

**A Thesis Submitted for the Degree of PhD at the University of Warwick**

**Permanent WRAP URL:**

<http://wrap.warwick.ac.uk/165832>

**Copyright and reuse:**

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

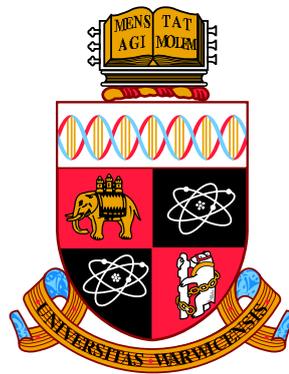
Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk)

# Carbon Nanodots as a Tool to Investigate Root-Microbe Interactions

Peter James Morrison



A thesis submitted for the degree of Doctor of Philosophy.

University of Warwick

School of Life Sciences

Synthetic Biology Centre for Doctoral Training

14th February 2020

# Contents

<b>Acknowledgements</b>	<b>1</b>
<b>Declaration</b>	<b>2</b>
<b>Abstract</b>	<b>3</b>
<b>Abbreviations</b>	<b>4</b>
<b>1. Introduction</b>	<b>22</b>
1.1. The root: a specialised transport and sensing system . . . . .	22
1.2. Cell type-specific analysis in plants . . . . .	24
1.3. Synthetic Biology: methods and mechanisms . . . . .	26
1.4. Automation, informatics, and the rise of data sets . . . . .	29
1.5. Applications of root biology . . . . .	30
1.5.1. Crops, Fertilisers, and the GMO debate . . . . .	30
<b>2. Nanoparticles as tools to deliver and trace molecules for biological investigation</b>	<b>32</b>
2.1. Carbon nanodots as an affordable and adaptable delivery system .	32
2.2. Methods . . . . .	38
2.2.1. Simple synthesis of carbon nanodot cores . . . . .	38
2.2.2. Adaption of the nanodot surface chemistry . . . . .	39
2.2.3. Attaching chemical payloads via amide linkers . . . . .	41
2.2.4. Testing the CuAAC reaction system for nanodot functionalisation . . . . .	42
2.3. Results . . . . .	43
2.3.1. Carbon nanodot core integrity confirmed by fluorescence analysis . . . . .	44
2.3.2. NMR analysis confirms previously published carbon nanodot properties. . . . .	46
2.3.3. NMR analysis confirms novel adaption of carbon dots with a long acid linker molecule . . . . .	48
2.3.4. NMR spectroscopy of flg-22 to functionalise carbon nanodots	51

2.3.5.	Proton NMR spectroscopy to analyse the bonding of flg-22 to adapted carbon nanodots . . . . .	52
2.3.6.	Two-dimensional spectroscopy confirms flg-22 and adapted carbon nanodots diffuse together . . . . .	55
2.3.7.	H-NMR analysis of size-exclusion chromatography fractions confirms significant crossover between nanodots and their oligopeptide payload post-reaction. . . . .	58
2.3.8.	Assessing the outcome of CuAAC reaction attempts using proton NMR . . . . .	61
2.4.	Conclusions . . . . .	62
<b>3.</b>	<b>Carbon Nanodot uptake in <i>Arabidopsis thaliana</i></b>	<b>64</b>
3.1.	Introduction . . . . .	64
3.2.	Methods . . . . .	67
3.2.1.	<i>Arabidopsis</i> growth and handling . . . . .	67
3.2.2.	Nanodot application . . . . .	69
3.2.3.	Confocal microscopy and image analysis . . . . .	69
3.2.4.	Root Protoplast generation protocol for <i>Arabidopsis</i> . . . . .	70
3.2.5.	Flow cytometry . . . . .	71
3.2.6.	RNA extraction and expression data . . . . .	71
3.3.	Results . . . . .	72
3.3.1.	Confocal microscopy of <i>Arabidopsis</i> seedlings indicates a route of systemic uptake . . . . .	72
3.3.2.	Flow Cytometry of protoplasts indicates possibility for cytosolic nanoparticles . . . . .	84
3.3.3.	Gene expression analysis suggests nanodots do not significantly alter transcriptional responses of flg-22 treated plant roots . . . . .	86
3.4.	Conclusion and Further Work . . . . .	89
<b>4.</b>	<b><i>Intelligate</i>: A software tool to tackle complex FCCS data</b>	<b>92</b>
4.1.	Introduction . . . . .	92
4.1.1.	Flow Cytometry and Cell Sorting: an introduction . . . . .	92
4.1.2.	Complexity challenges in flow cytometry and cell sorting experiments . . . . .	95
4.1.3.	Tackling repeatability and bias in FCCS results . . . . .	97
4.2.	Methods . . . . .	98
4.2.1.	Importing data to <i>R</i> ; pre-processing . . . . .	98
4.2.2.	Interpreting output; tuning parameters . . . . .	98

4.3. Results . . . . .	100
4.3.1. Using <i>Intelligate</i> to gate rare events in <i>Arabidopsis</i> protoplasts	100
4.3.2. Improving result robustness . . . . .	100
4.4. Conclusion and further work . . . . .	101
4.4.1. Post-hoc experiments . . . . .	101
4.4.2. Developing the <i>Intelligate</i> algorithm . . . . .	101
<b>5. Discussion</b>	<b>106</b>
5.1. Summary . . . . .	106
5.2. Further Work . . . . .	107
5.3. Conclusion . . . . .	108
<b>Appendix</b>	<b>109</b>
<b>A. <i>Intelligate</i> script</b>	<b>109</b>
A.1. Intelligate master script . . . . .	109
A.2. Data Import sub-script . . . . .	110
A.3. Primary clustering sub-script . . . . .	111
A.4. Covariance Heatmapping sub-script . . . . .	114
A.5. Plotting Sub-script . . . . .	114
<b>Bibliography</b>	<b>119</b>
<b>Afterword</b>	<b>130</b>

# Acknowledgements

This project is funded by EPSRC and BBSRC under the Synthetic Biology CDT grant (ref: EP/L016494/1), and used resources from WISB, a BBSRC/EPSRC Synthetic Biology Research Centre (grant ref: BB/M017982/1). Both are funded under the UK Research Councils' Synthetic Biology for Growth programme.

## **Principal Academic Supervisor**

Dr. Miriam L. Gifford

## **Secondary Academic Supervisors**

Prof. M. Carmen Galan

Dr. Heather Whitney

## **Referees**

Dr Alex M Eccles-Jones

Dr Giovanni Sena

# Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

Graphics from figures 1.1, 1.3, 1.4, and ?? - these are for reference only, and their production should not be considered for examination; sources for published material, where available, are given in the corresponding figure captions.

# Abstract

This project takes a synthetic biology approach into the data-driven sector of plant molecular biology. It is our aim to create a platform technology, which can be used to deliver a variety of synthetic effectors to whole plants, making subsequent use of fluorescent sorting methods to examine the effects on a single-cell scale: without using transgenic methods. My objective is to provide proof-of-concept for this technology, which will be highly useful in our group's progress in engineering nitrogen-fixing nodulation into non-legume plants.

# Abbreviations

AFM	Atomic force microscopy
API	Application Programming Interface
CDI	1,1'-Carbonyldiimidazole
crosstalk	Term (originally from radio communications) referring to interactions with systems other than the intended target.
CuAAC	Copper(I)-mediated azide-alkyne cycloaddition reaction. Established synthetic chemistry method adapted for biological purposes. Also known as "Click Chemistry".
DOSY	diffusion-ordered spectroscopy
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
ETI	Effector-Triggered Immunity
FC	Flow Cytometry
FCCS	Flow Cytometry and Cell Sorting / Flow Cytometric Cell Sorting
GO term	Gene Ontology term (see Ashburner et al., 2000)
GPU	Graphical Processing Unit ("Graphics card")
meta-data	"Data about data" e.g. data size, mean, variance, inter-quartile range.
MWCO	Molecular Weight Cut-Off
n+1 rule	Principle when interpreting proton NMR spectra: in aliphatic structures, number of NMR peaks corresponds to n+1, where n is the number of hydrogen atoms bonded to the neighbouring carbon atom.

NaN	"Not a Number" - an error value returned when a mathematical error has been made in computation e.g. attempting to divide by zero.
NaN	"Not a Number" - an error value returned when a mathematical error has been made in computation e.g. attempting to divide by zero.
PAMP	Pathogen associated molecular pattern
ppm	parts-per-million
PTI	Pattern-Triggered Immunity (a.k.a. non-host immunity)
THF	Tetrahydrofuran
TTDDA	4,7,10-trioxa-1,13-tridecanediamine
zeta-potential	Potential difference (typically measured in millivolts) between the dispersion medium of a particle, and the stationary layer of that medium attached to the dispersed particle. A measure on stability of the colloidal dispersion: larger magnitude is more stable.

# List of Figures

1.1. <b>Illustration of plant-pathogen interaction systems.</b> Pathogen associated molecular patterns (PAMPs) are detected in the extracellular spaces by pattern recognition receptors (PRR). This triggers a cascade of transcriptional changes known as pattern-triggered immunity (PTI). The bacterial flagellin component flg-22 is one such PAMP: it is sensed by the Flagellin-Sensitive 2 (FLS2) protein, which interacts with the BAK1 receptor-like kinase after binding flg-22 in order to effect intracellular activity. Some pathogens have mechanisms such as the type-3 secretion system (T3SS) pilus, which can inject pathogenic effector proteins directly into the intracellular space, and thereby suppress PTI: plants mostly rely on nucleotide-binding leucine-rich repeat receptor kinases (NB-LRR) to recognise these and initiate a response known as effector-triggered immunity (ETI). Figure created using BioRender.com (2020), based on Dodds and Rathjen (2010). . . .	23
1.2. <b>Photograph of a week-old <i>Arabidopsis thaliana</i> seedling.</b> Scale bar 10 mm . . . . .	25
1.3. <b>Fluorescence excitation-emission plot for carbon nanodot cores.</b> Colour scale represents intensity of emitted light at the corresponding wavelengths of excitation (x-axis) and emission (y-axis). Figure reproduced with author's permission from Swift et al. (2018). . . . .	28
1.4. <b>Illustration of root-microbe interactions in <i>Medicago truncatula</i>.</b> Adapted from G.Muscatt et al (personal communication).	30

2.1.	<b>Illustration (not to scale) depicting structure of carbon nanodots functionalised with a layer of (a) short acid linker and amino-glucose (b) long linker and flg-22 peptide, and (c) short alkyne linker and Fmoc-<math>\beta</math>-azido-Alanine-OH, respectively.</b> Here it can be seen that, through the use of a variety of linker chemicals (yellow), a variety of payloads (purple) may be loaded onto the particles (blue) with a greater or lesser degree of surface passivation; also, that the nature of the linker and payload in terms of polarity and molecular weight may affect overall physical properties of the final particle: especially since surface functionalisations may not always result in a homogenous corona surrounding the nanodot core, as seen in (Swift et al., 2018). . . .	34
2.2.	<b>Chemical structure diagram of the flg-22 peptide in zwitterionic form.</b> The amino acid sequence (from left to right) is QRLSTGSRINSAKDDAAGLQIA. . . . .	36
2.3.	<b>Reaction scheme for one-step microwave synthesis of carbon nanodots.</b> TTDDA is polymerised to form carbon nanodots, in the presence of glucosamine and microwave energy. Nitrogen-bearing groups are coloured blue; oxygen-bearing groups are in red. After Hill et al. (2016). . . . .	38
2.4.	<b>Reaction scheme for addition of a short acid linker to carbon nanodots.</b> The CDI enables a ring opening of succinic anhydride, allowing it to react with the primary amine on the carbon nanodot, resulting in a aliphatic carboxylic acid group coupled by an amide bond to the nanodot. . . . .	39
2.5.	<b>Reaction scheme for addition of a long acid linker to carbon nanodots.</b> The carboxylic acid group at the carbon nanodot surface is activated by CDI before the linker chemical is added, so as to minimise CDI-activated polymerisation of the linker. The target reaction is an amide coupling between the carboxylic acid group at the surface of the nanoparticle and the primary amine of the linker chemical, resulting in a nanoparticle with a long aliphatic linker, terminating in a carboxylic acid functional group. . . . .	40
2.6.	<b>Reaction scheme for addition of a short alkyne linker to carbon nanodots.</b> A condensation reaction occurs between the primary amine on the carbon nanodot and the hydroxyl group on the pentynoic acid, forming an amide bond between the two reactants, and resulting in a carbon nanodot presenting an alkyne functional group at the surface. . . . .	41

2.7.	<b>Reaction schemes for functionalizing carbon nanodots with the amino acid form of the flg-22 oligopeptide.</b> (a) by direct amide bond formation to acid-functionalized carbon dots (as seen above); and (b) using a longer "linker molecule" i.e. amide bond formation between flg-22 and a carbon nanodot functionalised with a TTDDA-based linker. . . . .	41
2.8.	<b>Reaction schemes for proposed copper(I)-mediated azide-alkyne cycloaddition (CuAAC) reactions with alkyne-adapted nanodots.</b> Each reaction involves the alkyne group on the nanodots reacting with the azide on the chemical payload to form a triazole compound coupling the two reactants. This reaction is catalysed by a copper (I) species which is relatively unstable compared to copper (II) in aqueous solutions: hence, the conditions required are quite specific. . . . .	42
2.9.	<b>(a) Optical absorbance curves for nanodots functionalised with flg-22 via short acid linker (black line) and flg-22 via long linker (red line) and fluorescence spectroscopy data for (b) nanodots with short acid linker and flg-22 and (c) nanodots with long acid linker and flg-22.</b> For this type of particle, changes in fluorescence depending on surface functionalisation are consistent with other published results (Swift et al., 2018). Note the response to 400 nm excitation, which is consistent between the two types of functionalised particle. . . . .	45
2.10.	<b>Analytical identification of carbon nanodot cores.</b> (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.3 with characteristic regions of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. The large peak near 4.7 ppm corresponds to the solvent (deuterium oxide) used to prepare the sample for NMR. Here are visible some peaks around 3.5 ppm from protons in the nanodot core, whose putative composition is a polymer comprised predominantly of SP3 and SP2 carbon, with a moderate degree of cross-linking occurring between the original SP3-hybrid constituents. Additionally, we note a pentet peak near 1.75 ppm, which corresponds to the ethylamine groups in the expected structure. . . . .	47

<p>2.11. <b>Analytical confirmation of the presence of the short acid linker.</b> (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.4, with characteristic regions of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. There are two pentat peaks annotated in (b) with an asterisk (*). Either (or both) of these could correspond to the amine group seen in the previous figure, suggesting that some free amine groups remain after the reaction: the splitting suggests some change in the chemical environment, for example one could infer that some (and not all) of the amines may be close enough to the added linker for spin-spin coupling. . . . .</p>	48
<p>2.12. <b>Analytical confirmation of the formation of the product in figure 2.5.</b> (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.5 with characteristic regions of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. Of note here is the shaded region near 2.35 ppm, where two formerly distinct triplet peaks have merged compared to the previous figure. As a result of the reaction, the environment of the corresponding protons has become more symmetrical: this causes their resonance peaks to become more similar, so much so that they become superimposed at this resolution. A slight skew remains visible, though it is possible for the required post-processing (phase correction) to produce a similar effect. The peak at 3.3 ppm correlates well with the four superimposed singlets from the CH<sub>2</sub> groups highlighted in the centre of the linker chain, and as expected integrates close to double the size of the peak at 2.35 ppm. . . . .</p>	50
<p>2.13. <b>NMR identification of the flg-22 synthetic oligopeptide.</b> (a) Structural diagram and (b-c) Proton NMR spectrum of the flg-22 oligopeptide, with characteristic region of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. NMR data was obtained with a 200MHz Bruker spectrometer, and analyzed using Mestrelabs MNova software. Flg-22 structure diagram was drawn using the PepDraw online tool. . . . .</p>	51

- 2.14. **Analytical confirmation of a reaction between short-linker nanodots and flg-22.** (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.7a, with characteristic regions of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. Between the two doublet peaks identified as coming from flg-22, we note the formation of an additional doublet, denoted with an asterisk (\*). It is difficult to assign this peak with confidence, since there are now a very large number of protons in the molecule, but its presence suggests a change near one of the methyl groups previously highlighted in the flg-22 expanded structure. . . . . 53
- 2.15. **Confirmation of one or more structural changes in the reaction between long-linker nanodots and flg-22.** (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.7b (carbon nanodots functionalised with flg-22 via long linker). Shaded areas indicate the depicted structural elements, and regions of the NMR spectrum where one or more corresponding peaks are expected. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. For a structure as depicted, we would expect to see a single peak in the green shaded area (around 3.3 ppm) as seen in prior spectra. The absence of this peak suggests that either the structural context of the corresponding protons has changed compared to the linker-only particle, or they are no longer present in the sample, having been displaced by other constituents. It is possible that flg-22 has hydrogen-bonded to the linker in such a manner as to shift of these protons downfield (to the left), causing them to overlap with other signals; alternatively, the reaction may have caused individual linker molecules on the surface of the nanoparticle to react with one another, rather than with the flg-22. An additional doublet peak, similar to that seen in the previous figure, is denoted with an asterisk (\*): overall, the proton spectrum is strikingly similar to that seen in the short-linker reaction, emphasizing the motivation towards correlative spectroscopy for greater resolving power. . . . . 54

- 2.16. **Two-dimensional analysis of the product shown in figure 2.7a.** (a) Expected structure, (b) annotated proton spectrum and (c) corresponding DOSY spectrum of the product of the reaction in figure 2.7a. X-axis shows frequency shift in parts per million; Y-axis shows diffusion constant in square centimetres per second. Z-axis (intensity) is represented by contours on a yellow-red scale, with darker colours representing larger values. Shaded areas indicate correspondence between the depicted structural elements and regions of the proton NMR spectrum. Dashed line is a horizontal guide for the eye: peaks on the same horizontal level in the lower contour plot have the same diffusion constant. Some peaks from the carbon nanodot are at similar horizontal levels to those from flg-22, indicating the latter may be covalently bonded to some quantity of the nanodot. Particularly, the lowest peak in the shaded area near 3.6 ppm is not seen in the spectrum of flg-22 alone, yet its estimated diffusion coefficient is very close to that of the flg-22 peaks. . . . . 56
- 2.17. **Diffusion analysis of the product of reaction 2.7b.** (a) Expected structure, (b) annotated proton spectrum and (c) corresponding DOSY spectrum of the product of the reaction in figure 2.7b. Shaded areas indicate correspondence between the depicted structural elements and regions of the proton NMR spectrum. Dashed line is a horizontal guide for the eye, at the same diffusion constant as the previous figure. Here the peaks identified as characteristic of flg-22 have less crossover on the horizontal level with those identifying carbon nanodot structures, when compared to the previous figure. This suggests that the two reactants are less likely to be bonded covalently as shown in the above structure. 57
- 2.18. **Fractionation by size exclusion chromatography confirms co-diffusion of short-acid nanodots with flg-22.** (a) Expected structure and (b) annotated stacked proton spectra of 0.5 ml size exclusion chromatography fractions of the product of the reaction in fig. 2.7a: each trace is a separate fraction of the reaction product. Shaded areas indicate correspondence between the depicted structural elements and regions of the proton NMR spectrum. The peak near 7.3 ppm corresponds to the activator chemical CDI, which it is desirable to remove from the system: emphasizing the requirement for size-exclusion purification post-reaction. . . . . 59

- 2.19. **Size-exclusion chromatography indicates diffusive similarity between long-linker nanodots and flg-22 after reaction.** (a) Expected structure and (b) annotated stacked proton spectra of size exclusion chromatography fractions of the product of the reaction in fig. 2.7b Shaded areas indicate correspondence between the depicted structural elements and regions of the proton NMR spectrum. Here the crossover between the flg-22 peptide, carbon nanodots, linker, and the activator chemical CDI can be seen. As indicated previously, a missing (possibly shifted) peak at 3.3 ppm calls into question that the structure is in fact as illustrated. . . . 60
- 2.20. **Analysis confirms the failure of our experimental conditions to effect azido-alkyne addition.** (a) Expected structures and annotated proton NMR spectra of the reactants in fig. 2.8a (b) before and (c) after attempting the CuAAC reaction. The fact that there has been no change in the spectrum, particularly around the distinctive peak at 2.2 ppm corresponding to the distal end of the alkyne, suggests that no reaction has taken place. This was attempted more than once, using buffers in three different pH levels, with similar results. Of additional note is the region around 3.6 ppm, which has become crowded with the addition of more organic chemical groups: this is one weakness of proton-only NMR when analysing larger molecules, and the original issue which led to the invention of two-dimensional correlative spectroscopy techniques such as diffusion-ordered spectroscopy (DOSY), seen above. Similarity of the proton spectra before and after reaction suggest that no reaction has taken place: most probably the (very specific) conditions required for this reaction were not met in this case, namely the formation of a Cu(I) ion complex in an aqueous environment. 62

- 3.1. **Illustration of uptake pathways in plant roots: apoplastic and symplastic.** Cell types are shown in colours (from right to left): epidermis in light grey, cortex in dark blue, endodermis in lilac, pericycle in yellow, and xylem and phloem in green. Although they are in practice two distinct tissues, the radial positions of xylem and phloem overlap: the outer files (dark green) are occupied by protoxylem, procambium, and the sieve elements and companion cells of the phloem; whereas the inner files (light green) are comprised of metaxylem and more procambium cells. White vertical arrows represent material transfer with other organs in the plant (e.g. leaves). In either pathway, a molecule must undergo two membrane crossings (black circles). In the symplastic route (lower black arrow) the molecule enters the cytoplasm at the epidermis, before being passed from cell to cell by protein pores called plasmodesmata (pink) through to the pericycle level, after which one further membrane crossing must take place in order to access the procambium and vascular cells. In the apoplastic route (upper black arrow) the molecule passes between cells, moving in the interstitial spaces, and may permeate through the walls of outer plant cells without crossing the membrane, up until the casparian strip (dark purple line): a waxy, impermeable strip which occupies the cell interstices of the endodermis. At this point, a molecule must cross the membrane of a endodermal cell in order to progress to the pericycle, and as with the symplastic route, at least one further membrane crossing is required to enter into vascular cells. Figure created using BioRender.com (2020); based on Zhou et al. (2013) . . . . . 66
- 3.2. **Photograph of cultures of *Arabidopsis* seedlings to be treated with carbon nanodots on liquid media (left) and solid media (right).** Liquid media can be changed at any time with relative ease, whereas solid media offers potential spatial stratification of conditions within a single root. Both of the pictured are nutrient-replete media (Murashige and Skoog formulation). 68

- 3.3. **Arabidopsis seedling growth protocols.** (a) Photograph of a two-week-old wild-type (Columbia-0 line) *Arabidopsis* seedling, grown vertically on solid media (MS agar). (b) (upper) Illustration and photograph of hydroponics apparatus used to culture *Arabidopsis* seedlings and (lower) illustration of “pulse-chase” experiment design used to infer nanodot uptake dynamics - white arrows indicate the passage of time, while black arrows indicate the hypothesised transport route of nanodots. . . . . 68
- 3.4. **Confocal micrograph of *Arabidopsis* root optical section treated with carbon nanodots.** Carbon nanodots (blue) were added to 7 day old plants in liquid MS growth media to a final concentration of 10 mg/ml for 72 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Fluorescence images at 420-450 nm (left) 580-650 nm (middle) and composite (right). Root axial optical section; scale bar is 38  $\mu\text{m}$ . . . . . 72
- 3.5. **Confocal micrograph of *Arabidopsis* leaf from seedling treated with acid-functionalised carbon nanodots.** Carbon nanodots (blue) were added to 7 day old plants in liquid MS growth media to a final concentration of 10 mg/ml for 72 hours prior to imaging. Filter parameters for propidium iodide (red, 580-650 nm) were used to illuminate chloroplasts and thus provide context, although no PI was added. Fluorescence images at 420-450 nm (lower left) 580-650 nm (upper left), transmitted light (lower-right) and composite (upper right). Scale bar 38  $\mu\text{m}$ . . . . . 73
- 3.6. **Confocal micrograph of *Arabidopsis* root optical section treated with (a) carbon nanodot cores and (b) control (DI water).** Carbon nanodots (blue) were added to 10 day old plants in liquid MS growth media to a final concentration of 10 mg/ml for 72 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Tangential optical section with fluorescence images at 420-450 nm (lower left) 580-650 nm (upper left) composite (upper right), and composite with transmitted light image (lower right). Scale bar 38  $\mu\text{m}$ . . . . . 75

- 3.7. **Confocal micrograph images of *Arabidopsis* roots** treated with (a-b) un-functionalised carbon nanodots (c-d) control treatment (DI water). Carbon nanodots (blue) were added to 10 day old plants in liquid MS growth media to a final concentration of 10 mg/ml for 72 hours prior to imaging. Images are sums of image ‘stacks’ from the same root at different optical planes: fluorescence at 420-450 nm (lower left), 580-650nm (upper left), composite (upper right), and composite with transmitted light image (lower right). (b) and (d) are data from (a) and (c), respectively, thresholded to exclude blue values less than 85 (arbitrary units) so as to suppress autofluorescence from lignified tissues in the stele. Bright areas in the cell interstices may suggest an apoplastic uptake route; bright regions in xylem (along the central axis) of this mature root may indicate that carbon nanodots are passively transported into aerial tissues of the plant. Scale bar is 38  $\mu\text{m}$ . . . . . 76
- 3.8. **Confocal micrograph of an *Arabidopsis* root tip, treated with un-functionalised nanodots.** Images are sums of image ‘stacks’ from the same root at different optical planes: fluorescence at 420-450 nm (lower left), 580-650nm (upper left), composite (upper right), and composite with transmitted light image (lower right). Note the bright regions in cell interstices, similar to 3.7, despite differences in cell development stage. Scale bar is 38  $\mu\text{m}$ . 77
- 3.9. **Fluorescence data from *Arabidopsis* root treated with non-functionalised carbon nanodots in liquid media.**(a)Wavelength-coloured confocal micrograph (axial optical section) from the meristem of an *Arabidopsis* root tip from a 7-day old seedling, treated with 10 mg/ml non-functionalised carbon nanodots in liquid MS media 48 hours prior to imaging. Image annotated with ROI indication crosshairs (A-C). Note the pale blue-coloured regions in the intercellular regions - this suggests the nanodots are present in the root apoplast. (b) Light emission plots corresponding to regions of interest in (a), obtained with the microscope’s array detector under 405 nm coherent light excitation. . . . . 78

- 3.10. **Confocal micrograph of *Arabidopsis* root optical section control treatment (water).** Treatments were added to 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence at 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection shown in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38  $\mu\text{m}$ . . . . . 79
- 3.11. **Confocal micrograph of *Arabidopsis* root (axial optical section) treated with acid-functionalised carbon nanodots.** Carbon nanodots (blue) suspended in DI water to concentration of 20 mg/ml were pipetted onto roots of 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection shown in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38  $\mu\text{m}$ . . . . . 80
- 3.12. **Confocal micrograph of *Arabidopsis* root (Axial optical section) treated with long-linker-functionalised carbon nanodots.** Carbon nanodots (blue) suspended in DI water to concentration of 20 mg/ml were added to 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38  $\mu\text{m}$ . . . . . 81

- 3.13. **Confocal micrograph of *Arabidopsis* root (axial optical section) treated with flg-22-functionalised carbon nanodots using short acid linker.** Carbon nanodots (blue) suspended in DI water to concentration of 20 mg/ml were added to 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection shown in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38  $\mu\text{m}$ . . . . 82
- 3.14. **Confocal micrograph of *Arabidopsis* root (axial optical section) treated with flg-22-functionalised carbon nanodots with long acid linker.** Carbon nanodots (blue) suspended in DI water to concentration of 20 mg/ml were added to 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection shown in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38  $\mu\text{m}$ . . . . 83
- 3.15. **Fluorescence cytometry data of carbon nanodot uptake in *Arabidopsis*.** (a) Scatter graph of log-fluorescence data collected from *Arabidopsis* root protoplasts illuminated with a 405nm laser, showing the carbon-dot fluorescence band (x-axis) against a yellow-green fluorescence band (y-axis) associated with off-target fluorescence (wide-band autofluorescence). The ratio between the two bands can be used to locate carbon nanodots (see Figure 3.9 on page 78); population of interest highlighted in red. (b) Bar graph showing the average proportion of cells identified in the highlighted population at each time point. These are data gathered from a flow cytometry experiment, - bars reflect the relative proportions of protoplasts considered to have a high concentration of carbon nanodots according to criteria defined using *Intelligate* software in Chapter 4. . . . . 85

3.16. <b>Fold-change expression data for Flg-22 key markers as identified by Beck et al. (2014).</b> Nine-day-old <i>Arabidopsis</i> seedlings were treated with flg-22 only (dark blue bars) carbon nanodots functionalised with flg-22 (mid-blue bars), and carbon nanodots with a short linker only (light blue bars). Yellow stars indicate significant differences ( $p \leq 0.05$ ) in expression between carbon nanodot-treated plants and the other treatments, according to a homoscedastic t-test. Error bars indicate plus/minus one mean standard error. . . . .	88
3.17. Fold-change expression data for key differentially expressed loci according to Stringlis et al. (2017). . . . .	89
4.1. <b>Illustration of practical workflow for collecting single-cell data from plant tissues.</b> Protoplast generation (1) involves the degradation of cell walls via the application of an enzyme cocktail - thus transforming the multicellular structure into a suspension of plant protoplasts. The protoplasts are subsequently sorted by their fluorescent and optical parameters via FCCS (2) in order to separate cells of interest from the bulk sample: these could be e.g. quantiles of GFP-expressing cells in a protein fusion experiment, or cells otherwise decorated with a selective fluorescent marker. The sorted protoplast samples can then be sent for further analysis via e.g. RNA-seq, allowing the collection of expression data. . . . .	93

- 4.2. **Annotated photograph of a FCCS machine (BD Influx).** Fluidic systems are represented in green; optical systems in orange; electrical/electronic components in cyan/blue respectively. Cells are loaded from the sample tube (A) into the nozzle (B) by positive pneumatic pressure, where the stream is hydrodynamically focussed with a secondary 'sheath' fluid, and subjected to ultrasonic stimulation by a piezoelectric element in the nozzle. The stream emerging from the nozzle is interrogated by a series of lasers (C) aligned by a series of mirrors (D) so as to meet perpendicularly at the interrogation point, marked with a yellow star. Hereafter, the stream is broken into drops, entrained by the periodic pressure field from the ultrasonic emitter. Transmitted and scattered light is filtered, detected (E), and electronically transmitted to a computer-connected realtime controller (F) which dynamically controls electric charge on the sorting plates (G) in order to deflect drops to a varying degree (as specified by the operator) in order that they align with various collection vessels placed at known locations in the collection area (H). . . . . 94
- 4.4. **Illustrations of scatter in FCCS.** Forward-scatter (FSC) is used as a measure on cell size; side-scatter (SSC) is a measure on a combination of internal optical density, and cell surface roughness. For each of forward- and side-scatter (as well as for every fluorescence channel) the computer records the peak amplitude of scatter within the event (known as scatter 'height', H), time duration of the scattering signal (scatter 'width', W), and area under the curve of the detector's electronic signal (scatter 'area', A). The latter is used as a measure on the total quantity of scattered light at low-angle (FSC) and high-angle (SSC) conditions. . . . . 95
- 4.5. **2D facet plot of scatter data collected within an FC experiment.** No fluorescence is included in this plot: the data shown describe scattered light only. The metrics being compared are (diagonally from top left): forward-scatter area (FSC.A), height (FSC.H) and width (FSC.W); side-scatter area (SSC.A), height (SSC.H) and width (SSC.W), as described in Figure 4.4. Although this chart is symmetric about the primary diagonal, it is evident that this parameter space will expand rapidly as fluorescence data is included, since each fluorescence channel will add up to three optical parameters to search. . . . . 96

4.6.	<b>Illustrative data-flow for Intelligate analysis software.</b> Data from one or more samples is imported via the FCS interchange file format (1) at which point it is log-transformed (2) and pre-processed, removing events with NaN (“Not a Number”) values, and with marginal values beyond the dynamic range of the photodetectors (since these would skew the distributions). . . . .	99
4.3.	<b>Overview of a FCCS apparatus.</b> Hydrodynamic focussing (1) takes a stream of suspended protoplasts (or single cells, polymer micro-beads, etc.) and aligns them in single file for interrogation by the optical system. The array of lasers, filters, and cameras (2) records optical parameters of each cell, which is plotted by the computer as a single dot on a scatter graph - the cumulative data is known as a ‘dot plot’. The user then selects one or more populations of interest from the graph with a geometric selection known as a ‘gate’. When a sort is started, the computer controls the electric field between two plates in the sort chamber (3) using electrostatic thrust to displace drops in one direction or according to their relation to the gate’s optical parameters. . . . .	103
4.7.	<b>CND-positive population of <i>Arabidopsis</i> protoplasts gated by <i>Intelligate</i> software</b> (red dots) as a subset of 100,000 protoplasts sampled from whole root tissue (black circles). Proportion of events falling within the gate is annotated in percent %. . . . .	104
4.8.	<b>Heatmap of data covariance matrix from FC data obtained in this study.</b> Lighter squares indicate higher covariance. This is a visual representation of the basis used for biclustering model formation. Off-diagonal comparisons with unusually low covariance (darker colours) indicate physical data comparisons where clusters may be most distinct. Using a 405nm violet laser excitation (denoted “V405”) the primary channel used to detect nanodots (dashed outline) is a 50nm band-pass filter with a 450nm peak transmittance - thus denoted “450.50”. The heatmap shows increased covariance between this and the side scatter ‘area’ channel (green solid outline) which is consistent with the observation that c-dots have high optical absorbance. The heatmap shows decreased covariance (blue solid outlines) with 610.25 and with 530.30 - the latter being excited by the blue 488nm laser, as opposed to the 405nm violet - although these latter two channels also show low covariance with one another, indicating the possibility of a multi-factor interaction between all three channels. . . . .	105

# List of Tables

2.1.	Table of experiments reported in Chapter 2 . . . . .	37
3.1.	Table of experiments reported in Chapter 3 . . . . .	67
3.2.	<b>Normalised read counts of four flg-22-associated genes from RNA sequencing of <i>Arabidopsis</i> roots treated with flg-22 and carbon nanodots.</b> Each column represents one technical replicate of fifty 9-day-old seedlings sampled 48 hours after treatment application with: mock solution (water, “M”), carbon dots with a short-acid linker (“C”), flg-22 peptide (“F”), carbon dots functionalised with a long linker and flg-22 (“CTF”), and carbon dots functionalised with flg-22 via a short linker (“CF”). . . . .	86
3.3.	<b>Significance values for 2-tailed homoscedastic t-test of RNAseq counts for four markers of flg-22 response in roots.</b> . . . .	86
3.4.	<b>MIPS Biomaps analysis of expression data versus the TAIR10 baseline transcriptome.</b> Gene list for the condition highlighted in yellow is seen in table 3.5. Adjusted p-values calculated using a false discovery rate of 5%. . . . .	87
3.5.	<b>Biological process GO term analysis of genes highlighted in table 3.4 (flg-22-functionalised carbon nanodots versus nanodots with the short linker only).</b> . . . . .	88

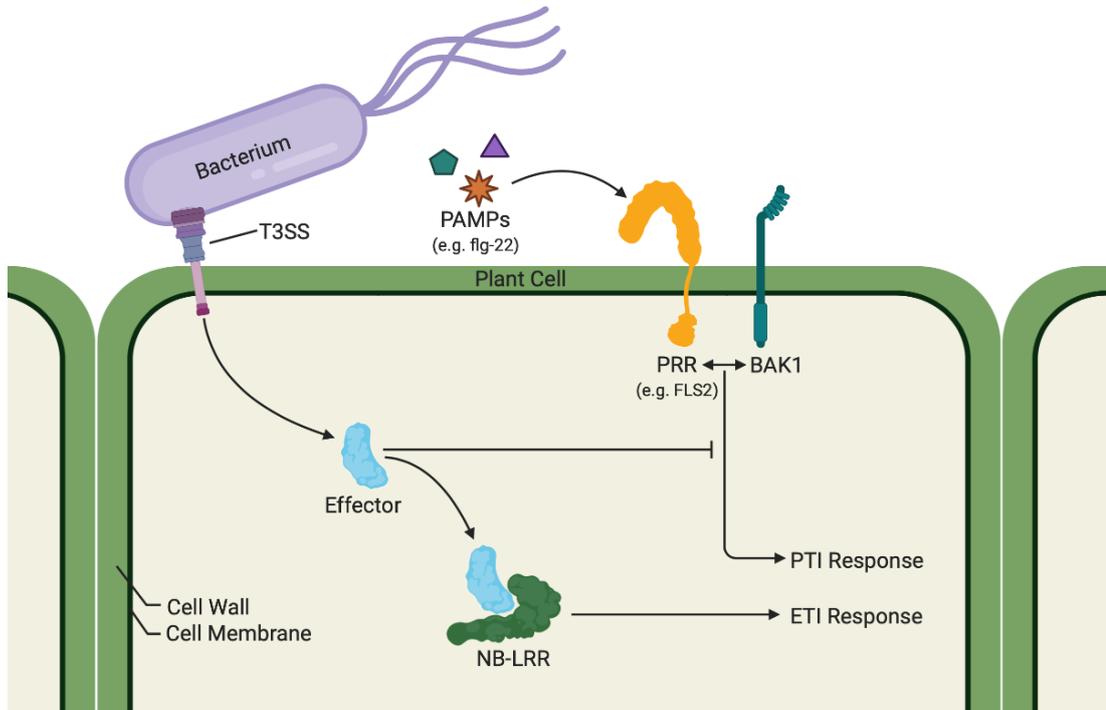
# 1. Introduction

## 1.1. The root: a specialised transport and sensing system

Unlike most other multicellular organisms, plants have no central nervous system: the entirety of their signalling and response to and from the outside world is governed by hormonal and gene expression changes. Because of plants' sedentary lifestyle, plant genomes (and meta-genome) must encode much of what would be 'active behaviour' in animals: for example feeding behaviour, response to threat, and so on. This goes some way towards explaining why plants tend to have longer genomes and higher ploidy than mammals - epigenetics in plants has a more complex role to fill than it does in mammals.

