Novel Approaches to the Production & Analysis of Biological Nanomaterials for Serial-Femtosecond X-ray Crystallography

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von
Thomas Seine

Hamburg
2018

Final version

Author e-mail: thomas.seine(at)protonmail.com

Gutachter/innen der Dissertation:	Prof. Dr. Henry Chapman Prof. Dr. Matthias Wilmanns
Zusammensetzung der Prüfungskommission:	Prof. Dr. Daniela Pfannkuche Prof. Dr. Arwen Pearson Prof. Dr. Henry Chapman Prof. Dr. Matthias Wilmanns Prof. Dr. Lars Redecke
Vorsitzende/r der Prüfungskommission:	Prof. Dr. Daniela Pfannkuche
Datum der Disputation	23.08.2018
Vorsitzender Fach-Promotionsausschuss PHYSIK:	Prof. Dr. Wolfgang Hansen
Leiter des Fachbereichs PHYSIK:	Prof. Dr. Michael Potthoff
Dekan der Fakultät MIN:	Prof. Dr. Heinrich Graener

Eidesstattliche Versicherung / Declaration on Oath

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium.

Die Dissertation wurde in der vorgelegten oder einer ähnlichen Form nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Hamburg, den 02.05.2018

Discere ne cessa, cura sapientia crescat.
-Disticha catonis

Zusammenfassung

Im Überschneidungsbereich der seriellen Kristallographie und der Anwendung ultrakurzer Röntgenpulse an Freie-Elektronen-Lasern gab es zuletzt wesentliche technische Fortschritte und Neuentwicklungen. Mit der hier vorgelegten Arbeit wird angestrebt, einen Beitrag zu eben diesen Entwicklungen zu leisten. Dies ist umgesetzt durch die Erforschung neuer Methodiken die die verlässliche Produktion von benötigten Proben im Nanometerbereich ermöglichen.
Im Speziellen widmet sich die Arbeit der Erforschung des peroxisomalen Systems der Hefen *H. polymorpha* und *S. cerevisiae*. Die damit verbundenen Möglichkeiten, dieses als Produktionsstätte und Schutzsystem für kristallisierte, heterolog-exprimierte Proteine zu nutzen, werden im Detail untersucht. Die bekannten Eigenschaften dieses Systems, speziell seine Größe, Robustheit und Vielseitigkeit, unterstreichen die gute Einsetzbarkeit dieses eukaryotischen Systems. Am Beispiel von natürlich vorkommenden AOX Kristallen in den Peroxisomen der Hefe *H. polymorpha* wird zunächst gezeigt, dass es möglich ist, klare Diffraktionsbilder von eben diesen zu erhalten, sogar wenn die Hefezellen als Ganzes auf den Röntgenstrahl treffen. Dabei wird die außerordentliche Fähigkeit des Systems deutlich, die fragilen Proteinkristalle auch unter den harschen experimentellen Bedingungen solcher Versuche zu erhalten. Aufbauend auf der Idee einer Hefe-basierten "Kristallfabrik" wird in dieser Arbeit ebenfalls gezeigt, wie sich *in vivo* Nanokristalle bereits durch Synchrotron-basierte Pulverdiffraktion charakterisieren lassen. Darüber hinaus wird eine Williamson-Hall Analyse eingesetzt, um eine Möglichkeit aufzuzeigen, die gewonnenen Daten hinsichtlich der Größe von Kristalldomänen und Verzerrungen im Kristallgitter bereits im Vorfeld von FEL Experimenten zu untersuchen.
Im Bestreben, auch an nicht-kristallinen Proben Forschung zu betreiben, wird in einem letzten Abschnitt auf die erfolgreiche Produktion von SOD1 Amyloidfibrillen eingegangen. Mit Bezug auf Ergebnisse aus verschiedensten Methodiken (DLS, EM, ThT Experiment) wird der Mechanismus diskutiert, durch den sich stabile Fibrillen produzieren lassen. Es wurde auch versucht, diese in Experimenten der seriellen Diffraktion an XFEL-Röntgenquelle zu nutzen.
Des Weiteren wird eine neue Methodik präsentiert, die es ermöglicht, die Akkumulation und ggf. Nukleation/Kristallisation von heterolog-exprimierten Proteinen im peroxisomalen System zu untersuchen. Sie basiert auf der Kombination von fluoreszenten Probemolekülen in Peroxisomen und der Messung von Fluoreszenzanisotropie und Resonanzenergietransfer. Dabei wird zunächst ein theoretisches Modell entwickelt, um diese Parameter genau zu beschreiben und um sie dann in einem weiteren Schritt mit experimentellen Daten zu vergleichen. Am Beispiel von EGFP ist es so möglich nachzuweisen, dass die aufgestellte Hypothese von sehr hohen Proteinkonzentrationen unter den experimentellen Bedingungen nicht verifiziert kann. Offenbar ist die Kristallisation von AOX in Peroxisomen nur bedingt abhängig von größeren Mengen des Proteins. Ebenso scheint es unerlässlich, weitere genetische Modifikationen am Hefesystem vorzunehmen, um auch Kristallisationsbedingungen für andere Proteine zu schaffen. Ausgehend von diesen Untersuchungen wurde auch ein kurzes *Gedankenexperiment* durchgeführt, in dem die Möglichkeit aufgezeigt wird, Proteinansammlungen von Proteinkristallen zu unterscheiden. Hierbei wird der Einfluss des FRET Orientierungsfaktors κ^2 diskutiert.

Summary

Tremendous effort has recently been put into the development of novel methods and techniques in the realm of serial-femtosecond X-ray crystallography (SFX) and X-ray free-electron lasers (XFEL). This thesis strives to contribute to current research approaches. This is done by focusing on the exploration of methods that can accommodate the increasing demand for nano-sized samples in serial data collection.
The presented research focuses on the investigation of the peroxisomal system of the budding yeast species *H. polymorpha* and *S. cerevisiae* in particular. Both are potential *in vivo* factories and carriers of crystalline material from heterologously expressed proteins. The outstanding features of the yeast eukaryotic systems, their size, robustness and versatility indicate optimal conditions for such efforts. Using the example of naturally-occuring peroxisomal AOX crystals in the yeast *H. polymorpha*, it is shown that meaningful diffraction data can be obtained at modern XFEL sources via the injection of whole yeast cells into the X-ray beam. The yeast systems demonstrate robustness under the applied harsh experimental conditions (injection via a GDVN nozzle). Thus, proof is given that yeast cells are a suitable envelope to protect fragile, nano-sized protein crystals. To build up on the idea of creating a productive *in vivo* 'crystal factory', it is further shown how to characterize *in vivo* crystals by means of synchrotron powder diffraction and a Williamson-Hall analysis. These methods enable to quantify crystal domain sizes and strain and can be used in future attempts to evaluate and optimize crystal quality ahead of FEL experiments.
It has also been demonstrated how to obtain superoxide dismutase 1 (SOD1) amyloid-fibers, so as to provide suitable samples for serial data collection from non-crystalline material. In relation to results from dynamic light scattering (DLS), electron microscopy (EM) and fluorescence-binding assays (ThT assay), possible mechanisms are discussed by which protein fibers are formed. Moreover, it has been attempted to utilize the optimized fiber sample for serial femtosecond fiber diffraction at a XFEL source.
In order to investigate the accumulation and possible nucleation/crystallization of heterologously expressed proteins in the peroxisomal environment, a new approach has been developed. It relies on the combination of intraperoxisomal, fluorescent probe molecules and the observation of their steady-state fluorescence anisotropy. A theoretical model is developed to which experimental data can be compared in terms of fluorescence anisotropy and energy transfer. Using the example of the protein EGFP it is shown that yeast cells require further genetic engineering to corroborate the hypothesis of an ever-increasing intraperoxisomal protein concentration. Consequently, the occurrence of AOX crystals seems also not to be induced by extremely high protein concentrations. Proceeding from the established anisotropy method, a *gedankenexperiment* is conducted to demonstrate the possibility to distinguish protein clusters from crystals in case the influence of the FRET orientation factor κ^2 is being considered in measurements of fluorescence anisotropy and homoFRET.

Acknowledgments

It seems to me that the scientific method, the constant search for answers, is the most important, fate-changing momentum in human history. We should never be too scared "to boldly go where no man has gone before"[1]. Likewise, we should also never forget that science is borne by humans, not just by itself. The past four years have been a long and interesting journey. Not only did I learn more about science, but I am also especially grateful for many friendly encounters, for numerous help- or cheerful words and conversations. Now, it is time for me to move on to meet new challenges. Therefore, I would like to thank those who accompanied me during the past years.

Firstly, I want to thank my two supervisors Prof. Dr. Henry Chapman and Prof. Dr. Matthias Wilmanns for all the given opportunities. It has been an exciting time and an honor to work in such a flourishing environment, to be able to collaborate with scientists from all over the world and to take part in the groundbreaking research in the field of X-ray diffraction. It has been especially fascinating to carry out scientific experiments at modern free-electron laser facilities in Japan and the USA. It has been an experience I will never forget.

Secondly, it is vital to thank Margret Fischer and Irmtraud Kleine, who both took care of countless organizational issues during my time at EMBL and CFEL. In the same breath, I also want to say thank you to my graduate school "The Centre for Ultrafast Imaging" (CUI) for their outstanding organizational support.

I would also like to thank many of my former/current colleagues and collaborators for their support. Most notably, Prof. Dr. Wolfgang Schliebs, Dr. Alke Meents, Dr. Miriam Barthelmeß and Dr. Leonard Chavas. All of you were able provide me with new perspectives. This is also true for many new and old friends. Especially Dr. Robin Schubert, Julius Nitsche and Philipp Hornburg who share my enthusiasm for science.

Concerning my private life I am also really thankful for the infinite support of my family, especially for the constant faith of my mother.
Finally, I want to thank the most important person in my life. Miriam, you have been the bastion of calm, you have taken away my fears and troubles whenever I was struggling. I would not have mastered the past journey without your unconditional love.

[1] Yes, a *Star Trek* quote.

Contents

List of Figures

List of Tables

List of Abbreviations

A.U. Arbitrary Units

AOX Alcohol Oxidase

CAT (Peroxisomal) Catalase

CNS Classical Nucleation Theory

DESY Deutsches Elektronen-Synchrotron

DHAS Dihydroxyacetone Synthase

DLS Dynamic Light Scattering

DTT Dithiothreitol

EGFP Enhanced Green Fluorescent Protein

EM electron microscopy

Em. Emission Wavelength

eV electron volt

Exc. Excitation Wavelength

FAD Flavin Adenine Dinucleotide

FEL (X-Ray) Free-Electron Laser

FRET Förster Resonance Energy Transfer

GDVN Gas Dynamic Virtual Nozzle

H.p. *Hansenula polymorpha*

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HIP Hansenula Integration Plasmid

KBr Potassium Bromide

kDa kilo Dalton

LCLS Linac Coherent Light Source

mAns Unitless Fluorescence Anisotropy

ME Mercaptoethanol

MW Molecular Weight

MX Macromolecular Crystallography

NaCl Sodium Chloride

NMR nuclear magnetic resonance spectroscopy

OD Optical Density

PCR Polymerase Chain Reaction

PDB Protein Data Bank

PETRA(III) Positron-Electron Tandem Ring Accelerator

PP Potassium Phosphate

PPMCC Pearson Product-Moment Correlation Coefficient

PTS Peroxisomal Targeting Signal

QY Quantum Yield

ROI Region of Interest

S.c. *Saccharomyces cerevisiae*

SACLA SPring-8 Angstrom Compact FEL

SDS-PAGE Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis

SEC Size-Exclusion Chromatography

SED Stokes-Einstein-Debye

SFX Serial Femtosecond X-ray Crystallography

SKL Serine-Lysine-Leucine

SLiCE Seamless Ligation Cloning Extract

SLS Static Light Scattering

ssDNA Single-Stranded Deoxyribonucleic Acid

TCEP tris(2-carboxyethyl)phosphine

TRIS 2-Amino-2-(hydroxymethyl)propane-1,3-diol

UV/vis Ultraviolet & Visible Light Spectrum

VPF Volume Packing Fraction

YP Yeast Peptone

I | Motivation

Recent advances in the development of hard X-ray radiation sources allow to obtain diffraction data on ultra-fast timescales. Pulses on the femtosecond scale facilitate the acquisition of data which is basically free from radiation damage. The access to such hard X-ray free-electron laser (FEL) facilities is extremely limited and is only granted in competition for the highest probability of experimental success. Although FELs are used for a wide variety of physical research, biological targets remain a key focus of novel approaches to push the understanding of molecular machines. Hence, current scientific progress still requires high-quality crystals of macromolecules to ensure that X-ray diffraction extends to sufficiently large angles. Unfortunately, such crystals can usually only be obtained by tedious optimization procedures.

To ensure future advancements in obtaining structural information of macromolecules at atomic resolution, the outstanding properties of next generation X-ray sources are combined with the new approach of serial femtosecond crystallography (SFX). The approach allows to collect as many scattered photons as possible from a single crystal without considering introduced radiation damage. Since any sample in the FEL X-ray pulse focus will also be immediately destroyed on the femtosecond scale, serial diffractive snapshots of thousands of randomly oriented crystals are required to obtain a single macromolecular structure.

To counteract on the demand for homogenous protein nanocrystals, this thesis strives to present a highly interdisciplinary and novel approach of yeast *in vivo* crystallization. The budding yeast species *H. polymorpha* & *S. cerevisiae* are explored as universal carriers and production facilities (*crystal factories*) of nanocrystalline protein material. Due to their proven potency in heterologous protein expression, concentration-induced nucleation in specific cell compartments may be achieved. This would yield endless amounts of highly similar crystals. Yeast cells also provide chemical and mechanical stability. Thus, they are ideally suited to withstand the typically harsh conditions which go hand in hand with efforts in the reduction of background scattering (e.g. vacuum conditions) and sample delivery procedures (e.g. GDVNs). It is the main focus of this thesis to investigate yeast cell properties with regards to crystal yield and as an active crystal carrier, not only by diffraction experiments, but also by means of other biophysical methods.

Hard X-ray FELs were also built with the aspiration to enable structural characterization of macromolecules without the need for crystallization. Although current FEL peak brilliances and experimental procedures might not satisfy all necessary requirements, scientists try to converge on this distant goal by reducing the complexity of diffracting objects. Diffraction data is measured from crystals on the nanoscale, but is extended to macromolecular fiber structures and already large single particles. In a second part, this thesis will present an approach to produce non-crystalline protein fiber structures as possible targets in X-ray diffraction at FELs.

II | Introduction

1 | Modern X-ray Crystallography

1.1 A Brief History

In retrospect, it has been a truly adventurous endeavor for humankind to arrive at the advanced research that is performed nowadays in terms of structural investigation at atomic resolution. Not only did humankind come a long way since the discovery of X-ray radiation by W. C. Röntgen in 1895 [3] to shine a light on the very basic structure of the most important molecules of life, proteins and nucleic acid. It also managed to constantly realign its focus on the progress of the development of new methods in X-ray crystallography and competing techniques. This led to the amazing 112'000+ protein structures deposited in the protein data bank (PDB) by September, 2017 [271], not counting structures from nuclear magnetic resonance spectroscopy (NMR) or electron microscopy (EM).

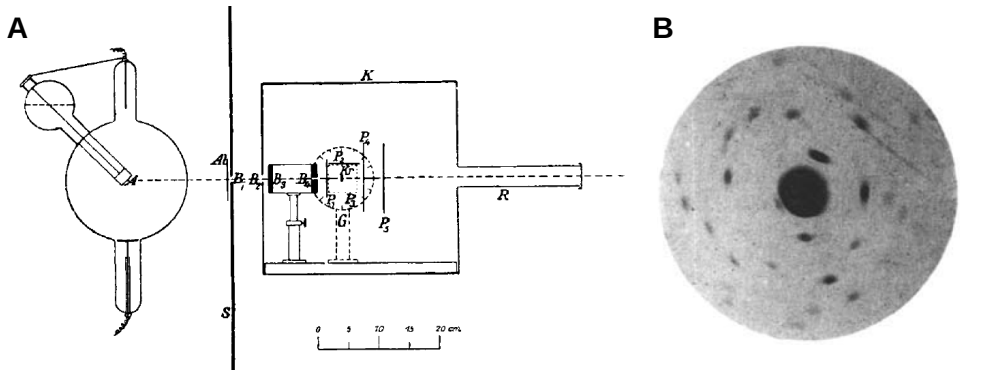

Figure 1.1: *First demonstration of X-ray crystallography* ***(A)*** *Experimental setup with X-ray bulb, goniometer and photographic film* ***(B)*** *First diffraction recorded after narrowing beam slit size; Image taken from: [5]*

At the time Röntgen discovered X-rays, it was still up to debate to define the true nature of light and arguments were raised for a particle (corpuscular) or a wave-like nature. Röntgen's findings about this new type of radiation stimulated further research into the topic. In fact, diffraction-like behaviour for X-rays was already postulated, but not convincingly proven in 1899 by H. Haga and C. H. Wind in Groningen. Their research is connected to the theoretical work of A. Sommerfeld who developed a mathematical theory of the diffraction of light, which he extended to the case of X-rays in 1901 [183]. W. Wien and J. Stark contributed to the scientific discussion, as they individually published results on the possible wavelength of X-rays in 1907 [4] by extending Plancks radiation theory to X-rays. Both estimated the wavelength to be about 0.6 – 0.7 Å. Following many attempts to further proof the wave nature of X-rays, M. von Laue postulated diffraction of X-rays from crystals in 1912. Together with P. Knipping and W. Friedrich, von Laue published the results

in 1913 [5]. They used an X-ray bulb emitting a white beam and recorded the diffraction from $CuSO_4 \cdot 5H_2O$- crystals on a photographic film, which was the first demonstration of X-ray crystallography (see Figure 1.1). At that time, the experiment was recognized as an important step forward in modern physics, and one may quote A. Einstein who sent a postcard to M. Laue in 1912 saying, "This experiment belongs to the most beautiful moments in physics." (Postcard to M. Laue, 10.06.1912, Archives of the Deutsches Museum, Munich, Germany).

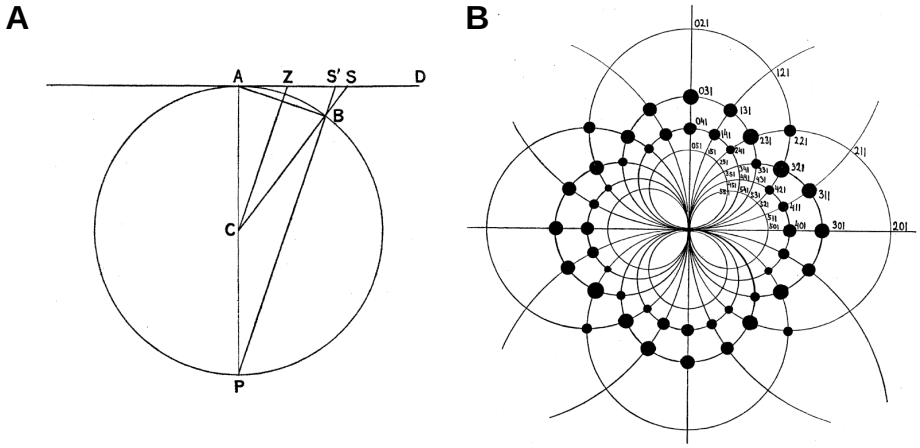

Figure 1.2: *Geometrical explanation of X-ray reflection spots via lattice planes in a stereographic projection* ***(A)*** *Geometrical derivation for diffraction from a lattice plane* ***(B)*** *Projected spots of a potassium chloride crystal with assigned* LAUE *indices hkl; Image taken from: [4]*

Furthermore, this experiment strongly supported the idea of X-rays being of wave-like nature, and also verified the space-lattice hypothesis of classical crystallography, established in 1850 by A. Bravais [1] which is homothetic to the reciprocal lattice introduced by J. W. Gibbs in 1881 [2]. Interestingly, Friedrich et al. used copper-containing crystals in their first demonstration due to their expectancy to rather see an effect related to fluorescence than a constructive interference originating from a crystal-based wavelength selection process [183]. Shortly after that, W. L. Bragg and his father W. H. Bragg, at that time both convinced of a corpuscular nature of X-rays, began to extend Laue's findings. It was the son W. L. Bragg who changed his attitude about the nature of X-rays in late 1912, when he explained Laue's experiment by the reflection of electromagnetic waves in a set of lattice planes [4]. In contrast to that, his father insisted that in order to have a full understanding of the X-ray's nature, one need to "find one theory which possesses the capacities of both" particle- and wave-theory. It was not before the discovery of the Compton effect and L. de Broglie's famous dualistic hypothesis in 1924 that the true nature of electromagnetic waves was understood and validity could be conceded to W. H. Bragg's assumption [183]. The discovery of electron- and neutron diffraction followed after extensive research in the field of X-ray diffraction in the years 1927 and 1936, respectively [183]. The discovery of X-ray diffraction provided humankind with an almost ubiquitous applicable tool to study the properties of matter at atomic resolution. Its impact can be felt in various aspects of modern world society: in material science, especially with computer- and nanotechnology, or in biotechnology with the development of new drugs. A good review about the technological developments in the

twentieth century that are related to protein crystallography is given by Su *et al.* [245].

1.2 Advanced X-ray Sources

Nowadays, groundbreaking experiments in hard X-ray diffraction require radiation sources with extreme peak brilliance. Modern synchrotrons like PETRAIII (Germany) or SPring-8 (Japan) can reach a peak brilliance of up to $10^{25 ph}/_{s \cdot mrad^2 \cdot mm^2}$ [158], while state-of-the-art free-electron lasers (FELs) even produce beams with a brilliance of $10^{33 ph}/_{s \cdot mrad^2 \cdot mm^2}$. To date, three major hard X-ray FELs are operational in biophysical research, the Linac Coherent Light Source (LCLS, USA) [151, 233], the SPring-8 Angstrom Compact FEL (SACLA, Japan) [175] and as of recently the European XFEL (Germany) [260]. Other free-electron lasers are operated in the soft X-ray regime [256].
Both synchrotron and FEL radiation are generated from accelerated electron bunches, either on a circular or linear trajectory [79]. Initially, free electrons are yielded by thermal emission and are accelerated to relativistic velocities, forced into a trajectory by magnetic fields. In synchrotrons, tangentially installed 'insertion devices' (bending magnets, undulators, wigglers), which supply strong magnetic fields, either deflect or manipulate the electron bunches (see fig. 1.3). They give rise to high energy *bremsstrahlung* whose energy is determined by the velocity of the charged particles. Usually, the deflection of X-rays by a bending magnets yields a broader bandwidth ($\sim 1.0\%$) and is thus of interest in Laue diffraction. Wiggler and undulator devices are rather based on the creation of harmonic oscillations of the electron bunches via sinusoidal magnetic fields. While wigglers give rise to a similarly broad bandwidth at higher photon energies, undulators will yield an initially narrower bandwidth ($\sim 0.1\%$) by utilizing weaker magnetic field strengths ($\sim 1T$) that allow for the interference of radiation cones. Still, both devices are based on the alternating arrangement

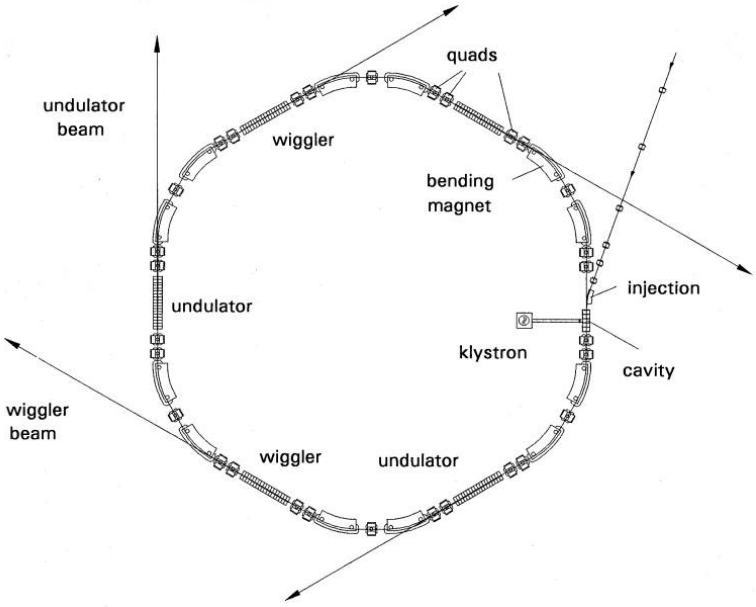

Figure 1.3: *General layout of a third-generation synchrotron facility utilizing bending magnets, wigglers and undulators for electron deflection and X-ray generation; Image from [51]*

of magnets on the cm scale. LORENTZ contraction and DOPPLER effect explain the yield of Å-wavelengths. Typically, modern synchrotrons are operated at electron energies of 1.5-7 GeV [51]. Although higher energies of up to 100 GeV are possible, physical limitations restrict economic operations to lower energies. This is due to the energy loss of electrons at relativistic speeds, which increases with the 4th-power of the particle's kinetic energy [133]. At the end, the radiation created in synchrotrons is transformed and deflected by X-ray optics to supply beamlines with a mono- or polychromatic beam at atomic wavelengths. Synchrotrons have been extremely useful facilities throughout the last decades but suffer from severe limitations in X-ray pulse lengths [89]. Modern X-ray free-electron lasers allow to advance research into the nm-fs space by creating highly-collimated ($\sim \mu rad$) and ultra-brilliant X-ray beams. Photons are squeezed into ultra-short pulses at the femtosecond scale ($10^{-15}s$) and allow for the observation of chemical processes and atomic motions in matter. Starting again with an electron source and acceleration process, electrons adopt longitudinal coherence by *FEL collective instability* in a long, linear undulator section [257], reaching energies of 15-20 GeV. The instability transforms randomly distributed electrons to a collective '1D-electron crystal' with a spacing at the order of X-ray wavelengths. Unlike spontaneous undulator radiation in a synchrotron, the electrons, and by that the EM waves, superimpose constructively yielding an intensity proportional to (nb. of electrons)^2. The instability mainly works by (I) an interaction of random electrons with the EM waves emitted by other electrons, to lead to changes in electron energy at the period of λ followed by (II) a stimulation of a λ-dependent sorting process with regards to the electrons by external magnetic fields (microbunching) and finally (III) an increased coherence that again stimulates constructive interference of EM waves yielding a larger amplitude. The rate at which the amplitude growth is called the 'gain length'. The whole process saturates when all electrons are well-ordered. If the instability simply starts from the random synchrotron radiation at the beginning of the undulator, the FEL is also denoted as SASE-FEL (self-amplified spontaneous emission free-electron laser). SASE-FEL radiation is usually randomly spiked with regards to intensity and wavelength distribution. Fortunately, the randomness is reduced in the exponential gain regime of the undulator, to finally yield FEL radiation with reduced pulse variations [91] (cp. fig. 1.4). The quality and gain of X-ray radiation also mainly depend on the quality of the initial electron beam [77].

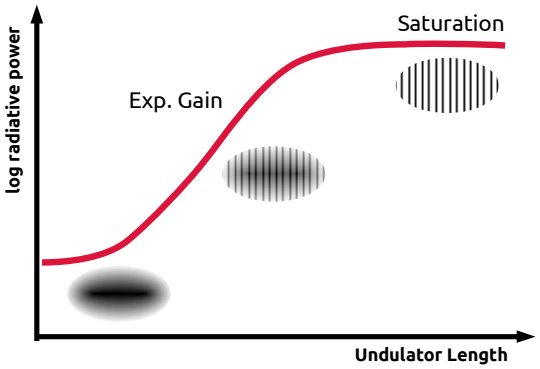

Figure 1.4: *Schematic representation of SASE microbunching of electrons and gain in radiative power via increasing coherence*

1.3 Theory of Protein Crystal Diffraction

Crystals of biological macromolecules are difficult to obtain, as most such molecules only tend to congregate at rare conditions. Still, large *in vitro* grown protein crystals are the backbone of today's approaches in structural investigation of biomolecules via X-rays. Therefore, a clear understanding of the basic definitions with regards to crystalline material is obligatory. Typically, a crystal is defined as a solid material of homogeneous nature which is made up of repeating units whose order extends repetitively in several dimensions. This 'unit cell' repeats only through translation, but might contain further symmetry elements. The lattice parameters consisting of the basis vectors a, b, c and their interaxial angles α, β, γ describe the general structure of the unit cell, so that all linear combinations of the basis vectors are lattice vectors. The smallest volume unit, which contains all of the structural information, is called 'asymmetric unit'. It can reproduce the unit cell by the application of certain symmetry operations.

Crystalline Order

Likewise to simple molecules, proteins can adapt various but limited orders within seven basic crystal (lattice) systems (cp. table 1.1). The translational symmetry is described by the 14 possible Bravais lattices [1]. They can be described with regards to their centering component, which can be either primitive (P), face (F), body (I), base (A,B,C) or rhombohedral (R) (cp. fig. 1.5). Additional symmetry is introduced via point symmetries (rotation, reflection, inversion). They add 32 possible crystallographic point groups, according to the *crystallographic restriction theorem*. Still, the Friedel law describes that Bragg reflections from a crystal always appear centrosymmetric. Hence, only the 11 Laue groups can be directly recognized in diffraction data.

