NTNU
Norwegian University of
Science and Technology

Investigating the prospect of using the yellow fluorescent protein mutant YFP3G as a pressure sensor in plants

Eskil Nodeland Indregard

MSc in Physics
Submission date: July 2018
Supervisor: Rita de Sousa Dias, IFY
Co-supervisor: Thorsten Hamann, IBI

Norwegian University of Science and Technology
Department of Physics

I would like to thank everyone!

Summary

Plants have no skeleton, but instead get their structural support from a hydrostatic pressure called turgor pressure pressing against a strong cell wall. This pressure can reach up to 2 MPa, and is also an important part in how plant cells grow. New research have also began to shed light on the possible role of turgor pressure in sensing cell wall damage and in the mechanisms maintaining the integrity of the cell wall. Current methods of measuring this pressure are not capable of measuring turgor pressure in vivo, without expensive time consuming techniques, and a better method is therefore needed. A pressure sensitive yellow fluorescent protein have recently been developed, and this project have therefore tried to test the feasibility of employing this protein as a turgor pressure sensor, as well as doing molecular dynamics simulations to try to better understand the conformational changes happening in the molecule with varying pressure. The results are inconclusive, with no significant difference in change measured between transgenic plants expressing the mutant yellow fluorescent protein versus plants expressing the wild type. The molecular dynamics simulations also did not produce any significant results

Preface

Write your preface here...

Table of Contents

List of Tables

List of Figures

Abbreviations

GFP	=	Green Fluorescent Protein
CWI	=	cell wall integrity
CWS	=	cell wall stress
AFM	=	atomic force microscopy
YFP	=	yellow fluorescent protein
YFP3G	=	a pressure sensitive version of YFP
ESPT	=	excited state proton transfer
MD	=	molecular dynamics
PDB	=	The Protein Data Bank
NMR	=	nuclear magnetic resonance
PME	=	particle mesh Ewald
NPT	=	a constant number of particles N, a constant temperature T, and
NVT	=	constant N, constant volume V and constant T
NVE	=	constant N, constant V and constant energy E
RMSD =	root mean square deviation	
MS	=	Murashige and Skoog basal medium
LB	=	Luria-Bertani medium
cDNA	=	complementary DNA
Promega	=	ImProm-II Reverse Transcription System
dsDNA	=	double stranded DNA
ACT2	=	ACTIN 2 gene
ROI	=	region of interest

Chapter 1

Introduction

A little over two decades ago, a small fluorescent protein found in the jellyfish Aequorea victoria began to spark interest. Discovered in 1961 and appropriately named Green Fluorescent Protein (GFP) due to its green fluorescence, it was mostly forgotten until the American molecular biologist Douglass Pracher began looking closer into the protein. He started sequencing the gene coding for the protein in 1988, and also proposed that it could be useful as a molecular tracer. When a few years later it was discovered that the protein was both remarkably stable and could fold correctly at room temperature, even without the presence of any cofactors native to the jellyfish, the idea of using the protein as a biological tool began to form. But although the wild type GFP could fluoresce in various environments, it also had a few drawbacks as a biological marker. This lead to a race in developing new mutated variants with improved photostability and quantum yield, or who could better handle certain environments. This work is still ongoing, and in recent years mutants with the capability of functioning as a temperature or pressure sensor have been published, allowing the protein to be used in entirely new ways.

One of the fields were this has opened up new possibilities is in plant biology. In plants, osmosis creates an internal pressure called turgor pressure. This pressure is pushing the plasma membrane against the cell wall, creating the rigidity of the plant structure. Recent studies have also linked turgor pressure to the cell wall integrity (CWI) mechanism, a mechanism responsible for monitoring the condition of the cell wall, as well as initiating responses to cell wall stresses. These stresses can arise both from damage to the cell, as well as cell growth and division, and so, measuring the turgor pressure in vivo can give new insight into how it affects growth and development of the cell. Unfortunately, current methods of measuring turgor pressure are either unable to measure the pressure in vivo or very difficult to use, with the most common method being the pressure probe. This method gives precise measurements and the cells to be analysed must be accessible close to the surface. In the atomic force microscopy (AFM) method, this is avoided, and instead tapping the cell with an AFM is combined with knowledge of the cell structure and used to calculate the pressure. But as this requires very detailed knowledge of the cell, the measurements take longer and can only give results for the epithelial cells, this too has its

disadvantages. The goal of the project presented here is therefore to develop a novel turgor pressure sensor based on targeted modification of a yellow fluorescent protein.

Chapter 2

Background

Unlike animals, plants have no internal skeleton consisting of nearly rock-hard cells maintaining the structure of the organism. Instead, it is the internal pressure of the plant cells and the strength of the cell wall that maintains the form of the plant (Hamant and Haswell, 2017). This hydrostatic pressure can in plants reach up to $2\,\mathrm{MPa}$ (Beauzamy et al., 2014), higher than in a car a tire, so in a sense, a plant is built up by many tiny car tires-bricks glued together. While this may be the most important function of the turgor pressure, it is also involved in several other important cell functions. The most researched of these is the important part it plays in cell growth, but resent studies have also hinted to an important role in detecting cell wall damage and in the cell wall integrity (CWI) maintenance mechanism. This is a mechanism in plants with the purpose of maintaining the integrity of the cell wall and induce responses in the cell to cell wall stress (CWS) (Hamann, 2015). This stress can be defined as any type of damage to the functional integrity of the cell wall, either it is coming from a herbicide, disease or abiotic stress such as drought. To understand why turgor pressure might change in a plant, and what roles it can play in the cell, some basic knowledge of the composition and structure of the plant cell, is needed.

2.1 Structure of the plant and the plant cell

The plant cell is like an animal cell composed of a plasma membrane that separates the various interior compartments from the outside environment. But in addition, the plant cell is surrounded by either one or two cell walls. Actively growing plants only have the primary cell wall, while in some cells, a secondary cell wall is developed inside the primary for extra strength (Cosgrove, 2005). The primary cell wall is a thin layer composed of cellulose microfibrils embedded and cross-linked in a matrix of hemicelluloses and pectins (Wolf and Greiner, 2012), which together combine both high tensile strength and flexibility.

When the cell is expanding and growing, initially, a loosening of the cell wall will occur. This allows the internal turgor pressure to expand the cell, effectively lowering the pressure (Zonia and Munnik, 2007). The osmotic difference across the plasma membrane that maintained the original turgor pressure have at the same time not changed. More water

will therefore move into the cell and either cause even further expansion or a build up of the pressure levels. The water moving into the cell is kept in a special compartment of the cell called the vacuole, which is a large intracellular organelle counting for up to 80 % of the cell volume (Haswell and Verslues, 2015). The vacuole has many different functions, one important of which is to main maintain the turgor pressure of the cell. This is done by creating an osmotic pressure difference across the vacuole membrane by the active uptake of ions (Hedrich, 2012).

The exact amount of the volume of the cell the vacuole makes up, is depending on the type of cell. The plant is made up of an outer layer of epidermal cells, with a layer of palisade and spongy mesophyll cells beneath (Xiao et al., 2016). The epidermal cells consists of stomatal guard cells and pavement cells. The stoma is an opening in the leaf allowing for gas to pass, while the pavement cells are jigsaw puzzle shaped cells constituting the main part of the epidermis (Nadeau and Sack, 2002). The palisade cells and spongy mesophyll constitutes the mesophyll beneath the epidermal layer, and contain the chloroplasts, the organelles containing chlorophyll.

2.2 Osmosis and turgor pressure

In contrast to animal cells, the plant cell is under constant high stress, coming from the internal turgor pressure. This hydrostatic pressure is a result of the osmotic difference between the vacuole, cytosol, and the surrounding environment. The theory behind osmosis is not one easily explained, so this section will only go so far as to say that a concentration difference of solutes across a semi-permeable membrane creates a flow of solvent molecules towards the side of the membrane with a higher solute concentration. If this side of the membrane is a sealed compartment, a hydrostatic pressure will build up. For a more thorough explanation, see Borg (2003).

The flexible nature of the primary cell wall means that an increase in the external level of solutes will lead to a lowering of the turgor pressure and thereby cause the cell to shrink. This causes the plant to wilt and lose its ability to keep itself upright, as well as preventing the other functions turgor pressure have in the cells. To combat this, plant cells have developed the ability to further increase their internal ion concentration if the environmental concentration is increased and thereby recover the turgor pressure (Shabala and Lew, 2002).

2.3 Yellow fluorescent protein

The yellow fluorescent protein (YFP) is a mutated version of the green fluorescent protein, GFP. The GFP is an 11-stranded β-barrel with each strand containing approximately 9-13 amino acids, as seen in Figure 1(TO DO). These strands are wrapped around a central helix containing a chromophore, the part of the protein responsible for the absorption and emission of visible light (Merabti et al., 2017). It is protruding from the middle of the central helix almost perpendicularly (60 degrees) to the direction of the barrel. The chromophore is formed spontaneously from the residues Ser65-Tyr66-Gly67, in a process beginning with the autocatalytic cyclization of the Ser65 and Gly67 residues, going through an eno-

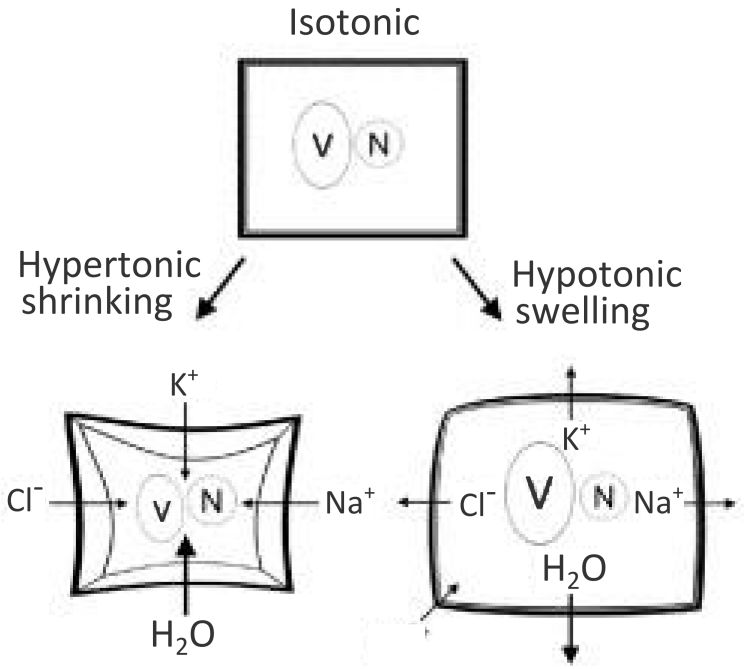


Figure 2.1: Modified from (Zonia and Munnik, 2007)

late intermediary stage, and ending with the oxidation of the $\alpha - \beta$ bond of Tyr66 Ormö et al. (1996). This process is happening under the influence of the Arg96 and Glu222 residues, with the arginine helping to stabilize the enolate and the glutamine acting as a general base for the deprotonation of the Tyr66. The encapsulation of the chromophore by the β-barrel means that it is shielded from the bulk water, presumably giving the protein its high quantum yield, low photo quenching and high resistance to pH titration.

The absorption spectrum of the GFP comes with two different absorption peaks, a main peak at 395 nm and a smaller peak at 475 nm. The two absorption peaks are attributed to two different ground states of the chromophore, an A-state attributed to the 395 nm peak and a B-state attributed to the 475 nm, with movement between the two states going through an intermediate state I (Ormö et al., 1996). The difference between the two ground states is caused by the protonation of the chromophore's phenolic ring (see Figure 2.3), with the protonated state, the A-state, being the most stable (Brejc et al., 1997). Absorption of a photon by this neutral state will lead to an excited state proton transfer (ESPT) via a hydrogen bonding network to a proton acceptor, believed to be either Ser205 or a combination of Glu222 and Thr203 (Shu et al., 2007). After the ESPT, the chromophore is now in the excited I*-state, which decays down to the I-state while emitting fluorescence with wavelength of 508 nm (Chattoraj et al., 1996). The anionic B-state, on the other hand, decays directly down from the B*-state, also producing an emission of fluorescence at 508 nm.

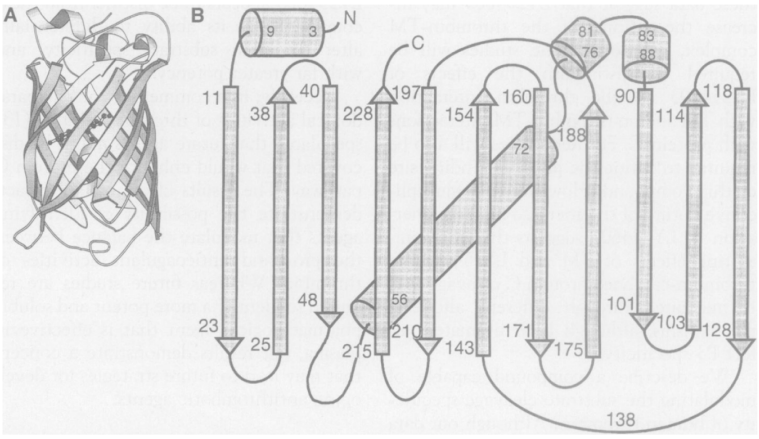


Figure 2.2: Schematic drawing of the backbone of GFP. Picture from Ormö et al. (1996)

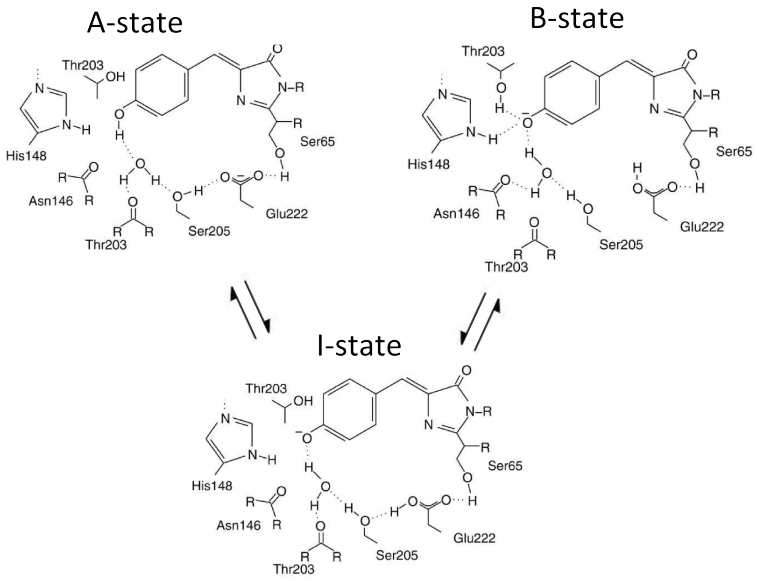


Figure 2.3: Model of the three states of the GFP chromophore. Hydrogen bonds are represented with dotted lines. Picture modified from Jung et al. (2005).

As is shown in Figure 2.3, a complex network of hydrogen bonding between the chromophore and the surrounding amino acids is involved in the absorption process. Mutations in these residues can therefore influence the fluorescence spectrum and change the equilibrium between these states. Out of these possible mutations, those introducing an aromatic ring in the place of Thr203, as well as exchanging Ser65 with either Gly or Thr, produce a red-shifted emission spectrum and have been named YFP (Tsien, 1998). The YFP used in this study has four mutations compared to the wild type GFP, one in the chromophore

and three in the β-barrel. The mutation of Thr203 is a T203Y mutation, i.e. introducing a tyrosine close to the chromophore. This allows π-π stacking with the phenolic ring of the chromophore, causing a significant redshift of both the excitation and emission maxima of the protein (Ormö et al., 1996). The mutation in the chromophore is an S65G mutation. This mutation removes the 395 nm absorption peak of the chromophore, indicating that the hydrogen bond between the hydroxyl group of Ser65 and the uncharged carboxyl of Glu222 is essential for the hydrogen bonding network supporting the protonated A-state (Jung et al., 2005). The other two mutations in the β-barrel, V68L and S72A only improve the folding of the protein at 37 degrees, but do not substantially alter the spectra Ormö et al. (1996). All in all, these mutations move the now singular absorption peak to 513nm and the emission maxima to 527nm.