One such role is the identification of sources of stress, including the recognition of pathogens. Lacking an adaptive immune system, plant immunity depends on two systems innate to every living cell in the organism, as depicted in figure 1.1. Broadly speaking, a variety of molecules known as pathogen-associated molecular patterns (PAMPs) are released by pathogens into the extracellular space - for example flg-22, a peptide fragment of bacterial flagellin, one of the protein components of flagella. A PAMP is detected by pattern recognition receptors at the cell surface: this activates a kinase cascade, triggering an immune response from the cell known as pattern-triggered immunity (PTI). PTI behaviour typically involves the generation of reactive oxygen species (ROS); changes in expression of immunity-related genes such as peroxidases and MYB genes (a diverse family of transcription factors with roles in metabolism, stress response, and developmental control); and down-regulation of growth-associated genes (Carter et al., 2013). In order to function correctly, many pathogen recognition receptors rely on interactions with other membrane-bound proteins such as the BAK1 (brl-associated receptor kinase 1), a leucine-rich repeat receptor-like kinase which interacts with BRI1 to mediate brassinosteroid signalling pathways, as well as sugar-related growth and development (Peng et al., 2018). Some pathogens can exploit pili such as the type-3 secretion system (T3SS) to inject effector proteins

directly into the plant cell, bypassing the cell membrane where pattern recognition receptors localise. These pathogenic effector proteins are recognised from within the plant cell by nucleotide-binding leucine-rich repeat receptor-like kinases (NB-LRR), which activate transcriptional changes in a behaviour known as effector-triggered immunity (ETI).



**Figure 1.1.: Illustration of plant-pathogen interaction systems.** Pathogen associated molecular patterns (PAMPs) are detected in the extracellular spaces by pattern recognition receptors (PRR). This triggers a cascade of transcriptional changes known as pattern-triggered immunity (PTI). The bacterial flagellin component flg-22 is one such PAMP: it is sensed by the Flagellin-Sensitive 2 (FLS2) protein, which interacts with the BAK1 receptor-like kinase after binding flg-22 in order to effect intracellular activity. Some pathogens have mechanisms such as the type-3 secretion system (T3SS) pilus, which can inject pathogenic effector proteins directly into the intracellular space, and thereby suppress PTI: plants mostly rely on nucleotide-binding leucine-rich repeat receptor kinases (NB-LRR) to recognise these and initiate a response known as effector-triggered immunity (ETI). Figure created using BioRender.com (2020), based on Dodds and Rathjen (2010).

Though some of the detail about plant immunity remains unknown, the first discoveries made often concern a plant known commonly as thale cress, or sometimes mouse-ear cress. Thale cress, known to scientists as *Arabidopsis thaliana* (pictured in figure 1.2, referred to as *Arabidopsis* for brevity) is a superficially unimpressive weed, whose characteristics of small size, short genome, low ploidy, short generation length, as well as its high degree of genetic commonality with

crop plants such as *brassica* and its ease of genetic transformation, have historically made it science's favourite model organism for flowering plants everywhere. When a new tool or method for plant sciences is tested, *Arabidopsis* is most often the first species it is tested on, and there are a number of information resources available, reviewed by Woodward and Bartel (2018).

For some of the experiments in this study, there are relatively few established biological protocols to use directly, but those that can be built from are mostly best-optimised for *Arabidopsis*. For example, the hydroponics culture apparatus in 3.2.1 is a method established by Gifford et al. (2008), and the protoplast generation protocol referred to in 3.2.4 is readily available for *Arabidopsis* cultures. The availability of mature genome annotations allows better analysis of any data collected from gene expression experiments, and the relatively thin, translucent roots of *Arabidopsis* lend themselves well to optical microscopy of all kinds.

## 1.2. Cell type-specific analysis in plants

Like all multicellular organisms, plants are comprised of a large number of cells whose structures and behaviours are specialised according to specific tasks within the organism. It has been shown that transcriptional profiles are typically different in differing cell types, and signalling cascades govern key factors in growth and development: cascades between tissues, which may be confounded if considering only the whole-root level (Rogers et al., 2012). For example, FLS2 expression levels in root are typically higher inside the innermost tissues than in the outermost (epidermal) cell files, in contrast to the intuitive expectation that cells closest to potential pathogens might have higher prevalence of an immune receptor (Gomez-Gomez and Boller, 2000). Thus it is important to measure changes in each cell type, in order to gain a detailed understanding of each interacting sub-system in the root as a whole organ.

A number of approaches have been taken to isolate cells from differing tissue types to investigate expression in more spatial detail. Laser-capture micro-dissection (Asano et al., 2002; Kerk et al., 2003; Nakazono et al., 2003) is a technique allowing different tissue types to be physically separated, and RNA can be subsequently extracted from the dissected cells; however, it relies on consistent visual recognition of the different cell types, and the ability to make each cell type mechanically accessible. The significant time involved in sample preparation also indicates that this method would not scale well to larger numbers of individuals.

Fluorescence Cell Sorting (FCS) is a relatively high-throughput technique for sorting cells into different types, mostly when fluorescent marker lines already

exist for these types. FCS can be coupled to expression experiments such as microarray or RNA-seq to determine expression changes in different cell types. Protoplast generation - the process required to prepare plant cells for FCS - has measurable effects on gene expression in *Arabidopsis* root cells (Birnbaum et al., 2005; Gifford et al., 2008). Nevertheless, this method can be used to generate high-quality expression data from cells with various fluorescent characteristics - this has been used extensively in conjunction with fluorophore-conjugated antibodies in mammalian immunology (Porter et al., 2017; Birnbaum et al., 2005). Single-cell RNA-sequencing (scRNA-seq) can produce gene expression data for each cell in a sample, and thus can garner expression data at a very high spatial resolution. Especially useful in tissues which are easily accessible, and for which fluorescent tags are not yet available, this technique has been used to create 'expression maps' of *Arabidopsis* root cells (Efroni and Birnbaum, 2016; Birnbaum et al., 2003). It is possible to combine the two technologies, creating a highly detailed dataset from each cell from a population of cells in a tissue with particular fluorescent characteristics (Freiherr von Boeselager et al., 2018).

The procedure for genetic transformation in plants is more complex than it is for micro-organisms, and is relatively time-intensive, requiring the transformation of at least one intermediate bacterial culture for use as a vector to infect the host plant (which is typically wounded in the procedure). Indeed, it is usual to transform the construct first into *E. coli*, in order to quickly grow a sizeable clonal culture, before transferring the construct to e.g. *Agrobacterium tumefaciens* for plant transformation, before transformants are regrown, and self-crossed if a homozygous line is required. Even if all the previous steps are entirely successful, the transgene may subsequently be lost from the plant genome in subsequent generations, or it may be 'silenced' - retaining its presence in the genome, but devoid of any measurable expression. For these reasons, the es-



**Figure 1.2.:** Photograph of a week-old *Arabidopsis thaliana* seedling. Scale bar 10 mm

establishment of stable transgenic plant

lines is often in itself a time-intensive and expensive process for even simple genetic constructs.

In the industrial and agricultural sectors, germline editing is seen as fairly permanent, and the inclusion of antibiotic resistance genes as selection factors is one of the reasons that the prospect of transgenic crops has sparked a reasonable quantity of ethical discussion, in the process garnering significant quantities of negative media attention, especially in Europe. Successful co-option of the CRISPR-Cas9 system for genetic editing of plant species has provided an alternative to antibiotic selection, but (at least for the time being) this raises its own issues of potential off-target editing, including insertion/deletion mutations up to one thousand base pairs either side of the target (Kosicki et al., 2018).

On the other hand, if a reliable, quantitative input-output model of plant gene expression could be developed, the desired behavioural or developmental response might be achieved by dynamically mediating the system inputs (e.g. light, heat, ambient levels of various volatile organic compounds) rather than directly modifying the host genome.

## **1.3. Synthetic Biology: methods and mechanisms**

Synthetic Biology is the application of engineering approaches towards (conventionally, molecular) biology: applying systems analysis and modelling techniques, as well as pursuing engineering principles such as modularity and standardisation. As the field develops, it is becoming increasingly clear that while simple techniques are enough to create basic synthetic 'gene circuits' within a controlled and stable environment, host factors and other more complex secondary factors preclude the stable operation of more complex (and therefore more useful) designs for synthetic constructs where they are based on (naïve) deterministic models with orthogonal components.

Considering plants as an object of study, the highly complex native architecture of interacting transcriptional systems present a synthetic biology conundrum, which is especially complex considering that current practical techniques used to validate computational models rely heavily on biological elements: the few which do not rely directly on transgenic constructs, often depend on other biological tools with natural origins in other organisms (e.g. engineered monoclonal antibodies). In theory, the fact that such elements come from foreign organisms is used as justification for the claim that they are perfectly orthogonal - and indeed, it would

seem sensible to assume that a plant, having no genetic context for a transgenic system, would be unlikely to have that system interact with the native intracellular environment. In practice, however, it is shown that (at the very least) native gene expression is reduced by the expression of a genetic construct - reducing the observability of native expression changes - and this effect is especially apparent using 'strong' promoters as often seen in transgenic constructs (Darlington et al., 2018; Kim et al., 2020; Nikolados et al., 2019). This is one reason that the vision of synthetic biology - the detailed engineering of complex biological behaviour - remains unattainable, and one of the reasons that the field has come under criticism in the past for its lack of apparent transferability into the industrial sector (Kwok, 2010).

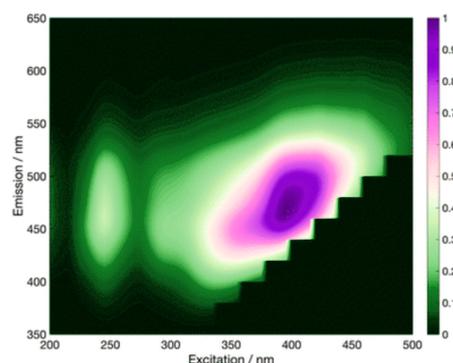
The above is an excellent example of the so-called 'observer effect' (also known as the 'probe effect') - familiar to quantum physicists, computer programmers, and behavioural scientists - whereby the practical act of testing has an impact on the phenomenon being observed. The effects become more and more problematic as the systems developed become more complex, which is why study of the highly complex regulatory networks in plant biology is among the first to be hampered by this obstacle. In specific terms - it is very difficult to isolate elements from a highly convoluted network such as the symbiotic nitrogen-fixation interactions between legumes and rhizobia; and this issue is further complicated by the probability that any practical test may affect the network behaviour. Using carbon nanodots, it may be possible to avoid generating transgenic plants, and thus, it might be possible to avoid crosstalk with native transcription. However, the introduction of nanodots may itself effect a change on the root transcriptome - as other varieties of nanomaterial have been known to do (Saharan and Pal, 2016; Hill et al., 2019; Jin and Gao, 2009; Li et al., 2015).

As with many fields encountering the observer effect, it is impractical if not impossible to preclude its impact on experimental results: however, it is at least conceptually possible to control for the observer effect by introducing a 'mock' experiment, which observes a null (control) phenomenon or treatment. For example, when using a transgenic plant, an ideal internal validity control would be the observation of an identical plant transformed with a 'dummy' transgene under the same promoter region. However, because of the difficulty of producing stable transgenic plant lines, such controls are rarely implemented. Moreover, in plants and other organisms with histone regulation, there is no guarantee that two like transgenes will garner the same level of expression within each cell type - and expression level and rate is both difficult to practically quantify, and varies stochastically from cell to cell (Araújo et al., 2017). An internal validity control

for carbon nanodots, on the other hand, is almost as simple practically as it is in theory - because a 'dummy' payload can be complexed to the nanodots by the same synthesis route as the treatment. Swift et al. (2019) have shown that payload on nanodots can affect uptake in plants - and this may present a potential issue if the treatment and control have very different uptake characteristics - but comparing uptake of two different fluorescent particle types is relatively quicker and easier than comparing the expression levels of two different transgenes.

Ideally, any system used to trace the location of another molecule has a minimal impact on its biological characteristics. Un-modified carbon nanodots are readily water-soluble, and the solution remains stable at biologically-relevant pH ranges, which makes them well-suited to a variety of biological contexts, which confers compatibility with a variety of delivery protocols. Functionalising carbon dots with large non-polar groups, unsurprisingly, reduces their aqueous solubility, although this can be managed by use of an appropriate solvent system.

Confocal microscopy produces high-quality fluorescence data in up to three spatial dimensions plus time, and this is a desirable characteristic for ascertaining where the fluorescent carbon dots tend to collect within the plant tissue. On the other hand, this is not a method that lends itself well to quantitative results: negative controls can be used to set the sensitivity of the photodetectors to minimise the masking effects of biospeckle and autofluorescence on the results, and laser power and detector gain can be recorded, but biological variability makes it difficult to establish a robust dynamic range for the detector which can be held consistent across several experiments. Even when such a range is established, results may be very clear on a digital display, but difficult to discern on paper (particularly so for dark-field fluorescence data). Flow cytometry (FC) is a method that rapidly accumulates fluorescence data on every single cell within a sample: as such, it is excellent for analysing a large quantity of cells in a way which is impractical to do with microscopy alone. FC also gives information about the distribution of fluorescence between different



**Figure 1.3.: Fluorescence excitation-emission plot for carbon nanodot cores.** Colour scale represents intensity of emitted light at the corresponding wavelengths of excitation (x-axis) and emission (y-axis). Figure reproduced with author's permission from Swift et al. (2018).

cell type: although because of the extensive pre-processing required, it becomes non-trivial to determine from where in the plant the cells originated, and because this process removes cell walls, it removes with it information about the cells' interstitial environment. It is possible to use flow cytometric cell sorting to isolate populations of interest, and then observe the sorted cells under a microscope: however, by the time they have been sorted, the already fragile cells have been stressed by hydrodynamic pressure and laser illumination, which leaves them relatively difficult to recognise visually under a microscope. On the other hand, the ability of FC to identify a distribution of fluorescence within cells throughout the plant allows for a number of like experiments to be conducted at different time points, providing information about the evolution of that fluorescence over time.

### **1.4. Automation, informatics, and the rise of data sets**

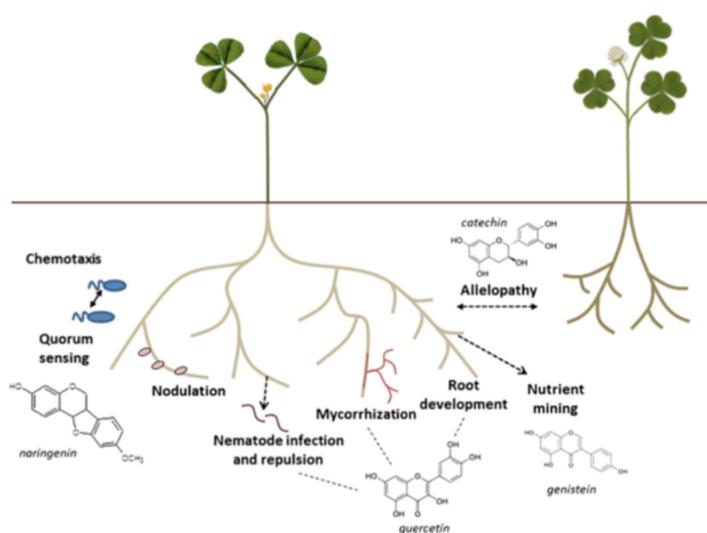
As the complexity of biological systems is made increasingly apparent, so too is the data collected in scientific experiments becoming more complex. Advances in computing performance have enabled scientific experiments to be performed, whose data is no longer comprehensible without computer use. Not only are the raw data generated in e.g. RNA-sequencing experiments too large to be stored in anything but a digital format, and the numerical pre-processing calculations impractical to perform manually: the high-dimensional data sets generated in such experiments makes direct comparison between samples a complex issue(Wang et al., 2017). The RNA-sequencing sector has handled this well, with a number of statistical analytics available to condense data into human-comprehensible forms (although the comparison itself is still often manual)(Guo et al., 2019). Flow cytometry (along with its sister technology, flow cytometric cell sorting ) is one example of a method which collects high-dimensional data. In Chapter 4, we outline a computational tool designed to aid in the translation of high-dimensional sensor data into a human-manageable analytical space: increasing analytical power within an experiment, facilitating machine operation, increasing repeatability between individual samples, and reducing the opportunity for unconscious operator bias to affect experiment outcomes. In principle, we thus enable subsequent high-complexity experiments to be conducted on cells pre-sorted using FCCS. This would allow data to be gathered from e.g. several deciles of effective dose from an applied absorptive or fluorescent elicitor, resulting in a greatly increased signal-to-noise ratio when examining transcriptional data from

those cells, and potentially increasing the overall statistical power of the experiment further using methods such as scRNAseq, in order to gain information about the variability as well as the level of response at each input dose.

While our prime focus has been to use this tool as a means to repeatably isolate rare events from a number of variable and highly autofluorescent plant cell samples, the cross-applicability of such a technique is apparent in sectors such as mammalian cell biology, where machine-learning approaches are already being adopted in e.g. imaging flow cytometry experiments (Eulenberg et al., 2017).

## 1.5. Applications of root biology

### 1.5.1. Crops, Fertilisers, and the GMO debate



**Figure 1.4.:** Illustration of root-microbe interactions in *Medicago truncatula*. Adapted from G.Muscatt et al (personal communication).

Sustainable use of fertilisers is one of several unanswered questions in the conundrum that is long-term global food security. In particular, cereal crops representing the majority of human food production (Tilman et al., 2002) still rely heavily on chemical fertilisers in order to generate profitable yields. Much effort to reduce this dependency has revealed rhizospheric nitrogen fixation as a potentially lucrative option: such as is found between barrelclover *Medicago truncatula* and its fungal symbiont *Sinorhizobium meliloti*. In this and many similar systems, a plant root typically offers flavonoid sugars to a microbe mutualist in exchange for

nitrate- and ammonium-rich organic compounds(Herrbach et al., 2016). Figure 1.4 shows some of the root exudates of interest in these interactions between the roots of this species and its biological environment.

Fertilisers are not the only synthetic chemicals being sprayed in bulk onto crop fields: pesticides are commonplace in modern agriculture as a means of phytopathogen control. Research into plant immunity is increasingly challenging this paradigm - suggesting that the priming of plant immune systems by non-transgenic means may produce better yields in the long-term than e.g. relying on the yield and ecological stability of R-gene transgenic crops(Lozano-Durán and Zipfel, 2015; Creissen et al., 2016; Kim et al., 2020). Such 'plant elicitors' may also mediate the fears that currently dominate legislative restrictions in this sector - that germline transgenes could find their way into the wider ecosystem via uncontrolled horizontal transfer - by reducing or altogether eliminating the requirement for transgenic constructs, perhaps in favour of more ethically palatable cysgenic varieties.

## **2. Nanoparticles as tools to deliver and trace molecules for biological investigation**

### **2.1. Carbon nanodots as an affordable and adaptable delivery system**

The overall scientific trend towards more detailed observation necessarily requires the incorporation of more advanced techniques and the gathering of larger quantities of data in order to demonstrate similar increments in scientific understanding. The object of biological study is intrinsically variable: at the cellular level, the process of transcription is stochastic (Araújo et al., 2017) and this makes it difficult to separate intrinsic variation from induced effects, as seen in prior attempts to construct genetic circuits in mammalian cells (Hebenstreit et al., 2012). This chapter outlines the primary steps in the creation of an abiotic, fluorescent marker, with a modular system of reactions that can combine it with a variety of chemicals of conventional biological interest: thus facilitating the identification of their induced effects, making the chemicals visible through conventional fluorescence imaging methods, including laser-scanning fluorescence microscopy.

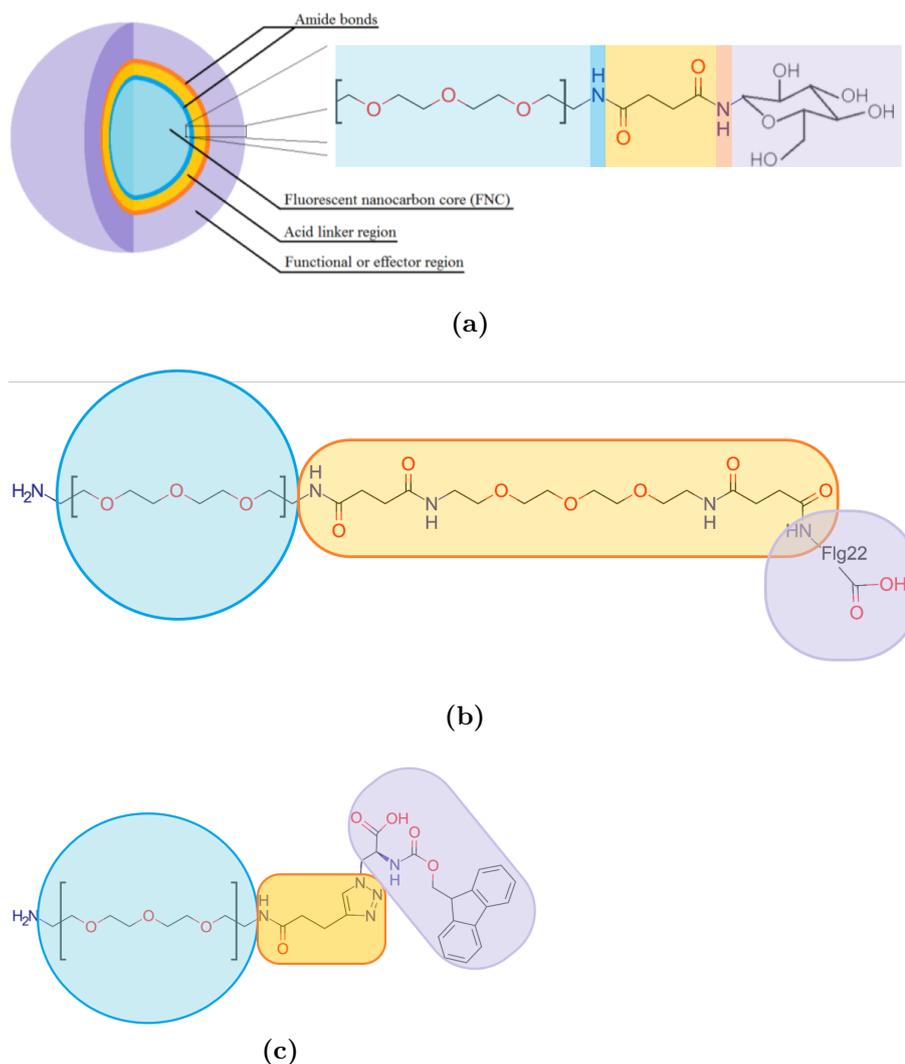
Carbon nanodots are a good choice for this marker, since they are easy to produce with basic equipment - such as a domestic microwave, a vacuum pump, and a magnetic stirrer - much of which is already available in a biology lab environment (Hill et al., 2016). Production and purification is feasible at practical scales using commercially available materials: for example the two chemicals required for the core formation reaction in 2.3, TTDDA and glucosamine, are sold in bulk quantities for around £0.40 and £0.70 per gram, respectively (Sigma-Aldrich). Additionally, many of the subsequent reactions are similarly simple in their required conditions: often requiring no more than stirring overnight at room temperature.

Carbon nanodot cores are hydrophilic, with a stable zeta-potential in water, which

is to say that their colloidal suspension in water is highly stable. Carbon dots also have high quantum yield and photostability - meaning that they fluoresce brightly, and are resistant to photobleaching - and an estimated shelf life of activity of 12 months or more, as reviewed in (Li et al., 2012a; Himaja et al., 2015). Because fluorescence is not dependent on the inclusion of any heavy-metallic elements (unlike cadmium-selenium and zinc sulfide based quantum dots), they can be used and disposed of by routes already familiar to biology labs for materials with a low- to medium-level of chemical hazard. All of these properties are convenient for use in biological research sectors.

Because the surface of the nanodot core is not chemically uniform, it is desirable to adapt the surface chemistry of the nanodot, in order to reduce the probability of forming undesired side-products when it comes to attaching the desired payload. Importantly, a family of simple synthetic chemistry reactions exists such that the surface chemistry of the nanoparticle in question can be adapted to support a variety of relevant molecular payloads. This results in a tripartite system: the core nanoparticle, a molecular adapter or linker molecule, and the desired payload, as seen in 2.1. The longer and more hydrophobic the linker molecule, the greater the degree of steric isolation between nanodot and payload; conversely, a shorter, more hydrophilic linker is more likely to maintain the water-solubility of the core particle. Once functionalised, the hydrodynamic radius of the complete particle may also play a role in its uptake and/or function in biological context. In the absence of an established standard, we selected a small variety of linker particles, as depicted in the figure. The largest of these, when functionalised, was slightly more difficult to work with when forming concentrated aqueous stock solutions: for this reason, we avoided using much larger linker molecules in order to avoid working with recalcitrant stocks.

The simplest of these reactions is an amide bond formation between the nanoparticle and the desired payload: a reaction which can take place overnight at room temperature, if an activator chemical such as CDI (1,1'-Carbonyldiimidazole) or EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride) is used. It is this reaction which is used to add the desired linker molecule to the nanoparticles depicted in figure 2.1.



**Figure 2.1.:** Illustration (not to scale) depicting structure of carbon nanodots functionalised with a layer of (a) short acid linker and amino-glucose (b) long linker and flg-22 peptide, and (c) short alkyne linker and Fmoc- $\beta$ -azido-Alanine-OH, respectively. Here it can be seen that, through the use of a variety of linker chemicals (yellow), a variety of payloads (purple) may be loaded onto the particles (blue) with a greater or lesser degree of surface passivation; also, that the nature of the linker and payload in terms of polarity and molecular weight may affect overall physical properties of the final particle: especially since surface functionalisations may not always result in a homogenous corona surrounding the nanodot core, as seen in (Swift et al., 2018).

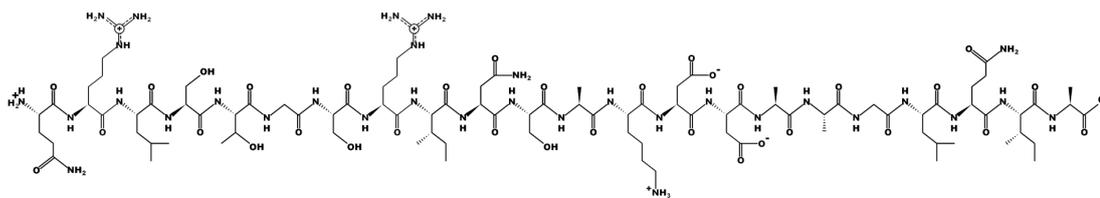
Moreover, in order to provide proof-of-concept for a truly modular system, compatibility with existing efforts at synthetic biology standards is highly desirable. For example, copper(I)-mediated azido-alkyne cycloaddition (CuAAC) reactions a.k.a. “CLICK” chemistry, which are one such standard: designed by the synthetic chemistry community to be bio-orthogonal, and compatible with the wide variety of molecules, oligomers, and polymers dealt with in biological sciences,

as reviewed in (Moses and Moorhouse, 2007). CuAAC reactions are highly site-specific in terms of the parts of the molecules being bonded together, which can be of great use when considering payloads with specific structure-function relationships. Furthermore, Marchand et al. (2019) have discovered a pathway for terminal-alkyne amino acid biosynthesis, which may pave the way for cellular production of alkyne-labelled proteins to be used in such reactions. In the meantime, the standard production of CuAAC reagents is by conventional means of synthetic chemistry.

### **Biological interactions often exhibit strong structure-function relation**

In theory, CLICK chemistry is a highly modular, highly workable system with unmatched specificity and yield. In practice, the reaction parameters can require quite a lot of tuning; however, this family of reactions is still the most specific option known, and as such is particularly useful when working with e.g. ligand-receptor systems, when the orientation and secondary structure of the functionalising molecule may be highly important for its functionality *in vivo*. For example, the secondary structure of the lipooligosaccharide NodL is critical to its function as a signalling molecule in interactions between *Sinorhizobium meliloti* and *Medicago truncatula*, which are part of the compatibility system that can lead to initiation of nodulation symbiosis (Oldroyd, 2013; Ferguson et al., 2010).

Another molecule for which conformation is important, flg-22 is a 22-residue amino-acid oligomer; a well-documented plant elicitor, with well-characterised effects and established treatment protocols in plants. The structure is shown in figure 2.2. A key component of bacterial flagellin, flg-22 is a so-called pathogen-associated molecular pattern (PAMP) detected at the cell surface by *Arabidopsis* and other higher plants via the Leu-rich repeat transmembrane receptor kinase Flagellin Sensitive-2 (FLS2) (Zipfel et al., 2004; Beck et al., 2014). This interaction elicits a response from plants, leading to regulation of immunity-related genes, and subsequent changes involved in pathogen recognition: a reaction known as pattern-triggered immunity.



**Figure 2.2.:** Chemical structure diagram of the flg-22 peptide in zwitterionic form. The amino acid sequence (from left to right) is QRLSTGS-RINSAKDDAAGLQIA.

### Using carbon nanodots to investigate biological interactions

The well-characterised resistance response in roots, combined with established treatment protocols, commercial availability, and solubility in water, make flg-22 an excellent candidate as a test molecular payload for carbon nanodots in the context of plant biological study. This chapter explores how a signalling molecule such as flg-22 can be attached to carbon nanodots for the first time. Both CuAAC and amide coupling approaches will be adopted, with a view to treating plants with the novel nanodots in a later chapter. Table 2.1, below, summarises the attempts taken in this chapter to achieve that goal.

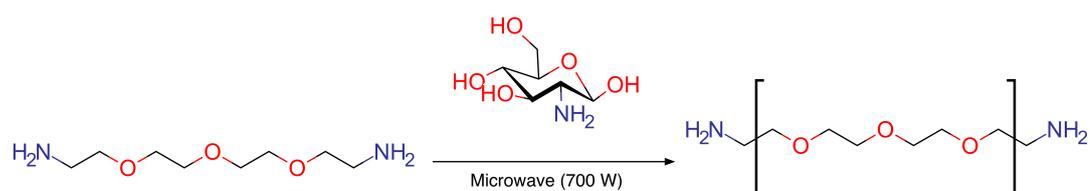
Description	Method(s) of characterisation	Reaction method in section	Results shown in figure(s)	Notes
Carbon nanodot core formation	Fluorimetry	2.2.1	2.9(a)	Replication of previously-published result
Analysis of the synthetic flg-22 oligopeptide	NMR spectroscopy	n/a	2.13	Characteristic resonance pattern identified
Nanodots surface adaption with short-acid linker	NMR spectroscopy	2.2.2.1	2.19	Replication of previously-published result
Nanodots surface adaption with long-acid linker	NMR spectroscopy	2.2.2.2	2.12	Linker present and intact
Nanodots surface adaption with alkyne linker	NMR spectroscopy	2.2.2.3	2.20	Linker present; alkyne intact
Alkyne-adapted nanodots functionalised with Fmoc-beta-azido-Ala-OH	NMR spectroscopy	2.2.4	2.20	Experiment failed (no reaction)
Short-acid-adapted nanodots functionalised with flg-22	2D NMR spectroscopy, Fluorimetry	2.2.3	2.16, 2.18, 2.9(b)	Linker intact; reagents diffuse together
Long-acid adapted nanodots functionalised with flg-22	2D NMR spectroscopy, Fluorimetry	2.2.3	2.19, 2.17, 2.9(c)	Linker resonance-shifted or displaced; reagents diffuse together

Table 2.1.: Table of experiments reported in Chapter 2

## 2.2. Methods

### 2.2.1. Simple synthesis of carbon nanodot cores

The reaction scheme for this step is shown below in fig. 2.3, first published by (Hill et al., 2016). 1.00 g (4.63 mmol) of glucosamine hydrochloride is dissolved in 20 ml deionised water in a 500 ml Ehrlenmeyer flask, and 1.35 ml (5.09 mmol) of 4,7,10-trioxa-1,13-tridecanediamine (TTDDA) is added. This mixture is then heated in a domestic microwave (3 mins, 700 Watts). The produced brown oil is then washed with chloroform several times: sonicating for 5 minutes each time, and discarding the supernatant; until the solvent runs clear. The results from this chapter were primarily obtained using chloroform washes: alternatively, tetrahydrofuran (THF) may be used. Chloroform tends to require fewer washes, but is generally a more hazardous substance than THF, and chloroform use is tightly regulated in some countries (e.g. USA). Additionally, chloroform tends to permeate through polypropylene vessels (e.g. microcentrifuge tubes), which can ultimately result in solvent leaking from plastic vessels: this would become an issue if glass alternatives were not available. After purification, the brown oil is then dissolved in 20 ml deionised water and centrifuged through a 10,000 molecular weight cut-off (MWCO) spin filter (GE Vivaspin® 20) at 5000 rcf for 1 hour. Material to be analysed at this stage is then passed through a 0.2 micron syringe filter. The resulting solution is decanted into either a 100 ml round-bottomed flask, or one or more 50 ml centrifuge tubes, and freeze-dried to yield approximately a gram of carbon nanodots.

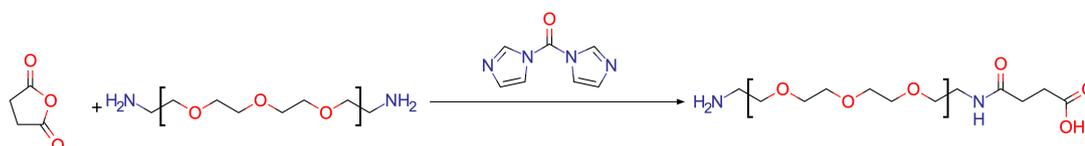


**Figure 2.3.:** Reaction scheme for one-step microwave synthesis of carbon nanodots. TTDDA is polymerised to form carbon nanodots, in the presence of glucosamine and microwave energy. Nitrogen-bearing groups are coloured blue; oxygen-bearing groups are in red. After Hill et al. (2016).

## 2.2.2. Adaption of the nanodot surface chemistry

### 2.2.2.1. Short acid adaption enables subsequent reaction with primary amines

A CDI-activated amide bond formation with succinic anhydride (IUPAC name: oxolane-2,5-dione) and the free amine on the surface of carbon nanodot produces a carbon-dot with a carboxylic acid functional group at the surface, as seen in fig. 2.4. The product of this reaction allows primary amine groups to be bonded to the nanoparticle by subsequent similar reactions, making it highly useful as an intermediate step. This method was first published in (Hill et al., 2018). 0.4 g of carbon dots dissolved in methanol were stirred overnight with half-equivalence by mass (0.2 g) of succinic anhydride, which had been pre-activated with equivalent mass (0.2 g) of CDI. The product was freeze-dried, and subsequently purified by size-exclusion chromatography, to remove the CDI along with any unreacted material. For size exclusion, the product was dissolved in methanol, then carefully added to the top of a 30 cm glass chromatography column, which had been set up with silica media and methanol as the solvent. This formed a brown-coloured band, which processed down as solvent passed through the column, until it was collected at the other end, the corresponding fraction of liquid being separated from the rest. Other polar solvents (e.g. ethanol, water) could have been used: methanol was chosen as a solvent which completes this process quickly.

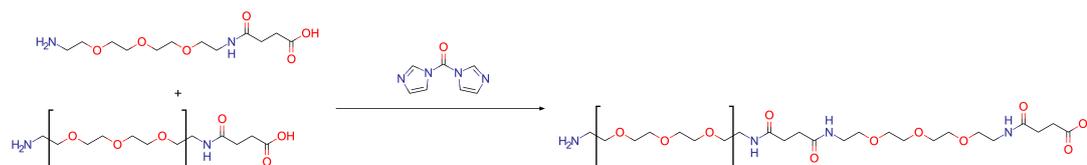


**Figure 2.4.:** Reaction scheme for addition of a short acid linker to carbon nanodots. The CDI enables a ring opening of succinic anhydride, allowing it to react with the primary amine on the carbon nanodot, resulting in a aliphatic carboxylic acid group coupled by an amide bond to the nanodot.