A combination of all possible symmetry operations under consideration of

Table 1.1: *The seven crystal systems and their geometrical restrictions, possible bravais lattices*

Lattice Systems	**Axial Vector Lengths**	**Axial Angles**	**Bravais Lattice Types**	**Lattice Symbols**	**Laue Groups**
Cubic	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$	Primitive	P	$m\bar{3}$
			Body-centered	I	$m\bar{3}m$
			Face-centered	F	
Tetragonal	$a = b \neq c$	$\alpha = \beta = \gamma = 90^\circ$	Primitive	P	$4/m$
			Body-centered	I	$4/mmm$
Orthorhombic	$a \neq b \neq c$	$\alpha = \beta = \gamma = 90^\circ$	Primitive	P	mmm
			Body-centered	I	
			Base-centered	C	
			Face-centered	F	
Rhombohedral	$a = b = c$	$\alpha = \beta = \gamma \neq 90^\circ$	Primitive	R	$\bar{3}$
					$\bar{3}m$
Hexagonal	$a = b \neq c$	$\alpha = \beta = 90^\circ$ $\gamma = 120^\circ$	Primitive	P	$6/m$
					$6/mmm$
Monoclinic	$a \neq b \neq c$	$\alpha \neq \gamma = 90^\circ \neq \beta$	Base-centered	C	$2/m$
Triclinic	$a \neq b \neq c$	$\alpha \neq \beta \neq \gamma \neq 90^\circ$	Primitive	P	$\bar{1}$

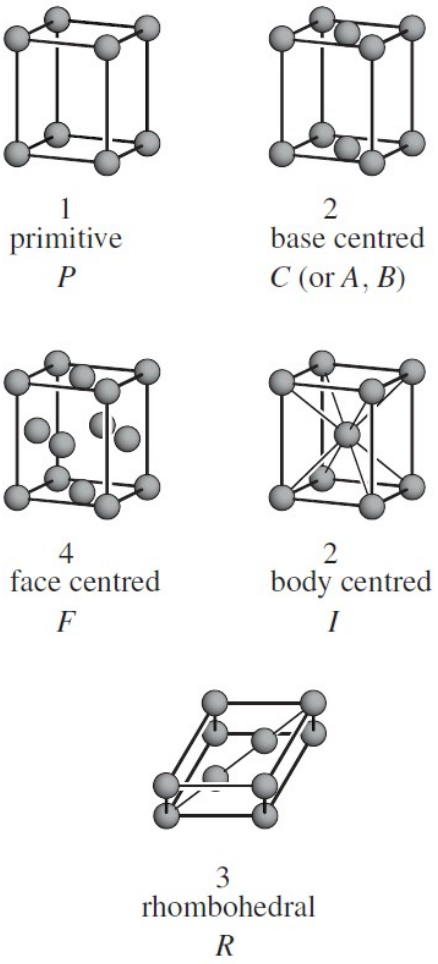

Figure 1.5: *Unit cells of centered bases and their symbols, The numbers specify by which factor the unit cell is enlarged relative to the corresponding primitive cell; Image from [190]*

screw axes and glide planes yield the 230 crystallographic space groups, firstly described by Fedorov and Schoenflies in 1891. Restrictions due to molecule chirality result in 65 possible enantiomorphic space groups for biomolecules. A detailed description of crystallographic relationships is presented by Müller [190]. The description of space groups follows different standards, but the HERMANN-MAUGUIN symbols intend to generate the most simplistic and clear description. They have been defined as an international standard in the *international tables for crystallography A* since 1935 [249] and are also used throughout this thesis.

X-ray Scattering by Atoms & Crystals

This paragraph briefly describes the fundamentals of X-ray scattering and mostly follows the descriptions given by Rupp [159]. The interaction of atoms in crystalline biomolecules with hard X-rays is described by elastic THOMSON scattering. Thereby, photons interact with inner shell electrons in such a way that wavelength and energy are conserved. The scattered wave from a single electron, has an intensity according to:

$$I_{2\theta} = I_0 \frac{1}{2} \left[\frac{e^2}{4\pi\epsilon_0 c^2 m_e} \right]^2 (1 + cos^2 2\theta) \tag{1.1}$$

In this equation I_0 is the intensity of the incident wave and $I_{2\theta}$ the scattered intensity at a given angle. e, and m_e are the charge and mass of an electron, respectively. ϵ_0 describes the permitivity of the vacuum, whereas c is the vacuum speed of light. The last bracket describes the cosine-dependent polarization factor. Obviously, scattering is strongest in forward and backward direction. The scattering amplitude from a single atom scales with its atomic number Z and thus depends on the electron density $\rho(r)$ (r is a position vector). The general distribution of electrons is assumed to be spherical in first approximation. The phases of all scattered waves from one atom are completely coincident in the forward direction but show an increasing dissonance with a larger scattering angle θ due to the atomic electron distribution. The scattering from a single atom in spherical approximation is described in terms of a wavelength-independent atomic scattering (form) factor:

$$f^0(S) = \int^{V(Atom)} \rho(r) e^{2\pi i S \cdot r} dr \tag{1.2}$$

S (sometimes also denoted as: $Q = 2\pi S$) is the momentum transfer or scattering vector and is also a vector in 3D reciprocal space to describe the change of momentum between incident and scattered electromagnetic wave. The atomic scattering vector falls rapidly off with increasing scattering angle and appears centrosymmetric due to the approximation of the electron cloud. It is related to the unit vectors of incident (k_0) and scattered wave (k_1) and the X-ray wavelength by:

$$S = \frac{1}{\lambda}(k_1 - k_0) \ , \quad S = \frac{2sin\theta}{\lambda} \tag{1.3}$$

As a conclusive generalization, the total diffraction from a molecule is then described by a summation:

$$F_M(S) = \sum_{j=1}^{Atoms} f_j^0 \cdot exp\left(2\pi i S r_j\right) \tag{1.4}$$

Here, r_j is the position vector of an individual atom in the molecule. F_M is called the molecular diffraction envelope or scattering function and describes scattering due to the superposition of waves from all atoms in the molecule. Since in such a case atoms are usually ordered in an irregular pattern, the scattering function is a continuous, decaying complex function with irregularly structured modulations. The phase difference between scattered waves can be described with regards to the separation vector r between two objects:

$$\Delta\phi = 2\pi i S \cdot r \tag{1.5}$$

On the contrary, the periodicity of a crystal lattice, in which each atom/molecule has exact translational (in-phase) equivalents in 3D space, leads to a linear intensity amplification that scales with the number of unit cells: $F^C = n \cdot F^{UC}$ (UC= unit cell). The amplification process is simply due to constructive interference (cp. eq. 1.5), thus it strongly depends on the accuracy of atomic positions in a repeating lattice. The accurate repetition of a lattice is usually disturbed by vibrations of atoms and molecules, which depends on their kinetic energy. Additionally, crystal quality can suffer due to general disorder. Both

effects are commonly described in terms of the wavelength-dependent DEBYE-WALLER factor (T_S), which also includes the general isotropic displacement factor (or B-factor). Hence, the total scattering of regularly placed atoms in a crystal lattice is then described by:

$$f_B^0 = f^0 \cdot T_S = f^0 \cdot exp\left(-B\left(\frac{sin\theta}{\lambda}\right)^2\right) \tag{1.6}$$

The conditions for constructive interference can also be expressed in terms of the LAUE equations, which summarizes the statement with regards to the unit cell basis vectors a,b,c and the integers $n_1 - n_3$:

$$S \cdot a = n_1 \ , \quad S \cdot b = n_2 \ , \quad S \cdot c = n_3 \tag{1.7}$$

These equations fully explain the occurrence of diffraction maxima in a diffraction pattern. It is possible to describe these equations in terms of the LAUE indices (h k l) of the lattice planes of the unit cell, which are also described in detail by Rupp [159]. The LAUE indices are just the n-th order indices of the MILLERIAN indices. Confusingly, these terms are sometimes used as equivalent. Considering the definition of the real space lattice vector $r_j = A^T x_j = (ax_j + by_j + cz_j)$ one can derive:

$$Sr_j = Sax_j + Sby_j + Scz_j = hx_j + ky_j + lz_j = hx_j \tag{1.8}$$

This definition can be used to derive a direct dependency of the structure factor on the reciprocal index vector h. With the definition from equation 1.8 the structure factor becomes:

$$F_h^C(S) = N \cdot \sum_{j=1}^{Atoms} f_j^0 \cdot exp\left(2\pi i h x_j\right) \tag{1.9}$$

which is in principle equivalent to equation 1.4, except that the summation is now over the atoms of the entire unit cell. It can be seen that the scattering function of the unit cell with regards to the fractional coordinates x_j and the reciprocal lattice indices h is the structure factor F_h^C of the reflection h.
The idea of lattice planes and their corresponding description via hkl indices is best understood by means of the BRAGG equation (Eq. 1.10). It utilizes the idea to normalize the scattering vector S so that diffraction from a crystal can be interpreted as the reflection from a set of perpendicular real space planes, which are separated by the distance d. The basic principle is schematically depicted in figure 1.7 in terms of the scattering angle θ and the lattice spacing d. The following equations give the corresponding mathematical relation with regards to the X-ray wavelength λ and define the conditions for constructive interference for d spacings:

$$n\lambda = 2d_{hkl} \cdot sin\theta \ , \quad d_{hkl}^* = \frac{1}{d_{hkl}} \tag{1.10}$$

The reciprocal lattice vector d_{hkl}^* allows to derive a direct relation between the real space lattice distances d_{hkl} and the scattering vector S. This relation is

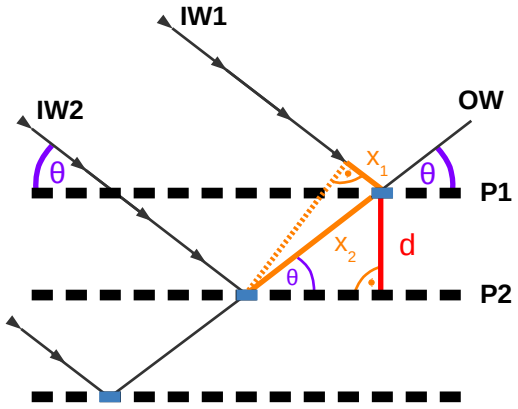

Figure 1.6: *Schematic depiction of diffraction interpreted as originating from crystal lattice planes (P1, P2) at an angle of incidence θ and lattice spacing d, IW: incoming wave, OW: outgoing wave; Usually described in terms of the* Bragg *equation (Eq. 1.10)*

usually described and visualized by means of the Ewald construction. For resolving a protein structure at atomic resolution, not only the positions of possible hkl reflections are of importance, but also the corresponding intensities and phases. The intensity of a given reflection is determined by the actual unit cell content. Therefore, it depends on the actual scattering intensities of the contributing atoms in combination with the conditions for constructive interference.
The structure factor can be subjected to an inverse Fourier transformation to yield the electron density $\rho(r)$. It is obvious from equation 1.9 that both structure factor amplitude and phase information are required for this operation. Using Euler's formula, the complex structure factor can be split to yield separate terms for the structure factor amplitude and the phase:

$$F_h^C = |F_h^C| \cdot exp\,(2\pi i h \alpha) \tag{1.11}$$

where α is the phase associated with the point h in reciprocal space. Thus, applying FT^{-1}, the electron density at point r is described by:

$$\rho(r) = \frac{1}{V} \sum_{h}^{\infty} |F_h^C| \cdot exp\,(i\alpha) \cdot exp\,(-2\pi i h) \tag{1.12}$$

Equation 1.12 is divided by the volume V of the unit cell for scaling purposes and could be readily used to calculate electron densities. Unfortunately, recorded diffraction patterns in X-ray crystallographic experiments only conserve amplitudes but not the phases of the structure factors, a problem commonly referred to as *the phase problem* of X-ray crystallography. For large molecules this problem can be solved via multiple approaches. One is molecular replacement, a completely computational approach where already known phases of a similar molecule are being utilized. Additionally, experimental techniques such as isomorphous replacement (SIR, MIR) or anomalous dispersion (SAD, MAD) can be used to retrieve the phase information to yield a correct electron density map.

Further Interactions of X-ray Radiation and Matter

Elastic scattering, as it was described above, is only one possible type of interaction between X-ray radiation and matter. Furthermore, X-rays can be scattered via inelastic COMPTON scattering, but can also be absorbed (photoelectric effect). The transmission (or attenuation) of an X-ray beam can be calculated according to the LAMBERT-BEER law:

$$T = \frac{I}{I_0} = exp(-\mu R) \tag{1.13}$$

Here, $sfracII_0$ is the ratio of incident and outgoing intensity, R is the penetration depth and μ is the attenuation (absorption) coefficient defined as:

$$\mu = \left(\frac{\rho N_A}{M}\right)\sigma \tag{1.14}$$

in which σ represents the absorption cross-section of the material, ρ is the mass density and M the molar mass of the investigated material. N_A is the AVOGADRO constant. The attenuation of X-rays approximately scales with the 4-th power of the atomic number Z.

1.4 Serial Femtosecond X-ray Crystallography (SFX)

Over the past decade new developments in the field of ultra-brilliant light sources (cp. sec. 1.2), fast-readout detectors and sample delivery techniques have stimulated new approaches in protein crystallography at synchrotrons and XFELs. Even though classical macromolecular crystallography (MX) with macroscopic crystals is responsible for tremendous progress in biological research the fundamental limitations in technical feasibility and sample preparation/deterioration largely prevent further groundbreaking research [268].
An alternative approach is serial crystallography (SX), which can circumvent radiation damage and difficulties in growth of large, well-ordered crystals. Instead of simply rotating a single large crystal, SX is based on acquiring consecutive diffraction patterns from thousands of single sub-micron sized and randomly-oriented crystals. The utilization of intense, femtosecond (20-50 fs) X-ray pulses at XFELs allows to apply ultra-high doses to single crystals in comparison to classical MX, whilst still avoiding signal deterioration from radiation damage (*diffraction-before-destruction* principle). Unfortunately, only three hard X-ray FELs are operational to date, which limits access for a broader scientific community. Furthermore, multidisciplinary teams are required to provide experimental chambers, suitable sample delivery methods, the samples themselves and a fast pipeline for data evaluation. Due to the high demand for serial data collection, the principle has been recently also transferred to synchrotron facilities as serial millisecond crystallography (SMX) or in a broader sense by taking serial diffraction snapshots along a larger crystal [207, 268, 269]. Even though millisecond exposures cannot circumvent the acquisition of diffraction data influenced by molecular damage from liberated electrons, they can outrun secondary radiation damage from molecule relaxation or diffusing radicals. Thus, all these approaches share the idea of distributing the radiation dose to allow for a more unbiased data acquisition. More recently, also serial LAUE diffraction has been performed at synchrotrons as shown by Perry *et al.* [215] and Meents *et al.* [264]. SFX with nano- or micro-crystals is also advantageous in terms of time-resolved measurements at FELs. Only

very small crystals enable a fast and complete diffusion of reactants within the crystal during mixing experiments or for a homogeneous reaction initiation in pump-probe experiments [232, 243].
The techniques and methods that are required to facilitate serial crystallography are under constant development. New studies provide the scientific community with new techniques for sample delivery, and to date, several approaches exist for this purpose. Injection via micro-sized nozzles is certainly the most versatile method, comprising standard gas dynamic virtual nozzles (GDVN) [129], grease jet [218] and related lipidic cubic phase injectors [221]. The principle of GDVNs has also been extended to nozzle systems that rely on double flow-focused jets and thus improve jet stability and sample consumption [265]. Newer approaches range from fixed-target delivery system [240, 267] to microfluidic devices [209]. Likewise, an infrastructure for data analysis has been provided in terms of the software *Cheetah* [204] and *CrystFEL*. Over the past years, also novel biological samples have been investigated by SFX, also comprising membrane proteins [163, 188, 212, 237] and *in vivo*-grown or *in cellulo*-situated crystals [194, 207, 217, 253].
Serial data collection approaches at FELs aim not only to utilize nanocrystalline samples but might open the route to single-particle imaging of biomolecules. So far, researchers have already made larger, non-crystalline samples like virus particles amenable for structural characterization [170, 267]. Still, for now most progress seems to be achieved with regards to method development and structural determination of crystalline protein samples. Especially membrane proteins or proteins that offer new insights with regards to reaction mechanisms are of interest, as they can fully exploit the advantages of extreme time resolution and beam brightness. Interesting reviews about serial data collection at FELs are presented by Schlichting [243] and Martin-Garcia *et al.* [254].

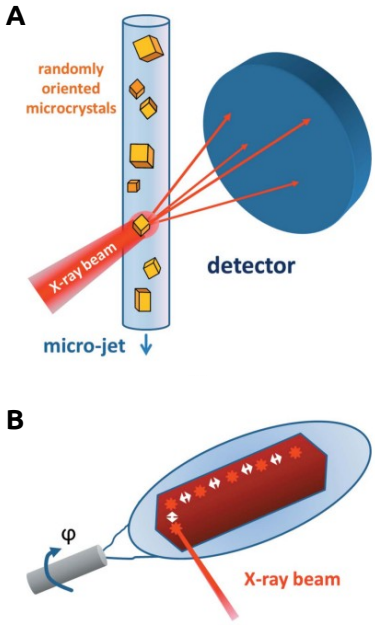

Figure 1.7: *Schematic representation of typical sample delivery methods in SFX* ***(A)*** *Injection in a liquid jet via nozzles* ***(B)*** *Sample delivery in a loop or fixed-target device; Image adapted from [254]*

1.5 Principles of Protein- Nucleation & Crystallization

Obtaining protein crystals for X-ray diffraction experiments remains mostly an empirical task. It is also a highly individual process, as every protein has a different size, surface charge distribution, amount of flexible domains or general stability. Aggregation or nucleation/crystallization is further influenced by different kinds of impurities. Since proteins must often maintain function and solubility at general protein packing fractions as high as 40%, it has been suggested by Doye *et al.* [94] to assume an evolutionary negative design pressure on proteins with regards to crystallizability. It might depend on the composition of general surface residues, the cellular environment, the protein's specific concentration and the task of the protein.
The crystallization process of proteins is characterized by two major processes, nucleation and crystal growth. While nucleation is the rate determining step and requires usually high concentrations of protein or precipitant, growth generally follows at lower concentrations (cp. fig. 1.8). But this static view is too often an oversimplified approximation of reality. The success of crystallization dramatically depends on the exact physical conditions of a solution. This influences strength and range of intermolecular interactions.
A general approach to describe protein nucleation is the utilization of the *classical nucleation theory* (CNS). Its most basic equation states that the rate (per unit volume) of homogenous nucleation follows [122, 179]:

$$k_N = \rho Z j \cdot exp\left(\frac{-\Delta G^*}{k_B T}\right) \tag{1.15}$$

Within the exponential factor, ΔG^* is the free energy cost of creating a critical nucleus and $k_B T$ is expressing the thermal energy of the system. The factor indicates a rather low probability for critical nucleus formation for most systems as it usually means to overcome a high energy barrier (activation energy). The prefactor is constituted of the number density of particles ρ, the rate of nucleus growth j and the ZELDOVICH factor, which determines whether a critical nucleus readily grows into a crystal. Both the rate of nucleation and of crystal growth can be rate- or diffusion-limited [122]. As mentioned above,

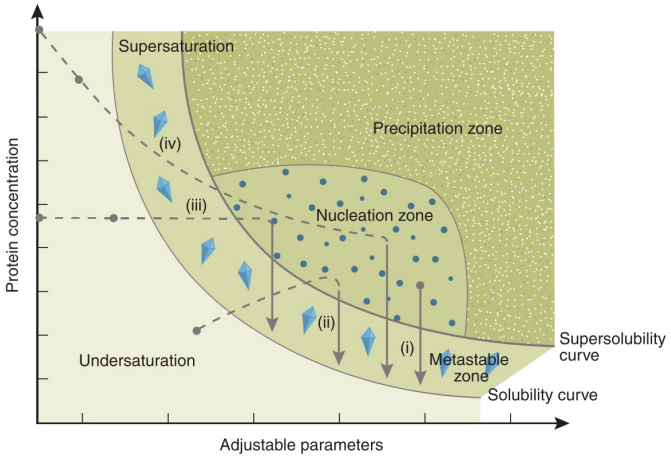

Figure 1.8: *Schematic representation of the protein crystallization phase diagram; Image from [127]*

impurities like dust particles, surfaces, nucleants and oligomerization processes can affect nucleation by providing an altered energy landscape. This can ultimately lead to heterogenous nucleation processes, in which equation 1.15 is also applicable in case all impurities have equal energy barriers. Researchers have also introduced additional techniques to achieve the creation of artifical impurities and stimulate nucleation at lower supersaturation. A good example is given by Yoshikawa *et al.* [224] with the usage of femtosecond laser pulses in the UV/vis range.
A rich body of experimental and theoretical research exists to elucidate the complex energy landscape of protein nucleation and crystallization in solution. In the pioneering work from George & Wilson [42] it is demonstrated via static light-scattering (SLS) that the success of *in vitro* protein crystallization correlates with a slightly negative second virial coefficient (B_2) according to:

$$\Pi = k_B T \left(\frac{n}{V}\right) \left[B_1 + B_2 \left(\frac{n}{V}\right) + B_3 \left(\frac{n}{V}\right)^2 + ...\right] \tag{1.16}$$

Equation 1.16 is an expansion of the ideal gas law, in which Π is the osmotic pressure (Pa), $^n/_V$ the protein concentration, k_B the BOLTZMANN constant and T the temperature of a solution. This correlation has been the first found rational predictor for macromolecular crystallization [74, 127].
The second virial coefficient provides a direct measure of the protein-protein pair potential and should be in the rather narrow range of -2 to $-8{\cdot}10^{-4} mL{\cdot}mol{\cdot}g^{-2}$ for successful crystallization. Positive values of B_2 represent mainly repulsive forces between molecules and are similarly to highly negative values not correlated with crystallization success. The latter ones typically yield protein aggregates. The range of attraction between proteins has a drastic effect on the phase diagram (cp. fig. 1.8). It is typcially situated in the short-range regime for proteins and mostly limited to V.D.WAALS interactions [56, 64]. Therefore, the influence of pH, ionic strength and non-adsorbing polymers can be crucial for crystallization. Still, temperature and particle density have a major influence on molecule behaviour, as indicated in the above equations and also experimentally shown by Astier & Veesler [124].
The mentioned range of B_2 values corresponds to a metastable liquid-liquid critical point in the vicinity of the liquid-solid equilibrium, which does not follow CNT predictions. In practical terms this means a coverage of the protein nucleus by a thin film of highly concentrated protein [69]. This situation resembles the amorphous phase formed by higher protein concentrations in which density fluctuations are enhanced. This will finally cause a lowering of the free energy barrier for the formation of critical size nuclei and it will also increase the rate of crystallization by several orders of magnitude [56, 74]. The whole process can basically be perceived as a two-step mechanism for nucleation. Usually, higher protein concentrations will not only show higher rates of nucleation but also cause the formation of protein aggregates. Furthermore, supersaturated solutions show an increased thermodynamic tendency for the crystallization process itself, yielding rapidly growing and thereby imperfect crystals. Hence, crystallization at the liquid-liquid critical point can yield crystals with less imperfections as only the nucleation rate, but not the crystallization rate is increased [56]. Experimental evidence of these theoretical considerations has been given for example by Galkin & Vekilov [68]. Their work demonstrates how non-adsorbing polymers like PEG molecules can specifically affect lysozyme nucleation rates at the liquid-liquid phase boundary. More recently, Sauter

et al. [242] have demonstrated the observation of two-step nucleation of proteins via optical microscopy and SAXS, which corresponds to the presented theory of nucleation at the liquid-liquid critical point. On the contrary, even though computer simulations by Wedekind *et al.* [248] corroborate the assertion of liquid patches in which the free energy barrier is lowered, their experiments also suggest that the two-step model does not apply for all metastable liquid-liquid phase transitions but is complemented by two additional mechanisms: (I) direct crystallization from a low-density phase and (II) crystallization from unstable liquid droplets at critical densities. It is additionally shown that higher nucleation rates are rather associated with the whole metastable region rather than just the metastable critical point itself.
In recent years, miniaturization and automation of crystallization experiments have been applied to facilitate easier high-throughput screening of suitable crystallization conditions. Intriguingly, nanoliter or even picoliter volumes represent a different environment as the same 'mother liquor' in a µL droplet. Experiments suggest that the number of nucleation events is linearly proportional to the droplet volume [80, 116]. Therefore, it takes longer for crystals to appear in smaller volumes. The decreased nucleation rate might be related to the increased surface tension of a small drop according to [80]:

$$\Delta\rho = \frac{\sigma}{r} \tag{1.17}$$

ρ is the pressure inside the droplet, whereas σ and r are the surface tension and radius of the droplet, respectively. Bodenstaff *et al.* [80] also identified different crystal shapes in dependency of the droplet volume, indicating a visible influence of surface tension on the crystallization process. These findings appear somewhat related to the influence of gravity on protein crystallization success and crystal shape/size as found for crystallization trials in the earth orbit [235]. In general, it is a widely accepted fact that nanoliter volumes render a different environment despite equal buffer conditions [127].
In experimental terms *in vitro* crystallization is based on many different techniques like microbatch crystallization, in which a paraffin oil works as protection against early evaporation, or vapor-diffusion, that is based on drop equilibration against a liquor reservoir (sitting-drop, hanging-drop). Other techniques like dialysis, free interface diffusion (FID) or the application of microfluidic devices have also been successfully used in protein crystallization. A decent review of such techniques is for example given by Chayen & Saridakis [127].
Biological systems like cells or organelles are much more complex than a simple drop of mother liquor. Still, they also have to obey general laws of physics, but are simultaneously more difficult to approach. Despite the known mechanisms of active turgor pressure regulation, cells and organelles might encounter further out-of-equilibrium fluctuations with regards to hydrostatic and osmotic pressure [205]. Considering equation 1.16, this might have a tremendous influence on *in vivo* crystallization success. On the other hand, the presence of a mixed solution of macromolecules and solutes might also act as mixture of natural nucleants and enhance nucleation. Unfortunately, there are no studies available to confirm such a claim.

2 | Budding Yeast & The Crystal Factory

2.1 Budding Yeast in Heterologous Gene Expression

Yeast can be defined as unicellular, *ascomycetous* or *basidiomycetous* fungi exhibiting vegetative reproduction by budding or fission. Most budding yeast species exist in different states, more precisely haploidic or diploidic cells, capable of sexual reproduction and sporulation. In a non-scientific sense the word 'yeast' usually solely relates to the budding yeast species *Saccharomyces cerevisiae*, which is not only popular in food production and general biotechnology, but is also one of the most important eukaryotes exploited in scientific research. It usually has its growth optimum at about 30°C and shows neither a strong resilience at low nor at high temperatures. Haploidic cells commonly reach diameters of 2-5 µm [28, 61, 149, 168, 193].
Hansenula polymorpha (also designated as *Ogataea angusta* or *Pichia angusta*) is a distant relative of *S. cerevisiae* in the family of *Saccharomycetaceae*. It shows distinct differences in comparison to its relative with regards to geno- and phenotype [193]. In fact, it belongs to the rare group of methylotrophic yeast and can be additionally characterized by its pronounced thermotolerance with possible growth between 20-50°C. Cells usually grow only to diameters of 1-3 µm. *H. polymorpha* and other closely related yeast like *Pichia pastoris* have been heavily exploited in industry, for example for the production of hepatitis B vaccine, insulin, IFNα-2a, or even biofuels [149, 191]. Indeed, yeasts are a superior eukaryotic system for heterologous gene expression. On the one hand, cultivation and growth are easier and faster, also less-expensive and can be facilitated at large-scale in comparison to mammalian or insect cell lines. On the other hand, yeasts still support many more posttranslational modifications which are not supported by bacterial expression systems. Additionally, most yeasts are also capable of fermentation [52, 61, 214]. Not surprisingly, a lot of scientific studies promote these outstanding properties in heterologous gene expression, either with regards to *S. cerevisiae* [25], *H. polymorpha* [37, 63] or other yeast species [201].
Peroxisomal research is usually closely associated with the issue of high expression yields due to the strong promoters that are involved in the expression of peroxisomal enzymes [222, 239]. Although *S. cerevisiae* is a more inconspicuous yeast, all of the mentioned species allow for an easily achievable induction of peroxisome proliferation by specific growth conditions. While oleic acid is inducing peroxisomal fatty acid β-oxidation enzymes in *S. cerevisiae*, methanol stimulates the expression of peroxisomal enzymes which target on carbon assimilation and energy extraction in *H. polymorpha*. Both processes foster the

development of a multitude of fully developed peroxisomes per cell and result in high yields of metabolic enzymes. Promoters that offer strong expression rates are either P_{FOX3} in *S. cerevisiae*, which is inducing the expression of 3-ketoacyl-CoA-thiolase (Fox3p) under oleate-inducing conditions, or P_{AOX} in *H. polymorpha*, responsible for the expression of alcohol oxidase (MOX, AOX) under methanol-inducing conditions. Still, some *S. cerevisiae* strains are not capable to solely grow on oleic acid. On the contrary, *H. polymorpha* can maintain cellular growth on methanol as the sole carbon- and energy source. Thus, the widespread application of these yeast species and the availability of a simple trigger for peroxisome biogenesis have also rendered both as suitable model organisms specifically in peroxisomal research [60, 196, 208, 222].

