.J. (2004) In vivo analysis of 3-phosphoinositide dynamics during Dictyostelium phagocytosis and chemotaxis. *Journal of Cell Science*, **117**, 6497-6509.
128. Loovers, H.M., Kortholt, A., de Groote, H., Whitty, L., Nussbaum, R.L. and van Haastert, P.J.M. (2007) Regulation of Phagocytosis in Dictyostelium by the Inositol 5-Phosphatase OCRL Homolog Dd5P4. *Traffic*, **8**, 618-628.
129. Buczynski, G., Bush, J., Zhang, L., Rodriguez-Paris, J. and Cardelli, J. (1997) Evidence for a recycling role for Rab7 in regulating a late step in endocytosis and in retention of lysosomal enzymes in Dictyostelium discoideum. *Mol Biol Cell*, **8**, 1343-1360.
130. Rupper, A., Grove, B. and Cardelli, J. (2001) Rab7 regulates phagosome maturation in Dictyostelium. *J Cell Sci*, **114**, 2449-2460.
131. Vieira, O.V., Bucci, C., Harrison, R.E., Trimble William, S., Lanzetti, L., Gruenberg, J., Schreiber Alan, D., Stahl Philip, D. and Grinstein, S. (2003) Modulation of Rab5 and Rab7 Recruitment to Phagosomes by Phosphatidylinositol 3-Kinase. *Molecular and Cellular Biology*, **23**, 2501-2514.
132. Gorvel, J.P., Chavrier, P., Zerial, M. and Gruenberg, J. (1991) rab5 controls early endosome fusion in vitro. *Cell*, **64**, 915-925.
133. Neuhaus, E.M., Almers, W. and Soldati, T. (2002) Morphology and Dynamics of the Endocytic Pathway inDictyostelium discoideum. *Molecular Biology of the Cell*, **13**, 1390-1407.
134. Sun-Wada, G.H., Tabata, H., Kawamura, N., Aoyama, M. and Wada, Y. (2009) Direct recruitment of H+-ATPase from lysosomes for phagosomal acidification. *J Cell Sci*, **122**, 2504-2513.
135. Clarke, M., Köhler, J., Arana, Q., Liu, T., Heuser, J. and Gerisch, G.n. (2002) Dynamics of the vacuolar H+-ATPase in the contractile vacuole complex and the endosomal pathway of Dictyostelium cells. *Journal of Cell Science*, **115**, 2893-2905.
136. Marchetti, A., Lelong, E. and Cosson, P. (2009) A measure of endosomal pH by flow cytometry in Dictyostelium. *BMC Res Notes*, **2**, 7.
137. Sattler, N., Monroy, R. and Soldati, T. (2013) Quantitative analysis of phagocytosis and phagosome maturation. *Methods Mol Biol*, **983**, 383-402.
138. Gopaldass, N., Patel, D., Kratzke, R., Dieckmann, R., Hausherr, S., Hagedorn, M., Monroy, R., Krüger, J., Neuhaus, E.M., Hoffmann, E. *et al.* (2012) Dynamin A, Myosin IB and Abp1 couple phagosome maturation to F-actin binding. *Traffic*, **13**, 120-130.
139. Yates, R.M., Hermetter, A. and Russell, D.G. (2005) The Kinetics of Phagosome Maturation as a Function of Phagosome/Lysosome Fusion and Acquisition of Hydrolytic Activity. *Traffic*, **6**, 413-420.
140. King, J.S., Gueho, A., Hagedorn, M., Gopaldass, N., Leuba, F., Soldati, T. and Insall, R.H. (2013) WASH is required for lysosomal recycling and efficient autophagic and phagocytic digestion. *Molecular biology of the cell*, **24**, 2714-2726.
141. Maniak, M. (2003) Fusion and fission events in the endocytic pathway of Dictyostelium. *Traffic*, **4**, 1-5.
142. Jenne, N., Rauchenberger, R., Hacker, U., Kast, T. and Maniak, M. (1998) Targeted gene disruption reveals a role for vacuolin B in the late endocytic pathway and exocytosis. *J Cell Sci*, **111 (Pt 1)**, 61-70.
143. Rauchenberger, R., Hacker, U., Murphy, J., Niewöhner, J. and Maniak, M. (1997) Coronin and vacuolin identify consecutive stages of a late, actin-coated endocytic compartment in Dictyostelium. *Current Biology*, **7**, 215-218.
144. Wienke, D., Drengk, A., Schmauch, C., Jenne, N. and Maniak, M. (2006) Vacuolin, a flotillin/reggie-related protein from Dictyostelium oligomerizes for endosome association. *Eur J Cell Biol*, **85**, 991-1000.

145. Charette, S.J. and Cosson, P. (2006) Exocytosis of late endosomes does not directly contribute membrane to the formation of phagocytic cups or pseudopods in Dictyostelium. *FEBS Letters*, **580**, 4923-4928.
146. Lima, W.C., Leuba, F., Soldati, T. and Cosson, P. (2012) Mucolipin controls lysosome exocytosis in Dictyostelium. *Journal of Cell Science*, **125**, 2315-2322.
147. Holt, O.J., Gallo, F. and Griffiths, G.M. (2006) Regulating Secretory Lysosomes. *The Journal of Biochemistry*, **140**, 7-12.
148. Klionsky, D.J., Cregg, J.M., Dunn, W.A., Jr., Emr, S.D., Sakai, Y., Sandoval, I.V., Sibirny, A., Subramani, S., Thumm, M., Veenhuis, M. *et al.* (2003) A unified nomenclature for yeast autophagy-related genes. *Dev Cell*, **5**, 539-545.
149. Takeshige, K., Baba, M., Tsuboi, S., Noda, T. and Ohsumi, Y. (1992) Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction. *J Cell Biol*, **119**, 301-311.
150. Cotter, D.A., Miura-Santo, L.Y. and Hohl, H.R. (1969) Ultrastructural changes during germination of Dictyostelium discoideum spores. *J Bacteriol*, **100**, 1020-1026.
151. Calvo-Garrido, J., Carilla-Latorre, S., Kubohara, Y., Santos-Rodrigo, N., Mesquita, A., Soldati, T., Golstein, P. and Escalante, R. (2010) Autophagy in Dictyostelium: Genes and pathways, cell death and infection. *Autophagy*, **6**, 686-701.
152. Mesquita, A., Cardenal-Muñoz, E., Dominguez, E., Muñoz-Braceras, S., Nuñez-Corcuera, B., Phillips, B.A., Tábara, L.C., Xiong, Q., Coria, R., Eichinger, L. *et al.* (2017) Autophagy in Dictyostelium: Mechanisms, regulation and disease in a simple biomedical model. *Autophagy*, **13**, 24-40.
153. Otto, G.P., Wu, M.Y., Kazgan, N., Anderson, O.R. and Kessin, R.H. (2003) Macroautophagy is required for multicellular development of the social amoeba Dictyostelium discoideum. *The Journal of biological chemistry*, **278**, 17636-17645.
154. Cabral, M., Anjard, C., Malhotra, V., Loomis, W.F. and Kuspa, A. (2010) Unconventional secretion of AcbA in Dictyostelium discoideum through a vesicular intermediate. *Eukaryot Cell*, **9**, 1009-1017.
155. Klionsky, D.J., Eskelinen, E.-L. and Deretic, V. (2014) Autophagosomes, phagosomes, autolysosomes, phagolysosomes, autophagolysosomes... wait, I'm confused. *Autophagy*, **10**, 549-551.
156. Mesquita, A., Tábara, L.C., Martinez-Costa, O., Santos-Rodrigo, N., Vincent, O. and Escalante, R. (2015) Dissecting the function of Atg1 complex in Dictyostelium autophagy reveals a connection with the pentose phosphate pathway enzyme transketolase. *Open Biology*, **5**, 150088.
157. Tekinay, T., Wu, M.Y., Otto, G.P., Anderson, O.R. and Kessin, R.H. (2006) Function of the Dictyostelium discoideum Atg1 kinase during autophagy and development. *Eukaryot Cell*, **5**, 1797-1806.
158. Fischer, S. and Eichinger, L. (2019) Dictyostelium discoideum and autophagy - a perfect pair. *The International journal of developmental biology*, **63 8-9-10**, 485-495.
159. Park, J.M., Seo, M., Jung, C.H., Grunwald, D., Stone, M., Otto, N.M., Toso, E., Ahn, Y., Kyba, M., Griffin, T.J. *et al.* (2018) ULK1 phosphorylates Ser30 of BECN1 in association with ATG14 to stimulate autophagy induction. *Autophagy*, **14**, 584-597.
160. Axe, E.L., Walker, S.A., Manifava, M., Chandra, P., Roderick, H.L., Habermann, A., Griffiths, G. and Ktistakis, N.T. (2008) Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. *J Cell Biol*, **182**, 685-701.
161. Calvo-Garrido, J., King, J.S., Muñoz-Braceras, S. and Escalante, R. (2014) Vmp1 regulates PtdIns3P signaling during autophagosome formation in Dictyostelium discoideum. *Traffic*, **15**, 1235-1246.
162. Tábara, L.C., Vicente, J.J., Biazik, J., Eskelinen, E.L., Vincent, O. and Escalante, R. (2018) Vacuole membrane protein 1 marks endoplasmic reticulum subdomains enriched in phospholipid synthesizing enzymes and is required for phosphoinositide distribution. *Traffic*, **19**, 624-638.
163. Lindmo, K. and Stenmark, H. (2006) Regulation of membrane traffic by phosphoinositide 3-kinases. *J Cell Sci*, **119**, 605-614.
164. Obara, K., Sekito, T., Niimi, K. and Ohsumi, Y. (2008) The Atg18-Atg2 complex is recruited to autophagic membranes via phosphatidylinositol 3-phosphate and exerts an essential function. *The Journal of biological chemistry*, **283**, 23972-23980.
165. Geng, J. and Klionsky, D.J. (2008) The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. 'Protein modifications: beyond the usual suspects' review series. *EMBO reports*, **9**, 859-864.
166. Matthias, J., Meßling, S. and Eichinger, L. (2016) The two Dictyostelium autophagy eight proteins, ATG8a and ATG8b, associate with the autophagosome in succession. *Eur J Cell Biol*, **95**, 15-25.
167. Meßling, S., Matthias, J., Xiong, Q., Fischer, S. and Eichinger, L. (2017) The two Dictyostelium discoideum autophagy 8 proteins have distinct autophagic functions. *Eur J Cell Biol*, **96**, 312-324.
168. Lamb, C.A., Yoshimori, T. and Tooze, S.A. (2013) The autophagosome: origins unknown, biogenesis complex. *Nature Reviews Molecular Cell Biology*, **14**, 759-774.
169. Stanley, R.E., Ragusa, M.J. and Hurley, J.H. (2014) The beginning of the end: how scaffolds nucleate autophagosome biogenesis. *Trends Cell Biol*, **24**, 73-81.

170. Wirawan, E., Vanden Berghe, T., Lippens, S., Agostinis, P. and Vandenabeele, P. (2012) Autophagy: for better or for worse. *Cell Res*, **22**, 43-61.
171. Jahreiss, L., Menzies, F.M. and Rubinsztein, D.C. (2008) The itinerary of autophagosomes: from peripheral formation to kiss-and-run fusion with lysosomes. *Traffic*, **9**, 574-587.
172. Nakamura, S. and Yoshimori, T. (2017) New insights into autophagosome-lysosome fusion. *J Cell Sci*, **130**, 1209-1216.
173. Stolz, A., Ernst, A. and Dikic, I. (2014) Cargo recognition and trafficking in selective autophagy. *Nature Cell Biology*, **16**, 495-501.
174. Wild, P., McEwan, D.G. and Dikic, I. (2014) The LC3 interactome at a glance. *Journal of Cell Science*, **127**, 3-9.
175. Kim, P.K., Hailey, D.W., Mullen, R.T. and Lippincott-Schwartz, J. (2008) Ubiquitin signals autophagic degradation of cytosolic proteins and peroxisomes. *Proceedings of the National Academy of Sciences*, **105**, 20567.
176. Cardenal-Muñoz, E., Arafah, S., López-Jiménez, A.T., Kicka, S., Falaise, A., Bach, F., Schaad, O., King, J.S., Hagedorn, M. and Soldati, T. (2017) Mycobacterium marinum antagonistically induces an autophagic response while repressing the autophagic flux in a TORC1- and ESX-1-dependent manner. *PLOS Pathogens*, **13**, e1006344.
177. Gerstenmaier, L., Pilla, R., Herrmann, L., Herrmann, H., Prado, M., Villafano, G.J., Kolonko, M., Reimer, R., Soldati, T., King, J.S. *et al.* (2015) The autophagic machinery ensures nonlytic transmission of mycobacteria. *Proceedings of the National Academy of Sciences*, **112**, E687.
178. Ding, S.-W. and Voinnet, O. (2007) Antiviral immunity directed by small RNAs. *Cell*, **130**, 413-426.
179. Singla-Rastogi, M., Charvin, M., Thiébeauld, O., Perez-Quintero, A.L., Ravet, A., Emidio-Fortunato, A., Mendu, V. and Navarro, L. (2019) Plant Small RNA Species Direct Gene Silencing in Pathogenic Bacteria as well as Disease Protection. *bioRxiv*, 863902.
180. Kuo, Y.-W. and Falk, B.W. (2020) RNA interference approaches for plant disease control. *BioTechniques*, **69**, 469-477.
181. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, **391**, 806-811.
182. Elbashir, S.M., Lendeckel, W. and Tuschl, T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev*, **15**, 188-200.
183. Tuschl, T., Zamore, P.D., Lehmann, R., Bartel, D.P. and Sharp, P.A. (1999) Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev*, **13**, 3191-3197.
184. Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494-498.
185. Cerutti, H. and Casas-Mollano, J.A. (2006) On the origin and functions of RNA-mediated silencing: from protists to man. *Current genetics*, **50**, 81-99.
186. Hammond, S.M., Bernstein, E., Beach, D. and Hannon, G.J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature*, **404**, 293-296.
187. Noland, C.L., Ma, E. and Doudna, J.A. (2011) siRNA repositioning for guide strand selection by human Dicer complexes. *Mol Cell*, **43**, 110-121.
188. Schwarz, D.S., Hutvágner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, **115**, 199-208.
189. Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L. and Hannon, G.J. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science*, **305**, 1437-1441.
190. Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G. and Tuschl, T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell*, **15**, 185-197.
191. Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N. and Maine, E.M. (2000) EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. *Curr Biol*, **10**, 169-178.
192. Maida, Y. and Masutomi, K. (2011) RNA-dependent RNA polymerases in RNA silencing. *Biol Chem*, **392**, 299-304.
193. Devert, A., Fabre, N., Floris, M., Canard, B., Robaglia, C. and Crété, P. (2015) Primer-dependent and primer-independent initiation of double stranded RNA synthesis by purified Arabidopsis RNA-dependent RNA polymerases RDR2 and RDR6. *PLoS One*, **10**, e0120100.
194. Voinnet, O. (2008) Use, tolerance and avoidance of amplified RNA silencing by plants. *Trends Plant Sci*, **13**, 317-328.
195. Pak, J., Maniar, J.M., Mello, C.C. and Fire, A. (2012) Protection from feed-forward amplification in an amplified RNAi mechanism. *Cell*, **151**, 885-899.
196. Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J. and Plasterk, R.H. (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev*, **15**, 2654-2659.