### 2.2.2.2. Long linker adaption permits greater steric isolation between core nanodot and functional payload

This reaction is the same mechanism and similar in method to the previous: an amide coupling is conducted between a primary amine and a carboxylic acid. This reaction, however, uses the product of fig. 2.4 in combination with a longer carboxylic acid linker, the latter also the product of an amide coupling between TTDDA and succinic anhydride. The purpose of a longer linker is to reduce the risk of steric interactions between the nanoparticle and the molecular payload -

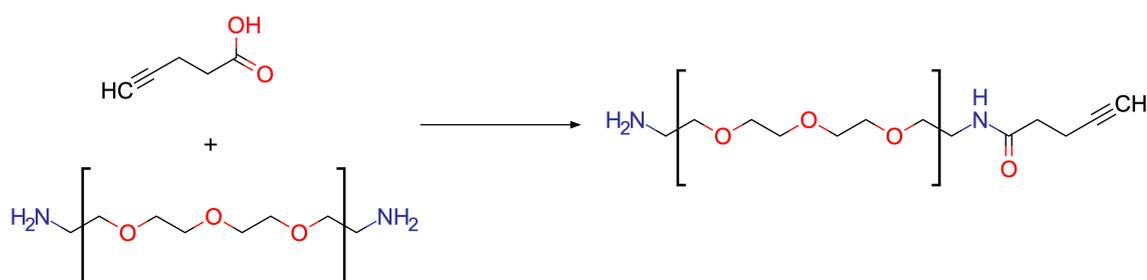
this is of particular importance where there is a structure-function relationship in the payload molecule. 100mg of acid-coated carbon dots (purified from the reaction shown in 2.4) in methanol at a concentration of 190 mg/ml were added to the reaction vessel. To this was added half-equivalence by mass of CDI (50 mg) in order to activate the carboxylic acid groups on the carbon dot surface. After 5 minutes, the same mass (50mg) of the linker molecule was added; the reaction was stirred overnight at room temperature. Products were freeze-dried, and purified by size exclusion using a 10 ml, 6000 MWCO dextran desalting column (Thermo Scientific).



**Figure 2.5.: Reaction scheme for addition of a long acid linker to carbon nanodots.** The carboxylic acid group at the carbon nanodot surface is activated by CDI before the linker chemical is added, so as to minimise CDI-activated polymerisation of the linker. The target reaction is an amide coupling between the carboxylic acid group at the surface of the nanoparticle and the primary amine of the linker chemical, resulting in a nanoparticle with a long aliphatic linker, terminating in a carboxylic acid functional group.

### 2.2.2.3. Alkyne adaption of nanodots in preparation for a copper(I)-mediated azido-alkyne cycloaddition (CuAAC) reaction.

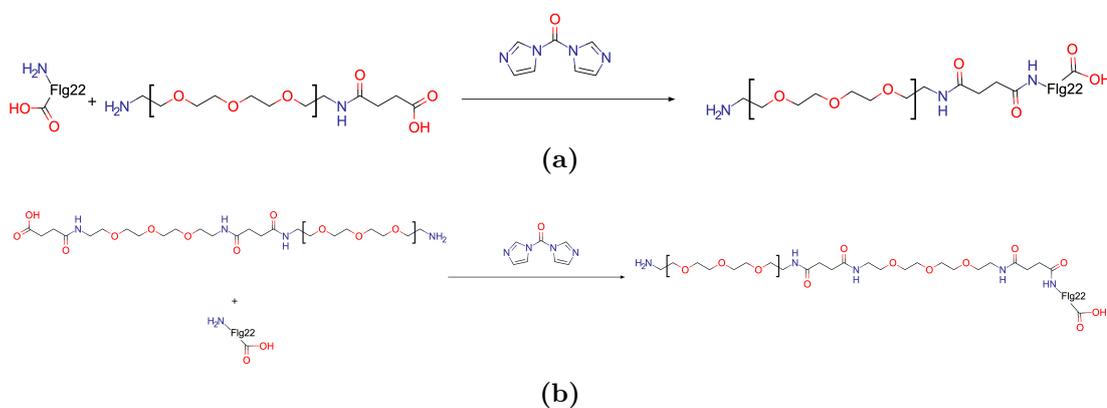
The CuAAC reaction requires an alkyne group and an azide group to react. We were able to choose which group to attach to the nanoparticle, since the reaction for either is equivalently simple. Since carbon-dots are likely to be the reagent in excess, and azides are not well-known for their biological compatibility, we choose to functionalise the C-dots with the alkyne group: thus, any unreacted linker groups on the nanoparticle surface are unlikely to interfere with its biocompatibility. This is performed via an amide bond formation with 4-pentynoic acid (see figure 2.6 below). 200 mg of unfunctionalised carbon dots in 2 ml methanol were stirred overnight with 420 mg of 4-pentynoic acid in a small glass vial. The resultant mixture was freeze-dried, and purified via size exclusion under methanol in a 30 cm silica column to remove unreacted reagents. A small-scale test showed the product of the reaction was soluble in chloroform: therefore, chloroform washes as in the above section were not attempted, although such a process may otherwise have helped purify the reaction product. This observation is another example of surface functionalisation altering the physical properties of the nanoparticle, as corroborated by (Swift et al., 2018).



**Figure 2.6.:** Reaction scheme for addition of a short alkyne linker to carbon nanodots. A condensation reaction occurs between the primary amine on the carbon nanodot and the hydroxyl group on the pentynoic acid, forming an amide bond between the two reactants, and resulting in a carbon nanodot presenting an alkyne functional group at the surface.

### 2.2.3. Attaching chemical payloads via amide linkers

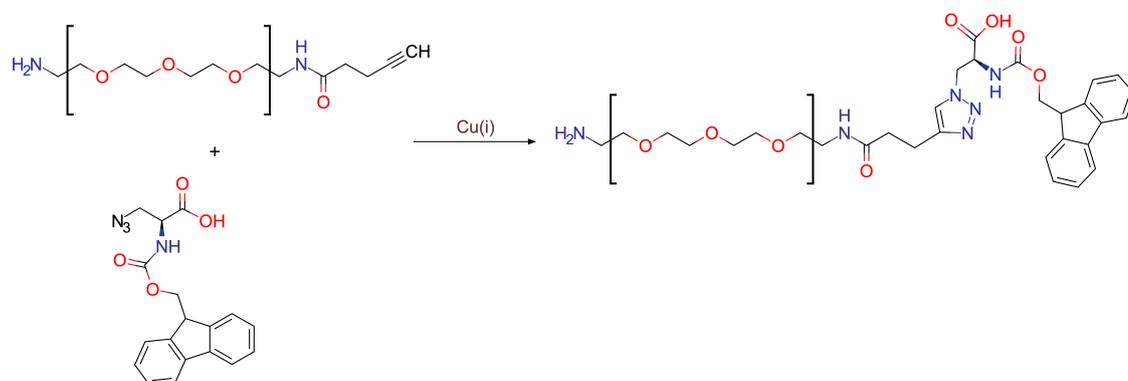
In a similar manner to amide coupling reactions in previous sections, 125 mg of flg-22 peptide (GenScript, HPLC 80% purity) was added under methanol to 100 mg of carbon dots pre-activated with 50 mg of CDI, functionalised with either a short acid linker (figure 2.7a) or a long acid linker (figure 2.7b). The reaction was stirred at room temperature overnight in a 10ml glass vial, and the results were purified via size exclusion using a 10 ml, 6000 MWCO dextran desalting column (Thermo Scientific).



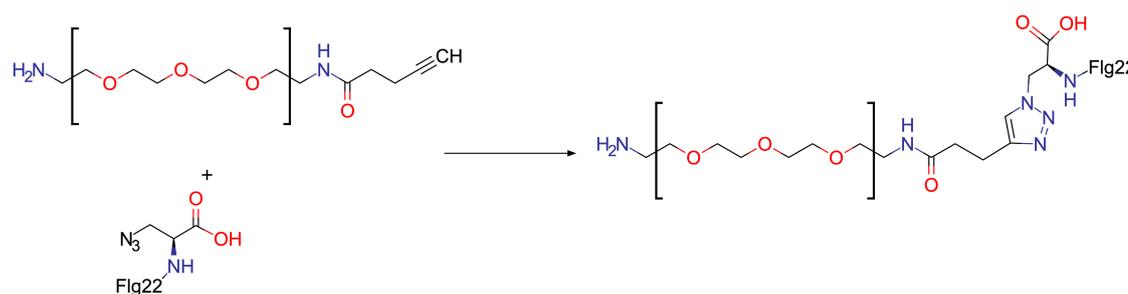
**Figure 2.7.:** Reaction schemes for functionalizing carbon nanodots with the amino acid form of the flg-22 oligopeptide. (a) by direct amide bond formation to acid-functionalized carbon dots (as seen above); and (b) using a longer "linker molecule" i.e. amide bond formation between flg-22 and a carbon nanodot functionalised with a TTDDA-based linker.

### 2.2.4. Testing the CuAAC reaction system for nanodot functionalisation

Since the conditions required for CuAAC reactions can be very specific, a test reaction was initially performed using a “dummy” payload of fmoc- $\beta$ -azido-alanine-OH. A buffer was prepared using 50:1 v/v phosphate-buffered saline (PBS) to dimethyl sulfoxide (DMSO), and to this was added: 0.5 mM copper sulfate, 2.5 mM Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), and 25 mM sodium ascorbate. To this was added first the fmoc- $\beta$ -azido-Alanine-OH to 0.1 mM concentration, then the alkyne-adapted carbon nanodots to a concentration of 0.1 mM. The reaction was warmed slightly to 34°C, using a heated stirrer plate and an aluminium thermal coupling mass, and stirred overnight. Results were freeze-dried before analysis.



(a) Reaction scheme for the proposed CuAAC (“Click chemistry”) reaction between alkyne-adapted carbon dots and fmoc- $\beta$ -azido-Alanine-OH.



(b) Reaction scheme for the proposed CuAAC (“Click chemistry”) reaction between alkyne-adapted carbon dots and an N-terminal azide-modified flg-22 oligopeptide.

**Figure 2.8.:** Reaction schemes for proposed copper(I)-mediated azide-alkyne cycloaddition (CuAAC) reactions with alkyne-adapted nanodots. Each reaction involves the alkyne group on the nanodots reacting with the azide on the chemical payload to form a triazole compound coupling the two reactants. This reaction is catalysed by a copper (I) species which is relatively unstable compared to copper (II) in aqueous solutions: hence, the conditions required are quite specific.

## 2.3. Results

We formed carbon nanodots, functionalised them with a short-acid and long-acid linker as above, and performed amide couplings with flg-22 using both linkers. Steady state fluorescence and absorption spectroscopy was conducted to measure the electronic properties of the functionalised carbon dots; this was carried out to confirm fluorescence parameters to be used for microscopy, and to compare with other data from carbon dots produced by the same methods (Swift et al., 2018). Nuclear Magnetic Resonance (NMR) spectroscopy was also conducted on all samples. NMR is a spectroscopy technique which involves aligning permanent nuclear dipoles using a strong magnetic field; these dipoles are then “flipped” at near-resonance frequency using a comparatively weak oscillating magnetic field, causing a phenomenon known as Larmor precession, in which the nuclei generate a periodic electromagnetic signal. That signal is then measured: its frequency domain is characteristic of the chemical environment of each nucleus, and is typically measured in parts-per-million of frequency shift versus a known standard. So many techniques have now been developed using this method for a variety of applications within analytical chemistry, that more than one journal has been devoted to the subject; (Marion, 2013) has published an extensive review of NMR principles from a biological perspective. Nevertheless, the overall result can be thought of as a “chemical barcode”, which is to say that each unique chemical has a unique NMR spectrum. Particles were characterised using a 500Mhz Bruker nuclear magnetic resonance spectroscope, with deuterium dioxide as the solvent in all cases. Mestrelabs MNova software was used to analyse NMR spectra, and the NMR Predict extension used to help inform predictions of NMR peaks on the basis of chemical structures (Mestrelab, 2020a). The MarvinSketch software package from ChemAxon (2017) was used to draw all chemical structures, with the exception of the flg-22 structural diagram, which was drawn with the Pep-Draw online tool(Freeman, 2015). As previously, MarvinSketch draws nitrogen atoms in blue and oxygen in red.

Analysis of the proton spectra was used to determine the presence or absence of key chemical groups on the nanoparticle. Two-dimensional DOSY analysis can be used to correlate bonded groups with one another, by additionally calculating diffusion coefficients to plot on a perpendicular axis: this was used in order to ascertain whether a particular reaction had been successful, by searching for correlations between known frequency shifts from each reactant.

### **2.3.1. Carbon nanodot core integrity confirmed by fluorescence analysis**

To verify that the reaction in Figure 2.3 on page 38 formed carbon nanodots as expected, the simplest means of testing is to check for the fluorescence properties characteristic of the variety of nanoparticle in question. This was affirmed at every stage of the process, to check that the reactions had not compromised the core property of the nanodots. A very quick check can be performed using a UV lamp, under which intact nanoparticles should appear to emit an aqua-coloured light; for best results, however, fluorescence spectroscopy was employed to check the emission and fluorescence curves of the reaction products. Figure 2.9 on page 45 shows results of this type obtained for the functionalisation with flg-22 using both short and long linker, as described in Figures 2.7a and 2.7b.

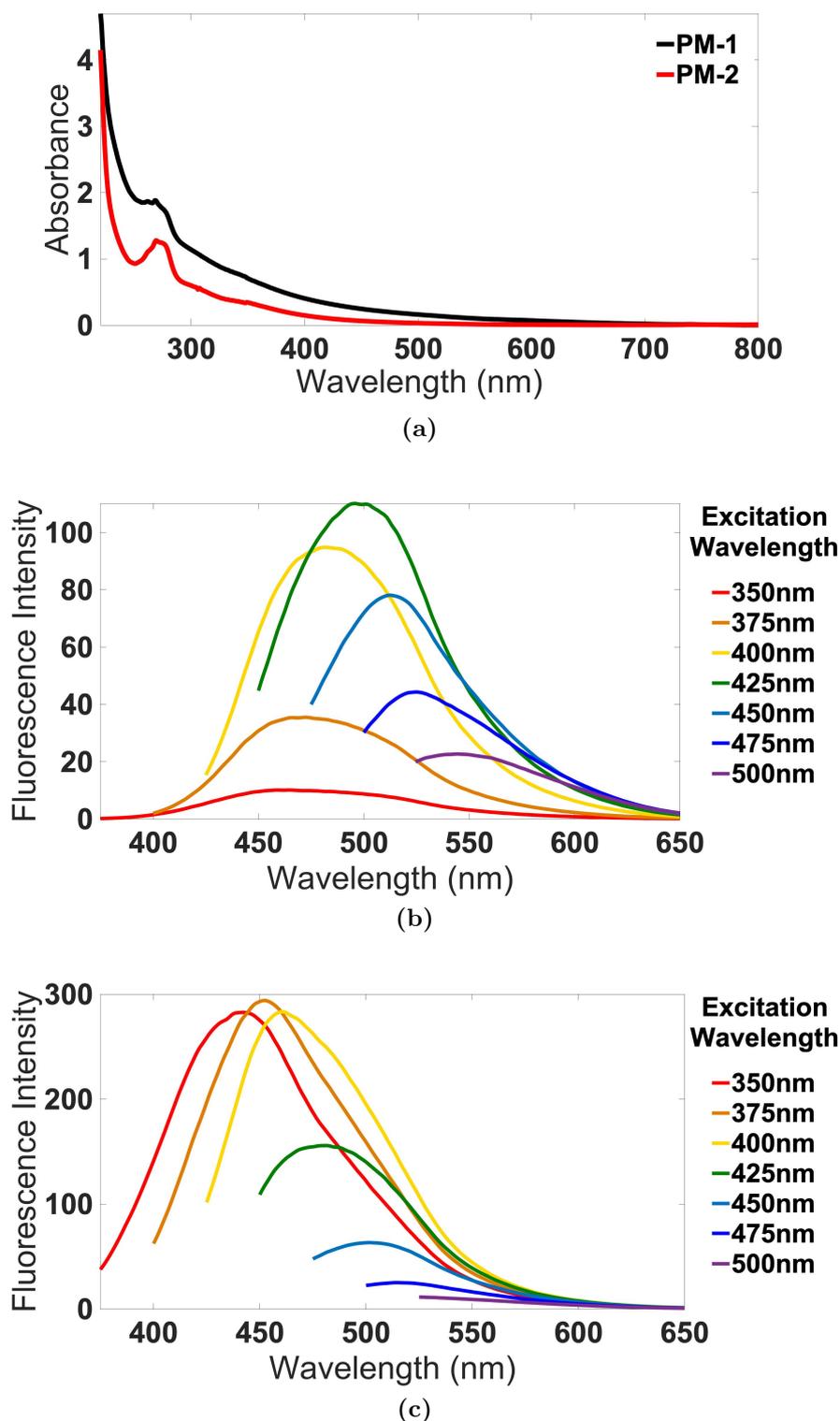
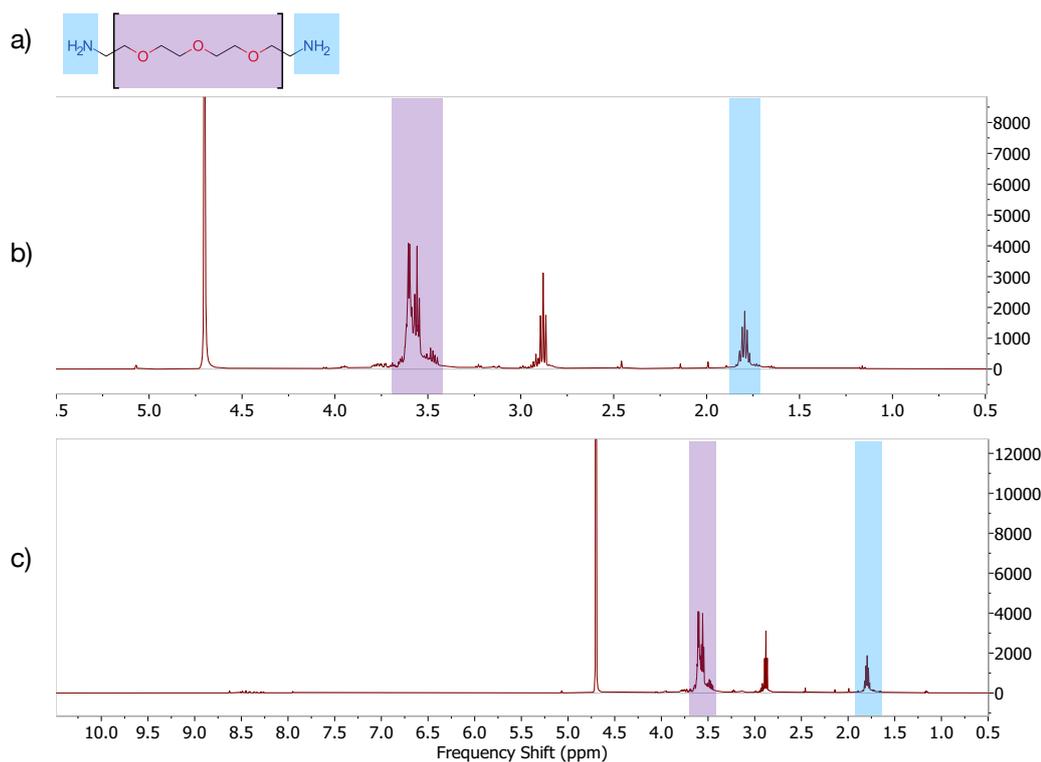


Figure 2.9.: (a) Optical absorbance curves for nanodots functionalised with flg-22 via short acid linker (black line) and flg-22 via long linker (red line) and fluorescence spectroscopy data for (b) nanodots with short acid linker and flg-22 and (c) nanodots with long acid linker and flg-22. For this type of particle, changes in fluorescence depending on surface functionalisation are consistent with other published results (Swift et al., 2018). Note the response to 400 nm excitation, which is consistent between the two types of functionalised particle.

We note the length of linker used affects some of the fluorescence properties of the nanodot, and this is consistent with other studies in this field (Swift et al., 2018; Loukanov et al., 2016). In this case, the particles with the short linker appear to have the higher frequency range of their fluorescence quenched: conversely, particles with the long linker have a lower amplitude of fluorescence in the ranges with longer wavelengths. In both varieties of particle shown (and in other similar functionalisations) the fluorescence behaviour under 400nm excitation appears reliable, and corresponds well with a 405nm diode laser line available on a wide variety of equipment - for this reason, we chose this behaviour as the optical property by which to identify the presence or absence of carbon nanodots in subsequent tests. Alternatively, a krypton laser at 416 nm might well prove ideal; perhaps a helium-cadmium or argon laser line near to 450 nm in wavelength would also be appropriate.

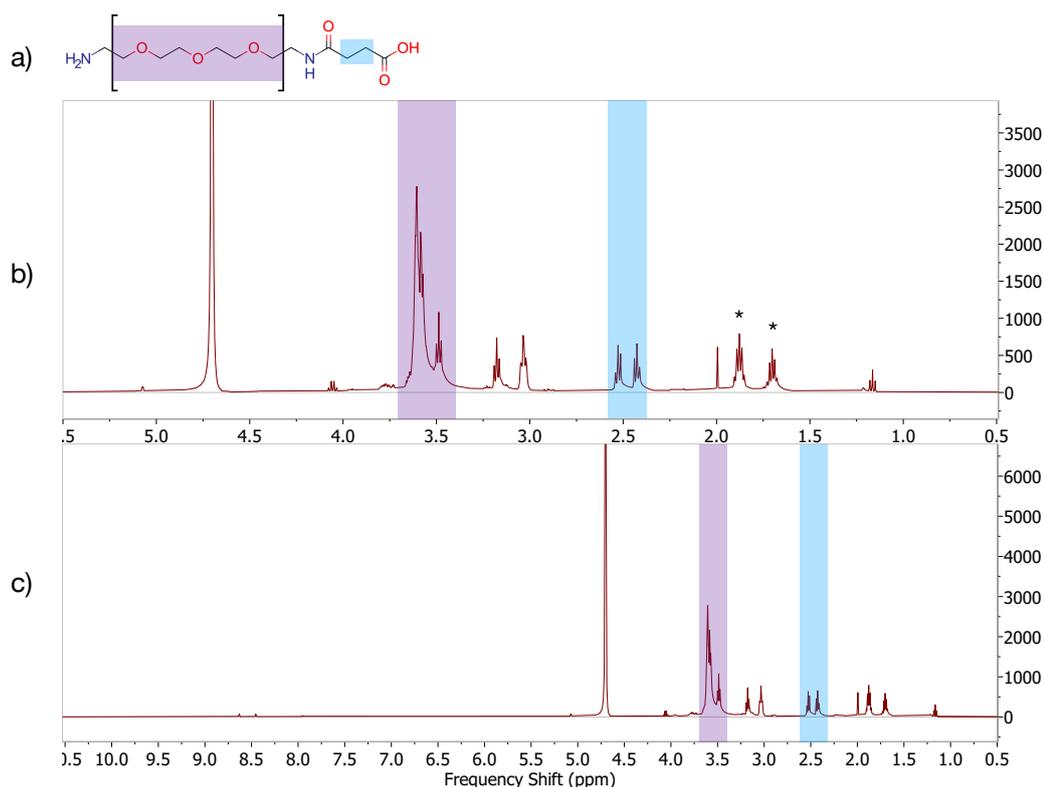
### **2.3.2. NMR analysis confirms previously published carbon nanodot properties.**

After the initial microwave reaction is performed, we are looking for signs of a polymeric core, with some primary amines. The exact degree of core cross-linking is uncertain, and the exact structure of the nanodot is unlikely to be deterministic, since there are no protecting groups on the original carbon chain. Therefore, as the signal from lots of similar individual structures overlap, the resulting spectrum of the sample is likely to have at least one region with clear signal but lacking distinct clearly-defined peaks: this is seen in figure 2.10(b-c) in the shaded region near 3.5 ppm. We also note a multiplet at 1.8 ppm, which in this case is most likely to originate from a distal ethylamine group as highlighted in figure 2.10a).



**Figure 2.10.:** Analytical identification of carbon nanodot cores. (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.3 with characteristic regions of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. The large peak near 4.7 ppm corresponds to the solvent (deuterium oxide) used to prepare the sample for NMR. Here are visible peaks around 3.5 ppm from protons in the nanodot core, whose putative composition is a polymer comprised predominantly of SP3 and SP2 carbon, with a moderate degree of cross-linking occurring between the original SP3-hybrid constituents. Additionally, we note a pentet peak near 1.75 ppm, which corresponds to the ethylamine groups in the expected structure.

Moving on to the short-acid adapter: we expect to see a move and/or split of the multiplet at 1.6 ppm from the previous figure. After performing the reaction shown in figure 2.4, we note that the aforementioned peak has indeed been affected. Additionally, the proton spectrum in 2.11 shows two opposed triplets at 2.5 ppm: these opposed triplets correlate very well with the two carbon atoms in the linker molecule when bonded. Moreover, the fact that they are distinct from one another suggests asymmetry in their chemical context, which supports the predicted structure in 2.11a.



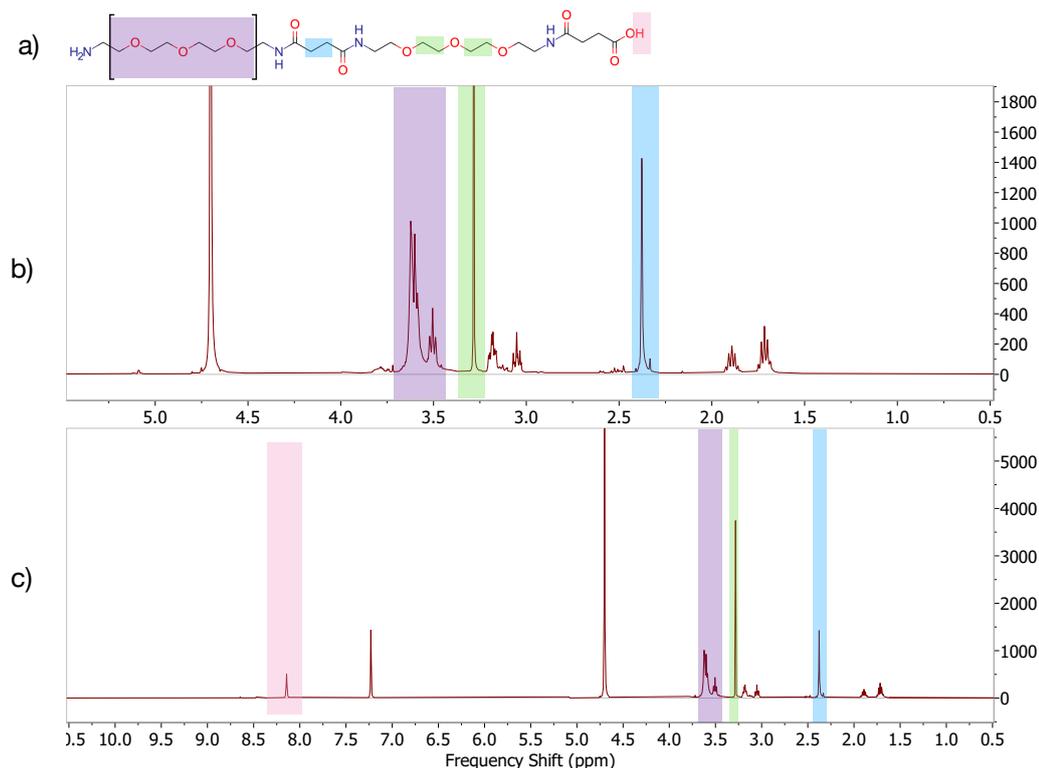
**Figure 2.11.: Analytical confirmation of the presence of the short acid linker.** (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.4, with characteristic regions of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. There are two pentat peaks annotated in (b) with an asterisk (\*). Either (or both) of these could correspond to the amine group seen in the previous figure, suggesting that some free amine groups remain after the reaction: the splitting suggests some change in the chemical environment, for example one could infer that some (and not all) of the amines may be close enough to the added linker for spin-spin coupling.

Overall, these results correspond well with those previously published by Swift et al. (2018): no less so than expected, since I performed this work using the same facilities and very similar equipment. Nevertheless, these findings form a reassuring starting point from which to progress to reactions which have never been attempted before.

### 2.3.3. NMR analysis confirms novel adaption of carbon dots with a long acid linker molecule

In order to ascertain that a reaction has taken place, we obtain a proton NMR spectrum of the reaction product, shown in figure 2.12. In the figure, we note a

single tall peak at 2.35 ppm in place of the opposed triplets seen in the previous figure. This is a very important finding, since the collapsing of the triplets into a tall (4H) singlet reflects very well the predicted structure, which would put those protons in symmetric environments (hence merging what we would otherwise expect to be triplet peaks). Continuing the theme of symmetry, a similar-looking peak near 3.3 ppm is also notable, and corresponds to CH<sub>2</sub> groups in the mid-linker region highlighted in 2.12a. Moreover, the peak's height is roughly double that of the singlet at 2.35 ppm, thus corresponding to 8H in size, which is as predicted by an intact linker. Finally, the single peak seen at (approx.) 8.15 ppm in fig. 2.12c correlates well with the distal hydroxyl group in the linker: this is a good indication that the linker structure has not been altered by the reaction.

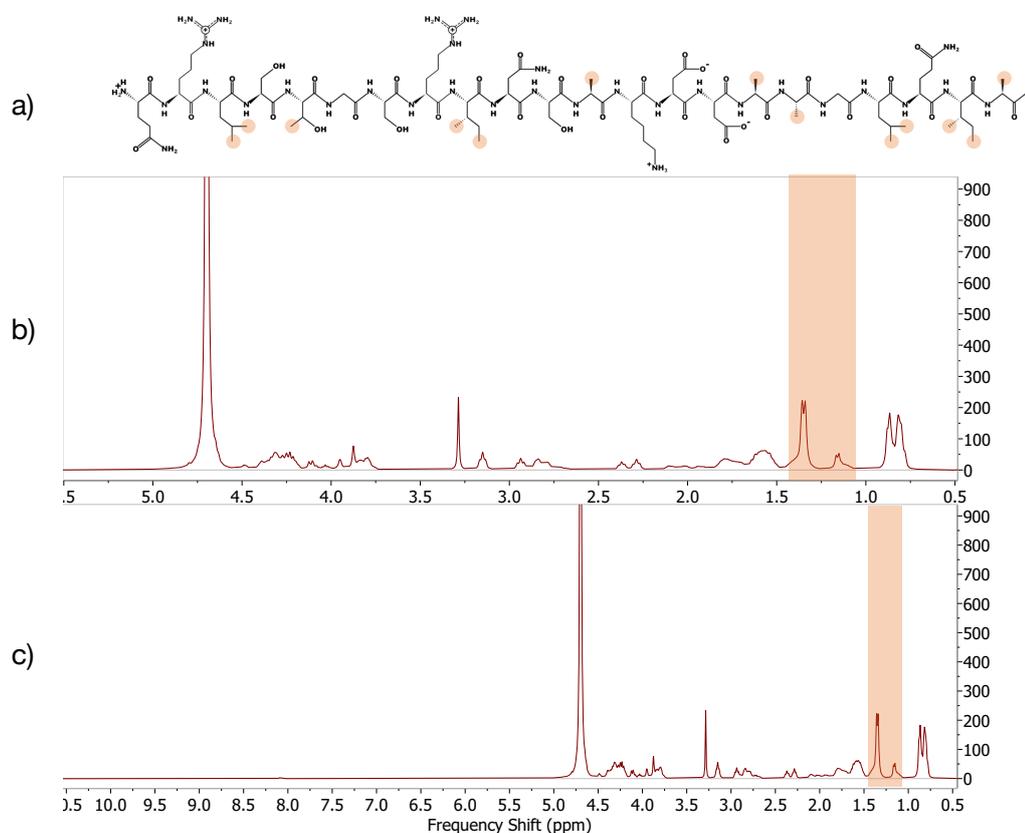


**Figure 2.12.: Analytical confirmation of the formation of the product in figure 2.5.** (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.5 with characteristic regions of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. Of note here is the shaded region near 2.35 ppm, where two formerly distinct triplet peaks have merged compared to the previous figure. As a result of the reaction, the environment of the corresponding protons has become more symmetrical: this causes their resonance peaks to become more similar, so much so that they become superimposed at this resolution. A slight skew remains visible, though it is possible for the required post-processing (phase correction) to produce a similar effect. The peak at 3.3 ppm correlates well with the four superimposed singlets from the CH<sub>2</sub> groups highlighted in the centre of the linker chain, and as expected integrates close to double the size of the peak at 2.35 ppm.

All in all, this data suggests that the amide coupling between the nanodot and the longer linker molecule was successful, without compromising the structure of the core nanoparticle or the linker molecule, and yielding a structure as predicted. While this may seem a trivial progression from the previous result, it is important to note the possibility that the linker would homo-polymerise in this reaction, or that the reaction would otherwise not progress as expected: due to a side-reaction with CDI and the linker reagent, for example.

### 2.3.4. NMR spectroscopy of flg-22 to functionalise carbon nanodots

Because NMR spectra become more complex as the analyte grows in molecular complexity, because flg-22 is relatively complex compared to the kind of synthetic analyte that is most easily interpreted by NMR spectroscopy, and because the proton spectrum of pure flg-22 in deuterium oxide is not something that is readily available in literature, we elected to obtain our own NMR spectrum from our reagent: a commercially-obtained flg-22 with an HPLC purity measured at 80%. The resulting spectrum is depicted below in 2.13, along with the structure of flg-22 for reference.



**Figure 2.13.:** NMR identification of the flg-22 synthetic oligopeptide.

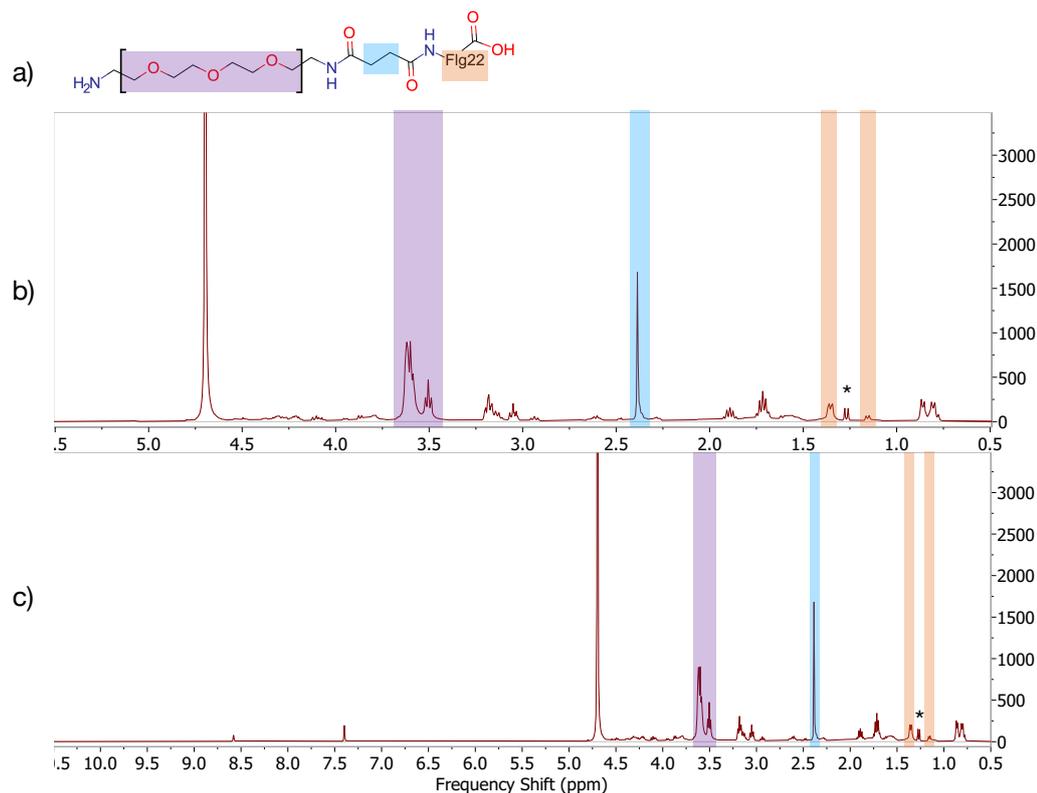
(a) Structural diagram and (b-c) Proton NMR spectrum of the flg-22 oligopeptide, with characteristic region of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. NMR data was obtained with a 200MHz Bruker spectrometer, and analyzed using Mestrelabs MNova software. Flg-22 structure diagram was drawn using the PepDraw online tool.

Rather than attempt to assign every peak in the spectrum, we used the pair of duplets between 1.0 and 1.5 ppm, highlighted in orange in figure 2.13, as a fea-

ture characteristic of the flg-22 payload. This characteristic feature was chosen as a means of identifying the molecule in subsequent NMR spectra since it was a distinctive pattern at a resonance reasonably distant from those formerly occupied by unfunctionalised nanodots, and because its magnitude made it relatively easy to identify. These doublets are associated with methyl groups such as those occurring in the structure of flg-22, which are similarly highlighted in the inset structural diagram: typically, doublets with chemical shifts between 1 and 2 indicate methyl groups bonded to another carbon. The doublet pattern would tend to indicate  $\text{CH}_3\text{CH}$  (rather than  $\text{CH}_3\text{CH}_2$ ) by virtue of the  $n+1$  rule, though this may be split by other factors. For example, some  $\text{CH}_3\text{CH}$  groups in the molecule are closer to peptide bonds than others: the highly electronegative nitrogen atom in those bonds affects the frequency shift of protons near them.