2.2 The Peroxisomal System of *S. cerevisiae* & *H. polymorpha*

As mentioned in section 2.1, *S. cerevisiae* and *H. polymorpha* are utilized to study peroxisome biogenesis and disease-related deficiencies of peroxisomal proteins in eukaryotic cells. Both species utilize a well-conserved and comparable protein family, mostly denoted as PEX proteins, to accomplish peroxisomal development. Most underlying processes are also partially comparable to such in mammalian cells. Under non-inducing growth conditions only about one small preperoxisome exists per yeast cell. This is a situation which changes dramatically upon induction of peroxisome proliferation. Two distinct mechanisms have been proposed to contribute to the multiplication of peroxisomes within cells. Firstly, *de novo* formation from the ER at an early stage of induction and secondly, direct fission from preexisting organelles as indicated in figure 2.1. A comprehensive discussion of this topic is provided by Veenhuis & van der Klei [220] or Kim & Hettema [229]. Multiplication of organelles is due to the increased demand for protein synthesis of peroxisomal matrix proteins, mostly specific metabolic enzymes, which are imported in vast amounts into peroxisomes to process metabolites in a confined environment. Upon prolonged incubation such proteins can substantially contribute to the total protein yield from yeast cells. When synthesis capacities are adjusted to counteract on the demand for matrix proteins, it is equally important for cells to provide an efficient import machinery for peroxisomes. For yeast, two distinct transient pores (PTSI & PTSII) are proposed which facilitate protein import. The cytosolic receptors Pex5p and Pex7p recognize specifically either PTSI, or PTSII motifs, respectively. These are short C- or N-terminal peptides sufficient for targeting and translocation. For example, the PTSI motif simply consists of a SKL tripeptide, or one of its many recognized variants [102]. The aforementioned transient pores are constituted of several PEX proteins and assemble within the peroxisomal membrane upon cargo recognition [121, 123]. Although a complete understanding of these pores is not yet in reach and no complete atomic model exists for either of them, it is well established that a minimalistic PTSI or PTSII tag might be sufficient for the peroxisomal import of heterologously-expressed and cytosolically-folded proteins [43, 46, 162].
It has been shown that a deletion of the protein Pex11p yields enlarged peroxisomes at an reduced number of organelles per cell, as it usually facilitates peroxisomal fission [147]. A deletion of the receptor protein Pex5p is similarly detrimental for matrix protein import. Both mutants result in specific phenotypes as indicated in figure 2.1. In the case of an absence of Pex5p, peroxisomal

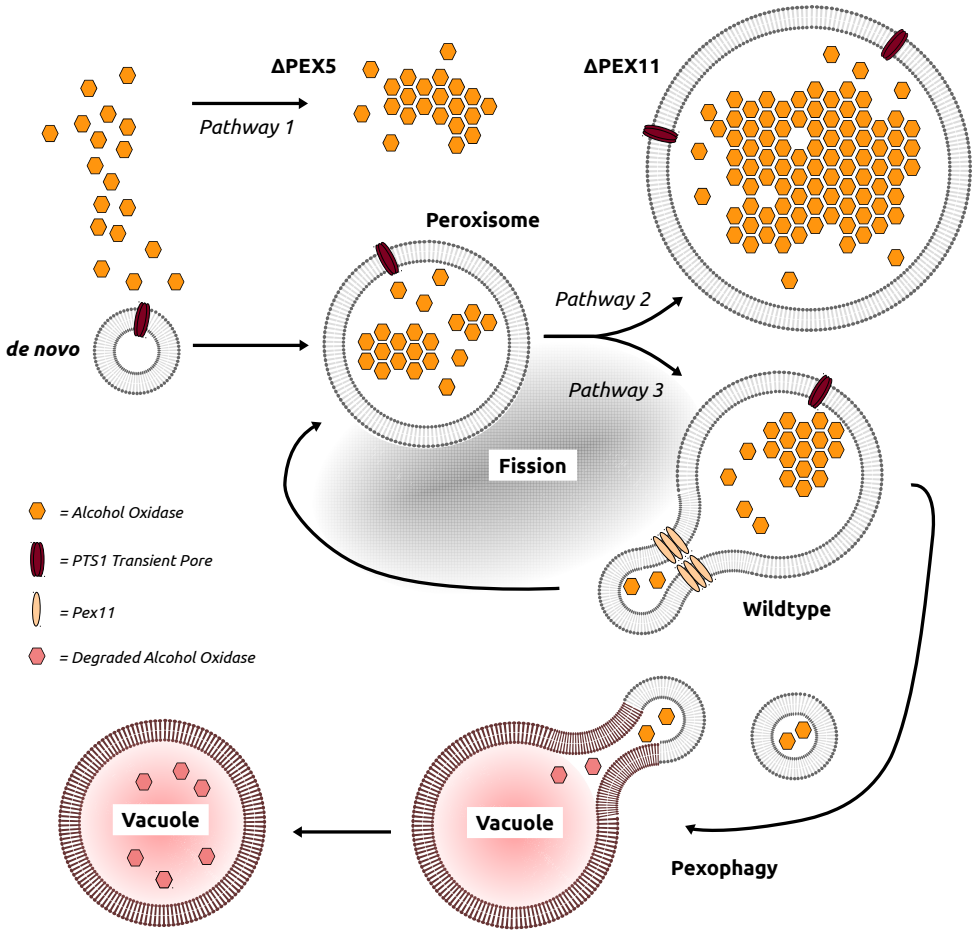

Figure 2.1: *Schematic representation of intracellular peroxisome biogenesis & pexophagy in H. polymorpha wildtype, ΔPEX5 and ΔPEX11 strains under methanol growth conditions*

matrix proteins, that are usually imported via the PTSI pathway, reside in the cytoplasm. Characteristic for *H. polymorpha* peroxisomes and their induction via methanol is the occurrence of tiny solid-state catalyst protein crystals that are constituted of 600 kDa AOX octamers[1] (cp. fig. 2.1). Protein crystals appear as soon as 6 hrs after induction (this thesis) in ΔPEX11, ΔPEX5 as well as wildtype cells and remain stable for prolonged times. It has mainly been suggested that the strong P_{AOX} promoter leads to an increased protein concentration within peroxisomes, finally yielding protein crystals. This is partially supported by findings of Roggenkamp *et al.* [32] who found that the overexpression of native AOX leads to an apparent strengthening of the peroxisomal crystalline matrix. Due to their natural tendency to contain protein crystals *H.p.* peroxisomes earned the soubriquet 'crystal factory'.
This is in contrast to *S. cerevisiae* cells, which do not possess similar enzymes and also do not indicate natural protein crystal formation by other proteins within peroxisomes. Still, peroxisome biogenesis is very comparable to *H. polymorpha*.
Much like cells in a viable culture, the peroxisome population is under the mercy of natural turnover. Old peroxisomes will be recycled by processes called

[1]Monomer MW: 75 kDa; AOX is a alcohol:oxygen oxidoreductase catalyzing the reaction: O_2+prim. alcohol $\longrightarrow$ an aldehyde $+H_2O_2$

macropexophagy and microautophagy. This usually happens under conditions of repression of the related carbon metabolism or nitrogen depletion [139].

2.3 *In Vivo* Protein Crystallization

As reviewed by Doye & Poon [105], *in vivo* crystallization is featured across many different species and also serves a lot of different purposes. Protein crystals may act as active enzymatic entities, as in the case of alcohol oxidase in *H. polymorpha* [24]. In the filamentous fungi *Neurospora crassa*, the protein Hex-1 forms a crystalline matrix in Woronin bodies which ultimately function as plugs for septal pores [72, 93]. These Woronin bodies are also directly derived from peroxisomes. Indeed, peroxisomes seem to frequently host natural protein crystals. It is known that catalase crystals appear in peroxisomes of tobacco leafs or sunflowers [14, 70]. A similar observation has been made for rat liver cells which also feature crystalline catalase within their peroxisomes [13]. Furthermore, the occurrence of urate oxidase and L-alpha hydroxyacid oxidase crystals has been reported for bovine liver and kidney cell peroxisomes [39]. It would seem that peroxisomes are predestined to host crystalline proteins although organellar function can vary dramatically. As opposed to this, natural protein crystals have also been reported to occur in other cellular loci. This is the case for insulin in β-cells of mice and human pancreas [57, 148], proteins in plant seeds [59] and for bacterial toxins [30]. Although the kinetic energy of particles is heavily increased at typical ambient temperatures of cells (∼37°C), *in vivo* crystallization occurs frequently and might be energetically favored due to high protein concentrations or other advantageous conditions.
More recently, researchers have tried to pick up on these possible opportunities for a more general generation of *in vivo* grown protein crystals, as reviewed by Duszenko *et al.* [227]. Mainly fostered by developments in the field of X-ray sources, the demand for nano- and micro-sized protein crystals increased. Intense, ultrashort X-ray pulses from free-electron lasers, as introduced in section 1.2, offer the opportunity to obtain damage-free diffraction data from nano-sized crystals. However, the required approach of serial data collection usually relies on the availability of vast amounts of tiny crystals (cp. subsec. 1.4). Although it is possible to obtain dense crystal solutions of some standard proteins like lysozyme or thaumatin [203, 255] and even more relevant targets [188, 212, 237, 251], production of such is never routine in a general sense. Consequently, *in vivo* systems were set under active investigation to best

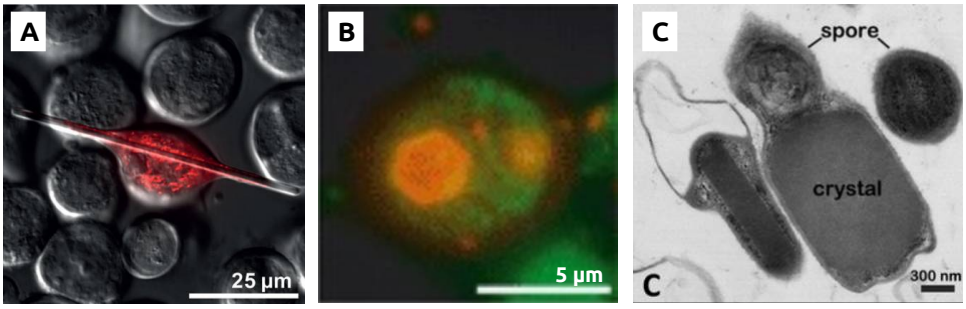

Figure 2.2: *Examples of successful in vivo crystallization upon heterologous expression* ***(A)*** *Firefly luciferase crystals in Sf9 insect cells [244]* ***(B)*** *Hex-1 crystals in S. cerevisiae cells [140]* ***(C)*** *CryA3 toxin crystals in B.t. cells [217]*

accommodate this demand, as they may offer easy, large-scale production and *in cellulo* maintenance of crystals, but also different degrees of direct posttranslational modifications of proteins.
Especially the Sf9/Baculovirus system has received a lot of attention recently, as it has demonstrated its advanced maturation in the generation of *in vivo* protein crystals. Following its application in general protein overexpression and more historic observations of calcineurin protein crystals by Fan *et al.* [49], the system yielded crystals of cathepsin B [194, 207], polyhedrin [202, 228], luciferase and GFP-tagged reovirus µNS [244]. Still, as represented by the last example, high-quality crystals are not necessarily obtained in every case. Intriguingly, crystallization has been achieved in various cellular compartments. Mammalian cell lines have also been exploited as carriers of protein crystals in several cases. For example, the overexpression of human IgG in hamster ovary cells [166] or of the protein KikGR in HEK293 cells [247] lead both to successful protein crystallization. Nonetheless, most of these efforts yielded micro-sized protein crystals which are still amenable to synchrotron-based data collection. It is even more sweeping that in many cases removal of the crystals from the cellular environment is required, not only to reach proper concentrations, but also to allow for an injection via liquid jets in SFX experiments. This removal can of course introduce additional damage to the crystals and thereby reduce the achievable resolution. A partial remedy for this drawback has recently been found in fixed-target delivery systems. On the contrary, Sawaya *et al.* [217] demonstrated that smaller bacterial cells can be directly injected into an X-ray beam.
Yeast cells have not yet been exploited as generalized 'crystal factories', even though they are in a superb starting position (cp. sec. 2.1 & 2.2). They are small enough to be directly used as crystal carriers, and protein crystal formation in peroxisomes has been reported after heterologous expression of a target protein [140].

3 | Fluorescence Anisotropy & Energy Transfer

3.1 Principles of Fluorescence Anisotropy

The understanding of light-matter interactions can be extended from the viewpoint of (elastic) THOMSON scattering as the low-energy limit of (inelastic) COMPTON scattering in X-ray diffraction experiments. In the visible frequency range of electromagnetic radiation both quantized interactions between energy states of an atom or molecule and an electromagnetic wave as well as RAMAN scattering are predominantly of interest in biophysical research. The quantized energy exchange happens via *absorption, spontaneous and stimulated emission*, three basic types of interaction that further depend on the specific properties of the incident wave and the molecule characteristics.
Depending on the electronic state most fluorophores absorb light along a preferred direction which is parallel to their absorption transition moment (absorption dipole). This is in contrast to the emission transition moment which has a fixed orientation due to internal conversion to the first excited singlet-state S_1 (see fig. 3.1-B). When linearly polarized light is targeted onto a set of fluorophores a subset with absorption dipoles parallel to the incident light will get excited with a probability proportional to the dot product $|\overrightarrow{D_A}| \cdot |\vec{E}| cos^2\Theta$. Here $\overrightarrow{D_A}$ and $\vec{E}$ are the absorption transmision moment, and the electric field vector, respectively. Obviously the probability is maximum when $\vec{E}$ is parallel to $\overrightarrow{D_A}$.
A spontaneous emitted photon from a fluorophore will maintain a polarization according to the axis of the emission transition moment. Any changes of this axis during the lifetime of the excited state will cause a deviation from the original polarization direction of $\vec{E}$. The deviations from a molecule's *fundamental anisotropy* r_0 (caused by non-parallel absorption and emission transition moments) by rotational motion or energy transfer to a differently oriented molecule is measured as a dimensionless quantity according to

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \tag{3.1}$$

with $I_{\parallel}$ and $I_{\perp}$ being the fluorescence intensity parallel and perpendicular to the original polarization, respectively [78]. Throughout this thesis anisotropy is denoted with r or the arbitrary unit mAns. There is no phase relation between the photons emitted by previously excited fluorophores in an ensemble measurement. Therefore, the resulting fluorescence intensity can be considered as distributed on three perpendicular axes from which two are summarized in the term $2I_{\perp}$. The anisotropy r is independent of the total fluorescence intensity.

In general it is differentiated between time-resolved and steady-state anisotropy. Time-resolved anisotropy is described as a single-exponential decay with regards to the rotational correlation time (θ) with $r(t) = r_0 exp(-t/\theta)$ and is measured with a pulsed excitation. On the contrary, steady-state anisotropy is measured under continuous illumination and its final value is thus a time average of the anisotropy corrected with the fluorescence intensity [78]:

$$\langle r \rangle = \frac{\int_0^\infty I(t) r(t) dt}{\int_0^\infty I(t) dt} = \frac{1}{\tau} \int_0^\infty r(t) exp(-t/\tau) dt \tag{3.2}$$

With an intensity decay following a single exponential it is possible to derive the PERRIN equation [110]:

$$\frac{r_0}{r} = 1 + t/\theta = 1 + 6 D_r \tau \tag{3.3}$$

with D_r as the rotational diffusion coefficient, defined by the STOKES-EINSTEIN-DEBYE (SED) relation: $D_r = k_B T \cdot (8\pi\eta r^3)^{-1}$ [88, 90].

3.2 Förster Resonance Energy Transfer (FRET)

FÖRSTER resonance energy transfer describes a long-range dipole-dipole interaction with very weak coupling between donor and accceptor molecule. In such a weak interaction the inter-molecular interactions are less pronounced in comparison to the interaction of electronic and nuclear motions within individual molecules. Thus, the conditions for energy transfer are in principle much more stringent than for strong coupling cases, but thermal and solvent effects will effectively broaden the vibronic bandwidths and lead to a more general occurrence of the effect. In FRET, a singlet-singlet energy transfer occurs without the appearance of a photon. As schematically depicted in figure 3.1-A for an energy transfer between identical molecules (homoFRET), the transfer does require a spectral overlap of absorption and emission. This view corresponds to the description of energy states in a molecule before and after excitation (resulting in a STOKES shift) as depicted in figure 3.1-B. As a

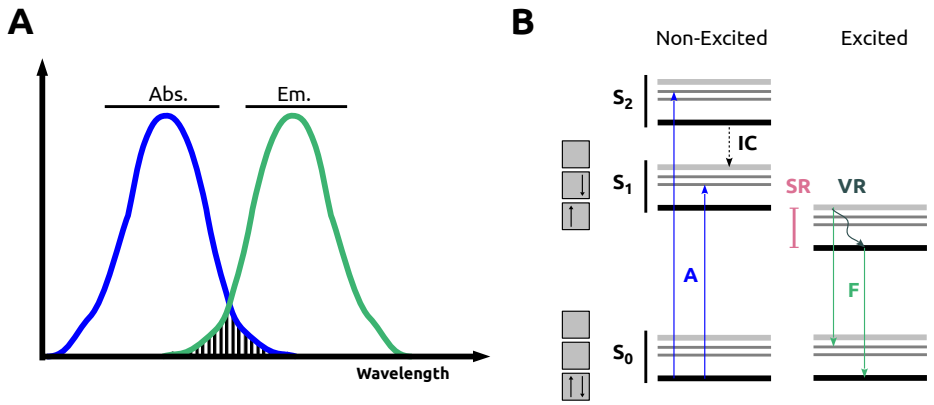

Figure 3.1: ***(A)*** *Schematic absorption and emission spectra of a molecule with their spectral overlap indicated* ***(B)*** JABLONSKI *diagram of a non-excited and excited molecule of the same kind, describing the cause of the* STOKES *shift in terms of solvent reorganization (SR) and vibrational relaxation (VR); IC: Internal conversion between singlet states, A: Photon Absorption, F: Fluorescence Photon Emission*

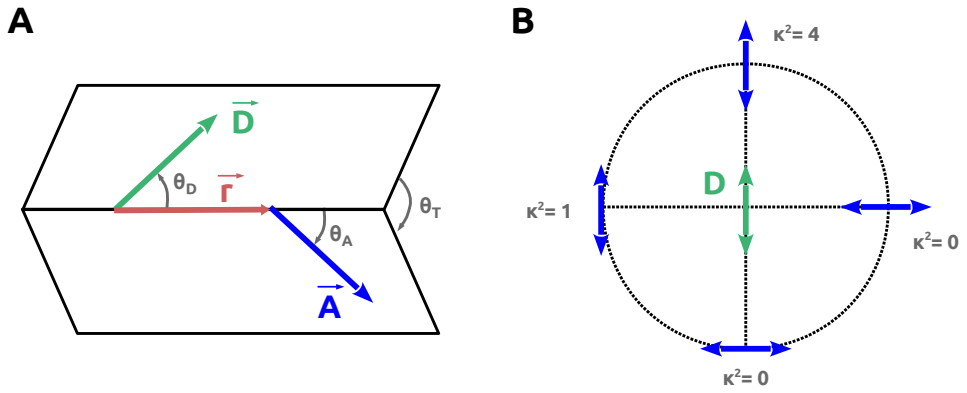

Figure 3.2: ***(A)*** *Relation of donor (D) and acceptor (A) dipoles expressed in terms of the angles θ_D, θ_A, θ_T* ***(B)*** *Absolute values of the orientation factor κ^2 in dependency of the relative donor and acceptor dipole orientation*

molecule absorbs energy and undergoes a transition from a vibronic ground state (S_0) to an excited state (e.g. S_1), the dipole moment of the molecule changes. Adjacent dipoles will only adapt on a slower timescale. This *solvent reorganization* (SR) leads to a an effective lowering of the energy levels. If the sum of such effects still yields overlapping energy states FRET occurs at a rate of:

$$k_T = k_D \left(\frac{R_0}{R}\right)^6 = \frac{1}{\tau_0}\left(\frac{R_0}{R}\right)^6 \tag{3.4}$$

Here k_D is the emission rate constant of the donor and τ_0 is the average lifetime of an excited state in the donor molecule in absence of acceptor. R represents the distance between donor and acceptor. The FÖRSTER radius R_0 gives the critical distance of molecules in FRET at which energy transfer and spontaneous decay are equally probable ($k_T = k_D$) and summarizes all critical terms of the molecular interaction [1]:

$$R_0^6 = \frac{9000(ln10)\kappa^2\Phi_D^0}{129\pi^5 N_A n^4}\int_0^\infty I_D(\lambda)\epsilon_A(\lambda)\lambda^4 d\lambda \tag{3.5}$$

In this equation Φ_D^0 is the fluorescence quantum yield of the donor in the absence of transfer. $I_D(\lambda)$ is the normalized fluorescence spectrum of the donor (with λin nm), $\varepsilon_A(\lambda)$ is the molar absorption coefficient of the acceptor (in $M^{-1}cm^{-1}$), n is the refractive index of the medium at $\Delta\lambda$ and κ^2 is the orientation factor according to:

$$\kappa^2 = (cos\theta_R - 3cos\theta_D cos\theta_A)^2 \tag{3.6}$$

θ_R is the angle between the vectors of donor and acceptor transition moment. θ_D and θ_A describe the angle between the donor or acceptor transition moment and the separation vector $\vec{r}$. The relation of these angles which describe the relative position of donor and acceptor dipole moments is schematically depicted in figure 3.2-A. Depending on this relation the orientation factor κ^2 can take values on the interval [0,4] as shown in figure 3.2-B. A detailed review of the orientation factor and its average value is given by van der Meer *et al.* [198].

[1] $R_0 = 0.211\left(\kappa^2\Phi_D^0 n^{-4}\int_0^\infty I_D(\lambda)\epsilon_A(\lambda)\lambda^4 d\lambda\right)^{1/6}$

HomoFRET in Solution

A solution to the description of FRET involving multiple acceptor molecules in a three-dimensional space has been given by Förster [6][7] in 1948. Following his derivations, the probability of a donor molecule of being in an excited state decreases over time according to:

$$-\frac{dP}{dt} = \frac{1}{\tau_0} + \frac{1}{\tau_0}\sum_{k=1}^{N}\left(\frac{R_0}{R_k}\right)^6 \tag{3.7}$$

Hereby $^1/_{\tau_0}$ describes the natural deactivation process of an excited molecule via radiation or other intramolecular processes with τ_0 as the fluorescence lifetime of the donor. R_0 is the FÖRSTER radius (or FÖRSTER distance) and R_k is the intermolecular distance. Acceptor molecules in the solution will offer additional ways for deactivation of the donor molecule. Those are condensed in the sum term. N denotes the total number of acceptor molecules in the solution. For homoFRET there is a probability of re-excitation of the original donor molecule. Under the assumption of a large number N of available acceptor molecules in a solution, this re-excitation can be assumed to approach zero. Integration of equation 3.7 yields:

$$P(t) = exp\left(-\left[1+\sum_{k=1}^{N}\left(\frac{R_0}{R_k}\right)^6\right]\frac{t}{\tau_0}\right) = exp\left(-\frac{t}{\tau_0}\right)\cdot\prod_{k=1}^{N} exp\left(-\left(\frac{R_0}{R_k}\right)^6\frac{t}{\tau_0}\right) \tag{3.8}$$

Since the exact distance of every acceptor molecule cannot be known, one can estimate the average probability of excitation (of the donor molecule) as follows:

$$\langle P(t)\rangle = exp\left(-\frac{t}{\tau_0}\right)\cdot\prod_{k=1}^{N}\int_0^{R_G} exp\left(-\left(\frac{R_0}{R_k}\right)^6\frac{t}{\tau_0}\right)\omega(R_k)dR_k \tag{3.9}$$

$\omega(R_k)$ describes the probability to find a certain acceptor molecule at the distance R+dR, whereas R_G simply describes the considered volume according to $V_{total} = {}^{4\pi R_G^3}/_3$. Considering a static situation ωwill be 1 at a distinct molecule distance R_k for every single molecule.
If one presupposes a random distribution of acceptor molecules the probability ω(R) to find an acceptor at distance R+dR will be:

$$\omega(R_k)dR_k = \frac{4\pi R_k^2 dR_k}{V_{total}} \tag{3.10}$$

If $R_G \gg R_0$ equation 3.9 can be rearranged to yield:

$$\langle P(t)\rangle = exp\left(-\frac{t}{\tau_0} - \sqrt{\frac{t}{\tau_0}}\cdot\frac{\sqrt{\pi}R_0^3 N}{R_G^3}\right) \tag{3.11}$$

Even though the decay of a single donor still follows an exponential decay, the average decay probability of an ensemble of excited donors in a solution, is rather additionally modulated by the total number of acceptor molecules N in the volume R_G^3, the particle density. The occurrence of emission events per time interval is proportionate to this decay probability:

$$\Phi = C\cdot\int_0^{\infty}\langle P(t)\rangle dt \tag{3.12}$$

Φ is the average quantum yield of the processes. If the decay is strictly exponential, it is possible to derive $\Phi_0 = C \cdot \tau_0$ and hence:

$$\frac{\Phi}{\Phi_0} = \frac{1}{\tau_0} \cdot \int_0^\infty \langle P(t) \rangle dt \tag{3.13}$$

Considering equation 3.11, one can define the constant $q = \sqrt{\pi} N R_0^3 / 2 R_G^3$ and the variable $s = t/\tau_0$ to yield:

$$\begin{aligned} \frac{\Phi}{\Phi_0} = \int_0^\infty \langle P(s) \rangle ds &= \int_0^\infty exp\left(-s - 2q\sqrt{s}\right) ds \\ &= 1 - q\sqrt{\pi} \cdot exp\left(q^2\right) \left[1 - \Sigma(q)\right] \end{aligned} \tag{3.14}$$

as a result of the integral, with Σ(q) being the GAUSSIAN error function. Obviously the defined constant q can be varied considering that it is proportional to the concentration of acceptors in the solution N/R_G^3. Hence it is:

$$q = \frac{\sqrt{\pi}}{2} \frac{C}{\left(\frac{4}{3}\pi R_G^3\right)^{-1}} = \frac{\sqrt{\pi}}{2} \frac{C}{C_0} \tag{3.15}$$

C_0 is entirely defined by the FÖRSTER radius (R_0) of the considered molecules. Expressions 3.14 and 3.15 describe the change in quantum yield of excited donors in a bath of potential acceptors in dependency of their concentration. The change in energy transfer between molecules will decrease the observed ensemble fluorescence anisotropy. This is due to the random orientation of acceptor molecules. The average fluorescence lifetime will not change.

HomoFRET in a Protein Cluster

FÖRSTER resonance energy transfer has the potential to resolve the oligomerization states of molecules due to the strong distance and orientation dependency of the process. It follows from equation 3.7 that the fluorescent excitation of a donor molecule will exponentially decrease:

$$P(t) = e^{-t/\tau} \left(\frac{1}{\tau} = k_T = \frac{1}{\tau_0} \left(\frac{R_0}{R} \right)^6 \right) \tag{3.16}$$

The quantum yield Φ of the excited molecule can be written as τ/τ_0. If one now assumes an oligomerization process of identical molecules that can undergo energy transfer (homoFRET), the probability for being in the excited state of an originally excited donor molecule can be expressed with the sum of de-excitation and re-excitation processes [45]:

$$-\frac{dP}{dt} = \frac{P(t)}{\tau} + \sum^N D - \sum^N R \tag{3.17}$$

Here, the first term is the de-excitation in absence of other molecules, whereas $\sum^N D$ and $\sum^N R$ describe the sum of all de-excitation and re-excitation (energy back transfer) processes, respectively, which occur due to N molecules in proximity to the original donor. From this Runnels & Scarlata [45] derived a description of the anisotropy and quantum yield readout for a cluster of N identical fluorescent molecules following the approach of Craver & Knox [15] for a situation with two molecules:

$$r_{tot} = r_D \frac{\Phi}{\Phi_{tot}} + \sum_{n \neq 1}^{N} r_{et}(n) \frac{\Phi_n}{\Phi_{tot}} \tag{3.18}$$

In this equation r_{tot} and r_{et} are the ensemble and transfer emission anisotropy, respectively. The latter one is simply the anisotropy arising from the n-th molecule in the cluster. Consequently, Φ is the quantum yield of the original donor and Φ_n the quantum yield of the molecule n in the cluster. These considerations finally yield the following equation [15, 45]:

$$r_{tot} = r_D \frac{(1 + k_T \cdot \tau)}{(1 + N \cdot k_T \cdot \tau)} + r_{et} \frac{(N - 1) \cdot (k_T \cdot \tau)}{(1 + N \cdot k_T \cdot \tau)} \tag{3.19}$$

to describe the change in anisotropy in a cluster of identical molecules. This theory has been applied e.g. by Bader *et al.* [113][143] to quantify *in cellulo* oligomerization states in fluorescence anisotropy microscopy. Another study by Ganguly *et al.* [165] similarly describes the utilization of homoFRET to quantify higher-order oligomers of serotonin 1A receptors in cells. Older studies have also tried to utilize simple fluorescence anisotropy measurements to observe early steps in protein crystallization [31].

III | Materials & Methods

4 | X-ray Diffraction Experiments

4.1 *In Cellulo* Powder Diffraction at Synchrotron X-ray Sources

Powder diffraction experiments were conducted at beamline P14 at the synchrotron radiation source PETRAIII (DESY, Hamburg, Germany) or at the ESRF (Grenoble, France). For data collection, yeast cells were obtained from yeast cell cultures at different developmental stages. Initially *MiTeGen MicroRT*^{TM} tubes were prepared with a small layer of 0.8% agarose in distilled water to bridge the non-uniform tip of the tube (See fig. 4.1). Afterwards tubes were infused with cooled cell suspension (4°C) and centrifuged at 1500 $\cdot g$ at 4°C for 5 min. Depending on the cell density this step has been repeated several times to obtain a dense cell pellet. The pellet was sealed with 0.8% agarose to prevent water evaporation. Lastly, tubes were glued onto a goniometer base. The samples were stored on ice for a maximum of 2 hrs. Measurements at the beamlines were always conducted at RT with a detached nitrogen-cryo stream. For measurements at MX beamlines the detector position has been

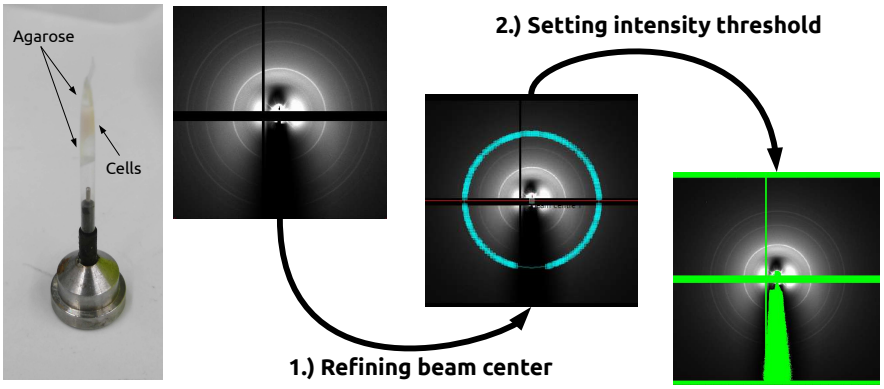

Figure 4.1: ***(Left)*** *Prepared yeast cells for powder diffraction with goniometer base* ***(Right)*** *Powder data processing procedure*

reallocated to distances of about 2000 mm to increase the visibility of low resolution reflections. The beam stop has beend moved as close as possible towards the sample. In case of beamline P14 in a distance of 34 mm. Usually measurements were conducted at a photon energy of 12.8 keV and at maximum beam flux. The beam diameter was refined via beam slits at 150 µm. Exposure times ranged from 1-30 s.
Data processing was done by utilizing the Eclipse-based data processing platform DAWN (Vers. 2.7.0), distributed and maintained by the Diamond Light Source (Oxfordshire, UK). Data processing comprised refinement of the beam

center, deposited in the original *.cbf* file, by fitting a ring to the powder pattern. Subsequently, blank data from the detector was removed by applying an appropriate intensity threshold. Afterwards an azimuthal integration with pixel splitting could be performed to obtain 1D powder patterns.
1D powder patterns were subjected to a rolling baseline correction. A defined background subtraction was not possible in every case due to varying degrees of pellet density. Peak fitting and further processing steps were performed using the software qtiplot (Vers. 0.9.8.9; Ion Vasilief, 2011). Peak fitting was performed under the assumption of a GAUSSIAN peak profile.