2.4 A glycine insertion makes yellow fluorescent protein sensitive to hydrostatic pressure

The paper forming the foundation for this thesis is the 2013 study by Watanabe et al. (2013) called "Glycine insertion makes yellow fluorescent protein sensitive to hydrostatic pressure". In this paper, the authors present a pressure sensitive version of YFP, called YFP3G. The name for the new protein comes from the mutation used to create the pressure sensitivity, namely inserting the insertion of three glycine molecules into the β-barrel. The three new residues were inserted between Asn144 and Tyr145, as this is close to where the chromophore is nearest the β-barrel. The hypothesis of the group was that allowing the chromophore to be more exposed to the outside of the barrel by making β-strand 7 more disordered would increase the pressure dependency. This is because the interaction between water molecules and the chromophore of small fluorescent proteins is known to quench fluorescence and that this effect is enhanced by increasing pressure (Zhang et al., 1985). Regular YFP will also begin to decrease in fluorescence after around 100 MPa, with current theories suggesting water permeation of the β-barrel to be the cause (Jacchetti et al., 2016). YFP3G has in contrast an increased fluorescence intensity with higher pressure, and therefore some other process must work against the increased water-chromophore interactions at higher hydrostatic pressure levels. Watanabe et al. (2013) speculates that the increased number of water molecules seen in the mutated protein might be expelled at higher hydrostatic pressure. It has been shown that proteins decrease in volume when the pressure is increased (Jacchetti et al., 2016), which might explain why water molecules would be expelled from the expanded cavity near the chromophore. There are also other changes done by the mutation, such as a blue shift in the fluorescence spectrum compared to the YFP and then a gradual red shift with higher pressure.

Chapter 3

Molecular dynamics simulation

In addition to trying to establish the feasibility of using YFP3G as a turgor pressure sensor in plants, this study also aims to understand the physical processes behind the proteins pressure sensitivity. To examine this, the molecular dynamics (MD) package GROMACS has been used. Molecular dynamics is a method for simulating how atoms and molecules move and interact over time by computationally solving the forces arising between them (Hospital et al., 2015). Todays improved computational power allow bigger systems to be simulated for longer than ever before, and systems simulating for instance ligand binding can now be performed [Hospital]. MD is therefore considered to be a valuable tool in biology, where the use ranges from complementing experiments to exploring processes that are not yet fully understood and cannot be examined with other methods, and in this way generate new knowledge (Hansson et al., 2002)(Meinhold et al., 2007).

GROMACS is an MD package primarily designed for simulating proteins, lipids and nucleic acids and is, according to the GROMACS team, the most efficient MD program (GROMACS, 2016). The following chapter will be an overview of the methodology used for performing the MD simulations, and at the same time describing the algorithms used and calculations done by GROMACS. To begin with, a short introduction to the MD algorithm is given, before being followed by a discussion on the initial preparations of the molecules to be simulated. This is then followed by an explanation of the periodicity of the simulation box as well as the atomic interactions. After this, the system conditions, such as temperature and pressure, are presented, before the simulation run itself and how the analysis might be done is covered. There are also some final remarks on the approximations used in GROMACS and the limitations of MD simulations. The source for the following chapter is the GROMACS manual, if not otherwise is specified.

3.1 The global MD algorithm and the basics

In short, a molecular dynamics program takes a set of particles and their positions, gives them an initial velocity and then calculates the forces acting on each particle. These forces are then the basis for updating the velocities and the positions of the particles simultane-

ously when moving one time step forward. Newtons equations of motion are the basis for these calculations. This is then repeated for a set number of steps, giving a time evolution of the system and a trajectory for each particle. The particles to be simulated are put in a box and in most cases, including this study, periodic boundary conditions are used. This means that the simulation box, or simulation cell, is surrounded by translated copies of itself with no boundary in between.

The forces in the algorithm are calculated by force fields. These are sets of parameters and functions determining how the potential energy between the particles of the simulation is to be described, i.e. the interactions between them, and thereby the forces. The potential energy is partly determined by the topology of the simulation, which, among a few other things covered in section 3.4, is what kind of atom each particle is and their specific position inside their molecule, given by their atom type. This means that a carbon atom interacts differently with its neighbouring particles depending on whether it is defined as for instance an sp2 aromatic carbon or an sp3 carbon. The long-range interactions between particles that are not bonded are also determined by the atom types and are calculated pairwise. To save computational power, a cut off radius, meaning a predetermined limit to how far apart two particles can be and still have their long-range interactions calculated, is employed.

3.2 The PDB file and structural modifications

The basis for any MD simulation is the structural data of the molecules that are to be examined, i.e. the positions of the atoms and how they connect, which is provided by a text file format called The Protein Data Bank (PDB). There exist databases providing PDB files for large amounts of macromolecules with data derived from for instance x-ray diffraction or nuclear magnetic resonance (NMR) studies, for instance the RCSB Protein Data Bank. The PDB file can then be read by GROMACS and used to create the topology and the initial coordinate file for the simulation.

The first thing to be considered when the using MD in a study is what kind of system needs to be designed. The PDB files can give the structures of many different macromolecules, but the system created by these needs to be as similar to the real-life conditions of the process one wants to study as possible. For a study on a ligand binding to a membrane protein, this does not mean that the entire cell needs to be part of the simulated system, but that an area of a membrane containing the membrane protein should be designed (Murail, 2017). Because of the periodic boundary conditions, the smaller area simulated will in practice become infinite. Often the structural data in the PDB files needs to be modified before a simulation. This can be because the macromolecule is incomplete, as proper positional data for flexible and highly mobile parts of molecules can be difficult to derive through experimental means. In this case, the structure will need to be completed by adding what is missing in a molecular viewing software, for instance PyMOL or Chimera. In addition, many structures can include solvent molecules not necessarily present, or even binding proteins used to crystallize the main structure, which will have to be removed before moving on.

3.3 Periodic boundary conditions

In a simulation of finite size, some of the particles in the simulation will have to interact with the boundary of the simulation box. This will of course create unwanted edge effects on the system, so to avoid this problem GROMACS use periodic boundary conditions. This implies that the simulation box, or the *unit cell*, is bordered by translated copies of itself called *images*, making the system infinite and with no proper border. A particle exiting one side of the unit cell will only emerge at the opposite side and interact with the particles there. This removal of edge effects is at the same time bringing in new problems with the periodicity of the system. These effects become smaller with an increased system size, and it is therefore recommended to test different box sizes to test for unwanted effects.

The periodicity also means that the cut off radius for long-range interactions is limited to be less than half of the shortest vector defining the unit cell. Otherwise, a particle at the centre of the unit cell could interact with a particle at the edge of the box twice at the same time, leading to effects like charge build-up at the boundaries and wrong energy levels in the simulation.

3.4 The force fields

The force fields used in MD is not part of the simulation package itself, but developed by independent teams. This also means that there are a few differences between the force fields and their conventions, and the MD packages must therefore be made compatible with the force fields. The various force fields are usually developed to be as accurate as possible for a given system, for instance the OPLS/AA force field was originally developed for solvated organic molecules. Other popular force fields include AMBER, CHARMM, and GROMOS. The main difference between the force fields stem from the parameters they use in describing the particle interactions. These parameters are developed from either high-level quantum-mechanical calculations or fitting to experimental data (Gonzáles, 2011), and are the values the force field use when it calculates the potential energies of the system. The other half of a force field, the equations called potential energy functions, are used in these calculations and they are for the most part similar when comparing force field. These functions are also divided into two main parts, non-bonded and bonded interactions, which combined give the total potential energy experienced by an atom. In addition to these potential functions, it is also possible to add restraints on the particles. This is often implemented as an energy penalty and can for example be position restraints keeping an atom fixed, or angle restraints on the angle between two particle pairs.

3.5 Non-bonded interactions

Non-bonded interactions are the interactions between the particles of the simulation that are not covalently bound and have a distance between them that is less than the cut off radius. As a result of this, the non-bonded interactions are calculated from a constantly updated neighbour list containing all the atoms within the cut off radius. These non-bonded interactions are divided into three terms, a repulsive term, an attractive term and a

Coulomb term covering charged interactions.

$$V_{non-bonded} = (V_{repulsion} + V_{attaction}) + V_{Coulomb}, \tag{3.1}$$

The repulsion and attractive term gives rise to the van der Waals forces, and are combined in either the Lennard-Jones or the Buckingham potential. The main difference between them when choosing which potential to use is that the Buckingham potential has a more flexible and realistic repulsion term, but it is also more expensive to compute than the Lennard-Jones potential. The function describing the Buckingham potential between particle i and j has the form

$$V_{BH}(r_{ij} = A_{ij}e^{-B_{ij}r_{ij}} - \frac{C_{ij}}{r_{ij}^6}, \tag{3.2}$$

with A_{ij},B_{ij} and C_{ij} parameters dependent on the atom types of the two atoms, and r_ij being the distance between them. In comparison, the Lennard-Jones potential has the form

$$V_{LJ}(r_{ij} = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^6}, \tag{3.3}$$

with $C_{ij}^{(12)}$ and $C_{ij}^{()}$ also being atom type dependent parameters. The cut off radius can either be implemented with an abrupt cut off or with the use of a shift function, shifting either parts of or the whole potential to be zero at the cut off radius. In any case, the use of a cut off radius will lead to some errors, but in GROMACS there is the possibility to add corrections to both the energy and pressure calculations.

The Coulomb interaction between two charged particles are given by

$$V_C(r_{ij} = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{\epsilon_r r_{ij}}, \tag{3.4}$$

where ϵ_r is the relative permittivity, ϵ_0 is the vacuum permittivity and q_i and q_j are the charges of the particles. When calculating the electrostatic potentials, considering only one image of the charged particles in the system may not be accurate enough, as the Coulomb interaction energy decays a lot slower than the van der Waals interaction energy. To solve this, GROMACS uses techniques that sums over the unit cell and all the infinite images, such as the Ewald sum and the particle mesh Ewald (PME) method. In GROMACS, the version of PME that is implemented uses cardinal B-spline interpolation, or smooth PME (Essmann et al., 1995). PME can also be used with the Lennard-Jones potential to avoid truncation errors in stead of the abrupt cut off or shift function mentioned before.

3.6 Bonded interactions

In contrast to the non-bonded interactions, the bonded interactions are calculated from a fixed list of atoms and are not always calculated between atom pairs but include interactions between 3 and 4 atoms. These interactions are between atoms that are covalently

bound to each other, and the fact that this is a fixed list means that covalent bonds cannot be broken and new bonds cannot be created while the simulation is running (Gonzáles, 2011). The potential energy function for a bonded atom is given by three different interactions, the *bond stretching* between 2 atoms, the bond angle of 3 atoms, and the *dihedral angle* of 4 atoms.

$$V_{bonded} = (V_{bond} + V_{angle}) + V_{dihedral}. \tag{3.5}$$

The bond stretching is usually represented by a harmonic potential on the form

$$V_{bond}(r_{ij} = \frac{1}{2}k^b_{ij}(r_{ij} - b_{ij})^2, \tag{3.6}$$

where r$_i j$ again is the distance between particle i and j, b$_i j$ is the equilibrium distance and k$^b_{ij}$ is the harmonic force constant. GROMACS also includes other potentials, such as the Morse potential and the cubic bond stretching potential for systems where it is important to use an anharmonic potential to get a more accurate representation.

The bond angle interactions cover the increased energy of bending the angle formed by 3 atoms away from its equilibrium. It is also usually represented by a harmonic potential,

$$V_{angle}(\theta_{ijk} = \frac{1}{2}k^\theta_{ijk}(\theta_{ijk} - \theta_{ijk})^0)^2, \tag{3.7}$$

now with θ^0_{ijk} being the equilibrium angle between the three atoms i,j and k. k$^\theta_{ijk}$ is the harmonic force constant and is together with the equilibrium angle dependent on the parameters of the force field. The bond angle vibrations can also be described by a trigonometric potential, and there are also additional terms that can be added if a more accurate potential is needed.

When doing MD simulations, it is usually the movements of the macromolecule that is of interest, and while the bond angles and bond stretching do add to this, the torsional movement described by the dihedrals are where the real conformational changes happens. Typically, they can be hundreds of times less stiff than bond stretching and getting them right is therefore of the utmost importance when the large conformational changes of a macromolecule is to be reproduced (Gonzáles, 2011). The dihedral angle is created by the angle between the intersecting planes of 4 neighbouring atoms, and as is illustrated in figure (TO DO NEED FIGURE), they come in two different forms, so

$$V_{dihedral} = V_{proper} + V_{improper}. \tag{3.8}$$

The proper dihedrals are used to restrict the torsional rotation of four consecutive atoms, while the improper dihedrals are meant to keep planar groups planar and to prevent molecules from flipping over to their mirror images.

The proper dihedral potential is either represented by a periodic function or, like with the Ryckaert-Bellemans potential, a function based on expansion in power of $\cos\phi$ on the form

$$V_{proper}(\phi_{ijkl} = \sum_{n=0}^{5} C_n(\cos\psi)^n. \tag{3.9}$$

The improper dihedral potential can also be formulated as a periodic function, but the simplest representation is a harmonic potential

$$V_{improper}(\xi_{ijkl} = \frac{1}{2}k_{\xi}^{(}\xi_{ijkl} - \xi_0)^2, \tag{3.10}$$

with ξ_{ijkl} being the improper dihedral angle, ξ_0 the reference value from the force field parameters and k_ξ the force constant.

3.7 Thermostats and barostats

When a system has been designed with the proper structure and a force field matching the system requirements has been chosen, it is time to think about what conditions the system is going to be simulated under. This would for most biological and chemical processes be with either a constant number of particles N, a constant temperature T, and a constant pressure P, giving an NPT ensemble, or constant N, constant volume V and constant T, giving a canonical (NVT) ensemble. However, a simulation left to run the basic MD algorithms would produce a microcanonical ensemble with constant N, constant V and constant energy E (NVE). To solve this, methods called temperature and pressure coupling have been introduced that alters some of the parameters of the system to create the required conditions.

In temperature coupling, it is the velocities of the particles that are altered to keep the total temperature of the system correct. Although all these, so called thermostats, alters the velocities, the methods they use can vary and with big implications for the system. The simplest is the Andersen thermostat that randomizes the velocities of the particles with a stochastic term producing a Maxwell-Boltzmann distribution. This can either be done all at once after some set time or with a small probability at every time step for each particle. The other available thermostats in GROMACS are the Berendsen, the improved Berendsen called the velocity rescaling, and the Nosé-Hoover.

The Berendsen thermostat uses a weak coupling scheme to introduce the coupling. This implies that an external heat bath with a temperature T_0 is coupled to the system with first order kinetics giving a slow correction of the temperature following the equation

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}, \tag{3.11}$$

where the temperature decays exponentially with the time constant τ. T is the temperature of the simulation. This is done by scaling the velocities of all particle every step, according to the formula

$$\lambda = \left[1 + \frac{n_{TC}\Delta t}{\tau_T}\left\{\frac{T_0}{T(t - \frac{1}{2}\Delta t} - 1\right\}\right]^{1/2}, \tag{3.12}$$

where λ is the scaling factor, τ_T is a parameter that is close to the time constant τ, t is time and n_{TC} is how many steps between each time the velocities are scaled. This ensures an efficient relaxation towards the desired temperature, but at the same time it does not create a true ensemble, i.e. the correct distribution of velocities among the atoms. This is however, fixed in the velocity rescaling thermostat by adding a stochastic term given by

$$dK = (K_0 - K)\frac{dt}{\tau_T} + 2\sqrt{\frac{KK_0}{N_f}}\frac{dW}{\sqrt{\tau_T}}, \tag{3.13}$$

where K is the kinetic energy, N_f the number of degrees of freedom, dW a weiner process. The Nosé-Hoover does also produce the correct ensemble but as it uses oscillatory relaxation, meaning the temperature will oscillate as it relaxes towards the desired temperature, it is very efficient. For those reasons, it is common to equilibrate the system using a Berendsen thermostat and then change to one of the other later.