### **2.3.5. Proton NMR spectroscopy to analyse the bonding of flg-22 to adapted carbon nanodots**

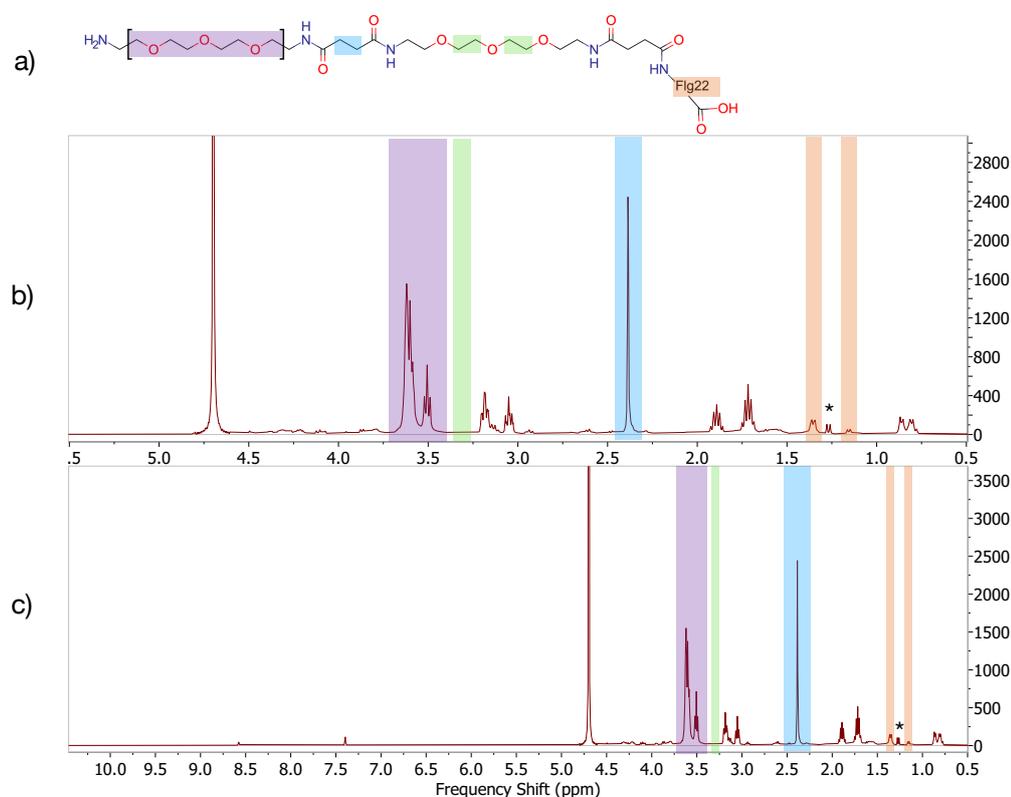
Having identified a characteristic feature of the proton NMR spectrum of flg-22, we have a means of locating it in the spectrum of the products formed from the reaction with the adapted carbon dots. Unlike the previous reaction, these are attempting to couple a relatively complex molecule, which incorporates many of its own peptide bonds: therefore, on the basis of the proton spectrum alone, it would be difficult to state with confidence whether or not the reaction has taken place. Nevertheless, we examine the proton spectrum alone first, before correlating it to the diffusion constant in a two-dimensional analysis, and beginning with the shorter of the two linker adaptations in figure 2.14.



**Figure 2.14.: Analytical confirmation of a reaction between short-linker nanodots and flg-22.** (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.7a, with characteristic regions of the spectrum (and corresponding structural groups) shaded in colour. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. Between the two doublet peaks identified as coming from flg-22, we note the formation of an additional doublet, denoted with an asterisk (\*). It is difficult to assign this peak with confidence, since there are now a very large number of protons in the molecule, but its presence suggests a change near one of the methyl groups previously highlighted in the flg-22 expanded structure.

In the above spectrum, we can see that the nanodots and the flg-22 are present in the same sample: this is encouraging, given that this data follows size-exclusion purification. However, in this case the proton spectrum alone does not provide strong enough support for our hypothesis: although much of the spectrum aligns with our expectations, some parts of the data are not as expected. In particular, the appearance of an additional doublet (or the splitting of an existing doublet) in the spectrum compared with our previous data for flg-22 suggest some degree of chemical change near one or more of the ethyl groups identified: for example, flg-22 could have formed a peptide bond with the particle at a location near these, rather than at the N-terminal end of the molecule. Therefore, it is decided to collect more data from this sample, using correlative spectroscopy to offer more information that may help elucidate the result: this data is seen later in section

2.3.6. In the meantime, we examine the proton spectrum from the alternative reaction using the longer linker, as seen in figure 2.7b.



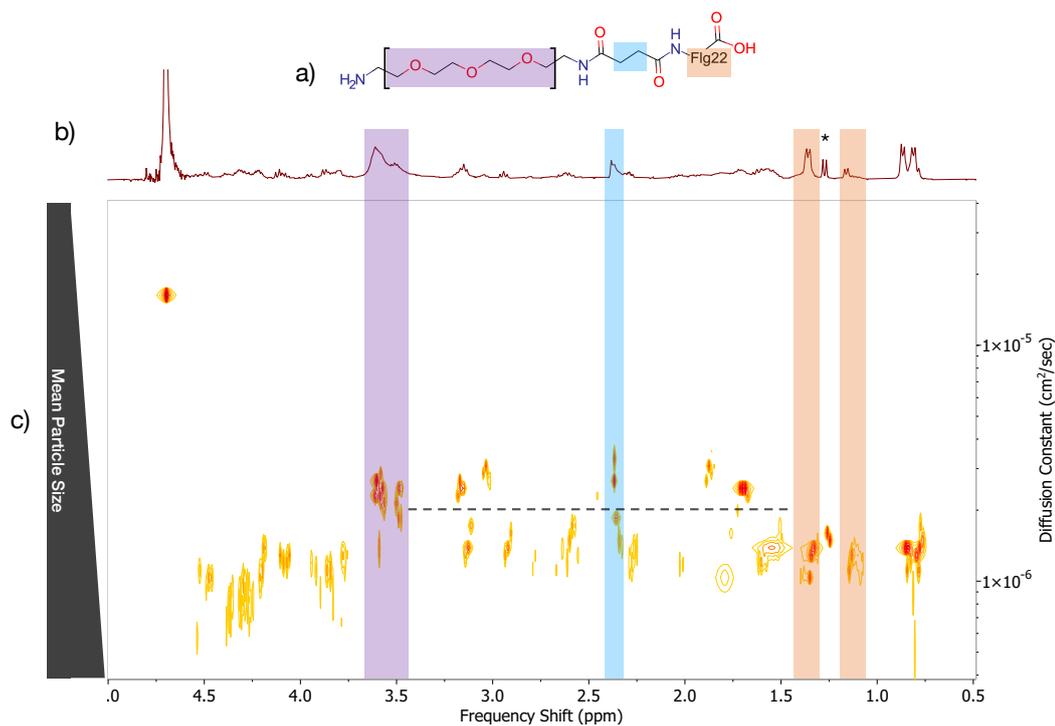
**Figure 2.15.: Confirmation of one or more structural changes in the reaction between long-linker nanodots and flg-22.** (a) Expected structure and (b-c) Annotated proton NMR spectrum of the product of the reaction in fig. 2.7b (carbon nanodots functionalised with flg-22 via long linker). Shaded areas indicate the depicted structural elements, and regions of the NMR spectrum where one or more corresponding peaks are expected. (b) is a detail view of (c) for resonances between 0.5 and 5.5 ppm. For a structure as depicted, we would expect to see a single peak in the green shaded area (around 3.3 ppm) as seen in prior spectra. The absence of this peak suggests that either the structural context of the corresponding protons has changed compared to the linker-only particle, or they are no longer present in the sample, having been displaced by other constituents. It is possible that flg-22 has hydrogen-bonded to the linker in such a manner as to shift of these protons downfield (to the left), causing them to overlap with other signals; alternatively, the reaction may have caused individual linker molecules on the surface of the nanoparticle to react with one another, rather than with the flg-22. An additional doublet peak, similar to that seen in the previous figure, is denoted with an asterisk (\*): overall, the proton spectrum is strikingly similar to that seen in the short-linker reaction, emphasizing the motivation towards correlative spectroscopy for greater resolving power.

Given that the spectra from figures 2.15 and 2.14 come from different reactions,

the results in NMR are disturbingly similar. Most disturbing is the apparent disappearance of the peak at 3.3 ppm: one would expect the singlet perhaps to split in response to a change in spin-spin coupling permutations, or even possibly to shift into a region overlapping with another peak. Considering the large size of the peak in figure 2.12c, however, there are not many places for it to be 'hidden' by other similarly large peaks: there is a (slim) chance that hydrogen bonding could shift these protons downfield (to the left) just enough to be covered by the solvent peak at 4.7 ppm. An alternative is that the longer linker molecule has been displaced by something else - the most obvious possibility being the flg-22 oligopeptide itself. If true, this suggests that the bond in this reaction is stronger than that between the nanodot and the long linker molecule: a relatively good sign for the formation of a payload-bearing nanodot product, but relatively bad news if we are relying on the linker to provide steric isolation between nanodot and payload. Using diffusion-ordered NMR spectroscopy (DOSY), we should be able to see a clear distinction between the solvent peak and anything else in the tube: since the solvent (deuterium oxide) is a much smaller molecule than anything else in our sample, and will hence have a much larger diffusion coefficient.

### **2.3.6. Two-dimensional spectroscopy confirms flg-22 and adapted carbon nanodots diffuse together**

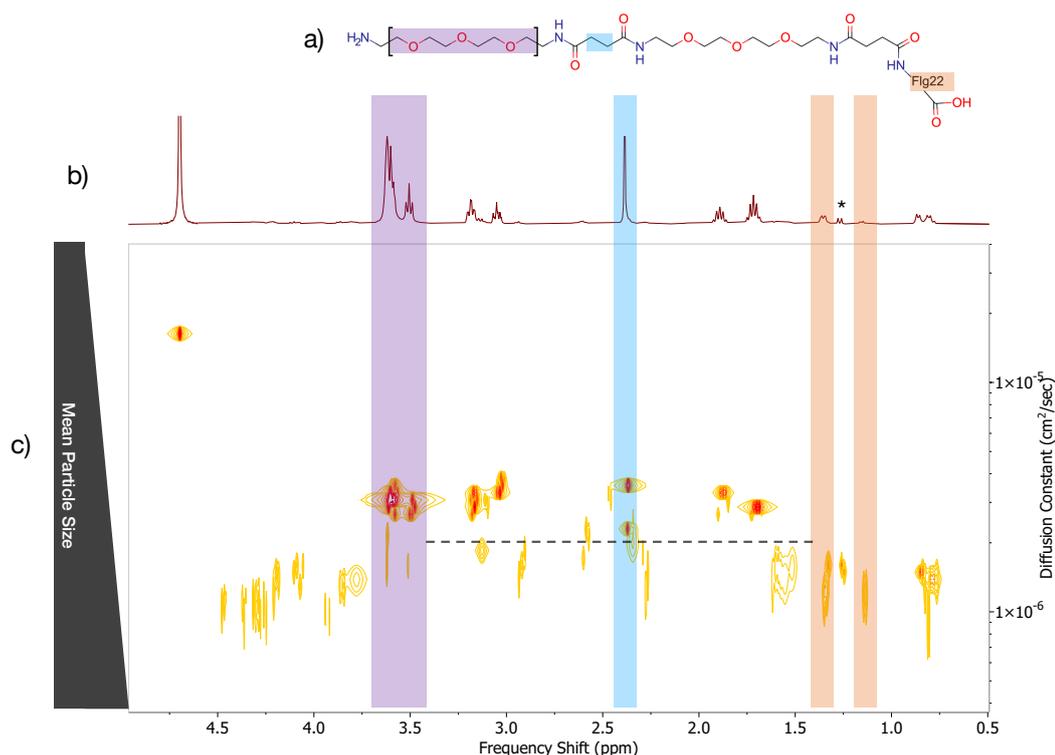
For these experiments, diffusion-ordered spectroscopy (DOSY) was used: this allows an additional dimension of resolution in the data, by correlating resonance peaks with diffusion constant. Since diffusion constant is inversely related to molecular size and weight, two peaks with equitable diffusion constants are likely to belong to the same molecule, or at least are bonded to each other tightly enough that they diffuse together. Since the results occur in a three dimensional format (of intensity versus resonance versus diffusion constant), the data is often shown as a contour plot, with a corresponding two-dimensional projection of the NMR trace placed above it on the same scale, such that each peak in the contour plot has a corresponding peak in the proton spectrum above it. In figures 2.16 and 2.17, a yellow-red scale is used for positive contours in the contour plot (c); additionally, shaded area annotations are continuous between the contour plot and the proton spectrum (b), and a reminder of the relationship between diffusion constant and particle size is added to the left-hand side of the contour plot (c).



**Figure 2.16.:** Two-dimensional analysis of the product shown in figure 2.7a. (a) Expected structure, (b) annotated proton spectrum and (c) corresponding DOSY spectrum of the product of the reaction in figure 2.7a. X-axis shows frequency shift in parts per million; Y-axis shows diffusion constant in square centimetres per second. Z-axis (intensity) is represented by contours on a yellow-red scale, with darker colours representing larger values. Shaded areas indicate correspondence between the depicted structural elements and regions of the proton NMR spectrum. Dashed line is a horizontal guide for the eye: peaks on the same horizontal level in the lower contour plot have the same diffusion constant. Some peaks from the carbon nanodot are at similar horizontal levels to those from fig-22, indicating the latter may be covalently bonded to some quantity of the nanodot. Particularly, the lowest peak in the shaded area near 3.6 ppm is not seen in the spectrum of fig-22 alone, yet its estimated diffusion coefficient is very close to that of the fig-22 peaks.

In figure 2.16, we can see that the position on the y-axis of the two doublets which are associated with our carbon dots, correlate well with the corresponding position of the duplets identified earlier as characteristic of the fig-22 payload. Conversely, in figure 2.17 one can see that the same fig-22 peaks are somewhat further apart on the horizontal from the bulk of the carbon nanodot signal: moreover, the particle which should have been of a larger overall size, had core resonance signals at a higher diffusion constant, which would tend to indicate a particle size that was smaller overall. From this we conclude that at least some of the fig-22 may be bonded to the nanodots with the short linker; however, although they diffuse at similar rates, the long linker nanodot was unlikely to have performed any

covalent bonding with the majority of the desired flg-22 payload. Nevertheless, it is still possible that some chemical coupling has taken place, perhaps through an interaction other than that predicted. There are possibilities such as hydrogen bonding, which might not be apparent in an NMR analysis. At least, the two reactants appear to diffuse together: this could be ascertained by analysis of each size-exclusion fraction obtained from the original reaction: an analysis which is conducted later in section 2.3.7.



**Figure 2.17.: Diffusion analysis of the product of reaction 2.7b.** (a) Expected structure, (b) annotated proton spectrum and (c) corresponding DOSY spectrum of the product of the reaction in figure 2.7b. Shaded areas indicate correspondence between the depicted structural elements and regions of the proton NMR spectrum. Dashed line is a horizontal guide for the eye, at the same diffusion constant as the previous figure. Here the peaks identified as characteristic of flg-22 have less crossover on the horizontal level with those identifying carbon nanodot structures, when compared to the previous figure. This suggests that the two reactants are less likely to be bonded covalently as shown in the above structure.

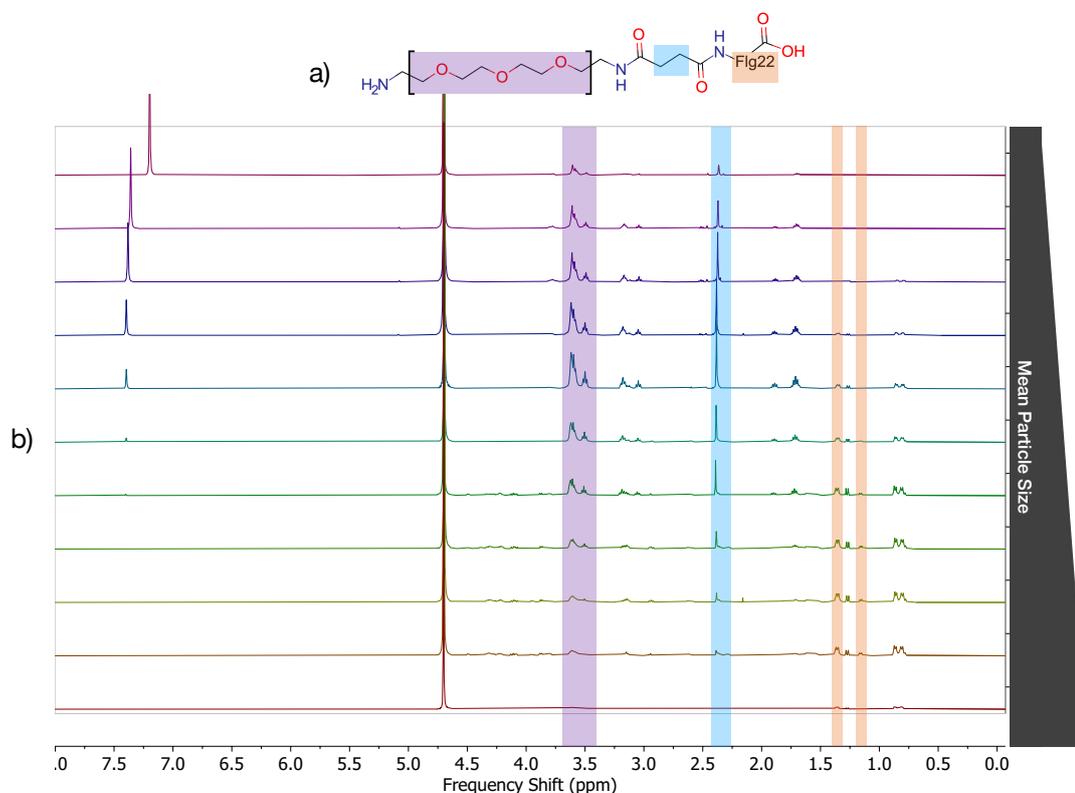
As previously in figure 2.15, we note the absence of a large (8H) expected peak at or near 3.3 ppm on the proton spectrum, which would be produced by an intact linker molecule. Using DOSY, we can establish that this peak has not shifted to a position that overlaps with the solvent, since there is only one peak in the contour plot at a corresponding frequency shift. Also of note are the two

vertically-aligned peaks in the blue-shaded region around 2.35 ppm: the fact that they are present at all would indicate that the correspondingly shaded structural element has remained coupled to the nanodot, rather than being washed out during size exclusion purification. Two vertically aligned peaks suggest two sets of near-identical structures, each belonging to molecules of a different overall size: this might support a hypothesis that the long linker molecule has been displaced by the flg-22 oligomer, and that the final structure may be much closer to that of the short-linker nanodot seen in the prior figure. The relative proximity of the lower peak to other peaks corresponding to flg-22 might also suggest that some portion of the groups are correlating with flg-22-loaded nanodots; alternatively, the lower peak might be caused by the distal portion of the long linker (a portion which is identical to the short linker) coupling to flg-22 having been decoupled from the nanodot. To summarise: the structure of the end product is unclear, and it is likely that at least some portion of the longer linker molecule has become decoupled from the nanodot and lost; more investigation will be necessary to ascertain whether the end product will remain useful in the context of a biological experiment.

### **2.3.7. H-NMR analysis of size-exclusion chromatography fractions confirms significant crossover between nanodots and their oligopeptide payload post-reaction.**

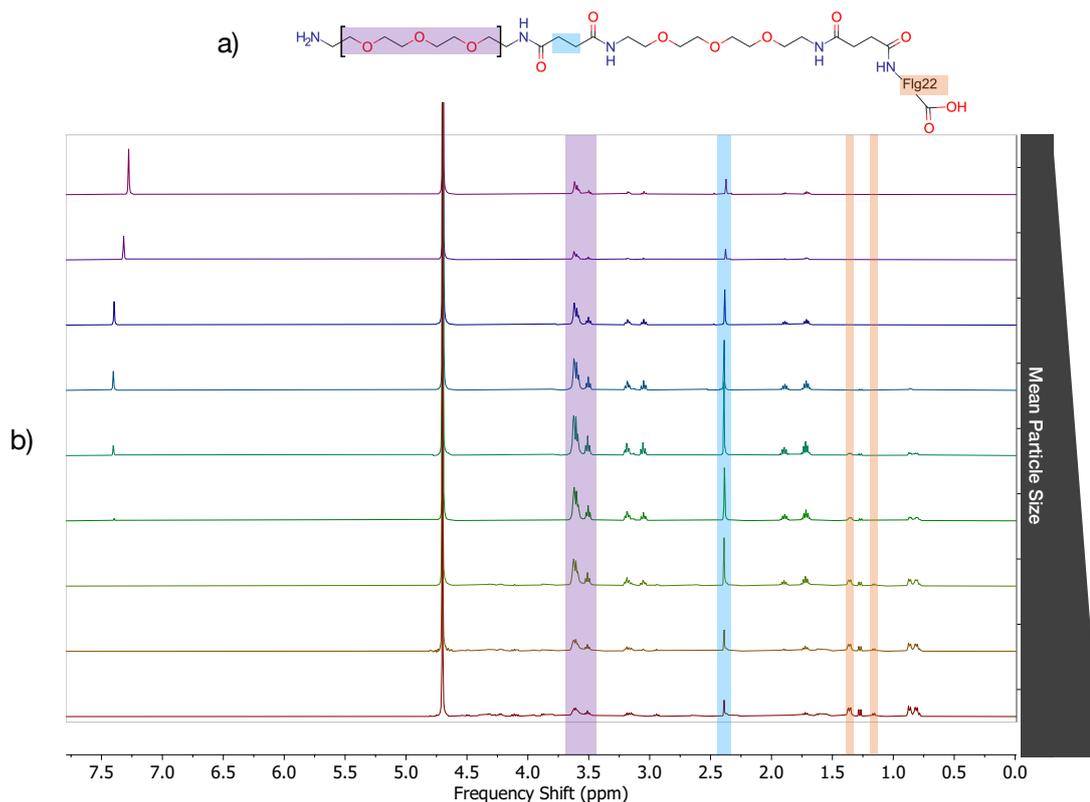
Size-exclusion chromatography fractions were obtained from each reaction product from figures 2.7a and 2.7b, and proton NMR spectra obtained from the fractions, with the aim of removing the CDI activator from the final product. There is a characteristic peak for CDI between 7 and 7.5 ppm, seen on the left-hand side of both figures, which shows the motivation for this step in the process. In figure 2.18 there is a visible correlation between the heights of the peaks corresponding to the nanodot core, and those of the short-acid linker. Moreover, there is clear crossover between these peaks and those identified as characteristic of the flg-22 payload, which indicates that at least some fraction of the two are diffusing at consistent rates together; however, clearly there are some nanodots which are diffusing before any fraction of flg-22, and a significant number of fractions in which both appear together. Since the reaction was constructed with nanodots as the reagent in stoichiometric excess, a remainder of un-functionalised nanodots is to be expected. Ideally, a greater distinction could be achieved in the size-exclusion process between payload-bearing nanodots, and those with linker alone: indeed, on the basis of molecular weight alone, the payload ought to alter the character-

istics of the nanodot significantly. With that said, the dynamics of size-exclusion are non-linear, i.e. a doubling of molecular weight does not imply a doubling of diffusion time.



**Figure 2.18.: Fractionation by size exclusion chromatography confirms co-diffusion of short-acid nanodots with flg-22.** (a) Expected structure and (b) annotated stacked proton spectra of 0.5 ml size exclusion chromatography fractions of the product of the reaction in fig. 2.7a: each trace is a separate fraction of the reaction product. Shaded areas indicate correspondence between the depicted structural elements and regions of the proton NMR spectrum. The peak near 7.3 ppm corresponds to the activator chemical CDI, which it is desirable to remove from the system: emphasizing the requirement for size-exclusion purification post-reaction.

Overall, by comparing between fractions the peaks for the flg-22 payload and the nanodots, the above figure suggests a one-way correlation, such that the presence of flg-22 in any fraction indicates the presence of nanodots in that fraction, though the reverse is not always true. Though not an ideal outcome, we can nevertheless proceed with caution into a useful biological experiment, bearing in mind that the majority of the nanodots are diffusing together with the flg-22 payload, and given that a size exclusion fraction is carefully selected to maximise the proportion of identical relationship between payload and nanodot.

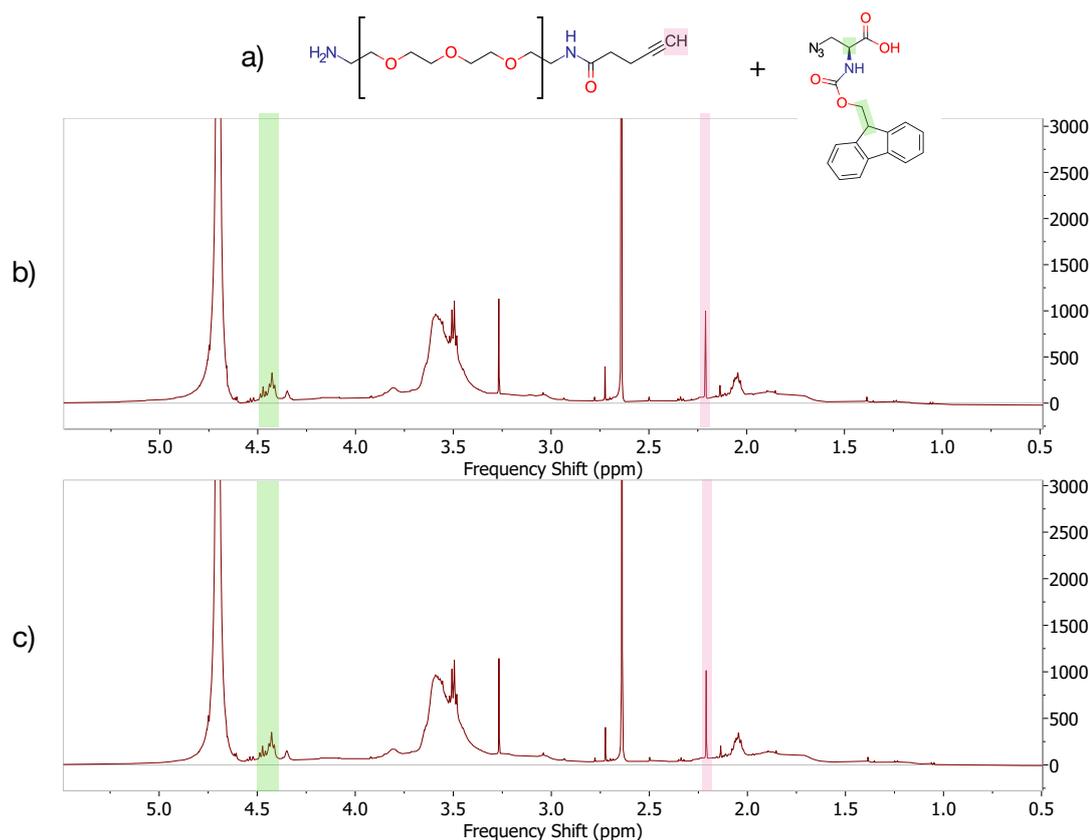


**Figure 2.19.: Size-exclusion chromatography indicates diffusive similarity between long-linker nanodots and flg-22 after reaction.** (a) Expected structure and (b) annotated stacked proton spectra of size exclusion chromatography fractions of the product of the reaction in fig. 2.7b Shaded areas indicate correspondence between the depicted structural elements and regions of the proton NMR spectrum. Here the crossover between the flg-22 peptide, carbon nanodots, linker, and the activator chemical CDI can be seen. As indicated previously, a missing (possibly shifted) peak at 3.3 ppm calls into question that the structure is in fact as illustrated.

In comparing figure 2.19 to the previous figure, we see a spread of the nanodot-associated peaks across a larger range of size-exclusion fractions: perhaps an indication of a greater variation in stoichiometric ratio between nanodots and their flg-22 payload. Also of note is the larger fraction of nanodot-associated signal which does not correspond to a similar value for flg-22: this suggests a weaker relationship overall between nanodot and oligopeptide. Overall, there is less confidence in the bonding effect, even though the data indicates that the long-acid linker section has been displaced.

### **2.3.8. Assessing the outcome of CuAAC reaction attempts using proton NMR**

Finally, the outcome of the CuAAC reactions was tested by NMR - there should be a distinct change in the proton spectrum between the free azide and amine groups on the reactants, and the cyclic group of the proposed product: that is to say, there ought to be a distinct visible change when comparing the spectrum before the reaction (Figure 2.20b) with the spectrum after the reaction (Figure 2.20c). In all attempts at a range of buffered pH values, and using more than one alternate buffer formulation, the result was the same as seen in figure 2.20: the spectra for the carbon nanodots with the alkyne linker molecule, purified by size exclusion before and after the attempted reaction were practically identical. Though unfortunate, this is a clear indication that no measurable reaction took place in this case.



**Figure 2.20.: Analysis confirms the failure of our experimental conditions to effect azido-alkyne addition.** (a) Expected structures and annotated proton NMR spectra of the reactants in fig. 2.8a (b) before and (c) after attempting the CuAAC reaction. The fact that there has been no change in the spectrum, particularly around the distinctive peak at 2.2 ppm corresponding to the distal end of the alkyne, suggests that no reaction has taken place. This was attempted more than once, using buffers in three different pH levels, with similar results. Of additional note is the region around 3.6 ppm, which has become crowded with the addition of more organic chemical groups: this is one weakness of proton-only NMR when analysing larger molecules, and the original issue which led to the invention of two-dimensional correlative spectroscopy techniques such as diffusion-ordered spectroscopy (DOSY), seen above. Similarity of the proton spectra before and after reaction suggest that no reaction has taken place: most probably the (very specific) conditions required for this reaction were not met in this case, namely the formation of a Cu(I) ion complex in an aqueous environment.

## 2.4. Conclusions

The results shown in figure 2.9 suggest a 405nm excitation frequency with fluorescence imagery collected in a band-pass filter around 450 nm will be suitable for further study. This is consistent with similar published data (Swift et al., 2018).

According to the NMR data, almost all reactions were successful: the functionalisation of carbon nanodot cores with short- and long-acid linkers, and with 4-pentynoic acid, as well as evidence to support the covalent amide coupling of flg-22 to carbon dots functionalised with a short-acid linker. Longer linkers have been tried before using polyethylene glycol (Doyle et al., 2019); however, to the author's knowledge this is the first example of flg-22 being bonded to carbon nanoparticles: a reaction which appears to be favourable enough to have displaced part of the linker of the nanoparticle adapted with the longer linker. This is an important result, not only in its analytical verification of the materials attained, but also in demonstrating the versatility of the amide coupling reactions: their utility remaining even without extensive optimisation of the reaction conditions or stoichiometry. Conversely, these results represent a caution against the use of longer self-similar linker molecules, which although usefully highly identifiable in NMR may well collapse or be displaced in subsequent reactions.

The exception to the general trend of qualified success was the CuAAC reaction: despite more than one attempt, altering reaction temperature, and using phosphate-saline buffers in a range of pH, we were not able to activate the reaction: most likely because the reaction would not remain within the stability range of the copper(I) catalyst's oxidation state, which is very narrow in aqueous conditions. Additionally, the characteristic peaks of nanodot and payload in figure 2.16 are not on a horizontal level with one another: this suggests that the two reactants are probably not covalently bonded, although they have similar enough diffusion coefficients to suggest they may be hydrogen-bonded, which is enough to warrant further study: it is possible that in this particular case, the tenacious nature of the amide coupling reaction could have been counter-productive, for example by forming oligomeric combinations of the linker molecule. Without the benefit of site-specificity that a CuAAC reaction provides, amide bond formation with flg-22 may occur with primary amines on the molecular side-chains, rather than the N-terminus. It may be seen in further chapters whether this affects the biological activity of the flg-22 once the nanoparticle has been bonded to it.

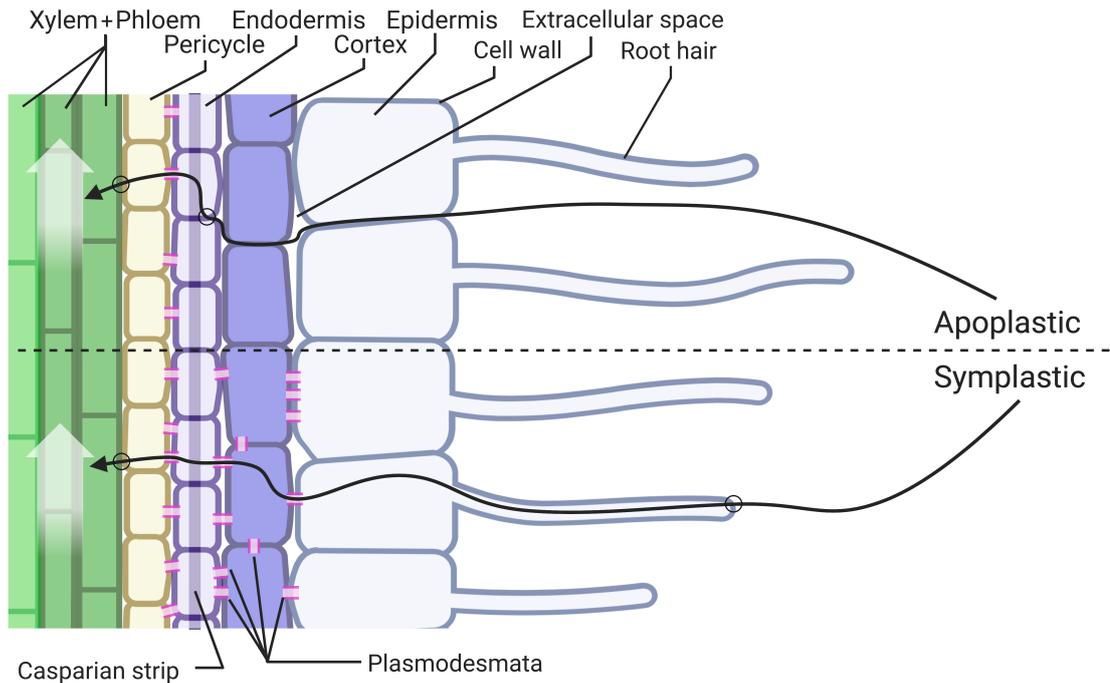
# 3. Carbon Nanodot uptake in *Arabidopsis thaliana*

## 3.1. Introduction

When observing effects in plant roots - whether focusing on development and structure, expression of a particular gene of interest, or even examining more general changes in gene expression patterns - typically data is measured from a whole root, or a single identifiable sub-system or tissue type. While this approach of isolating particular tissue types and comparing to whole-root data is useful in terms of understanding the role of each tissue in a particular interaction, it requires that lots of experiments be performed in parallel, and at worst it can limit tissue-specific understanding to those tissues which can be easily isolated, or for which e.g. stable fluorescent marker lines already exist. Additionally, such an approach does not address the difficulty of ascertaining interactions between cells, which may be highly important when considering apparently co-ordinated responses such as lateral root initiation (Dubrovsky et al., 2000). Moreover, for e.g. ligand-receptor interactions, often the plant is treated with a relatively high concentration of ligand, and then e.g. transcriptional responses are measured on the basis that the treatment affects every change it possibly can in the plant over a particular period of time. While this approach is valid, it assumes that time periods for endogenous levels of ligand are the same as they are for exogenous treatments - and, especially, that the level and timing of co-ordinated responses is uniform enough to be seen at a whole-root level, even after accounting for intrinsic variation in response.

There are methods such as single-cell RNA sequencing which make it possible to analyse the response of each cell in a particular sample, but in an environment where different cell types may be reacting differently to the same treatment, an internal validation issue of the per-cell level of input remains. In this chapter we attempt to provide the first steps to show that carbon nanodots may provide a means of verifying this level, thereby increasing the power of single-cell experiments.

When considering the possibilities for uptake of a material via plant roots, it is important to consider the known paths available. These fall into two archetypes, depicted in Figure 3.1, known as symplastic and apoplastic. It is expected that smaller, more polar molecules will follow an apoplastic uptake pattern, permeating through pores in cell walls and extracellular space until the waxy casparian strip forces a membrane crossing into the endodermal cell file; whereas larger, less-polar molecules will follow a symplastic pattern, crossing the cell membrane at the outermost epidermal layer and being passed from cell to cell by intercellular pores known as plasmodesmata.