4.2 Williamson-Hall Analysis

Powder diffraction data has been processed according to section 4.1. 1D patterns were subjected to manual peak search and fitting to obtain parameters for a WILLIAMSON-HALL plot analysis.
The utilized plot method is attributed to G.K. Williamson and W.H. Hall [9] and relies on the principle that line broadening due to size and strain varies differently with the BRAGG angle according to [135]:

$$\beta_{tot} = \beta_{size} + \beta_{strain} = C\epsilon \cdot tan\theta + \frac{K\lambda}{L \cdot cos\theta} \tag{4.1}$$

Obviously one contribution varies with $^1/_{cos\theta}$ and the other one with $tan\theta$. If both effects are present they are somewhat convoluted. Williamson and Hall simplified the deconvolution by assuming that the resulting convolute is a simple sum of both sources. In equation 4.1 we find K to be the SCHERRER constant, λ the X-ray wavelength and L the crystallite size. The exact value of the SCHERRER constant depends on the exact peak shape, crystal shape and habit, but is usually assumed to equal 0.9 for spherical crystals [135].

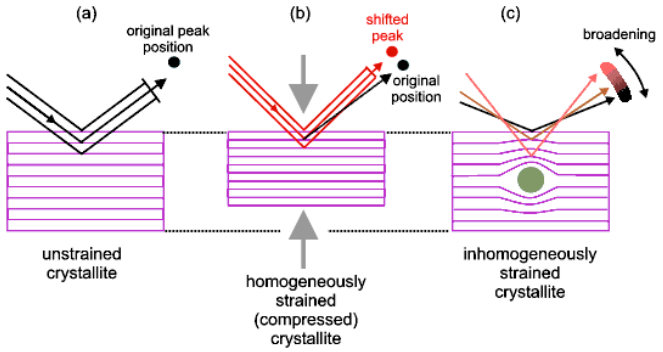

Figure 4.2: *Schematic depiction of crystal strain, Diffraction from* ***(A)*** *an unstrained* ***(B)*** *a homogeneously compressed and* ***(C)*** *an inhomogeneously strained crystal lattice; Picture taken from: Barnes, Jacques, Vickers (University of London)*

Strain is a measure of compression within a crystal. If a crystal is homogeneously compressed a simple peak shift (with regards to diffraction angle) would be measured. On the contrary, inhomogenous compression, as schematically shown in figure 4.2-C[1], yields a peak broadening. Strain can be introduced via interstitials, vacancies, dislocations or layer faults. Strain can be described by means of a mean inhomogeneity parameter ϵ and a related constant C. The exact value of the constant is based on assumptions concerning the strain, but

[1] http://pd.chem.ucl.ac.uk/pdnn/peaks/size.htm

is typically assumed to be $\sim 4-5$ [216].
Equation 4.1 can be transformed to be equivalent to a standard equation for a straight line:

$$\beta_{tot} cos\theta = C\epsilon \cdot sin\theta + \frac{K\lambda}{L} \tag{4.2}$$

From this it is possible to derive the slope $C\epsilon$ and the y-intercept $^{K\lambda}/_{L}$.

5 | Sample Purification & Characterization

5.1 General Laboratory Equipment

Table 5.1: *List of utilized laboratory devices*

Manufacturer	Device	Description
eppendorf (Germany)	5810 R	centrifuge
eppendorf (Germany)	miniSpin	centrifuge
eppendorf (Germany)	5424	centrifuge
eppendorf (Germany)	MasterCycler (Personal)	thermocycler
eppendorf (Germany)	BioPhotometer	OD_{600}
eppendorf (Germany)	Thermomixer C	heating- & stirring device
Heidolph (Germany)	MR 3001 K	magnetic stirrer
Scientific Industries (USA)	Vortex-Genie 2	vortex device
New Brunswick Scientific (USA)	Innova 44 / 40	batch incubator
Mettler Toledo (USA)	PL303	micro scale
Bio-Rad (USA)	ChemiDoc	gel imaging system
Bio-Rad (USA)	MicroPulser	electroporator device
Consort (Belgium)	EV245	electrophoresis power supply
Thermo Fisher Scientific (USA)	Power Ease 500	electrophoresis power supply
Thermo Fisher Scientific (USA)	Sorvall RC26 Plus, SS-34	centrifuge
Tecan (Austria)	Infinite M1000	plate reader
Agilent (USA)	Varian Cary 50 Scan UV-Vis	Spectrophotometer
Avestin (Canada)	Emulsiflex	high pressure homogenizer
Olympus (Japan)	IX-71, Power Supply 232	Microscopy Equipment
LEj (Germany)	HXP 120 V	compact light source
Serva (Germany)	Implen OD600	OD_{600}
GE Healthcare (USA)	Amersham plc Aekta Purifier, Frac-950	chromatography system
Beckman Coulter (USA)	Avanti I-20 XP, JLA-8.1	centrifuge

Table 5.2: *List of further chemicals and laboratory supplies*

Manufacturer	Device/Product	Description
MiTeGen (USA)	RT-T1 MicroRT	polymer tubing
WJM Glas, W. Mueller (Germany)	Borokapillaren	borosilicate capillaries
Kisker (Germany)	HR-PR-50, 50 mL	centrifuge tubes
Quiagen (The Netherlands)	Quiaprep Spin Miniprep Kit	dna purification
NEB (USA)	Monarch DNA Gel Extraction Kit	dna purification
Thermo Fisher Scientific (USA)	Slide-A-Lyzer Dialysis cassetes G2	dialysis supplies
Spectrum Laboratories (USA)	SpectralPor membrane tubings	dialysis supplies
Merck (Germany)	Corning Spin-X UF Concentrator	protein concentrator
GE Healthcare (USA)	Hitrap Q HP 5 mL	anion exchange column
GE Healthcare (USA)	Superdex 200 Increase GL	gel filtration column
Roth (Germany)	Rotiphorese Gel 40% acrylamide solution	gel supply
Roth (Germany)	TEMED	gel supply
Roth (Germany)	APS	gel supply
Sigma-Aldrich	Instant Blue	protein gel staining

5.2 Chromatography for Batch Protein Purification

Anion-exchange (IEX) and size-exclusion (SEC) chromatography have been conducted on an ÄKTAexplorer chromatography system (GE Healthcare, USA). For the preparation of purified batches of *H.p.* alcohol oxidase a HiTrap® Q HP 5 mL and a Superdex S200 column (both GE Healthcare, USA) have been used in IEX and SEC, respectively. Used buffers were filtered and degased beforehand. Sample flow was set to $1ml \cdot min^{-1}$, which was in agreement with the pressure specification of the used columns and chromatogrpahy systems.

5.3 Dynamic Light Scattering (DLS)

Dynamic Light Scattering epxeriments were either performed by using the DLS plate reader SpectroLight 600, or the cuvette-based setup SpectroSize300 (Xtal Concepts, Germany). The SepctroLight 600 setup utilizes a 100 mW laser with at $\lambda_{Em} = 660nm$ and operates at a fixed scattering angle of 142°with an APD. A similar laser source is used in the cuvette-based setup, but variable scattering angles could be set. Typically, measurements were conducted with a scattering angle of 90°. Measurements were always conducted at room temperature (293.15 K) for 30 s. The build-in autocorrelator (Xtal Concepts, Germany) covers a time interval from 0.4 μs to 30 s. The Xtal software allowed for a direct CONTIN analysis of the obtained autocorrelation functions.

5.4 Electron Microscopy (EM)

If not stated otherwise, transmission electron microscopy has been performed in collaboration with Dr. Rudolph Reimer at the Heinrich-Pette institute (HPI) for experimental virology (Hamburg, Germany). Protein fiber samples

were prepared over night at varying temperatures according to section 13.1. Afterwards, samples were transported to the HPI at ambient temperatures. In order to visualize all different types of fiber or protein structures the solution was mildly vortexed right before processing. At first, $400 \cdot 400 \mu m$ copper square meshes (Electron Microscopy Sciences, USA) were glow-discharged for 1 min at 25 mV (EmiTech, KX100). Thereafter, 2 μL of dense protein solution were pipetted onto the grid and incubated for 60 s. The Sample solution was immediately removed with *Whatman* paper and the grid was directly covered with 2-3% (w/v) uranyl acetate solution. After an incubation for 30 s the grid was washed two times with ultrapure water. Imaging was performed at a FEI Tecnai G20 transmission electron microscope operated at 120 kV in combination with a FEI Eagle 4k CCD camera.

5.5 SDS-Polyacrylamide Gel Electrophoresis

SDS gels were prepared according to table 5.3. 5x gel buffer was prepared by dissolving 74.4 g BisTris in 200 mL ultrapure water. Afterwards the pH has been adjusted to a value o about 6.5-6.8 with HCl. Usually MES running buffer has been used for standard SDS-PAGE. It was prepared according to 5.4. Protein samples were first treated with loading buffer (see tab. 5.5) and heated up to 75°C for several minutes. Afterwards gel electrophoresis was started by applying voltage of about 150-200 V. After a successful SDS-PAGE run, gels were washed and afterwards stained with Instant Blue over night.

Table 5.3: *Recipe for SDS-polyacrylamide gel preparation*

For 2 gels: 10 mL	**12%**	**10%**	**8%**	**Stacking**
5x gel buffer [mL]	2	2	2	0.8
40% acrylamide [mL]	3	2.5	2	0.5
Ultrapure water [mL]	4.9	5.4	5.9	2.7
20% APS [μL]	50	50	50	50
TEMED	10	10	10	10

Table 5.4: *20x SDS-PAGE running buffer composition*

Compound	**Conc.**
MES	97.6 g/500 mL
Tris base	60.6 g/500 mL
SDS	10 g
EDTA	3 g

Table 5.5: *4x SDS-PAGE loading buffer composition*

Compound	**Conc.**
Tris-HCl	200 mM
DTT	400 mM
SDS	8% (w/v)
Bromphenol blue	0.4% (w/v)
Glycerol	40% (v/v)

6 | Yeast Cultivation & Genetic Engineering

6.1 *H. polymorpha* & *S. cerevisiae* Yeast Strains

Table 6.1 holds information about the basic classification of all yeast strains which were obtained and used throughout this thesis. New *H.p.* strains (integration strains) were derived by the utilization of integration plasmids pHIPX4 or pHIPZ4. *S. cerevisiae* strains were transformed with the simple multi-copy vectors pMSC13, pCW6 or pCW12, which are described elsewhere. Original strains were obtaind either from the workgroup of Prof. Dr. Ida v.d. Klei at RUG Groningen, The Netherlands or from the workgroup of Prof. Dr. Ralf Erdmann at the RUB Bochum, Germany. *H.p.* 12DAC4 is a triple-deletion strain carrying deletions for all major peroxisomal matrix proteins: alcohol oxidase (AOX), dihydroxyacetone synthase (DHAS) and peroxisomal catalase (CAT). It still carries a leucine deficiency and is susceptible to zeocin treatment. Both markers can be used for transformation.

Table 6.1: *Yeast Strains and origins as obtained from collaborators, Δ indicates a gene deletion; RUG Groningen: workgroup of Prof. Dr. Ida. v.d. Klei, RUB Bochum: workgroup of Prof. Dr. Ralf Erdmann*

Species	Strain	Origin
H. polymorpha	NCYC495 wildtype	RUG Groningen
H. polymorpha	NCYC495 ΔPEX5	RUG Groningen
H. polymorpha	NCYC495 ΔPEX11	RUG Groningen
H. polymorpha	NCYC495 ΔAOX	RUG Groningen
H. polymorpha	NCYC495 ΔDHAS	RUG Groningen
H. polymorpha	NCYC495 ΔDHAS ΔCAT	RUG Groningen
H. polymorpha	NCYC495 PMP47-EGFP	RUG Groningen
H. polymorpha	NCYC495 12DAC4	RUG Groningen
H. polymorpha	NCYC495 12DAC4 EGFP-SKL	RUG Groningen
H. polymorpha	NCYC495 12DAC4, T4 Lysozyme(SKL)	RUG Groningen
S. cerevisiae	UTL7A	RUB Bochum
S. cerevisiae	BY4742	RUB Bochum

6.2 Nutritional Supplements & Cellular Growth

If not stated otherwise all chemicals used for yeast growth were obtained from Sigma-Aldrich, USA.

Yeast Rich Medium (Unselective)

Yeast rich medium is mainly constituted of two compounds. First, yeast extract, to supplement cells with a complex formulation of nutrients and secondly, peptone, to supply cells with easily consumable nitrogen. Both compounds are usually supplied in amounts of 1-2% (w/v). Additionally a carbon source needs to be present to enable growth. This is usually D-Glucose (Dextrose) but may vary depending on the experiment. If dextrose is supplied as the sole carbon source, the medium is usually called *Yeast Extract-Peptone-Dextrose* (YPD). The wildtype of the yeast *H. polymorpha* can instead also be supplied with methanol, or a combination of methanol and glycerol if P_{AOX} is intended to be induced. In such specific cases also other complementary sugars like D-Xylose or D-Ribose [107] can be used as de-repressing carbon source to replace glycerol. Depending on the intended growth pattern the carbon source concentration may be varied from 0.1-2% (w/v). The formulation of all compounds is summarized in table 6.2. Kanamycin may be added to prevent bacterial contamination. For preparation purposes peptone and yeast extract

Table 6.2: *YPD medium formulation for unselective yeast cultivation*

Compound	**Relative Amount [%]**	**Conc. [g/L]**
D-Glucose	0.5	5
Yeast Extract	1	10
Peptone	1	10
Sorbitol	1M	182.2

can easily be dissolved in water at RT. This solution is autoclaved overnight. A carbon source is prepared separately as described later in this section and added freshly whenever medium is used. For the maintenance of spheroblasts or fragile cells 1M sorbitol can be added additionally.

Synthetic-Defined (SD) Medium (Selective)

Synthetic defined medium comprises synthetic-complete as well as synthetic-minimal medium. Typically a variant is used that enables for the selection for certain markers. Most often enzymes necessary for the synthesis of tryptophan, uracil and leucine are knocked out in a cell line. If these cells are transformed with a plasmid containing the required enzyme sequences, colonies can be screened for successful transformation/ genomic integration. If the cells are not complemented with regards to a selection marker, the nutrient need to be supplied in the medium (cp. table 6.4).
The basis of SD medium is constituted of YNB whose compounds are listed in table 6.3. Freshly prepared media may be stored at a dark and cold place for 1 month. Once opened it should be used within 1-3 days. Kanamycin may be added to prevent bacterial contamination.

Table 6.3: *YNB without tryptophan, uracil & leucine (Sigma-Aldrich, Y0626), Medium Formulation*

Compound	Type	Conc.
Ammonium sulfate	Nitrogen Source	5 g/L
Biotin	Vitamin	2 μg/L
Calcium pantothenate	Vitamin	400 μg/L
Folic acid	Vitamin	2 μg/L
Inositol	Vitamin	2 mg/L
Nicotinic acid	Vitamin	400 μg/L
p-Aminobenzoic acid	Vitamin	200 μg/L
Pyridoxine HCl	Vitamin	400 μg/L
Riboflavin	Vitamin	200 μg/L
Thiamine HCl	Vitamin	400 μg/L
Citric acid	Vitamin	0.1 g/L
Boric Acid	Trace Element	500 μg/L
Copper sulfate	Trace Element	40 μg/L
Potassium iodide	Trace Element	100 μg/L
Ferric chloride	Trace Element	200 μg/L
Manganese sulfate	Trace Element	400 μg/L
Sodium molybdate	Trace Element	200 μg/L
Zinc sulfate	Trace Element	400 μg/L
Potassium phosphate (mono)	Salt	1 g/L
Magnesium sulfate	Salt	0.5 g/L
Sodium chloride	Salt	0.1 g/L
Calcium chloride	Salt	0.1 g/L

Table 6.4: *100x stock solution of selective markers leucine, tryptophan & uracil*

Compound	Conc.
Leucine	600 mg/50 mL
Tryptophan	400 mg/50 mL
Uracil	200 mg/50 mL

Synthetic-defined medium is prepared via the following steps:

1. Dissolve 6.7 g YNB w/o amino acids (1.7 g/L of YNB w/o AA w/o AS) in 845 mL ultrapure water to reach a final concentration 6.7 g/L
2. Autoclave solution, Let bottle cool down to 70°C
3. Add glucose to its final concentration (see above)
4. Add dropout supplements (see table 6.5) →1.47 g/50 mL, filter-sterilized (0.2 μm), Add completely to medium
5. Add selective markers (see table 6.4), 10 mL from 100x stock solution
6. Add potassium phosphate (K_2HPO_4,KH_2PO_4) from autoclaved 1 M stock solutions to a final concentration of 0.1 M at pH 6.

Table 6.5: *Dropout Supplements (Sigma-Aldrich, Y1771), Final compound concentrations when used at 1.47 g/L*

Compound	**Conc. [mg/L]**
Adenine	18
p-Aminobenzoic acid	8
Alanine	76
Arginine	76
Asparagine	76
Aspartic acid	76
Cysteine	76
Glutamic acid	76
Glutamine	76
Glycine	76
Histidine	76
myo-Inositol	76
Lysine	76
Methionine	76
Phenylalanine	76
Proline	76
Serine	76
Threonine	76
Tyrosine	76
Valine	76

Mineral Medium (Selective)

Mineral medium has been prepared as previously described by Van Dijken *et al.* [20]. The full medium contains basic nutrients as listed in table 6.6 and additional complex additives as listed in tables 6.7 & 6.8. The latter ones comprise trace elements and vitamins which are essential for growth. Full mineral medium also needs to be supplied with selective markers in case of non-selective growth. It can only be used for growth in liquid batch or chemostat cultures but not for the preparation of agar plates.

Table 6.6: *Mineral medium formulation for cultivation of H. polymorpha in a chemostat according to [20]*

Compound	**Type**	**Conc. [g/L]**
Ammonium Sulfate	Nitrogen Source	2.5
Magnesium Sulfate	Trace Element	0.2
Potassium Dihydrogen Phosphate	Salt	0.7
Monosodium Phosphate	Salt	3.0
Yeast Extract	Complex Nutrients	0.04
Vishniac Solution	Trace Element Solution	1 mL
Vitamin Solution	Vitamin Solution	1 mL

Table 6.7: *Vishniac solution according to [139]*

Compound	Conc. (v.d. Klei) [g/L]	Conc. (Original) [g/L]
$ZnSO_4$	4.4	22
$CaCl_2$	1.47	5.54
$MnCl_2$	1.01	5.06
$FeSO_4$	1.0	4.99
$CoCl_2$	0.32	1.61
$CuSO_4$	0.315	1.57
$(NH_4)_6Mo_7O_{24}$	0.22	1.1
EDTA	10	50

Table 6.8: *Vitamin solution according to [139]*

Compound	Conc. [g/L]
Biotin	0.1
Thiamine HCl	0.2
Riboflavin	0.1
Nicotinic acid	5
p-Aminobenzoic acid	0.3
Pyridoxine HCl	0.1
Ca-Panthothenate	2
Inositol	10

Carbon Source Stock Solutions

1. 50% (w/v) Glucose Solution
   (a) Heat up 20 mL ultrapure water in microwave
   (b) Dissolve 25 g of D-glucose by shaking
   (c) Add water to a final volume of 50 mL and shake until all glucose is dissolved
   (d) Use sterile 0.2 μm filter and a fresh syring for filter sterilization into new 50 mL falcon tube next to a bunsen burner or under the fume hood
2. 50% (v/v) Methanol Solution
   (a) Mix pure 25 mL methanol with 25 mL ultrapure water under the fume hood to obtain a 50% methanol solution
   (b) *(Optional)* Sterilize with 0.2 μm filter
   (c) *(Remarks)* Do not work next to a flame. Wear protective gloves and eye protection
3. 50% (v/v) Glycerol Solution
   (a) Mix pure 25 mL pure glycerol with 25 mL ultrapure water to obtain a 50% glycerol solution under the fume hood or next to a bunsen burner
   (b) Mix vigorously to get a clear solution

4. 100x Oleic Acid Stock Solution
   (a) Pipet 5 mL oleic acid from 100% stock into a 50 mL falcon tube
   (b) Add 10 mL Tween-40 (or -80)
   (c) Add ultrapure water to a total volume of 50 mL
   (d) Sterilize with 0.2 μm filter. The created suspension should turn milky by now
5. 100x Rytka Solution
   (a) Pipet 5 mL oleic acid from 100% stock into a 50 mL falcon tube
   (b) Add 10 mL Tween-40 (or -80)
   (c) Add 10 mL from 50% D-glucose solution
   (d) Add ultrapure water to a total volume of 50 mL
   (e) Sterilize with 0.2 μm filter. The created suspension should turn milky by now

6.3 Molecular Cloning

H.p. Integration Plasmids

The integration plasmids pHIPZ4 and pHIPX4 were obtained from the workgroup of Ida. v.d. Klei (RUG Groningen) and are described by Saraya *et al.* [177]. The plasmids have been altered via standard and SLiCE-cloning approaches to yield full plasmids for integration. Both plasmids can be conveniently used as shuttle vectors, to perform cloning with laboratory *E. coli* strains. For this purpose Ampicillin (pHIPZ4) or Kanamycin (pHIPX4) selection markers are integrated into the vector backbone. Final plasmids carry the

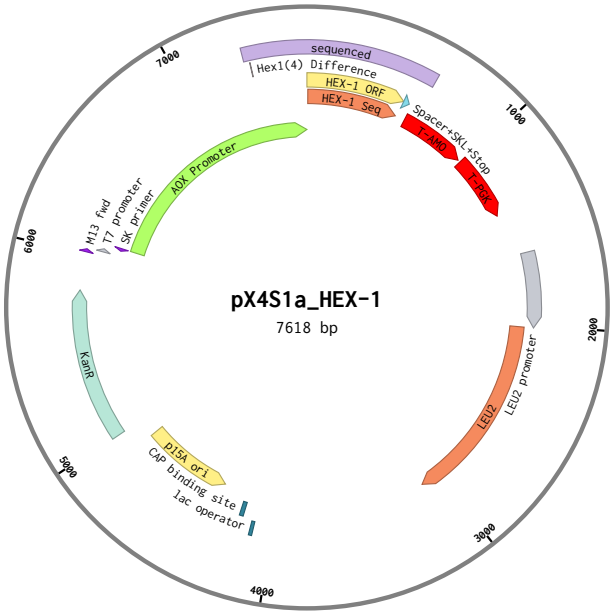

Figure 6.1: *Schematic depiction of the integration plasmid pX4S1 carrying the sequence of N.c. Hex-1 for integration into the yeast genome; The region that was sequenced after cloning is indicated in light purple; Image was created via the Benchling website (https://benchling.com)*

sequence of target proteins, a spacer peptide (GMDELYK) and an additional SKL tag. Transcription is initiated via the P_{AOX}-promoter and terminated either via both the T_{AMO}- and T_{PGK}-terminator, or only the latter one in case of the pHIPZ4 vector backbone. Selection is facilitated via the auxotrophic marker LEU2 (pHIPX4) or the Zeocin-resistance cassette (pHIPZ4). Successful cloning was verified via DNA sequencing (Eurofins Genomics, Germany). Obtained vectors were named either pX4S1 or pZ4S1 with regards to the introduced spacer peptide.

Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) is used to artificially generate copies of dsDNA fragments and is described in detail in literature. If not stated otherwise, a protocol as schematically described in table 6.9 has been used for the generation of DNA fragments for cloning and/or transformation. Annealing and elongation times have been individually adapted. Usually 30-35 cycles of heating, annealing and elongation have been applied. For all reactions Q5® High-Fidelity polymerase and the related buffers have been used (New England Biolabs, USA). The DNA primers used for SliCE cloning of target proteins are listed in table A.4.

Table 6.9: *Simplistic PCR protocol to obtain dsDNA fragments*

Description	**Time [s]**	**Temp. [°C]**
Initial heating	20-120	99
Heating	20-120	99
Annealing	10-20	50-72
Elongation	15-120	72
Final Elongation	120	72

SLiCE Cloning

SLiCE (Seamless Ligation Cloning Extract)- cloning has been used to facilitate creation of new plasmids for yeast transformation. The technique was first described by Zhang *et al.* [182] and is based on the usage of bacterial cell extracts to achieve site-directed ligation of PCR fragments and vector DNA. Cloning is facilitated via the annealing of homologous DNA regions (15-30 bps) between fragments and vector. At first, the vector backbone is linearized via digestion with one or two restriction enzymes and purified via agarose-gel electrophoresis. DNA fragments are created by PCR (see above; cp. tab. A.4). Afterwards purified PCR-fragments and vector are mixed with T4 ligase buffer and SLiCE medium. The solution is incubated for 30 min at 37°C and is followed by transformation into competent *E. coli* $DH5\alpha$ cells.

Transformation of *E. coli DH5α* Cells

1. Thaw competent cells on ice
2. Mix 1-5 µL of ligation reaction with competent cells
3. Leave mixture on ice for 20-30 min
4. Perform heat shock in water bath at 42°C
5. Place cell suspension immediately back on ice for 2 min
6. Add 500 µL of LB media at RT
7. Inc ubate suspension for 1 hr at 37°C
8. Plate out on selective LB agar plates

Agarose Gel-Electrophoresis

Agarose-gel electrophoresis is a technique to separate DNA in a agarose matrix with regards to size and charge of particles. Typcial gels were casted by using 0.8% (w/v) agarose dissolved in 1x TAE buffer (cp. table 6.10) adjusted to pH 8.

Table 6.10: *50x TAE buffer for agarose gel elctrophoresis*

Compound	**Conc.**
Tris-HCl	2 M
Acetic Acid	1 M
EDTA	50 mM

6.4 Electroporation of Budding Yeast

Electroporation of budding yeast is conducted for transformation of circular multi-copy plasmids in case of *S. cerevisiae* and of linearized integration-plasmids in case of *H. polymorpha*.

Buffers for Yeast Electroporation

Table 6.11: *TED Buffer; Used for the preparation of yeast cells for transformation*

Compound	Conc. [mM]
Tris-HCl, pH 8	100
EDTA, pH 8	50
DTT (fresh!)	25

Table 6.12: *STM Buffer; Transformation buffer for yeast*

Compound	Conc. [mM]
Tris-HCl, pH 8	10
Sucrose	270
$MgCl_2$	1

Preparation of Competent Cells

1. Grow cells overnight at 30°C/37°C
2. Innoculate 50 mL (*pre-heated*) YPD or mineral medium with 0.5 mL from overnight culture and allow further growth at the appropriate temperature
3. Allow growth until an OD_{600} of 1.2-1.5 is reached
4. Pellet cells 10 min at 6000 rpm (table-top centrifuge) at RT
5. Resuspend cells in 0.2 vol. *pre-heated* TED buffer
6. Incubate cells 15 min at 30°C/ 37°C, while shaking at 200 rpm
7. Pellet cells 10 min at 6000 rpm (table-top centrifuge) at RT
8. Wash cells twice with 1.0 vol. in ice-cold STM buffer
9. Pellet cells 10 min at 6000 rpm (large table-top centrifuge) at 4°C
10. Resuspend cells in 0.25 mL ice-cold STM buffer
11. Divide the suspension with competent cells into batches 60 µL (Use immediately or freeze in liquid nitrogen, followed by storage at -80°C)

Electroporation-Protocol

1. *(H. polymorpha-specific)* Linearize integration plasmid pHIPX/pHIPZ or others by using an appropriate restriction enzyme or PCR-reaction
2. *(H. polymorpha-specific)* Purify linearized plasmid via a DNA purification kit and elute in smallest possible volume for transformation
3. Add linear/plasmid DNA (max. 3-4 µL with 100-10,000 $\mu g \cdot \mu L$) to cooled competent cells and incubate for 10 min on ice
4. Transfer cell/DNA mixture to a 2 mm electroporation cuvette
5. Apply electroporation pulse (1.5 - 3 $kV \cdot cm^{-1}$)
6. Immediately add 940 µL YPD (RT)
7. Incubate cells 1-2 hr at 37°C while shaking at 200 rpm
8. Pellet cells 2 min at 7000 rpm (small table-top centrifuge) at 4°C
9. Wash cells twice with 1.0 vol. water or YNB
10. Resuspend cells in 100 µL water or YNB and plate out on selective plates

Clone Selection

After transformation of yeast cells (see above) yeast cells were plated out either on selective YPD or YNB-based agar plates containing 2%(w/v) agar-agar. Plates were incubated at 30°C (*S. cerevisiae*) or 37°C (*H. polymorpha*). Depending on the utilized plasmid, either 200 µg/mL Hygromycin B, or Zeocin have been used in for yeast clone selection on YPD plates. For auxotrophic strains, YNB plates without the specific selection marker have been used. To prevent evaporation of water, agar plates were sealed with *Parafilm*. Usually colonies appeared after 3-5 days. Colonies were restreaked again on fresh selective plates (as depicted in fig. 6.2) to avoid false-positive transformants. Those were again incubated for about 2-3 days to yield enough cell material for the purification of genomic DNA. Due to many falsely integrated plasmids usually about 100 hundred colonies needed to be obtained and screened to yield a correctly integrated plasmid clone.