This is also the case for the pressure coupling techniques, or barostats. The Berendsen barostat functions in the same way as the thermostat, except it is now the lengths of vectors defining the simulation box and every coordinate inside it that is scaled. But as with the thermostat, it does not create the true ensemble of pressure fluctuations. The other commonly used barostat is the Parrinello-Rahman barostat. This is a barostat that is similar to the Nosé-Hoover thermostat and does produce the correct ensemble. With the Parrinello-Rahman pressure coupling, the box vectors are included in the matrix $\boldsymbol{b}$ that obey the matrix equation of motion

$$\frac{d\boldsymbol{b}^2}{dt^2} = V\boldsymbol{W}^{-1}\boldsymbol{b}'^{-1}(\boldsymbol{P} - \boldsymbol{P}_{ref}). \tag{3.14}$$

Here, $\boldsymbol{P}$ is the pressure of the simulation, while $\boldsymbol{P}_{ref}$ is the matrix defining the pressure of the heat bath. V is the volume of the box and W is a matrix parameter defining the strength of the coupling. In addition to the box vectors, the equations of motion for the particles are also changed

$$\frac{d^2\boldsymbol{r}_i}{dt^2} = \frac{\boldsymbol{F}_i}{m_i} - \boldsymbol{M}\frac{dr_i}{dt} \tag{3.15}$$

where the $\boldsymbol{M}$ matrix follows

$$\boldsymbol{M} = \boldsymbol{b}^{-1}\left[\boldsymbol{b}\frac{d\boldsymbol{b}'}{dt} + \frac{d\boldsymbol{b}}{dt}\boldsymbol{b}'\right]\boldsymbol{b}'^{-1}. \tag{3.16}$$

The mass parameter matrix $\boldsymbol{W^{-1}}$ mentioned earlier are defined as

$$(\boldsymbol{W}^{-1})_{ij}?\frac{4\pi^2\boldsymbol{\beta}_{ij}}{3\tau_p^2 L}, \tag{3.17}$$

with τ_p being the pressure time constant, L the largest box matrix element, and $\boldsymbol{\beta}$ the approximate isothermal compressibilities.

There is also a third option for pressure control in GROMACS, the MTTK algorithms. This algorithm has been shown to be the most accurate of the barostats (Shirts, 2013), but it is also much more computationally consuming than the others. This comes as a result of its inability to be used together with *constraints*, a widely used technique for calculating the bonds allowing for up to four times bigger time steps than with a harmonic potential.

3.8 The simulation run

If a thermostat and possibly a barostat have been chosen, these are put into the *molecular dynamics parameter file* (.mdp) of the system, covering all the parameters the simulation will follow. In addition to the coupling techniques, this file includes the decisions on, among other things, the length of the simulation and the time steps, how often to save simulation data, the bonded and non-bonded interaction functions and the cut offs, and the periodic boundary conditions. There is also the important choice of which integrator to use. The integrator are the algorithms that calculates the Newtonian equations of motion,

$$\boldsymbol{F}_i = m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2}, \quad i = 1...N, \tag{3.18}$$

for every atom i with mass m_i, coordinates r_i and $\boldsymbol{F}_i$ the total force acting on the atom. The forces are derived from the equation

$$\boldsymbol{F}_i = -\frac{\partial V_i}{\partial \boldsymbol{r}_i} \tag{3.19}$$

with V_i being the total potential energy experienced by atom i. All equations are solved simultaneously by the integrator with a numerical scheme, most commonly either the leapfrog or the velocity Verlet algorithm. The leapfrog algorithm is the default integrator in GROMACS and updates the position r_i at time t, but the velocity v_i at time $t + \frac{1}{2}$ is calculated using the velocity at time $t - \frac{1}{2}$, so at half time steps ahead and behind the positions. The velocities at the whole time steps is therefore calculated as an average of the preceding and following half time step velocities. The velocity Verlet avoids this problem by integrating both the position and velocity at the same time, following the equations

$$\boldsymbol{r}_i(t + \Delta t == \boldsymbol{r}_i(t) + \boldsymbol{v}_i + \frac{\Delta t^2}{m_i} \boldsymbol{F}_i(t), \tag{3.20}$$

$$\boldsymbol{v}_i(t + \Delta t) = \boldsymbol{v}_i(t) + \frac{\Delta t}{2m_i}[\boldsymbol{F}_i(t) + \boldsymbol{F}_i(t + \Delta t)]. \tag{3.21}$$

The velocity Verlet algorithm does need a few more calculations than the leap frog, but when doing precise pressure control simulations, it is the only option when choosing the integrator. This is because of the errors introduced by calculating the position and velocity at different times, and can lead to large errors in pressure (Cuendet and van Gunsteren, 2007).

Before beginning with the simulation itself, it is also important to do an energy minimization on the system to reduce the thermal noise of the structure. This is also crucial if the structure comes from a PDB file that was bound or otherwise far from equilibrium, as the resulting forces without energy minimization might be large enough to crash the simulation. For this, the *steepest descent method* is often the simplest and best choice.

If the system is to be run with temperature or pressure coupling, it is normal to run a short simulation where the system is gradually equilibrated towards the chosen temperature and pressure levels. When this is done, the production run itself can be initiated. It is

this simulation that will give the output data that can be analysed. This data, coordinates, velocities and energy, are saved in *trajectory files* that show the time evolution of these values. As mentioned earlier, it is possible in the parameter file to choose how often the data should be saved. There are many different methods to analyse the data, and the choice should depend on what the study is trying to examine. It is also common when using large molecules with long run times to check the root mean square deviation (RMSD) of the system to check whether the system has reached an equilibrium state and then only analyse the data from after this has been reached (Schreiner et al., 2012).

3.9 Approximations and limitations

The method behind molecular dynamics is to study a system by solving Newtons equations of motion, or with other words, using classical mechanics. The implications of this is that all electronic motions are ignored and only the movement of the nuclei is considered, which builds on the *Born-Oppenheimer approximation*. The effect of the electrons is still present in the simulation in the covalent bonds and van der Waals forces, but their motions are not considered. Instead, they are always in their ground state, and react instantaneously to changes in the atomic positions. At most temperature levels and for most studies, this should create no problems, but this is still an approximation and must be considered when deciding if MD can properly simulate the system in question. The lack of any electronic motion prevents the simulation of any electron transfer processes, and the use of classical harmonic oscillators instead of quantum oscillators mean that the high frequency vibrations in covalent bonds will not be accurately replicated in MD.

When talking about limitations in MD, the force fields also add to the uncertainty of the method. The parameters making up the force fields are of course constantly getting better, but the parameters are still only approximate, exemplified by the variation in results between them (Mu et al., 2003) (Smith et al., 2015). In addition, the calculations done by the force fields on van der Waals interactions are pair-additive, meaning that the potentials calculated between two atoms don't change when the conditions around them is different from the conditions the parameters were calculated under. For the most part, this should not be too great a concern, but this lack of effective pair potentials, that is, pair potentials that change depending on the surroundings, means that atomic polarizability is not possible to simulate. Fine-tuning of bonded interactions are also not possible given the present force fields.

Chapter 4

Materials and method

4.1 Plant material and tissue cultures

The plant used in this study was the model plant Arabidopsis thaliana (A. thaliana). The wild type seeds were Col-0, which were available in the host research group. Seeds grown on soil were not sterilized and spread on moist soil (S-JORD, Hasselfors Garden AB) that had been pre-autoclaved at 121 °C for 20 min. Each seed was grown in its own pot and each pot covered with cling foil after spreading to retain moisture and avoid seed contamination from nearby plants. The pots were placed in a growth room with standard conditions (22 °C for 16 h with 150μ mol light intensity, and for 8 h with dark conditions). Relative humidity kept constant at 50 % and the foil removed when the plants developed the first true leaves after the. Seeds grown in liquid media was first sterilized by vortexing the seeds in 1 ml of a 50 % bleach solution and leaving them for 10 min. This was followed by removal of the bleach solution and adding 1 ml Milli-Q water before removing the water after 1 min. This was repeated four times to properly wash the seeds. The sterilized seeds were then grown in $\frac{1}{2}$ Murashige and Skoog (MS) basal medium (0.21 % MS, 2.6 mM MES sodium salt, 1 % sucrose at pH 5.7) in sterile conditions in either a multi-well plate or Erlenmeyer flask. In the multi-well plates 20 seeds were grown in 2 ml of the $\frac{1}{2}$MS for 7-10 days, while approximately 100 seeds were grown in 100 ml of $\frac{1}{2}$MS in the Erlenmeyer flasks for 6-8 days. These flasks were put on a IKA KS501 flask shaker on 130rpm.

4.2 Synthesising the DNA sequences

DNA sequences coding for both the proteins of interest, the YFP and YFP3G from the study by Watanabe et al. (2013), were ordered from GeneArt, a company specializing in the synthesis of genetic sequences. The genes were designed to be flanked by special attachment sites to be used with the Gateway cloning system and were delivered from GeneArt in a line fragment/a basic plasmid ready to be processed using recombinant DNA

technology.

4.3 The Gateway® cloning technique

To generate the final DNA constructs, the Gateway® cloning system (Termo Fisher) was used. This is a method for molecular cloning that exploits the recombination process happening when the lambda phage, a bacterial virus, integrates its DNA with the victims. In the bacterial genome, specific attachment sites called attB will recombine with complementary sites in the circular viral DNA called attP. After recombination, the phage's genome is integrated in the bacterial genome and flanked by two new sites, the attL and attR (left and right). These two sites sometimes recombine again and release the viral DNA as well as recreating the original attB and attP sites. In the Gateway® cloning system, the recombination site of the phage's incision is mutated into two new ones. As the new attB1 and attB2 sites only recombine with their matching attP1/2 sites, a gene flanked by these sites will have the correct orientation after the recombination. The cloning technique is split up in two separate reactions, the BP reaction (attB $\times$ attP $\rightarrow$ attL $\times$ attR) and the LR reaction (attL $\times$ attR $\rightarrow$ attB $\times$ attP). The direction of the reaction is controlled by adding additional enzymes to catalyse the reaction, namely the BP Clonase™ and LR Clonase™ enzyme mixes. Following the Gateway® cloning method, 70 ng of the linearized DNA fragments flanked by attB sites were recombined with the pDONR™/Zeo donor vector through a BP reaction. To account for any unrecombined vectors, the donor contains the ccdB gene coding for a toxin.

4.4 The pUB-DEST and pFAST-G02 destination vectors

The expression vector pUB-DEST was used to generate the finished constructs. This is a Gateway compatible vector containing the ubiquitin-10 gene promoter (Grefen et al., 2010). This promoter is native to Arabidopsis plants and its accompanying gene is in the top percentile of highly expressed genes throughout the plant.

The pFAST destination vectors are a set of gateway compatible binary vectors using the Fluorescence-Accumulating Seed Technology (FAST) for faster screening of transformed Arabidopsis plants (Shimada et al., 2010). This is a screening marker technology based on the oil body membrane protein OLEOSIN 1 (OLE1) found in A. thaliana seeds. This protein is only expressed during seed dormancy and is quickly degraded during germination. A fluorescent protein fused with the OLE1 protein following the OLE1 promoter will therefore only be expressed in the seeds, allowing for the screening of transformed seeds without affecting the adult plant. Shimada et al. (2010) developed a range of different pFAST vectors with either green or red fluorescent proteins, and in this work, it is the pFAST-G02 vector that has been used. Vector maps are presented in appendix (TO DO: NEED TO ADD APPENDIX) (TO DO: Present antibiotics)

4.5 Transformation of E. coli

After each recombination reaction, 2 μl of the reaction volume were used for transformation of chemically competent One Shot® TOP10 E. coli cells to amplify the recombined plasmids. The protocol for transformation from Invitrogen was followed (TO DO), starting with incubation on ice for 30 min. This was followed with heat-shock treatment using a $42\,^{\circ}\mathrm{C}$ water bath for 45 sec before recovering the cells in 700 μl Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, 5 g/L sodium chloride) in a shaker at 220 rpm for 1 h in $37\,^{\circ}\mathrm{C}$. After incubation, the cells were centrifuged at 3500xg for 6 min, and all but 50 μl of the supernatant discarded, which was used to resuspend the E. coli cells. To grow colonies from the transformed cells, sterile LB agar plates (TO DO: check word from Thorsten) (LB medium with 1 $\%$ agar) were prepared, containing the proper antibiotics for selection of the vectors. During the last step, cells were spread on plates and incubated at $37\,^{\circ}\mathrm{C}$ overnight.

4.6 Plasmid DNA extraction from E. coli colonies

From the incubated LB-agar plates, single colonies were picked using sterile pipette-tips and used to inoculate 5 ml of sterile LB medium containing the proper antibiotics. The cells were then put overnight in an incubator shaker set at 220rpm and with $37\,^{\circ}\mathrm{C}$. Extraction of plasmid DNA from the E. coli cells were performed with miniprep using the E.Z.N.A Plasmid DNA Mini Kit I (Omega) and following the accompanying protocol.

4.7 Sanger DNA Sequencing of plasmid DNA

Extracted plasmid DNA was sequenced using the Sanger DNA sequencing method to ensure that the YFP and YFP3G genes were inserted in the vectors, and that no mutations had occurred. Following the miniprep, 200 ng of the extracted plasmid DNA was mixed with 2.5 μl of primer and enough Milli-Q water as to have a final volume of 10 μl. Samples were sent to GATC Biotech AG in Cologne, Germany for sequencing. The results of the sequencing were usually received within a week and analysed using the FinchTV (Geospiza, Inc.) and MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle/) tools.

4.8 Transformation of *Agrobacterium tumefaciens* (*A. tumefaciens*)

The bacteria A. tumefaciens was used to transform the A. thaliana plants. This is the bacterium responsible for crown gall disease and during infection transfer parts of its plasmid DNA into its victim's genome (Pitzschke, 2013). To transform the A. tumefaciens with the expression plasmids, the electro competent A. tumefaciens C58 strain were used. 300 ng of the extracted expression plasmids where mixed with bacterial cells and transferred to a cold pulser-cuvette. After inserting the cuvette in a Gene Pulser® II (BioRad), the cuvette was subjected to 1.8 kV for 5 sec on 25 μF and 200Ω. After transformation, the

cells were recovered in 700 μl of LB medium in an incubator shaker set at 180rpm and 28 °C for 3h. The recovered cells where then spread on sterile LB-agar plates containing the proper antibiotics for the expression vector and incubated at 28 °C for two days.

4.9 Transformation of *A. thaliana* using the floral dip method

A. tumefaciens mediated transformation of the A. thaliana plants was performed following a modified version of the protocol Clough and Bent described in their 1998 paper (Clough and Bent, 1998). When the growing wild type A. thaliana plants were at the beginning of the flowering stage with many unopened flower buds, the transformed A. tumefaciens cells containing the expression plasmids were incubated overnight in 5 ml of sterilized LB medium containing the proper antibiotics at 28 °C and shaken at 180rpm. The bacteria were then transferred to a new container with 200 ml LB medium and incubated with the same conditions for a few hours until the stationary phase was reached. To determine this, the SmartSpec™ Plus Spectrophotometer (BioRad) was used to measure the OD600. A value between 0.8 and 1 indicated a correct culture density, and the LB solution with the cells was centrifuged at 6000xg for 6 min. Following this, the pellet of bacterial cells was resuspended in a beaker using (TO DO WHAT AMOUNT) of the floral dip cultivation medium (5 % sucrose, 10mM MgCl2, 200 μM acetosyringone and Milli-Q water) and incubated at room temperature for 2 h. Then 0.01 % Silwet L-77 was added before the plants were dipped into the mixture for 1 min. This was done by turning the plants upside down and letting the entire plant above a few centimetres of the soil be submerged. As a high level of humidity for the first 12-24 h leads to a doubling of transformations (Clough and Bent, 1998), the plants were covered with cling film immediately after dipping and placed horizontally on a tray. The tray was then covered with aluminium foil and the plants were left like this in a growth chamber for 48 h. After the 48 h had passed, the plants were unwrapped and left to grow until they were ready for harvest.

4.10 Harvest and selection of transformed seeds

Seeds were screened using agar plates with the proper antibiotics. T1 generation from growing seedlings, created x lines. T2 generation by checking the % of resistant of each line. T3 generation wanted homozygous lines.