197. Pak, J. and Fire, A. (2007) Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science*, **315**, 241-244.
198. Blevins, T., Rajeswaran, R., Shivaprasad, P.V., Beknazariants, D., Si-Ammour, A., Park, H.-S., Vazquez, F., Robertson, D., Meins, F., Jr., Hohn, T. *et al.* (2006) Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic acids research*, **34**, 6233-6246.
199. Hemmings-Mieszczak, M., Steger, G. and Hohn, T. (1997) Alternative structures of the cauliflower mosaic virus 35 s RNA leader: implications for viral expression and replication11Edited by I. Tinoco. *Journal of Molecular Biology*, **267**, 1075-1088.
200. Malicki, M., Iliopoulou, M. and Hammann, C. (2017) Retrotransposon Domestication and Control in *Dictyostelium discoideum*. *Frontiers in Microbiology*, **8**, 1869.
201. Kjellin, J., Pränting, M., Bach, F., Vaid, R., Edelbroek, B., Li, Z., Hoeppner, M.P., Grabherr, M., Isberg, R.R., Hagedorn, M. *et al.* (2019) Investigation of the host transcriptional response to intracellular bacterial infection using Dictyostelium discoideum as a host model. *BMC Genomics*, **20**, 961.
202. Davis, F.F. and Allen, F.W. (1957) Ribonucleic acids from yeast which contain a fifth nucleotide. *The Journal of biological chemistry*, **227**, 907-915.
203. Hotchkiss, R.D. (1948) THE QUANTITATIVE SEPARATION OF PURINES, PYRIMIDINES, AND NUCLEOSIDES BY PAPER CHROMATOGRAPHY. *Journal of Biological Chemistry*, **175**, 315-332.
204. Boccaletto, P., Machnicka, M.A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T.K., de Crécy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A. *et al.* (2018) MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Research*, **46**, D303-D307.
205. Helm, M. and Motorin, Y. (2017) Detecting RNA modifications in the epitranscriptome: predict and validate. *Nat Rev Genet*, **18**, 275-291.
206. Boo, S.H. and Kim, Y.K. (2020) The emerging role of RNA modifications in the regulation of mRNA stability. *Experimental & Molecular Medicine*, **52**, 400-408.
207. Delaunay, S. and Frye, M. (2019) RNA modifications regulating cell fate in cancer. *Nat Cell Biol*, **21**, 552-559.
208. Dimitrova, D.G., Teysset, L. and Carré, C. (2019) RNA 2'-O-Methylation (Nm) Modification in Human Diseases. *Genes (Basel)*, **10**, 117.
209. Abou Assi, H., Rangadurai, A.K., Shi, H., Liu, B., Clay, M.C., Erharter, K., Kreutz, C., Holley, C.L. and Al-Hashimi, Hashim M. (2020) 2′-O-Methylation can increase the abundance and lifetime of alternative RNA conformational states. *Nucleic Acids Research*, **48**, 12365-12379.
210. Jonkhout, N., Tran, J., Smith, M.A., Schonrock, N., Mattick, J.S. and Novoa, E.M. (2017) The RNA modification landscape in human disease. *Rna*, **23**, 1754-1769.
211. Nachtergaele, S. and He, C. (2018) Chemical Modifications in the Life of an mRNA Transcript. *Annu Rev Genet*, **52**, 349-372.
212. Kadumuri, R.V. and Janga, S.C. (2018) Epitranscriptomic Code and Its Alterations in Human Disease. *Trends Mol Med*, **24**, 886-903.
213. Roundtree, I.A., Evans, M.E., Pan, T. and He, C. (2017) Dynamic RNA Modifications in Gene Expression Regulation. *Cell*, **169**, 1187-1200.
214. Shi, H., Wei, J. and He, C. (2019) Where, When, and How: Context-Dependent Functions of RNA Methylation Writers, Readers, and Erasers. *Mol Cell*, **74**, 640-650.
215. Singh, G., Pratt, G., Yeo, G.W. and Moore, M.J. (2015) The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annu Rev Biochem*, **84**, 325-354.
216. Heiss, M., Hagelskamp, F., Marchand, V., Motorin, Y. and Kellner, S. (2021) Cell culture NAIL-MS allows insight into human tRNA and rRNA modification dynamics in vivo. *Nature Communications*, **12**, 389.
217. Jackman, J.E. and Alfonzo, J.D. (2013) Transfer RNA modifications: nature's combinatorial chemistry playground. *Wiley interdisciplinary reviews. RNA*, **4**, 35-48.
218. Helm, M. (2006) Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Research*, **34**, 721-733.
219. Helm, M., Giegé, R. and Florentz, C. (1999) A Watson−Crick Base-Pair-Disrupting Methyl Group (m1A9) Is Sufficient for Cloverleaf Folding of Human Mitochondrial tRNALys *Biochemistry*, **38**, 13338-13346.
220. Motorin, Y. and Helm, M. (2010) tRNA Stabilization by Modified Nucleotides. *Biochemistry*, **49**, 4934-4944.
221. Sakurai, M., Ohtsuki, T. and Watanabe, K. (2005) Modification at position 9 with 1-methyladenosine is crucial for structure and function of nematode mitochondrial tRNAs lacking the entire T-arm. *Nucleic Acids Research*, **33**, 1653-1661.
222. Ontiveros, R.J., Stoute, J. and Liu, K.F. (2019) The chemical diversity of RNA modifications. *The Biochemical journal*, **476**, 1227-1245.

223. Hayase, Y., Jahn, M., Rogers, M.J., Sylvers, L.A., Koizumi, M., Inoue, H., Ohtsuka, E. and Söll, D. (1992) Recognition of bases in Escherichia coli tRNA(Gln) by glutaminyl-tRNA synthetase: a complete identity set. *The EMBO Journal*, **11**, 4159-4165.
224. Nawrot, B., Malkiewicz, A., Smith, W.S., Sierzputowska-Gracz, H. and Agris, P.F. (1995) RNA Modified Uridines VII: Chemical Synthesis and Initial Analysis of tRNA D-Loop Oligomers with Tandem Modified Uridines. *Nucleosides and Nucleotides*, **14**, 143-165.
225. Dalluge, J.J., Hashizume, T., Sopchik, A.E., McCloskey, J.A. and Davis, D.R. (1996) Conformational Flexibility in RNA: The Role of Dihydrouridine. *Nucleic Acids Research*, **24**, 1073-1079.
226. Dyubankova, N., Sochacka, E., Kraszewska, K., Nawrot, B., Herdewijn, P. and Lescrinier, E. (2015) Contribution of dihydrouridine in folding of the D-arm in tRNA. *Org Biomol Chem*, **13**, 4960-4966.
227. Helm, M. and Alfonzo, Juan D. (2014) Posttranscriptional RNA Modifications: Playing Metabolic Games in a Cell's Chemical Legoland. *Chemistry & Biology*, **21**, 174-185.
228. Schäck, M.A., Jablonski, K.P., Gräf, S., Klassen, R., Schaffrath, R., Kellner, S. and Hammann, C. (2020) Eukaryotic life without tQCUG: the role of Elongator-dependent tRNA modifications in Dictyostelium discoideum. *Nucleic Acids Research*, **48**, 7899-7913.
229. El Yacoubi, B., Bailly, M. and de Crécy-Lagard, V. (2012) Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annu Rev Genet*, **46**, 69-95.
230. Grosjean, H., Sprinzl, M. and Steinberg, S. (1995) Posttranscriptionally modified nucleosides in transfer RNA: Their locations and frequencies. *Biochimie*, **77**, 139-141.
231. Agris, P.F. (2008) Bringing order to translation: the contributions of transfer RNA anticodon-domain modifications. *EMBO reports*, **9**, 629-635.
232. Li, X., Xiong, X. and Yi, C. (2016) Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat Methods*, **14**, 23-31.
233. Huang, H., Weng, H. and Chen, J. (2020) The Biogenesis and Precise Control of RNA m(6)A Methylation. *Trends in genetics : TIG*, **36**, 44-52.
234. Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.M., Li, C.J., Vågbø, C.B., Shi, Y., Wang, W.L., Song, S.H. *et al.* (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell*, **49**, 18-29.
235. Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.G. *et al.* (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol*, **7**, 885-887.
236. Zhao, B.S., Wang, X., Beadell, A.V., Lu, Z., Shi, H., Kuuspalu, A., Ho, R.K. and He, C. (2017) m(6)A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. *Nature*, **542**, 475-478.
237. Yoon, K.J., Ringeling, F.R., Vissers, C., Jacob, F., Pokrass, M., Jimenez-Cyrus, D., Su, Y., Kim, N.S., Zhu, Y., Zheng, L. *et al.* (2017) Temporal Control of Mammalian Cortical Neurogenesis by m(6)A Methylation. *Cell*, **171**, 877-889.e817.
238. Paris, J., Morgan, M., Campos, J., Spencer, G.J., Shmakova, A., Ivanova, I., Mapperley, C., Lawson, H., Wotherspoon, D.A., Sepulveda, C. *et al.* (2019) Targeting the RNA m(6)A Reader YTHDF2 Selectively Compromises Cancer Stem Cells in Acute Myeloid Leukemia. *Cell Stem Cell*, **25**, 137-148.e136.
239. Borland, K.M., Diesend, J., Ito-Kureha, T., Heissmeyer, V., Hammann, C., Buck, A.H., Michalakis, S. and Kellner, S.M. (2019) Production and Application of Stable Isotope-Labeled Internal Standards for RNA Modification Analysis. *Genes (Basel)*, **10**.
240. Sloan, K.E., Warda, A.S., Sharma, S., Entian, K.D., Lafontaine, D.L.J. and Bohnsack, M.T. (2017) Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA biology*, **14**, 1138-1152.
241. Noller, H.F., Hoffarth, V. and Zimniak, L. (1992) Unusual resistance of peptidyl transferase to protein extraction procedures. *Science*, **256**, 1416-1419.
242. Nissen, P., Hansen, J., Ban, N., Moore, P.B. and Steitz, T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science*, **289**, 920–930.
243. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. *et al.* (2000) Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell*, **102**, 615-623.
244. Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T. and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327-339.
245. Bassler, J. and Hurt, E. (2019) Eukaryotic Ribosome Assembly. *Annu Rev Biochem*, **88**, 281-306.
246. Decatur, W.A. and Fournier, M.J. (2002) rRNA modifications and ribosome function. *Trends Biochem Sci*, **27**, 344-351.
247. Polikanov, Y.S., Melnikov, S.V., Söll, D. and Steitz, T.A. (2015) Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. *Nat Struct Mol Biol*, **22**, 342-344.
248. Liang, X.-h., Liu, Q. and Fournier, M.J. (2007) rRNA Modifications in an Intersubunit Bridge of the Ribosome Strongly Affect Both Ribosome Biogenesis and Activity. *Molecular Cell*, **28**, 965-977.

249. Henras, A.K., Plisson-Chastang, C., Humbert, O., Romeo, Y. and Henry, Y. (2017) Synthesis, Function, and Heterogeneity of snoRNA-Guided Posttranscriptional Nucleoside Modifications in Eukaryotic Ribosomal RNAs. *Enzymes*, **41**, 169-213.
250. Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J. and Nielsen, H. (2015) Profiling of ribose methylations in RNA by high-throughput sequencing. *Angew Chem Int Ed Engl*, **54**, 451-455.
251. Krogh, N., Jansson, M.D., Häfner, S.J., Tehler, D., Birkedal, U., Christensen-Dalsgaard, M., Lund, A.H. and Nielsen, H. (2016) Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res*, **44**, 7884-7895.
252. Ramachandran, S., Krogh, N., Jørgensen, T.E., Johansen, S.D., Nielsen, H. and Babiak, I. (2020) The shift from early to late types of ribosomes in zebrafish development involves changes at a subset of rRNA 2'-O-Me sites. *Rna*.
253. Locati, M.D., Pagano, J.F.B., Girard, G., Ensink, W.A., van Olst, M., van Leeuwen, S., Nehrdich, U., Spaink, H.P., Rauwerda, H., Jonker, M.J. *et al.* (2017) Expression of distinct maternal and somatic 5.8S, 18S, and 28S rRNA types during zebrafish development. *RNA*, **23**, 1188-1199.
254. Georgeson, J.M. and Schwartz, S. (2021) The ribosome epitranscriptome: Inert - or a platform for functional plasticity? *RNA*.
255. Krogh, N., Kongsbak-Wismann, M., Geisler, C. and Nielsen, H. (2017) Substoichiometric ribose methylations in spliceosomal snRNAs. *Organic & Biomolecular Chemistry*, **15**, 8872-8876.
256. Taoka, M., Nobe, Y., Yamaki, Y., Sato, K., Ishikawa, H., Izumikawa, K., Yamauchi, Y., Hirota, K., Nakayama, H., Takahashi, N. *et al.* (2018) Landscape of the complete RNA chemical modifications in the human 80S ribosome. *Nucleic Acids Res*, **46**, 9289-9298.
257. Dai, Q., Moshitch-Moshkovitz, S., Han, D., Kol, N., Amariglio, N., Rechavi, G., Dominissini, D. and He, C. (2017) Nm-seq maps 2'-O-methylation sites in human mRNA with base precision. *Nat Methods*, **14**, 695-698.
258. Somme, J., Van Laer, B., Roovers, M., Steyaert, J., Versées, W. and Droogmans, L. (2014) Characterization of two homologous 2'-O-methyltransferases showing different specificities for their tRNA substrates. *Rna*, **20**, 1257-1271.
259. Lapeyre, B. and Purushothaman, S.K. (2004) Spb1p-Directed Formation of Gm2922 in the Ribosome Catalytic Center Occurs at a Late Processing Stage. *Molecular Cell*, **16**, 663-669.
260. Kiss-László, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M. and Kiss, T. (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell*, **85**, 1077-1088.
261. Erales, J., Marchand, V., Panthu, B., Gillot, S., Belin, S., Ghayad, S.E., Garcia, M., Laforêts, F., Marcel, V., Baudin-Baillieu, A. *et al.* (2017) Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes. *Proc Natl Acad Sci U S A*, **114**, 12934-12939.
262. Clay, M.C., Ganser, L.R., Merriman, D.K. and Al-Hashimi, H.M. (2017) Resolving sugar puckers in RNA excited states exposes slow modes of repuckering dynamics. *Nucleic Acids Res*, **45**, e134.
263. Kawai, G., Yamamoto, Y., Kamimura, T., Masegi, T., Sekine, M., Hata, T., Iimori, T., Watanabe, T., Miyazawa, T. and Yokoyama, S. (1992) Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry*, **31**, 1040-1046.
264. Kool, E.T. (1997) Preorganization of DNA: Design Principles for Improving Nucleic Acid Recognition by Synthetic Oligonucleotides. *Chem Rev*, **97**, 1473-1488.
265. Yildirim, I., Kierzek, E., Kierzek, R. and Schatz, G.C. (2014) Interplay of LNA and 2'-O-methyl RNA in the structure and thermodynamics of RNA hybrid systems: a molecular dynamics study using the revised AMBER force field and comparison with experimental results. *J Phys Chem B*, **118**, 14177-14187.
266. Didychuk, A.L., Butcher, S.E. and Brow, D.A. (2018) The life of U6 small nuclear RNA, from cradle to grave. *Rna*, **24**, 437-460.
267. Montemayor, E.J., Curran, E.C., Liao, H.H., Andrews, K.L., Treba, C.N., Butcher, S.E. and Brow, D.A. (2014) Core structure of the U6 small nuclear ribonucleoprotein at 1.7-Å resolution. *Nat Struct Mol Biol*, **21**, 544-551.
268. Mefford, M.A. and Staley, J.P. (2009) Evidence that U2/U6 helix I promotes both catalytic steps of pre-mRNA splicing and rearranges in between these steps. *Rna*, **15**, 1386-1397.
269. Choi, J., Indrisiunaite, G., DeMirci, H., Ieong, K.W., Wang, J., Petrov, A., Prabhakar, A., Rechavi, G., Dominissini, D., He, C. *et al.* (2018) 2'-O-methylation in mRNA disrupts tRNA decoding during translation elongation. *Nat Struct Mol Biol*, **25**, 208-216.
270. Cohn, W.E. (1959) 5-Ribosyl uracil, a carbon-carbon ribofuranosyl nucleoside in ribonucleic acids. *Biochimica et Biophysica Acta*, **32**, 569-571.
271. Charette, M. and Gray, M.W. (2000) Pseudouridine in RNA: What, Where, How, and Why. *IUBMB Life*, **49**, 341-351.
272. Hamma, T. and Ferré-D'Amaré, A.R. (2006) Pseudouridine synthases. *Chem Biol*, **13**, 1125-1135.