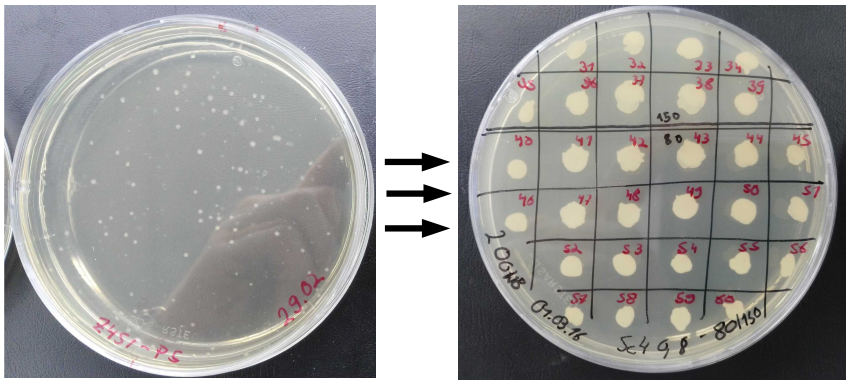

Figure 6.2: *Selection of yeast transformants on selective growth media and restreaking after the appearance of the first colonies*

Isolation of genomic DNA

Isolation of genomic DNA was necessary to allow for reliable PCR conditions. A protocol from Harju *et al.* [95] has been adapted to obtain pure genomic DNA:

1. Pick a yeast colony from a selective plate after two days growth (preferentially freshly grown cells)
2. Resuspend cells in 100 µL lysis buffer (see table 6.13)
3. Rapidly freeze cells at -80°C for 2-3 min
4. Rapidly heat cells to 98°C for 1 min
5. Repeat freezing/heating 3 times
6. Vortex vigorously for 30 s
7. Add 200 µL chloroform and vortex for 2 min
8. Centrifuge 5 min at $20000 \cdot g$ at RT (in table-top centrifuge)
9. Transfer **upper** aqueous layer into a tube which contains 400 µL **ice-cold** 100% ethanol
10. Allow samplpe precipitation for 5 min on ice
11. Centrifuge again 5 min at $20000 \cdot g$ at RT
12. Remove supernatant and let pellet dry for several minutes
13. Wash DNA pellet with 0.5 mL 70% (v/v) ethanol
14. Repeat drying step
15. Resuspend pellet in 50-100 µL ultrapure water
16. Use about 1 µL (should be about 150 ng) for a PCR reaction

Table 6.13: *Lysis buffer for purification of yeast genomic DNA from plate cultures*

Compound	**Conc.**
Triton X-100	2%
SDS	1%
NaCl	100 mM
Tris-HCl, pH 8.0	10 mM
EDTA, pH 8.0	1 mM

6.5 Purification of Peroxisomes

Protocols of this section were partially obtained from the workgroup of Prof. Dr. Ida v.d. Klei (RUG Groningen, The Netherlands).

Buffers for Peroxisome Purification

Table 6.14: *Pre-Incubation Buffer; Used for the removal of growth media*

Compound	Conc. [mM]
Tris-HCl, pH 8	100
EDTA	50

Table 6.15: *Protoplast Buffer; Used for the removal of the cell wall via zymolyase*

Compound	Conc. [mM]
Potassium phosphate, pH 7.2	50
Sorbitol	1200

For the preparation of protoplast buffer the following volumes from prepared potassium phosphate stock solutions can be used (for 1 L): 35.85 mL 1 M K_2HPO_4 & 14.15 mL KH_2PO_4.

Table 6.16: *Homogenization Buffer, final pH 5.5; Used to maintain osmolytic equilibrium to prevent protoplast disruption; without Sorbitol this buffer yields 'buffer B'*

Compound	Conc. [mM]
MES, pH 5.5	5
EDTA	0.1
KCl	1
Sorbitol	1200

Protocol for the Preparation of Peroxisomes

1. Protoplast-preparation
   (a) Determine the weight of the empty centrifuge tubes
   (b) Harvest cells by centrifugation, 5 min at 3800 $\cdot g$ at RT (4000 rpm large table-top centrifuge)
   (c) Determine wet weight of pellets
   (d) Resuspend cells at 0.06 $g \cdot mL^{-1}$ in pre-incubation buffer and add 140 mM β-mercaptoethanol (or 70 mM fresh DTT)
   (e) Incubate for 15 min at 37°C
   (f) Pellet cells 2 min at 4000 rpm at RT
   (g) Wash cells twice in *pre-heated* protoplast buffer
   (h) Resuspend cells in *pre-heated* protoplast buffer and add Zymolyase ($1mg \cdot mL^{-1}$)

(i) Incubate at 37°C with marginal shaking (50 rpm)

(j) Verify protoplasting via light microscopy

(k) Collect protoplasts for 10 min at 4000 rpm at 4°C

2. Homogenization of protoplasts

(a) Add 2.5 µg $\cdot mL^{-1}$ Leupeptin and 1 mM PMSF (or other protease inhibitors) to homogenization buffer

(b) Wash protoplast pellet once with ice-cold homogenization buffer

(c) Carefully detach pellet by using a spatula

(d) Centrifuge for 10 min at 4000 rpm at 4°C

(e) Discard supernatant

(f) Resuspend pellet vigorously in ice-cold homogenization buffer by using a syringe (a potter homogenizer might be used)

(g) Check disruption of protoplast by light microscopy

3. Differential centrifugation

(a) Centrifuge cell homogenate for 10 min at 3000 $\cdot g$ at 4°C

(b) Collect supernatant and repeate centrifugation (to obtain post-nuclear supernatant)

(c) Collect supernatant and centrifuge for 30 min at 30000 $\cdot g$ at 4°C

(d) Resuspend organellar pellet in 40% (w/v) sucrose solution in buffer B (see below)

Table 6.17: *Mixing ratios of buffer B and 65% (w/v) sucrose buffer B to yield solutions for the preparation of layered sucrose density gradients*

Sucrose [%]	**65% sucr. buffer B [mL]**	**Buffer B [mL]**	**Total vol. [mL]**
52	22.6	7.4	30
50	21.6	8.4	30
48	33.9	15.8	50
46	26	14	40
44	24.6	15.4	40
40	16.5	13.5	30
35	14.1	15.9	30

Table 6.18: *Volumes for single sucrose density gradients for organelle purification either with an organellar pellet (OP) or post-nuclear supernatant (PNS); OP gradients are overlayed by 35% sucrose solution while PNS gradients are overlayed by HOMO-buffer; For both, not more than 10 mL sample (depending on conc.)*

OP		**PNS**	
Sucrose [%]	**Vol. [mL]**	**Sucrose [%]**	**Vol [mL]**
65	6.5	65	6.5
52	2.5	52	5
50	2.5	50	5
48	7.5	48	7
46	4	-	-
44	5	-	-

4. Sucrose density centrifugation

   (a) Prepare sufficient amounts of buffer B/ 65% (w/v) sucrose in buffer B (For buffer composition see table 6.16)

   (b) Prepare gradient according to table 6.17 and table 6.18

   (c) Centrifuge gradient for 10 hrs at 36000 $\cdot g$ at 4°C (Slow acceleration and deceleration; Turn off centrifuge brakes)

   (d) Harvest peroxisomes from highest density fraction interface with a serological pipette (Remove top fractions without mixing)

   (e) Dilute peroxisome fraction 1:1 in homogenization buffer (See table 6.16)

   (f) Centrifuge peroxisomes for 30 min at 30000 $\cdot g$ at 4°C

   (g) Resuspend peroxisome pellet in 1 mL homogenization buffer

   (h) Fixation might be conducted with 0.2% glutaraldehyde, overnight in the cold room in a slowly rotating manner

5. Remarks

   - Timing on harvesting of protoplasts is crucial: It can take 5 min to 2 hrs and should be constantly monitored
   - The used protease inhibitor cocktail contained 12.7 mg$\cdot mL^{-1}$ of AEBSF-HCl, Aprotinin, E-64 and Leupeptin
   - Both post-nuclear supernatant and organellar pellet can be used for the purification of peroxisomes
   - Air bubbles need to be avoided on preparation of sucrose gradients
   - gradient can easily be poored by freezing them layer-by-layer in a -80°C freezer
   - Preparation of 65% sucrose solution works best by heating up the solution in a waterbath under constant stirring. Only small amounts of sucrose should be added successively
   - 1L of batch culture at OD 1-1.5 should yield approx. 2.5 mg of peroxisomes
   - Peroxisomes should be stored at 4°C

Catalase-Assay

To determine peroxisomal fractions a simple qualitative but efficient assay has been utilized, which is based on the detection of catalase. As shown by Iwase *et al.* [186] catalase activity can be determined and compared via foaming in a solution containing hydrogen peroxide and Triton X-100. While hydrogen peroxide acts as a direct substrate for the enzyme, the produced oxygen is trapped in small bubbles formed by the detergent. The stated assay reproducibility at 100U activity is about 4.7%. As shown in figure 6.3-A/B the foaming is linearly proportional to the catalase activity over a wide range of activities.

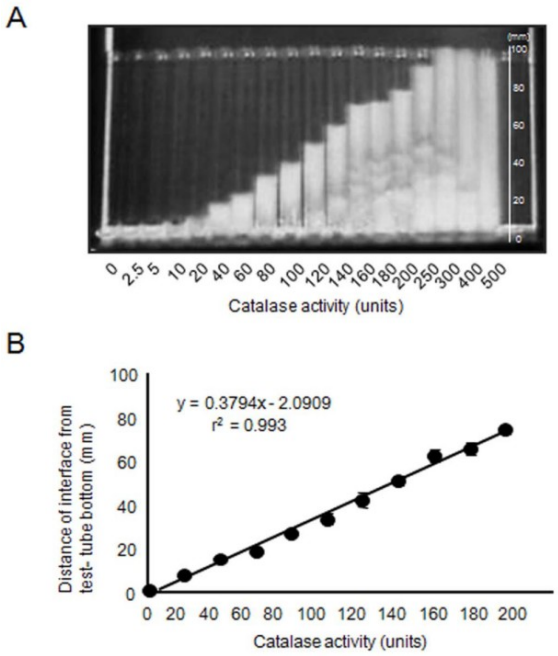

Figure 6.3: *Qualitative catalase assay* ***(A)*** *Image of test tubes showing foam development in 100 mL catalase solution* ***(B)*** *Linear fit between foam heights and catalase activity observed over a range of 20-200U; Image taken from: [186]*

7 | Fluorescence-based Assays

7.1 Fluorescence Microscopy

Fluorescence microscopy experiments were carried out on an Olympus IX-71 inverted microscope, in combination with a HXP 120 C metal halide lamp (Leistungselektronik JENA GmbH) and an Olympus FITC filter set. All experiments were carried out at 100% lamp intensity. Recorded image intensities were regulated via absolute exposure times. All image acquisition was done with a 100x immersion oil, super-apochromat and coverslip-corrected objective (UPLSAPO 100XO, WD: 0.13 mm, NA: 1.4). For image acquisition a Zeiss AxioCam mRm CCD camera has been used in combination with the AxioVision software suite (Vers. 4.8). The setup is schematically depicted in figure 7.1.

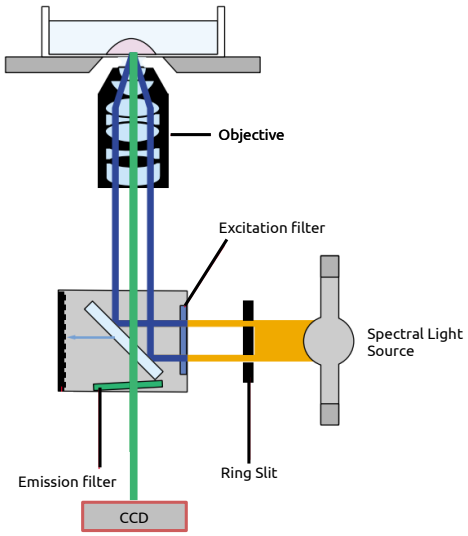

Figure 7.1: *Microscopic setup utilized for fluorescence and whitelight imaging (Blue) Excitation- (Green) Emission light path; Image adapted from manual: "Olympus Research Inverted Microscope IX71/IX81, IX2 Series" (Olympus Corp., Japan)*

Approximate Resolution and Theoretical Setup-specific PSF

For the incorporation of a scale into the recorded microscopy images a theoretical pixel size has been calculated via the following formula:

$$PxSize = \frac{ChipPxSize \cdot binningF}{ObjMagn \cdot LensMagn} \tag{7.1}$$

Here, *ChipPxSize* is the size of a single pixel on the CCD detector chip, *binningF* is the software binning factor, *ObjMagn* is the stated magnification of the used

objective and *LensMagn* is the magnification factor of further installed lenses (e.g. ocular). With a pixel edge length of 6.49 μm and a 100x magnification of the used objective an image pixel edge length of 65 nm can be calculated. It has been used as a scale for microscopic imaging and data analysis throughout this thesis.
For deconvolution of recorded images a theoretical point spread function (PSF) has been calculated via the SVI Nyquist calculator[1] as shown in figure 7.2-A/B. The PSF is shown in the xy- and xzplane in subfigure A and B, respectively. Its FWHM in x,y is 91 nm and 274 nm in the z dimension.

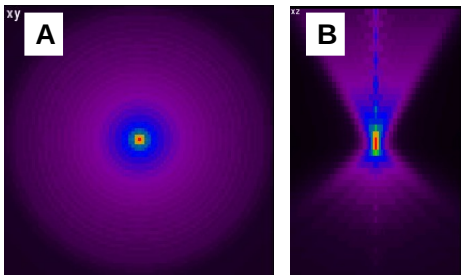

Figure 7.2: *Theoretical PSF for microscope and specimen specifications as described in table 7.1;* ***(A)*** *xy plane* ***(B)*** *xz plane*

Table 7.1: *Microscope and specimen specifications for the calculation of a theoretical PSF*

Parameter	**Value**
Microscope Type	Widefield
Numerical Aperture	1.4
$\lambda_{\text{Exc.}}$	488 nm
$\lambda_{\text{Em.}}$	512 nm
$n_{\text{Immersion Oil}}$	1.518
$n_{\text{Specimen Medium}}$	1.34
Acq. Depth	0.0

Sample & Glass Slide Preparation

For the acquisition of whitelight and fluorescence microscopy images, yeast cells were harvested and concentrated via centrifugation at 3000 rpm for 2 min in a table-top centrifuge. Glass-slides were cleaned with ultrapure water and 70% (v/v) ethanol to remove dust and dirt particles. Afterwards 1-3 μL of cell suspension were pipetted onto the slide and immediately covered by a cover slide. Slow evaporation of the liquid allowed for unbiased measurements of cell suspensions for about 30 min.

Z-Stack Recording & Data Processing

Z-stacks were manually recorded by single image acquisition in tiff format every 300 nm, according to the z-dimension of the theoretical PSF (see above). Usually 7-9 images were recorded per stack to cover the full diameter of a yeast cell. Afterwards images were converted into full stacks by utilizing the

[1] https://svi.nl/NyquistCalculator

software Fiji [178], which has also been used to perform all following analysis steps. Subsequently, a bleaching correction following a simple ratio method has been conducted. This step was necessary to allow for a homogenous analysis of images in later steps. In a further step the setup-specific PSF has been used for constrained iterative deconvolution with parameters as listed in table 7.2. Afterwards an averaged intensity image has been created from processed

Table 7.2: *Parameters for constrained iterative image deconvolution with the DAMAS3 plugin for the Fiji software suite [178]*

Parameter	**Value**
WIENER filter γ	0.001
Low pass filter (xy)	1 px
Low pass filter (z)	1 px
Maximum iterations	5
Termination at $\delta < x\%$	0.001

z-stacks. This has been subjected to a CANNY edge detection analysis (Canny Edge Detector Plugin, Tom Gibara, 2013) and subsequent transformation into a binary image. The binary image has been used to analyze sizes and intensities of the underlying regions of interest (ROI). The same processing parameters have been used for all data points.

7.2 Thioflavin T- Protein Fiber Assay

Thioflavin T (ThT) is a fluorescent marker dye commonly used for the detection of amyloid fibrils. It belongs to the benzothiazole family, which exhibits strong fluorescence enhancement upon binding to insoluble βsheet-rich, fibrillar protein structures [206].
The photophysical mechanism underlying the fluorescence enhancement is not fully understood, but it is postulated that ThT (fig. 7.3) behaves as a molecular rotor, which is quenched due to rotational motion [206]. In viscous environments the increasing hindrance of rotational motion leads to an increase in emitted fluorescence photons. The mode of binding to fibrillar structures is not yet elucidated and explanations have been given that involve monomeric, or multimeric binding modes [150, 154]. ThT lacks strong interaction with most soluble protein, amyloid oligomers and protofibrils.

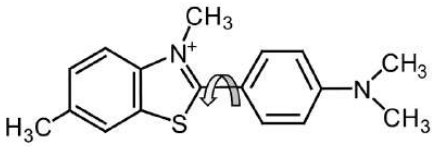

Figure 7.3: *2D representation of thioflavin T with main rotation axis; Image taken from [206]*

A 1 M ThT stock solution was prepared by dissolving 0.16 g of dry ThT (318.8 $g{\cdot}mol^{-1}$) in 500 µL distilled water. The solution was heated up to 80°C if needed for an assay and stored at 4°C.
For monitoring protein fiber formation, ThT was mixed 1:2 (protein:ThT) with a protein solution in the µM range and in a total volume of 150 µL. ThT

fluorescence emission was monitored in a spectrometric plate reader at $\lambda_{Em.}$: 485 nm, under excitation at $\lambda_{Exc.}$: 442 nm.

7.3 Fluorescence Anisotropy & Plate Reader-based Screening

Standard fluorescence intensity, -anisotropy and optical density measurements were conducted on a spectrometric Tecan Infinite® M1000 plate reader in combination with the corresponding i-control software (Tecan, Austria) and non-transparent 96 well plates. The setup is driven by a Xe-flash lamp and two tunable double monochromators for excitation and emission bandwidth selection. The final spot size FWHM is approximately 2 mm. Detection is achieved by utilizing a photo-multiplier tube (PMT) detector[1]. The intensity-related setup optics are depicted in figure 7.4. The schematic setup of anisotropy-related optics, including dichroic mirrors ,rotator and polarizer are shown in figure 7.5. Intensity differences between pure intensity and anisotropy measurements are due to the difference in the utilized light sources. While intensity measurements are conducted via the integrated Xe-flash lamp, anisotropy data is obtained via a specific blue LED with an emission maximum at $\lambda_{Em} = 470nm$. Anisotropy is a dimensionless quantity as seen from equation 3.1. It is denoted either as r or in artifical units of Ans/mAns throughout this thesis.

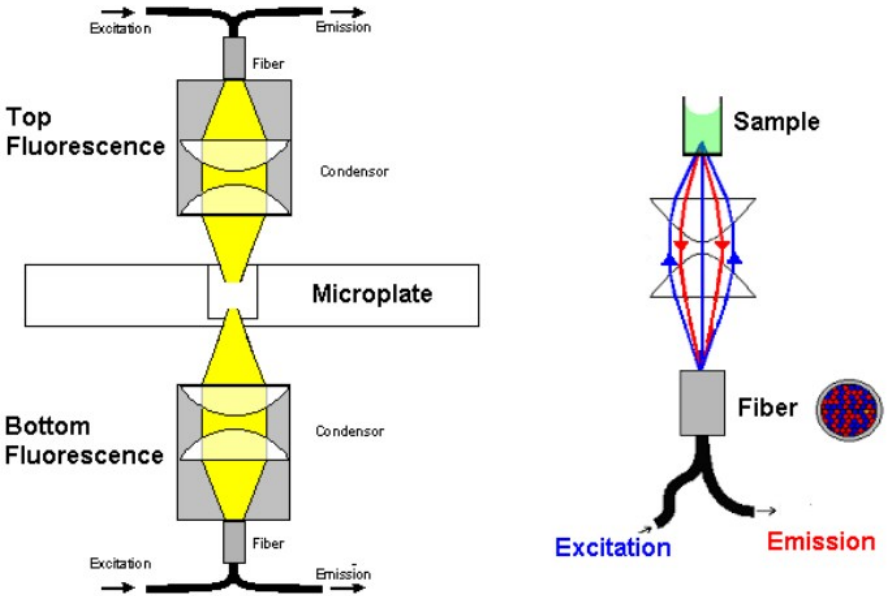

Figure 7.4: *Schematic representation of Tecan Infinite M1000 plate reader optics for fluorescence intensity measurements; Image taken from the official Tecan manual (see footnote)*

For anisotropy measurements a setup-specific G factor of 1.150 was applied throughout all measurements, which determines the weighting of parallel, perpendicular polarized light for the calculation of anisotropy values. Gain factors were used as indicated in the results section. The gain factor is a detection specific parameter to alter the detector voltage. Additionally, a z-optimization has been conducted to determine the optimum working distance for a solution of 200 μL cell suspension or protein solution. Different distances were tested for EGFP fluorescence intensity to yield a optimum parameter of 2.3 cm. All

[1]Instructions for Use for Infinite M1000 No. 30036266 Rev. No. 1.6, 2011

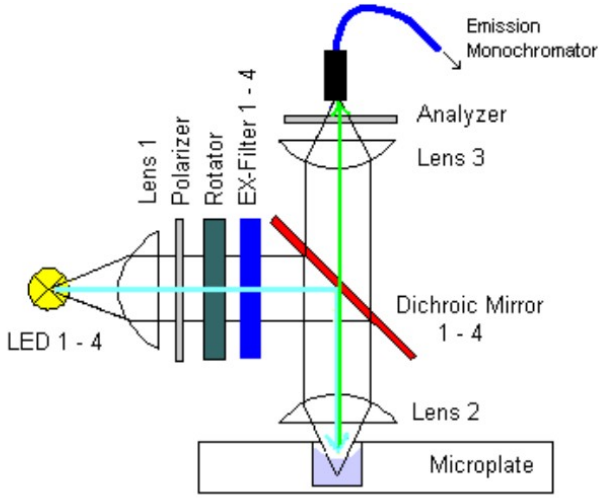

Figure 7.5: *Schematic representation of Tecan Infinite M1000 plate reader optics for fluorescence anisotropy measurements; Image taken from the official Tecan manual (see footnote on the page before)*

measurements were conducted with a flash rate of 400 Hz and 50 flashes in total. Data post-processing and analysis was done by utilizing the software qtiplot (Vers. 0.9.8.9; Ion Vasilief, 2011).

Fluorescence Bleaching

Bleaching of intracellular EGFP fluorescence or in solution has been conducted in 1 mL UV-transparent cuvettes in a self-made cuvette holder (see fig. 11.16). Bleaching was realized by applying full intensity radiation of a HXP 120 C metal halide lamp (see above). To counteract on sample deterioration due to heating, samples were continuously cooled with ice. Bleaching was usually conducted gradually in steps of 1-2 minutes, follwed by extra cooling and plate reader measurements.

Perrin Plot Analysis

A Perrin plot analysis was only conducted on fully bleached samples (subsec. 11.1.3). The plate reader setup could not automatically cool down to temperatures below 20°C. Therefore, test samples at 4°C were filled into 96 well plates and the time which is needed to heat up to 10, 15 and 20°C at ambient temperatures has been measured. The approximate time was used to measure cell samples at temperatures below 20°C. The setup-internal heating has been utilized for temperatures above 20°C. Samples were allowed to equilibrate at a new temperature for 10 min before a measurement was conducted.

Plate Reader Assay

Optical density, fluorescence intensity and anisotropy were measured in the setup described above and as schematically depicted in figure 7.6. Optical densities were measured in transparent 96 well plates. Afterwards cell suspensions were pipetted into non-transparent plates to measure fluorescence intensity and anisotropy.

For all measurements 100µL of KBr (or another quenching agent) were thoroughly mixed with 100 µL of cell suspension at various optical densities. Cell suspensions were grown for 24-36 hrs in 5 mL falcon tubes with perforated lids. Media and growth conditions are described section 6.2.

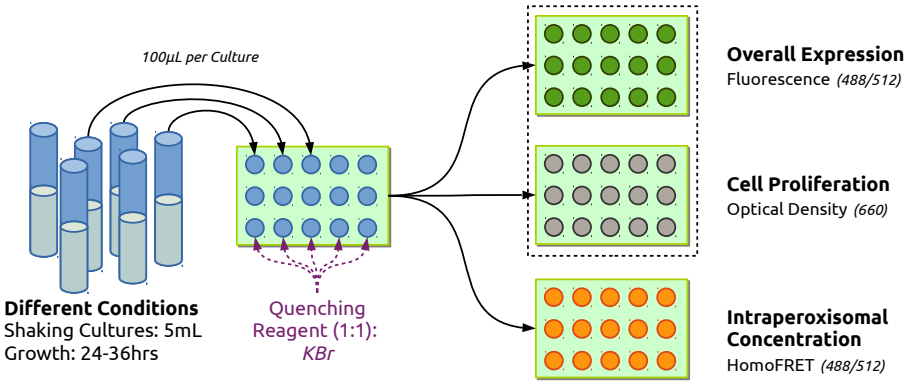

Figure 7.6: *Schematic representation of a plate reader-based screening method for yeast in vivo crystallography, used to track system properties upon introduction of system perturbations*

IV | Results & Discussion

8 | Engineering the Peroxisomal Crystal Factory

Heterologous expression of proteins in different budding yeast species is utilized in scientific research and industrial applications for different purposes. Here, the yeast *H. polymorpha* has been chosen as the prime target organism for cloning and expression attempts due to its unique features with regards to promoter strength and peroxisomal induction (sec. 2.1). This chapter starts with a verification of the induction of P_{AOX} under new laboratory conditions (sec. 8.1). Additionally, *in vitro* crystallization attempts of purified *H.p.* AOX are presented to complement data from powder diffraction and SFX experiments (sec. 8.2). Lastly, section 8.4 exemplary describes the successful heterologous expression and peroxisomal targeting of proteins in the organisms *H. polymorpha* and in *S. cerevisiae*.

8.1 AOX Purification & Verification of P_{AOX} Induction

In a first step, it was necessary to verify the induction of 75 kDa alcohol oxidase (AOX)[1] expression in *H. polymorpha* under new laboratory conditions. Without having the advantage of a specific antibody the protein was directly purified from yeast cells and identified via mass spectrometry. This was a necessary approach due to the extremely weak expression in comparison to bacterial systems.
Furthermore, an attempt was made to establish a purification protocol to obtain functional protein for an *in vitro* crystallization attempt (see sec. 8.2). This was done in a multi-step approach comprising ammonium sulfate (AS) precipitation, anion exchange- and size-exclusion chromatography (SEC). The protocol for AOX purification from yeast cells has been adapted from van der Klei *et al.* [35]. It has to be remarked that cultivation of yeast cells had been difficult with its usually complex media requirements (see sec. 6.2). To simplify the cultivation it was tried to grow cells in unselective YP media supplemented with glucose/methanol as the sole carbon source.

Purification & Identification via Mass Spectrometry

For the purification of AOX protein from yeast cells the following steps were conducted. In addition, proof is given that pure, monodisperse protein was obtained which was subjected to crystallization trials.
Yeast cells were harvested from 6×200 mL batch cultures, inoculated from a YPD pre-culture and grown for 24 hrs in 0.5% YPM media. All media

[1] Usually assembles into an octameric structure of ∼ 600 kDa

contained 0.1 M KH_2PO_4/K_2HPO_4 at pH 6.0 to prevent early acidification during growth. Cells were spun down in a table-top centrifuge at 5000·g for 10 min at 4°C. The pellet was washed twice in 50 mM potassium phosphate (PP) buffer pH 7.5. All following steps were equally performed at 4°C to prevent degradation. In the next step, cells were subject to 5 cycles of emulsifier treatment (1 g cells/5 mL, $1.4 \cdot 10^5 kN/m^2$) to break most of the cell walls. The breakage was monitored via microscopy. To remove cell remains, the lysate was centrifuged for 30 min at 30000·g. The obtained supernatant was used for ammonium sulfate precipitation. To collect the fraction containing AOX at first a 45% (w/v) saturation was reached by adding 66.67 g AS to 250 mL buffer. The solution was kept on ice in the cold-room for about 30 min and afterwards centrifuged for 20 min at 10000·g. In a further precipitation step 17.46 g AS were added to the supernatant to obtain a 55% (w/v) concentration of AS. The total volume of the lysate was now about 295 mL. The solution was again kept on ice for 20 min.