4.11 RNA extraction

In order to perform gene expression analysis of the plants, the total RNA of the plants was extracted using a Spectrum™ Plant Total RNA Kit (Sigma-Aldrich). 10-day old seedlings were placed in Eppendorf tubes together with a stainless-steel bead and placed in liquid nitrogen for a couple of minutes to be snap-frozen. After this, the plants were homogenized with a TissueLyser (Qiagen) at 21Hz for 2 min. The instructions in the protocol provided with the kit were followed, to extract the total RNA. Finally, a Nanodrop-100

spectrophotometer (ThermoFisher) was used to determine the amount and quality of the extracted RNA.

4.12 cDNA synthesis

Two step qRT-PCR method was used for gene expression analysis, and the RNA therefore had to be converted into a suitable template, a more stable complementary DNA (cDNA) by cDNA synthesis, or reverse transcription. To avoid the possibility of stray genomic DNA having contaminated the RNA, RQI RNase-Free DNase (Promega) was, following the protocol from the manufacturer, used to treat the RNA samples before the reverse transcription.

The reverse transcription was performed using the ImProm-II Reverse Transcription System (Promega), using a modified version of the accompanying protocol. While the RNA was treated with DNase, a mastermix giving 15 μl to each sample were created using the following ingredients:

A master mix were created with a volume large enough to give 15 μl to each well, using the following ingredients: The mastermix were then divided into PCR tubes, and,

Component	**Volume**
Milli-Q water	5.75 μl
ImPromII™ 5x Reaction buffer	4 μl
ImPromII™ Reverse transcriptase	1 μl
dNTP Mix (10 mM) (Thermo Fisher)	1.25 μl
$MgCl_2$ (25 mM)	2 μl
Oligo$(dT)_{15}$Primer (500 μg/ml)	0.5 μl
Random Primers (500 μg/ml	0.5 μl

immediately after the inactivation step of the DNase digestion combined with 5 μl of the total RNA now also containing the DNase mix. As the RNA was treated to $65\,^{\circ}\mathrm{C}$ for 10 min while inactivating the DNase and then added directly to the PCR tubes, no denaturation step was included. The settings for the PCR was a $25\,^{\circ}\mathrm{C}$ annealing temperature for 10 min, followed by elongation at $42\,^{\circ}\mathrm{C}$ for 1 hour, before a heat inactivation step for the reverse transcriptase polymerase at $72\,^{\circ}\mathrm{C}$ for 15 min. The finished product were then stored at $4\,^{\circ}\mathrm{C}$.

4.13 Quantitative real-time PCR

The fluorescent marker molecule SYBR® Green was used in the qRT-PCR, as the marker has a >1000-fold increase in fluorescence when binding to double stranded DNA (dsDNA) (Dragan et al., 2012). This means that for each amplification cycle, the fluorescence intensity at the end of the cycle increases. And by correlating the fluorescence intensity with the number of cycles run, the relative initial number of transcripts in the samples can be calculated (VanGuilder et al., 2008).

A 96-well plate were used to hold the samples, and three technical replicates done for each sample to control for technical variaton. The reference gene used for normalization was ACTIN 2 (ACT2), as this is a gene that is known to have a strong expression in the entire plant (Yong-Qiang et al., 1996).

A master mix was created with a volume large enough to give 18 μl to each well, using the following ingredients: The forward and reverse primers used for the YFP and YFP3G

Component	**Volume**
Milli-Q water	3 μl
LightCycler® 480 SYBT Green I Master 2X	10 μl
cDNA (diluted 1:10)	5 μl

are described in section (TO DO CREATE SECTION). 2 μl of both these primers, as well as the ones for ACT2, were added to the side of their respective wells using a multi-stepper pipette, for a total reaction volume of 20 μl. A sample known not to contain the YFP or YFP3G genes were also added, as a negative control. A LightCycler® 96 System (Roche) was used with a program consisting of a pre-incubation step at $95\,^{\circ}\mathrm{C}$ for 5 min to activate the DNA polymerase, followed by 45 cycles of $45\,^{\circ}\mathrm{C}$ denaturation for 10 sec, $457\,^{\circ}\mathrm{C}$ annealing for 10 sec and $72\,^{\circ}\mathrm{C}$ elongation for 15 sec, with a measurement of the fluorescence intensity being taken at the end of each cycle.

The expression levels (TO DO)

4.14 Preliminary confocal fluorescence microscopy study of cotyledons and roots from T_3 plants

To explore how strong the fluorescent signals from the transgenic plants were, a preliminary study was set up. Around 100 sterilized T_3 seeds from the selected lines were grown in Erlenmeyer flasks containing 100 ml liquid $\frac{1}{2}$MS for 6 days in a shaker at 130 rpm. The seedlings were then split equally into new flasks of 100 ml $\frac{1}{2}$MS, with a concentration of either 300 mM DMSO or 300 mM sorbitol, in which the seedlings were left for three hours. The DMSO was in this case used as a mock control, while the sorbitol treatment was used to lower the turgor pressure level of the plant cells. After the treatment, the cotyledons and roots of the seedlings were cut and placed on glass slides, so the sample would be flat when it was analysed under the microscope.

The study examined the roots and cotyledons of 6 days old seedlings using a Zeiss LSM 800 confocal microscope. The images taken of the cotyledons used dry objectives with either 2.5x or 20x magnification, while the roots were imaged with a dry objective with 10x, giving total magnifications of 25x, 200x and 100x respectively. Both YFP and YFP3G was excited using a wavelength set to 508 nm and the emission was collected with a 511-558 nm band pass filter. In addition, a digital gain of 3 was used.

The fluorescence coming from the roots of the plants expressing YFP3G was very weak, so a decision was made not to do further analysis on this part of the seedlings. Also, in spite of the preparations, the cotyledons investigated in the images using 25x magnification were not entirely flat. This caused the focal plane to go through the epithelial

cells on the edge of the cotyledons, and the palisade mesophyll cells in the middle. In the images of the cotyledons with 200x magnification, the focal plane was set to only pass through the epithelial cells.

4.14.1 Analysis

Based on the images, two methods for analysing the fluorescence of the seedlings were chosen, one analysing the 200x in images of the epithelial cells, and one on the whole cotyledon. Both methods used the ImageJ software tool (Schindelin et al., 2017) for analysis. When analysing the whole cotyledon, only the area of the cotyledon where the focus cut through the epithelial cells was analysed. This decision was made to ensure that the fluorescence intensity of the cotyledons could be compared, as the size of the area of weaker fluorescence in the middle of the cotyledon was varying from image to image, affecting the average intensity of the whole cotyledon. In ImageJ, the strongly fluorescing areas on the edge of the cells were then chosen as a region of interest (ROI) using the freehand selection tool. The average fluorescence intensity of the ROI was then calculated using the Measure tool. The method for analysing the epithelial cells in the pictures with 200x magnification, a different method for analysing the fluorescence intensity was chosen. Here, the cells were manually counted, before the mean intensity of the entire image calculated by the Measure tool was divided by the number of cells. This was done to minimize the effect the variations in cell size, and thereby the number of cells in a picture, would have on the mean fluorescence.

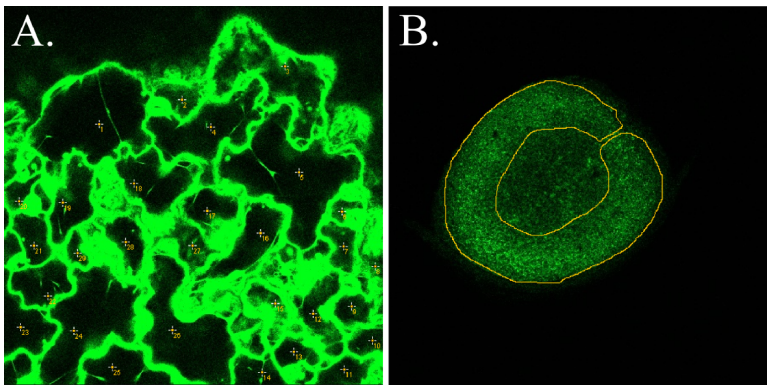


Figure 4.1: A.: Example of how the cells were counted. B.: Example of how the ROI manually created when analysing the mean intensity of the cotyledons

4.15 Confocal fluorescence microscopy study of cotyledons and roots from T_3 plants

Based on the experiences from the preliminary study, a few changes were made for the final study. Like in the preliminary study, 100 sterilized T3 seeds were grown for 6 days

in Erlenmeyer flasks containing 100 ml liquid $\frac{1}{2}$MS in a shaker set at 130 rpm. But the seedlings were now divided between seven different treatments, $\frac{1}{2}$MS with either DMSO, 150 mM, 300 mM or 600mM sorbitol, or 150 mM, 300 mM or 600 mM sodium chloride. The length of the treatment was also shortened to 30 min, as the discovery of an article presenting how the epidermal root cells of A. thaliana have the ability to recover turgor pressure levels by taking up more ions [Shabala]. The inclusion of sodium chloride in the final study was done to get a wider sample set, and see if there was a similar trend for the two different solutes. After the treatment the cotyledons were cut and put on microscope slides for imaging.

The microscope was this time a Leica DMI6000B CS confocal microscope, and for all images a 10x objective without immersion was used. The main difference from the preliminary studies was that a z-stack of minimum 5 images were taken of all the samples. Based on some trial and error, a fixed distance of 2.41 μm between each image gave three images per epithelial cell, which when combined increases the amount of data gathered. The top and bottom image was placed just above and below the epithelial cells. The excitation wavelength was set to 514 nm and the detection range to 522-550 nm. This time, the gain was also set differently for the seedlings expressing YFP and the ones expressing YFP3G, with a gain of 733 and 981 respectively. This was to ensure pictures of as high a quality as possible for both the plants expressing YFP and the ones expressing YFP3G.

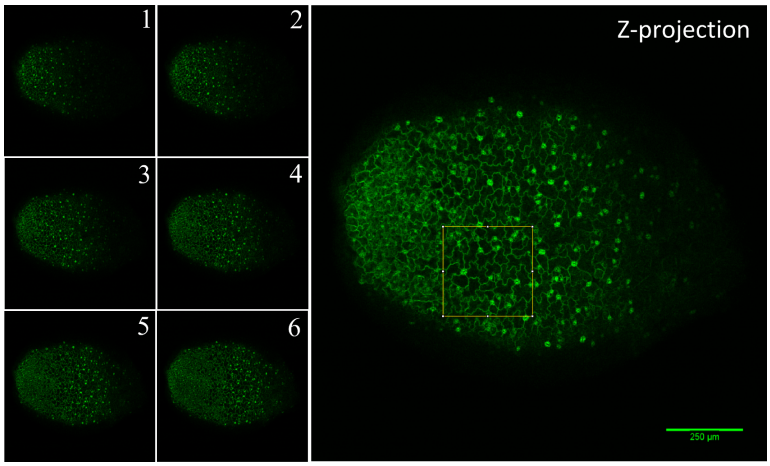


Figure 4.2: On the left are the six images moving down into the cotyledon. On the right is the Z-projection created by image 3-5.

4.15.1 Analysis

Of the images in the Z-stack, three consecutive images cutting through the epithelial cells of a large area of the cotyledon was combined into a Z-projection using the Z-project tool in ImageJ with a Maximum Intensity setting. For each Z-projection, a square ROI was created with the Specify selection tool, and width and height set to 200. The ROI was then manually placed in a location containing epithelial cells with clear borders and no sign of

the top or bottom layer of the cell, and then analysed with the Measure tool. Following this, the number of cells in each picture was counted, and the mean intensity of the ROI was divided by the number of cells in the picture.

Chapter 5

Results

5.1 The constructs

The nucleotide sequence coding for the YFP and YFP3G genes used in the paper by Watanabe et.al. were obtained from the authors. Both genes are nearly identical except for the three glycine molecules inserted in the β-barrel as residue number 145-147. The total nucleotide sequence of YFP3G, as well as the translated residue sequence, can be seen in Figure 5.1.

5.2 Completing the PDB file structure

To create the amino acid sequence for the simulations, PDB files from the RCSB Protein Data Bank were used. The structure used to construct the YFP3G was the structure resulting from the X-ray crystallography performed in the study by the Watanabe group (PDB ID 3W1D), while the structure used for the YFP was the PDB file employed by the group as a search model for the crystallography (PDB ID 3DQ7). The finished structures can be seen in Figure 5.2. Both PDB files were incomplete, or in other ways differing from the amino acid sequence presented in Figure 5.1. The PDB files are presented in Appendix (TO DO) and as can be seen, the last seven amino acids in the sequence, residue 231-238 in the PDB file description of YFP or 234-241 for YFP3G, are not present. These residues are known not to be essential for the fluorescence of the protein and also to cause excessive free energy and were therefore not added to the PDB files (Cetinkaya et al., 2005).

A missing residue sequence that had to be repaired was found in the YFP3G PDB file. The four amino acid residues following the first inserted glycine (146-149, GGYN) in β-strand 7 were all missing from the PDB file, essentially dividing the protein in two separate parts. This sequence therefore had to be manually repaired by adding the missing residues to the PDB file. This was done using the PyMOL Molecular Graphics System, and care was taken to orient Tyr148 away from the barrel, following the descriptions from the Watanabe group.

```
0   M__V__S__K__G__E__E__L__F__T__G__V__V__P__I__L__V__E__L__D__
    ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGAC
20  G__D__V__N__G__H__K__F__S__V__S__G__E__G__E__G__D__A__T__Y__
    GGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTAC
40  G__K__L__T__L__K__F__I__C__T__T__G__K__L__P__V__P__W__P__T__
    GGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACC
60  L__V__T__T__F__G__Y__G__L__Q__C__F__A__R__Y__P__D__H__M__K__
    CTCGTGACCACCTTCGGCTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAG
80  Q__H__D__F__F__K__S__A__M__P__E__G__Y__V__Q__E__R__T__I__F__
    CAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTC
100 F__K__D__D__G__N__Y__K__T__R__A__E__V__K__F__E__G__D__T__L__
    TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTG
120 V__N__R__I__E__L__K__G__I__D__F__K__E__D__G__N__I__L__G__H__
    GTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCAC
140 K__L__E__Y__N__G__G__G__Y__N__S__H__N__V__Y__I__M__A__D__K__
    AAGCTGGAGTACAACGGAGGTGGATACAACAGCCACAACGTCTATATCATGGCCGACAAG
160 Q__K__N__G__I__K__V__N__F__K__I__R__H__N__I__E__D__G__S__V__
    CAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTG
180 Q__L__A__D__H__Y__Q__Q__N__T__P__I__G__D__G__P__V__L__L__P__
    CAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCC
200 D__N__H__Y__L__S__Y__Q__S__A__L__S__K__D__P__N__E__K__R__D__
    GACAACCACTACCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGAT
220 H__M__V__L__L__E__F__V__T__A__A__G__I__T__L__G__M__D__E__L__
    CACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTG
240 Y__K__*__
    TACAAGTGA
```

Figure 5.1: Nucleotide and residue sequence for the gene coding for the YFP3G protein. On the left is the number of the first residue of each row, following the numbering used in PDB files, i.e. the first residue is Met0. The three glycine residues inserted into the β-barrel of the YFP is highlighted in grey and is the only difference between the two genes.

5.3 Modifying the AMBER99SB*-ILDN force field

The chromophore used in YFP and YFP3G is essentially a single residue formed from the residues Ser65-Tyr66-Gly67. In the PDB file format, this is represented as a single residue called CR2. As this is a residue only native to this protein, the force field used in this study, the AMBER99SB*-ILDN, didn't have the parameters for this residue. Therefore, the partial charges of the chromophore needed to be computed and added to the force field. This was done using the PyRED Server, which is a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments (Vanquelef et al., 2011) (Dupradeau et al., 2010) (Bayly et al., 1993) (F. Wang, 2013). The approach when calculating these charges was, following the tutorials on the PyRED website, to begin by flanking the N-terminal of the bare and neutral CR2 residue by an acetate residue (ACE) and the C-terminal by a methylamine residue (NME). The adding of these residues were performed using a custom script in the PyMOL software. The resulting molecule was uploaded to the PyRED server and the partial charges calculated. The result of these calculations can be seen in Figure 5.3.