**Figure 3.1.: Illustration of uptake pathways in plant roots: apoplastic and symplastic.** Cell types are shown in colours (from right to left): epidermis in light grey, cortex in dark blue, endodermis in lilac, pericycle in yellow, and xylem and phloem in green. Although they are in practice two distinct tissues, the radial positions of xylem and phloem overlap: the outer files (dark green) are occupied by protoxylem, procambium, and the sieve elements and companion cells of the phloem; whereas the inner files (light green) are comprised of metaxylem and more procambium cells. White vertical arrows represent material transfer with other organs in the plant (e.g. leaves). In either pathway, a molecule must undergo two membrane crossings (black circles). In the symplastic route (lower black arrow) the molecule enters the cytoplasm at the epidermis, before being passed from cell to cell by protein pores called plasmodesmata (pink) through to the pericycle level, after which one further membrane crossing must take place in order to access the procambium and vascular cells. In the apoplastic route (upper black arrow) the molecule passes between cells, moving in the interstitial spaces, and may permeate through the walls of outer plant cells without crossing the membrane, up until the casparian strip (dark purple line): a waxy, impermeable strip which occupies the cell interstices of the endodermis. At this point, a molecule must cross the membrane of a endodermal cell in order to progress to the pericycle, and as with the symplastic route, at least one further membrane crossing is required to enter into vascular cells. Figure created using BioRender.com (2020); based on Zhou et al. (2013)

Description	Method(s) of characterisation	Method(s) in section(s)	Results shown in figure(s)
Identifying uptake characteristics of non-functionalised carbon nanodots in <i>Arabidopsis</i> roots	Confocal microscopy	3.2.3	3.4,3.6, 3.7, 3.8, 3.9
Identifying uptake characteristics of functionalised carbon nanodots in <i>Arabidopsis</i> roots	Confocal microscopy	3.2.3	3.10, 3.11, 3.12, 3.13, 3.14
Assessing uptake dynamics of non-functionalised carbon nanodots in plants	Flow cytometry	3.2.4, 3.2.5	3.15
Analysis of impact of carbon nanodots on root transcriptome	RNA sequencing	3.2.6	3.16, 3.17
Assessing carbon nanodots as a delivery tool for the flg-22 peptide	Confocal microscopy, RNA sequencing	3.2.3, 3.2.6	3.14, Table 3.5

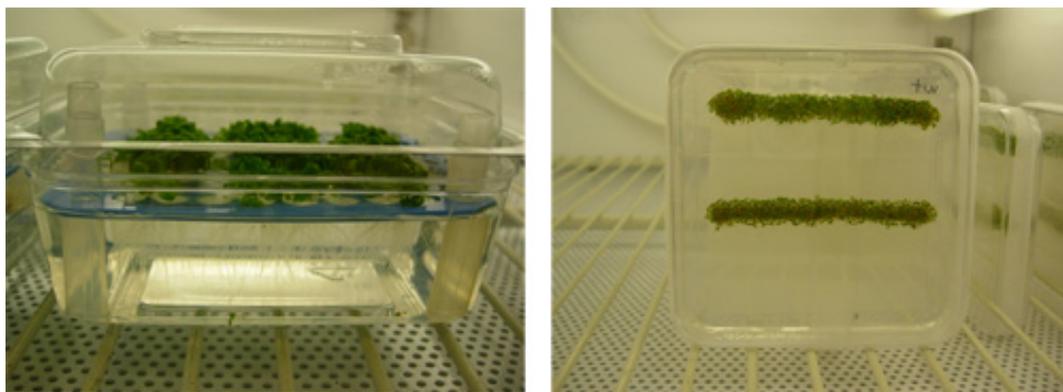
**Table 3.1.:** Table of experiments reported in Chapter 3

## 3.2. Methods

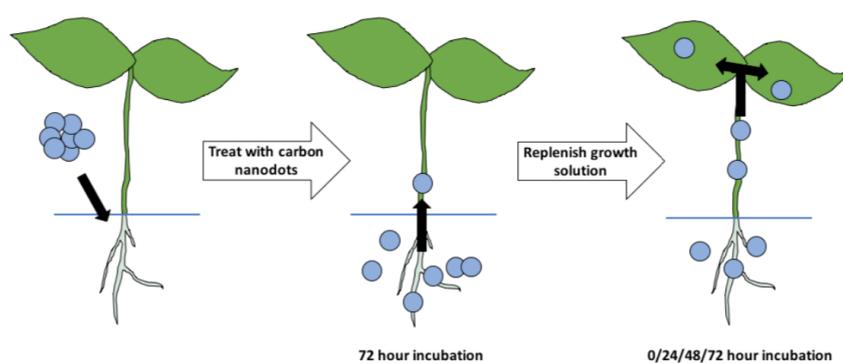
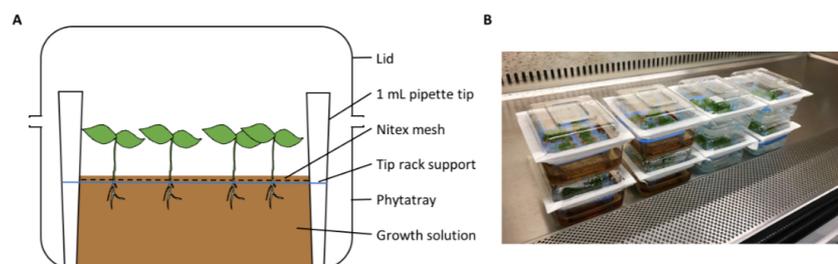
### 3.2.1. *Arabidopsis* growth and handling

*Arabidopsis* seeds were sterilised by washing with 70% ethanol, 10% sodium hypochlorite, and then rinsed twice with sterile water before being sown. Replete Murashige and Skoog media formulation buffered to pH 5.7 was used for both liquid and solid growth media.

For liquid media, the required volume of seeds in sterile water was aspirated into a 1ml micropipette, before being deposited directly onto a Nitex nylon mesh, supported by a plastic assembly and enclosed within a Phytatray (Sigma). For solid media, seeds in water were aspirated in 12-16 $\mu$ l increments, into a 20 $\mu$ l micropipette with a cut pipette tip. Seeds were then deposited individually onto the surface of the agar in a 150mm square Petri dish (Sarstedt) and the dishes oriented vertically for growth. Once sown, the containers were sealed with Micropore tape (3M) and seeds were grown for 7-10 days in long-day conditions (16-hour light) at 22 degrees Celsius before being treated with carbon nanodots.



**Figure 3.2.:** Photograph of cultures of *Arabidopsis* seedlings to be treated with carbon nanodots on liquid media (left) and solid media (right). Liquid media can be changed at any time with relative ease, whereas solid media offers potential spatial stratification of conditions within a single root. Both of the pictured are nutrient-replete media (Murashige and Skoog formulation).



(a)

**Figure 3.3.:** *Arabidopsis* seedling growth protocols. (a) Photograph of a two-week-old wild-type (Columbia-0 line) *Arabidopsis* seedling, grown vertically on solid media (MS agar). (b) (upper) Illustration and photograph of hydroponics apparatus used to culture arabidopsis seedlings and (lower) illustration of “pulse-chase” experiment design used to infer nanodot uptake dynamics - white arrows indicate the passage of time, while black arrows indicate the hypothesised transport route of nanodots.

### 3.2.2. Nanodot application

Carbon nanodots were simply added to the Murashige and Skoog formulation (MS) growth media in which *Arabidopsis* seedlings were grown - by pipetting in (or on) a 0.5ml volume. Results were similar from plants grown in liquid media and on solid media. The latter has the advantage that different parts of the same plant root can be differently treated; growth on a vertical or near-vertical plane (as seen in Figure 3.3 on page 68) allows simple study of root architecture, without requiring the seedling to be transplanted for observation.

In principle similar to a radiolabeling pulse-chase experiment, a high concentration (20 mg/ml) of non-functionalised nanodots were added to the growth media of hydroponically cultured *Arabidopsis* seedlings for 72 hours before the first sample was processed for flow cytometry. After this initial 'pulse' step, the media was replaced with fresh media of the same MS formulation (without any carbon nanodots). Subsequent samples were taken at 24 hour intervals to provide information about the decay dynamics of carbon nanodots in *Arabidopsis* root tissues as micronutrients were 'chased' (eliminated) from plant tissues by the refreshed media.

### 3.2.3. Confocal microscopy and image analysis

Laser scanning confocal microscopy was used to ascertain the uptake of carbon nanodot cores into *Arabidopsis* tissue samples. Roots and leaves were imaged with a Zeiss LSM 880 microscope. Carbon nanodot cores manufactured by the method used in this study have a Stokes shift of around 50nm, and a peak absorbance around 400nm (Hill et al., 2016). Therefore, the sample was illuminated with a UV 405nm laser, and light collected in the 420-450nm range. Samples were counter-stained with propidium iodide to highlight cell walls, and emitted light images from 580-650nm were collected from samples excited with 488nm laser excitation. An array detector in wavelength mode was used to validate this choice of optical filter. The properties of confocal microscopy allow the collection of light exclusively from a shallow tangential section through the root a few microns in depth: this is small enough to gather light from a cell's cytoplasm without including light from the cell interstices above and below the optical plane. In some cases, the brightness of each pixel in the image was then calculated from the sum value of a vertical "stack" of shallow confocal slices (a process referred to as 'z-projection').

In order to determine whether nanodots in this experiment are cytosolic, the array detector on the confocal microscope was set to wavelength mode, allowing

the gathering of coarse fluorescence emission spectra at different points on the micrograph. a single shallow optical section is then taken through the root, and the ratio of emitted light at 450nm and at 550nm was compared, to indicate the relative abundance of carbon nanodots in each region of the image.

To better understand the distribution of nanodot-associated fluorescence between tissues in the plant root, radial optical sections were taken from treated plant roots by confocal microscopy. Pixel values (brightnesses) from the 'blue' channel (420-450nm) were extracted from line selections perpendicular to the root axis. These values were summed and plotted against their position on the line: patterns in the resulting plot were used to provide information about the most likely uptake route of nanodots.

### 3.2.4. Root Protoplast generation protocol for *Arabidopsis*

*Adapted from (Birnbaum et al., 2005)*

Protoplast solutions A and B were prepared in advance. Protoplasts were generated from *Arabidopsis* root and leaf tissues separately.

**Solution A** Solution A is an aqueous buffer solution comprising 500 mmol/l of mannitol, 2 mmol/l of magnesium chloride, 0.1% (w/v) bovine serum albumen, 2 mmol/l calcium chloride, 2 mmol/l MES hydrate, and 10 mmol/l potassium chloride, adjusted to pH 5.5 with potassium hydroxide. This was typically made in a batch of 200 ml, and stored at -20°C in 10ml aliquots.

**Solution B** Solution B was prepared by adding the following enzymes to Solution A, above: 1.2% (w/v) Cellulase RS (Duchefa); 1.5% Cellulase R10 (Duchefa); 0.2% Macerozyme R10 (Duchefa); and 0.12% Pectinase (Sigma).

#### **Tissue harvesting and protoplast extraction**

A 30 mm petri dish with a 70 micrometre cell strainer inside (VWR, 734-0003 - white) was filled 2/3rds full with solution B using a Pasteur pipette. Place a cell strainer with a 70 micron mesh (e.g. Falcon 352350) inside the petri dish being careful not to trap air bubbles below the mesh. Prepare 1 petri dish per phytatray or agar plate.

Roots or leaves were cut from 10 day old *Arabidopsis* seedlings, dried briefly, then sliced with a sterile scalpel or razor blade into 3mm sections. Sliced material was added to the Solution B inside the cell strainer and agitated gently.

Tissues were left to incubate for 2 hours at room temperature on an orbital shaker at 120 rpm. The release of protoplasts is indicated by the enzyme solution turning turbid after gentle swirling motion. The cell strainer was removed, with the remnants of intact tissues. Protoplast solution was then gently transferred into a 50 ml tube (using a serological or pasteur pipette) and centrifuged (300 ref; 4°C; 5min). The supernatant was discarded, and the cell pellet resuspended in 0.5 ml Solution A very gently, using a 1 ml micropipette, or pasteur pipette to mix until the suspension appeared homogeneous. After resuspending protoplasts, the cell suspension aliquots were passed through a 40 micron cell strainer (VWR, 734-0002 - blue) to separate any aggregates. The suspension may need to be diluted further depending on the cell sorter used: in this case, a twofold dilution with Solution A proved sufficient to avoid exceeding the recommended event rate range.

### 3.2.5. Flow cytometry

A BD Fortessa flow cytometer was used in this study - the lasers and photodetectors are calibrated automatically according to a proprietary standard method. Samples were run using a 70 µm nozzle at 30 psi sample pressure, and illuminated using 405nm and 420nm laser lines. Data from 100,000 events per biological sample was recorded in scatter channels, as well as all available fluorescence channels.

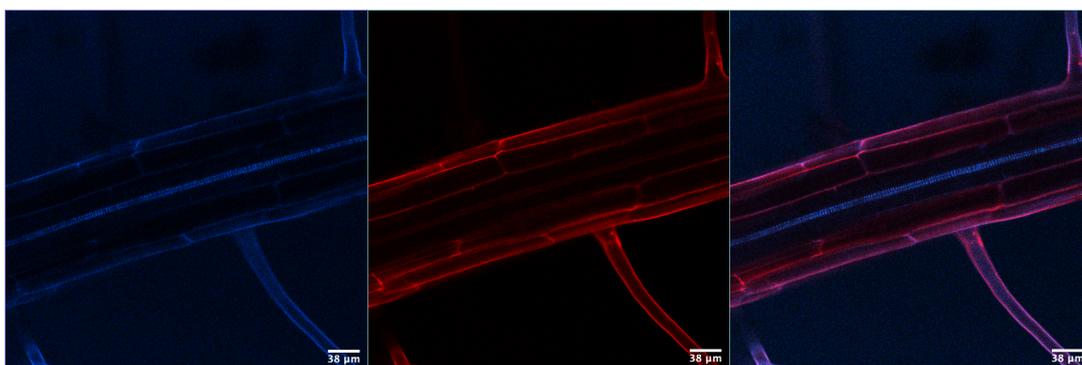
### 3.2.6. RNA extraction and expression data

Whole roots from 9-day old *Arabidopsis* seedlings were treated with flg-22 and functionalised and unfunctionalised carbon nanodots as seen in table . Fifty seedlings per plate were treated with 2 ml per plate at a concentration of 1 µM for 2-3 hours on solid media, sampled, and flash-frozen with liquid nitrogen. RNA was extracted using Monarch(R) Total RNA miniprep kit (New England Biolabs) and stored at -80°C before being sent to an internal facility for sequencing (Illumina MiSeq). The *Arabidopsis thaliana* reference genome version TAIR 10 was downloaded using the Ensembl informatics portal (Cunningham et al., 2018), and RNA sequencing reads were aligned to the genome using STAR (Dobin et al., 2012). DeSeq was used to normalise read counts and conduct principle component analysis of whole-transcriptome data. Normalised counts were exported for differential expression analysis, and Virtual Plant was used to analyse gene enrichment versus the TAIR 10 baseline, as well as gene ontology (GO) term analysis (Katari et al., 2010).

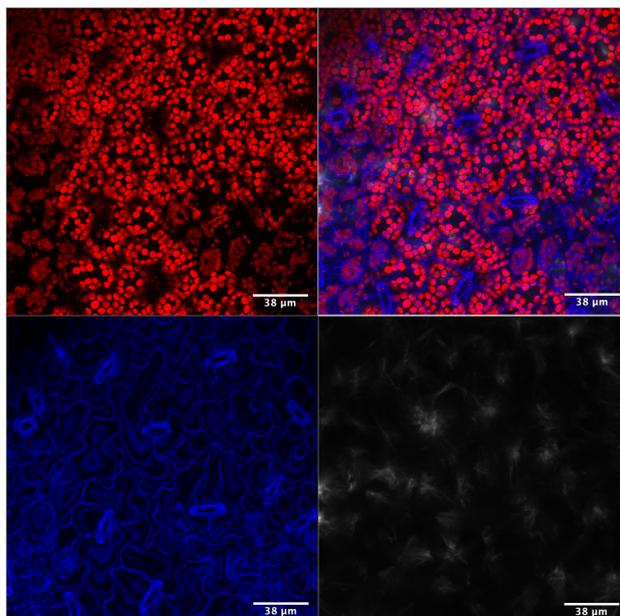
### 3.3. Results

#### 3.3.1. Confocal microscopy of *Arabidopsis* seedlings indicates a route of systemic uptake

Nanodot cores were discernible in *Arabidopsis* roots under the confocal microscope after incubation for 48 hours, with the result becoming steadily clearer and brighter as they continued co-incubation over the following 24 to 36 hours. Initial observations were used to determine that *Arabidopsis* seedlings were best imaged at 9-12 days old, with the fluorescence of the carbon nanodots reaching its peak after roughly 60 hours of co-incubation. After this time, fluorescent signals consistent with carbon nanodots were noticeable in the whole root, (Fig. 3.4) as well as in the cell interstices in aerial plant tissue, particularly around the guard cells surrounding the stomata in the leaves (see Figure 3.5).



**Figure 3.4.:** Confocal micrograph of *Arabidopsis* root optical section treated with carbon nanodots. Carbon nanodots (blue) were added to 7 day old plants in liquid MS growth media to a final concentration of 10 mg/ml for 72 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Fluorescence images at 420-450 nm (left) 580-650 nm (middle) and composite (right). Root axial optical section; scale bar is 38  $\mu\text{m}$ .



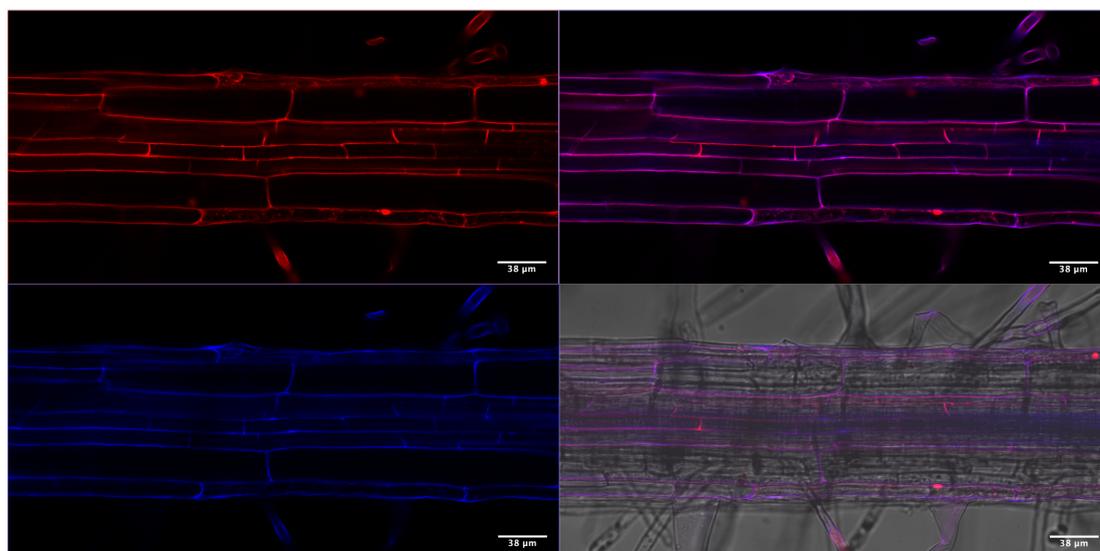
**Figure 3.5.:** Confocal micrograph of *Arabidopsis* leaf from seedling treated with acid-functionalised carbon nanodots. Carbon nanodots (blue) were added to 7 day old plants in liquid MS growth media to a final concentration of 10 mg/ml for 72 hours prior to imaging. Filter parameters for propidium iodide (red, 580-650 nm) were used to illuminate chloroplasts and thus provide context, although no PI was added. Fluorescence images at 420-450 nm (lower left) 580-650 nm (upper left), transmitted light (lower-right) and composite (upper right). Scale bar 38  $\mu\text{m}$ .

Given the similarity in size and polarity of carbon nanodot cores to a small soluble protein, it would be reasonable to expect them to be transferred by an apoplastic route through the root. Moreover, examining these alongside the images in Figure 3.6, areas at cell perimeters are particularly bright, which might corroborate the idea that nanodots are moving in the interstices of cells i.e. the apoplast.

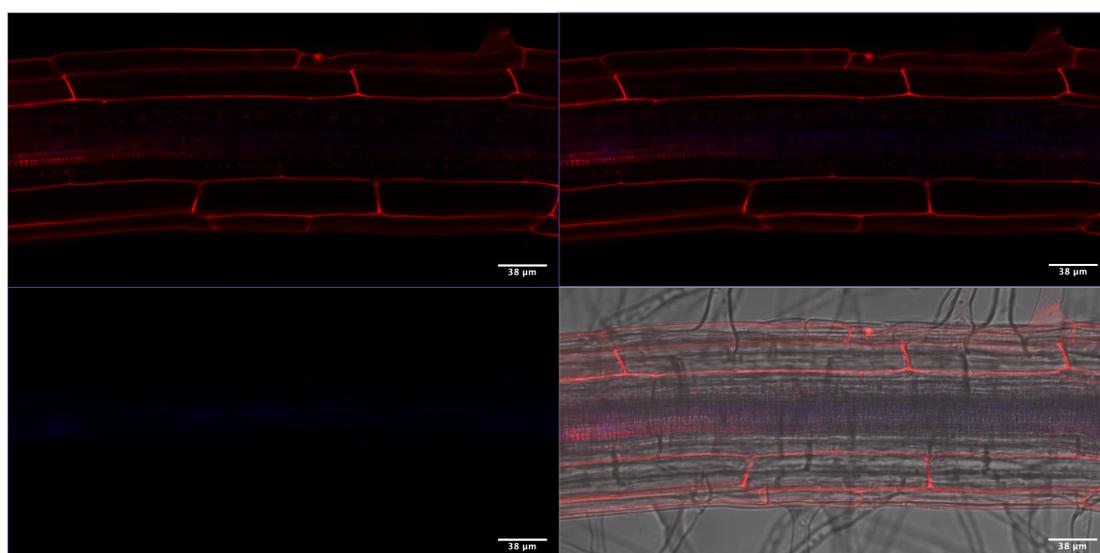
However, as can be seen in Figure 3.4: although carbon nanodots appear to be present in the stele (and also in the aerial tissues, despite having been applied solely to the plant roots) there is no evident concentration around the Casparian strip in mature plant roots. This is an observation of particular note since, for the apoplastic route, it is the waxy Casparian strip that would normally obstruct polar complexes from entering the plant vasculature, forcing the molecule through at least one transmembrane region.

Altogether, this leaves two possibilities: either the Casparian strip presents an insignificant barrier in this case; or (more likely) the nanoparticles are not entirely apoplastic. In practice it is comparatively difficult to test the former hypothesis, whereas testing the latter requires altering a few settings and more closely analysing data from the microscope, so this is the method selected.

There is a particularly bright area in the root xylem in Figure 3.4, which suggests that nanodots may be taken up into the xylem and transported passively into aerial tissues. However, the xylem is comprised of lignified tissues, which are known to autofluoresce in response to violet light: in order to clarify the situation, it is necessary to threshold the images, such that low-level autofluorescence can be suppressed. An example of this process can be seen in Figure 3.7.

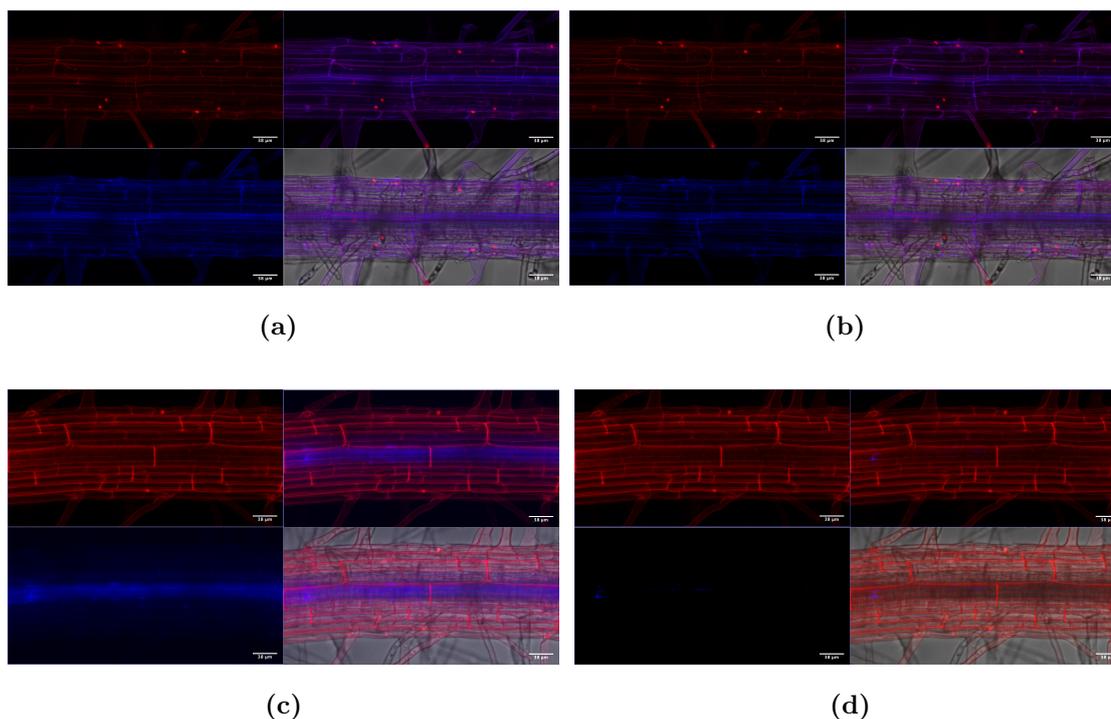


(a)



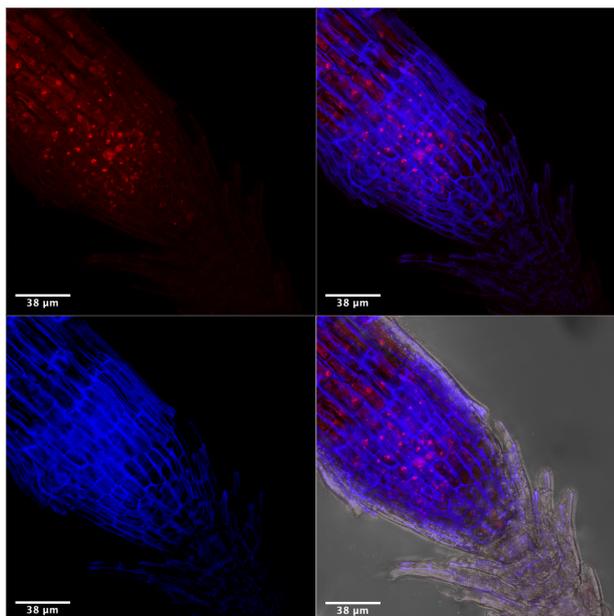
(b)

**Figure 3.6.:** Confocal micrograph of *Arabidopsis* root optical section treated with (a) carbon nanodot cores and (b) control (DI water). Carbon nanodots (blue) were added to 10 day old plants in liquid MS growth media to a final concentration of 10 mg/ml for 72 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Tangential optical section with fluorescence images at 420-450 nm (lower left) 580-650 nm (upper left) composite (upper right), and composite with transmitted light image (lower right). Scale bar 38  $\mu$ m.



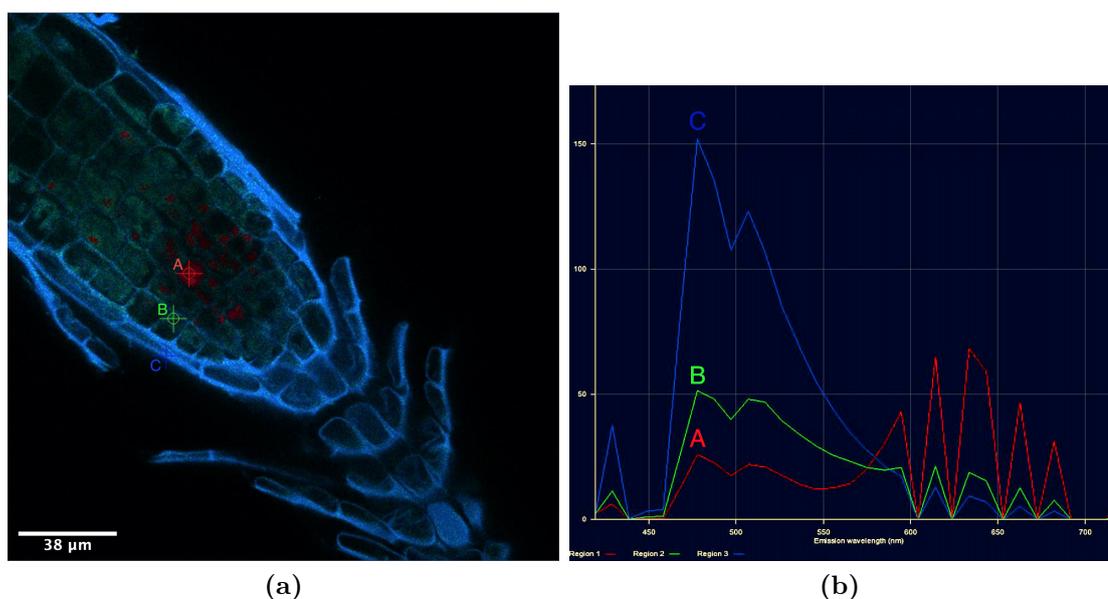
**Figure 3.7.:** Confocal micrograph images of *Arabidopsis* roots treated with (a-b) un-functionalised carbon nanodots (c-d) control treatment (DI water). Carbon nanodots (blue) were added to 10 day old plants in liquid MS growth media to a final concentration of 10 mg/ml for 72 hours prior to imaging. Images are sums of image ‘stacks’ from the same root at different optical planes: fluorescence at 420-450 nm (lower left), 580-650nm (upper left), composite (upper right), and composite with transmitted light image (lower right). (b) and (d) are data from (a) and (c), respectively, thresholded to exclude blue values less than 85 (arbitrary units) so as to suppress autofluorescence from lignified tissues in the stele. Bright areas in the cell interstices may suggest an apoplastic uptake route; bright regions in xylem (along the central axis) of this mature root may indicate that carbon nanodots are passively transported into aerial tissues of the plant. Scale bar is 38  $\mu\text{m}$ .

Figure 3.7 shows that the bright regions visible at cell boundaries and in the xylem are persistent, in spite of thresholding to a level which suppresses the image in the control-treated root. Parts of this are consistent with the idea that, being relatively small and highly hydrophilic, nanodots would follow an uptake path through the plant similar to that of water: progressing predominantly through the root apoplast, and transported in the xylem to other tissues. However, a bright xylem does not constitute proof of an apoplastic uptake route. Figure 3.8, shows a root tip which has been similarly treated.



**Figure 3.8.: Confocal micrograph of an Arabidopsis root tip, treated with un-functionalised nanodots.** Images are sums of image ‘stacks’ from the same root at different optical planes: fluorescence at 420-450 nm (lower left), 580-650nm (upper left), composite (upper right), and composite with transmitted light image (lower right). Note the bright regions in cell interstices, similar to 3.7, despite differences in cell development stage. Scale bar is 38  $\mu\text{m}$ .

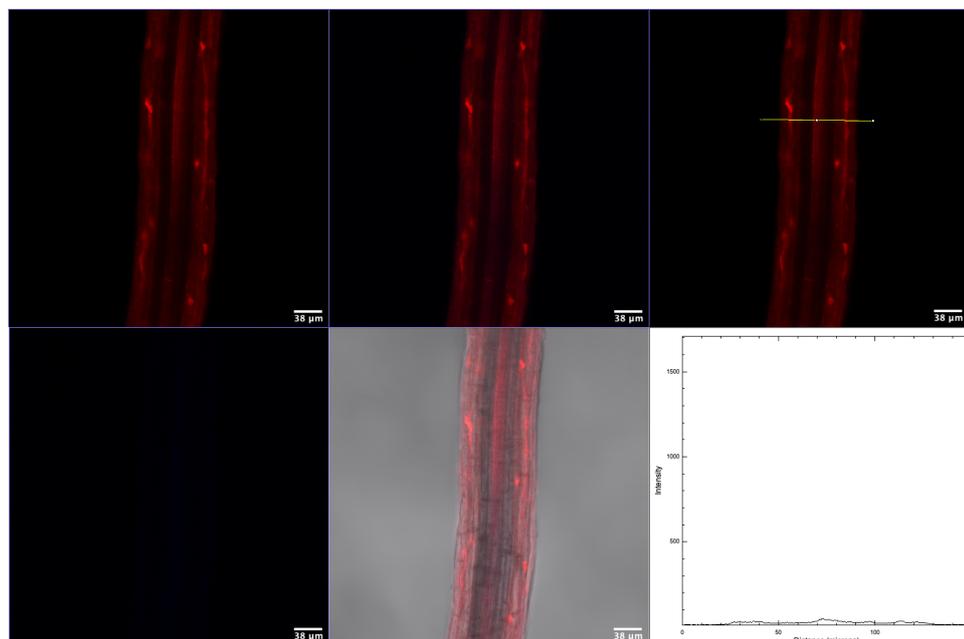
The root tip in Fig. 3.8 is particularly bright - although this should come as no surprise, since root meristems are comprised of rapidly dividing cells, whose activity will promote autofluorescence and biospeckle (Zdunek et al., 2014), and cells in the root cap are typically coated in autofluorescent waxes. Nevertheless, the cell boundaries are particularly bright, as seen previously, and there appear to be some brighter intracellular regions as well. This part of the plant is of interest as the cells at the root tip are at an earlier developmental stage than those further up the root. In order to investigate further, the microscope is put into wavelength-scanning mode, and the image is changed from z-projection to a single axial optical section: this result is seen in Figure 3.9.



**Figure 3.9.: Fluorescence data from *Arabidopsis* root treated with non-functionalised carbon nanodots in liquid media.** (a) Wavelength-coloured confocal micrograph (axial optical section) from the meristem of an *Arabidopsis* root tip from a 7-day old seedling, treated with 10 mg/ml non-functionalised carbon nanodots in liquid MS media 48 hours prior to imaging. Image annotated with ROI indication crosshairs (A-C). Note the pale blue-coloured regions in the intercellular regions - this suggests the nanodots are present in the root apoplast. (b) Light emission plots corresponding to regions of interest in (a), obtained with the microscope’s array detector under 405 nm coherent light excitation.

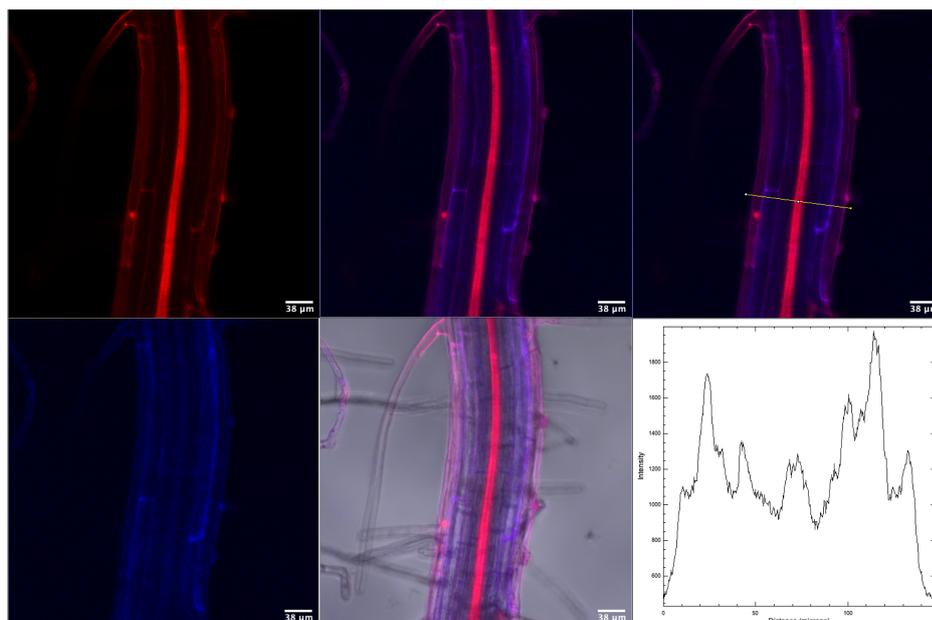
As hinted at by the previous figure, there is a bright fluorescent peak at 450 nm at the cell boundaries, even in the middle of the root tip. There is also a similar (though less prominent) peak in the intracellular regions of the image (represented by a greenish colour in 3.9a), which contraindicates the idea of an apoplastic uptake route. In summary, the overall pattern is not yet clear; moreover, it has been shown by Swift et al. (2018) that functionalisation of nanodots may affect their physical properties, and therefore perhaps their uptake through plant roots. With the intention of clarifying the previous findings, and also of increasing the relevance to functionalised nanodots, a collection of optical micrographs from plant roots treated with various carbon nanodots is established below. At thirty intervals along the root, a transverse line of 150  $\mu\text{m}$  was drawn: the pixel values from the blue channel (420-450 nm) were extracted, and their sum is plotted against position on the line. The pattern in the line will evidence the uptake route of the nanoparticles: a symplastic uptake should produce a smoother curve, more uniform across the root, with perhaps small peaks at either extremity of the root as the first membrane crossing occurs, and a slight peak near the middle as material crosses from the pericycle into the vascular tissue; whereas an apoplastic

uptake would be indicated by a trace with numerous peaks, with perhaps a more pronounced trough towards the middle as the material is impeded by the endodermis. In either case, it has been established that a peak will appear in the central area of the root where the xylem resides. The results are shown below.