273. Rintala-Dempsey, A.C. and Kothe, U. (2017) Eukaryotic stand-alone pseudouridine synthases - RNA modifying enzymes and emerging regulators of gene expression? *RNA biology*, **14**, 1185-1196.
274. Yu, Y.-T. and Meier, U.T. (2014) RNA-guided isomerization of uridine to pseudouridine--pseudouridylation. *RNA biology*, **11**, 1483-1494.
275. Guthrie, C. and Patterson, B. (1988) Spliceosomal snRNAs. *Annu Rev Genet*, **22**, 387-419.
276. Ofengand, J. and Fournier, M.J. (1998), *Modification and Editing of RNA*, pp. 229-253.
277. Jack, K., Bellodi, C., Landry, D.M., Niederer, R.O., Meskauskas, A., Musalgaonkar, S., Kopmar, N., Krasnykh, O., Dean, A.M., Thompson, S.R. *et al.* (2011) rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol Cell*, **44**, 660-666.
278. King, T.H., Liu, B., McCully, R.R. and Fournier, M.J. (2003) Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol Cell*, **11**, 425-435.
279. Baudin-Baillieu, A., Fabret, C., Liang, X.-H., Piekna-Przybylska, D., Fournier, M.J. and Rousset, J.-P. (2009) Nucleotide modifications in three functionally important regions of the Saccharomyces cerevisiae ribosome affect translation accuracy. *Nucleic acids research*, **37**, 7665-7677.
280. Liang, X.-H., Liu, Q. and Fournier, M.J. (2009) Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA (New York, N.Y.)*, **15**, 1716-1728.
281. Sumita, M., Jiang, J., SantaLucia, J., Jr. and Chow, C.S. (2012) Comparison of solution conformations and stabilities of modified helix 69 rRNA analogs from bacteria and human. *Biopolymers*, **97**, 94-106.
282. Jiang, J., Kharel, D.N. and Chow, C.S. (2015) Modulation of conformational changes in helix 69 mutants by pseudouridine modifications. *Biophys Chem*, **200-201**, 48-55.
283. Jiang, J., Aduri, R., Chow, C.S. and SantaLucia, J., Jr. (2014) Structure modulation of helix 69 from Escherichia coli 23S ribosomal RNA by pseudouridylations. *Nucleic Acids Res*, **42**, 3971-3981.
284. Desaulniers, J.P., Chang, Y.C., Aduri, R., Abeysirigunawardena, S.C., SantaLucia, J., Jr. and Chow, C.S. (2008) Pseudouridines in rRNA helix 69 play a role in loop stacking interactions. *Org Biomol Chem*, **6**, 3892-3895.
285. Zhao, X. and Yu, Y.T. (2004) Pseudouridines in and near the branch site recognition region of U2 snRNA are required for snRNP biogenesis and pre-mRNA splicing in Xenopus oocytes. *Rna*, **10**, 681-690.
286. Yu, Y.T., Shu, M.D. and Steitz, J.A. (1998) Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *Embo j*, **17**, 5783-5795.
287. Yang, C., McPheeters, D.S. and Yu, Y.T. (2005) Psi35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in Saccharomyces cerevisiae. *The Journal of biological chemistry*, **280**, 6655-6662.
288. Lin, Y. and Kielkopf, C.L. (2008) X-ray structures of U2 snRNA-branchpoint duplexes containing conserved pseudouridines. *Biochemistry*, **47**, 5503-5514.
289. Davis, D.R. (1995) Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res*, **23**, 5020-5026.
290. Deb, I., Popenda, Ł., Sarzyńska, J., Małgowska, M., Lahiri, A., Gdaniec, Z. and Kierzek, R. (2019) Computational and NMR studies of RNA duplexes with an internal pseudouridine-adenosine base pair. *Scientific Reports*, **9**, 16278.
291. Kierzek, E., Malgowska, M., Lisowiec, J., Turner, D.H., Gdaniec, Z. and Kierzek, R. (2014) The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Res*, **42**, 3492-3501.
292. Arnez, J.G. and Steitz, T.A. (1994) Crystal structure of unmodified tRNA(Gln) complexed with glutaminyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry*, **33**, 7560-7567.
293. Vaidyanathan, P.P., AlSadhan, I., Merriman, D.K., Al-Hashimi, H.M. and Herschlag, D. (2017) Pseudouridine and N(6)-methyladenosine modifications weaken PUF protein/RNA interactions. *Rna*, **23**, 611-618.
294. Chen, C., Zhao, X., Kierzek, R. and Yu, Y.T. (2010) A flexible RNA backbone within the polypyrimidine tract is required for U2AF65 binding and pre-mRNA splicing in vivo. *Mol Cell Biol*, **30**, 4108-4119.
295. Kolev, N.G. and Steitz, J.A. (2006) In vivo assembly of functional U7 snRNP requires RNA backbone flexibility within the Sm-binding site. *Nat Struct Mol Biol*, **13**, 347-353.
296. Kligun, E. and Mandel-Gutfreund, Y. (2015) The role of RNA conformation in RNA-protein recognition. *RNA biology*, **12**, 720-727.
297. Hudson, G.A., Bloomingdale, R.J. and Znosko, B.M. (2013) Thermodynamic contribution and nearest-neighbor parameters of pseudouridine-adenosine base pairs in oligoribonucleotides. *RNA (New York, N.Y.)*, **19**, 1474-1482.
298. Meroueh, M., Grohar, P.J., Qiu, J., SantaLucia, J., Jr., Scaringe, S.A. and Chow, C.S. (2000) Unique structural and stabilizing roles for the individual pseudouridine residues in the 1920 region of Escherichia coli 23S rRNA. *Nucleic Acids Res*, **28**, 2075-2083.
299. Höfler, S. and Carlomagno, T. (2020) Structural and functional roles of 2'-O-ribose methylations and their enzymatic machinery across multiple classes of RNAs. *Curr Opin Struct Biol*, **65**, 42-50.

300. Ganot, P., Bortolin, M.L. and Kiss, T. (1997) Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell*, **89**, 799-809.
301. Lafontaine, D.L., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. and Tollervey, D. (1998) The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev*, **12**, 527-537.
302. Tollervey, D., Lehtonen, H., Jansen, R., Kern, H. and Hurt, E.C. (1993) Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell*, **72**, 443-457.
303. Huang, G.M., Jarmolowski, A., Struck, J.C. and Fournier, M.J. (1992) Accumulation of U14 small nuclear RNA in Saccharomyces cerevisiae requires box C, box D, and a 5', 3' terminal stem. *Mol Cell Biol*, **12**, 4456-4463.
304. Caffarelli, E., Fatica, A., Prislei, S., De Gregorio, E., Fragapane, P. and Bozzoni, I. (1996) Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA. *Embo j*, **15**, 1121-1131.
305. Henras, A.K., Soudet, J., Gerus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A. and Henry, Y. (2008) The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci*, **65**, 2334-2359.
306. Klein, D.J., Schmeing, T.M., Moore, P.B. and Steitz, T.A. (2001) The kink-turn: a new RNA secondary structure motif. *Embo j*, **20**, 4214-4221.
307. Watkins, N.J., Dickmanns, A. and Lührmann, R. (2002) Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP. *Mol Cell Biol*, **22**, 8342-8352.
308. Kuhn, J.F., Tran, E.J. and Maxwell, E.S. (2002) Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein. *Nucleic Acids Res*, **30**, 931-941.
309. Charron, C., Manival, X., Cléry, A., Senty-Ségault, V., Charpentier, B., Marmier-Gourrier, N., Branlant, C. and Aubry, A. (2004) The archaeal sRNA binding protein L7Ae has a 3D structure very similar to that of its eukaryal counterpart while having a broader RNA-binding specificity. *J Mol Biol*, **342**, 757-773.
310. Szewczak, L.B., DeGregorio, S.J., Strobel, S.A. and Steitz, J.A. (2002) Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP. *Chem Biol*, **9**, 1095-1107.
311. Schultz, A., Nottrott, S., Watkins, N.J. and Lührmann, R. (2006) Protein-protein and protein-RNA contacts both contribute to the 15.5K-mediated assembly of the U4/U6 snRNP and the box C/D snoRNPs. *Mol Cell Biol*, **26**, 5146-5154.
312. Moore, T., Zhang, Y., Fenley, M.O. and Li, H. (2004) Molecular basis of box C/D RNA-protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA. *Structure*, **12**, 807-818.
313. Nottrott, S., Hartmuth, K., Fabrizio, P., Urlaub, H., Vidovic, I., Ficner, R. and Lührmann, R. (1999) Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA. *Embo j*, **18**, 6119-6133.
314. Lafontaine, D.L. and Tollervey, D. (2000) Synthesis and assembly of the box C+D small nucleolar RNPs. *Mol Cell Biol*, **20**, 2650-2659.
315. Aittaleb, M., Rashid, R., Chen, Q., Palmer, J.R., Daniels, C.J. and Li, H. (2003) Structure and function of archaeal box C/D sRNP core proteins. *Nat Struct Biol*, **10**, 256-263.
316. Cahill, N.M., Friend, K., Speckmann, W., Li, Z.H., Terns, R.M., Terns, M.P. and Steitz, J.A. (2002) Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP. *Embo j*, **21**, 3816-3828.
317. Lin, J., Lai, S., Jia, R., Xu, A., Zhang, L., Lu, J. and Ye, K. (2011) Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature*, **469**, 559-563.
318. van Nues, R.W., Granneman, S., Kudla, G., Sloan, K.E., Chicken, M., Tollervey, D. and Watkins, N.J. (2011) Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. *The EMBO Journal*, **30**, 2420-2430.
319. van Nues, R.W. and Watkins, N.J. (2017) Unusual C′/D′ motifs enable box C/D snoRNPs to modify multiple sites in the same rRNA target region. *Nucleic Acids Research*, **45**, 2016-2028.
320. Yang, Z., Lin, J. and Ye, K. (2016) Box C/D guide RNAs recognize a maximum of 10 nt of substrates. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 10878-10883.
321. Leontis, N.B., Stombaugh, J. and Westhof, E. (2002) The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res*, **30**, 3497-3531.
322. Ni, J., Tien, A.L. and Fournier, M.J. (1997) Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. *Cell*, **89**, 565-573.
323. Lange, T.S., Ezrokhi, M., Amaldi, F. and Gerbi, S.A. (1999) Box H and box ACA are nucleolar localization elements of U17 small nucleolar RNA. *Molecular biology of the cell*, **10**, 3877-3890.
324. Narayanan, A., Lukowiak, A., Jády, B.E., Dragon, F., Kiss, T., Terns, R.M. and Terns, M.P. (1999) Nucleolar localization signals of box H/ACA small nucleolar RNAs. *The EMBO journal*, **18**, 5120-5130.

325. Bousquet-Antonelli, C., Henry, Y., G'Elugne, J.P., Caizergues-Ferrer, M. and Kiss, T. (1997) A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *The EMBO journal*, **16**, 4770-4776.
326. Henras, A., Dez, C., Noaillac-Depeyre, J., Henry, Y. and Caizergues-Ferrer, M. (2001) Accumulation of H/ACA snoRNPs depends on the integrity of the conserved central domain of the RNA-binding protein Nhp2p. *Nucleic acids research*, **29**, 2733-2746.
327. Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gélugne, J.P. and Caizergues-Ferrer, M. (1998) Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *The EMBO journal*, **17**, 7078-7090.
328. Zebarjadian, Y., King, T., Fournier, M.J., Clarke, L. and Carbon, J. (1999) Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA. *Mol Cell Biol*, **19**, 7461-7472.
329. Gu, X., Liu, Y. and Santi, D.V. (1999) The mechanism of pseudouridine synthase I as deduced from its interaction with 5-fluorouracil-tRNA. *Proc Natl Acad Sci U S A*, **96**, 14270-14275.
330. Kammen, H.O., Marvel, C.C., Hardy, L. and Penhoet, E.E. (1988) Purification, structure, and properties of Escherichia coli tRNA pseudouridine synthase I. *The Journal of biological chemistry*, **263**, 2255-2263.
331. Mueller, E.G. and Ferrè-d Amarè, A.R. (2013).
332. Veerareddygari, G.R., Singh, S.K. and Mueller, E.G. (2016) The Pseudouridine Synthases Proceed through a Glycal Intermediate. *Journal of the American Chemical Society*, **138**, 7852-7855.
333. Huang, L., Pookanjanatavip, M., Gu, X. and Santi, D.V. (1998) A conserved aspartate of tRNA pseudouridine synthase is essential for activity and a probable nucleophilic catalyst. *Biochemistry*, **37**, 344-351.
334. Kiss, D.J., Oláh, J., Tóth, G., Menyhárd, D.K. and Ferenczy, G.G. (2018) Quantum chemical calculations support pseudouridine synthase reaction through a glycal intermediate and provide details of the mechanism. *Theoretical Chemistry Accounts*, **137**, 162.
335. Charpentier, B., Muller, S. and Branlant, C. (2005) Reconstitution of archaeal H/ACA small ribonucleoprotein complexes active in pseudouridylation. *Nucleic acids research*, **33**, 3133-3144.
336. Wu, H. and Feigon, J. (2007) H/ACA small nucleolar RNA pseudouridylation pockets bind substrate RNA to form three-way junctions that position the target U for modification. *Proc Natl Acad Sci U S A*, **104**, 6655-6660.
337. Li, L. and Ye, K. (2006) Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature*, **443**, 302-307.
338. Hamma, T., Reichow, S.L., Varani, G. and Ferré-D'Amaré, A.R. (2005) The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nat Struct Mol Biol*, **12**, 1101-1107.
339. Rashid, R., Liang, B., Baker, D.L., Youssef, O.A., He, Y., Phipps, K., Terns, R.M., Terns, M.P. and Li, H. (2006) Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol Cell*, **21**, 249-260.
340. Liang, B., Zhou, J., Kahen, E., Terns, R.M., Terns, M.P. and Li, H. (2009) Structure of a functional ribonucleoprotein pseudouridine synthase bound to a substrate RNA. *Nature structural & molecular biology*, **16**, 740-746.
341. Duan, J., Li, L., Lu, J., Wang, W. and Ye, K. (2009) Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase. *Mol Cell*, **34**, 427-439.
342. Ge, J. and Yu, Y.-T. (2013) RNA pseudouridylation: new insights into an old modification. *Trends in biochemical sciences*, **38**, 210-218.
343. Dieci, G., Preti, M. and Montanini, B. (2009) Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics*, **94**, 83-88.
344. Brown, J.W., Echeverria, M. and Qu, L.H. (2003) Plant snoRNAs: functional evolution and new modes of gene expression. *Trends Plant Sci*, **8**, 42-49.
345. Piekna-Przybylska, D., Decatur, W.A. and Fournier, M.J. (2007) New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. *Rna*, **13**, 305-312.
346. Aspegren, A., Hinas, A., Larsson, P., Larsson, A. and Soderbom, F. (2004) Novel non-coding RNAs in Dictyostelium discoideum and their expression during development. *Nucleic Acids Res*, **32**, 4646-4656.
347. Chen, C.L., Chen, C.J., Vallon, O., Huang, Z.P., Zhou, H. and Qu, L.H. (2008) Genomewide analysis of box C/D and box H/ACA snoRNAs in Chlamydomonas reinhardtii reveals an extensive organization into intronic gene clusters. *Genetics*, **179**, 21-30.
348. Li, S.-G., Zhou, H., Luo, Y.-P., Zhang, P. and Qu, L.-H. (2005) Identification and functional analysis of 20 Box H/ACA small nucleolar RNAs (snoRNAs) from Schizosaccharomyces pombe. *The Journal of biological chemistry*, **280**, 16446-16455.
349. Huang, Z.P., Chen, C.J., Zhou, H., Li, B.B. and Qu, L.H. (2007) A combined computational and experimental analysis of two families of snoRNA genes from Caenorhabditis elegans, revealing the expression and evolution pattern of snoRNAs in nematodes. *Genomics*, **89**, 490-501.