In addition to adding new partial charges, the atom types (see section 3.1) in the chromophore needed to be added to the force field, as well as parameters regarding bonded interactions. This parametrization was done by the R.E.D. server in the same calculations that derived the partial charges, and new parameters for bond stretching, angle bending, and proper and improper dihedrals were added to the force field. The changes are available

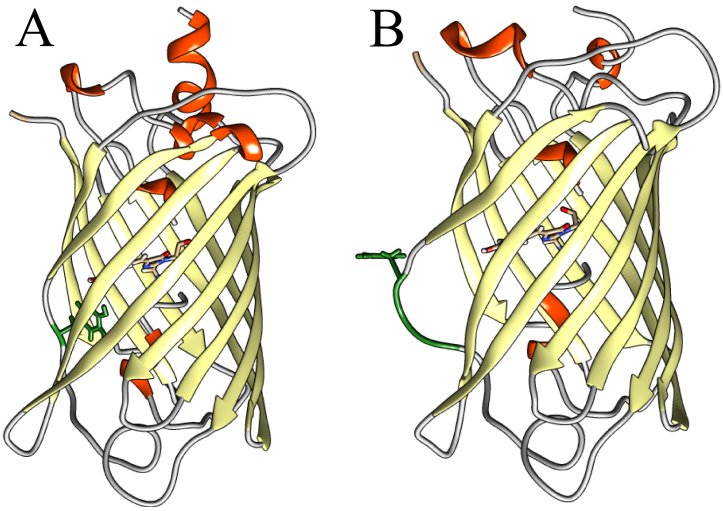


Figure 5.2: Ribbon representation of the protein structures simulated in this study. A: The YFP structure with the chromophore visible in the middle. Tyr145 is shown in green. B: The YFP3G structure with the chromophore visible in the middle. The three inserted glycine residues, Gly145-147, and Tyr148 are shown in green.

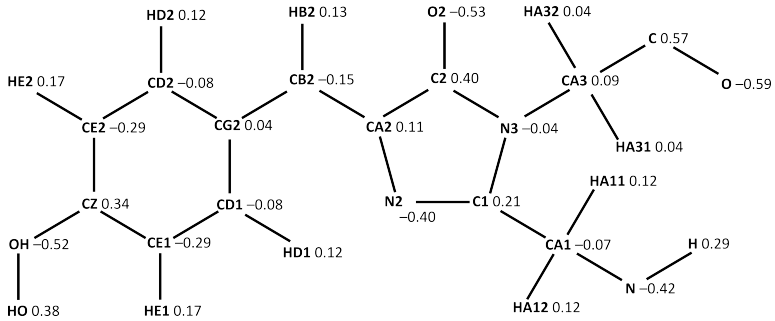


Figure 5.3: Partial charges derived for the neutral CR2 chromophore using the R.E.D. server together with the PDB names of the atoms.

in Appendix (TO DO).

5.4 The simulated systems and input parameters

In total, six molecular dynamics simulations were performed, three on each version of the YFP. These simulations were set with the protein floating freely in water inside a rhombic dodecahedron box with initial side lengths of 79.3 Å. With this length, the system had $\cong$ 8600 water molecules, as well as 6 Na^+ ions that were added to make the system neutral. The simulations had periodic boundary conditions and the AMBER99SB*-ILDN force field was used. The three simulations of each protein were set in NPT conditions,

with a temperature of 300 K (the protein and the solvent were connected to separate heat baths) and a pressure of either 0.1, 10, or 50 MPa. To control these properties, the v-rescale thermostat and the Parrinello-Rahman barostat were applied with coupling values of τ_T=0.1 ps and τ_P=5 ps respectively. Each simulation lasted for 90 ns, with energy minimization and three 100 ps simulations performed prior to the production run to relax the system towards the proper temperature and pressure values. For analysis, the first 30 ns of the production run were dismissed. The simulations were performed with the GROMACS 5.1.4 package. LINCS was used to constrain the covalent bond lengths of hydrogen atoms, allowing for a time step of 2 fs. The neighbour list was updated every 10 steps. To describe the van der Waals interactions, a Lennard-Jones potential with a cut-off of 1 nm were used, and the particle mesh Ewald method described the electrostatic interactions, also with a cut-off of 1nm.

5.5 RMSD and flexibility results of the simulations

To inspect the stability of the proteins, the root-mean squared deviation (RMSD) of the proteins were calculated. This is a measurement that compares the deviation of the structure to a reference structure, where a high RMSD value indicates a large deviation from the reference structure, and a value of zero means that the structure is identical to the reference structure. The stability of the protein cannot be derived directly from the RMSD, but an RMSD moving from zero to fluctuating around a mean value is a strong evidence for a stable structure (Pitera, 2014). To calculate the RMSD, the GROMACS tool gmx rms was employed, and the calculations were based on the protein backbone atoms, defined in the GROMACS manual as the N, C_α, and C atoms.

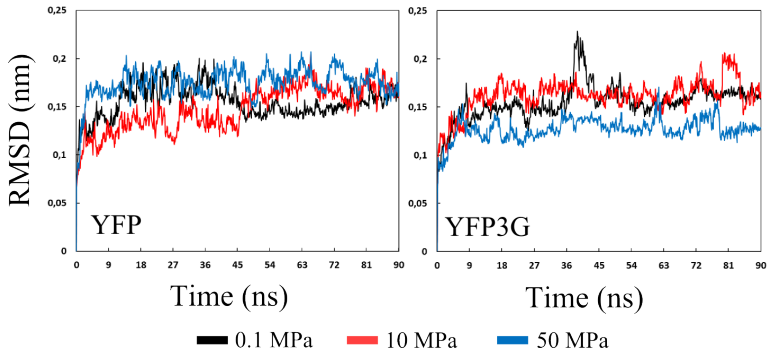


Figure 5.4: The RMSDs of the wild type YFP and the mutated YFP3G, either simulated in 0.1 MPa, 10 MPa, or 50 MPa.

The results of these calculations are presented in Figure 5.4. Based on these plots, the first 30 ns of the simulation run were omitted from further analysis, as the rapid changes in the RMSD indicate that the structures haven't completely settled yet. There are also large fluctuations after 30 ns, but especially for the YFP the values are generally more stable.

To look further into the flexibility of the structures, the root-mean square fluctuation (RMSF) were calculated using the GROMACS tool gmx rmsf. The RMSF is a value

for the magnitude of the motion of atoms by calculating the fluctuation away from their average positions. It was calculated for every atom in the protein, and then averaged for each residue. The average

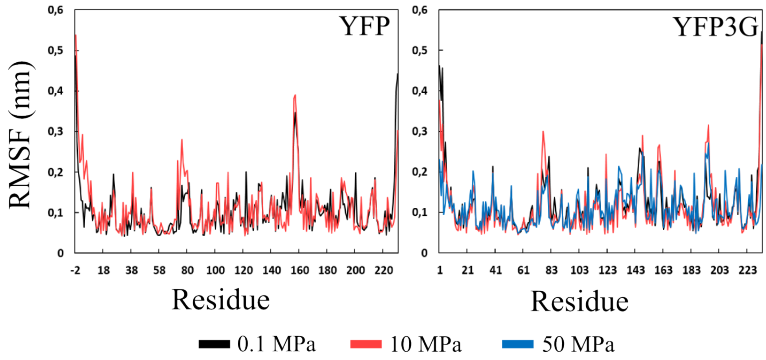


Figure 5.5: The average RMSF value for each residue in the two atoms, calculated after 30 ns.

5.6 Distance between chromophore and Tyr203/206

The mutation introducing the aromatic ring of Tyr203 close to the aromatic ring of the chromophore is believed to be the main reason for the redshift in YFP compared to GFP. The study by Watanabe indicated that the chromophore in YFP3G moved away from the corresponding Tyr206 as a result of the opening of a whole inside the β-barrel by Tyr148 flipping towards the outside of the barrel. To investigate this, the distance between the two aromatic rings were calculated at every timepoint (after 30 ns) of the simulation. This was done by looking at the distance between the CG2 atom in the chromophore and the CZ atom of Tyr145/148 with the GROMACS tool gmx distance, the results of which are presented in Figure 5.6.

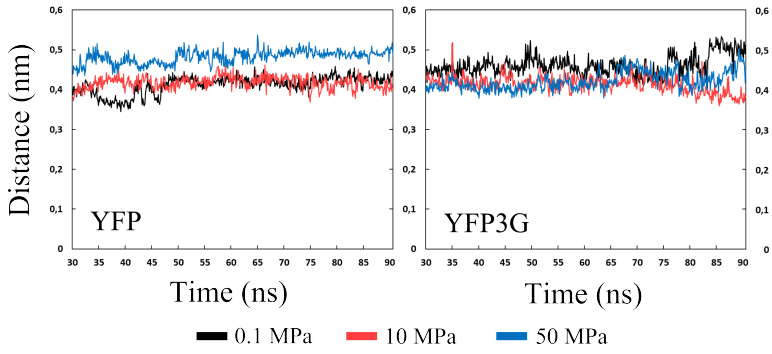


Figure 5.6: The distance between the CZ atom of the chromophore and the CG2 atom in the aromatic ring of Tyr203 for YFP and Tyr206 for YFP3G.

	0.1 MPa	**10 MPa**	**50 MPa**
YFP	0.412 ± 0.031	0.417 ± 0.029	0.482 ± 0.027
YFP3G	0.461 ± 0.040	0.414 ± 0.032	0.422 ± 0.034

Table 5.1: : The distance between the CZ atom of the chromophore and the CG2 atom in the aromatic ring of Tyr203 for YFP and Tyr206 for YFP3G as an average over the last 60 ns of the simulations $\pm$ standard deviations.

5.7 Water molecules inside the β-barrel

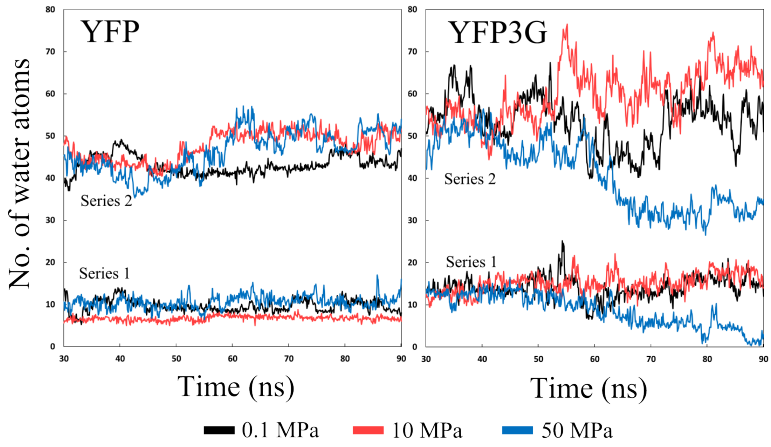


Figure 5.7: The number of atoms belonging to a water molecule, i.e. both hydrogens and oxygens that are part of a water molecule. Series 1 is the number of atoms that are within 0.6 nm of the CG2 atom of the chromophore. Series 2 is the number of water atoms that are within 1 nm of the N2 atom of the chromophore.

The number of water molecules close to the chromophore was estimated using the gmx select tool. The number of atoms part of a water molecule was calculated within either 0.6 nm from the CG2 atom of the chromophore, or 1 nm from the N2 atom of the chromophore. The first calculation is to determine the number of water molecules close to the chromophore. The second measurement gives a more general view of the number of water molecules inside the β-barrel, as the N2 atom is located almost in the middle of the protein. The diameter of the barrel was fluctuating between 2.3 and 2.6 nm, and radius of 1 nm will therefore cover a large part of the volume inside the barrel. The number of water molecules inside the barrel remained relatively constant for the YFP molecule, with no distinct pattern emerging. The YFP3G molecule, on the other hand, showed a steep drop in water close to the chromophore

5.8 Cloning of the constructs

The nucleotide sequence shown in Figure 5.1 was synthesised by GeneArt and afterwards recombined with the pDONR™/Zeo donor vector using the BP reaction of the Gateway® cloning system. This system for cloning were chosen because it is a fast and efficient method for producing expression clones that only need two reactions. The middle step entry clones can also easily be recombined with a large set of destination vectors. The recombinant plasmids were then transformed into heat-shock treated E. coli cells to amplify the plasmids. After the recombination, cells were spread on zeocin agar plates, from where single colonies were picked and the plasmids purified. Sanger sequencing were then performed on the plasmids from each colony to confirm that they contained the correct construct without mutations.

After the entry clones had been validated, the second stage of the Gateway® cloning system, the LR reaction, were performed. In this reaction, the two entry clones were recombined with the chosen destination vectors, the pUB-DEST and the pFAST-G02. Both vector maps are available in Appendix (TO DO: CREATE APPENDIX) The reason for working with two different promoters were the interesting properties of the pFAST vectors in screening transformed seeds, while at the same time having a promoter that were native to A. thaliana. As the 35S promoter used in the pFAST vectors is not native to the plant, the pUB-DEST containing the native ubiqutin-10 (UBQ10) promoter was chosen as a complement. Then the same amplification and purification and validation process as done after the BP reaction were performed, only now with spectinomycin agar plates for all four constructs.

When the final constructs had been generated, amplified, purified and validated they were transformed into the bacteria A. tumefaciens using electroporation. The bacteria were then spread on agar plates with spectinomycin, ampicillin and rifampicin, and the selected colonies were then grown in liquid cultures. These bacteria were then used to create transgenic A. thaliana seeds through the floral dip method.

5.9 Selection and screening of A. thaliana seeds

The A. thaliana seeds produced from dipped plants were T_1 seeds. After harvesting, the seeds were screened to identify the transformed seeds which could be used for growing the T1 plants. Because of the new screening method allowed by the pFAST vector, the screening was different between the two expression vectors.

The T_1 seeds containing the constructs with the UBQ10 promoter were screened by spreading the seeds on sterile agar plates containing the herbicide Basta and leaving them in dark conditions for a week. The seedlings displaying a strong root penetrating down into the agar were then transferred to pots with soil and left to grow. The T_2 seeds produced from these plants were then harvested and once again spread on agar plates containing Basta following the same procedure. To determine the number of insertions that had occurred in the genome of the transgenic seeds resulting from the floral dip, the number of Basta resistant versus susceptible plants was determid, and a segregation analysis were performed. Based on the result of this analysis, displayed below, the lines displaying close to 75 $\%$ resistance were considered as having a single insertion and used to grow the T_2

plants.

The plants with the pFAST vector, on the other hand, could be identified directly by looking at the dry harvested seeds in a fluorescence microscope. Seeds that were glowing are seeds that contain the OLE1-GFP fusion protein, and are markers of a successful transformation. These seeds were spread on soil and grown to produce T_2 seeds. To perform a seed segregation analysis, a minimum of 100 randomly selected seeds were imaged using a fluorescence microscope. The number of glowing seeds in these pictures were then counted using the Fiji software tool, and the lines with close Following the same procedure of imaging the seeds in a fluorescence microscope, a segregation analysis was performed using a minimum of 100 randomly selected seeds the glowing seeds in the Fiji software tool and, again, the lines displaying close to 75 % resistance were considered as having a single insertion and used to grow the T_2 seeds.

Construct	Counted glowy seeds	Counted non-glowing seeds	Total number of seeds counted	Ratio of glowing seeds
35S::YFP line 1	90	21	111	0.81
35S::YFP line 2	0	0	100	0
35S::YFP line 3	0	0	100	0
35S::YFP line 4	91	53	144	0.63
35S::YFP line 5	147	40	187	0.79
35S::YFP line 6	230	56	286	0.8
35S::YFP3G line 1	0	0	100	0
35S::YFP3G line 2	0	0	100	0
35S::YFP3G line 3	0	0	100	0
35S::YFP3G line 4	166	5	220	0.75
35S::YFP3G line 5	0	0	100	0
35S::YFP3G line 6	0	0	100	0
35S::YFP3G line 7	170	44	214	0.79
35S::YFP3G line 8	0	0	100	0

Table 5.2: The lines that have a 3:1 ratio (75 %
of glowy seeds were considered to only have a single insertion.

14 of the surviving T_2 seedlings from each of the homozygous lines were transferred to soil for both the UBQ::YFP lines 2 and 3, and the UBQ::YFP3G lines 2 and 4, as well as the 35S::YFP lines 1,4,5, and 6 and 35S::YFP3G lines 4 and 7. This selection was also made based on the qRT-PCR results displayed in the next section. The T3 seeds resulting from these plants were screened as before. Only the lines displaying a 100 % survival rate were considered homozygous.