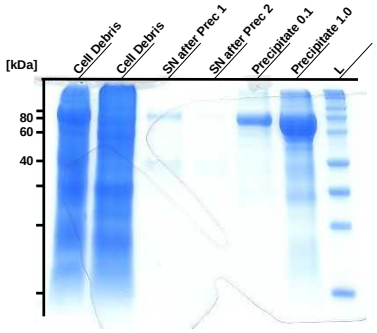

Figure 8.1: *12% SDS-PAGE with samples representing cell debris, supernatant and precipitate during ammonium sulfate precipitation as a first step in AOX purification; L= Protein Weight Marker (Roti®-Mark 10-150)*

The precipitated fraction between 45% and 55% was collected as it contains most of the AOX activity and was resuspended in 1 mL 50 mM PP buffer pH 7.5. The solution was subject to an overnight dialysis step against the same buffer (2 L) to remove most of the ammonium sulfate. The results were checked on a SDS gel as depicted in figure 8.1. The gel shows initial cell debris, cell debris after centrifugation, two lanes supernatant after the precipitation steps and two lanes from the resuspended AOX fraction between 45% and 55% AS from left to right. It is apparent that a strong band appears for the resuspended fraction at around 75 kDa which corresponds to the monomer weight of AOX. The band is slightly visible in the cell debris, too, indicating that not all cells/organelles were broken during purification. The concentration of the rightmost lanes were 0.1 and 1.0 mg·mL^{-1}, respectively, both in 20 µL loading volume. The other samples were applied as obtained.
Hereafter, the solution was first concentrated and then purified further by using a HiTrap® Q HP 5mL column. Equilibration was done with 50 mM PP pH 7.5 buffer and 0 M AS. This represents the low salt buffer A. A second buffer denoted as buffer B (or elution buffer) is used for 'salting out' and contains additional 0.5M AS. This buffer is used in a linear gradient to change the ionic strength and elute AOX and other proteins separately from the column. The results are shown in figure 8.2-A and -C. As it can be seen in figure 8.2-A, a certain fraction of molecules elutes as nonretained from the column without an increased salt concentration (peaks *X1* and *X2*). After increasing buffer B

content two major peaks elute from the column (peak *B-C* and *D-E*). As it can be seen in figure 8.2-C, only peak *D-E* primarily contains a protein of the desired size of 75 kDa. In addition, this peak has a uniform shape indicating the same partition coefficient of the contained molecules. However, all other peaks show a peak tailing effect. A closer look at figure 8.2-C reveals a band of higher molecular weight (above 150 kDa) in fractions of peak *D-E* even under the denaturing conditions of a SDS-PAGE gel and might correspond to a dimeric sub fraction of the target protein. As the results indicated a second successful purification step the collected fractions from peak *D-E* were pooled, concentrated and subject to size-exclusion chromatography.
Gel filtration was performed on a *Superdex* S200 column and the related results are depicted in figure 8.2-B and -D. As it can be seen in sub-figure B, a single uniform peak emerges on the gel filtration column indicating protein in a very homogenous state. The elution volume corresponds approximately to a molecular weight of 70-80 kDa. Intriguingly, a possible dimeric sub fraction is again visible under SDS gel conditions. A further explanation for this observation could be a minor impurity that is eluting in a similar pattern.

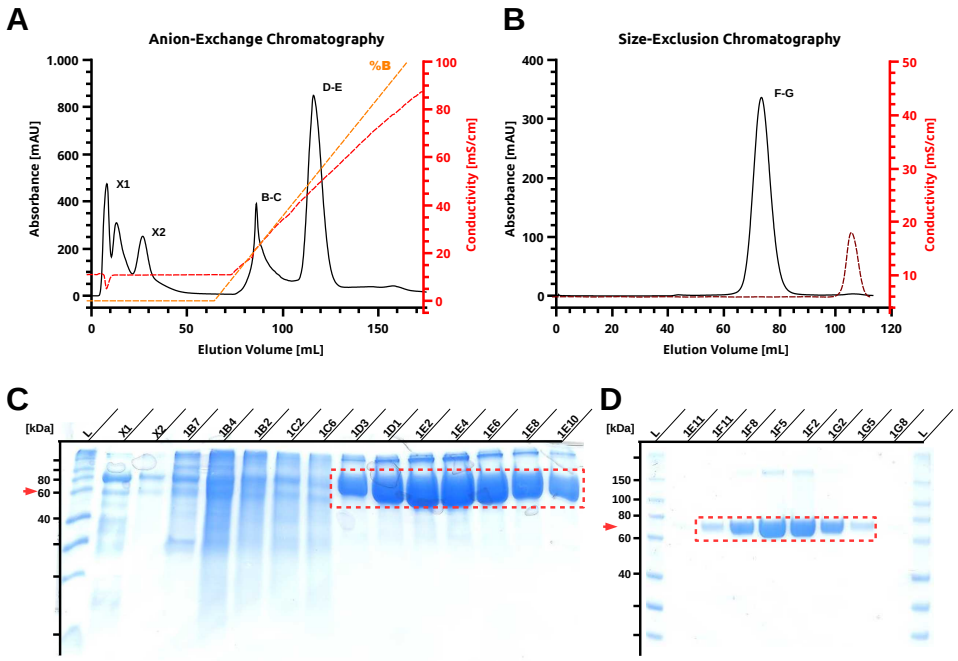

Figure 8.2: ***(A)*** *Chromatogram from anion-exchange chromatography with resuspended protein from AS precipitation on a HiTrap® Q HP 5mL column* ***(B)*** *Chromatogram from size-exclusion chromatography with fraction D-E (See A) on a Superdex S200 column* ***(C)*** *12% SDS-PAGE showing protein content of fractions from A* ***(D)*** *8% SDS-PAGE showing protein content of fractions from B; L= Protein Weight Marker (Roti®-Mark 10-150)*

To further investigate the purity of the sample an attempt was made to break up all possible oligomeric sub-species by a harsh treatment and verify the present state by a DLS denaturation curve (SPC, EMBL Hamburg). In figure 8.3-A it is shown how the purified sample reacts to a treatment with 32% β-ME & 200 mM DTT. In every case, the sample was incubated at 98°C for 30 min. Protein was loaded at 1.1 mg·mL^{-1} in 16 µL onto the gel. As it can be seen, a weak band is still present at a higher molecular weight above 150 kDa. This is in fact reasonable evidence that the band rather represents an impurity than a multimeric state of the AOX, since it cannot be broken up by

the harsh treatment. To conclusively identify all bands of the protein sample, mass spectrometry was conducted on bands A, B, C as shown in figure 8.3-C. The shown image of a SDS-gel has been amplified via contrast enhancement to visualize a visible band at molecular weights even higher than 150 kDa (band A). Bands B and C were exclusively identified as AOX by a *MASCOT* search [66], with a found sequence coverage of 67% and 92%, respectively. The band with the highest molecuar weight was identified as AOX with still 23% sequence coverage. In addition, some remains of fatty acid synthase A and B subunits were found with a 11% sequence coverage in this band. All proteins were identified for the host organism *H. polymorpha*. This disproofs the earlier proposed hypothesis that the bands at higher molecular weights might only be remains of other proteins. Indeed, at very high molecular weights impurities are found, but the vast majority of the contained protein is identified AOX at different oligomeric states. This is giving positive evidence for the successful induction of the P_{AOX} promoter and the expression of AOX under the applied growth conditions.

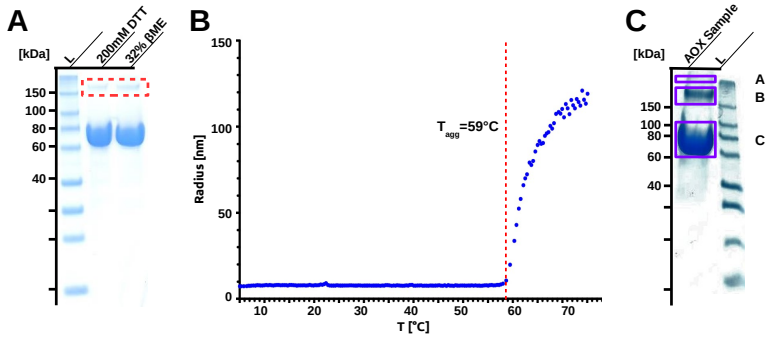

Figure 8.3: ***(A)*** *12% SDS-PAGE showing AOX samples treated with 32% β-ME or 200 mM DTT and heated to 98°C for 30 min* ***(B)*** *DLS denaturation curve of purified AOX protein* ***(C)*** *Contrast-enhanced 12% SDS-PAGE showing samples A-C used for mass spectrometric protein identification*

In addition to the presented results, the purified protein was tested for stability via a temperature ramp in DLS as depicted in figure 8.3-B. The protein solution was tested at a concentration of 3 mg·mL^{-1} in a heated cuvette. The protein seemed mostly monodisperse and stable up to a temperature of 59°C, which fits to the thermotolerant nature of the host organism and tops off the analysis of the AOX protein solution.

Spectrophotometric Determination of Equimolar Cofactor Binding

AOX requires bound FAD as a cofactor to be functional. In order to verify that mostly complete protein (protein + cofactor) has been purified an assay was developed to compare protein and cofactor concentration. FAD can be released from the protein by denaturation as it is not covalently bound. Hence, absorbance at 450 nm[1] will increase due to a change in the FAD local environment [252].
A small fraction (30 µL) of the protein solution (50 mM PP pH 7.5) with an approximate concentration of 10 mg·mL^{-1} got heated up to 98°C in a thermo mixer without shaking. Both absorbance at 280 nm and 450 nm were monitored successively with a spectrophotometer. It was expected to see slight

[1] The neutral form of FAD shows intrinsic fluorescence with an excitation maxima at 450 nm [252]

changes in the UV-absorbance of the aromatic amino acids due to denaturation. This process is obviously initiated by changes in the local environment of the residues. Usually, the absorbance at 280 nm will give more precise results with calculated extinction coefficients if the protein is denatured. Additionally, an increase in absorbance at 450 nm was expected due to freed FAD molecules. Table 8.1 shows the obtained values from the assay at the timepoints 0, 10, 20 and 30 min. In addition, figure 8.4-A and -B depict these values in a plot for a better comparison. As it can be seen in figure 8.4-B a lot of FAD seems to be released within the first 10 min of the assay and its concentration slowly reaches saturation after 30 min. In contrast, the absorbance at 280 nm (Fig. 8.4-A) seems to decrease between 20 and 30 min which is why a complete denaturation was assumed. Accurate calculation of the concentration was done with an extinction coefficients of 93500 $M \cdot cm^{-1}$ for AOX and 11300 Mcm^{-1} for FAD [62].

Table 8.1: *Data of FAD binding assay to determine protein/cofactor binding ratio; Conducted with a NanoDrop™ 1000 Spectrophotometer*

Time [min]	Abs. at 280 nm	Abs. at 450 nm
0	1.172 ± 0.013	0.051 ± 0.003
10	1.962 ± 0.098	0.219 ± 0.004
20	1.985 ± 0.049	0.241 ± 0.010
30	1.785 ± 0.102	0.245 ± 0.018

By using the *Lambert-Beer law*, a path length of 0.1 cm and approximate molar masses of 75 kDa (AOX) and 0.785 kDa (FAD), it was possible to calculate the molar- and mass concentration of both species in the solution. Since no crucial changes occurred after an incubation time of 10 min an averaged value of the timepoints 10-30 min was used for this calculation. As a result the concentrations were calculated to be 204±12.9μM (15.32 mg·mL^{-1}) and 208 ±13.8μM (0.16 mg·mL^{-1}) for AOX and FAD, respectively. The implementation of a SEC purification step ensured that no free FAD was present in the solution before the experiment started. In conclusion, it is apparent that both molecule species are present at equimolar concentrations and a 1:1 binding can be assumed. Hence, the purified protein seems to be fully assembled [48].

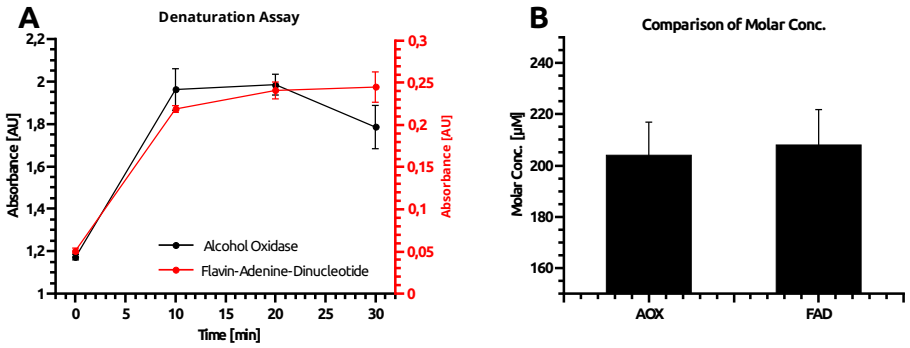

Figure 8.4: ***(A)*** *FAD binding assay to determine protein/cofactor binding ratio* ***(B)*** *Verification of equimolar concentrations of protein and cofactor*

8.2 *In vitro* Re-Crystallization of Alcohol Oxidase

As it was described in section 8.1 a purification procedure has been established for assembled alcohol oxidase. The purified protein was subsequently subjected to *in vitro* crystallization trials to obtain diffraction data for comparison with powder and SFX data.
The protein was concentrated to either 10, 20 or 30 mg·mL^{-1} in 50mM potassium phosphate pH 7.5 buffer by centrifugation. Afterwards one drop of each concentration was applied as *sitting drop* to the commercially available screens *MD Clear Strategy I*, *QIAGEN Classics I* and *QIAGEN pHClear I*. Drops were mixed 1:1 (0.5/0.5 µL) with the respective screen buffers. Immediately after preparation, plates were transferred to an automated storage device, maintained at 18°C at the SPC Facility of EMBL Hamburg. Whitelight and UV images were taken automatically on a daily basis.
Cubic protein crystals appeared already after 24-48 hrs in conditions with the highest protein concentration (30 mg·mL^{-1}), but only in combination with HEPES/ TRIS buffers at high pH between 7-8.5. Additionally, it was found that MPD was required as a compound to obtain protein crystals at all. Figure 8.5-A exemplarily depicts obtained crystals in 0.1 M HEPES at pH 7 and with 40% (v/v) MPD. The shown UV image verifies protein content within the crystals, whereas the yellow color of the crystals in the whitelight image indicates the presence of the cofactor FAD. In a second step, an attempt was made to optimize crystal growth conditions. A manual *hanging drop* screen was set up to cover various conditions in 0.1 M HEPES and 0.1 M TRIS buffer, ranging from pH 7 to 8 at MPD concentrations of 30-50% (v/v). An initial protein concentration of 30 mg·mL^{-1} has been used throughout the manual screen. Plates were stored again at 18°C. Crystal growth could be verified for

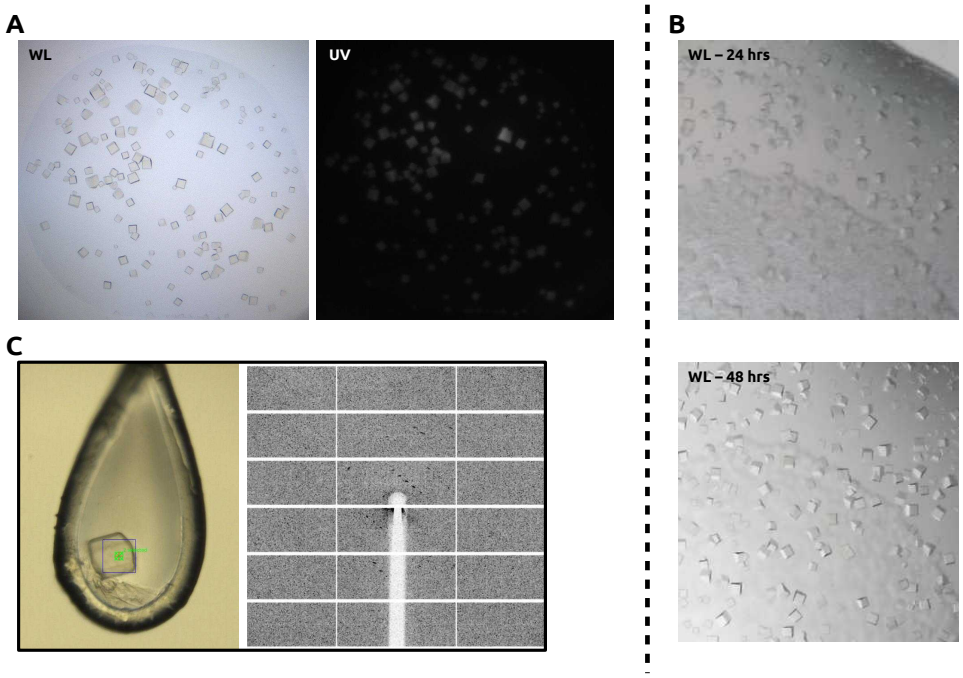

Figure 8.5: ***(A)*** *Crystallization trial at a conc. of 30 mg·mL^{-1} (1:1) in 0.1 M HEPES at pH 7, 40% (v/v) MPD after 24 hrs, whitelight and UV image* ***(B)*** *Crystals in manual screen at a protein conc. of 30 mg·mL^{-1} (1:1) in 0.1 M HEPES at pH 7, 40% (v/v) MPD after 24 and 48 hrs* ***(C)*** *Plunge-frozen AOX crystals in loop & low res. diffraction at 12.8 keV ($5.5 \cdot 10^9$ ph·s^{-1}, full flux, exp. time: 0.018 s)*

HEPES buffer at pH 7 & 7.5 and for TRIS buffer at pH 8 & 8.5. Furthermore all conditions contained between 31 to 46% MPD (v/v). Figure 8.5-B depicts crystals (0.1 M HEPES, pH 7, 40% MPD) grown after 24 hrs. Crystal size and shape remained stable between 24 and 48 hrs of storage.
Both, crystals from the initial screen and the manual screen were harvested with crystal loops and immediately cooled in liquid nitrogen to avoid evaporation of buffer and further damage. Finally, 8 crystals were tested for diffraction at beamline P14 (PETRAIII, DESY, Hamburg) at 12.8 keV (λ: 0.976 Å) and a flux of about $5.5 \cdot 10^9$ ph$\cdot s^{-1}$ (Exp. time: 0.018 s). The detector edge resolution was at 3.5 Å. None of the tested crystals yielded high resoluion diffraction as emphasized by diffraction data shown in figure 8.5-C. Bragg diffraction did not extend beyond 40-50 Åsimilarly to *in vivo* powder diffraction or SFX experiments.

8.3 Creation of Genetically-modified *H. polymorpha* Strains

Various target proteins were attempted to be cloned into *H.p.* specific shuttle vectors for genomic integration. At first, the vectors pHIPZ4 and pHIPX4[1] were altered by the incorporation of contiguous flanking regions at their MCS to fit the intended strategy in SLiCE-cloning (sec. 6.3). For that reason, original pHIPZ4 and pHIPX4 plasmids were digested with the restriction enzymes HindIII and SalI to linearize the plasmid without any inserts. Linearized plasmids were purified by agarose gel electrophoresis. Short ssDNA sequences were designed to yield a complementary dsDNA strand and which comprise 2x20 bp long flanking regions with the end of the P_{AOX} promoter and the beginning of the T_{PGK} terminator. A unique HindIII restriction site was introduced to allow for an easy digestion of the resulting vector for cloning attempts. Furthermore, a gene sequence has been introduced, that encodes for a heptameric spacer-peptide (GMDELYK) and an adjacent SKL tag. The ssDNA strands were annealed at a concentration of 50 pmol$\cdot\mu L^{-1}$. First, oligonucleotides were heated up to 98°C for 2 min. After that, the solution was gradually cooled down to room temperature over 15 min to allow for the formation of dsDNA. 100 ng of purified vector (in 1 µL) were mixed 1:500 with the diluted dsDNA insert. Cloning and bacterial transformation has been conducted afterwards as described (sec. 6.3). The resulting vectors were named pZ4S1 and pX4S1, respectively. Plasmids denoted with Z contain a Zeocin™ resistance cassette, whereas those with an X contain a cassette which counteracts on leucine (LEU2) auxotrophy by the expression of β-isopropylmalate dehydrogenase. Successful creation of the integration vectors has been verified by DNA sequencing.
Target proteins were chosen according to availability and based on an assessment of crystallizability. SLiCE-cloning was performed with PCR fragments comprising the protein's DNA sequence and additional flanking regions, which are compatible to the vector backbone. All used primers are listed in table A.4. Vectors (pX4S1 & pZ4S1) were linearized via the introduced unique HindIII restriction site and purified via agarose gel electrophoresis. Following the protocols described in section 6.3 complete plasmids for integration could be created. The plasmids carry the following DNA sequences in 5'-3' direction: *P_{AOX}, Gene of Interest (Target Protein) w/o stopcodon, Spacer, SKL tag with*

[1]Thankfully supplied by the work group of Ida. v. d. Klei (RUG Groningen, The Netherlands)

stopcodon, T_{PGK}. Linearization for integration of plasmids was done either at the unique restriction site of PflMI, StuI or SpHI, respectively, which are all located within the sequence of P_{AOX}. Due to the linearization within the promoter region, homologous recombination at the P_{AOX} genome locus is enabled. Plasmids were purified and subjected to yeast electroporation and plate-based clone selection (sec. 6.3). Usually colonies appeared after 3-5 days. Single colonies were re-streaked on selective plates and exploited for the purification of genomic DNA. PCR reactions on purified genomic DNA were used to verify stable integration of the plasmids at the location of the P_{AOX} promoter. Stable integration was achieved for target proteins listed in

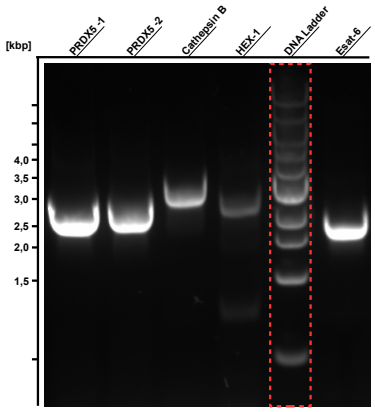

Figure 8.6: *PCR reactions on genomic DNA to verify pX4S1 integration containing different target proteins, Sizes are due to PCR on 1900 bp of genome/plasmid plus the size of the target gene; DNA Ladder: GeneRuler 1kb DNA Ladder (Thermo Fisher Scientific)*

Table 8.2: *Overview of successful cloning attempts for target proteins; 12DAC4: triple-deletion strain (ΔAOX, ΔDHAS, ΔCAT); MBP: maltose-binding protein, DAO: D-aspartate oxidase*

Target Protein	**Species of Origin**	**Created Vectors**	**Yeast Int. Strain**
HEX-1	*N. crassa*	pZ4S1, pX4S1	12DAC4
Cathepsin B	*T. brucei*	pZ4S1, pX4S1	12DAC4
PRDX5	*H. polymorpha*	pZ4S1, pX4S1	12DAC4
Esat-6	*S. agalactiae*	pZ4S1, pX4S1	12DAC4
MBP	*E. coli*	pZ4S1, pX4S1	-
Luciferase	*P. pyralis* (Firefly)	pZ4S1, pX4S1	-
DAO	*H. polymorpha*	pZ4S1, pX4S1	-

table 8.2. Exemplary PCR results with primers from table A.4 for successfully obtained integration clones (all pX4S1-based) are shown in figure 8.6. The figure shows two obtained clones for the protein PRDX5 (510 bp) and one for each of the proteins Cathepsin B (966 bp), HEX-1 (533 bp) and Esat-6 (285 bp). The final fragment length results from reaction design that involved a forward primer, starting on a genomic upstream location and a reverse primer starting on the terminator sequence of the vector. This results in a total fragment length of 1900 bp plus length of the DNA sequence of a target protein.
Two major problems could be identified during the cloning process. Firstly, the P_{AOX} promoter seems to exhibit a strong secondary structure, which decreases

the success rate of cloning in its vicinity. Secondly, transformation of *H.p.* cells was accompanied by a high rate of false-positive clones. Original strains as well as competent *H.p.* cells were unable to grow on the respective selective media. Hence, the vector DNA seems to integrate at different locations in the genome. Attempts to optimize the utilized protocols resulted in the yield of strains with a correctly integrated vector at a low rate of about 1-5%.

8.4 Heterologous Expression & Peroxisomal Targeting of Proteins in the Yeast *S. cerevisiae* & *H. polymorpha*

This section exemplarily describes the verification of heterologous expression of target proteins and their peroxisomal targeting in different budding yeast species. Firstly, the expression and peroxisomal targeting of *H.p.* alcohol oxidase (AOX) is shown in *S. cerevisiae* cells under oleate-inducing conditions. Secondly, the expression of *N.c.* HEX-1 is verfied for both, *H. polymorpha* and *S. cerevisiae* cells under methanol- or oleate-inducing conditions, respectively. AOX crystals appear for several mutant strains of *H. polymorpha* that were altered with regards to peroxisomal fission (ΔPEX11) or peroxisomal import (ΔPEX5). Especially the occurrence of cytosolic AOX crystals in import-deficient mutants should receive special attention since it implies a strong tendency of AOX to form *in vivo* crystals under unambiguously different conditions. On the other hand, HEX-1 is the protein thate makes up the crystalline core of Woronin bodies in filamentous fungi like *Neurospora crassa*. Likewise to AOX, the protein shows a natural tendency to crystallize. Consequently, AOX and HEX-1 has been chosen as representative target proteins among potential candidates in heterologous expression trials to achieve intraperoxisomal crystallization. Although similarities in peroxisome biogenesis cannot be denied, and both species *S. cerevisiae* and *H. polymorpha* belong to the class of *Saccharomycetes*, very distinctive features renders them easily distinguishable. For example, differences are found with regards to size, optimal growth temperature and carbon source utilization. These phenotypic features are based on huge differences in the genetic makeup with protein families unique to every species [193].
The full-length gene of *H. p* AOX has been cloned into the multi-copy vector *pMSC13*, which utilizes the strong, oleate-inducible promoter P_{FOX3} (ch. 2). In fact, the promoter represents one of the strongest expression promoters known for *S. cerevisiae* to date [41, 83] and is related to the expression of 3-ketoacyl-CoA-thiolase (Fox3p, involved in fatty acid β-oxidation) in wildtype strains. For expression trials the *S.c.* strain UTL7A has been used. Section 6.2 is giving a detailed overview of general vector transformation and batch growth of mutant and control strains[1] at 30°C. As described elsewhere, oleate induces peroxisome proliferation in *S. cerevisiae* and thus leads to the resemblance of similar cellular conditions as in methanol-induced *H. polymorpha* cells. Alcohol oxidase has been expressed with two different peroxisomal PTSI-import signals. Both, the classical -ARF, as well as the -SKL signal are C-terminal tripeptides usually sufficient for peroxisomal import in various species [102].

[1] Experiments have been conducted together with Thomas Walter in the labs of Prof. Dr. R. Erdmann (Ruhr-Universität Bochum, Germany); The creation of *S.c.* expression vectors was done by T.W.