350. Huang, Z.P., Zhou, H., He, H.L., Chen, C.L., Liang, D. and Qu, L.H. (2005) Genome-wide analyses of two families of snoRNA genes from Drosophila melanogaster, demonstrating the extensive utilization of introns for coding of snoRNAs. *Rna*, **11**, 1303-1316.
351. Tanaka-Fujita, R., Soeno, Y., Satoh, H., Nakamura, Y. and Mori, S. (2007) Human and mouse protein-noncoding snoRNA host genes with dissimilar nucleotide sequences show chromosomal synteny. *Rna*, **13**, 811-816.
352. Chanfreau, G., Legrain, P. and Jacquier, A. (1998) Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism11Edited by J. Karn. *Journal of Molecular Biology*, **284**, 975-988.
353. Petfalski, E., Dandekar, T., Henry, Y. and Tollervey, D. (1998) Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol Cell Biol*, **18**, 1181-1189.
354. Kufel, J. and Grzechnik, P. (2019) Small Nucleolar RNAs Tell a Different Tale. *Trends in Genetics*, **35**, 104-117.
355. Tycowski, K.T., Aab, A. and Steitz, J.A. (2004) Guide RNAs with 5' caps and novel box C/D snoRNA-like domains for modification of snRNAs in metazoa. *Curr Biol*, **14**, 1985-1995.
356. Deng, W., Zhu, X., Skogerbø, G., Zhao, Y., Fu, Z., Wang, Y., He, H., Cai, L., Sun, H., Liu, C. *et al.* (2006) Organization of the Caenorhabditis elegans small non-coding transcriptome: genomic features, biogenesis, and expression. *Genome Res*, **16**, 20-29.
357. Leader, D.J., Clark, G.P., Watters, J., Beven, A.F., Shaw, P.J. and Brown, J.W. (1997) Clusters of multiple different small nucleolar RNA genes in plants are expressed as and processed from polycistronic pre-snoRNAs. *Embo j*, **16**, 5742-5751.
358. Qu, L.-H., Henras, A., Lu, Y.-J., Zhou, H., Zhou, W.-x., Zhu, Y.-Q., Zhao, J., Henry, Y., Caizergues-Ferrer, M. and Bachellerie, J.-P. (1999) Seven Novel Methylation Guide Small Nucleolar RNAs Are Processed from a Common Polycistronic Transcript by Rat1p and RNase III in Yeast. *Molecular and Cellular Biology*, **19**, 1144-1158.
359. Guffanti, E., Ferrari, R., Preti, M., Forloni, M., Harismendy, O., Lefebvre, O. and Dieci, G. (2006) A minimal promoter for TFIIIC-dependent in vitro transcription of snoRNA and tRNA genes by RNA polymerase III. *The Journal of biological chemistry*, **281**, 23945-23957.
360. Kruszka, K., Barneche, F., Guyot, R., Ailhas, J., Meneau, I., Schiffer, S., Marchfelder, A. and Echeverría, M. (2003) Plant dicistronic tRNA-snoRNA genes: a new mode of expression of the small nucleolar RNAs processed by RNase Z. *Embo j*, **22**, 621-632.
361. Isogai, Y., Takada, S., Tjian, R. and Keleş, S. (2007) Novel TRF1/BRF target genes revealed by genome-wide analysis of Drosophila Pol III transcription. *Embo j*, **26**, 79-89.
362. Makarova, J.A. and Kramerov, D.A. (2011) SNOntology: Myriads of novel snoRNAs or just a mirage? *BMC Genomics*, **12**, 543.
363. Lykke-Andersen, S., Chen, Y., Ardal, B.R., Lilje, B., Waage, J., Sandelin, A. and Jensen, T.H. (2014) Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. *Genes Dev*, **28**, 2498-2517.
364. Schulz, D., Schwalb, B., Kiesel, A., Baejen, C., Torkler, P., Gagneur, J., Soeding, J. and Cramer, P. (2013) Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell*, **155**, 1075-1087.
365. Porrua, O., Hobor, F., Boulay, J., Kubicek, K., D'Aubenton-Carafa, Y., Gudipati, R.K., Stefl, R. and Libri, D. (2012) In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. *Embo j*, **31**, 3935-3948.
366. Grzechnik, P. and Kufel, J. (2008) Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. *Mol Cell*, **32**, 247-258.
367. Vasiljeva, L. and Buratowski, S. (2006) Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Mol Cell*, **21**, 239-248.
368. Ooi, S.L., Samarsky, D.A., Fournier, M.J. and Boeke, J.D. (1998) Intronic snoRNA biosynthesis in Saccharomyces cerevisiae depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. *Rna*, **4**, 1096-1110.
369. Ghazal, G., Ge, D., Gervais-Bird, J., Gagnon, J. and Abou Elela, S. (2005) Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. *Mol Cell Biol*, **25**, 2981-2994.
370. Rasmussen, E.B. and Lis, J.T. (1993) In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. *Proc Natl Acad Sci U S A*, **90**, 7923-7927.
371. Lee, C.Y., Lee, A. and Chanfreau, G. (2003) The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5'-end processing of S. cerevisiae box C/D snoRNAs. *Rna*, **9**, 1362-1370.
372. Ojha, S., Malla, S. and Lyons, S.M. (2020) snoRNPs: Functions in Ribosome Biogenesis. *Biomolecules*, **10**, 783.
373. Kass, S., Tyc, K., Steitz, J.A. and Sollner-Webb, B. (1990) The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. *Cell*, **60**, 897-908.
374. Li, H.D., Zagorski, J. and Fournier, M.J. (1990) Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in Saccharomyces cerevisiae. *Mol Cell Biol*, **10**, 1145-1152.

375. Tollervey, D. (1987) A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. *Embo j*, **6**, 4169-4175.
376. Morrissey, J.P. and Tollervey, D. (1993) Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol Cell Biol*, **13**, 2469-2477.
377. Tycowski, K.T., Shu, M.D. and Steitz, J.A. (1994) Requirement for intron-encoded U22 small nucleolar RNA in 18S ribosomal RNA maturation. *Science*, **266**, 1558-1561.
378. Schmitt, M.E. and Clayton, D.A. (1993) Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in Saccharomyces cerevisiae. *Mol Cell Biol*, **13**, 7935-7941.
379. Watkins, N.J., Ségault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C. and Lührmann, R. (2000) A Common Core RNP Structure Shared between the Small Nucleoar Box C/D RNPs and the Spliceosomal U4 snRNP. *Cell*, **103**, 457-466.
380. Canzler, S., Stadler, P.F. and Hertel, J. (2017) Evolution of Fungal U3 snoRNAs: Structural Variation and Introns. *Non-Coding RNA*, **3**.
381. Phipps, K.R., Charette, J. and Baserga, S.J. (2011) The small subunit processome in ribosome biogenesis—progress and prospects. *Wiley Interdiscip Rev RNA*, **2**, 1-21.
382. Hughes, J.M. and Ares, M., Jr. (1991) Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. *Embo j*, **10**, 4231-4239.
383. Dragon, F., Gallagher, J.E., Compagnone-Post, P.A., Mitchell, B.M., Porwancher, K.A., Wehner, K.A., Wormsley, S., Settlage, R.E., Shabanowitz, J., Osheim, Y. *et al.* (2002) A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature*, **417**, 967-970.
384. Dutca, L.M., Gallagher, J.E.G. and Baserga, S.J. (2011) The initial U3 snoRNA:pre-rRNA base pairing interaction required for pre-18S rRNA folding revealed by in vivo chemical probing. *Nucleic acids research*, **39**, 5164-5180.
385. Hellung-Larsen, P., Kulamowicz, I. and Frederiksen, S. (1980) Synthesis of low molecular weight RNA components in cells with a temperature-sensitive polymerase II. *Biochim Biophys Acta*, **609**, 201-204.
386. Chandrasekharappa, S.C., Smith, J.H. and Eliceiri, G.L. (1983) Biosynthesis of small nuclear RNAs in human cells. *J Cell Physiol*, **117**, 169-174.
387. Kiss, T., Marshallsay, C. and Filipowicz, W. (1991) Alteration of the RNA polymerase specificity of U3 snRNA genes during evolution and in vitro. *Cell*, **65**, 517-526.
388. Antal, M., Mougin, A., Kis, M., Boros, E., Steger, G., Jakab, G., Solymosy, F. and Branlant, C. (2000) Molecular characterization at the RNA and gene levels of U3 snoRNA from a unicellular green alga, Chlamydomonas reinhardtii. *Nucleic Acids Research*, **28**, 2959-2968.
389. Reddy, R., Ro-Choi, T.S., Henning, D., Shibata, H., Choi, Y.C. and Busch, H. (1972) MOdified nucleosides of nuclear and nucleolar low molecular weight ribonucleic acid. *The Journal of biological chemistry*, **247**, 7245-7250.
390. Jawdekar, G.W. and Henry, R.W. (2008) Transcriptional regulation of human small nuclear RNA genes. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1779**, 295-305.
391. Nabavi, S. and Nazar, R.N. (2008) U3 snoRNA promoter reflects the RNA's function in ribosome biogenesis. *Current genetics*, **54**, 175-184.
392. Tollervey, D., Lehtonen, H., Carmo-Fonseca, M. and Hurt, E.C. (1991) The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. *Embo j*, **10**, 573-583.
393. Parker, K.A. and Steitz, J.A. (1987) Structural analysis of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA. *Mol Cell Biol*, **7**, 2899-2913.
394. Venema, J., Vos, H.R., Faber, A.W., van Venrooij, W.J. and Raué, H.A. (2000) Yeast Rrp9p is an evolutionarily conserved U3 snoRNP protein essential for early pre-rRNA processing cleavages and requires box C for its association. *Rna*, **6**, 1660-1671.
395. Lübben, B., Marshallsay, C., Rottmann, N. and Lührmann, R. (1993) Isolation of U3 snoRNP from CHO cells: a novel 55 kDa protein binds to the central part of U3 snoRNA. *Nucleic Acids Res*, **21**, 5377-5385.
396. Clerget, G., Bourguignon-Igel, V., Marmier-Gourrier, N., Rolland, N., Wacheul, L., Manival, X., Charron, C., Kufel, J., Méreau, A., Senty-Ségault, V. *et al.* (2020) Synergistic defects in pre-rRNA processing from mutations in the U3-specific protein Rrp9 and U3 snoRNA. *Nucleic acids research*, **48**, 3848-3868.
397. Jansen, R., Tollervey, D. and Hurt, E.C. (1993) A U3 snoRNP protein with homology to splicing factor PRP4 and G beta domains is required for ribosomal RNA processing. *Embo j*, **12**, 2549-2558.
398. Dunbar, D.A., Wormsley, S., Agentis, T.M. and Baserga, S.J. (1997) Mpp10p, a U3 small nucleolar ribonucleoprotein component required for pre-18S rRNA processing in yeast. *Mol Cell Biol*, **17**, 5803-5812.
399. Marchand, V., Pichot, F., Neybecker, P., Ayadi, L., Bourguignon-Igel, V., Wacheul, L., Lafontaine, D.L.J., Pinzano, A., Helm, M. and Motorin, Y. (2020) HydraPsiSeq: a method for systematic and quantitative mapping of pseudouridines in RNA. *Nucleic Acids Research*, **48**, e110-e110.