5.10 qRT-PCR of T_2 seedlings

During transformation, insertion of the constructs into the genome of the A. thaliana plants is random [Azpiroz-Leehan]. It has been shown that DNA topology has a strong regulatory effect on transcription (Kouzine et al., 2014), so the different locations of the insertions

Construct	Resistant seeds	Susceptible seeds	Total number of seeds counted	Ratio of resistant seeds
UBQ::YFP line 1	0	0	62	0
UBQ::YFP line 2	71	1	72	0.99
UBQ::YFP line 3	59	8	67	0.88
UBQ::YFP line 4	0	0	100	0
UBQ::YFP3G line 1	38	15	53	0.72
UBQ::YFP3G line 2	40	18	58	0.69
UBQ::YFP3G line 3	42	18	60	0.70
UBQ::YFP3G line 4	50	18	68	0.74
UBQ::YFP3G line 5	53	2	55	0.96
UBQ::YFP3G line 6	51	10	61	0.84

Table 5.3: The lines that have a 3:1 ratio (75 %) of resistant seeds were considered to only have a single insertion.

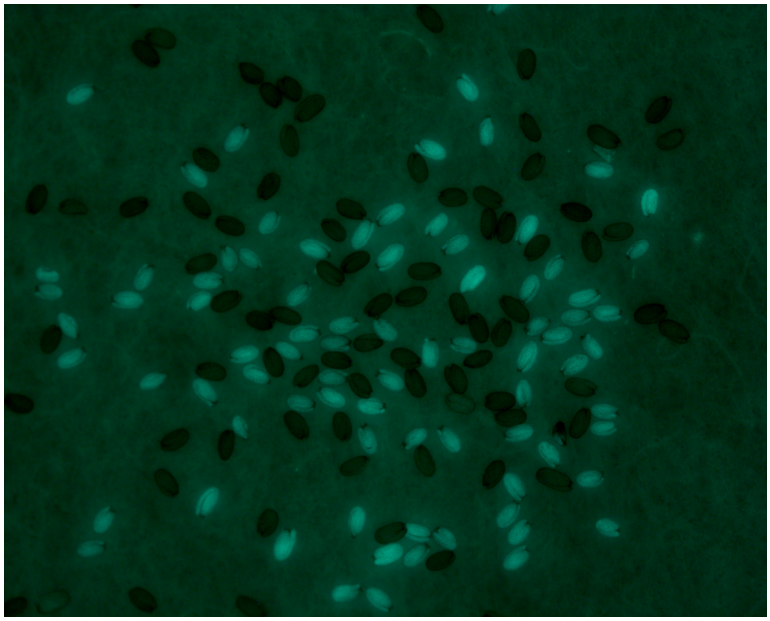


Figure 5.8: T_2 seeds of 35S::YPF3G line 7.

in the seeds could have a strong effect on the expression levels of the genes. As the fluorescence emitted by the YFP3G is only one sixth of the levels emitted by the YFP, stronger levels of expression would be positive for the further experiments, and therefore a qRT-PCR gene expression analysis was performed on the T2 seeds of the transgenic lines, the results of which are presented in Figure 5.10.

As can be seen in Figure 5.10, the expression levels of all the YFP constructs with

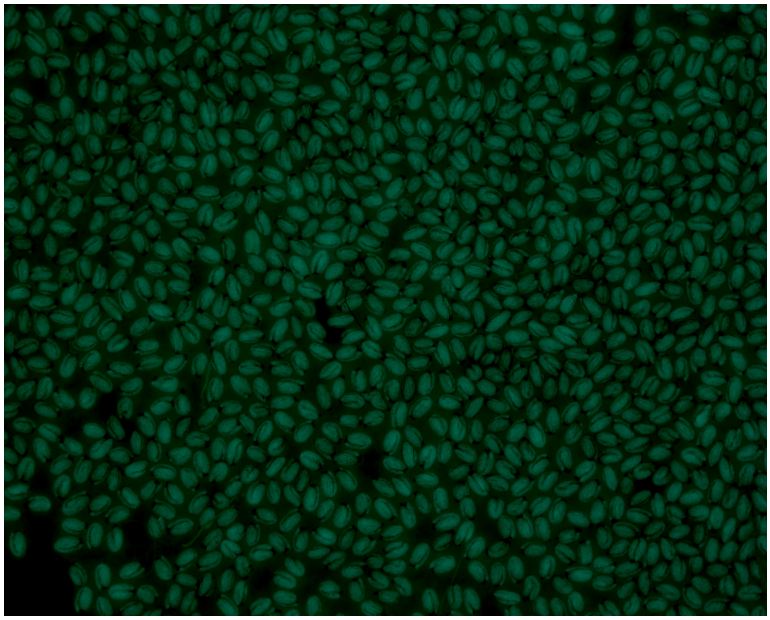


Figure 5.9: T_3 seeds of the 35S::YFP line 6.1. All seeds are glowing, indicating homozygosity.

the 35S promoter were zero. T_3 plants were still grown of these lines, but on imaging the seedlings with a fluorescence microscopy, no fluorescence was found.

5.11 Preliminary confocal fluorescence microscopy study of cotyledons from T_3 plants

To explore how strong the fluorescent signals from the plants were, as well as how large the differences between individual plants of the same line would be, a preliminary study was set up. In the study, 6 days old seedlings grown in $\frac{1}{2}$MS from the transgenic lines UBQ::YFP line3.5, UBQ::YFP3G line2.3, and UBQ::YFP3G line4.2 were subjected to a 3 h treatment of either a $\frac{1}{2}$MS with DMSO used as a mock control or $\frac{1}{2}$MS with 300 mM sorbitol. The latter used to decrease the osmotic difference between the inside of the plant cell and the surroundings, and thereby lower the turgor pressure of the plant cells. Examples of the images from this study are found in Figure 5.11.

There was observed a large difference in fluorescence intensity between the seedlings expressing YFP and YFP3G, which, based on the difference in fluorescence intensity between the wild type YFP and the mutant reported by Watanabe et.al., as well as the large difference in basal expression levels revealed by the qRT-PCR, was expected. In part C, F and I of Figure 5.11, though, the intensity was remarkably lower in the root tips of the seedlings expressing YFP3G than in both those expressing YFP and the wild type Col-0 seedlings. Based on this, the root tips were excluded from further analysis.

It was also observed that the bulges in the cotyledons made some of the images capture

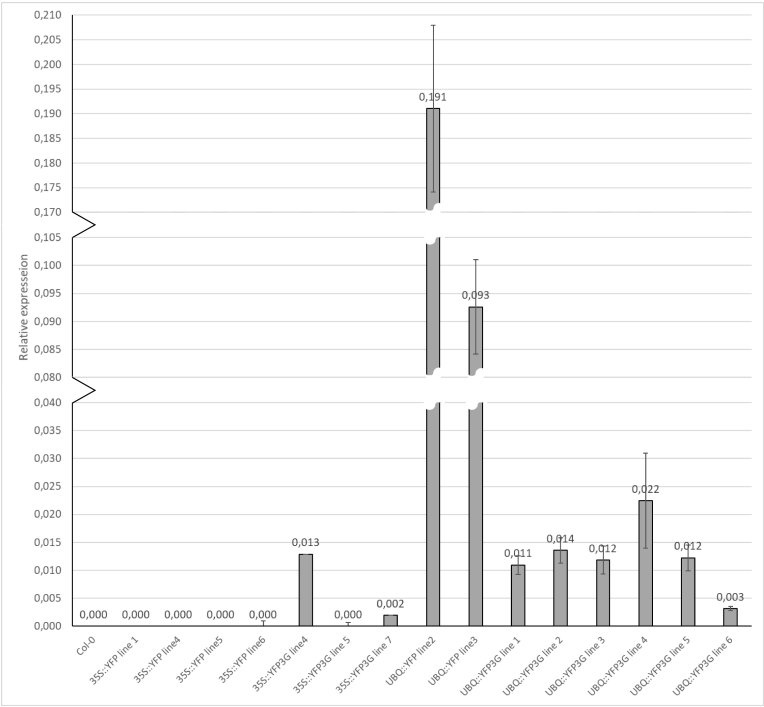


Figure 5.10: Gene expression analysis with qRT-PCR.

both the epithelial cells and what appears to be the palisade mesophyll cells. This can be seen in Figure 5.12, where a "ridge" of palisade mesophyll cells are going diagonally through the image, with epithelial cells on both sides.

The images were analysed in the software tool ImageJ with two different methods. One analysed the images taken of the cotyledons using a 2.5x objective, and one analysed the epithelial cells imaged with a 20x objective, both explained in section 4.14.1. The results of the first method, presented in Figure 5.13, indicated an about 7 times higher fluorescence intensity in the YFP expressing seedlings than in the two lines expressing YFP3G. This method also gave a slight increase in mean fluorescence for all the lines, but using the Student's t-test revealed that only YFP3G line 2.3 had a p-value of less than 0.05. The variations in fluorescence between the individual plants of one line were also not very big, indicated by the standard deviations in Figure 5.13.

The second method, dividing the average fluorescence intensity by the number of cells, produced contradictory results to the first method. The insignificant increase in YFP line 3.5 and YFP3G line 4.2 had changed to a significant decrease (p ¡ 0.05), while YFP3G line 2.3 now had an insignificant increase. The variations were roughly the same, with some of the lines having an increased standard deviation, and some having a decrease.

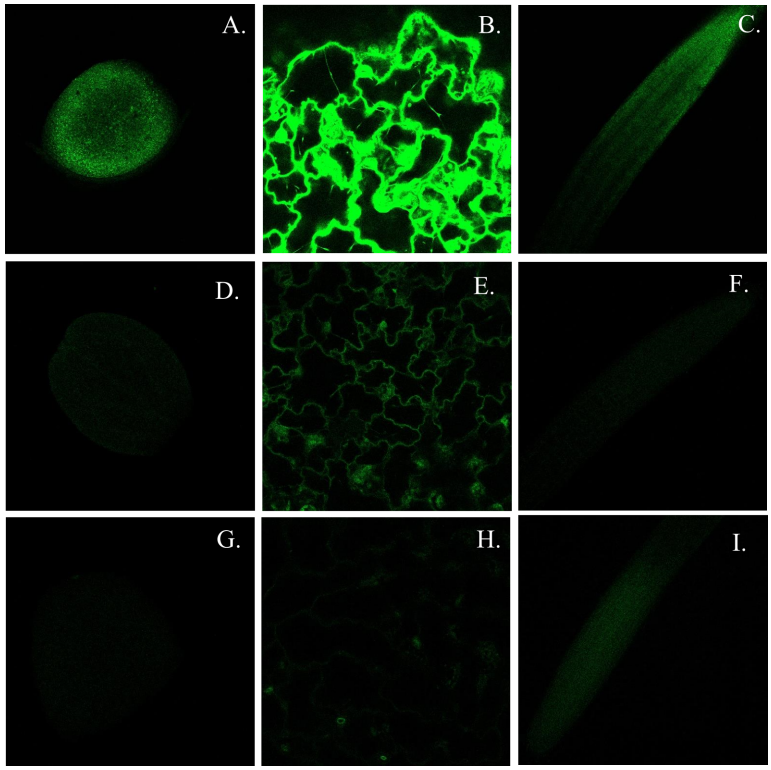


Figure 5.11: Comparison between the fluorescence from UBQ10::YFP line 3.5, UBQ10::YFP3G line 4.2. and the wild type Col-0. All pictures are from mock treated plants. A-C: The whole cotyledon, epithelial cells, and the root tip of seedlings expressing YFP. D-F: The whole cotyledon, epithelial and the root tip of seedlings expressing YFP3G. G-I: The whole cotyledon, epithelial cells and the root tip of seedlings expressing for the wild type Col-0.

5.12 Confocal fluorescence microscopy study of cotyledons from T_3 plants

The final microscopy study used the transgenic lines UBQ::YFP line 3.5, UBQ::YFP3G line 4.2 and UBQ::YFP3G line 2.3. The study used 6 days old seedlings grown in $\frac{1}{2}$MS, and exposed them to either a DMSO mock treatment, or a sorbitol or sodium chloride treatment of either 150 mM, 300 mM or 600 mM. YFP3G line 2.3 was only exposed to 300 mM treatments, while the two other lines were exposed to all concentrations. All treatments lasted for 30 min. The varying sorbitol and sodium chloride treatments was used to induce a gradual decrease in the turgor pressure of the plant cells, which, according to Watanabe et.al. should lower the fluorescence intensity of the YFP3G. After the treatment, the cotyledons of the seedlings were cut and imaged as a Z-stack. Out of the images taken of each cotyledon, three consecutive images were combined into a Z-projection. In this projection, an area where all three images give clear outlines of the same epithelial

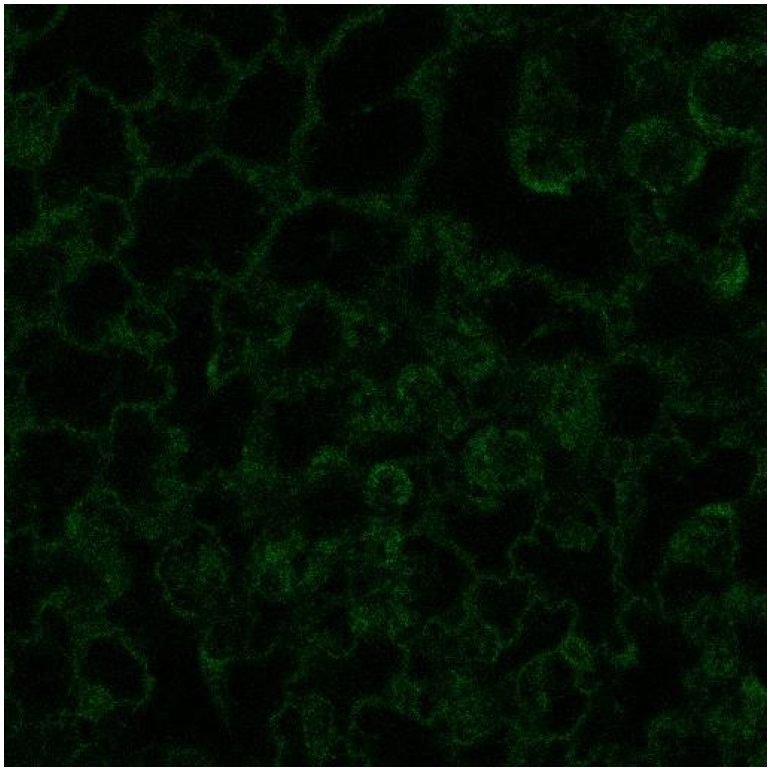


Figure 5.12: Image taken with 20x objective of YFP3G line 4.2 showing palisade mesophyll cells between epithelial cells.

cells was analysed in ImageJ and the average fluorescence was divided by the manually counted number of cells in the image. For more details see section 4.15. Examples of the Z-projections created with ImageJ are given in Figure 5.15. A trend seen in these images is that the mean fluorescence increases with higher concentrations of solutes. This trend was also backed up by the data of mean fluorescence per area. The fluorescence is also more spread throughout the cells, with the dark imprint of the vacuole seen in the DMSO treated cells blurred out by fluorescence in the solute treated images. This was especially bad in YFP3G line 4.2 as the combination of blurred cells after treatment with 600 mM NaCl and low fluorescence intensity made it impossible to distinguish the cells. These images were therefore not analysed.

The data for fluorescence intensity per cell can be found in Figures 5.16, 5.17, and 5.18. The only significant results found in this data was between the DMSO treatment and the 300 mM sodium chloride treatment for both lines expressing YFP3G. For line 4.2 the intensity dropped from 0.506 ± 0.036 with the mock treatment to 0.346 ± 0.050 with the 300 mM NaCl treatment (Student's t-test, p ¡ 0.001), while the line 2.3 dropped from 0.472 ± 0.037 to 0.394 ± 0.023 (Student's t-test, p ¡ 0.01).