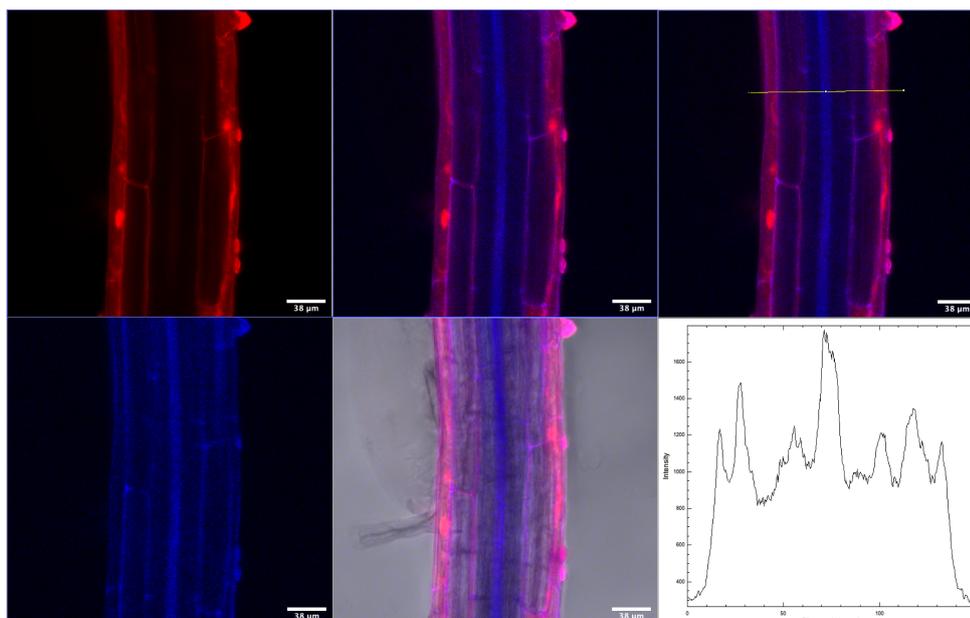
**Figure 3.10.:** Confocal micrograph of *Arabidopsis* root optical section control treatment (water). Treatments were added to 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence at 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection shown in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38  $\mu\text{m}$ .

Figure 3.10 above, the mock treatment shows little except that the thresholding process is working fairly well to suppress autofluorescence from the xylem, and that the xylem is indeed where the peak autofluorescent signal is located. The graph scales are consistent with the subsequent figure 3.11, showing carbon nanodots functionalised with a short-acid linker.



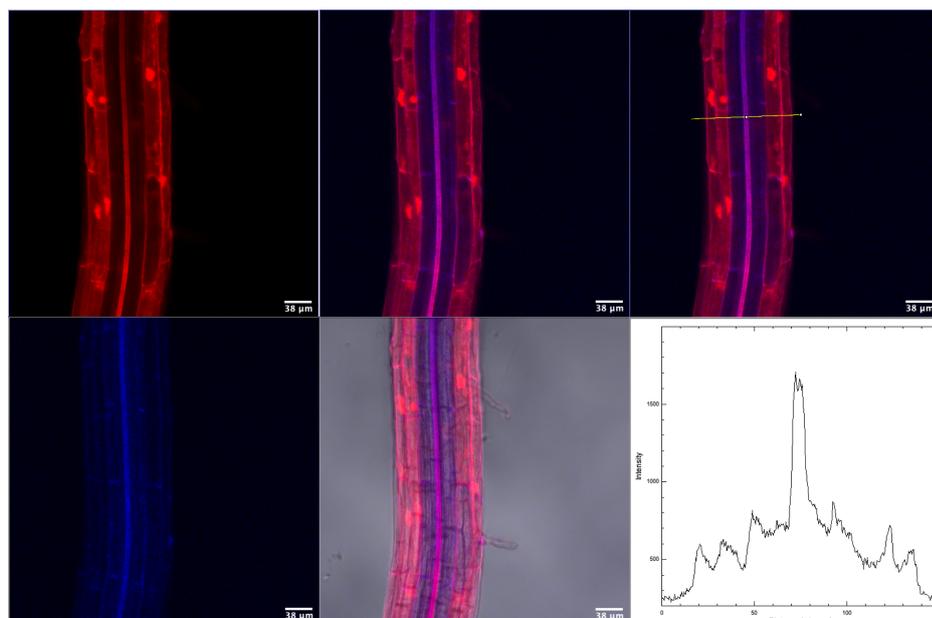
**Figure 3.11.:** Confocal micrograph of *Arabidopsis* root (axial optical section) treated with acid-functionalised carbon nanodots. Carbon nanodots (blue) suspended in DI water to concentration of 20 mg/ml were pipetted onto roots of 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection shown in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38  $\mu\text{m}$ .

The result of Figure 3.11 is very interesting - the curve appears to be somewhat of a hybrid of the descriptions given above. There is somewhat of a trough in the area corresponding to tissues immediately inside the endodermis, but the overall shape is smoother than might be expected of an apoplastic uptake, and the absence of clear peaks at either edge of the root would indicate that some degree of permeation into the apoplast is taking place. The central peak corresponding to the xylem is also smaller than expected (although still much larger than the autofluorescence peak seen in the mock treated root). The next figure, 3.12, shows a nanodot with similar surface chemistry but a larger overall size.



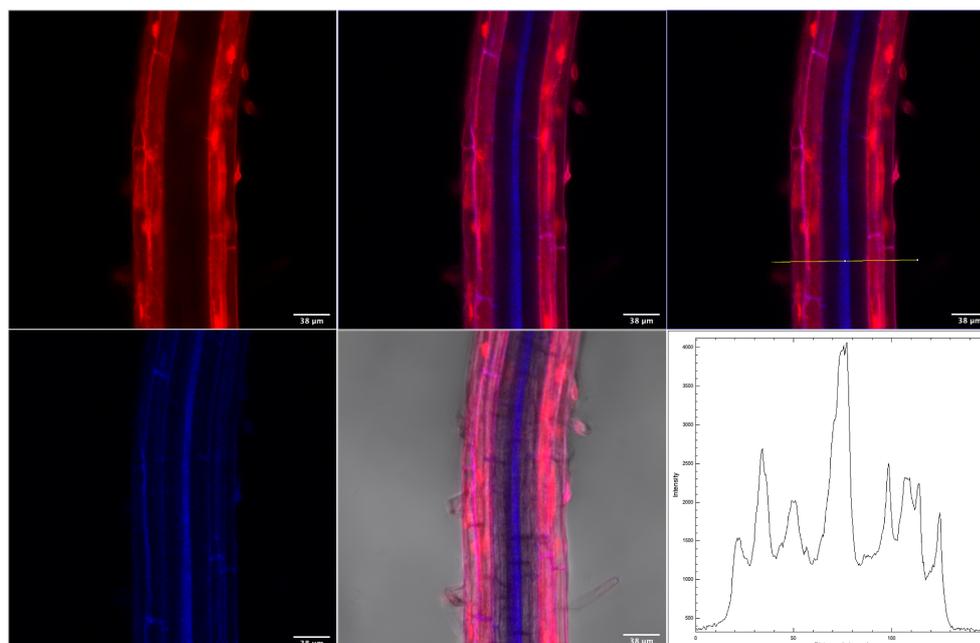
**Figure 3.12.:** Confocal micrograph of *Arabidopsis* root (Axial optical section) treated with long-linker-functionalised carbon nanodots. Carbon nanodots (blue) suspended in DI water to concentration of 20 mg/ml were added to 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38  $\mu\text{m}$ .

Compared with the figure before it, the result in Figure 3.12 shows a more consistent signal across the entire root; in particular, there is no central trough to speak of, which supports a model of predominantly symplastic uptake, although the relative amplitude of the periodic peaks in the graph suggest some accumulation may be occurring at cell boundaries. The next figure shows carbon nanodots functionalised with flg-22: a larger and more hydrophobic molecule than the long-acid linker.



**Figure 3.13.:** Confocal micrograph of *Arabidopsis* root (axial optical section) treated with flg-22-functionalised carbon nanodots using short acid linker. Carbon nanodots (blue) suspended in DI water to concentration of 20 mg/ml were added to 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection shown in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38  $\mu\text{m}$ .

There are several aspects of note in Figure 3.13, of which the first is the much lower signal in the root overall. Since flg-22 is an indicator of pathogenic threat, and it makes sense that active transport should slow in the presence of a known pathogen, this may be an indication not only that the flg-22-functionalised nanodots are transported symplastically, but also that the flg-22 retains some activity even after chemical attachment to nanodots. Also of note is the very pronounced central peak, and the fact that fluorescence trends steadily upward towards the centre of the root: the former observation may indicate a signalling response by the root to flg-22.

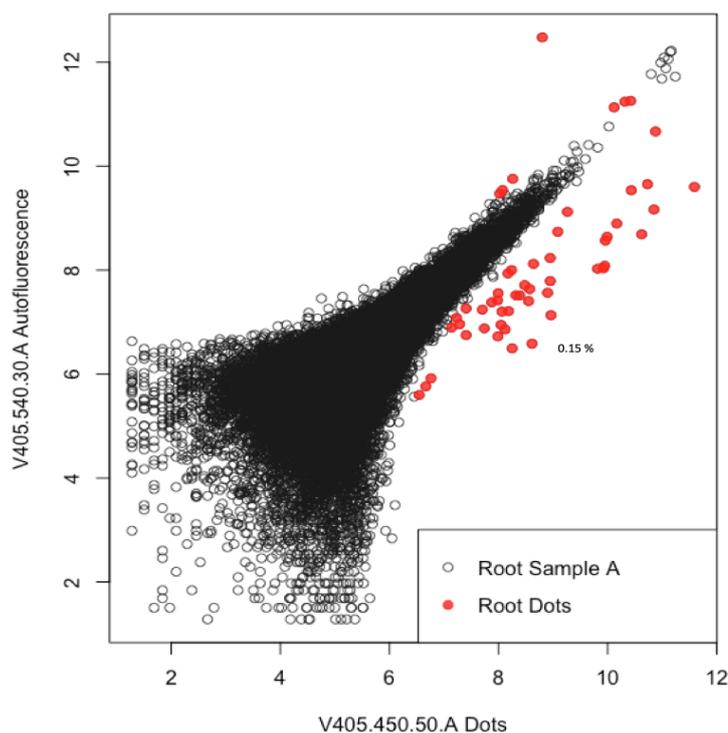


**Figure 3.14.:** Confocal micrograph of *Arabidopsis* root (axial optical section) treated with flg-22-functionalised carbon nanodots with long acid linker. Carbon nanodots (blue) suspended in DI water to concentration of 20 mg/ml were added to 9-day-old seedlings on replete MS solid growth media for 24 hours prior to imaging. Propidium iodide (red) was added five minutes prior to imaging, in order to stain cell walls. Axial optical section: red fluorescence at 580-650nm (upper left), blue fluorescence 420-450 nm (lower left), composite (upper middle), composite with transmitted light image (lower middle), composite with exemplar line selection shown in yellow (upper right) and plot of blue fluorescence intensity versus position (lower right). Scale bar 38 µm.

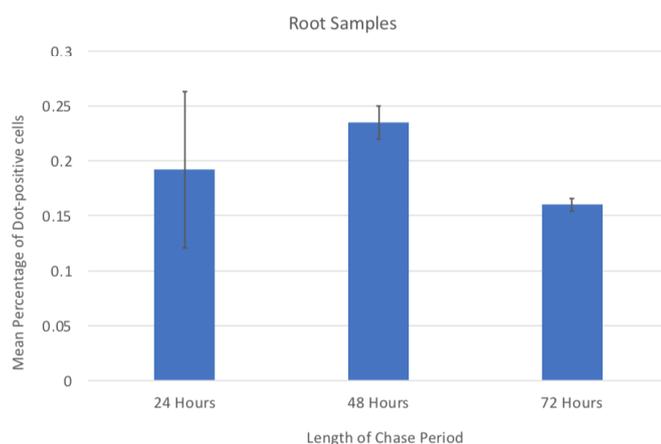
Compared to the particle in the previous figure, the nanodot in 3.14 had a larger linker region, and recall from conclusions of the previous chapter that there was low confidence in the establishment of chemical bonding between the peptide and the nanodot. Note that the graph appears as somewhat of a hybrid of Figures 3.12 and 3.13: there is a series of sharp peaks across the root (perhaps the most pronounced of any treatment) and a pronounced central peak as seen in Fig. 3.13. From this result, it seems likely that some fraction of nanodots have not bound flg-22, and have been able to make some progress into the apoplast, although the central peak consistent with the previous figure would suggest that flg-22 retains its activity.

#### **3.3.2. Flow Cytometry of protoplasts indicates possibility for cytosolic nanoparticles**

Fluorescence of plant protoplasts treated with carbon nanodots was shifted compared to the control group - this indicates that carbon nanodots remain in the cytoplasm throughout the preparation protocol required for protoplast generation, which requires several cell washing steps and incubation with an enzyme cocktail for over an hour. Of note is a small sub-population that appear to have retained much more nanodots than the surrounding cells - highlighted in red in Figure 3.15a on page 85 - whose fluorescence signals are one or more orders of magnitude more fluorescent than their peers. The reason for this is as yet uncertain, but the author hypothesises that this may be due to cell wall fragments remnant from the protoplast generation step, which may have retained a large proportion of the interstitial carbon nanodot signal seen earlier (although the reason for the retention is unclear). The highlighted population, however, serves as a metric for overall presence of carbon nanodots in the root tissue, and is used to produce the bar graph in Figure 3.15b on page 85.



(a)



(b)

**Figure 3.15.: Fluorescence cytometry data of carbon nanodot uptake in *Arabidopsis*.** (a) Scatter graph of log-fluorescence data collected from *Arabidopsis* root protoplasts illuminated with a 405nm laser, showing the carbon-dot fluorescence band (x-axis) against a yellow-green fluorescence band (y-axis) associated with off-target fluorescence (wide-band autofluorescence). The ratio between the two bands can be used to locate carbon nanodots (see Figure 3.9 on page 78); population of interest highlighted in red. (b) Bar graph showing the average proportion of cells identified in the highlighted population at each time point. These are data gathered from a flow cytometry experiment, - bars reflect the relative proportions of protoplasts considered to have a high concentration of carbon nanodots according to criteria defined using *Intelligate* software in Chapter 4.

### 3.3.3. Gene expression analysis suggests nanodots do not significantly alter transcriptional responses of flg-22 treated plant roots

STAR found 33602 transcripts across the 11 samples shown in table 3.2. The whole-genome data for all samples did not cluster well initially, as seen in figure 3.17: this may be due to complications during the experimental setup leading to different numbers of technical repeats. Four key identifiers of flg-22 responses in *Arabidopsis* were isolated from the data, and when individual treatments were compared to one another, two-sample homoscedastic t-tests showed some differences in their expression between treatments containing flg-22 and treatments without it, although not enough to present a robust test statistic, as seen in table 3.3. Moreover, the same test found high similarity in expression levels with and without nanodots, which is reassuring evidence that the transcriptional responses are not significantly affected by the presence of the nanoparticle, nor in this case the length of the linker molecule when flg-22 was present.

Gene name	Gene ID	M	C			F			CTF		CF		Key
		null	Cdots	Cdots	Cdots	null	null	null	Cdots	Cdots	Cdots	Cdots	Carbondots
		null	short	short	short	null	null	null	long	long	short	short	Linker
		null	null	null	null	flg22	Payload						
FRK1	AT2G19190	411	1775	780	120	1928	1331	1801	671	2231	1451	1920	
NHL10	AT2G35980	1583	1810	1848	774	1979	2131	1644	2052	1550	1989	2525	
WRKY22	AT4G01250	346	629	286	300	459	444	277	670	699	601	455	
MYB51	AT1G18570	2930	7020	3375	876	6106	5835	5206	3041	4280	5188	5662	

**Table 3.2.:** Normalised read counts of four flg-22-associated genes from RNA sequencing of *Arabidopsis* roots treated with flg-22 and carbon nanodots. Each column represents one technical replicate of fifty 9-day-old seedlings sampled 48 hours after treatment application with: mock solution (water, “M”), carbon dots with a short-acid linker (“C”), flg-22 peptide (“F”), carbon dots functionalised with a long linker and flg-22 (“CTF”), and carbon dots functionalised with flg-22 via a short linker (“CF”).

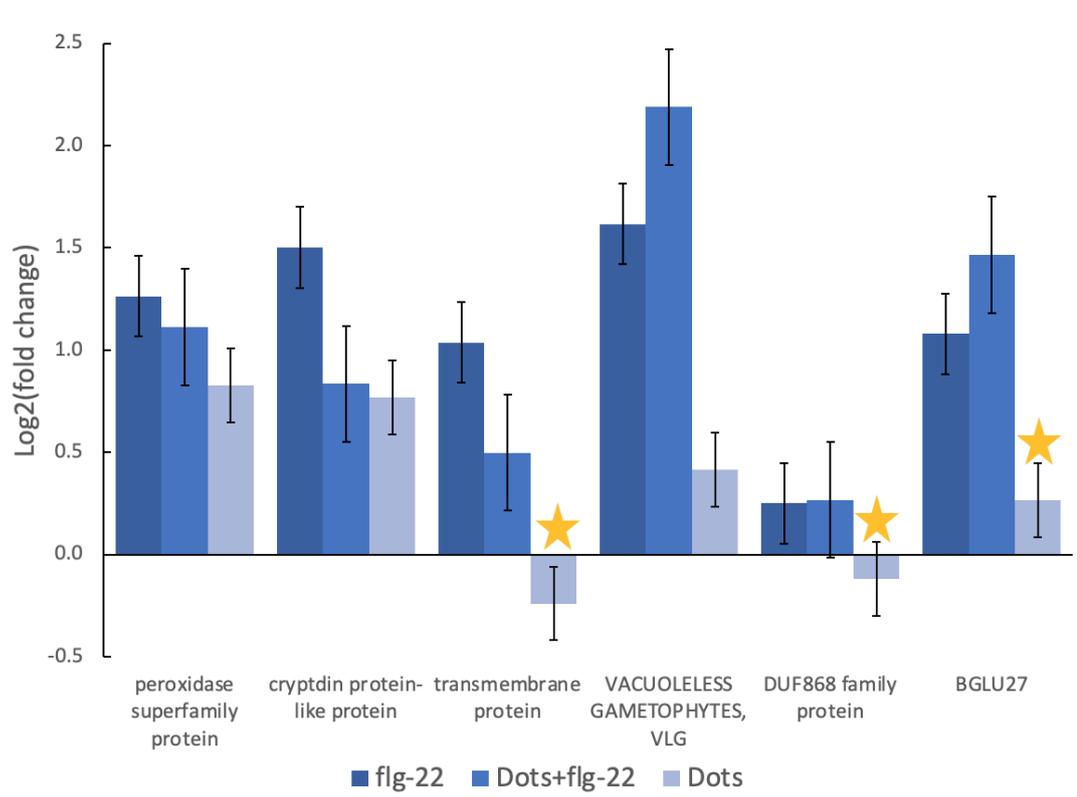
Gene name	Significance value for 2-tailed t-test			
	flg22 vs. null all	flg22 vs. null for short dots	Long vs. short linker	Dots vs. no dots
FRK1	0.049	0.306	0.801	0.852
NHL10	0.083	0.214	0.340	0.890
WRKY22	0.227	0.484	0.170	0.175
MYB51	0.197	0.522	0.117	0.498

**Table 3.3.:** Significance values for 2-tailed homoscedastic t-test of RNAseq counts for four markers of flg-22 response in roots.

Analysis set	Term	Observed Frequency	Expected Frequency	p-value
fig22 vs. null for short dots only	INTERACTION WITH THE ENVIRONMENT	33 out of 310 genes, 10.6%	1451 out of 26623 genes, 5.5%	0.02
fig22 vs. null for short dots only	cellular sensing and response to external stimulus	32 out of 310 genes, 10.3%	1296 out of 26623 genes, 4.9%	0.02
fig22 vs. null for short dots only	chemoperception and response	17 out of 310 genes, 7.1%	770 out of 26623 genes, 2.9%	0.02
fig22 vs. null for short dots only	plant hormonal regulation	17 out of 310 genes, 5.5%	539 out of 26623 genes, 2%	0.02
fig22 vs. null for short dots only	SYSTEMIC INTERACTION WITH THE ENVIRONMENT	19 out of 310 genes, 6.1%	664 out of 26623 genes, 2.5%	0.02
fig22 vs. null for short dots only	plant / fungal specific systemic sensing and response	17 out of 310 genes, 5.5%	599 out of 26623 genes, 2.2%	0.03
fig22 vs. null for all dots	cell wall	19 out of 369 genes, 5.1%	251 out of 26623 genes, 0.9%	1.90E-06
fig22 vs. null for all dots	cell wall	20 out of 369 genes, 5.4%	428 out of 26623 genes, 1.6%	0.000576
fig22 vs. null for all dots	cellular sensing and response to external stimulus	35 out of 369 genes, 9.5%	1296 out of 26623 genes, 4.9%	0.02
fig22 vs. null for all dots	INTERACTION WITH THE ENVIRONMENT	36 out of 369 genes, 9.8%	1451 out of 26623 genes, 5.5%	0.03
fig22 vs. null for all dots	chemoperception and response	23 out of 369 genes, 6.2%	770 out of 26623 genes, 2.9%	0.03
fig22 vs. null for all dots	SYSTEMIC INTERACTION WITH THE ENVIRONMENT	21 out of 369 genes, 5.7%	664 out of 26623 genes, 2.5%	0.03
fig22 vs. null for all samples	INTERACTION WITH THE ENVIRONMENT	64 out of 562 genes, 11.4%	1451 out of 26623 genes, 5.5%	7.76E-06
fig22 vs. null for all samples	cellular sensing and response to external stimulus	60 out of 562 genes, 10.7%	1296 out of 26623 genes, 4.9%	7.76E-06
fig22 vs. null for all samples	SYSTEMIC INTERACTION WITH THE ENVIRONMENT	36 out of 562 genes, 6.4%	664 out of 26623 genes, 2.5%	6.58E-05
fig22 vs. null for all samples	response to biotic stimulus	18 out of 562 genes, 3.2%	220 out of 26623 genes, 0.8%	0.000205
fig22 vs. null for all samples	chemoperception and response	34 out of 562 genes, 6%	770 out of 26623 genes, 2.9%	0.00472
fig22 vs. null for all samples	plant / fungal specific systemic sensing and response	28 out of 562 genes, 5%	599 out of 26623 genes, 2.2%	0.00685
fig22 vs. null for all samples	nucleotide/nucleoside/nucleobase binding	57 out of 562 genes, 10.1%	1643 out of 26623 genes, 6.2%	0.01
fig22 vs. null for all samples	plant hormonal regulation	25 out of 562 genes, 4.4%	539 out of 26623 genes, 2%	0.01
fig22 vs. null for all samples	ATP binding	44 out of 562 genes, 7.8%	1209 out of 26623 genes, 4.5%	0.01
fig22 vs. null for all samples	immune response	9 out of 562 genes, 1.6%	99 out of 26623 genes, 0.4%	0.01
fig22 vs. null for all samples	response to wounding	8 out of 562 genes, 1.4%	79 out of 26623 genes, 0.3%	0.01
fig22 vs. null for all samples	animal specific systemic sensing and response	9 out of 562 genes, 1.6%	105 out of 26623 genes, 0.4%	0.02
fig22 vs. null for all samples	disease, virulence and defense	20 out of 562 genes, 3.6%	424 out of 26623 genes, 1.6%	0.03
fig22 vs. null for all samples	CELL RESCUE, DEFENSE AND VIRULENCE	46 out of 562 genes, 8.2%	1349 out of 26623 genes, 5.1%	0.03
fig22 vs. null for all samples	C-compound binding	7 out of 562 genes, 1.2%	77 out of 26623 genes, 0.3%	0.04
fig22 vs. null for all samples	plant defense response	7 out of 562 genes, 1.2%	81 out of 26623 genes, 0.3%	0.04
fig22 vs. null for all samples	plant organ	12 out of 562 genes, 2.1%	209 out of 26623 genes, 0.8%	0.04
fig22 vs. null for all samples	BIOGENESIS OF CELLULAR COMPONENTS	49 out of 562 genes, 8.7%	1514 out of 26623 genes, 5.7%	0.04
Long vs. short linker	peptide transport	7 out of 431 genes, 1.6%	66 out of 26623 genes, 0.2%	0.02
Long vs. short linker	drug/toxin transport	7 out of 431 genes, 1.6%	64 out of 26623 genes, 0.2%	0.02
Long vs. short linker	biosynthesis of glutamate	3 out of 431 genes, 0.7%	7 out of 26623 genes, 0%	0.03
Long vs. short linker	transport facilities	22 out of 431 genes, 5.1%	596 out of 26623 genes, 2.2%	0.03

**Table 3.4.: MIPS Biomaps analysis of expression data versus the TAIR10 baseline transcriptome.** Gene list for the condition highlighted in yellow is seen in table 3.5. Adjusted p-values calculated using a false discovery rate of 5%.

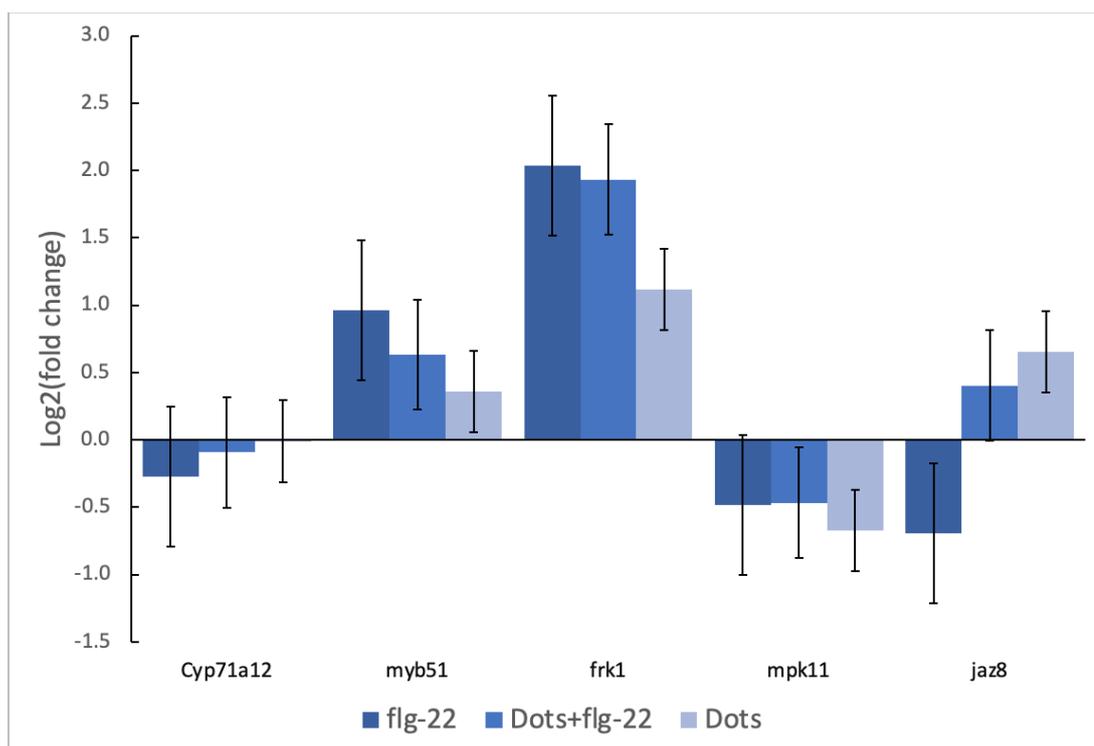
### 3.3 Results



**Figure 3.16.:** Fold-change expression data for Flg-22 key markers as identified by Beck et al. (2014). Nine-day-old *Arabidopsis* seedlings were treated with flg-22 only (dark blue bars) carbon nanodots functionalised with flg-22 (mid-blue bars), and carbon nanodots with a short linker only (light blue bars). Yellow stars indicate significant differences ( $p \leq 0.05$ ) in expression between carbon nanodot-treated plants and the other treatments, according to a homoscedastic t-test. Error bars indicate plus/minus one mean standard error.

flg22 vs. null: carbon dots with short linker	
Gene	GO terms
At4g14400:	ACD6, ankyrin repeat family protein
At3g50970:	LT130, XERO2, dehydrin family protein
At5g43860:	ATCLH2, CLH2, chlorophyllase 2
At4g34750:	SAUR-like auxin-responsive protein family
At4g28160:	hydroxyproline-rich glycoprotein family protein
At3g26830:	CYP71B15, PAD3, Cytochrome P450 superfamily protein
At3g16500:	IAA26, PAP1, phytochrome-associated protein 1
At1g74890:	ARR15, response regulator 15
At1g20450:	ERD10, LTI29, LTI45, Dehydrin family protein
At2g32020:	Acyl-CoA N-acyltransferases (NAT) superfamily protein
At4g34760:	SAUR-like auxin-responsive protein family
At5g19880:	Peroxidase superfamily protein
At4g28110:	AtMYB41, MYB41, myb domain protein 41
At1g35515:	HOS10, MYB8, high response to osmotic stress 10
At1g18710:	AtMYB47, MYB47, myb domain protein 47
At2g46680:	ATHB-7, ATHB7, HB-7, homeobox 7
At1g74430:	ATMYB95, ATMYBCP66, MYB95, myb domain protein 95

**Table 3.5.:** Biological process GO term analysis of genes highlighted in table 3.4 (flg-22-functionalised carbon nanodots versus nanodots with the short linker only).



**Figure 3.17.:** Fold-change expression data for key differentially expressed loci according to Stringlis et al. (2017).

### 3.4. Conclusion and Further Work

From the data in Figure Figure 3.15b on page 85, it is inferred that “fast-up, slow-down” dynamics are the dominant paradigm in this case. This has some influence on subsequent studies on plant roots, since the effective “washout time” of treatment is likely to be longer than it is in leaves. A delivery method via aerial tissues is expected to present different kinetics of response than are found for direct application to the root Doyle et al. (2019). It is unclear as yet the exact mechanism by which carbon nanodots are transported to aerial tissue - this may impact the results obtained in such studies - but observations made so far suggest that unfunctionalised carbon dots dissolved in the aqueous components of the growth media may be transported passively by the same route as water is transpired. This theory is corroborated by the higher concentration of carbon dots observed around leaf guard cells - as water evaporates and passes through the stomata, carbon nanodots (having a much higher boiling point) would presumably precipitate out and gather around the guard cells. Our collaborators have replicated this observation independently, and this corroborates published findings in maize (Chen et al., 2016). This corroborates similar findings concerning particles of a diameter around 5 nm (Carpita et al., 1979).

For larger particles, there appears to be slightly more clustering about the axial

ends of root cortical cells - there are some transport proteins known to locate here, such as the PIN family of auxin transporters (Xi et al., 2016; Sawchuk et al., 2013). Involvement of such transport proteins would be corroborated by the findings of several auxin response genes during GO term analysis, although it should be noted that *flg-22* also regulates auxin (Navarro et al., 2004). The indications that some symplastic transport may be taking place require an exploration of what transport proteins may be associated with this, particularly at the membrane crossing stage. Examining interactions with proteins in known pathways for the receptor-mediated endocytosis of nitrogen may form a useful starting point: some of these are identified by Noguero and Lacombe (2016).

More time points in the pulse-chase experiment would be ideal - although the fairly lengthy preparation process required for FC makes this difficult to do at scale, since preparation of each small batch of samples takes at least half a working day. Since the flow cytometry itself takes a relatively short time by comparison, it would be ideal to scale up the protoplast generation protocol such that it is feasible for one researcher to prepare, say, ten samples at once.

It would be ideal to be able to image more of the plant in one operation, and ideally this could be done with minimal intrusion to the sample. Light sheet microscopy would be a viable alternative, which might provide the possibility for in-situ microscopy of samples. However, as it stands, this would require design and development of a custom culture enclosure, with known optical characteristics: the disposable containers used for this study are made of PETG, which is excellent for chemical resistance and sterility, but would block the ultraviolet illuminating light, interfering with the image gathered.

Imaging flow cytometry would provide useful insight as to the distribution of nanodots within sorted protoplasts. Of particular interest would be the intracellular location of the fluorescence in the population highlighted in Figure 3.15a on page 85. The microscopy results in this chapter indicate that nanodots are mostly present in the apoplast and this underlines their utility for the study of interactions happening at the cell surface, including root-microbe interactions, and other receptor-triggered behaviours: indeed, the concept of a useful tool for data generation is effectively already proven for the effector mapping projects taking place in the lab. Our findings also corroborate previous studies, which show that particles less than 5 nm in size may have the capacity to traverse the intact cell wall efficiently (Carpita et al., 1979). However, ideally it would be desirable to have the majority of carbon nanodots on the inside of the cell membrane, since this would allow traceable interactions at the transcriptional level as well as the receptor level. This would allow the possibility to provide alternative regulatory

pathways dynamically, in parallel with the host genome, traceably, and without editing the germline (therefore more cost-efficiently than transgenic plants).

Imaging Raman spectroscopy is a method which might be of great interest to this study, despite fluorescence being a phenomenon that is typically avoided in Raman spectra where possible. Raman illuminates a sample with a monochromatic source, and uses the wavelengths of returning scattered light (which have been shifted by a quantum effect known as Raman scattering) to attempt to identify the material being illuminated. Since carbon nanodots are fluorescent, they should display a very asymmetric Raman spectrum (with Stokes-Raman scattering much greater in magnitude than anti-Stokes-Raman) and so imaging Raman spectroscopy (which produces an image on the micro-scale) should clearly show the location of carbon nanodots within a plant root.

Overall, this work shows that plant cells' treatment with carbon nanodots can be detected via flow cytometry; that carbon nanodots applied to plant roots can be used to deliver molecular payloads to aerial plant tissues, that nanodots appear to be taken up by plants more rapidly than they diffuse out of plant tissues, and that the overall gene expression profile of plant roots is not significantly altered by the presence of carbon nanodots.

## **4. *Intelligate*: A software tool to tackle complex FCCS data**

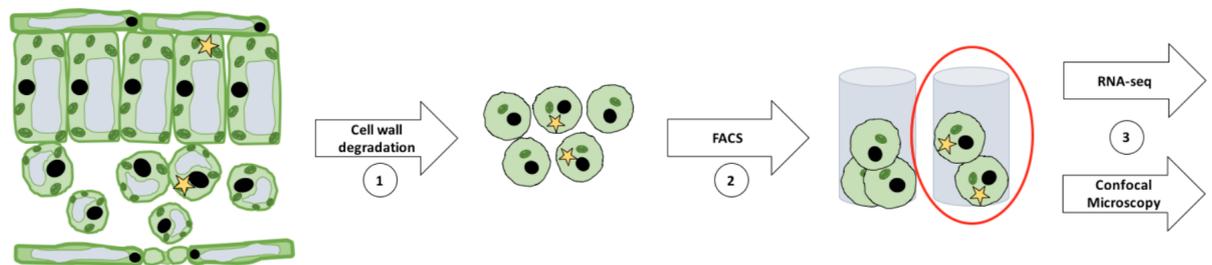
### **4.1. Introduction**

#### **4.1.1. Flow Cytometry and Cell Sorting: an introduction**

Flow cytometry (FC) is an experimental method that allows the individual collection of optical data from a large population of particles or cells. FC is a method which allows the collection of optical data (scattering and fluorescence) from each cell within a sample. Flow cytometric cell sorting (FCCS) allows these cells to be physically sorted into two or more groups for subsequent experiments, based on these optical parameters. This can be performed on a per-cell basis, allowing hundreds of thousands of cells to be recorded, analysed, and sorted within a few minutes.

While FC technology is originally associated with mammalian immunology, it has great potential in the field of plant sciences: the ability to collect and analyse data from a large number of cells with resolution down to the individual is particularly appealing, since it allows the gathering data about the spatial resolution of plant processes. This in turn facilitates the inference of systems interaction models of plant cells.

Unlike mammalian immune cells, plant cells tend to form cohesive multicellular structures, using biopolymers such as cellulose and pectin in order to retain their mechanical resilience. Thus, in order to prepare plant cultures for FC/CS, it is necessary to isolate the cells: in this case, we use a method employing an enzyme cocktail containing cellulase, pectinase, and others in order to degrade the cell walls, while leaving the cell membranes intact. This form of modified plant cell is known as a protoplast; the method (shown in Appendix 3.2.4) is known as protoplast generation.