400. Lowe, T.M. and Eddy, S.R. (1999) A Computational Screen for methylation Guide snoRNAs in Yeast. *Science*, **283**, 1168–1171.
401. Lorenz, R., Bernhart, S.H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F. and Hofacker, I.L. (2011) ViennaRNA Package 2.0. *Algorithms for Molecular Biology*, **6**, 26.
402. Schäck, M. (2018) Elongator dependent anticodon modifications in *Dictyostelium discoideum* and the effect on glutamine codon translation.
403. Kruse, J.M. (2015) Dicer-like proteins in *Dictyostelium discoideum*.
404. Malicki, M. (2017) The Retrotransposon Silencing Complex (RSC) is a key repressor of retrotransposons in *Dictyostelium discoideum*.
405. Avesson, L., Reimegård, J., Wagner, E.G. and Söderbom, F. (2012) MicroRNAs in Amoebozoa: deep sequencing of the small RNA population in the social amoeba Dictyostelium discoideum reveals developmentally regulated microRNAs. *Rna*, **18**, 1771-1782.
406. Cornillon, S., Pech, E., Benghezal, M., Ravanel, K., Gaynor, E., Letourneur, F., Brückert, F. and Cosson, P. (2000) Phg1p is a nine-transmembrane protein superfamily member involved in dictyostelium adhesion and phagocytosis. *The Journal of biological chemistry*, **275**, 34287-34292.
407. Wiegand, S. and Hammann, C. (2013) The 5' spreading of small RNAs in Dictyostelium discoideum depends on the RNA-dependent RNA polymerase RrpC and on the dicer-related nuclease DrnB. *PLOS ONE*, **8**, e64804.
408. Boesler, B., Meier, D., Foerstner, K.U., Friedrich, M., Hammann, C., Sharma, C.M. and Nellen, W. (2014) Argonaute proteins affect siRNA levels and accumulation of a novel extrachromosomal DNA from the Dictyostelium retrotransposon DIRS-1. *The Journal of biological chemistry*, **289**, 35124-35138.
409. Park, L., Thomason, P.A., Zech, T., King, J.S., Veltman, D.M., Carnell, M., Ura, S., Machesky, L.M. and Insall, R.H. (2013) Cyclical action of the WASH complex: FAM21 and capping protein drive WASH recycling, not initial recruitment. *Dev Cell*, **24**, 169-181.
410. Peracino, B., Buracco, S. and Bozzaro, S. (2013) The Nramp (Slc11) proteins regulate development, resistance to pathogenic bacteria and iron homeostasis in Dictyostelium discoideum. *J Cell Sci*, **126**, 301-311.
411. Somesh, B.P., Neffgen, C., Iijima, M., Devreotes, P. and Rivero, F. (2006) Dictyostelium RacH regulates endocytic vesicular trafficking and is required for localization of vacuolin. *Traffic*, **7**, 1194-1212.
412. Benghezal, M., Fauvarque, M.O., Tournebize, R., Froquet, R., Marchetti, A., Bergeret, E., Lardy, B., Klein, G., Sansonetti, P., Charette, S.J. *et al.* (2006) Specific host genes required for the killing of Klebsiella bacteria by phagocytes. *Cell Microbiol*, **8**, 139-148.
413. Müller, I., Subert, N., Otto, H., Herbst, R., Rühling, H., Maniak, M. and Leippe, M. (2005) A Dictyostelium mutant with reduced lysozyme levels compensates by increased phagocytic activity. *The Journal of biological chemistry*, **280**, 10435-10443.
414. Loovers, H.M., Veenstra, K., Snippe, H., Pesesse, X., Erneux, C. and van Haastert, P.J. (2003) A diverse family of inositol 5-phosphatases playing a role in growth and development in Dictyostelium discoideum. *The Journal of biological chemistry*, **278**, 5652-5658.
415. Fang, J., Brzostowski, J.A., Ou, S., Isik, N., Nair, V. and Jin, T. (2007) A vesicle surface tyrosine kinase regulates phagosome maturation. *Journal of Cell Biology*, **178**, 411-423.
416. Abrahão, J., Silva, L., Silva, L.S., Khalil, J.Y.B., Rodrigues, R., Arantes, T., Assis, F., Boratto, P., Andrade, M., Kroon, E.G. *et al.* (2018) Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nature Communications*, **9**, 749.
417. Andreani, J., Aherfi, S., Bou Khalil, J.Y., Di Pinto, F., Bitam, I., Raoult, D., Colson, P. and La Scola, B. (2016) Cedratvirus, a Double-Cork Structured Giant Virus, is a Distant Relative of Pithoviruses. *Viruses*, **8**, 300.
418. Neuhaus, E.M., Horstmann, H., Almers, W., Maniak, M. and Soldati, T. (1998) Ethane-freezing/methanol-fixation of cell monolayers: a procedure for improved preservation of structure and antigenicity for light and electron microscopies. *J Struct Biol*, **121**, 326-342.
419. Ravanel, K., de Chassey, B., Cornillon, S., Benghezal, M., Zulianello, L., Gebbie, L., Letourneur, F. and Cosson, P. (2001) Membrane sorting in the endocytic and phagocytic pathway of Dictyostelium discoideum. *Eur J Cell Biol*, **80**, 754-764.
420. Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841-842.
421. Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T. *et al.* (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nature methods*, **12**, 115-121.
422. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403-410.
423. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357-359.

424. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011*, **17**, 3.
425. Gandrud, C. (2016) DataCombine: Tools for Easily Combining and Cleaning Data Sets.
426. The EndNote Team. (2013). EndNote X9 ed. Clarivate, Philadelphia, PA.
427. Andrews, S. (2015) FastQC: A Quality Control Tool for High Throughput Sequence Data.
428. Liao, Y., Smyth, G.K. and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923-930.
429. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nature Methods*, **9**, 676-682.
430. Becton Dickinson and Company. (2021). 10.8.0 ed. Becton, Dickinson and Company, Ashland.
431. GraphPad Software. (2021). 9.1.1 ed. GraphPad Software, San Diego, California, USA.
432. Ramakers, C., Ruijter, J.M., Deprez, R.H. and Moorman, A.F. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett*, **339**, 62-66.
433. Ruijter, J.M., Ramakers, C., Hoogaars, W.M., Karlen, Y., Bakker, O., van den Hoff, M.J. and Moorman, A.F. (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res*, **37**, e45.
434. Nawrocki, E.P. and Eddy, S.R. (2013) Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*, **29**, 2933-2935.
435. Katoh, K. and Standley, D.M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, **30**, 772-780.
436. Bailey, T.L., Johnson, J., Grant, C.E. and Noble, W.S. (2015) The MEME Suite. *Nucleic Acids Research*, **43**, W39-W49.
437. Anaconda Inc. (2020). Anaconda Inc., Anaconda Documentation.
438. Kang, W., Eldfjell, Y., Fromm, B., Estivill, X., Biryukova, I. and Friedländer, M.R. (2018) miRTrace reveals the organismal origins of microRNA sequencing data. *Genome Biology*, **19**, 213.
439. Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D. *et al.* (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res*, **47**, W636-W641.
440. Weinberg, Z. and Breaker, R.R. (2011) R2R - software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinformatics*, **12**, 3.
441. R Core Team. (2020) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
442. RStudio Team. (2020) *RStudio: Integrated Development Environment for R*. RStudio, PBC.
443. Gu, Z., Eils, R. and Schlesner, M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, **32**, 2847-2849.
444. Sakai, R. (2015) dendsort: Modular Leaf Ordering Methods for Dendrogram Nodes.
445. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.
446. Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T. and Carey, V.J. (2013) Software for Computing and Annotating Genomic Ranges. *PLOS Computational Biology*, **9**, e1003118.
447. Arora, S.M., Martin; Carlson, Marc; Pagès, H. (2020) GenomeInfoDb: Utilities for manipulating chromosome names, including modifying
them to follow a particular naming style.
448. Wickham, H. (2016) *Ggplot2: Elegant graphics for data analysis*. 2 ed. Springer International Publishing.
449. Neitmann, T. (2020) *mdthemes: Markdown Themes for 'ggplot2'*.
450. Neuwirth, E. (2014) *RColorBrewer: ColorBrewer Palettes*.
451. Müller, K.W., Hadley. (2020). R package version 3.0.3 ed.
452. Rehmsmeier, M., Steffen, P., Hochsmann, M. and Giegerich, R. (2004) Fast and effective prediction of microRNA/target duplexes. *RNA*, **10**, 1507-1517.
453. De Rijk, P., Wuyts, J. and De Wachter, R. (2003) RnaViz 2: an improved representation of RNA secondary structure. *Bioinformatics*, **19**, 299-300.
454. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.
455. GSL Biotech LLC. (2021).
456. Aylward, F.O. and Moniruzzaman, M. (2021) ViralRecall—A Flexible Command-Line Tool for the Detection of Giant Virus Signatures in 'Omic Data. *Viruses*, **13**, 150.
457. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res*, **14**, 1188-1190.
458. Hagedorn, M., Neuhaus, E.M. and Soldati, T. (2006) Optimized fixation and immunofluorescence staining methods for Dictyostelium cells. *Methods Mol Biol*, **346**, 327-338.

459. Green, M.R. and Sambrook, J. (2012) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.
460. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol*, **51 Pt 1**, 263-273.
461. Laemmli, U.K. (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4. *Nature*, **227**, 680-685.
462. Liao, Z., Kjellin, J., Hoeppner, M.P., Grabherr, M. and Soderbom, F. (2018) Global characterization of the Dicer-like protein DrnB roles in miRNA biogenesis in the social amoeba Dictyostelium discoideum. *RNA biology*, **15**, 937-954.
463. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*, **10**, R25.
464. Krogh, N., Birkedal, U. and Nielsen, H. (2017) RiboMeth-seq: Profiling of 2'-O-Me in RNA. *Methods Mol Biol*, **1562**, 189-209.
465. Pichot, F., Marchand, V., Ayadi, L., Bourguignon-Igel, V., Helm, M. and Motorin, Y. (2020) Holistic Optimization of Bioinformatic Analysis Pipeline for Detection and Quantification of 2'-O-Methylations in RNA by RiboMethSeq. *Front Genet*, **11**, 38.
466. Kalvari, I., Nawrocki, E.P., Ontiveros-Palacios, N., Argasinska, J., Lamkiewicz, K., Marz, M., Griffiths-Jones, S., Toffano-Nioche, C., Gautheret, D., Weinberg, Z. *et al.* (2021) Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Research*, **49**, D192-D200.
467. Marz, M. and Stadler, P.F. (2009) Comparative analysis of eukaryotic U3 snoRNA. *RNA biology*, **6**, 503-507.
468. Lamrabet, O., Melotti, A., Burdet, F., Hanna, N., Perrin, J., Nitschke, J., Pagni, M., Hilbi, H., Soldati, T. and Cosson, P. (2020) Transcriptional Responses of Dictyostelium discoideum Exposed to Different Classes of Bacteria. *Frontiers in microbiology*, **11**, 410-410.
469. Mantzouranis, L., Bagattini, R. and Souza, G.M. (2010) KeaA, a Dictyostelium kelch-domain protein that regulates the response to stress and development. *BMC Developmental Biology*, **10**, 79.
470. Pan, M., Xu, X., Chen, Y. and Jin, T. (2016) Identification of a Chemoattractant G-Protein-Coupled Receptor for Folic Acid that Controls Both Chemotaxis and Phagocytosis. *Developmental cell*, **36**, 428-439.
471. Schrad, J.R., Abrahão, J.S., Cortines, J.R. and Parent, K.N. (2020) Structural and Proteomic Characterization of the Initiation of Giant Virus Infection. *Cell*, **181**, 1046-1061.e1046.
472. Giorgione, J. and Clarke, M. (2008) Heterogeneous modes of uptake for latex beads revealed through live cell imaging of phagocytes expressing a probe for phosphatidylinositol-(3,4,5)-trisphosphate and phosphatidylinositol-(3,4)-bisphosphate. *Cell Motil Cytoskeleton*, **65**, 721-733.
473. Drengk, A., Fritsch, J., Schmauch, C., Rühling, H. and Maniak, M. (2003) A coat of filamentous actin prevents clustering of late-endosomal vacuoles in vivo. *Curr Biol*, **13**, 1814-1819.
474. Bosmani, C., Leuba, F., Hanna, N., Bach, F., Burdet, F., Pagni, M., Hagedorn, M. and Soldati, T. (2020) Vacuolins and myosin VII are required for phagocytic uptake and phagosomal membrane recycling in Dictyostelium discoideum. *Journal of Cell Science*, **133**.
475. Balest, A., Peracino, B. and Bozzaro, S. (2011) Legionella pneumophila infection is enhanced in a RacH-null mutant of Dictyostelium. *Commun Integr Biol*, **4**, 194-197.
476. Buracco, S., Peracino, B., Andreini, C., Bracco, E. and Bozzaro, S. (2017) Differential Effects of Iron, Zinc, and Copper on Dictyostelium discoideum Cell Growth and Resistance to Legionella pneumophila. *Front Cell Infect Microbiol*, **7**, 536.
477. Weber, S.S., Ragaz, C. and Hilbi, H. (2009) The inositol polyphosphate 5-phosphatase OCRL1 restricts intracellular growth of Legionella, localizes to the replicative vacuole and binds to the bacterial effector LpnE. *Cell Microbiol*, **11**, 442-460.
478. Otto, G.P., Wu, M.Y., Kazgan, N., Anderson, O.R. and Kessin, R.H. (2004) Dictyostelium macroautophagy mutants vary in the severity of their developmental defects. *The Journal of biological chemistry*, **279**, 15621-15629.
479. Wiegand, S., Meier, D., Seehafer, C., Malicki, M., Hofmann, P., Schmith, A., Winckler, T., Foldesi, B., Boesler, B., Nellen, W. *et al.* (2014) The Dictyostelium discoideum RNA-dependent RNA polymerase RrpC silences the centromeric retrotransposon DIRS-1 post-transcriptionally and is required for the spreading of RNA silencing signals. *Nucleic Acids Res*, **42**, 3330-3345.
480. Hebras, J., Krogh, N., Marty, V., Nielsen, H. and Cavaillé, J. (2020) Developmental changes of rRNA ribose methylations in the mouse. *RNA biology*, **17**, 150-164.
481. Krogh, N. and Nielsen, H. (2019) Sequencing-based methods for detection and quantitation of ribose methylations in RNA. *Methods*, **156**, 5-15.
482. Krogh, N., Asmar, F., Côme, C., Munch-Petersen, H.F., Grønbæk, K. and Nielsen, H. (2020) Profiling of ribose methylations in ribosomal RNA from diffuse large B-cell lymphoma patients for evaluation of ribosomes as drug targets. *NAR Cancer*, **2**.