For the sorbitol treated plants, a slight increase was observed for YFP line 3.5 illus-

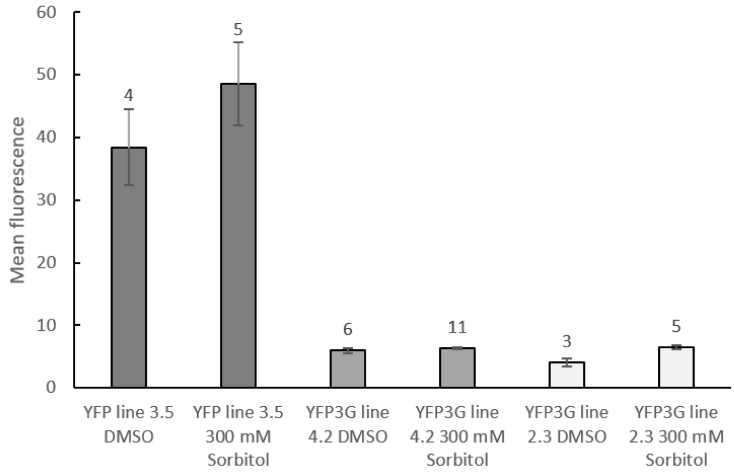


Figure 5.13: The results from analysing the mean fluorescence intensity a selected ROI of the cotyledons after 3 h of DMSO or Sorbitol treatment, with y-axis in arbitrary units. Data sets are created by an average of the images analysed for each line and treatment, with the labels indicating the number of images analysed for each set. The error bars are ± SD.

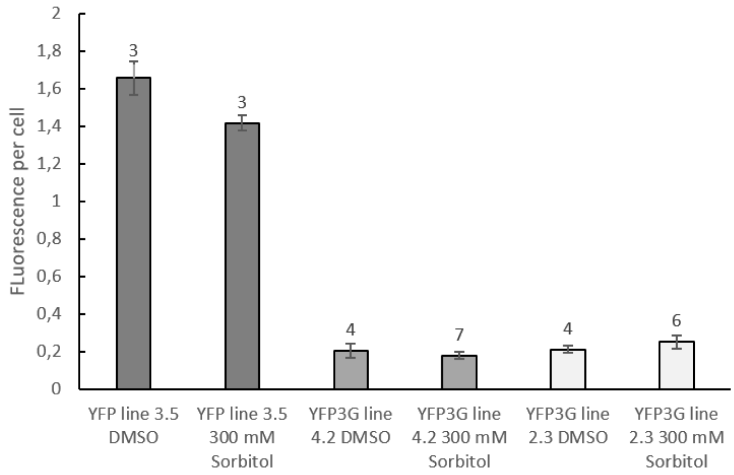


Figure 5.14: The results from analysing the fluorescence intensity per cell of the cotyledons after 3 h of DMSO or Sorbitol treatment, with y-axis in arbitrary units. Data sets are created by an average of the images analysed for each line and treatment, with the labels indicating the number of images analysed for each set. The error bars are ± SD. It is clear that the defined "borders" of the cells, constituted by the cytosol containing the fluorescent proteins pressed up against the cell wall, become more diffuse with the higher concentrations.

trated by the trendline shown in Figure 5.16. The YFP3G line 4.2, on the other hand, showed neither an increase nor a decrease. This result was also found for YFP3G line 2.3

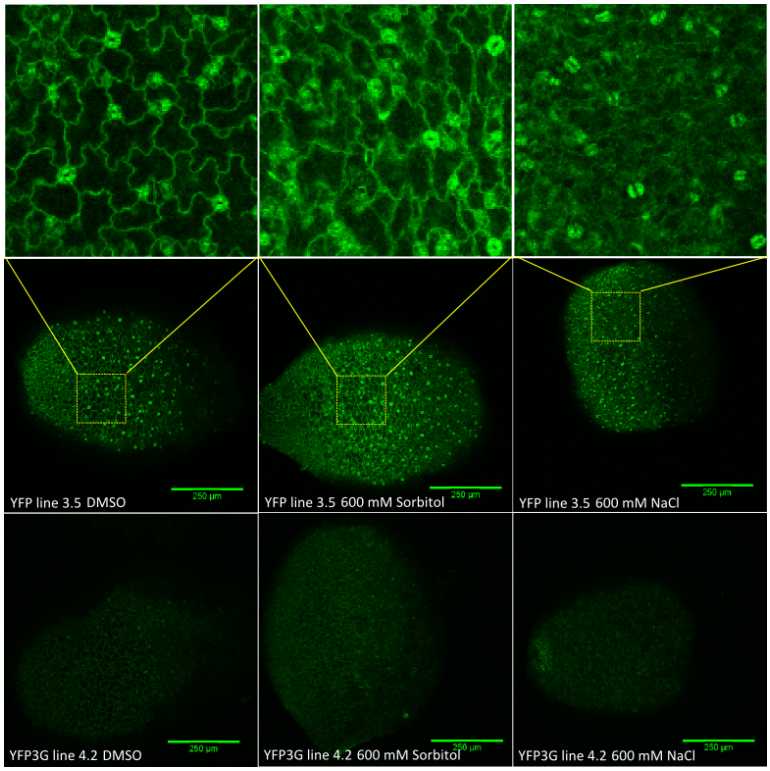


Figure 5.15: Pictures showing Z-projection images of YFP line 3.5 and YFP3G line 4.2. The top part is an excerpt of the Z-projection images the same size as the ROI that were analysed and in which the number of cells were counted.

which had a slight decrease of 5 % between the mock treatment and the 300 mM sorbitol treatment, that is well within the standard deviations.

These results are contrasted with the sodium chloride treated plants, where the trend-lines show a decrease for all lines. As already mentioned, both lines expressing YFP3G had a significant decrease in fluorescence intensity per cell, while the decrease seen in YFP line 3.5 was not.

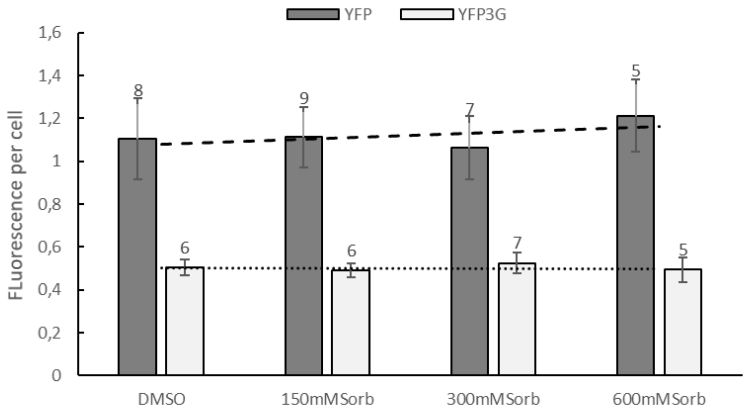


Figure 5.16: Results from analysing the fluorescence intensity per cell of YFP line 3.5 and YFP3G line 4.2, with y-axis in arbitrary units. Data sets are created by an average of the images analysed for each line and treatment, with the labels indicating the number of images analysed for each set. The error bars are ± SD. Dotted lines are trendlines calculated by Excel.

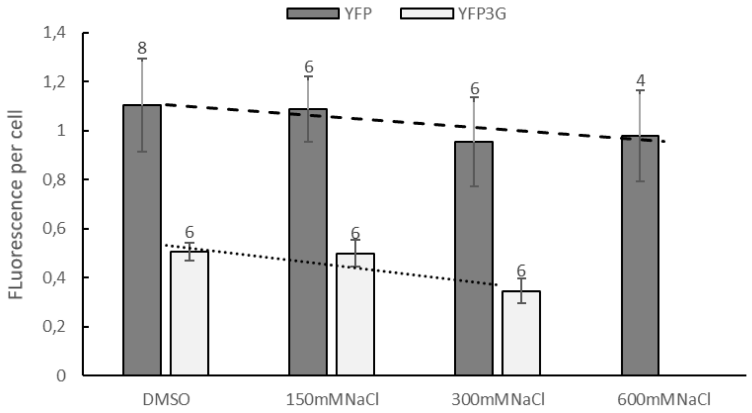


Figure 5.17: Results from analysing the fluorescence intensity per cell of YFP line 3.5 and YFP3G line 4.2, with y-axis in arbitrary units. The final data set at 600 mM NaCl for the YFP3G line was not analysed, as the cells were too blurred to be properly counted. Data sets are created by an average of the images analysed for each line and treatment, with the labels indicating the number of images analysed for each set. The error bars are ± SD. Dotted lines are trendlines calculated by Excel.

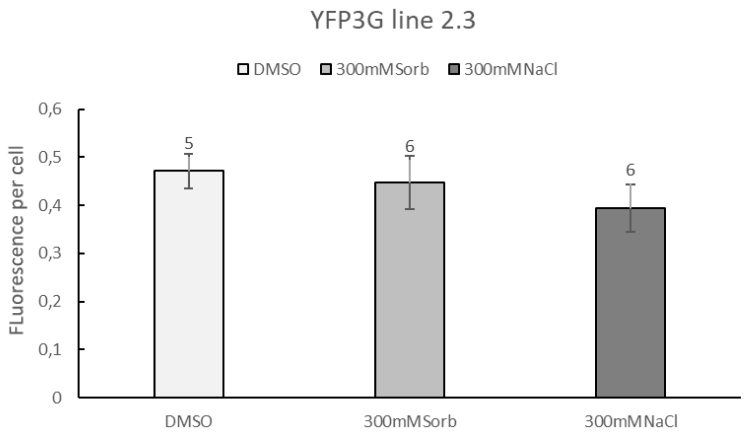


Figure 5.18: Results from analysing the fluorescence intensity per cell of YFP3G line 2.3, with y-axis in arbitrary units. Data sets are created by an average of the images analysed for each line and treatment, with the labels indicating the number of images analysed for each set. The error bars are $\pm$ SD.

Chapter 6

Discussion

6.1 MD simulations

An MD simulation is only as good as the parameters you put into it, so the final results of the simulations are heavily influenced by the initial work done on the completion of the PDB structures and the parameters calculated for the new CR2 chromophore added to the amino acid list in the force field. The changes made to the PDB structure can be criticized in that the structure of the proteins in the PDB files are not the full YFP protein. On the other hand, in an effort to reduce the amount of movement of the simulated proteins, combined with studies indicating that the missing residues don't affect the fluorescence of the protein (Cetinkaya et al., 2005) this was believed to be justified. Much care was taken with the preparations before generating the new force fields parameters, but the risk of something going wrong is big. A mistake done at this stage, combined with the final preparations of the PDB file, was calculating the parameters for the protonated chromophore instead of the anionic form discussed in section 2.3 and using this form of the chromophore in the simulations. This can perhaps not explain the major movements seen in the chromophore, but would in a more detailed study of YFP3G be a major flaw.

The fluctuations of the RMSD values seem to be well within what is normal, indicating that the stability of the protein is reasonable. As shown by Sargsyan et al. (2017), a small protein simulated over 100 ns can have rmsd value fluctuations between 0.15-0.21. This is about the same as the RMSD values calculated here for the medium sized YFP molecule. The RMSF values on the other contain much larger fluctuations than comparable studies, like the one done by Jacchetti et al. (2016) on the GFP. The shorter simulations run before the production run was 200ps in the Jacchetti study, compared to the 100ps used here. Combined with a more gradual lowering of the restraints on the protein, this could be a reason for more energetic movements of the protein seen in this study.

A shortening of the distance between the chromophore and Tyr206 was by Watanabe et.al. believed to be the main contributor to the shift to longer wavelengths seen in the fluorescence of both YFP and YFP3G at higher pressure levels. The results from these simulations does not give a conclusive answer to this. For YFP, the distances are relatively

stable for the last 40 ns of the simulation, while some larger changes can be seen in the last 10 ns of the simulations for the YFP3G molecules. It is difficult to say from the data presented here whether these changes are the result of conformational changes in the protein, or just random movements within an already established conformation.

The explanation given by the Watanabe group for the pressure sensitivity of YFP3G was the accessibility of water molecules from the environment to the chromophore. The results from the simulations indicate that the number of water molecules inside the β-barrel and close to the chromophore is much more fluctuating for YFP3G than for YFP, which may be attributed to a larger exchange with the bulk water outside the protein. This cannot be said with total confidence, as the molecules might move out of the 1 nm and 0.6 nm spheres where water atoms are counted, but still be inside the protein. There generally seems to be more water molecules inside the β-barrel for the YFP3G than for YFP3G up to 10 MPa, but no lowering of the number of molecules can be seen with increasing pressure up to this point. At 50 MPa, on the other hand,a change is clearly seen after 60 ns, with the number of water molecules close to the chromophore falling to nearly zero, with a large drop for the general number inside the β-barrel as well. This drop corresponds closely in time with an increase in the distance between the chromophore and Tyr206 in the same simulation. These two changes happening at the same time might be an indication that the structure have moved into a new conformation, but more in depth analysis is required to say anything for certain.

All in all, the major changes to the proteins in the simulations seem to happen with the structures simulated at 50 MPa. Normally, proteins would become more rigid and less flexible at higher pressure levels, as seen in the study by Jacchetti et al. (2016), but this is not seen in this study. The results therefore have to be considered suspicious.

6.2 Biological part

When the constructs were being cloned, tests were done to confirm the correct recombination, so that all of the transgenic plants transformed with 35S::YFP did not express is not easily explained. Still, transgenic plants expressing both YFP and YFP3G with the ubiquitin-10 promoter were created and fluoresced.

The results of the preliminary confocal fluorescence microscopy study confirmed that the cells were fluorescing, and that the fluorescence of the plants expressing YFP3G was not to weak to measured. The images taken in this session was quite crude and not meant to be part of any final analysis, but the results were still quite consistent. The data also hinted towards a trend seen in the final study, namely that the mean fluorescence of the ROI increases when the seedlings where exposed to higher concentration of solutes, but the average intensity per cell does not. The method of manually counting the cells can at the same time be thought to introduce a new error in the analysis of the images. What constitutes a cell in grainy images (as they became with the more concentrated solutions) is not a clear answer. It was attempted to use automatic cell counting software (MIPAR) to do the counting, but because of the low and varying quality of the images, this was difficult to do. Still, especially in the preliminary test where the number of cells in each image were so small, a few large cells might skew the average of the image by introducing larger darker areas. As surface divided by area gets lower with increasing size, the average

would fall. In addition to this, the average number of cells per image was higher in the solute treated cotyledons. Considering that the treatments was meant to lower the osmotic difference across the cell membrane and therefore cause the cells to lose water and shrink (Veley et al., 2012), this increase should be expected. The counting of cells is therefore a way to remove the effect this increase will have on the average intensity of the images. At the same time, there is also the risk that narrow parts of the jigsaw puzzle-shaped pavement epithelial cells can shrink and, in the more blurry images, be counted as two cells. Still, as the exact definition of what constituted a cell in the picture required a bit of subjectivity, great care was taken to maintain a consistent counting between the images. In addition, the cells were counted before the fluorescence measurements were taken to avoid any subconscious altering of the numbers.

The results do not show any significant difference in the fluorescence between seedlings expressing YFP and YFP3G. Both have practically no change between the DMSO mock treatment and the sorbitol treatments, while the decrease in the NaCl treatments are also of similar size. At the same time, there is a slight increase in the sorbitol treated YFP line, while a flat or lowered trend for the YFP3G expressing lines, and combined with the fact that the only significant decreases in fluorescence per cell came from the lines expressing YFP3G, this might be a hint towards an effect from the lower turgor pressure levels. There are large uncertainties in this, though, especially from the increased blurriness of the outline of the cells cytosol. In Watanabe et al. (2013), the calculated increase in fluorescence over 1 MPa is only 1 %, so that no effect is clear here is perhaps not surprising.

Bibliography

Bayly, C. I., Cieplak, P., Cornell, W., Kollman, P. A., 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. The Journal of Physical Chemistry 97 (40), 10269–10280.

Beauzamy, L., Nakayama, N., Boudaoud, A., 2014. Flowers under pressure: Ins and outs of turgor regulation in development. Annals of Botany 114 (7), 1517–1533.

Borg, F. G., May 2003. What is osmosis? Explanation and understanding of a physical phenomenon. ArXiv Physics e-prints.

Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormö, M., Remington, S. J., 1997. Structural basis for dual excitation and photoisomerization of the aequorea victoria green fluorescent protein. Proceedings of the National Academy of Sciences of the United States of America 94 (6), 2306–2311.