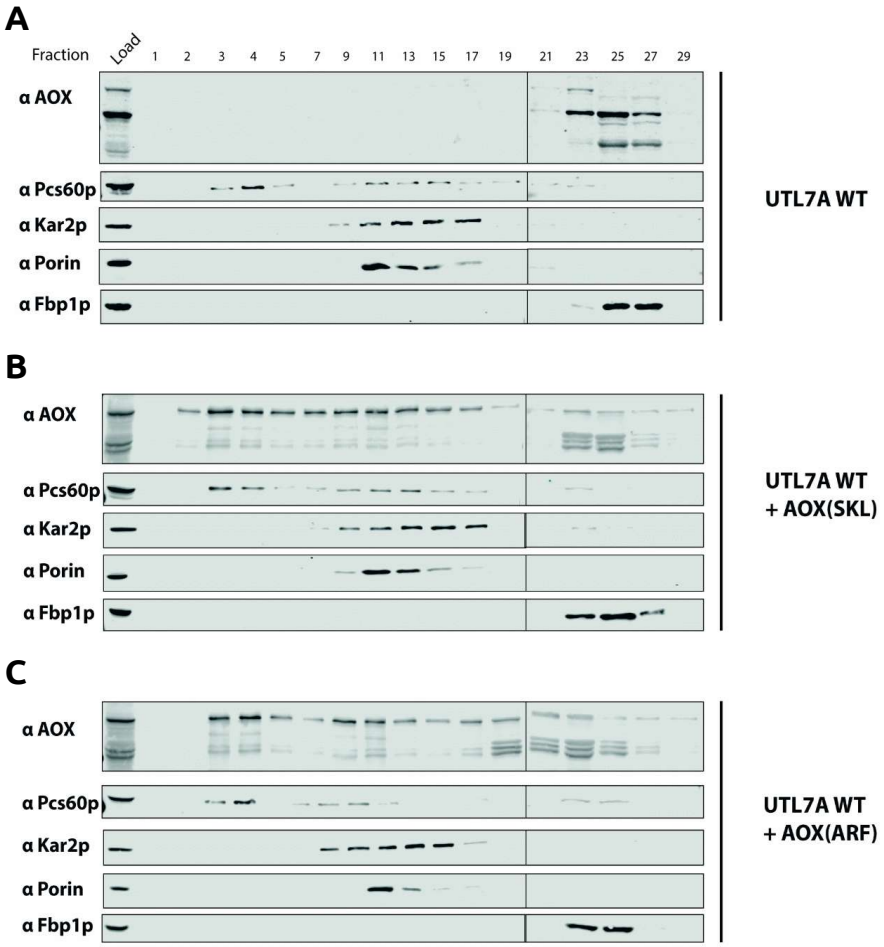

Figure 8.7: *Western blot with organellar marker antibodies to verify peroxisomal targeting of AOX in fractions of sucrose-density gradient* ***(A)*** *Wildtype control* ***(B)*** *Expression of AOX with a C-terminal -SKL tag* ***(C)*** *Expression of AOX with a C-terminal -ARF tag; (Polyclonal antibodies) αPcs60p: peroxisomes, αKar2p: endoplasmic reticulum, αPorin: mitochondria, αFbp1p: cytosol αAOX: alcohol oxidase*

Protein expression has been allowed under oleate conditions for about 16 hrs. After that, cells were harvested and a organellar pellet was obtained according to section 6.5. Subsequently, organelles were separated by sucrose-density gradient centrifugation. The final gradients were divided into 1 mL fractions which could be analyzed by western blotting. The obtained results are shown in figure 8.7-A-C for a wildtype control and both targeting signal trials.
Polyclonal antibodies were used as specific primary markers to identify organelle separation within gradients. αPcs60p is a typical peroxisome marker, whereas αKar2p, αPorin and αFbp1p are antibody markers for the ER, mitochondria and cytosol, respectively. Firstly, a comparison of the control blots in A to the AOX expression blots in B and C reveals successful expression of the target protein. Secondly, the AOX distribution patterns indicates for both mutants a co-localization with the peroxisomal marker protein Pcs60p. This implies a direct targeting of *H.p.* AOX into *S.c.* peroxisomes. A minor epitope binding of the αAOX antibody[1] can be found within the cytosolic fraction for control cells (fig. 8.7-A), but not in any other fractions, indicating either a problem

[1]Thankfully supplied by the work group of Ida. v. d. Klei (RUG Groningen, The Netherlands)

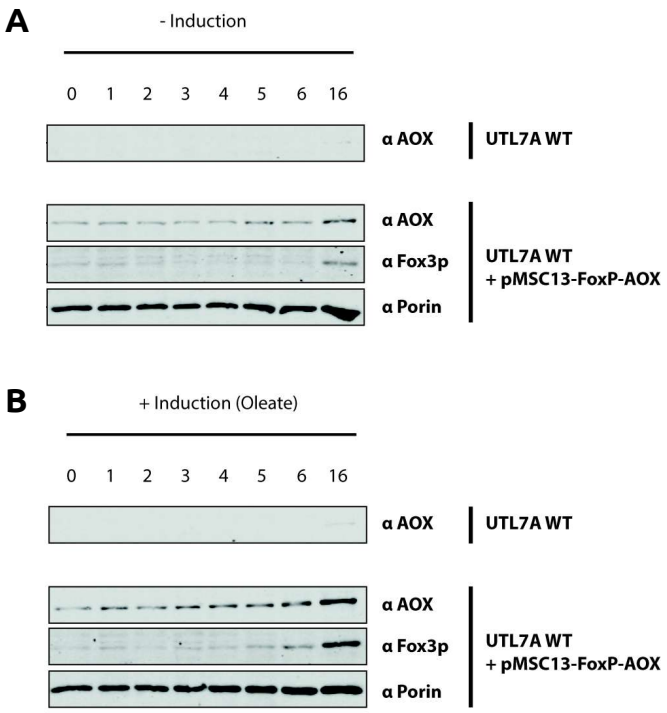

Figure 8.8: *Western blot with marker antibodies to verify quantitiative expression of AOX upon oleate induction over time (hrs);* ***(A)*** *Without induction* ***(B)*** *With induction; (Polyclonal antibodies) αPcs60p: peroxisomes, αFox3p: 3-ketoacyl-CoA-thiolase, αPorin: mitochondria*

during blotting or unspecific binding of the used antibody. Still, the presence of AOX epitopes in specific organelles can be precluded. Intriguingly, binding of the αAOX antibody is seen throughout different fractions upon expression of AOX for both tag variants. Since the obtained signals always co-localize with the Pcs60p marker the targeting of AOX to other organelles seems unlikely. It might rather be assumed, that pre-mature peroxisomes are scattered throughout the density gradient. The blots representative for the strains carrying the AOX gene, indicate a minor degradation of AOX protein within *S.c.* cells, both in the peroxisomal, but also in the cytosolic fraction by means of several degradation bands. This result might indicate an incomplete folding of the alcohol oxidase e.g. due to a lack of the cofactor FAD, or incomplete expression. On the other hand, a strong band at a molecular weight of approx. 75 kDa seems indicative for protein expressed at full-length. A comparison between blots B (-SKL) and C (-ARF) cannot substantiate differences in targeting efficiency with regards to the used signaling peptides.
In an additional experiment the induction of AOX expression via oleate has been investigated in whole cell lysates. Figure 8.8 shows a comparison of non-inducing to inducing conditions among the strain carrying a -SKL-tagged AOX and a wildtype UTL7A strain. While the wildtype strain does not exhibit a response to the αAOX antibody at any timepoint and condition, the mutant strain shows a clear induction of AOX and the related Fox3p protein, both under the control of the P_{FOX3} promoter. The mitochondrial marker Porin is visible with similar characteristics at any timepoint. Gels were always loaded with the same concentration of total protein, thus indicating a quantitative increase of intracellular AOX concentration by oleate induction (cp. fig. A.1). The highest concentration of target protein is apparently achieved when cellular

growth is allowed until stationary phase (16 hrs). Peroxisomal targeting via the PTSI pathway and the -SKL tripeptide is a widely applied tool in peroxisomal research, which has proven to function with proteins at different sizes and oligomeric states [121]. The conducted experiments verify that both targeting signals (-SKL,-ARF) can be used for successful targeting of heterologously expressed *H.p.* AOX into peroxisomes in *S. cerevisiae* batch cultures. Furthermore quantitiative expression of the target protein is achieved during growth for 16 hrs.

The full-length gene of *N.c.* HEX-1 has been cloned either into the multi-copy vectors pCW6 (P_{PGK}) and pCW12 (P_{FOX3}), or the integration vector pHIPX4 (P_{AOX}) for the expression in *S. cerevisiae* or *H. polymorpha*, respectively. In all three cases a C-terminal -SKL tag has been attached to the protein for peroxisomal targeting (see sec. 8.3). The strains BY4742 (*S.c.*) and NCYC495-12DAC4 (*H.p.*) have been used for expression trials. Likewise to the above experiments the related methods are presented in section 6.3. Cells at OD 2 were subjected to TCA precipitation, SDS-PAGE and western blotting with specific polyclonal antibodies. For comparison *S. cerevisiae* cells were grown either on glucose-, Rytka-, or pure oleate-containing medium, whereas *H. polymorpha* was grown on a combination of methanol and glycerol as sole carbon sources. As it can be seen in the upper panel of figure 8.9, a constitutive expression of HEX-1 is found for the P_{PGK} promoter under glucose-, oleate- or Rytka- conditions in *S. cerevisiae*. On the contrary, the P_{FOX3} promoter only allows expression of HEX-1 under oleate-inducing conditions. After 16 hrs of growth the expression under pure oleate conditions is more pronounced as for the Rytka- condition. This is correlated with the earlier expression of Fox3p, simply indicating a stronger initial derepression of the P_{FOX3} promoter. The advantage of the Rytka medium is a smooth transition to inducing conditions under oleate, but seems to differ from the direct derepression of the P_{AOX} promoter in *H. polymorpha* under methanol/glycerol conditions [108]. Usage of the P_{PGK} promoter might be advantagous in terms of achieving higher peroxisomal target

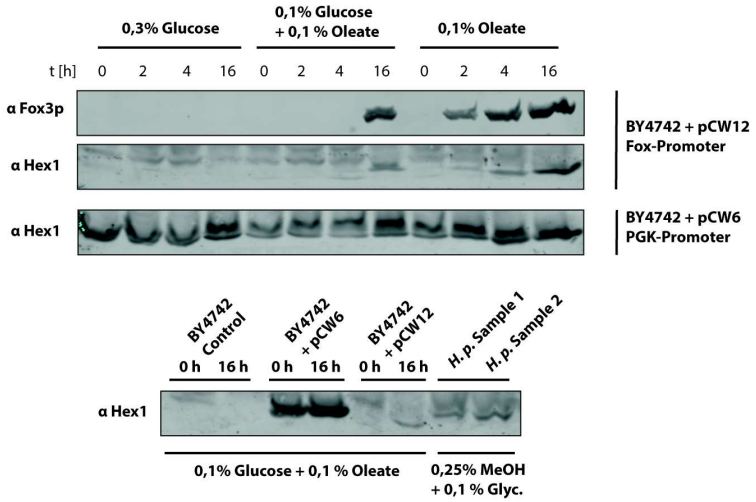

Figure 8.9: *Western blot with marker antibodies to verify quantitiative expression of HEX-1 upon oleate- or methanol induction; (Polyclonal antibodies) αFox3p: 3-ketoacyl-CoA-thiolase, αPorin: mitochondria*

protein concentrations, as peroxisome proliferation not necessarily needs to be induced for cell growth and heterologous protein expression. Still, this a purely tendencious assumption, since no studies exist to date that have tested peroxisomal import capacities under conditions that do not favor the expression of the peroxisomal proteome.
The lower panel of 8.9 summarizes the results of heterologous HEX-1 expression in two host organisms. As discussed above, a strong expression of *N.c.* HEX-1 is found for *S. cerevisiae* cells under oleate conditions, whereas the wildtype control cells do not show any antibody binding.

9 | Probing Target Crystallization in the Peroxisomal Crystal Factory

The analysis of nano-sized, *in vivo* crystals in yeast cells is a difficult task. Not only is the cellular environment restricting the access, but also the sheer size of the crystals exacerbates the analysis. Originally intended to supply thousands of nano-crystals for serial femtosecond X-ray crystallography (see sec. 1.4), it is difficult to conduct preliminary studies on yeast cells with regards to their crystal content. SFX with small nano-crystals is only possible at ultra-brilliant X-ray light sources like the free-electron lasers LCLS (SLAC Stanford, USA), SACLA (SPring-8, Japan) or the recently opened European XFEL (Hamburg, Germany). A preliminary investigation of the crystal content of yeast cells needs to hark back on newly developed approaches (see ch. 11 & 12) or more established methods like *in cellulo* powder diffraction, or electron microscopy (see sec. 10.3).

9.1 Synchrotron-based Powder Diffraction for *In Cellulo* Nano-Crystal Analysis

Synchrotron-based powder diffraction allows for an easy verification of crystalline material, even in a nano-sized *in cellulo* environment, for whom a low signal-to-background ratio prevails. Furthermore simple characterization of crystals can be conducted by means of a Williamson-Hall analysis, to characterize domain size and strain. This chapter exemplary summarizes efforts to conduct a powder diffraction analysis on the yeast *H. polymorpha*, containing alcohol oxidase crystals within peroxisomes (sec. 9.1). The analysis is conducted with regards to different aspects of crystal growth, e.g. cultivation temperature or time.
Additionally, an overview is given of yeast strains, that were tested for possible *in vivo* crystal formation in peroxisomes (sec. 9.2). The samples comprise *H. polymorpha* strains for heterologous expression of various target proteins, whose creation is described in chapter 8. Likewise, *S. cerevisiae* strains are listed, which were created together with collaborators (cp. sec. 8.4). All diffraction experiments were conducted on the micro-focus beamline P14 (macromolecular crystallography), located at the PETRAIII storage ring (DESY, Germany).

9.1.1 Optimal X-ray Powder Diffraction From Yeast *In Vivo* Samples

For powder diffraction experiments in tramsmission geometry, it is easy to calculate the optimum sample thickness. The attenuation of X-rays follows

a simple LAMBERT-BEER law and is therefore exponentially increasing. On the other hand a linear dependency exists between THOMSON scattering and sample thickness (equiv. crystal content). Thus the maximum of scattered signal relative to beam attentuation is obtained at a sample thickness of $^1/_\mu$, with μ being the attenuation coefficient of the investigated material. By using the CROMER-LIBERMAN algorithm [23], integrated in an available *Python* script [171], it has been possible to estimate the attenuation coefficient of yeast powder in a capillary.

Table 9.1: *Anomalous scattering coefficients & absorption of the atomic composition $CH_{1.8}O_{0.5}N_{0.2}$ at a density of 1.09 $^g/_{cm^3}$ for an X-ray beam at 12.8 keV*

Element	**Ø Comp.**	**f'**	**f"**	**μ [barn]**
C	1.0	0.007	0.003	17.64
H	1.8	0.000	0.000	0.00
O	0.5	0.021	0.012	32.57
N	0.2	0.013	0.007	7.15

The calculation considered a yeast-specific average atomic composition of $CH_{1.8}O_{0.5}N_{0.2}$ [67] and average mass density of about 1.09 $^g/_{cm^3}$ [8]. If a high energy beam at 12.8 keV (0.97 Å) is assumed, an approximate attenuation coefficient of 1.53 cm^{-1} is obtained. Assuming the usage of large *MiTeGen MicroRT*^{TM} tubes with 2 mm diameter, a beam transmission of 73.67% is retrieved. For the considered elements different anomalous scattering coefficients were calculated, which are listed in table 9.1. The variation of μ with the beam energy is shown in an additional plot in figure A.2.
Obviously the attenuation length $^1/_\mu$ approximates to 6.5 mm to reach a transmission of $^1/_e$ ($\approx 37\%$). Thus for the investigated *in vivo* sample, an increased sample thickness could still be beneficial for the detection of DEBYE-SCHERRER rings. On the contrary samples will tend to reabsorb X-rays at low 2Θ angles and effectively reduce the visibility of DEBYE-SCHERRER diffraction. Capillaries with increased diameter were not conveniently available for measurements and hence the 2 mm diameter can be considered as the best achievable parameter for these studies. Fig. 9.1-A shows an exemplary powder diffraction pattern recorded from a cell pellet composed of wildtype yeast cells 24 hrs after

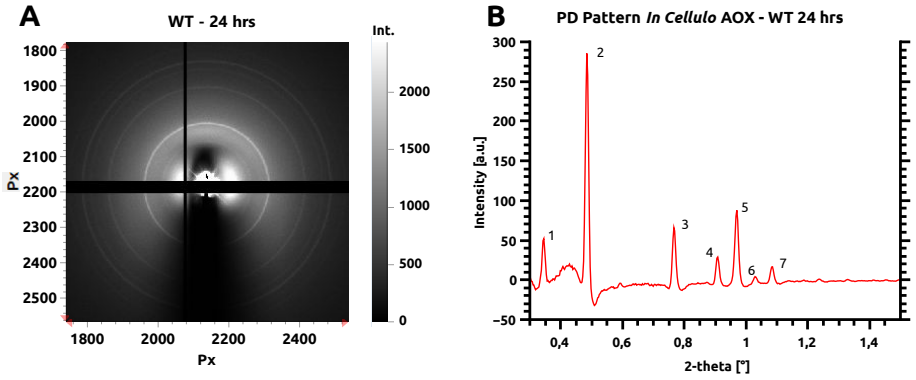

Figure 9.1: ***(A)*** *2D powder diffraction pattern and the* ***(B)*** *corresponding post-processed 1D pattern from wildtype H. polymorpha yeast cells with inducded AOX crystal formation after 24hrs of growth; Data was obtained according to sec. 4.1 with setup parameters listed in table 9.2, exposure time: 1 s*

P_{AOX} induction (P14, DESY). A detailed description of sample production is given in section 6.2. Setup-specific parameters for data collection are listed in table 9.2. Fig. 9.1-B shows an azimuthal sum of the data. The data has been additionally processed in terms of a rolling ball baseline subtraction (see sec. 4.1).

Table 9.2: *Setup Parameters for Powder Diffraction Experiments at beamline P14*

Parameter	**Value**
Photon Energy	12.8 keV (0.97Å)
Total Beam Flux	$1.2 \cdot 10^{12} ph \cdot s^{-1}$
Detector Distance	2000 mm
Beamstop Distance	30 mm
Oscillation Range	90°
Slit Size	$150 \mu m$
Capillary Diam.	2 mm

It is an apparent feature of methanol-induced wildtype yeast cells that smooth DEBYE-SCHERRER rings appear at low resolution of about 160-50 Å, corresponding to 2Θ angles of 0.3°to 1.2°. These rings appear due to the presence of nano-sized alcohol oxidase (AOX) crystals that are enclosed in peroxisomes (see sec. 2.2). Their appearance is related to the extremely increased intensity contribution of the Lorentz factor at low diffraction angles, that follows the relation $1/(sin(2\Theta)sin(\Theta)$ [19, 26]. The smoothness of the rings indicate a perfectly random orientation of specimen, as it can be expected for *in cellulo* crystals. Since the diffracted signal on the detector obeys *Poisson* statistics and the related SNR increases only with $N/\sqrt{N}$ (N=photon count), an increase of intensity (or likewise a longer exposure time) will not be sufficient to visualize atomic resolution diffraction in this setup. Still, since only extremely low resolution rings are visible, the repetition of a powder experiment with an increased exposure time might still be realistic to visualize DEBYE-SCHERRER rings at slightly higher resolution. A prerequisite is of course, that the increased exposure time does not exhaust the dynamic range of the detector. Radiation damage will most likely not be a cause of signal deterioration at higher exposure times since only extremely low 2Θ angles are being investigated. AOX crystal-containing wildtype cells were subjected to a second powder diffraction experiment at an increased exposure time of 15 s. All other measurement-related parameters were kept fixed. Instead, 10 consecutive measurements were conducted on

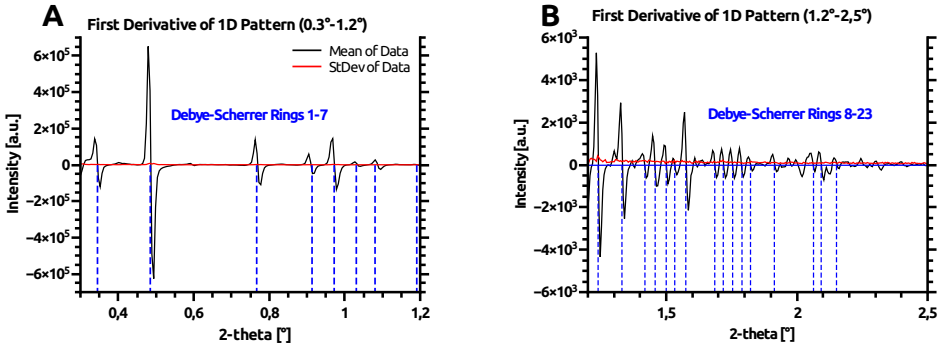

Figure 9.2: *First derivative of a 1D powder diffraction pattern for* 2Θ *angles between* ***(A)*** *0.3°to 1.2°and* ***(B)*** *1.2°to 2.5°; Setup parameters are listed in table 9.2, exposure time: 15 s*

different areas of the cell pellet. Data was processed in terms of azimuthal integration and background subtraction. Afterwards 2Θ-dependent first derivative mean and standard deviation were calculated to simplify the search for high resolution peaks.
Fig. 9.2-A/B show the results of a manual analysis of 2Θ ranges from 0.3°to 1.2°, showing the already known diffraction peaks, and from 1.2°to 2.5°, revealing additional 17 DEBYE-SCHERRER peak positions (blue dashed lines). A peak position is identified by a downward-going zero crossing. The red curve represents the corresponding standard deviation and can be used to estimate false zero-crossing positions. In an automated search a simple threshold would be used.
It is obvious from a comparison with control data (*H.p.*-12DAC4 cells, fig. A.3), that the newly identified peak positions do not arise from artifacts during processing and indeed must be related to crystal diffraction. The peak positions are listed in table 9.3. Peak intensities have been determined from GAUSSIAN fits of the corresponding peak areas. The extraction of peak intensities was increasingly problematic for higher 2Θ angles. Thus data has been processed three times with the first 7, 12, and 23 peaks by using the *Monte Carlo* indexing software *McMaille* [98]. Results were only obtained for calculations considering 7 and 12 peaks, respectively. For both calculations a body(I)-centered cubic lattice with equilateral unit cell dimensions of $\sim$ 228 Å is given as the most plausible solution. However a consideration of more than 12 peaks did not yield any solution. hkl indices are assigned according to the systematic absence of odd h+k+l reflections in the presumed Bravais lattice. Intriguingly some even reflections like (2 1 1), (2 2 0), (2 2 2), (3 3 2) and (4 4 0) are missing as well. Their absence could be indicative for anisotropic disorder in the crystals (cp. fig. A.4). A further increase of the exposure time beyond 15 s did not reveal any new features of the diffraction pattern (data not shown). Furthermore it has been tried to increase the detectable diffraction signal by different growth conditions, to expedite crystal growth. Unfortunately these efforts did not yield any better results.
Therefore it seems plausible to suspect, that AOX crystals in yeast cells are imperfect crystals and show only very weak high resolution diffraction, as it

Table 9.3: 2Θ *angles, d-spacing and millerian indices for AOX in cellulo crystal reflections in a 1D powder patter from H. polymorpha wildtype cells, indexed as a body(I)-centered cubic lattice*

Nb.	**2Θ** [°]	**d-Spacing** [Å]	**hkl Refl.**
1	0.35	163	(1 1 0)
2	0.49	115	(2 0 0)
3	0.77	73	(3 1 0)
4	0.91	61	(3 2 1)
5	0.97	57	(4 0 0)
6	1.03	54	(3 3 0) / (4 1 1)
7	1.09	51	(4 2 0)
8	1.19	47	(4 2 2)
9	1.24	45	(4 3 1) / (5 1 0)
10	1.33	42	(5 2 1)
11	1.42	40	(4 3 3) / (5 3 0)
12	1.46	38	(4 4 2) / (6 0 0)

has already been shown by Van der Klei *et al.* [33] for *in vitro* crystals of AOX. This hypothesis is in fact in agreement with the biological functioning of the AOX crystal, as being a solid-state catalyst crystal. Consequently diffraction experiments under optimized conditions at a free-electron laser X-ray source might provide a solution to obtain higher resolution data (see ch. 10).

9.1.2 Estimation of the Minimal Detectable Volume Fraction of Crystalline Material in a Yeast Powder

It is possible to approximate the minimum detectable crystalline volume fraction, composed of AOX crystals in yeast powder by a simple addition method. Purely amorphous phase is mixed with crystal-containing phase. The crystalline phase consists of *H. polymorpha* wildtype cells, which were grown according to sec. 6.2 for 24 hrs on methanol to have a strong induction of crystal formation. The amorphous sample consists of *H. polymorpha*-12DAC4 cells, that lack the ability to form *in vivo* crystals, but otherwise resemble P_{AOX}-specific cellular structures, when grown on a combination of methanol/glycerol. Amorphous and crystalline phase were mixed at different ratios to yield 50, 20, 10, 5 and 2% of crystalline phase. Powder samples were prepared according to section 4.1 and subjected to measurements at beamline P14. The obtained data has been corrected for small differences in sample packing.
The intensity of the 100%-peak (strongest reflection) is monitored in terms of its integrated peak areas (peak fits shown in fig. A.5-A-F). The peaks are fitted by assuming the shape of a GAUSSIAN distribution function. From the above procedure a linear fit with negative slope is yielded, that indicates no strong bias of the diffracted intensities (fig. 9.3-B). From the x-intercept a minimal detectable crystalline fraction of 3.3% can be inferred. Indeed, a dilution at 2% crystalline volume fraction does not show any visible peak (fig. A.5-F).

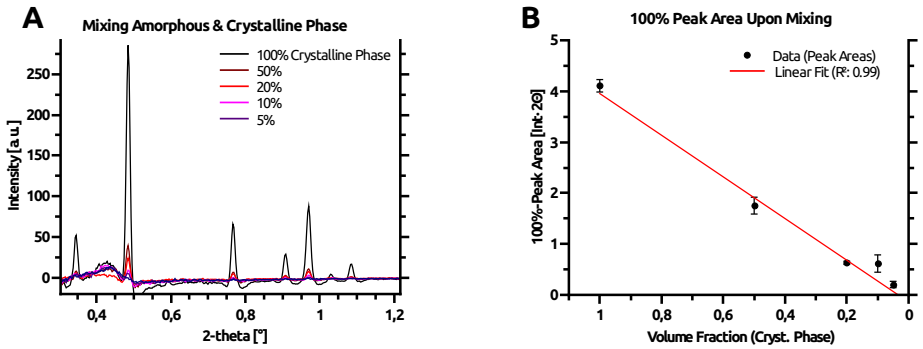

Figure 9.3: ***(A)*** *1D powder pattern of mixed phases: wildtype cells (crystalline phase) / control cells (amorphous phase)* ***(B)*** *Integral breadth decrease of 100% peak (peak 2) upon mixing with amorphous phase; Setup parameters are listed in table 9.2, exposure time: 1 s*

Section 11.2 exemplary describes the development of the intracellular peroxisomal volume upon induction of P_{AOX} metabolism. Further the approximate volume of a *H.p.* cell can as well be inferred from microscopic images. Under the assumption of a sphere-like cell shape an approximate volume of 3.450 ± 1.838 μm^3 can be calculated for a cell culture 24 hrs after P_{AOX}-induction. The total peroxisome-related volume at this timepoint approximates to 0.038 ± 0.003 μm^3. It is known from electron microscopy (cp. fig. 10.7) that alcohol oxidase crystals occupy most of the peroxisomal volume. Their volume fraction can hence be estimated by equating peroxisome volume and crystals size. Taking

the number of peroxisomes per cell into account, only about 1.2% of the cell volume is occupied by crystals in wildtype cells after 24 hrs. Considering the above calculations the minimal detectable volume fraction of *in cellulo* AOX crystals approximates to 0.40‰. If a higher frequency of peroxisomes per cell is assumed, as found for cells in late stationary phase a volume fraction of 1.15‰ can be inferred. Therefore the detection of *in cellulo* crystals can be rated as highly sensitive under the investigated conditions [99, 130].

9.1.3 Williamson-Hall Analysis to Characterize Domain Size & Strain

Powder diffraction allows for an in-depth analysis of the properties of nano-crystals, like domain size and strain via the Williamson-Hall plot [112, 141, 246, 263]. A prerequisite is a well-aligned diffractometer or beamline with little beam divergence [27]. Domain size and strain are both determined from the peak broadening across several reflections. The measured peak broadening is a convolution of instrument-based and sample-based broadening effects. The instrument broadening may arise from an imperfect beam in terms of wavelength dispersion and divergence, or the active diffracting volume within the sample. On the sample part, especially the domain size and crystal-lattice distortion will affect broadening. The technique can supply useful information about *in vivo* grown and *in cellulo* located nano-crystals prior to sophisticated experiments. More details about this methods are presented in section 1.3.
The beamline P14 provides helpful measures for a peak analysis with an approximate beam divergence of only $20\mu rad$[1]. Diffraction data from *in vivo* nano-crystals was obtained similarly to subsection 9.1.1. Here powder diffraction data from *H. polymorpha* wildtype cells, grown for 50 hrs in fed-batch mode, is exemplary used for analysis.
A sample-to-detector distance of 2000 mm and a slit size of $150 \cdot 150\mu m$ yield an instrument-based 2Θ line broadening effect of $5.4 \cdot 10^{-3} 2\Theta$. It is calculated via simple geometrical considerations according to equation 9.1.

$$\Theta = 2arctan\left(\frac{w_2 - w_1}{2d}\right) \quad \leftrightarrow \quad w_2 = 2 \cdot 2m \cdot tan\left(\frac{20\mu rad}{2}\right) + 150 \cdot 10^{-6}m \tag{9.1}$$

Here w_1 and w_2 are the width of the beam waist before and after the distance d. The geometrical spread can be converted via the detector's pixel size of 0.172 mm to give the final 2Θ value. The instrumental broadening will non-linearly increase with 2θ angles. Thus a more detailed analysis would require determination of the exact instrumental broadening at higher angles [50, 106]. Since here only extremely low 2Θ angles are analyzed the increasing broadening is neglected. The strongest diffraction peaks 1-5 and 7 (cp. fig. 9.1-B) have been used for data evaluation, by determining the FWHM and exact peak positions. Afterwards the theoretical instrumental broadening has been subtracted from all peaks. The obtained Williamson-Hall plot is shown in figure 9.2. The related data is listed in table 9.4. From the y-intercept of a

[1] At the time of the experiments (2015) beam divergence at P14 was anisotropic ($20\mu rad$ horizontally & $10\mu rad$ vertically); Information obtained from beamline scientist Gleb Bourenkov prior to experiments

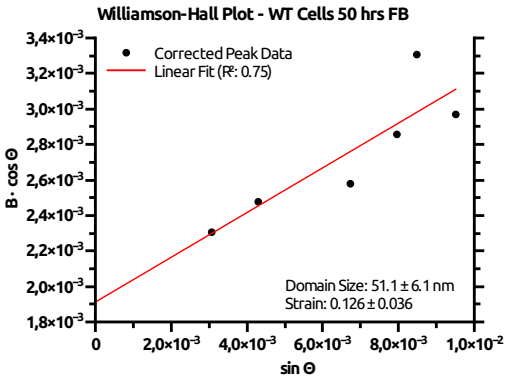

Figure 9.4: WILLIAMSON-HALL *plot of H. polymorpha wildtype cells grown under* P_{AOX} *inducing conditions in fed-batch mode for 50 hrs; Experimental parameters are listed in table 9.2, exposure time: 15 s; Data listed in table 9.4*

linear fit to the data points an average diffracting domain diameter of $51 \pm 7nm$ can be estimated, if the SCHERRER constant (K) is assumed to equal 1 (e.g. for spherical crystals). This indeed very low value is indicative for a highly imperfect crystal with domain irregularities at the order of about 2 unit cells. Due to the systematic absence of odd reflections it remains difficult to estimate a directional preference of disorder. But as mentioned above, the complete absence of certain even reflections might already indicate strong anisotropic disorder. The determined domain size is not necessarily correlated with the particle size estimated from electron microscopy (see sec. 10.3). It is rather a readout to estimate the average distance of crystallographically ordered, and diffracting domains in a paracrystalline material. If defects are being too extended, it is easier to consider the domain size as the size of a diffracting crystal, embedded with other such domains in a polycrystalline material [50]. The positive slope of the linear fit indicates the presence of microstrain at a magnitude of about 0.128. Intensive microstrain will dominate the peak broadening and thus can lead to a false approximation of the domain size [50]. Consequently it can be assumed, that the diffracting domain size is in fact larger. Both readouts are supporting the hypothesis of disordered AOX crystals [33] (subsec. 9.1.1).