483. Azevedo-Favory, J., Gaspin, C., Ayadi, L., Montacie, C., Marchand, V., Jobet, E., Rompais, M., Carapito, C., Motorin, Y. and Saez-Vasquez, J. (2021) Mapping rRNA 2'-O-methylations and identification of C/D snoRNAs in Arabidopsis thaliana plants. *RNA biology*, 1-18.
484. Kruse, J., Meier, D., Zenk, F., Rehders, M., Nellen, W. and Hammann, C. (2016) The protein domains of the *Dictyostelium* microprocessor that are required for correct subcellular localization and for microRNA maturation. *RNA biology*, **13**, 1000-1010.
485. Meier, D., Kruse, J., Buttlar, J., Friedrich, M., Zenk, F., Boesler, B., Förstner, K.U., Hammann, C. and Nellen, W. (2016) Analysis of the Microprocessor in *Dictyostelium*: The Role of RbdB, a dsRNA Binding Protein. *PLoS Genetics*, **12**, e1006057.
486. Weis, F., Giudice, E., Churcher, M., Jin, L., Hilcenko, C., Wong, C.C., Traynor, D., Kay, R.R. and Warren, A.J. (2015) Mechanism of eIF6 release from the nascent 60S ribosomal subunit. *Nature Structural & Molecular Biology*, **22**, 914-919.
487. Ramesh, M. and Woolford, J.L., Jr. (2016) Eukaryote-specific rRNA expansion segments function in ribosome biogenesis. *RNA*, **22**, 1153-1162.
488. Petrov, A.S., Bernier, C.R., Gulen, B., Waterbury, C.C., Hershkovits, E., Hsiao, C., Harvey, S.C., Hud, N.V., Fox, G.E., Wartell, R.M. *et al.* (2014) Secondary structures of rRNAs from all three domains of life. *PLoS One*, **9**, e88222.
489. Westhof, E., Yusupov, M. and Yusupova, G. (2019) The multiple flavors of GoU pairs in RNA. *Journal of molecular recognition : JMR*, **32**, e2782.
490. Carlile, T.M., Rojas-Duran, M.F. and Gilbert, W.V. (2015) Pseudo-Seq: Genome-Wide Detection of Pseudouridine Modifications in RNA. *Methods Enzymol*, **560**, 219-245.
491. Sun, L., Xu, Y., Bai, S., Bai, X., Zhu, H., Dong, H., Wang, W., Zhu, X., Hao, F. and Song, C.-P. (2019) Transcriptome-wide analysis of pseudouridylation of mRNA and non-coding RNAs in Arabidopsis. *Journal of Experimental Botany*, **70**, 5089-5600.
492. Bowman, J.C., Petrov, A.S., Frenkel-Pinter, M., Penev, P.I. and Williams, L.D. (2020) Root of the Tree: The Significance, Evolution, and Origins of the Ribosome. *Chemical Reviews*, **120**, 4848-4878.
493. Baxter-Roshek, J.L., Petrov, A.N. and Dinman, J.D. (2007) Optimization of ribosome structure and function by rRNA base modification. *PloS one*, **2**, e174-e174.
494. Katoh, M., Chen, G., Roberge, E., Shaulsky, G. and Kuspa, A. (2007) Developmental commitment in Dictyostelium discoideum. *Eukaryotic cell*, **6**, 2038-2045.
495. Rinke-Appel, J., Jünke, N., Osswald, M. and Brimacombe, R. (1995) The ribosomal environment of tRNA: crosslinks to rRNA from positions 8 and 20:1 in the central fold of tRNA located at the A, P, or E site. *Rna*, **1**, 1018-1028.
496. Bhangu, R. and Wollenzien, P. (1992) The mRNA binding track in the Escherichia coli ribosome for mRNAs of different sequences. *Biochemistry*, **31**, 5937-5944.
497. Wollenzien, P., Expert-Bezançon, A. and Favre, A. (1991) Sites of contact of mRNA with 16S rRNA and 23S rRNA in the Escherichia coli ribosome. *Biochemistry*, **30**, 1788-1795.
498. Meyer, B., Wurm, J.P., Sharma, S., Immer, C., Pogoryelov, D., Kötter, P., Lafontaine, Denis L.J., Wöhnert, J. and Entian, K.-D. (2016) Ribosome biogenesis factor Tsr3 is the aminocarboxypropyl transferase responsible for 18S rRNA hypermodification in yeast and humans. *Nucleic Acids Research*, **44**, 4304-4316.
499. Samarsky, D.A., Balakin, A.G. and Fournier, M.J. (1995) Characterization of three new snRNAs from Saccharomyces cerevisiae: snR34, snR35 and snR36. *Nucleic Acids Res*, **23**, 2548-2554.
500. Wise, J.A. and Weiner, A.M. (1980) Dictyostelium small nuclear RNA D2 is homologous to rat nucleolar RNA U3 and is encoded by a dispersed multigene family. *Cell*, **22**, 109-118.
501. Loomis, W.F., Jr. (1971) Sensitivity of Dictyostelium discoideum to nucleic acid analogues. *Exp Cell Res*, **64**, 484-486.
502. Sussman, R. and Sussman, M. (1967) Cultivation of Dictyostelium discoideum in axenic medium. *Biochem Biophys Res Commun*, **29**, 53-55.
503. Borovjagin, A.V. and Gerbi, S.A. (2004) Xenopus U3 snoRNA docks on pre-rRNA through a novel base-pairing interaction. *Rna*, **10**, 942-953.
504. Granneman, S., Kudla, G., Petfalski, E. and Tollervey, D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proceedings of the National Academy of Sciences*, **106**, 9613-9618.
505. Loomis, W.F. (2014) Cell signaling during development of Dictyostelium. *Dev Biol*, **391**, 1-16.
506. Sucgang, R., Kuo, A., Tian, X., Salerno, W., Parikh, A., Feasley, C.L., Dalin, E., Tu, H., Huang, E., Barry, K. *et al.* (2011) Comparative genomics of the social amoebae Dictyostelium discoideum and Dictyostelium purpureum. *Genome biology*, **12**, R20.
507. Glöckner, G., Lawal, H.M., Felder, M., Singh, R., Singer, G., Weijer, C.J. and Schaap, P. (2016) The multicellularity genes of dictyostelid social amoebas. *Nature communications*, **7**, 12085-12085.

508. Heidel, A.J., Lawal, H.M., Felder, M., Schilde, C., Helps, N.R., Tunggal, B., Rivero, F., John, U., Schleicher, M., Eichinger, L. *et al.* (2011) Phylogeny-wide analysis of social amoeba genomes highlights ancient origins for complex intercellular communication. *Genome Res*, **21**, 1882-1891.
509. Urushihara, H., Kuwayama, H., Fukuhara, K., Itoh, T., Kagoshima, H., Shin, I.T., Toyoda, A., Ohishi, K., Taniguchi, T., Noguchi, H. *et al.* (2015) Comparative genome and transcriptome analyses of the social amoeba Acytostelium subglobosum that accomplishes multicellular development without germ-soma differentiation. *BMC Genomics*, **16**, 80.
510. Schilde, C., Lawal, H.M., Kin, K., Shibano-Hayakawa, I., Inouye, K. and Schaap, P. (2019) A well supported multi gene phylogeny of 52 dictyostelia. *Molecular Phylogenetics and Evolution*, **134**, 66-73.
511. Hinas, A., Larsson, P., Avesson, L., Kirsebom, L.A., Virtanen, A. and Söderbom, F. (2006) Identification of the major spliceosomal RNAs in Dictyostelium discoideum reveals developmentally regulated U2 variants and polyadenylated snRNAs. *Eukaryot Cell*, **5**, 924-934.
512. Larsson, P., Hinas, A., Ardell, D.H., Kirsebom, L.A., Virtanen, A. and Söderbom, F. (2008) De novo search for non-coding RNA genes in the AT-rich genome of Dictyostelium discoideum: performance of Markov-dependent genome feature scoring. *Genome research*, **18**, 888-899.
513. Kjellin, J., Avesson, L., Reimegard, J., Liao, Z., Eichinger, L., Noegel, A., Glockner, G., Schaap, P. and Soderbom, F. (2021) Abundantly expressed class of noncoding RNAs conserved through the multicellular evolution of dictyostelid social amoebae. *Genome Res*.
514. Hinas, A. and Söderbom, F. (2007) Treasure hunt in an amoeba: non-coding RNAs in Dictyostelium discoideum. *Current genetics*, **51**, 141-159.
515. Richard, P. and Manley, J.L. (2009) Transcription termination by nuclear RNA polymerases. *Genes Dev*, **23**, 1247-1269.
516. Weisman, R.A. and Korn, E.D. (1967) Phagocytosis of Latex Beads by Acanthamoeba. I. Biochemical Properties*. *Biochemistry*, **6**, 485-497.
517. Chambers, J.A. and Thompson, J.E. (1976) Phagocytosis and Pinocytosis in Acanthamoeba castellanii. *Microbiology*, **92**, 246-250.
518. Andrade, A.C.D.S.P., Rodrigues, R.A.L., Oliveira, G.P., Andrade, K.R., Bonjardim, C.A., La Scola, B., Kroon, E.G. and Abrahão, J.S. (2017) Filling Knowledge Gaps for Mimivirus Entry, Uncoating, and Morphogenesis. *Journal of virology*, **91**, e01335-01317.
519. Gruenheit, N., Baldwin, A., Stewart, B., Jaques, S., Keller, T., Parkinson, K., Salvidge, W., Baines, R., Brimson, C., Wolf, J.B. *et al.* (2021) Mutant resources for functional genomics in Dictyostelium discoideum using REMI-seq technology. *BMC Biology*, **19**, 172.
520. Weber, S., Wagner, M., Hilbi, H., Swanson, J. and Swanson, M. (2014) Live-Cell Imaging of Phosphoinositide Dynamics and Membrane Architecture during Legionella Infection. *mBio*, **5**, e00839-00813.
521. Lemmon, M.A. and Schlessinger, J. (2010) Cell Signaling by Receptor Tyrosine Kinases. *Cell*, **141**, 1117-1134.
522. Aderem, A. and Underhill, D.M. (1999) Mechanisms of phagocytosis in macrophages. *Annu Rev Immunol*, **17**, 593-623.
523. Cougoule, C., Hoshino, S., Dart, A., Lim, J. and Caron, E. (2006) Dissociation of recruitment and activation of the small G-protein Rac during Fcgamma receptor-mediated phagocytosis. *The Journal of biological chemistry*, **281**, 8756-8764.
524. Stuart, L.M. and Ezekowitz, R.A. (2005) Phagocytosis: elegant complexity. *Immunity*, **22**, 539-550.
525. Lamrabet, O., Jauslin, T., Lima, W.C., Leippe, M. and Cosson, P. (2020) The multifarious lysozyme arsenal of Dictyostelium discoideum. *Developmental & Comparative Immunology*, **107**, 103645.
526. Nasser, W., Santhanam, B., Miranda, E.R., Parikh, A., Juneja, K., Rot, G., Dinh, C., Chen, R., Zupan, B., Shaulsky, G. *et al.* (2013) Bacterial discrimination by dictyostelid amoebae reveals the complexity of ancient interspecies interactions. *Curr Biol*, **23**, 862-872.
527. Schleifer, K.H. and Kandler, O. (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriological Reviews*, **36**, 407-477.
528. Boyer, M., Azza, S., Barrassi, L., Klose, T., Campocasso, A., Pagnier, I., Fournous, G., Borg, A., Robert, C., Zhang, X. *et al.* (2011) Mimivirus shows dramatic genome reduction after intraamoebal culture. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 10296–10301.
529. Koller, B., Schramm, C., Siebert, S., Triebel, J., Deland, E., Pfefferkorn, A.M., Rickerts, V. and Thewes, S. (2016) Dictyostelium discoideum as a Novel Host System to Study the Interaction between Phagocytes and Yeasts. *Frontiers in Microbiology*, **7**.
530. Ostrowski, E.A., Shen, Y., Tian, X., Sucgang, R., Jiang, H., Qu, J., Katoh-Kurasawa, M., Brock, D.A., Dinh, C., Lara-Garduno, F. *et al.* (2015) Genomic signatures of cooperation and conflict in the social amoeba. *Curr Biol*, **25**, 1661-1665.
531. Filée, J. and Chandler, M. (2010) Gene Exchange and the Origin of Giant Viruses. *Intervirology*, **53**, 354-361.

532. Monier, A., Pagarete, A., de Vargas, C., Allen, M.J., Read, B., Claverie, J.-M. and Ogata, H. (2009) Horizontal gene transfer of an entire metabolic pathway between a eukaryotic alga and its DNA virus. *Genome research*, **19**, 1441-1449.
533. Schulz, F., Alteio, L., Goudeau, D., Ryan, E.M., Yu, F.B., Malmstrom, R.R., Blanchard, J. and Woyke, T. (2018) Hidden diversity of soil giant viruses. *Nature Communications*, **9**, 4881.
534. Zaringhalam, M. and Papavasiliou, F.N. (2016) Pseudouridylation meets next-generation sequencing. *Methods*, **107**, 63-72.
535. Kos, M. and Tollervey, D. (2010) Yeast pre-rRNA processing and modification occur cotranscriptionally. *Mol Cell*, **37**, 809-820.
536. Rajan, K.S., Zhu, Y., Adler, K., Doniger, T., Cohen-Chalamish, S., Srivastava, A., Shalev-Benami, M., Matzov, D., Unger, R., Tschudi, C. *et al.* (2020) The large repertoire of 2'-O-methylation guided by C/D snoRNAs on Trypanosoma brucei rRNA. *RNA biology*, **17**, 1018-1039.
537. Marchand, V., Blanloeil-Oillo, F., Helm, M. and Motorin, Y. (2016) Illumina-based RiboMethSeq approach for mapping of 2'-O-Me residues in RNA. *Nucleic Acids Res*, **44**, e135.
538. Jeeninga, R.E., Van Delft, Y., de Graaff-Vincent, M., Dirks-Mulder, A., Venema, J. and Raué, H.A. (1997) Variable regions V13 and V3 of Saccharomyces cerevisiae contain structural features essential for normal biogenesis and stability of 5.8S and 25S rRNA. *RNA (New York, N.Y.)*, **3**, 476-488.
539. Gómez Ramos, L.M., Smeekens, J.M., Kovacs, N.A., Bowman, J.C., Wartell, R.M., Wu, R. and Williams, L.D. (2016) Yeast rRNA Expansion Segments: Folding and Function. *Journal of Molecular Biology*, **428**, 4048-4059.
540. Parker, M.S., Balasubramaniam, A., Sallee, F.R. and Parker, S.L. (2018) The Expansion Segments of 28S Ribosomal RNA Extensively Match Human Messenger RNAs. *Front Genet*, **9**, 66.
541. Rocchi, L., Pacilli, A., Sethi, R., Penzo, M., Schneider, R.J., Treré, D., Brigotti, M. and Montanaro, L. (2013) Dyskerin depletion increases VEGF mRNA internal ribosome entry site-mediated translation. *Nucleic Acids Res*, **41**, 8308-8318.
542. Montanaro, L., Calienni, M., Bertoni, S., Rocchi, L., Sansone, P., Storci, G., Santini, D., Ceccarelli, C., Taffurelli, M., Carnicelli, D. *et al.* (2010) Novel dyskerin-mediated mechanism of p53 inactivation through defective mRNA translation. *Cancer Res*, **70**, 4767-4777.
543. Penzo, M., Rocchi, L., Brugiere, S., Carnicelli, D., Onofrillo, C., Couté, Y., Brigotti, M. and Montanaro, L. (2015) Human ribosomes from cells with reduced dyskerin levels are intrinsically altered in translation. *Faseb j*, **29**, 3472-3482.
544. Carlile, T.M., Rojas-Duran, M.F., Zinshteyn, B., Shin, H., Bartoli, K.M. and Gilbert, W.V. (2014) Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature*, **515**, 143-146.
545. Wu, G., Xiao, M., Yang, C. and Yu, Y.T. (2011) U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *Embo j*, **30**, 79-89.
546. Cocucci, S.M. and Sussman, M. (1970) RNA in cytoplasmic and nuclear fractions of cellular slime mold amebas. *J Cell Biol*, **45**, 399-407.
547. Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S. *et al.* (2002) The genome sequence of Schizosaccharomyces pombe. *Nature*, **415**, 871-880.
548. Kupfer, D.M., Drabenstot, S.D., Buchanan, K.L., Lai, H., Zhu, H., Dyer, D.W., Roe, B.A. and Murphy, J.W. (2004) Introns and splicing elements of five diverse fungi. *Eukaryot Cell*, **3**, 1088-1100.
549. Fair, B.J. and Pleiss, J.A. (2017) The power of fission: yeast as a tool for understanding complex splicing. *Current genetics*, **63**, 375-380.
550. Neuvéglise, C., Marck, C. and Gaillardin, C. (2011) The intronome of budding yeasts. *C R Biol*, **334**, 662-670.
551. Swinburne, I.A. and Silver, P.A. (2008) Intron delays and transcriptional timing during development. *Dev Cell*, **14**, 324-330.
552. FlyBase. ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/current/fasta/dmel-all-intron-r6.16.fasta.gz.
553. Malko, D.B., Makeev, V.J., Mironov, A.A. and Gelfand, M.S. (2006) Evolution of exon-intron structure and alternative splicing in fruit flies and malarial mosquito genomes. *Genome Res*, **16**, 505-509.
554. Grzybowska, E.A. (2012) Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. *Biochem Biophys Res Commun*, **424**, 1-6.
555. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.
556. Sakharkar, M.K., Chow, V.T. and Kangueane, P. (2004) Distributions of exons and introns in the human genome. *In Silico Biol*, **4**, 387-393.
557. Yan, H., Dai, X., Feng, K., Ma, Q. and Yin, T. (2016) IGDD: a database of intronless genes in dicots. *BMC Bioinformatics*, **17**, 289.