Cetinkaya, M., Zeytun, A., Sofo, J., Demirel, M. C., 2005. How do insertions affect green fluorescent protein? Chemical Physics Letters 419 (1-3), 48–54.

Chattoraj, M., King, B. A., Bublitz, G. U., Boxer, S. G., 1996. Ultra-fast excited state dynamics in green fluorescent protein: Multiple states and proton transfer. Proceedings of the National Academy of Science of the United States of America 93 (16), 8362–8367.

Clough, S. J., Bent, A. F., 1998. Floral dip: A simplified method for agrobacterium-mediated transformation of arabidopsis thaliana. The Plant Journal 16 (6), 735–743.

Cosgrove, D. J., 2005. Growth of the plant cell wall. Nature Reviews Molecular Cell Biology 6 (11), 850–861.

Cuendet, M. A., van Gunsteren, W. F., 2007. On the calculation of velocity-dependent properties in molecular dynamics simulations using the leapfrog integration algorithm. The Journal of Chemical Physics 127 (18), 184102.

Dragan, A., Pavlovic, R., McGivney, J. B., Casas-Finet, J. R., Bishop, E. S., Strouse, R. J., Schenerman, M. A., Geddes, C. D., 2012. Sybr green i: Fluorescence properties and interaction with dna. Journal of Fluorescence 22 (4), 1189–1199.

Dupradeau, F.-Y., Pigache, A., Zaffran, T., Savineau, C., Lelong, R., Grivel, N., Lelong, D., Rosanski, W., Cieplak, P., 2010. The r.e.d. tools: advances in resp and esp charge derivation and force field library building. Phys. Chem. Chem. Phys. 12, 7821–7839.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., Pedersen, L. G., 1995. A smooth particle mesh ewald method. The Journal of Chemical Physics 103 (19), 8577–8593.

F. Wang, J.-P. Becker, P. C. . F.-Y. D., 2013. R.e.d. python: Object oriented programming for amber force fields. Université de Picardie, Sanford Burnham Prebys Medical Discovery Institute 1 (1), 1.

Gonzáles, M. A., 2011. Force fields and molecular dynamics simulations. Collection SFN 12, 169–200.

Grefen, C., Donald, N., Hashimoto, K., Kudla, J., Schumacher, K., Blatt, M. R., 2010. A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. The Plant Journal 64 (2), 355–365.

GROMACS, 2016. GROMACS Reference Manual. 2016th Edition.

Hamann, T., 2015. The plant cell wall integrity maintenance mechanism – concepts for organization and mode of action. Plant and Cell Physiology 56 (2), 215–223.

Hamant, O., Haswell, E. S., 2017. Life behind the wall: Sensing mechanical cues in plants. BMC Biology 15 (1), 59.

Hansson, T., Oostenbrink, C., van Gunsteren, W. F., 2002. Molecular dynamics simulations. Current Opinion in Structural Biology 12, 190–196.

Haswell, E. S., Verslues, P. E., 2015. The ongoing search for the molecular basis of plant osmosensing. Journal of General Physiology 145 (5), 389–394.

Hedrich, R., 2012. Ion channels in plants. Physiological Reviews 92 (4), 1777–1811.

Hospital, A., Goñi, J. R., Orozco, M., Gelpí, J. L., 2015. Molecular dynamics simulations: Advances and applications. Advances and Applications in Bioinformatics and Chemistry 8, 37–47.

Jacchetti, E., Gabellieri, E., Cioni, P., Bizzarri, R., R., N., 2016. Temperature and pressure effects on gfp mutants: Explaining spectral changes by molecular dynamics simulations and td-dft calculations. Physical Chemistry Chemical Physics 18 (18), 12828–12838.

Jung, G., Wiehler, J., Zumbusch, A., 2005. The photophysics of green fluorescent protein: Influence of the key amino acids at positions 65, 203, and 222. Biophysical Journal 88 (3), 1932–1947.

Kouzine, F., Levens, D., Baranello, L., 2014. Dna topology and transcription. Nucleus 5 (3), 195–202.

Meinhold, L., Smith, J. C., Kitao, A., Zewail, A. H., 2007. Picosecond fluctuating protein energy landscape mapped by pressure-temperature molecular dynamics simulation. Proceedings of the National Academy of Sciences of the United States of America 104 (44), 17261–17265.

Merabti, K. E., Azizi, S., Ridard, J., Lévy, B., Demachy, I., 2017. π -Stacking interactions in YFP, quantum mechanics and force field evaluations in the S_0 and S_1 states. Chemical Physics 493, 157–165.

Mu, Y., Kosov, D. S., Stock, G., 2003. Conformational dynamics of trialanine in water. 2. comparison of amber, charmm, gromos, and opls force fields to nmr and infrared experiments. Journal of Physical Chemistry B 107, 5064–5073.

Murail, S., 2017. Simulation of Ligand Binding to Membrane Proteins. Springer New York.

Nadeau, J. A., Sack, F. D., 2002. Stomatal development in arabidopsis. The Arabidopsis Book/American Society of Plant Biologists 1 (1), 1.

Ormö, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., J., R. S., 1996. Crystal structure of the aequorea victoria green fluorescent protein. Science 273 (5280), 1392–1395.

Pitera, J. W., 2014. Expected distributions of root-mean-square positional deviations in proteins. The Journal of Physical Chemistry B 118 (24), 6526–6530.

Pitzschke, A., 2013. Agrobacterium infection and plant defense – transformation success hangs by a thread. Frontiers in Plant Science 4 (519), 1–12.

Sargsyan, K., Grauffel, C., Lim, C., 2017. How molecular size impacts rmsd applications in molecular dynamics simulations. Journal of Chemical Theory and Computation 13 (4), 1518–1524.

Schindelin, J., Rueden, C. T., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., Eliceiri, K. W., 2017. Imagej2: Imagej for the next generation of scientific image data. BMC Bioinformatics 18 (1), 529.

Schreiner, W., Karch, R., Knapp, B., Ilieva, N., 2012. Relaxation estimation of rmsd in molecular dynamics immunosimulations. Computational and Mathematical Methods in Medicine 2012 (173521), 1–9.

Shabala, S. N., Lew, R. R., 2002. Turgor regulation in osmotically stressed arabidopsis epidermal root cells. direct support for the role of inorganic ion uptake as revealed by concurrent flux and cell turgor measurements. Plant Physiology 129 (1), 290–299.

Shimada, T. L., Shimada, T., Hara-Nishimura, I., 2010. A rapid and non-destructive screenable marker, fast, for identifying transformed seeds of arabidopsis thaliana. The Plant Journal 61 (3), 519–528.

Shirts, M. R., 2013. Simple quantitative tests to validate sampling from thermodynamic ensembles. Journal of Chemical Theory and Computation 9 (2), 909–926.

Shu, X., Leiderman, P., Gepshtein, R., Smith, N. R., Kallio, K., Huppert, D., Remington, S. J., 2007. An alternative excited-state proton transfer gateway in green fluorescent protein variant s205v. Protein Science 16 (12), 2703–2710.

Smith, M. D., Rao, J. S., Segelken, E., Cruz, L., 2015. Force-field induced bias in the structure of aβ21−30: A comparison of opls, amber, charmm, and gromos force fields. Journal of Chemical Information and Modeling 55 (12), 2587–2595.

Tsien, R. Y., 1998. The green fluorescent protein. Annual Review of Biochemistry 67, 509–544.

VanGuilder, H. D., Vrana, K. E., Freeman, W. M., 2008. Twenty-five years of quantitative pcr for gene expression analysis. Biotechniques 44 (5), 619–626.

Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J. C., Cieplak, P., Dupradeau, F.-Y., 2011. R.e.d. server: a web service for deriving resp and esp charges and building force field libraries for new molecules and molecular fragments. Nucleic Acids Research 39 (2), W511–W517.

Veley, K. M., Marshburn, S., Clure, C. E., Haswell, E. S., 2012. Mechanosensitive channels protect plastids from hypoosmotic stress during normal plant growth. Current Biology 22 (5), 408–413.

Watanabe, T. M., Imada, K., Yoshizawa, K., Nishiyama, M., Kato, C., Abe, F., Morikawa, T. J., Kinoshita, M., Fujita, H., Yanagida, T., 2013. Glycine insertion makes yellow fluorescent protein sensitive to hydrostatic pressure. PLOS ONE 8, 1.

Wolf, S., Greiner, S., 2012. Growth control by cell wall pectins. Protoplasma 249 (2), 169–175.

Xiao, Y., Tholen, D., Zhu, X.-G., 2016. The influence of leaf anatomy on the internal light environment and photosynthetic electron transport rate: exploration with a new leaf ray tracing model. Journal of Experimental Botany 67 (21), 6021–6035.

Yong-Qiang, A., McDowell, J. M., Huang, S., McKinney, E. C., Chambliss, S., Meagher, R. B., 1996. Strong, constitutive expression of the arabidopsis act2/act8 actin subclass in vegetative tissues. The Plant Journal 10 (1), 107–121.

Zhang, B., Chandrasekhar, M., Chandrasekhar, H. R., 1985. Pressure-induced shifts of the fluorescence spectrum of rhodamine 6g in solution. Applied Optics 24 (17), 2779–2782.

Zonia, L., Munnik, T., 2007. Life under pressure: hydrostatic pressure in cell growth and function. Trends in Plant Science 12 (3), 90–97.

Appendix A

Generated force field parameters

CD 12.01000 ;

C CC 1 0.14620 315808.3 ;
CK CT 1 0.14990 282252.6 ;
CA CD 1 0.14340 344510.6 ;
CC CD 1 0.14290 350033.4 ;
CD HA 1 0.10850 290537.0 ;

CK CT N 1 111.710 559,317 ;
CK CT H1 1 111.620 393.547 ;
CT CK NB 1 120.940 551.451 ;
CT CK N* 1 120.940 551.451 ;
CC NB CK 1 104.340 589.944 ;
CD CC NB 1 121.690 565.091 ;
C CC NB 1 123.470 553.878 ;
CD CC C 1 122.690 533.209 ;
CC CD HA 1 120.860 394.468 ;
CC CD CA 1 111.040 565.091 ;
HA CD CA 1 124.040 388.275 ;
CA CA CD 1 120.100 552.204 ;
CC C O 1 125.710 576.639 ;
CC C N* 1 113.750 574,463 ;
C N* CK 1 120.480 553.794 ;
C CT N* 1 106.510 570.698 ;

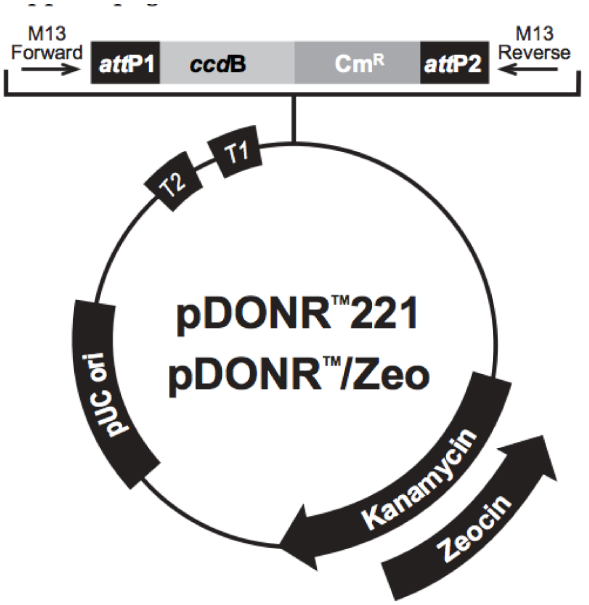
C CK N* CT 4 180.00 4.60240 2 ;

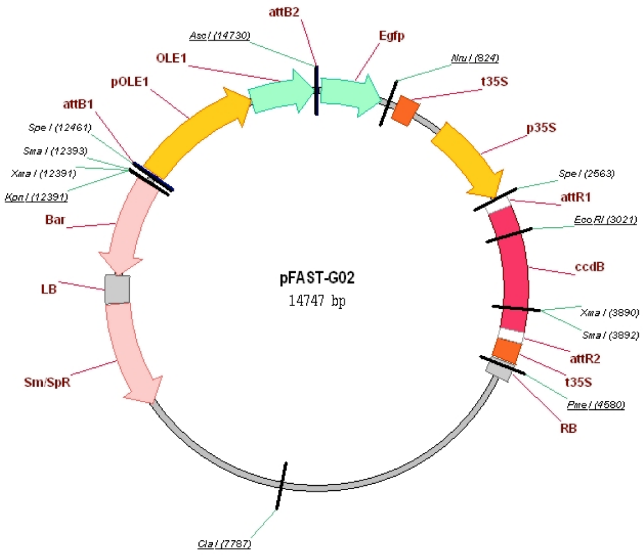
```
CC N* C O 4 180.00 43.93200 2 ;
CT N* CK NB 4 180.00 43.93200 2 ;
C CD CC NB 4 180.00 4.60240 2 ;
CA CA CA HA 4 180.00 4.60240 2 ;
CA CA CA CD 4 180.00 4.60240 2 ;



    NB CK CT N 9 0.0 0.00000 3 ;
N* CK CT N 9 0.0 0.00000 3 ;
NB CK CT H1 9 0.0 0.00000 3 ;
N* CK CT H1 9 0.0 0.00000 3 ;
NB CC CD HA 9 180.0 16.73600 2 ;
NB CC CD CA 9 180.0 16.73600 2 ;
C CC CD HA 9 180.0 16.73600 2 ;
C CC CD CA 9 180.0 16.73600 2 ;
O C CC NB 9 180.0 12.02900 2 ;
N* C CC NB 9 180.0 12.02900 2 ;
O C CC CD 9 180.0 12.02900 2 ;
N* C CC CD 9 180.0 12.02900 2 ;
CA CA CD CC 9 180.0 2.92880 2 ;
CA CA CD HA 9 180.0 2.92880 2 ;



    [ CR2 ]
[ atoms ]
N N -0.42070 1
H H 0.29420 2
CA1 CT -0.06750 3
HA11 H1 0.12240 4
HA12 H1 0.12240 5
C1 CK 0.20830 6
N1 NB -0.40760 7
CA2 CC 0.10770 8
CB CD -0.15640 9
HB HA 0.13470 10
CG CA 0.04260 11
CD1 CA -0.08230 12
HD1 HA 0.12230 13
CE1 CA -0.29560 14
HE1 HA 0.17440 15
CE2 CA -0.29560 16
HE2 HA 0.17440 17
CD2 CA -0.08230 18
HD2 HA 0.12230 19
CZ C 0.34350 20
```

OH OH -0.52130 21
HH HO 0.37910 22
C2 C 0.39840 23
O2 O -0.52880 24
N2 N* -0.04750 25
CA3 CT 0.09210 26
HA31 H1 0.04440 27
HA32 H1 0.04440 28
C C 0.56630 29
O O -0.58830 30

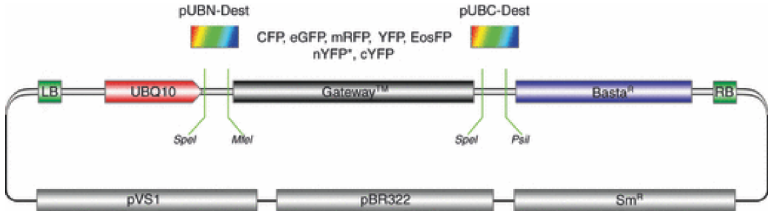
Appendix B

Vector maps




attB2
Asc I (14730)
Egfp
OLE1
Nru I (824)
pOLE1
t35S
attB1
p35S
Spe I (12461)
Sma I (12393)
Spe I (2563)
Xma I (12391)
attR1
Kpn I (12391)
EcoRI (3021)
Bar
ccdB
pFAST-G02
14747 bp
LB
Xma I (3890)
Sma I (3892)
attR2
Sm/SpR
t35S
Pme I (4580)
RB
Cla I (7787)


pFAST-G02
attR1
attR2
LB
bar
OLE1-GFP
35S promoter
CmR-ccdB
RB
500 bp
35S terminator


pUBN-Dest
pUBC-Dest
CFP, eGFP, mRFP, YFP, EosFP
nYFP*, cYFP
LB
UBQ10
Gateway™
Basta^R
RB
SpeI
MfeI
SpeI
PsiI
pVS1
pBR322
Sm^R