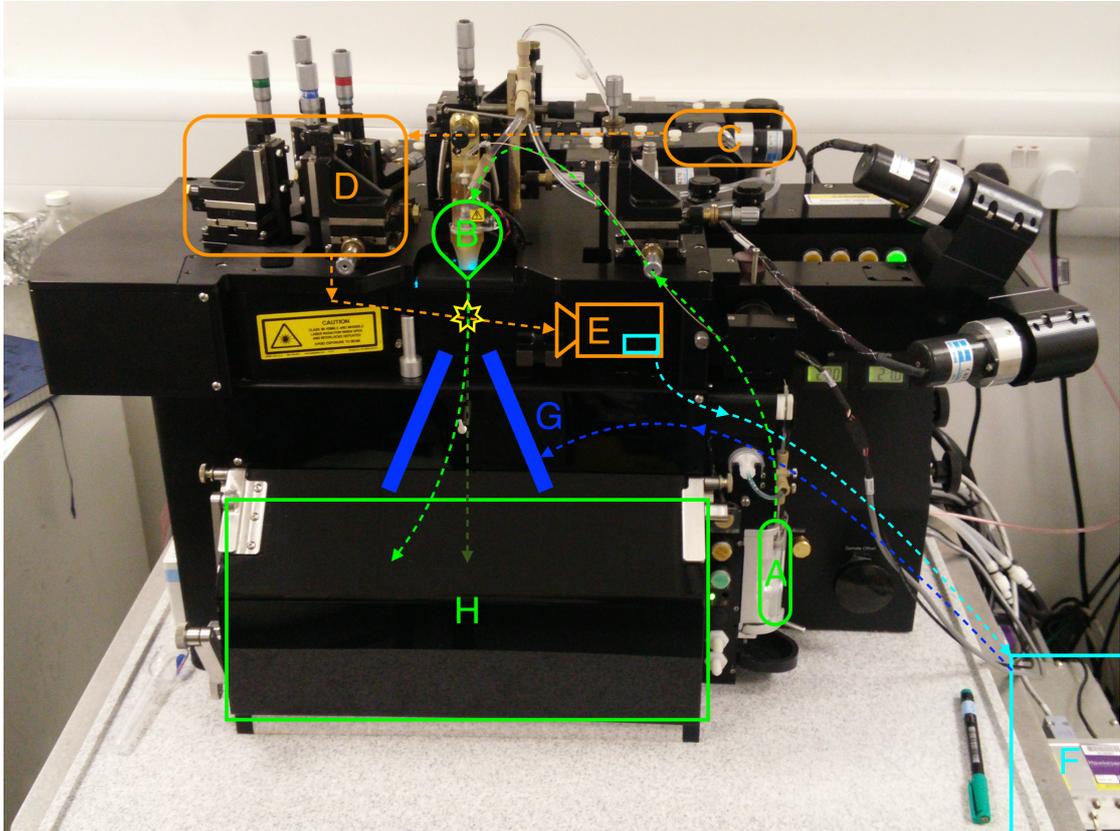


**Figure 4.1.: Illustration of practical workflow for collecting single-cell data from plant tissues.** Protoplast generation (1) involves the degradation of cell walls via the application of an enzyme cocktail - thus transforming the multicellular structure into a suspension of plant protoplasts. The protoplasts are subsequently sorted by their fluorescent and optical parameters via FCCS (2) in order to separate cells of interest from the bulk sample: these could be e.g. quantiles of GFP-expressing cells in a protein fusion experiment, or cells otherwise decorated with a selective fluorescent marker. The sorted protoplast samples can then be sent for further analysis via e.g. RNA-seq, allowing the collection of expression data.

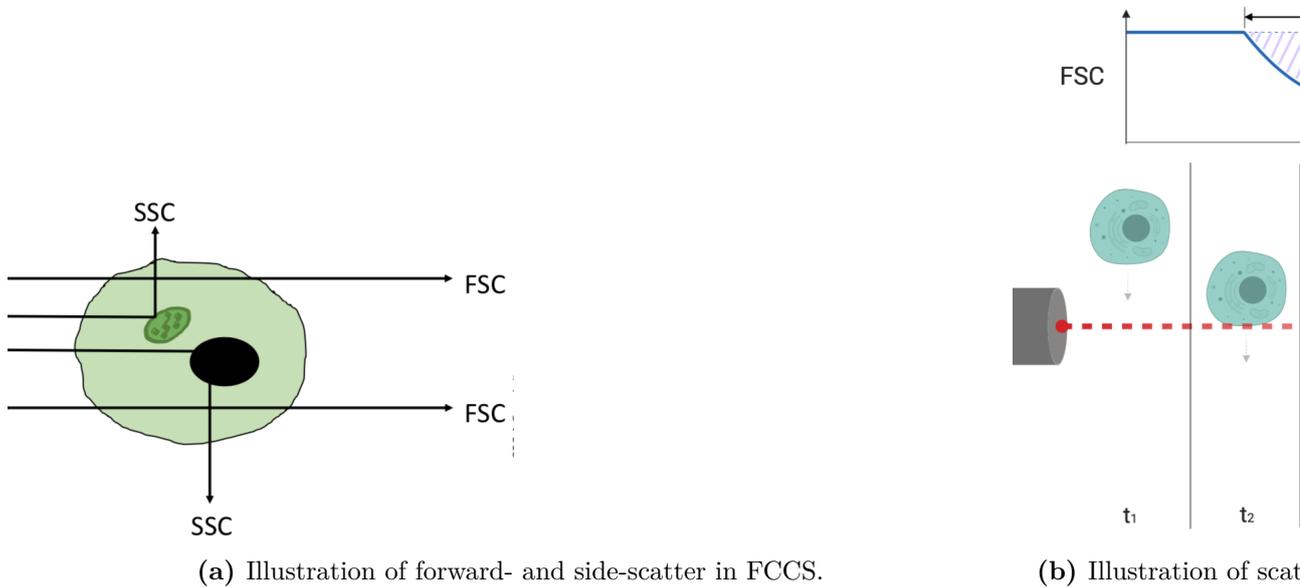
At the operator level, FCCS is conceptually identical to selecting a subset of data in a flow cytometry sample for further analysis. Although the additional practical complexity compared to flow cytometry is significant, a corresponding amount of 'information hiding' (encapsulation of technical details in order to provide a simplified user experience) takes place to obscure the real-time computing involved in the sorting process, and provide an interface with which flow cytometrists will be familiar.

Flow cytometry utilises a fluidics system to file cells into a stream, which is ejected from a nozzle and interrogated by a laser(s) and photodetector(s) at high frequency. This data is sent to a computer for storage and analysis. Cell sorting, by contrast, requires additional systems to be included in this apparatus, namely: acoustics to drive the stream into consistent-sized droplets; real-time processing to compare each droplet's data to criteria set electronically by the

operator; and electrostatics, to charge the stream as it emerges from the nozzle, and to electrically charge two plates such that the generated electric field deflects each charged droplets according to data provided by the real-time controller.



**Figure 4.2.: Annotated photograph of a FCCS machine (BD Influx).** Fluidic systems are represented in green; optical systems in orange; electrical/electronic components in cyan/blue respectively. Cells are loaded from the sample tube (A) into the nozzle (B) by positive pneumatic pressure, where the stream is hydrodynamically focussed with a secondary 'sheath' fluid, and subjected to ultrasonic stimulation by a piezoelectric element in the nozzle. The stream emerging from the nozzle is interrogated by a series of lasers (C) aligned by a series of mirrors (D) so as to meet perpendicularly at the interrogation point, marked with a yellow star. Hereafter, the stream is broken into drops, entrained by the periodic pressure field from the ultrasonic emitter. Transmitted and scattered light is filtered, detected (E), and electronically transmitted to a computer-connected realtime controller (F) which dynamically controls electric charge on the sorting plates (G) in order to deflect drops to a varying degree (as specified by the operator) in order that they align with various collection vessels placed at known locations in the collection area (H).

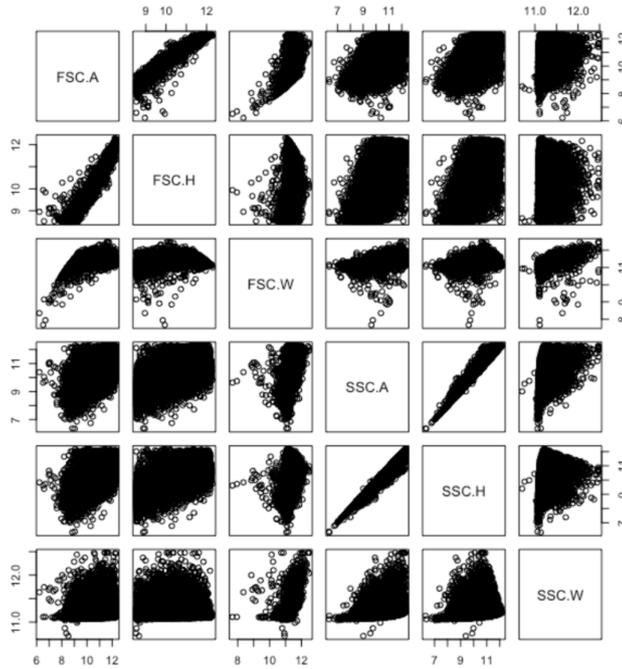


**Figure 4.4.: Illustrations of scatter in FCCS.** Forward-scatter (FSC) is used as a measure on cell size; side-scatter (SSC) is a measure on a combination of internal optical density, and cell surface roughness. For each of forward- and side-scatter (as well as for every fluorescence channel) the computer records the peak amplitude of scatter within the event (known as scatter 'height',  $H$ ), time duration of the scattering signal (scatter 'width',  $W$ ), and area under the curve of the detector's electronic signal (scatter 'area',  $A$ ). The latter is used as a measure on the total quantity of scattered light at low-angle (FSC) and high-angle (SSC) conditions.

#### 4.1.2. Complexity challenges in flow cytometry and cell sorting experiments

The “curse of dimensionality” as originally described by mathematician Richard Bellman in 1957, states the increasing difficulty of meaningful data analysis as more parameters are measured in a single experiment. As cytometers (and, by the same token, cell sorters) are developed to measure more and more optical parameters simultaneously, so the space in which different populations are characterised grows sub-exponentially: the number of comparisons that can be made follows the expression  $\frac{n(n-1)}{2}$  where  $n$  is the number of optical parameters collected for each datum. As such, the capacity of human operators to optimally sort and categorise FC data sets is being outstripped by a combination of the growth of optical parameters being collected, and the greater capacity of cytometer hardware

and modern computers to rapidly collect and store large volumes of data.



**Figure 4.5.: 2D facet plot of scatter data collected within an FC experiment.** No fluorescence is included in this plot: the data shown describe scattered light only. The metrics being compared are (diagonally from top left): forward-scatter area (FSC.A), height (FSC.H) and width (FSC.W); side-scatter area (SSC.A), height (SSC.H) and width (SSC.W), as described in Figure 4.4. Although this chart is symmetric about the primary diagonal, it is evident that this parameter space will expand rapidly as fluorescence data is included, since each fluorescence channel will add up to three optical parameters to search.

In the current state-of-the-art, establishing high-quality gates to run an effective flow cytometric cell sort requires in-depth understanding of the optical and mechanical behaviour of the machine, and the significance of various optical characteristic data acquired for each event, as well as how this interacts with the characteristics of the sample to be studied. This presents a significant barrier to entry for scientists unfamiliar with FCCS, who might not have access to resources enough to tackle the steep learning curve required to establish effective gating strategy, although they would otherwise find FC data and/or a FCCS experiment highly useful for their scientific enquiry. This issue is particularly noticeable in a case where a particular event of interest may only occur in a small fraction of cells at any one time.

### 4.1.3. Tackling repeatability and bias in FCCS results

Because FCCS gating is performed manually, there is an inherent risk of bias in the gating process: that is to say, whether the bias is conscious or unconscious, the operator often has a pre-conceived expectation of their results. Behavioural science studies conclude that such pre-conceptions can influence the results of an experiment significantly (Daniels and McCusker, 2010). Despite efforts to counter this with extensive meta-data reporting and post-hoc statistical tests, others show that operator bias in FCCS gating remains a significant issue (Yu et al., 2016). Similarly, the number of degrees of freedom in the gating of a multi-channel FACS sample tend to create disparity between the approaches (and hence the results) of experienced FCCS operators versus new users.

This issue is particularly relevant when considering the problem of establishing gates to distinguish between subsets of interest within a sample - in addition to the complexity associated with searching through numerous optical parameters by which to resolve two or more desired populations, it becomes difficult to independently justify particular gating strategies, particularly when examining 'meta-quality' parameters such as selectivity and yield. Hidden circular reasoning is an eminent hazard when conducting post-hoc tests in order to validate gate geometry, particularly when testing complex sets in a many-dimensional space, and as target events become a smaller fraction of the overall sample, number of degrees of freedom increases, and manual gating strategies become more heuristic in nature, the quality-control tests on these strategies become more and more difficult to identify. While computer assisted FCCS apparatus alignment and setup is increasing the robustness of results to systematic variance over time due to e.g. temperature changes, an ideal FCCS sort criteria would also be robust to intra- and inter-sample variability due to e.g. intrinsic noise in fluorophore expression or uptake. In practice this is rarely attainable except by the most experienced of operators, and often such robustness trades off against other desirable characteristics such as sample selectivity.

Our aim is to produce a tool integrating a theory-agnostic approach: that is to say, identifying distinct sub-populations within a data set without any prior knowledge of the shape of that data. A range of R packages are available which will provide suitable data processing functionality: our objective is to build a software interface to integrate this functionality into an intuitive flow cytometry data processing environment. We use the sorting of a nanodot-rich plant cell sub-sample as a test case, to generate robust, repeatable sort criteria, while presenting minimal complexity to the operator.

## 4.2. Methods

### 4.2.1. Importing data to *R*; pre-processing

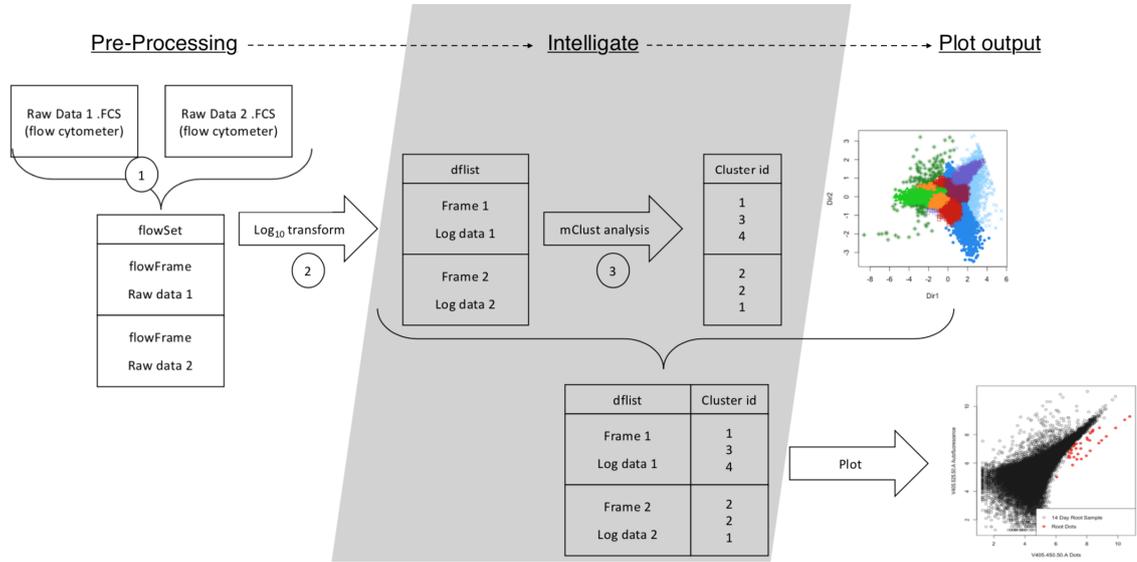
In this chapter, we use the open-source statistical analysis software package *R*, using *RStudio* as a development environment (Team, 2015, 2018). *RStudio* is a graphical scripting environment, which includes many high quality well-documented packages for various methods of data analysis. In this chapter we employ the *FlowCore* and *FlowViz* packages from the *Bioconductor* suite to respectively import and visualise flow cytometry data from the industry-standard FCS interchange file format (Huber et al., 2015). The data consisted of protoplast samples from whole *Arabidopsis* seedlings in pulse-chase experiments as detailed in 3.2.1. As shown in figure, this was first linearised via log-transform, allowing subsequent processes to be more resilient when considering sparse, skewed, or log-normally distributed data.

The dataset was also cleaned in *R* for out-of-range values: the machine's photo-detectors have 12 decades of dynamic range, which is calibrated to a standard as part of the machine set-up, and data falling outside that range (being either extremely bright or extremely dark) was discarded during this data-cleaning stage. Extremely 'dark' events have very high absorbance, so the machine records their low-angle scatter and fluorescence values as near-zero or negative, and the log-values in the data table are therefore either large-negative or not defined (since the logarithm for a negative number would be non-real). An opaque particle such as a protein aggregate would be a good example of a 'dark' event. Extremely 'bright' events, on the other hand, pass a very large quantity of light, typically across several low-angle scatter channels. These have values beyond the detectable maximum, and are recorded by the machine at the maximum value: hence, they are discernible as a number of events with the identical maximum value in one or more channels. A micro-bubble of air (or other transparent refractive material) would be a good example of a 'bright' event.

### 4.2.2. Interpreting output; tuning parameters

Once clusters have been assigned, we use *R* to plot the clusters in the statistical variance space, using a different colour for each cluster, as seen in the upper right corner of figure 4.6. The operator then selects one or more clusters as their subset of interest - these can then be highlighted as a subset of the data in a scatter-plot of any optical channels of the operator's choosing for validation. The operator

may adjust the sensitivity parameter  $\lambda$  if required, in order to generate results most valid according to their prior understanding of the system.



**Figure 4.6.: Illustrative data-flow for Intelligate analysis software.** Data from one or more samples is imported via the FCS interchange file format (1) at which point it is log-transformed (2) and pre-processed, removing events with NaN (“Not a Number”) values, and with marginal values beyond the dynamic range of the photodetectors (since these would skew the distributions).

Once the decision to proceed is made, the user must select a yield/purity coefficient  $\beta$ , which corresponds to the ratio between the size of the gate geometry and the standard deviation of the Gaussian model in each channel: so e.g. when  $\beta = 2$ , the gate geometry returned will have a diameter of two standard deviations in any direction, equitable to (approx.) 95% coverage of the modelled distribution. According to the parameter  $\beta$  selected by the operator, the tool then returns the gate geometry of the selected subset(s) of interest, consisting of an ellipse centred about the mean, with major and minor diameters in each fluorescence channel. The intention is to use a small fraction of the sample to analyse with *Intelligate*, using the gate geometry it establishes as a basis to sort the remaining majority of the sample.

## 4.3. Results

### 4.3.1. Using *Intelligate* to gate rare events in *Arabidopsis* protoplasts

*Intelligate* uses a dimension reduction algorithm from the *mclust* package (Scrucca et al., 2016) to re-parse the many-dimensional data set into co-ordinates of statistical variance and co-variance, building a Gaussian mixed-methods model of the data, and allocating the data to clusters on that basis. The sensitivity of the clustering algorithm can be tuned at run-time by a single numerical parameter,  $\lambda$ , which can be set to values between 0 and 1 to reflect a more homogenous or more granular output, respectively. The intention is to provide the end-user with a basic control to disallow over-fitting of sparse sub-sets (e.g. in a situation where two sparse subsets are initially identified as four tighter ones). We found that the fitting performance improved when we ran the mixed-methods model on its own output, and incorporated this approach into the workflow. Others have used similar techniques to group mammalian immune cells (Sørensen et al., 2015). Ultimately, our algorithm appends a category value to each point in the data set - as seen in figure 4.6. This value is an integer representation of the unique meta-cluster of which the point is a member.

We used *Intelligate* to analyse and cluster data from an experiment using protoplasts prepared by the method in Appendix 3.2.4 from *Arabidopsis* plants treated with carbon nanodots as in 3.2.1 on page 67. Using the tool, it is possible to isolate a small population of rare cells from a whole-tissue plant protoplast sample. The results, visible in Figure 4.7 on page 104 have gate geometry consistent with what has been established to be consistent over several months of sorting protoplasts from similarly treated seedlings of different ages.

### 4.3.2. Improving result robustness

*Intelligate* minimises complexity for the user: taking a space in many dimensions, and allowing it to be split using a single value. This minimises all possibility for operator bias to influence experimental outcome. The expression of the yield/selectivity trade off in terms of standard deviation is a transparent single value which data to be selected on a statistical basis *a priori*, rather than selecting a number of populations manually and performing post-hoc tests to verify their statistical validity. This is particularly useful where samples may have particular optical parameters that are distributed non-normally.

Additionally, we have transformed what is traditionally a many-step heuristic process into a single-step “wide” parameter search, simplifying the process, and providing a quasi-naïve statistical basis for gate geometry, rather than an empirically-derived basis.

### 4.4. Conclusion and further work

#### 4.4.1. Post-hoc experiments

Since this method allows the sorting of rare events within almost any population of cells, it could be very powerfully coupled with an experiment to determine expression data from those cells. Comparing this to the ‘baseline’ data from the same sample would provide information as to whether and how these rare events are related to changes in expression profile. This is particularly relevant where we have e.g. a large number protoplasts from an entire root, a small fraction of whom have interacted with cells from a symbiont micro-organism in the growth medium.

With a very little modification, it would be possible to divide the gated subset into quantiles of primary variance - in collaboration with expression data from those cells, this would provide the basis for a simple ‘black box’ type input-output model of e.g. the plant root interactome, allowing hypotheses obtained from in silico experiments to be tested empirically.

Experienced operators have an understanding of every sub-system within the FCCS method, and are able to tune parameters on the hardware for optimal results within a particular experiment: however, our software tool makes a clearer distinction between the setup and operation of the apparatus itself, and the establishment of metrics used for scientific enquiry. In combination with the computer-assisted alignment and setup protocols available in modern machines such as the BD FacsAria used in this study, our software allows a scientist new to FCCS with fundamental training on the machine’s operation to obtain FCCS results of a quality comparable to a competent regular user of the apparatus - lowering the experience-based barriers to entry between new users and the power of FCCS.

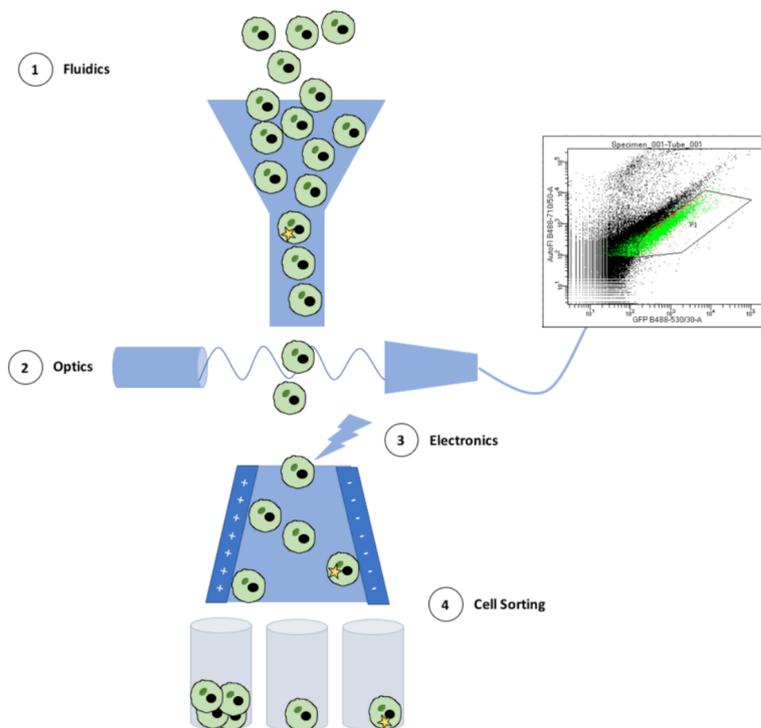
#### 4.4.2. Developing the *Intelligate* algorithm

**Parallel Computing** The aim is to develop *Intelligate* so that it may be executed in parallel - this would reduce the execution time significantly on computers

with mid-tier or better hardware (including most enterprise-specified computers). For a reasonable number of events, it should become possible to run the entire algorithm in a matter of seconds on a high-performance system (although this might require the repurposing of some resources from e.g. a high-performance GPU). Ideally, this could be executed semi-automatically, and integrated into FCCS software with an API for the most seamless user experience.

**Machine Learning approaches** Machine learning has garnered much attention in many analytical sectors, including bioinformatics. An approach based on a machine learning classifier algorithm would be interesting from a technical point of view, and might perhaps perform better for particularly odd-shaped clusters - but it is difficult to imagine designing and programming an effective metric for regression targets, whilst minimising the probability of over-fitting. Moreover, and apart from their often extended development time, and their requirement for training, machine-learning approaches tend to be processor-intensive, and it is arguable that at the current point in time, the extra resources spent on building and training a classifier instance for each experiment would not be justified by the benefits compared to e.g. a bi-clustering approach.

**Bi-Clustering** Similar to clustering, which is a numerical calculation to group data based on the similarity of their values, bi-clustering is a calculation which groups data on the basis of their co-variances. Although this requires that we initially generate a matrix whose size scales with the square of the number of data (and this is a relatively computationally expensive step) it allows us to reduce a high-dimension data set into that (large) 2D covariance matrix for the subsequent clustering step. This approach lends itself well to small data sets, and covariance scales well with large samples - we are taking a large quantity of data and clustering it based on a relatively small covariance matrix.



**Figure 4.3.: Overview of a FCCS apparatus.** Hydrodynamic focussing (1) takes a stream of suspended protoplasts (or single cells, polymer micro-beads, etc.) and aligns them in single file for interrogation by the optical system. The array of lasers, filters, and cameras (2) records optical parameters of each cell, which is plotted by the computer as a single dot on a scatter graph - the cumulative data is known as a 'dot plot'. The user then selects one or more populations of interest from the graph with a geometric selection known as a 'gate'. When a sort is started, the computer controls the electric field between two plates in the sort chamber (3) using electrostatic thrust to displace drops in one direction or according to their relation to the gate's optical parameters.

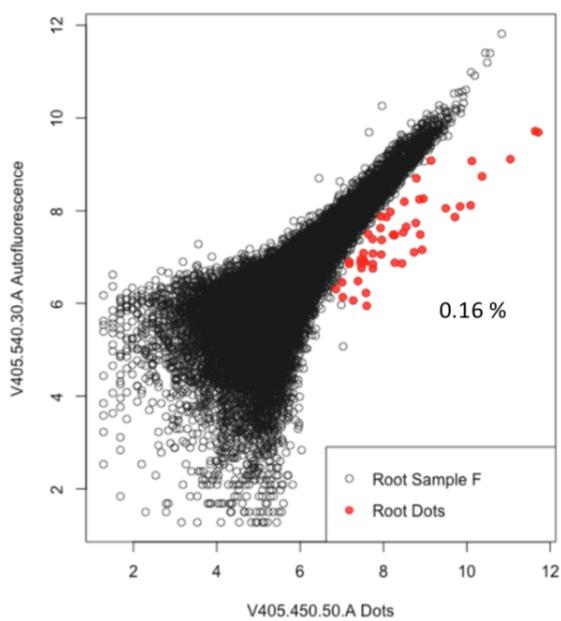
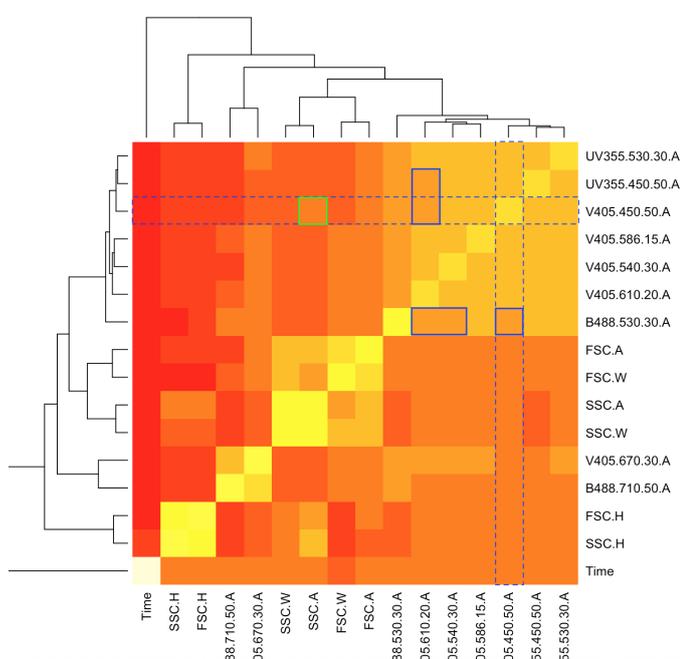


Figure 4.7.: CND-positive population of *Arabidopsis* protoplasts gated by *Intelligence* software (red dots) as a subset of 100,000 protoplasts sampled from whole root tissue (black circles). Proportion of events falling within the gate is annotated in percent %.



**Figure 4.8.: Heatmap of data covariance matrix from FC data obtained in this study.** Lighter squares indicate higher covariance. This is a visual representation of the basis used for biclustering model formation. Off-diagonal comparisons with unusually low covariance (darker colours) indicate physical data comparisons where clusters may be most distinct. Using a 405nm violet laser excitation (denoted “V405”) the primary channel used to detect nanodots (dashed outline) is a 50nm band-pass filter with a 450nm peak transmittance - thus denoted “450.50”. The heatmap shows increased covariance between this and the side scatter ‘area’ channel (green solid outline) which is consistent with the observation that c-dots have high optical absorbance. The heatmap shows decreased covariance (blue solid outlines) with 610.25 and with 530.30 - the latter being excited by the blue 488nm laser, as opposed to the 405nm violet - although these latter two channels also show low covariance with one another, indicating the possibility of a multi-factor interaction between all three channels.

# 5. Discussion

## 5.1. Summary

In Chapter 2, we set out the opportunity for the use of carbon nanodots as a delivery and tracing system for use in plant biology. We identified and performed a modular set of chemical reactions, capable of adapting the surface chemistry of the particle to accommodate a variety of subsequent surface functionalisations, including successfully adapting nanodots for use in a click reaction. We also successfully functionalised carbon nanodots with the flg-22 peptide, using only amide coupling reactions, and verified the structure of the functionalised nanodots using 2D NMR spectroscopy.

In Chapter 3, we examined the behaviour of nanodots in *Arabidopsis*, first outlining methods for nanodot application in solid and liquid media, and then examining the uptake of functionalised and un-functionalised nanodots into *Arabidopsis* roots. We found that both functionalised and un-functionalised nanodots reside predominantly in the root apoplast, and concluded that while un-functionalised nanodots may be small enough to move freely through pores in plant cell walls, some form of active transport was more likely to take place in the case of the larger functionalised particles. We also conducted flow cytometry experiments on arabidopsis roots treated with un-functionalised nanodots. Our observations revealed a small population of cells with a measurably elevated carbon nanodot fluorescence. Using this population, we examined the uptake dynamics of un-functionalised nanodots using a time-series flow cytometry experiment with a pulse-chase design. Finally, we examined the transcriptional effects of carbon nanodots on plant roots, using RNA-seq to examine the responses of treated roots to flg-22-functionalised carbon nanodots, as well as un-functionalised nanodots and flg-22 without nanodots. From our results, we concluded that there are no significant differences between the perception and response of root cells to flg22, and their perception and response to flg-22-functionalised carbon nanoparticles. When examining data from un-functionalised nanodots, we observed enrichment in a number of auxin transport and response genes, which may corroborate our speculation that transport proteins such as the PIN family may be involved in

the transport of larger nanoparticles.

In Chapter 4, we outline the work done developing the algorithm we used for our flow cytometry experiment, eventually finding that a Bayesian statistical meta-clustering approach was effective in conducting semi-automatic selection of rare events from mixed populations, in order to robustly identify the small population we first observed in Chapter 3.

## 5.2. Further Work

One of the more inspiring aspects of this project is the number of avenues of inquiry that remain open. As well as the questions that arise directly from the findings in this thesis, this project has identified several experiments and development projects that, if successful, may prove fruitful from a number of points of view.

For example, it is desirable to examine further the transcriptional response of plant cells to nanoparticles - this may help elucidate the mechanisms of transport, which is important for understanding the potential strengths and limitations of nanodots as a platform technology. A single-cell RNA-seq experiment, for example, could cross-reference responses to carbon nanodots with different cell type identification markers, identifying whether (and how) different cell types within the organ respond differently to the addition of the nanomaterial. This may also shed light on which proteins are involved in the transport of nanodots. Alternatively, FCCS could be coupled with RNAseq to separate populations with different concentrations of nanodots, thereby titrating out the treatment on a single-cell level, before examining the responses at each concentration. More experiments can be done using our existing stock of flg-22-functionalised nanodots, monitoring their progress through root tissues non-invasively using fluorescence imaging, and then sorting cells from different tissue types using FCCS, performing expression analysis to assess cellular responses to flg22 in different tissue types.

Similarly, using *Intelligate*, it would be possible to repeat the dynamics experiment performed with other varieties of nanodots - with some development, a coherent mathematical model could then potentially be used to estimate uptake dynamics on the basis of, say, average diameter of the nanoparticle as determined by AFM. Alternatively, the algorithm could be adapted to process data from imaging flow cytometry, and an experiment could be run to examine what intracellular localisation, if any, is exhibited by nanodots functionalised with various molecules.

We have shown success with one payload of interest to plant sciences, but there are many other molecules of interest when examining plant root science: synthetic nodulation factors, flavonoids, and other chemicals of interest in the rhizosphere are all yet to be studied. Our attempts to complete a click chemistry reaction on the surface of the nanoparticle were unsuccessful - but if this can be performed successfully, it potentially allows any or all available molecular payloads to be affixed in a single reaction. This may be a significant development effort, but could ultimately result in the co-functionalisation of the nanoparticle with multiple molecular payloads.

An equally interesting experiment from a fundamental systems biology point of view, following the methods outlined in this thesis, it should be relatively straightforward to attach a transcription factor active site to a nanoparticle, and then treat a plant with a GFP or GUS transgene under a native inducible promoter: this could be used to measure the ratio of input to output optically, and thereby begin to quantify the performance of the promoter in question. This would lead towards experiments with treatments much closer to their endogenous levels - which would begin to address issues surrounding external validity of experiments in the lab.

## **5.3. Conclusion**

At the beginning of this project - and of this thesis - we suggested carbon nanoparticles may become a platform technology for plant science, adding an extra dimension of observability to experiments, and potentially providing alternatives to time-consuming genetic transformation of plants. We have used a wide variety of techniques from different scientific disciplines in the attempt to provide sufficient fundamental investigation to constitute a proof of concept for this claim. As seen in the previous paragraph, much remains to be done: however, we believe this thesis provides enough evidence to justify an interest in the use and development of carbon nanoparticles in future plant science experiments, and that their incorporation into existing lines of scientific inquiry may begin to provide additional insight.