558. The Arabidopsis Genome, I. (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, **408**, 796-815.
559. Atambayeva, S.A., Khailenko, V.A. and Ivashchenko, A.T. (2008) Intron and exon length variation in Arabidopsis, rice, nematode, and human. *Molecular Biology*, **42**, 312.
560. Tatosyan, K.A., Stasenko, D.V., Koval, A.P., Gogolevskaya, I.K. and Kramerov, D.A. (2020) TATA-Like Boxes in RNA Polymerase III Promoters: Requirements for Nucleotide Sequences. *Int J Mol Sci*, **21**, 3706.
561. Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M. and Pagano, A. (2007) The expanding RNA polymerase III transcriptome. *Trends in Genetics*, **23**, 614-622.
562. Halbig, K.M., Lekven, A.C. and Kunkel, G.R. (2008) Zebrafish U6 small nuclear RNA gene promoters contain a SPH element in an unusual location. *Gene*, **421**, 89-94.
563. Chen, C.L., Perasso, R., Qu, L.H. and Amar, L. (2007) Exploration of pairing constraints identifies a 9 base-pair core within box C/D snoRNA-rRNA duplexes. *J Mol Biol*, **369**, 771-783.
564. Bachellerie, J.-P., Michot, B., Nicoloso, M., Balakin, A., Ni, J. and Fournier, M.J. (1995) Antisense snoRNAs: a family of nucleolar RNAs with long complementarities to rRNA. *Trends in Biochemical Sciences*, **20**, 261-264.
565. Dunbar, D.A., Chen, A.A., Wormsley, S. and Baserga, S.J. (2000) The genes for small nucleolar RNAs in Trypanosoma brucei are organized in clusters and are transcribed as a polycistronic RNA. *Nucleic Acids Res*, **28**, 2855-2861.
566. Darzacq, X. and Kiss, T. (2000) Processing of Intron-Encoded Box C/D Small Nucleolar RNAs Lacking a 5′,3′-Terminal Stem Structure. *Molecular and Cellular Biology*, **20**, 4522-4531.
567. Watkins, N.J., Leverette, R.D., Xia, L., Andrews, M.T. and Maxwell, E.S. (1996) Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D. *Rna*, **2**, 118-133.
568. Watkins, N.J. and Bohnsack, M.T. (2012) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip Rev RNA*, **3**, 397-414.
569. Sharma, S., Yang, J., van Nues, R., Watzinger, P., Kötter, P., Lafontaine, D.L.J., Granneman, S. and Entian, K.D. (2017) Specialized box C/D snoRNPs act as antisense guides to target RNA base acetylation. *PLoS Genet*, **13**, e1006804.
570. Martin, R., Hackert, P., Ruprecht, M., Simm, S., Brüning, L., Mirus, O., Sloan, K.E., Kudla, G., Schleiff, E. and Bohnsack, M.T. (2014) A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase. *Rna*, **20**, 1173-1182.
571. Khatter, H., Myasnikov, A.G., Natchiar, S.K. and Klaholz, B.P. (2015) Structure of the human 80S ribosome. *Nature*, **520**, 640-645.
572. Agrisani, A., Tafer, H., Stadler, P.F. and Furia, M. (2015) Unusual Novel SnoRNA-Like RNAs in Drosophila melanogaster. *Noncoding RNA*, **1**, 139-150.
573. Kiss, T. (2004) Biogenesis of small nuclear RNPs. *J Cell Sci*, **117**, 5949-5951.
574. Guy, M.P., Podyma, B.M., Preston, M.A., Shaheen, H.H., Krivos, K.L., Limbach, P.A., Hopper, A.K. and Phizicky, E.M. (2012) Yeast Trm7 interacts with distinct proteins for critical modifications of the tRNAPhe anticodon loop. *RNA*, **18**, 1921-1933.
575. Vitali, P. and Kiss, T. (2019) Cooperative 2'-O-methylation of the wobble cytidine of human elongator tRNA(Met)(CAT) by a nucleolar and a Cajal body-specific box C/D RNP. *Genes Dev*, **33**, 741-746.
576. Ito, S., Horikawa, S., Suzuki, T., Kawauchi, H., Tanaka, Y., Suzuki, T. and Suzuki, T. (2014) Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). *The Journal of biological chemistry*, **289**, 35724-35730.
577. Shi, J., Huang, C., Huang, S. and Yao, C. (2018) snoRNAs associate with mRNA 3' processing complex: New wine in old bottles. *RNA biology*, **15**, 194-197.
578. Huang, C., Shi, J., Guo, Y., Huang, W., Huang, S., Ming, S., Wu, X., Zhang, R., Ding, J., Zhao, W. *et al.* (2017) A snoRNA modulates mRNA 3' end processing and regulates the expression of a subset of mRNAs. *Nucleic Acids Res*, **45**, 8647-8660.
579. Taft, R.J., Glazov, E.A., Lassmann, T., Hayashizaki, Y., Carninci, P. and Mattick, J.S. (2009) Small RNAs derived from snoRNAs. *RNA (New York, N.Y.)*, **15**, 1233-1240.
580. Scott, M.S. and Ono, M. (2011) From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie*, **93**, 1987-1992.
581. Bratkovic, T., Bozic, J. and Rogelj, B. (2020) Functional diversity of small nucleolar RNAs. *Nucleic Acids Res*, **48**, 1627-1651.

Appendix

Table S1. Oligonucleotides used in the detection of giant viruses.

Oligonucleotide	Sequence (5´ -> 3´)
R656 for	TTATTGGTCCCAATGCTACTC
R656 rev	TAATTACCATACGCAATTCCTG
MCP for	ATCGTGCTCATGCAACAAAA
MCP rev	TGCGAATCCATGCAGTTTTA
GpdA for	GGTTGTCCCAATTGGTATTAATGG
GpdA rev	CCGTGGGTTGAATCATATTTGAAC

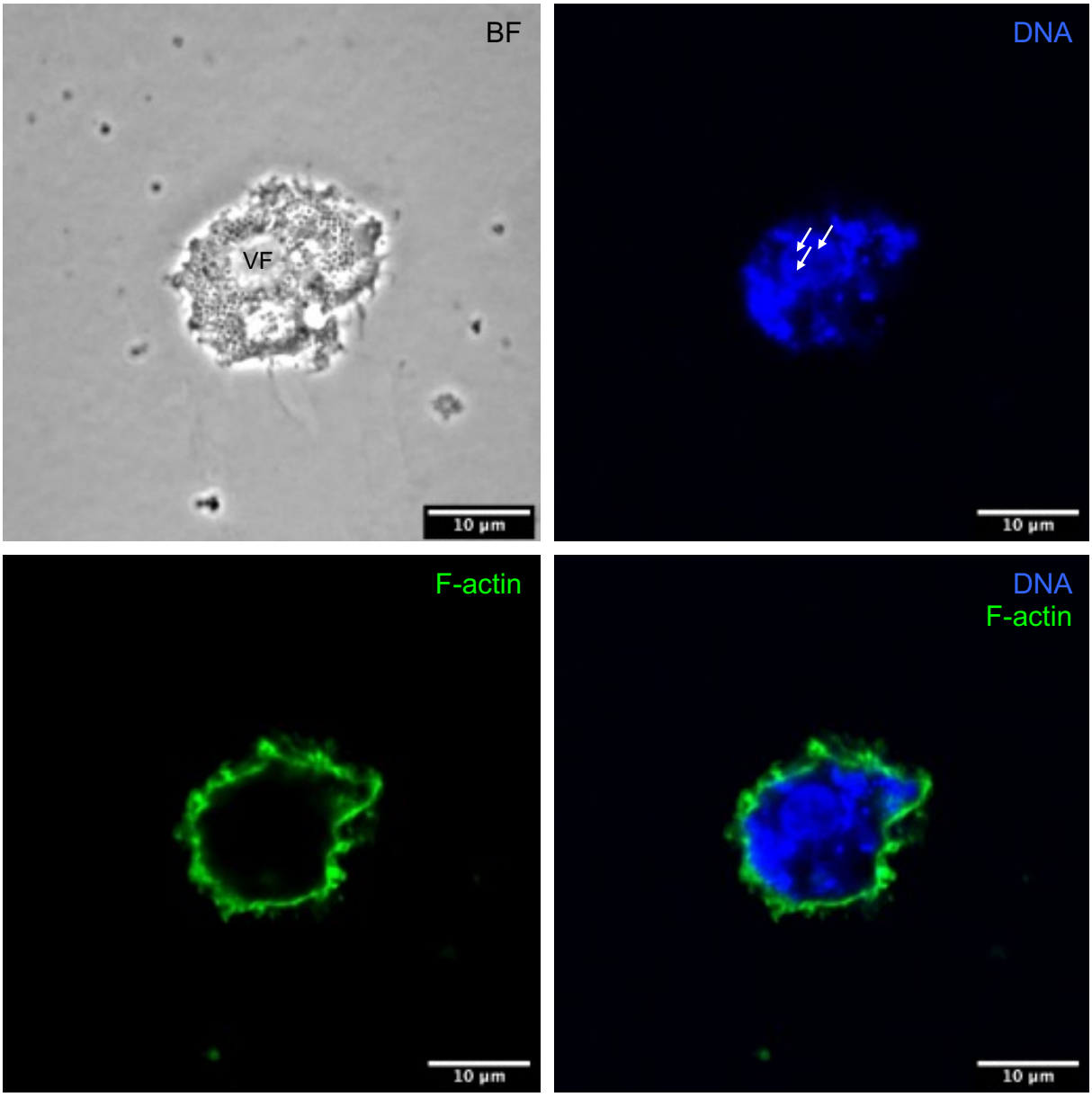


Figure S1. Fluorescence microscopy of APMV-infected *A. polyphaga* cells at 24 hpi. The cells were infected at MOI 1 and fixed at 24 hpi using 4% PFA. F-actin in the cell cortex was stained using phalloidin-488 (green) and DNA was stained using DRAQ5™ (blue). The virus factory was labelled in the brightfield (BF) image on the upper left. The white arrows point to novel capsids inside the virus factory.

Table S2. Oligonucleotides used in 'Chemical modifications in the rRNAs of *D. discoideum*.

Oligonucleotide	Sequence (5´ -> 3´)
CD1	AACTTCTTAGTTTTGGTCAAC
CD2	AGTTTTTGGCTTAATTAAAAATTAAA
CD3	AGTTTTTGGCTTAATAAAATATTAAA
CD4	AGTAACTATGAATATAGAATCAC
CD5	AGGAAAATTTAGATAACGCAAA
CD6	GTTAAAGACCTTACCACAG
CD7	ATTTGTGCGAACACGGA
CD8	AGACAATAAATGATCGATCAAA
CD9	GTCAGAAGCAAAACTGG
CD10	TTAAATGGAAAATCGTTATAATCA
CD11a/b	AGACAATAAAAATAAGGAACAAA
CD12	TCAATTTTCGTCATAGGTTATA
CD13	AAGCAAAACTGGCGTGT
CD14	GGCTAAAGATCATCAACAG
CD15	AAGACTGTCGTGAGAAATC
CD16	AAATTTCCATTCAGCATGATT
CD17	TTGTAATTTTTAAGAATCGTCTAT
CD18	ACGTGTTAAAAAAGATGTCC
CD19	AGATCTTGATGTAAATTGGAAA
CD20	ATTCTATAAGTGTAAAATTAAATGAT
CD21	TGTTTTAAAAAAACAGCCAAAG
CD22	GAAAGAATAATGCATAGTCTC
CD23	AAGACTGTCGTGAGAAATC
CD25	TATTAAAAAGCTCGTTGTTTTTTT
CD26	ATGTTATAGAAGAAGTAATGTTTT
CD27	AGTCAAATGTATTATGTAGAATT
CD28	ACAAAAATTTTGTACACGTAATG
CD29	ATGGTGTTGCATGGTAAATA
CD30	TGCTTGACTACTAGATAGG
CD31	AGTTATTATAATTTGAATCAGCAA
CD32	GTCATCTATCATAAGTTTCAGC
CD33	TCGATTCGGTATCAATGAAGCT
CD34	ATTTCCGCATGACGATTCTCA
CD35	TCAACATGGCTTAATAGAATG
CD36	AGGGAGGCCGTTCTTATCAAT
CD37	GACAGTCATGACAAAGGTGCT
OR6	GATCGATCCCTCCCTGCCAGT
17S seq for	ACACAATTGGAGGGCAAGTC
17S seq rev	AATTTCACCTCTCGCCCACCs
NRO_ETS for	ATTTAGGTGACACTATAGGATATCACAAGAAGAGTGAGCAAGCAGATGC
NRO_ETS rev	GAAATTAATACGACTCACTATAGGGAGCATGTAGATAACACGTATCATGATATTAATCAC
NRO_GpdA for	ATTTAGGTGACACTATAGGATATCGGTTGTCCCAATTGGTATTAATGG
NRO_GpdA rev	GAAATTAATACGACTCACTATAGGGCCGTGGGTTGAATCATATTTGAAC

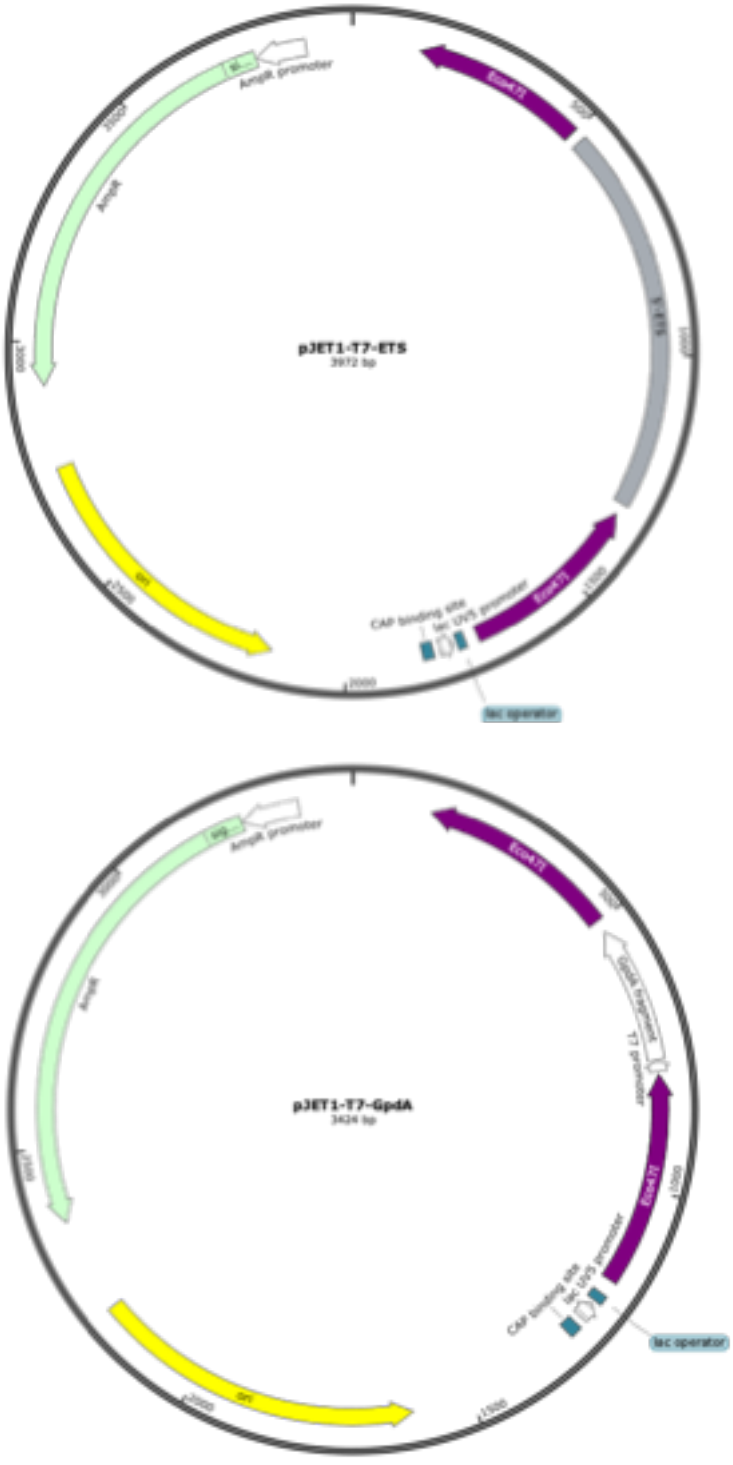


Figure S2. Plasmid maps for the vectors generated for *in vitro* transcription.