# A. *Intelligate* script

## A.1. Intelligate master script

```
# Install Software:
# source("http://bioconductor.org/biocLite.R")
# biocLite("flowCore")

### Initialise R: ###

## Load Libraries:
library(flowCore)
library(tidyverse) # for ggplot
library(mclust)
library(IDPmisc) # for "inf.omit()"
## Set the working directory:
# setwd("~/Documents/Data/Scripts/Intelligate")
## Include Intelligate core functions:
source("DataFrames.R")
source("mclust_dots.R")
source("FCoVar.R")

### Start workflow: ###

## Query path to FCS data:
liPrompt <- readline(prompt = "Enter name of folder containing
data for analysis: Data/")
fcPath = paste("Data/",liPrompt, sep = "")
print("OK! Working: ")
logExprs <- dataFrames(fcPath)
print("Complete.")
```

```
liPrompt <- userQuery("Begin mclust? ", default = "n")

if(liPrompt == "y"){
  print("OK, clustering:")
  clusterList <- mclustDots(logExprs)
}

liPrompt <- userQuery("Generate default plots? ", default = "n")

if(liPrompt == "y"){
  print("Plotting may take awhile. Please wait...")
  source("Figure_plotting.R")
  print("Done!")
}
```

## A.2. Data Import sub-script

```
dataFrames <- function(dataPath){

## Import and pre-process data: ##
# Load FlowFrames from path into FlowSet:
fs <- read.flowSet(path = dataPath, pattern = ".fcs",
  alter.names = TRUE, phenoData = list(name = "SAMPLE ID",
  Filename = "$FIL"))

## Create Raw DataFrames
# create empty list to hold each DataFrame
dfList <- list(length(fs))

print("Importing files...")

for (i in 1:(length(fs))) {
  dfList[[i]] <- data.frame(as.data.frame(
  na.omit(exprs(fs[[i]]))))
  #"na.omit()" handles NaN's
}
```

```
  for (j in 1:length(fs@colnames)) {
dflist[[i]] <- dflist[[i]][dflist[[i]][,j] != "-Inf",]
} # remove any "-Inf" values
}

print("Creating Log-transforms...")
# log-transform all channels (except time i.e. column 1)
fs_log <- transform(fs, transformList(fs@colnames[2:16], log))

# create list to hold each DataFrame, and name list items:
dflist_log <- list(length(fs_log))

print("Populating DataFrames...")
for (i in 1:(length(fs_log))) {
  dflist_log[[i]] <- data.frame(as.data.frame(
na.omit(exprs(fs_log[[i]]))) # delete inf/NaN values inline

  for (j in 1:length(fs_log@colnames)) {
    dflist_log[[i]] <- dflist_log[[i]][dflist_log[[i]][,j] > 0,]
# only keeping expression values greater than 0
  }
}

return(dflist_log)
}
```

### A.3. Primary clustering sub-script

```
mclustDots <- function (datalist_log){

  # Anonymous function gets data for the current experiment:
  dataFetch <- function (x) return (datalist_log[[x]])

  ## Initialise input variables:
  this_data <- dataFetch(1)
  events <- dim(this_data)[1] # number of events in the current FCS
```

```
channels <- dim(this_data)[2]
experiments <- length(datalist_log) # number of FCS's input
lambD <- 1 # Cluster tuning parameter (range ]0,1])

# plot nanodot fluorescence vs. green autofluorescence
plot(x = datalist_log[[8]][,"V405.450.50.A"],
     y = datalist_log[[8]][,"V405.540.30.A"],
     xlab = "log 405-450/50 nm", ylab = "log 405-540/30 nm")
# dot-positive cell subpopulation lies in high V405.450.50.A cluster

## Density Estimation:
# Produces a density estimate for each data point,
# using a Gaussian finite mixture model from Mclust
densE<- vector("list", experiments)
for (i in 1:experiments){
  densE[[i]] <- densityMclust(this_data[,c("V405.450.50.A",
                                           "V405.540.30.A")])
}

## Dimension reduction:
dimR <- vector("list",experiments)
for (i in 1:experiments){
  dimR[[i]] <- MclustDR(as.Mclust(densE[[i]]), lambda = lambD)
}

# Annotate with cluster number:
clustID <- list(experiments)
for (i in 1:experiments){
  clustID[[i]] <- c(dimR[[i]]$class)
}
# Column bind cluster_id to the datalist:
dots_cluster <- list(experiments)
for (i in 1:experiments){
  dots_cluster[[i]] <- cbind(clustID[[i]], datalist_log[[i]])
}

# Point density estimation via Mclust:
```

```
mod_dots2 <- list(events)
for (i in 1:events){
  mod_dots2[[i]] <- densityMclust(plot_dots[[i]][,c(
    "V405.450.50.A", "V405.540.30.A")])
}

## Validation plots:
# 3D perspective density plot
# plot(mod_dots2[[1]], what = "density", type = "persp", data =
# plot_dots[,c("V405.450.50.A", "V405.540.30.A")])
# 2D scatter plot
# plot(mod_dots2[[1]], what = "density", data =
# plot_dots[,c("V405.450.50.A", "V405.540.30.A")])

# Dimension reduction with MclustDR:
moddr_dots2 <- list(events)
for (i in 1:events){
  moddr_dots2[[i]] <- MclustDR(mod_dots2[[i]], lambda = lambD)
}

# Store the cluster IDs ----
cluster_id_dots2 <- list(length(mod_dots2))
for (i in 1:length(mod_dots2)){
  cluster_id_dots2[[i]] <- c(moddr_dots2[[i]]$class)
}
# Column bind cluster_id_dots2, datalist to isolate cluster(s)
dots_cluster2 <- list(experiments)

for (i in 1:experiments){
  dots_cluster2[[i]] <- cbind(cluster_id_dots2[[i]],
    plot_dots[[i]][,2:17])
}

return (c(dots_cluster, dots_cluster2))
}
```

## A.4. Covariance Heatmapping sub-script

```
FCoVar <- function (logExprs){  
  
  # Anonymous function: pull data for n-th experiment:  
  dataFetch <- function (n) return (logExprs[[n]])  
  
  # Get data from first FCS:  
  this_data <- dataFetch(1)  
  
  # Create covariance matrix:  
  FCov = cov(this_data)  
  FCor = cov2cor(FCov)  
  heatmap(FCor)  
  
}
```

## A.5. Plotting Sub-script

```
# Cluster plotting: ----  
# 2D scatter plot the cluster representing the subpopulation  
  
# save events corresponding to desired cluster in plot_dots  
plot_dots <- list(length(mod_dots))  
  
# 0 hour mock  
plot_dots[[1]] <- dots_cluster[[1]] [dots_cluster[[1]][,1]==8,]  
plot(plot_dots[[1]][, "V405.450.50.A"],  
plot_dots[[1]][, "V405.540.30.A"],  
xlab = "V405.450.50.A", ylab = "V405.540.30.A")  
  
# 0 hour treated  
plot_dots[[2]] <- dots_cluster[[2]] [dots_cluster[[2]][,1]==9,]  
plot(plot_dots[[2]][, "V405.450.50.A"],  
plot_dots[[2]][, "V405.540.30.A"],  
xlab = "V405.450.50.A", ylab = "V405.540.30.A")
```

```
# 24 hour mock
plot_dots[[3]] <- dots_cluster[[3]][dots_cluster[[3]][,1]==1,]
plot(plot_dots[[3]][, "V405.450.50.A"],
plot_dots[[3]][, "V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 24 hour treated
plot_dots[[4]] <- dots_cluster[[4]][dots_cluster[[4]][,1]==9,]
plot(plot_dots[[4]][, "V405.450.50.A"],
plot_dots[[4]][, "V405.540.30.A"],
xlab = "V405.450.50.A",
ylab = "V405.540.30.A")

# 48 hour mock
plot_dots[[5]] <- dots_cluster[[5]][dots_cluster[[5]][,1]==8,]
plot(plot_dots[[5]][, "V405.450.50.A"],
plot_dots[[5]][, "V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 48 hour treated
plot_dots[[6]] <- dots_cluster[[6]][dots_cluster[[6]][,1]==9 |
dots_cluster[[6]][,1]==5,]
plot(plot_dots[[6]][, "V405.450.50.A"],
plot_dots[[6]][, "V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 72 hour mock
plot_dots[[7]] <- dots_cluster[[7]][dots_cluster[[7]][,1]==3,]
plot(plot_dots[[7]][, "V405.450.50.A"],
plot_dots[[7]][, "V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 72 hour treated
plot_dots[[8]] <- dots_cluster[[8]][dots_cluster[[8]][,1]==7,]
plot(plot_dots[[8]][, "V405.450.50.A"],
plot_dots[[8]][, "V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")
```

```
# # Leaves
# 0 hour mock
plot_dots[[9]] <- dots_cluster[[9]][dots_cluster[[9]][,1]==8,]
plot(plot_dots[[9]][,"V405.450.50.A"],
plot_dots[[9]][,"V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 0 hour treated
plot_dots[[10]] <- dots_cluster[[10]][dots_cluster[[10]][,1]==9,]
plot(plot_dots[[10]][,"V405.450.50.A"],
plot_dots[[10]][,"V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 24 hour mock
plot_dots[[11]] <- dots_cluster[[11]][dots_cluster[[11]][,1]==1,]
plot(plot_dots[[11]][,"V405.450.50.A"],
plot_dots[[11]][,"V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 24 hour treated
plot_dots[[12]] <- dots_cluster[[12]][dots_cluster[[12]][,1]==9,]
plot(plot_dots[[12]][,"V405.450.50.A"],
plot_dots[[12]][,"V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 48 hour mock
plot_dots[[13]] <- dots_cluster[[13]][dots_cluster[[13]][,1]==8,]
plot(plot_dots[[13]][,"V405.450.50.A"],
plot_dots[[13]][,"V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 48 hour treated
plot_dots[[14]] <- dots_cluster[[14]][dots_cluster[[14]][,1]==9 |
dots_cluster[[14]][,1]==5,]
plot(plot_dots[[14]][,"V405.450.50.A"],
plot_dots[[14]][,"V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")
```

```
# 72 hour mock
plot_dots[[15]] <- dots_cluster[[15]][dots_cluster[[15]][,1]==3,]
plot(plot_dots[[15]][,"V405.450.50.A"],
plot_dots[[15]][,"V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# 72 hour treated
plot_dots[[16]] <- dots_cluster[[16]][dots_cluster[[16]][,1]==7,]
plot(plot_dots[[16]][,"V405.450.50.A"],
plot_dots[[16]][,"V405.540.30.A"],
xlab = "V405.450.50.A", ylab = "V405.540.30.A")

# Confirm this is the subpopulation by highlighting in original plot
plot(mod_dots[[1]], what = "density", data = dflist_log[[1]][,c(
"V405.450.50.A", "V405.540.30.A")])
points(plot_dots[[1]][,"V405.450.50.A"],
plot_dots[[1]][,"V405.540.30.A"], col = "red")

# Plot coloured clusters: ----
plot(moddr_dots[[1]], what = "scatterplot") # dark green
plot(moddr_dots[[2]], what = "scatterplot") # dark blue
plot(moddr_dots[[3]], what = "scatterplot") # blue
plot(moddr_dots[[4]], what = "scatterplot") # dark blue and maroon
plot(moddr_dots[[5]], what = "scatterplot") # dark green
plot(moddr_dots[[6]], what = "scatterplot") # orange and dark blue
plot(moddr_dots[[7]], what = "scatterplot") # green
plot(moddr_dots[[8]], what = "scatterplot") # maroon

# Plot clusters in different colours: ----
plot(moddr_dots2[[1]], what = "scatterplot") # blue and green
plot(moddr_dots2[[2]], what = "scatterplot") # not orange
plot(moddr_dots2[[3]], what = "scatterplot") # not red
plot(moddr_dots2[[4]], what = "scatterplot") # not maroon
plot(moddr_dots2[[5]], what = "scatterplot") # not l.blue or maroon
plot(moddr_dots2[[6]], what = "scatterplot") # not d.blue & d.green
```

```
plot(moddr_dots2[[7]], what = "scatterplot") # all
plot(moddr_dots2[[8]], what = "scatterplot") # not blue

# # Plot the subpopulation ----
plot(mod_dots[[1]], what = "density",
data = dflist_log[[1]][,c("V405.450.50.A", "V405.540.30.A")])
points(plot_dots2[[1]][, "V405.450.50.A"],
plot_dots2[[1]][, "V405.540.30.A"], col = "red")
```

# Bibliography

- Almagro, L., Ros, L. V. G., Belchi-Navarro, S., Bru, R., Barceló, A. R., and Pedreño, M. A. (2009). Class iii peroxidases in plant defence reactions. *Journal of Experimental Botany*, 60(2):377–390.
- Altenbach, D. and Robatzek, S. (2007). Pattern recognition receptors: From the cell surface to intracellular dynamics. *Molecular Plant-Microbe Interactions*, 20(9):1031–1039.
- Araújo, I. S., Pietsch, J. M., Keizer, E. M., Greese, B., Balkunde, R., Fleck, C., and Hülkamp, M. (2017). Stochastic gene expression in arabidopsis thaliana. *Nature Communications*, 8(1).
- Asano, T., Masumura, T., Kusano, H., Kikuchi, S., Kurita, A., Shimada, H., and Kadowaki, K.-i. (2002). Construction of a specialized cDNA library from plant cells isolated by laser capture microdissection: toward comprehensive analysis of the genes expressed in the rice phloem. *The Plant Journal*, 32(3):401–408.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., and et al. (2000). Gene ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29.
- Beck, M., Wyrsh, I., Strutt, J., Wimalasekera, R., Webb, A., Boller, T., and Robatzek, S. (2014). Expression patterns of flagellin sensing 2 map to bacterial entry sites in plant shoots and roots. *Journal of Experimental Botany*, 65(22):6487–6498.
- Berg, J. M., Tymoczko, J. L., and Stryer, L. (2007). *Biochemistry*. W. H. Freeman ; Basingstoke : Palgrave [distributor], New York, 6th ed. edition.
- BioRender.com (2020). Biorender.
- Birnbaum, K., Jung, J. W., Wang, J. Y., Lambert, G. M., Hirst, J. A., Galbraith, D. W., and Benfey, P. N. (2005). Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. *Nature Methods*, 2(8):615–619.
- Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W., and Benfey, P. N. (2003). A gene expression map of the arabidopsis root. *Science*, 302(5652):1956–1960.

- Burrell, T., Fozard, S., Holroyd, G. H., French, A. P., Pound, M. P., Bigley, C. J., Taylor, C. J., and Forde, B. G. (2017). The microphenotron: a robotic miniaturized plant phenotyping platform with diverse applications in chemical biology. *Plant Methods*, 13.
- Carpita, N., Sabularse, D., Montezinos, D., and Delmer, D. (1979). Determination of the pore size of cell walls of living plant cells. *Science*, 14(205):1144–7.
- Carter, A. D., Bonyadi, R., and Gifford, M. L. (2013). The use of fluorescence-activated cell sorting in studying plant development and environmental responses. *The International Journal of Developmental Biology*, 57(6-7-8):545–552.
- Chang, F.-P., Kuang, L.-Y., Huang, C.-A., Jane, W.-N., Hung, Y., Yue-ie, C. H., and Mou, C.-Y. (2013). A simple plant gene delivery system using mesoporous silica nanoparticles as carriers. *Journal of Materials Chemistry B*, 1(39):5279–5287.
- ChemAxon (2017). MarvinSketch v.17.1.2.
- Chen, J., Dou, R., Yang, Z., Wang, X., Mao, C., Gao, X., and Wang, L. (2016). The effect and fate of water-soluble carbon nanodots in maize (*zea mays* l.). *Nanotoxicology*, 10(6):818–28.
- Chinchilla, D., Bauer, Z., Regenass, M., Boller, T., and Felix, G. (2006). The arabidopsis receptor kinase fls2 binds flg22 and determines the specificity of flagellin perception. *The Plant Cell*, 18(2):465–476.
- Chis, O.-T., Banga, J. R., and Balsa-Canto, E. (2011). Structural identifiability of systems biology models: A critical comparison of methods. *PLoS ONE*, 6(11):e27755.
- Creissen, H. E., Jorgensen, T. H., and Brown, J. K. (2016). Impact of disease on diversity and productivity of plant populations. *Funct Ecol*, 30(4):649–657.
- Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M. R., Armean, I. M., Bennett, R., Bhai, J., Billis, K., Boddu, S., Cummins, C., Davidson, C., Dodiya, K. J., Gall, A., Girón, C. G., Gil, L., Grego, T., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O. G., Janacek, S. H., Juettemann, T., Kay, M., Laird, M. R., Lavidas, I., Liu, Z., Loveland, J. E., Marugán, J. C., Maurel, T., McMahon, A. C., Moore, B., Morales, J., Mudge, J. M., Nuhn, M., Ogeh, D., Parker, A., Parton, A., Patricio, M., Abdul Salam, A. I., Schmitt, B. M., Schuilenburg, H., Sheppard, D., Sparrow, H., Stapleton, E., Szuba, M., Taylor, K., Threadgold, G., Thormann, A., Vullo, A., Walts, B., Winterbottom, A., Zadissa, A., Chakiachvili, M., Frankish, A., Hunt, S. E., Kostadima, M., Langridge, N., Martin, F. J., Muffato, M., Perry, E., Ruffier, M., Staines, D. M.,

- Trevanion, S. J., Aken, B. L., Yates, A. D., Zerbino, D. R., and Flicek, P. (2018). Ensembl 2019. *Nucleic Acids Research*, 47(D1):D745–D751.
- Dalman, M. R., Deeter, A., Nimishakavi, G., and Duan, Z.-H. (2012). Fold change and p-value cutoffs significantly alter microarray interpretations. *BMC Bioinformatics*, 13(S2).
- Daniels, M. D. and McCusker, M. H. (2010). Operator bias characterizing stream substrates using wolman pebble counts with a standard measurement template. *Geomorphology*, 115(1):194 – 198.
- Darlington, A. P. S., Kim, J., Jiménez, J. I., and Bates, D. G. (2018). Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes. *Nature Communications*, 9(1).
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2012). Star: ultrafast universal rna-seq aligner. *Bioinformatics*, 29(1):15–21.
- Dodds, P. N. and Rathjen, J. P. (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. *Nature Reviews Genetics*, 11(8):539–548.
- Doyle, A., Lechler, R., and Turka, L. (2004). Organ transplantation: Halfway through the first century. *JASN; Journal of the American Society of Nephrology*, 15(12):2965–2971.
- Doyle, C., Higginbottom, K., Swift, T., Winfield, M., Bellas, C., Benito-Alifonso, D., Fletcher, T., Galan, M., Edwards, K., and Whitney, H. (2019). A simple method for spray-on gene editing in planta. *bioRxiv*.
- Dubrovsky, J., Doerner, P., Colón-Carmona, A., and Rost, T. (2000). Pericycle cell proliferation and lateral root initiation in arabidopsis. *Plant Physiology*, 124(4):1648–1657.
- Efroni, I. and Birnbaum, K. D. (2016). The potential of single-cell profiling in plants. *Genome Biology*, 17(1).
- Eulenberg, P., Köhler, N., Blasi, T., Filby, A., Carpenter, A. E., Rees, P., Theis, F. J., and Wolf, F. A. (2017). Reconstructing cell cycle and disease progression using deep learning. *Nature Communications*, 8(1).
- Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M. H., Lin, Y. H., Reid, D. E., and Gresshoff, P. M. (2010). Molecular analysis of legume nodule development and autoregulation. *J Integr Plant Biol*, 52(1):61–76.
- Fiers, M., Golemic, E., Xu, J., van der Geest, L., Heidstra, R., Stiekema, W., and Liu, C.-M. (2005). The 14–amino acid clv3, cle19, and cle40 peptides trigger

- consumption of the root meristem in arabidopsis through a clavata2-dependent pathway. *The Plant Cell*, 17(9):2542–2553.
- Freeman, T. (2015). Pepdraw.
- Freiherr von Boeselager, R., Pfeifer, E., and Frunzke, J. (2018). Cytometry meets next-generation sequencing – rna-seq of sorted subpopulations reveals regional replication and iron-triggered prophage induction in corynebacterium glutamicum. *Scientific Reports*, 8(1).
- Fu, P. (2013). Grand challenges in synthetic biology to be accomplished. *Frontiers in Bioengineering and Biotechnology*, 1.
- Gifford, M. L., Dean, A., Gutierrez, R. A., Coruzzi, G. M., and Birnbaum, K. D. (2008). Cell-specific nitrogen responses mediate developmental plasticity. *PNAS; Proceedings of the National Academy of Sciences*, 105(2):803–808.
- Gomez-Gomez, L. and Boller, T. (2000). Fls2: an lrr receptor-like kinase involved in the perception of the bacterial elicitor flagellin in arabidopsis. *Molecular Cell*, 5(6):1003–1011.
- Guo, W., Tzioutziou, N., Stephen, G., Milne, I., Calixto, C., Waugh, R., Brown, J., and Zhang, R. (2019). 3d rna-seq - a powerful and flexible tool for rapid and accurate differential expression and alternative splicing analysis of rna-seq data for biologists. *bioRxiv*.
- Hamant, O., Heisler, M. G., Jönsson, H., Krupinski, P., Uyttewaal, M., Bokov, P., Corson, F., Sahlin, P., Boudaoud, A., Meyerowitz, E. M., Couder, Y., and Traas, J. (2008). Developmental patterning by mechanical signals in arabidopsis. *Science*, 322(5908):1650–1655.
- Hao, Y., Yang, X., Shi, Y., Song, S., Xing, J., Marowitch, J., Chen, J., and Chen, J. (2013). Magnetic gold nanoparticles as a vehicle for fluorescein isothiocyanate and dna delivery into plant cells. *Botany*, 91(7):457–466.
- Hartmann, C. H. and Klein, C. A. (2006). Gene expression profiling of single cells on large-scale oligonucleotide arrays. *Nucleic Acids Research*, 34(21):e143.
- Havrdova, M., Hola, K., Skopalik, J., Tomankova, K., Petr, M., Cepe, K., Polakova, K., Tucek, J., Bourlinos, A. B., and Zboril, R. (2016). Toxicity of carbon dots – effect of surface functionalization on the cell viability, reactive oxygen species generation and cell cycle. *Carbon*, 99:238–248.
- Hebenstreit, D. (2012). Methods, challenges and potentials of single cell rna-seq. *Biology*, 1(3):658–667.
- Hebenstreit, D., Deonarine, A., Babu, M. M., and Teichmann, S. A. (2012). Duel

- of the fates: the role of transcriptional circuits and noise in cd4+ cells. *Current Opinion in Cell Biology*, 24(3):350 – 358. Nucleus and gene expression.
- Herrbach, V., Chirinos, X., Rengel, D., Agbevenou, K., Vincent, R., Pateyro, S., Huguet, S., Balzergue, S., Pasha, A., Provart, N., Gough, C., and Bensmihen, S. (2016). Nod factors potentiate auxin signaling for transcriptional regulation and lateral root formation in *medicago truncatula*. *Journal of Experimental Botany*.
- Hill, S. A., Benito-Alifonso, D., Davis, S. A., Morgan, D. J., Berry, M., and Galan, M. C. (2018). Practical three-minute synthesis of acid-coated fluorescent carbon dots with tuneable core structure. *Scientific Reports*, 8(1).
- Hill, S. A., Benito-Alifonso, D., Morgan, D. J., Davis, S. A., Berry, M., and Galan, M. C. (2016). Three-minute synthesis of sp<sup>3</sup>nanocrystalline carbon dots as non-toxic fluorescent platforms for intracellular delivery. *Nanoscale*, 8(44):18630–18634.
- Hill, S. A., Sheikh, S., Zhang, Q., Sueiro Ballesteros, L., Herman, A., Davis, S. A., Morgan, D. J., Berry, M., Benito-Alifonso, D., and Galan, M. C. (2019). Selective photothermal killing of cancer cells using led-activated nucleus targeting fluorescent carbon dots. *Nanoscale Advances*, 1(8):2840–2846.
- Himaja, A., Karthik, P., and Singh, S. (2015). Carbon dots: The newest member of the carbon nanomaterials family. *Chemical Record (New York, N.Y.)*, 15(3):595–615.
- Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., and et al. (2015). Orchestrating high-throughput genomic analysis with bioconductor. *Nature Methods*, 12(2):115–121.
- Iriti, M. and Faoro, F. (2007). Review of innate and specific immunity in plants and animals. *Mycopathologia*, 164(2):57–64.
- Jin, Y. and Gao, X. (2009). Plasmonic fluorescent quantum dots. *Nature Nanotechnology*, 4(9):571–576.
- Karlsson, J., Anguelova, M., and Jirstrand, M. (2012). An efficient method for structural identifiability analysis of large dynamic systems\*. *IFAC Proceedings Volumes*, 45(16):941 – 946. 16th IFAC Symposium on System Identification.
- Katari, M. S., Nowicki, S. D., Aceituno, F. F., Nero, D., Kelfer, J., Thompson, L. P., Cabello, J. M., Davidson, R. S., Goldberg, A. P., Shasha, D. E., Coruzzi, G. M., and Gutiérrez, R. A. (2010). Virtualplant: A software platform to support systems biology research. *Plant Physiology*, 152(2):500–515.

- Kerk, N. M., Ceserani, T., Tausta, S. L., Sussex, I. M., and Nelson, T. M. (2003). Laser capture microdissection of cells from plant tissues. *Plant Physiology*, 132(1):27–35.
- Kim, B. H., Kim, S. Y., and Nam, K. H. (2012). Assessing the diverse functions of BAK1 and its homologs in arabidopsis, beyond BR signaling and PTI responses. *Molecules and Cells*, 35(1):7–16.
- Kim, J., Darlington, A., Salvador, M., Utrilla, J., and Jiménez, J. I. (2020). Trade-offs between gene expression, growth and phenotypic diversity in microbial populations. *Current Opinion in Biotechnology*, 62:29 – 37.
- Kleber, R. and Kehr, J. (2006). Preparation and quality assessment of rna from cell-specific samples obtained by laser microdissection. *Arabidopsis Protocols*, 323:367–378.
- Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by crispr–cas9 leads to large deletions and complex rearrangements. *Nature Biotechnology*, 36(8):765–771.
- Kwok, R. (2010). Five hard truths for synthetic biology. *Nature*, 463(7279):288–290.
- Leach, J. E., Triplett, L. R., Argueso, C. T., and Trivedi, P. (2017). Communication in the phytobiome. *Cell*, 169(4):587 – 596.
- Li, H., Kang, Z., Liu, Y., and Lee, S.-T. (2012a). Carbon nanodots: synthesis, properties and applications. *Journal of Materials Chemistry*, 22(46):24230–24253.
- Li, J., Wen, J., Lease, K. A., Doke, J. T., Tax, F. E., and Walker, J. C. (2002). BAK1, an arabidopsis LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell*, 110(2):213–222.
- Li, K.-E., Chang, Z.-Y., Shen, C.-X., and Yao, N. (2015). Toxicity of nanomaterials to plants. *Nanotechnology and Plant Sciences*, pages 101–123.
- Li, S., Liu, Z., Ji, F., Xiao, Z., Wang, M., Peng, Y., Zhang, Y., Liu, L., Liang, Z., and Li, F. (2012b). Delivery of quantum dot-sirna nanoplexes in sk-n-sh cells for bace1 gene silencing and intracellular imaging. *Molecular Therapy—Nucleic Acids*, 1(4):e20.
- Li, S., Yamada, M., Han, X., Ohler, U., and Benfey, P. N. (2016). High-resolution expression map of the arabidopsis root reveals alternative splicing and lincrna regulation. *Developmental Cell*, 39(4):508 – 522.
- Liu, C., Zhang, P., Zhai, X., Tian, F., Li, W., Yang, J., Liu, Y., Wang, H., Wang, W., and Liu, W. (2012). Nano-carrier for gene delivery and bioimaging

- based on carbon dots with pei-passivation enhanced fluorescence. *Biomaterials*, 33(13):3604–3613.
- Lopez-Gomez, M., Sandal, N., Stougaard, J., and Boller, T. (2011). Interplay of flg22-induced defence responses and nodulation in lotus japonicus. *Journal of Experimental Botany*, 63(1):393–401.
- Loukanov, A., Sekiya, R., Yoshikawa, M., Kobayashi, N., Moriyasu, Y., and Nakabayashi, S. (2016). Photosensitizer-conjugated ultrasmall carbon nanodots as multifunctional fluorescent probes for bioimaging. *The Journal of Physical Chemistry C*, 120(29):15867–15874.
- Lozano-Durán, R. and Zipfel, C. (2015). Trade-off between growth and immunity: role of brassinosteroids. *Trends in Plant Science*, 20(1):12 – 19.
- Marchand, J. A., Neugebauer, M. E., Ing, M. C., Lin, C.-I., Pelton, J. G., and Chang, M. C. Y. (2019). Discovery of a pathway for terminal-alkyne amino acid biosynthesis. *Nature*, 567(7748):420–424.
- Marion, D. (2013). An introduction to biological nmr spectroscopy. *Molecular and Cellular Proteomics*, 12(11):3006–3025.
- Mestrelab (2020a). Mnova nmr.
- Mestrelab (2020b). Nmr predict.
- Morris, E. R. and Walker, J. C. (2003). Receptor-like protein kinases: the keys to response. *Current Opinion in Plant Biology*, 6(4):339–342.
- Moses, J. E. and Moorhouse, A. D. (2007). The growing applications of click chemistry. *Chem. Soc. Rev.*, 36(8):1249–1262.
- Munsky, B., Neuert, G., and van Oudenaarden, A. (2012). Using gene expression noise to understand gene regulation. *Science*, 336(6078):183–187.
- Mustafa, G. and Komatsu, S. (2016). Toxicity of heavy metals and metal-containing nanoparticles on plants. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1864(8):932–944.
- Nakazono, M., Qiu, F., Borsuk, L. A., and Schnable, P. S. (2003). Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: Identification of genes expressed differentially in epidermal cells or vascular tissues of maize. *The Plant Cell*, 15(3):583–596.
- Navarro, L., Zipfel, C., Rowland, O., Keller, I., Robatzek, S., Boller, T., and Jones, J. D. (2004). The transcriptional innate immune response to flg22. interplay and overlap with avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiology*, 135(2):1113–1128.

- Nemhauser, J. L. and Torii, K. U. (2016). Plant synthetic biology for molecular engineering of signalling and development. *Nature Plants*, 2:1–7.
- Nikolados, E.-M., Weiße, A. Y., Ceroni, F., and Oyarzún, D. A. (2019). Growth defects and loss-of-function in synthetic gene circuits. *ACS Synthetic Biology*, 8(6):1231–1240.
- Noguero, M. and Lacombe, B. (2016). Transporters involved in root nitrate uptake and sensing by arabidopsis. *Frontiers in Plant Science*, 7.
- Nolan, O., Perez, D., and Garry, P. (2002). Simultaneous measurement of multiple active kinase states using polychromatic flow cytometry. *Nature Biotechnology*, 20(2):155–162.
- Oldroyd, G. E. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat Rev Microbiol*, 11(4):252–63.
- Ostadhossein, F. and Pan, D. (2017). Functional carbon nanodots for multiscale imaging and therapy. *Wiley interdisciplinary reviews. Nanomedicine and Nanobiotechnology*, 9(3).
- Peng, Y., Chen, L., Li, S., Zhang, Y., Xu, R., Liu, Z., Liu, W., Kong, J., Huang, X., Wang, Y., Cheng, B., Zheng, L., and Li, Y. (2018). BRI1 and BAK1 interact with g proteins and regulate sugar-responsive growth and development in arabidopsis. *Nature Communications*, 9(1).
- Pinaud, F., Clarke, S., Sittner, A., and Dahan, M. (2010). Probing cellular events, one quantum dot at a time. *Nature Methods*, 7(4):275–285.
- Porter, J. R., Telford, W. G., and Batchelor, E. (2017). Single-cell gene expression profiling using facs and qpcr with internal standards. *Journal of Visualized Experiments*, (120).
- Pound, M. P., Fozard, S., Torres, M. T., Forde, B. G., and French, A. P. (2017). Autoroot: open-source software employing a novel image analysis approach to support fully-automated plant phenotyping. *Plant Methods*, 13(12).
- Reineke, J. E. (2012). *Nanotoxicity : methods and protocols*. Number 978-1-62703-002-1 in *Methods in molecular biology*,. Humana Press, New York, 1 edition.
- Rogers, E. D., Jackson, T., Moussaieff, A., Aharoni, A., and Benfey, P. N. (2012). Cell type-specific transcriptional profiling: implications for metabolite profiling. *The Plant Journal*, 70(1):5–17.
- Rounds, C. M., Lubeck, E., Hepler, P. K., and Winship, L. J. (2011). Propidium iodide competes with ca<sup>2+</sup> to label pectin in pollen tubes and arabidopsis root hairs. *Plant Physiology*, 157(1):175–187.

- Rubbens, P., Props, R., Boon, N., and Waegeman, W. (2017). Flow cytometric single-cell identification of populations in synthetic bacterial communities. *PLOS ONE*, 12(1).
- Saharan, V. and Pal, A. (2016). *Chitosan based nanomaterials in plant growth and protection*, volume 1 of *SpringerBriefs in Plant Science*. Springer International Publishing, New York, NY, 1 edition.
- Sawchuk, M. G., Edgar, A., and Scarpella, E. (2013). Patterning of leaf vein networks by convergent auxin transport pathways. *PLoS Genetics*, 9(2):e1003294.
- Schaumberg, K. A., Antunes, M. S., Kassaw, T. K., Xu, W., Zalewski, C. S., Medford, J. I., and Prasad, A. (2016). Quantitative characterization of genetic parts and circuits for plant synthetic biology. *Nat Methods*, 13(1):94–100.
- Scrucca, L., Fop, M., Murphy, T. B., and Raftery, A. E. (2016). mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, 8(1):205–233.
- Sharifi, R., Lee, S.-M., and Ryu, C.-M. (2017). Microbe-induced plant volatiles. *New Phytologist*, 220(3):684–691.
- Sörensen, T., Baumgart, S., Durek, P., Grützkau, A., and Häupl, T. (2015). immunoclust—an automated analysis pipeline for the identification of immunophenotypic signatures in high-dimensional cytometric datasets. *Cytometry Part A*, 87(7):603–615.
- Strano, J. P. G., Landry, M. P., Faltermeier, S. M., and S., T. P. M. N. M. I. A. A. B. N. F. R. A. J. H. F. S. J. A. B. M. (2014). Plant nanobionics approach to augment photosynthesis and biochemical sensing. *Nature Materials*, 13(5):530.
- Stringlis, I. A., Proietti, S., Hickman, R., Van Verk, M. C., Zamioudis, C., and Pieterse, C. M. J. (2017). Root transcriptional dynamics induced by beneficial rhizobacteria and microbial immune elicitors reveal signatures of adaptation to mutualists. *The Plant Journal*, 93(1):166–180.
- Sun, D., Hussain, H. I., Yi, Z., Siegele, R., Cresswell, T., Kong, L., and Cahill, D. M. (2014). Uptake and cellular distribution, in four plant species, of fluorescently labeled mesoporous silica nanoparticles. *Plant cell reports*, 33(8):1389–1402.
- Swift, T., Fagan, D., Benito-Alifonso, D., Hill, S., Yallop, M., Oliver, T., Lawson, T., Galan, M., and Whitney, H. (2019). Photosynthesis and crop productivity is enhanced by glucose-functionalized fluorescent carbon dots. *bioRxiv*.
- Swift, T. A., Duchì, M., Hill, S. A., Benito-Alifonso, D., Harniman, R. L., Sheikh, S., Davis, S. A., Seddon, A. M., Whitney, H. M., Galan, M. C., and et al. (2018).

- Surface functionalisation significantly changes the physical and electronic properties of carbon nano-dots. *Nanoscale*, 10(29):13908–13912.
- Team, R. (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.
- Team, R. C. (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, 418(6898):671–677.
- Torney, F., Trewyn, B. G., Lin, V. S.-Y., and Wang, K. (2007). Mesoporous silica nanoparticles deliver dna and chemicals into plants. *Nature Nanotechnology*, 2(5):295–300.
- Wang, J., Liu, X., and Shen, H.-W. (2017). High-dimensional data analysis with subspace comparison using matrix visualization. *Information Visualization*, 18(1):94–109.
- Wang, L., Wang, X., Bhirde, A., Cao, J., Zeng, Y., Huang, X., Sun, Y., Liu, G., and Chen, X. (2014). Carbon-dot-based two-photon visible nanocarriers for safe and highly efficient delivery of sirna and dna. *Advanced Healthcare Materials*, 3(8):1203–1209.
- Wang, P., Lombi, E., Zhao, F. J., and Kopittke, P. M. (2016). Nanotechnology: A new opportunity in plant sciences. *Trends Plant Sci*, 21(8):699–712.
- Wilson, M. E., Mixdorf, M., Berg, R. H., and Haswell, E. S. (2016). Plastid osmotic stress influences cell differentiation at the plant shoot apex. *Development*, 143(18):3382–93.
- Woodward, A. W. and Bartel, B. (2018). Biology in bloom: A primer on the arabidopsis thaliana model system. *Genetics*, 208(4):1337–1349.
- Wu, Y. F., Wu, H. C., Kuan, C. H., Lin, C. J., Wang, L. W., Chang, C. W., and Wang, T. W. (2016). Multi-functionalized carbon dots as theranostic nanoagent for gene delivery in lung cancer therapy. *Sci Rep*, 6:21170.
- Xi, W., Gong, X., Yang, Q., Yu, H., and Liou, Y.-C. (2016). Pin1at regulates pin1 polar localization and root gravitropism. *Nature Communications*, 7(1).
- Yadav, R., Tavakkoli, M., Xie, M., Girke, T., and Reddy, G. (2014). A high-resolution gene expression map of the arabidopsis shoot meristem stem cell niche. *Development*, 141(13):2735–2744.

- Yu, J. S., Pertusi, D. A., Adeniran, A. V., and Tyo, K. E. J. (2016). CellSort: a support vector machine tool for optimizing fluorescence-activated cell sorting and reducing experimental effort. *Bioinformatics*, 33(6):909–916.
- Zdunek, A., Adamiak, A., Pieczywek, P. M., and Kurenda, A. (2014). The biospeckle method for the investigation of agricultural crops: A review. *Optics and Lasers in Engineering*, 52:276 – 285.
- Zhou, J., Wang, X., Lee, J.-Y., and Lee, J.-Y. (2013). Cell-to-cell movement of two interacting at-hook factors in arabidopsis root vascular tissue patterning. *The Plant Cell*, 25(1):187–201.
- Zipfel, C. (2014). Plant pattern-recognition receptors. *Trends in Immunology*, 35(7):345–351.
- Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E. J., Jones, J. D. G., Felix, G., and Boller, T. (2004). Bacterial disease resistance in arabidopsis through flagellin perception. *Nature*, 428(6984):764–767.

# Afterword

On a personal note, there are many to whom I owe thanks for the completion of this project. Science can be a stressful occupation, and I remain humbled by the keenness with which people offer support, even when times are difficult:

To Tom Swift, who bought into this project from day one, and has supported it ever since, despite changing research groups along the way; Beatriz Lagunas, for sharing excellent scientific guidance and practical help, with equal generosity. To Beth Richmond, Luke Richards, Imke Sittel, and Cantug Bar, for pitching in their time and expertise when I needed it most, and far more so than I had any right to expect. To my family and friends, who have given their support throughout this project, without hesitation. And, of course, to my supervisors: Miriam Gifford, whose faith in this project was the reason it could become real in the first place, and whose patience and positivity has been endless; Heather Whitney, who invested a lot of effort into this project very early on; and Carmen Galan, who remains my chemistry oracle: I couldn't ask for better directors for such a project.