Table S3. Genomic location of C/D box snoRNAs with predicted rRNA methylation sites for CD RNAs[a].

RNA	Chrom.[b]	Start nt	End nt	Strand	Length [nt]	GC%	Predicted methylation site(s)
CD1	3	4409414	4409479	+	66	36.36	26S-Gm2132; 17S-Um1456; 17S-Gm1506
CD2	5	2599916	2599999	-	84	34.52	26S-Gm2661
CD3	1	3858016	3858099	-	84	33.33	26S-Gm2661
CD4	3	1341617	1341723	+	107	32.71	26S-Am2522
CD5	5	2585362	2585443	-	82	31.71	26S-Gm2984
CD6	2	7681951	7682036	-	86	34.88	26S-Gm3148
CD7	5	2249625	2249721	-	97	44.33	26S-Am711; 17S-Cm991
CD8	4	4444847	4444936	-	90	38.89	17S-Am432; 17S-Am466
CD9	3	4409131	4409211	+	81	40.74	26S-Am1370; 26S-Am1463
CD10	2	5045624	5045704	+	81	24.69	17S-Am1133
CD11a	2	7539516	7539591	-	76	38.16	26S-Am3279
CD11b	2	7539236	7539311	-	76	36.84	26S-Am3279
CD12	5	1152274	1152354	-	81	28.40	26S-Am841; 26S-Um2580
CD13	5	2600248	2600331	-	84	44.05	26S-Am1370; 26S-Am1463; 26S-Gm2952
CD14	3	3885476	3885562	+	87	31.03	26S-Gm2554
CD15	4	59280	59392	+	113	36.28	26S-Cm3281
CD16	5	2586030	2586108	-	79	32.91	26S-Gm1686; 17S-1588
CD17	2	2854096	2854165	-	70	27.14	26S-Um2170
CD18	4	5364580	5364650	-	71	29.58	17S-Am28
CD19	4	1888032	1888103	+	73	32.88	26S-Cm1673; 26S-Cm1685; 17S-Am796
CD20	6	2022190	2022263	+	74	27.03	17S-Um1255
CD21	5	3090313	3090388	+	77	31.17	17S-Um571
CD22	5	3090012	3090090	+	79	32.91	26S-Um2687
CD23	5	2585693	2585770	-	78	33.33	26S-Cm3292
CD24	3	2159499	2159574	+	76	34.21	26S-Am844
CD25	2	5722386	5722463	-	78	28.21	26S-Um3254; 17S-Am612
CD26	4	5364890	5364967	-	78	28.21	26S-Cm2603
CD27	1	4758315	4758393	+	79	22.78	26S-Am2159
CD28	4	5364289	5364369	-	81	35.80	17S-Cm1715
CD29	5	3859231	3859312	+	82	31.71	17S-Um1264
CD30	3	2037101	2037191	-	91	35.16	26S-Um2683
CD31	3	5036106	5036180	-	75	18.67	26S-Um2164
CD32	2	3177912	3177988	+	71	28.17	26S-Um2170
CD33	5	479167	479234	+	68	29.41	26S-Am1689
CD34	5	4272206	4272283	-	78	21.79	26S-Am2592
CD35	5	2434788	2434879	+	92	32.61	17S-Cm38
CD36	2	5722103	5722190	-	91	32.97	26S-Gm3124
CD37	1	3858279	3858361	-	83	39.76	17S-Gm1266
CD38	4	474437	474527	+	91	25.27	26S-Cm3212
OR1	2	4524799	4524900	+	102	37.25	-
OR2	4	4445171	4445247	-	77	36.36	-
OR3	4	5385832	5385918	-	87	33.33	-
OR4	4	5384392	5384479	-	88	31.82	-
OR5	2	6880357	6880444	+	88	38.64	-
OR6	4	1682178	1682262	-	85	30.59	-
OR7	4	1496745	1496814	-	70	37.14	-
OR8	3	5321597	5321673	-	77	22.08	-
OR9	4	1663005	1663104	-	100	35.00	-

[a]All box C/D snoRNAs are encoded intergenically, except for CD38, which is encoded in an intron of DDB_G0283293. Sequences with predicted methylation sites in rRNA are named CDx, and those without ORx for ORphan (x: natural number).

[b]Chromsomal accessions at dictybase.org are DDB0232428 (chr. 1), DDB0232429 (chr. 2), DDB0232430 (chr. 3), DDB0232431 (chr. 4), DDB0232432 (chr. 5), and DDB0232433 (chr. 6).

```
DdiU3-1         atgaccaaactcttaggatcatttctagagtatcgtctattaaaattattcatcaataat
DdiU3-2         atgaccaaactcttaggatcatttctagagtatcgtctattaaaattattcatcaataat
DdiU3-3         atgaccaaactcttaggatcatttctagagtatcgtctattaaaattattcatcaataat
DdiU3-4         atgaccaaactcttaggatcatttctagagtatcgtctattaaaattattcatcaataat
DdiU3-5         atgaccaaactcttaggatcatttctagagtatcgtctattaaaattattcatcaataat
DdiU3-6         atgaccaaactcttaggatcatttctagagtatcgtctattaaaattattcatcaataat
DdiU3-7         atgaccaaactcttaggatcatttctagagtatcgtctattaaaattattcatcaataat
                ************************************************************

DdiU3-1         ttttcctctttcacagctaggatgatgatatacactcactacacgaaagcgtgaaaccgt
DdiU3-2         ttttcctctttcacagctaggatgatgatatacactcactacacgaaagcgtgaaaccgt
DdiU3-3         ttttcctctttcacagctaggatgatgatatacactcactacacgaaagcgtgaaaccgt
DdiU3-4         ttttcctctttcacagctaggatgatgatacacactcactacacgaaagcgtgaaaccgt
DdiU3-5         ttttcctctttcacagctaggatgatgatacacactcactacacgaaagcgtgaaaccgt
DdiU3-6         ttttcctctttcacagctaggatgatgatacacactcactatacgaaagcgtgaaaccgt
DdiU3-7         ttttcctctttcacagctaggatgatgatacacactcactatacgaaagcgtgaaaccgt
                *******************************.**********.******************

DdiU3-1         tattatcgaatgattcatttatttttattaacattgatgaccgtctaattcagggatga
DdiU3-2         tattatcgaatgattcatttattctttattaacattgatgaccgtctaattcagggatga
DdiU3-3         tattatcgaatgattcatttattctttattaatattgatgaccgtctaattcagggatga
DdiU3-4         tattatcgaatgattcatttattctttattaatattgatgaccgtctaattcagggatga
DdiU3-5         tattatcgaatgattcatttattctttattaacattgatgaccgtctaattcagggatga
DdiU3-6         tattatcaaatgattcgtttatttgttattaacattgatgaccgtctaattcagggatga
DdiU3-7         tattatcaaatgattcatttatttgttattaacattgatgaccgtctaattcagggatga
                *******.********.******. *******.****************************

DdiU3-1         attggttgtgtggtgggattcgtactggc
DdiU3-2         attggttgtgtggtgggattcgtactggc
DdiU3-3         attggttgtgtggtgggattcgtactggc
DdiU3-4         attggttgtgtggtgggattcgtactggc
DdiU3-5         attggttgtgtggtgggattcgtactggc
DdiU3-6         attggttgtatggtgggattcgtactggc
DdiU3-7         attggttgtatggtgggattcgtactggc
                *********.******************
```

Figure S3. Sequence alignment of the U3 snoRNA genes of *D. discoideum*.

```
DpuU3-1         tagaccttactctttggatcatttctattagtatcgtctattaaaattattcatcaataa
DpuU3-2         cagaccttactctttggatcatttctattagtatcgtctattaaaattattcatcaataa
DpuU3-3         cagaccttactctttggatcatttctattagtatcgtctattaaaattattcatcaataa
DpuU3-4         cagaccttactctttggatcatttctattagtatcgtctattaaaattattcatcaataa
DpuU3-5         cagaccttactctttggatcatttctattagtatcgtctattaaaattattcatcaataa
DpuU3-6         cagaccttactctttggatcatttctattagtatcgtctattaaaattattcatcaataa
                .************************************************************

DpuU3-1         tttttctcaaaaaccccacggatgaagatgcacattcactttacataagcgggaagccgt
DpuU3-2         tttttctcaaaaaccccacggatgaagatgcacattcactttacataagcgggaagccga
DpuU3-3         tttttctcaaaaaccccacggatgaagatgcacattcactttacataagcgggaagccgt
DpuU3-4         tttttctcaaaaaccccacggatgaagatgcacattcactttacataagcgggaagccgt
DpuU3-5         tttttctcaaaaaccccacggatgaagatgcacattcactttacataagcgggaagccgt
DpuU3-6         tttttctcaaaaaccccacggatgaagatgcacattcactttacataagcgggaagccgt
                ************************************************************

DpuU3-1         gtaatagttttgattctatcatttctattatacattgatgatcgtctatctcagggatga
DpuU3-2         gtaatagttttgattctatcatttctattattcattgatgatcgtctatctcagggatga
DpuU3-3         gtaatagttttgattctatcatttctattatacattgatgatcgtctatctcagggatga
DpuU3-4         gtaatagttttgattctatcatttctattatacattgatgatcgtctatctcagggatga
DpuU3-5         gtaatagttttgattctatcatttctattatacattgatgatcgtctatctcagggatga
DpuU3-6         gtaatagttttgattctatcatttctattatacattgatgatcgtctatttcagggatga
                ****************************** ****************.**********

DpuU3-1         gatggttgtaaggtgaaattcttacgtgg
DpuU3-2         gatggttgtaaggtgaaattcttacgtgg
DpuU3-3         gatggttgtaaggtgaaattcttacgtgg
DpuU3-4         gatggttgtaaggtgaaattcttacgtgg
DpuU3-5         gatggttgtaaggtgaaattcttacgtgg
DpuU3-6         gatggttgtaaggtgaaattcttacgtgg
                ****************************
```

Figure S4. Sequence alignment of the U3 snoRNA genes of *D. purpureum*.

```
DlaU3-1         atgaccaaactcttaggattcatttcaagaggaacgtcaaagtatttttattctgaaaa
DlaU3-2         atgaccaaactcttaggattcatttcaagaggaacgtcaaagtatttttattctgaaaa
DlaU3-3         atgaccaaactcttaggattcatttcaagaggaacgtcaaagtatttttattctgaaaa
                ************************************************************

DlaU3-1         tactcaaaatctcagcaaggatgttgatccacactttcacagctttagcgggaagccgac
DlaU3-2         tactcaaaatctcagcaaggatgttgatccatactttcacagctttagcgggaagccgac
DlaU3-3         tactcaaaatctcagcaaggatgttgatccacactttcacagctttagcgggaagccgac
                ******************************.***************************

DlaU3-1         tttgattattgatttctatcaacagtcactgtcattgatgatcgtctttatagaaagtag
DlaU3-2         tttgattattgatttctatcaacagtcactgtcattgatgatcgtctttataggaagtag
DlaU3-3         tttgattattgatttctatcaacagtcactgtcattgatgatcgtctttataggaagtag
                ************************************************.******

DlaU3-1         gatttctatatagttgttgtgaaagtttctgacttgc
DlaU3-2         gatttctatttagttgttgtgaaagtttctgacttgc
DlaU3-3         gatttctatttagttgttgtgaaagtttctgacttgc
                ********* **************************
```

Figure S5. Sequence alignment of the U3 snoRNA genes of *D. lacteum*.

```
DfaU3-1         tagaccatactttaaggatctcttaacagtgtgcgacaaactatcaagtcgaaaacttga
DfaU3-2         tagaccatactttaaggatctcttaacagtgtgcgacaaactatcaagtcgaaaacttga
                ************************************************************

DfaU3-1         atcatatttaagtaaggatgatgaattgaaatcgcttgacctgagcgtgaaacggttagt
DfaU3-2         atcatatttaagtaaggatgatgaattgaaatcgcttgacctgagcgtgaaacggttagt
                ************************************************************

DfaU3-1         atgacagatgttcaatcttcatagctaacgatgatgatcgccaggaaatgagtagggctc
DfaU3-2         atgacagatgttcaatcttcatagctaacgatgatgatcgccaggaaatgagtagggctc
                ************************************************************

DfaU3-1         acttccaccggtcatgcgataggcaaacttac
DfaU3-2         acttccaccggtcatgcgata-----------
                *********************
```

Figure S6. Sequence alignment of the U3 snoRNA genes of *D. fasciculatum*.

Table S4. Genomic locations and upstream sequence elements of the U3 snoRNA in representatives of all major groups of Dictyostelia.

Species	Name	Sequence	Start	End	Strand	[nt]	Upstream motifs			
							DUSE	Motif	TATA-like	Motif
D. discoideum	U3-1	DDB0232430	3806685	3806893	+	209	✓	TCCCATAA	✓	TATAAATA
	U3-2	DDB0232431	2904295	2904503	+	209	✓	TCCCATAA	✓	TTAAAATA
	U3-3	DDB0232431	2927177	2927385	+	209	✓	TCCCATAA		
	U3-4	DDB0232431	2931648	2931856	+	209	✓	TCCCATAA	✓	TATAAATA
	U3-5	DDB0232431	2936427	2936635	+	209	✓	TCCCATAA	✓	TATAAATA
	U3-6	DDB0232433	2268599	2268807	+	209	✓	TCCCATAA		
	U3-7	DDB0232433	2270802	2271010	+	209	✓	TCCCATAA		
D. purpureum	U3-1	scaffold_688	3987	4195	-	209	✓	ACCCATAA	✓	TATATTCA
	U3-2	scaffold_140	37498	37706	-	209	✓	ACCCATAA		
	U3-3	scaffold_73	89977	90185	+	209	✓	ACCCATAA	✓	TATAATCA
	U3-4	scaffold_101	57185	57393	-	209	✓	ACCCATAA	✓	TATAAACA
	U3-5	scaffold_202	34893	35101	+	209	✓	TCCCATAA	✓	TATAAACA
	U3-6	scaffold_688	3598	3806	-	209	✓	ACCCATAA	✓	TATAAACA
D. lacteum	U3-1	LODT01000037.1	192486	192702	-	217	✓	TTCCATAA		
	U3-2	LODT01000037.1	423377	423592	+	216	✓	TTCCATAA		
	U3-3	LODT01000035.1	825082	825298	-	217	✓	AACCATAA		
P. pallidum	U3-1	GL290983.1	372856	373065	-	210	✓	TCCCATAA		
A. subglosum	U3-1	DF837589.1	656447	656660	-	214	✓	TACCATAA		
D. fasciculatum	U3-1	GL883010.1	1346903	1347114	-	212	✓	ACCCATAA		
	U3-2	GL883010.1	1325083	1325283	+	201	✓	ACCCATAA		