Table 9.4: WILLIAMSON-HALL *plot data; Fitted and corrected FWHM (2Θ) of incorporated hkl reflections;* $B = {}^{FWHM}/_{2}$

hkl Refl.	**FWHM** $[10^{-3}\ °]$	$\boldsymbol{sin\Theta}$ $[10^{-3}]$	$\boldsymbol{B \cdot cos\Theta}$ $[10^{-3}]$
(1 1 0)	4.617	3.054	2.309
(2 0 0)	4.950	4.276	2.475
(3 1 0)	5.162	6.719	2.581
(3 2 1)	5.724	7.941	2.862
(4 0 0)	6.613	8.465	3.306
(4 2 0)	5.936	9.512	2.968

Comparison of AOX Crystal Quality in Wildtype & ΔPEX11 Cells

To emphasize the capabilities of such a simple analysis method the crystalline content of *H. polymorpha* wildtype and ΔPEX11 cells were compared 6 and

24 hrs after P_{AOX} induction (YPM media, 0.5% methanol). A detailed description of the differences between both cell lines can be found in chapter 2. Powder patterns were again obtained similarly to subsections 9.1.1 and 9.1.3 by using the setup-specific parameters listed in table 9.2. The resulting WILLIAMSON-HALL plots are depicted in figure 9.5-A/B. The related data is listed in table A.1.
The quality of the linear fits is poor in both cases for data obtained 6 hrs after induction. In contrast diffraction data after 24 hrs gives reasonably good fits for both wildtype and ΔPEX11 cells, respectively. The RSS (residual sum of squares) of linear fits to data from 6, 8, 10 and 24 hrs is indeed constantly decreasing with time (see fig. A.7). This trend is indicating a homogeniztaion process along different crystallographic directions during cell growth / crystal maintenance. It could also indicate enzymatic inactivity that is associated with an ubiquitous conformation.
The plots indicate an increase in diffracting domain size for wildtype and ΔPEX11 cells, from 6 to 24 hrs of growth. Wildtype-specific domain size increases from 22.3 ± 4.0 nm to 46.3 ± 3.7 nm. Similarly one can find an increase from 24.8 ± 11.0 nm to 118.9 ± 10.3 nm for ΔPEX11 cells. The strain, as indicated by the slope of the fits, is increasing at the same time as well. Data for prolonged incubation (50 hrs) of wildtype cells, which is presented in figure 9.4, is in line with the observed trend of an increase in domain size and strain with the incubation time. Interestingly a distinct difference in domain size is apparent between both cell lines, as it is expected due to the biological differences. ΔPEX11 cells show an apparent domain size more than twice as big as the wildtype cells. This difference in size is similarly linked to a difference in strain on the same order of magnitude. The differences in peak FWHM

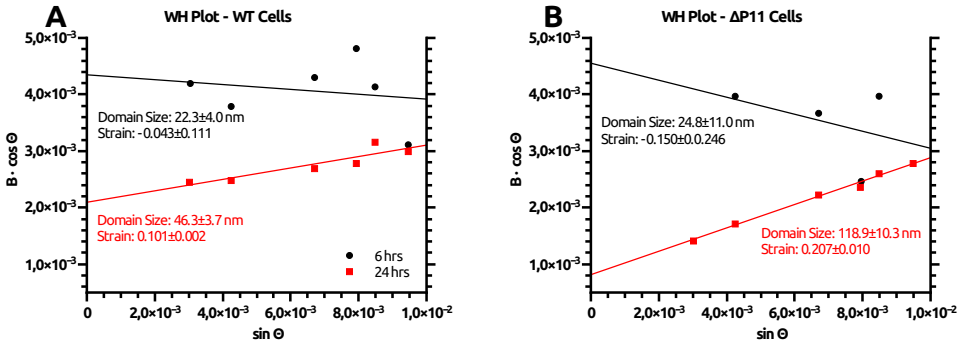

Figure 9.5: WILLIAMSON-HALL *plot of* ***(A)*** *H. polymorpha wildtype- and* ***(B)*** *ΔPEX11 cells grown under P_{AOX} inducing conditions for 6 hrs (black) & 24 hrs (red); Experimental parameters are listed in table 9.2, exposure time: 15 s; Data listed in table A.1*

do not arise from differences in diffraction intensity as shown in a control experiment (see fig. A.6). Hence from the presented results it is possible to derive the following conclusions. First, cellular growth of *H. polymorpha* cells under P_{AOX} inducing conditions is increasing the apparent domain size, but is also increasing strain within the crystals. Thus a prolonged growth seems to counteract on defect-free crystal formation. Second, the enlarged crystals found in ΔPEX11 cells do not possess advantageous properties in comparison to crystals from wildtype cells. Their prolonged maintenance within cells seem to have a negative impact on the diffraction properties. Intriguingly, the storage of crystal-containing cells at 4°C for longer time periods, does not seem to tremendously affect crystal quality (see fig. A.8).

Crystal Quality at Different Growth Temperatures

The yeast *H. polymorpha* is a thermotolerant organism, capable of growth at temperatures between 10-50°C [146]. Besides the manifold cellular changes upon incubation at different temperatures, AOX crystal quality itself should differ with growth temperature. Thus wildtype cells were incubated at 25, 37 and 45°C in mineral medium and 0.5% methanol as sole carbon source. Growth was allowed for 24-30 hrs to reach comparable cell densities (OD 6). Data was processed as explained above, to yield a Williamson-Hall plot. Even though cell densities were comparable, the total crystal content seem to differ. In a comparison of the total peak area of the (2 0 0) reflection, which was earlier also denoted as 100% peak, wildtype cells grown at 37°C diffracted 40 and 50 times stronger than cells grown at 45°C and 25°C, respectively.
Due to the growth on methanol as the sole carbon source, not only the sheer induction of the P_{AOX} promoter can be assumed, but also that AOX protein is successfully expressed to facilitate carbon metabolism. Therefore it seems plausible to conclude that the total crystal content is mainly influenced, not by the amount of produced alcohol oxidase, but rather other stimuli, like a faster rate of pexophagy at lower growth temperatures [96]. Nonetheless it has been shown in section 11.3, that peroxisomal development appears to be only slowed down at lower temperatures. Regardless of possible changes related to cellular metabolism, a Williamson-Hall analysis might reveal distinct quality differences among AOX crystals of different origin. And indeed, those crystals accrued at 25°C show considerably less strain than crystals from a cell batch grown at 37°C (fig. 9.6). Not surprisingly is the amount of strain highest in cells grown at 45°C. Interestingly is the apparent crystal domain sizes largest (lowest y-intercept) for the highest temperature and smallest for the lowest temperature. For cells grown at 25°C one can obtain a value of 41.4 ± 5.2 nm, whereas domain sizes of 63.0 ± 18.4 nm and 82.3 ± 44.3 nm are obtained for cells grown at 37°C and 45°C, respectively. In general this means that cells grown at higher temperatures tend to have larger crystal domains, but also exhibit more strain. The occurrence of crystal formation at temperatures as high as 45°C is already quite remarkable. The intrinsic tendency of alcohol oxidase to form octamers and crystals appears to be very high. This is supported by the apparent thermostability and resistance against reducing agents exhibited by multimeric alcohol oxidase, as shown in figure 8.3,

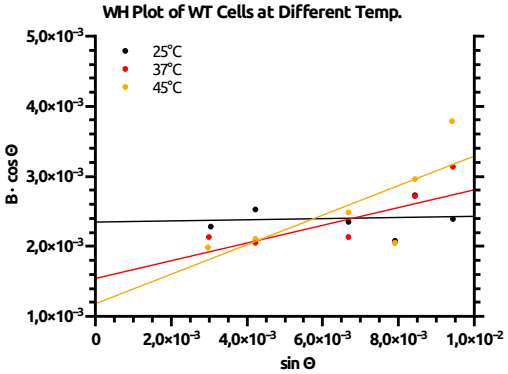

Figure 9.6: Williamson-Hall *plot of H. polymorpha wildtype cells grown under P_{AOX} inducing conditions for 24-30 hrs at different temperatures; Experimental parameters are listed in table 9.2, exposure time: 15 s; Data listed in table A.2*

but also by the occurrence of cytosolic crystals in ΔPEX5 mutants (sec. 10.5).

9.2 Powder Diffraction Analysis of Various Yeast Samples

Chapter 8 presents efforts and difficulties in order to obtain modified yeast strains that express various target proteins for *in vivo* crystallization in peroxisomes. The obtained clones were subjected to *in cellulo* powder diffraction analysis after growth under different conditions, in order to examine possible crystal formation.
Table 9.5 holds information about all tested samples and sample growth conditions. *H. polymorpha* strains were typically grown in non-selective YP media due to the stable genomic integration of target sequences. Additionally, 0.5% methanol was used to induce peroxisome proliferation and P_{AOX}-related pathways. 0.1-0.2% glycerol was added to supply yeast cells with a utilizable carbon source. Furthermore, batch cultures were supplemented with FAD in some cases to stimulate protein expression. It was also tried to grow *H.p.* cells as fed-batch cultures. In such case, cells were quickly transferred into new medium after 24 and again after 12 (36) hrs. *S. cerevisiae* strains were grown in selective synthetic-defined (SD) medium on the basis of YNB (LEU2,URA axuotrophy). For the induction of protein expression either 0.2% oleic acid/ Rytka (P_{FOX3}), or 0.3% glucose (P_{PGK}) were supplemented. Growth was typically allowed until batch cultures reached the stationary growth phase or as indicated. Powder diffraction experiments were conducted as presented in section 4.1 in *MiTeGen MicroRT*^{TM} tubes at beamline P14 (PETRAIII, DESY, Germany). Unfortunately, none of the tested growth conditions yielded crystal diffraction in powder diffraction experiments at beamline P14 (DESY, Germany).

Table 9.5: *Overview of samples tested in powder diffraction experiments at synchrotron beamline P14 (PETRAIII, DESY, Germany) to verify in vivo crystallization; Yeast specifications and their respective growth conditions; YPM and MM contained always a combination of methanol and glycerol as described; SD: synthetic-defined medium, O: oleic acid, R: Rytka, G: glucose*

Nb.	Species	Strain	Type	Promoter Expression	Gene Origin	Media	Inc. Time [hrs]	Temp. [°C]
1	*H. polymorpha*	NCYC495	12DAC4	T4 Lysozyme	*T4 bacteriophage*	YPM	24, 36, 72	25, 37
2	*H. polymorpha*	NCYC495	12DAC4	T4 Lysozyme	*T4 bacteriophage*	MM	24, 36	37
3	*H. polymorpha*	NCYC495	12DAC4	Hex-1	*N. crassa*	YPM	24, 48, 72	15, 20, 25, 30, 37
4	*H. polymorpha*	NCYC495	12DAC4	Hex-1	*N. crassa*	MM	24, 48	25, 37
5	*H. polymorpha*	NCYC495	12DAC4	Esat-6	*S. agalactiae*	YPM	24, 72	25, 37
6	*H. polymorpha*	NCYC495	12DAC4	Esat-6	*S. agalactiae*	MM	24	37
7	*H. polymorpha*	NCYC495	12DAC4	Cathepsin B	*T. brucei*	YPM	24, 72	25, 37
8	*H. polymorpha*	NCYC495	12DAC4	Cathepsin B	*T. brucei*	MM	24	37
9	*H. polymorpha*	NCYC495	12DAC4	PRDX5	*H. polymorpha*	YPM	24, 72	25, 37
10	*H. polymorpha*	NCYC495	12DAC4	EGFP	*A. victoria*	YPM	24, 72	25, 37
11	*H. polymorpha*	NCYC495	12DAC4 / ΔPEX11	EGFP	*A. victoria*	YPM	24	37
12	*S. cerevisiae*	UTL7A	wildtype	AOX	*H. polymorpha*	SD+O/R	24, 48, 72	30
13	*S. cerevisiae*	BY4742	wildtype	Hex-1	*N. crassa*	SD+O/R	24, 48	30
14	*S. cerevisiae*	BY4742	wildtype	Hex-1	*N. crassa*	SD+G	24, 48	30

10 | Proof of Principle Experiments for Yeast *In Vivo* Diffraction at XFELs

The preceding chapters delineate a conceptual approach to investigate properties of the yeast peroxisomal system. Ranging from genetic engineering (ch. 8) to synchrotron-based diffraction studies (ch. 9), effort has been put into the investigation of the utilizability of *H. polymorpha* cells as a potential 'crystal factory' for SFX studies at XFELs.
This chapter comprises the description of successful proof-of-principle SFX experiments with AOX nano-crystals contained in whole *H. polymorpha* yeast cells, at the free-electron laser LCLS (SLAC Stanford, USA). First, sections 10.1 & 10.2 show the purification of yeast peroxisomes and their characterization via DLS to estimate crystal sizes. Additionally section 10.3 presents results from a sample investigation via transmission electron microscopy, to verify AOX crystal formation and analyze sample integrity for batches of purified peroxisomes. The testing of sample injection into a X-ray beam via a liquid microjet is described in sec. 10.4. Finally diffraction studies are described in section 10.5. The presented results demonstrate the general feasibility of a yeast-based, *in vivo* sample delivery-approach in SFX.

10.1 Purification of Mature Peroxisomes from Yeast Cells

In this section the preparation of peroxisomal fractions from two mutant yeast strains is demonstrated. Wildtype-like peroxisomal characteristics are embodied by one strain (WT, PMP47-GFP), whereas a PEX11 deficiency leads to different phenotype in the other mutant (ΔPEX11, PMP47-GFP). Both mutants express the fusion protein PMP47-GFP, which is constituted of PMP47 as a peroxisome-specific adenosine triphosphate transporter and monomeric GFP. The abundance of PMP47 enables visual inspections in a microscope, but was not subjected to quantification of total fluorescence signals.
Crystal-containing yeast cells were produced according to a P_{AOX}-induction protocol (sec. 6.2). For each cell type a total of 600 mL culture was grown in YPM media. For efficient incubation the cells were split into 3×200 mL and poured in 2 L flasks. Growth was allowed over night until early stationary phase up to an optical density of 10.
In a second step peroxisomes were purified via a two-step purification of differential- and sucrose density-centrifugation. Cells were harvested at room temperature for 5 min at$3800 \cdot g$ and subjected to successive treatment. First protoplasting was achieved by zymolyase treatment. After disrupting the

protoplasts the total cell lysate was deployed in differential centrifugation as described, to obtain a post-nuclear supernatant (PNS). The initial differential centrifugation step is thought to separate the cell organelles from the nucleus fraction and cellular debris like membranes and cell wall remains. The subsequently conducted sucrose density gradient can yield cleanly separated organelles according to their density. In a first trial the sucrose density centrifugation was conducted at 30000·g for 5 h. A detailed description of all methods is given in section 6.5.
To obtain a pure peroxisomal fraction a smooth linear sucrose gradient usually yields good results. This is accomplished by pouring a linear gradient in a single-use centrifugation tube from which fractions in the milliliter range can be extracted. The result is then verified with a variety of organelle-specific marker antibodies in a western blot (see sec. 8.4). As these requirements could not be fulfilled at the time of the experiments an alternative approach was developed to identify the peroxisomal fraction in layered gradients.
Instead a catalase assay based on foam formation was used for the identification of the peroxisomal fraction (sec. 6.5). The assay itself was already applied to assess catalase activities in various bacterial strains and human cells [186]. As it is known that catalase activity is a reasonable marker for peroxisomal fractions [36, 181], the foam-based assay will be exploited to optimize the purification of peroxisomes from yeast. The resulting gradient is shown in

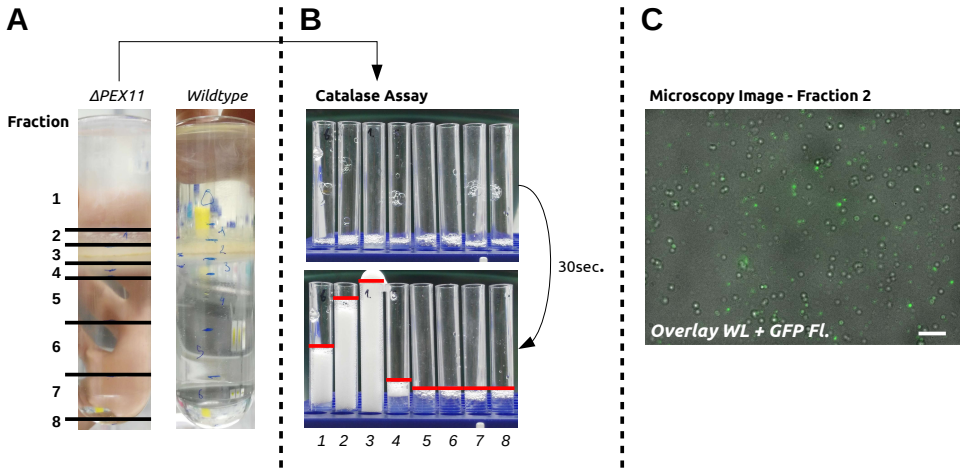

Figure 10.1: ***(A)*** *Sucrose-density gradient of PNS from wildtype and ΔPEX11 cells after 5 hrs at 30000·g* ***(B)*** *Catalase activity assay conducted on fractions 1-7 (ΔPEX11 cells)* ***(C)*** *fluorescence microscopy image, overlay of white light and FITC channel (350 ms exposure time); Scale bar: 5 µm*

figure 10.1-A for both, wildtype and ΔPEX11 mutant. Fractions are indicated by black lines and were chosen based on the gradient-internal density interfaces and visible agglomerations. The density gradient is based on a protocol with sucrose concentrations of 48%, 50%, 52% and 65%, which approximately corresponds to the densities 1221, 1232, 1243 and 1319 $kg \cdot m^{-3}$ [97]. Figure 10.1-B shows the results of the catalase assay for the ΔPEX11 mutant, that was conducted after crude fractionation of the gradient. After an incubation time of 30 s the assay was evaluated with regards to catalase activity as indicated by the drawn red lines. For each reaction 100 µL of extracted fraction material was diluted 1:1 with homogenization buffer to decrease the overall sucrose concentration and centrifuged for 30 min at 30000·g. Thereafter the pelleted organelles were vigorously resuspended in 250 μL homogenization buffer and

incubated with each 100 µL 30% H_2O_2 and 10% Triton X-100. The resulting foaming due to the presence of catalase was investigated. From the obtained results no distinct difference between wildtype and ΔPEX11 mutant could be found with regards to separation in the gradient. Mature peroxisomes are

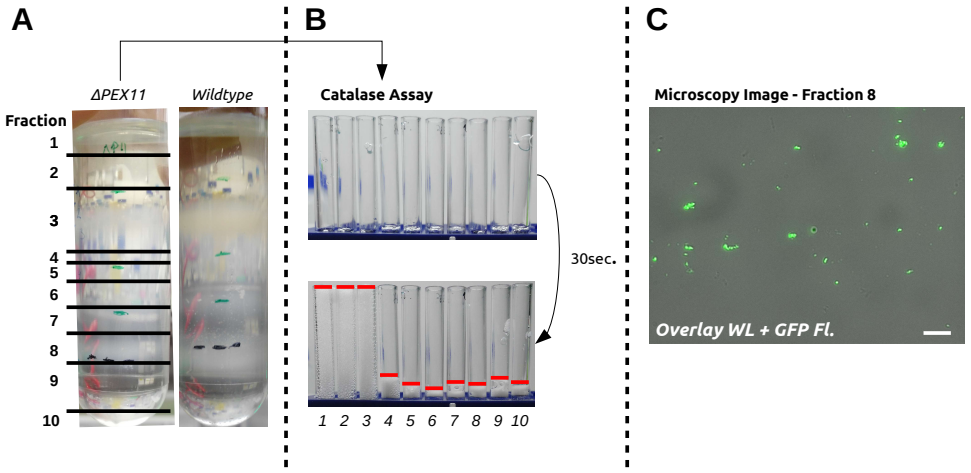

Figure 10.2: ***(A)*** *Sucrose-density gradient of PNS from wildtype and ΔPEX11 cells after 10 hrs at 30000·g* ***(B)*** *Catalase activity assay conducted on fractions 1-10 (ΔP11 cells)* ***(C)*** *fluorescence microscopy image, overlay of white light and FITC channel (350 ms exposure time); Scale bar: 5 µm*

thought to have the highest density of all cell organelles in the PNS and hence were expected to accumulate at the interface between 52% and 65% sucrose solution, or eventually migrate to the bottom of the centrifugation tube [81, 104]. Interestingly, the catalase assay strongly indicated an accumulation of peroxisomes at lower densities corresponding to a sucrose concentration of 0-48% (fractions 2-4). Figure 10.1-C shows organelles from fraction 2 in an overlay of white light image and FITC channel. It clearly shows a large amount of peroxisomes apparent from the PMP47-GFP fluorescence, but unfortunately co-localized with other non-fluorescent organellar structures. Strong foaming for low density fractions is also partially explained by the presence of catalase from the cytosolic fraction and the overall protein concentration.
Furthermore the following explanations for the results can be adduced. First, mostly immature peroxisomes could populate the cells and hence end up at lower densities. A certain fraction of immature peroxisomes is always expected in a viable culture (sec. 2.2), but no visible organelles could be seen in fraction 7 or 8 when they were subjected to microscopic investigation. The latter observation would be conclusive with the second explanation, which is that the peroxisomes did not properly migrate through the gradient. Either because they stick to other cellular components or because the time of centrifugation was too short.
Hence the protocol was adapted to this new hypothesis. The time for centrifugation for the sucrose density gradient was increased to 10 h and the cell lysate has been treated by more extensive pipetting to achieve a stronger homogenization of the sub-cellular fractions. An inspection under the microscope indicated still intact organelles (data not shown). The results of the new approach are shown in figure 10.2. As expected and visible in 10.2-A the gradient appears more fractionated. Due to the appearance of a more complex purification pattern 10 layers were defined for extraction and subjected to the catalase assay (fig. 10.2-B). In both cases the same amount of protein of

about 7,5 mg·mL^{-1} in 5 mL was loaded onto the gradient for comparability. To contain the foaming the total volume of reagents has been reduced by half to 50 µL each. Intriguingly it was impossible to decrease the amount of foaming reaction in lower density fractions 1-3, but it slightly increased in all higher density fractions 4-10 (see fig. 10.2-B), which is indicative for a better migration of peroxisomes. A qualitative comparison of the purifications yield a slightly increased catalase reaction at higher density fractions.
It is apparent that a lot of catalase- or catalase-like activity can be located in the low density region of the gradient. This can be due to cytosolic enzyme activities or immature peroxisomes, but might also be attributed to disrupted peroxisomes. In addition it can be assumed that a certain amount of peroxisomes gets stuck during the separation process. Prolonged centrifugtation seem to counteract on these problems by increasing the amount of catalase activity in higher density fractions, indicative for peroxisome migration.
Hereupon it was necessary to verify peroxisomal content in higher density fractions for the newly derived protocol. Therefore fraction 8-10 were pooled, centrifuged and resuspended in 100 μL homogenization buffer. In a microscopic investigation of both wildtype and ΔPEX11 mutant, intact peroxisomes could be verified via the PMP47-GFP marker. Figure 10.3-A and -B show examples from the pooled fractions for wildtype and ΔP11 mutant, respectively, in white light and FITC channel. Even though the findings indicate a rather pure fraction of intact peroxisomal material, many of the peroxisomes seem to cluster, which does fit to the described problems during purification. It can be assumed that the purified peroxisomes at high densities represent a subset of mature peroxisomes. Pure peroxisome fractions obtained in the same manner were subject to experimental investigation of the size distribution via dynamic light scattering (DLS) as described in section 10.2.

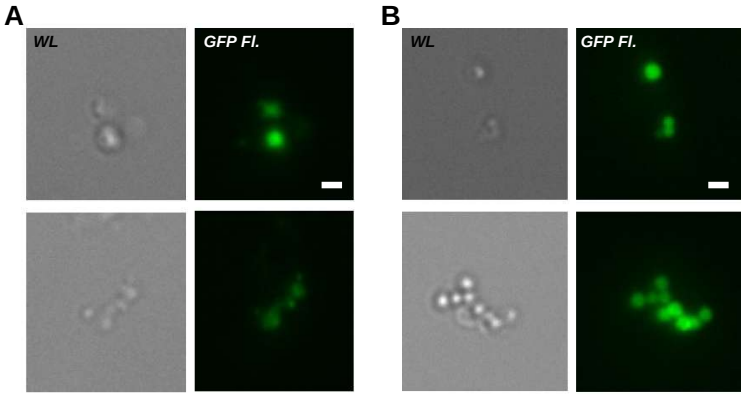

Figure 10.3: *Images of purified peroxisomes (fractions 8-10) from* ***(A)*** *wildtype and* ***(B)*** *ΔPEX11 cells under white light illumination and in FITC channel (200 ms exposure time); Scale bar: 1 µm*

10.2 Size Distribution Analysis of Purified Yeast Peroxisomes via DLS

This section presents a size distribution analysis of purified yeast peroxisomes (see sec. 10.1) to validate sample quality and approximate particle sizes for SFX data collection.
The analysis of purified peroxisomes was difficult due to their small scale and

fragile nature. At first it was tried to analyze purified batches via fluorescence microscopy which would hold the advantage of excluding remains and impurities. Unfortunately it was not possible to reliably embed larger fractions of the organelles in a stabilizing matrix like agarose to prevent diffusion during exposure times (data not shown). Hence it was tried to analyze purified batches via a non-specific, but very sensitive technique like dynamic light scattering (DLS). Two different types of preparation procedure were tested. To prevent early degradation of samples first fixated peroxisomes were subjected to measurements. Subsequently untreated peroxisomes were measured to obtain a complete view on sample stability and behaviour.
For all DLS measurements a setup (*SpectroSize 300*, Xtal Concepts, Germany) with a diode laser at a wavelength of 660 nm has been used (sec. 5.3). The setup is theoretically limiting reliable measurements to particles sizes within a range of 0.001-6 µm. Measurements were conducted in a quartz cuvette with a minimum sample volume of 10 µL and at a temperature of 20°C. Data was recorded in a series of measurements (n=30, 30 s) in each case. For all mea-

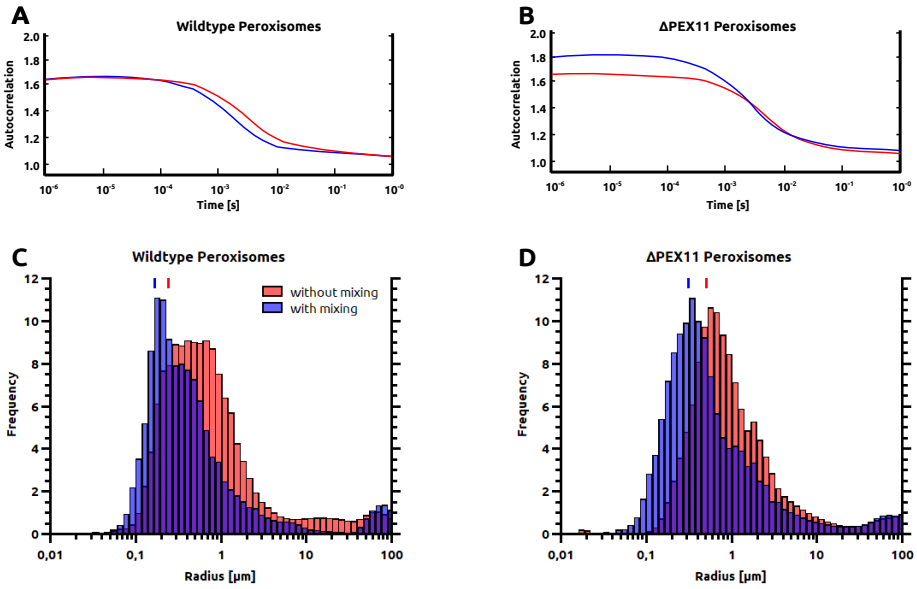

Figure 10.4: *Mean autocorrelation curve for fixated* ***(A)*** *wildtype and* ***(B)*** *ΔP11 peroxisome solutions and likewise the corresponding mean hydrodynamic radius distributions for* ***(C)*** *wildtype and* ***(D)*** *ΔP11 peroxisome solutions; n=30, 30 s; Mean and standard deviation can be found in table 10.1, data from a* CONTIN *analysis in table A.3; red=without vigorous mixing, blue=with vigorous mixing*

surements peroxisomes were purified as described in section 10.1. A first batch of organelles was mildly fixated with 0.2% glutaraldehyde at 4°C overnight to prevent degradation. Fixation was carried out in a rotating manner to counteract on sedimentation and clustering processes. The next day organelles were extracted from the sorbitol-containing homogenization buffer and gently resuspended in 5 mM MES buffer pH 5.5 containing additionally 0.1 M EDTA. This allowed the usage of an 'aqueous' viscosity parameter of 1.006 $mPa \cdot s$ to correctly determine the sizes of particles (cp. fig. A.9). Visible contaminations were spun down for 30 s at 500·g and the resulting supernatant has been used for measurements.
Figure 10.4-A/B show exemplary the mean autocorrelation functions for different experiments with fixated peroxisomes corresponding to the mean radius

distributions shown in figure 10.4-C/D. Both graphs are compiled from several successive measurements as described before. The autocorrelation function is normalized in y on the interval [0,2] and hence values on [1,2] mathematically represent positive correlations. As it can be seen for all measurements, the autocorrelation curves fall into this range over the investigated orders of time. As it is represented by the red-colored histograms in figure 10.4-C/D a reasonable distribution function is obtained for both strains in the expected range of sizes. The distribution is uniform in both cases with a skew to larger peroxisome radii. About 72% of the wildtype and 81% of the ΔP11 data fall onto the interval between 0.1 and 1 µm. The mean particle sizes and polydispersity parameters from a CONTIN analysis are shown in table A.3. For the specified interval [0.1,1] a mean of about 356 nm can be found for wildtype peroxisomes, whereas the radius distribution of ΔPEX11 peroxisomes has a mean of about 377 nm on the same interval (see table 10.1). It is known that organelles tend to stick together (see sec. 10.1) and hence DLS measurements might represent data that is biased towards larger particles. In fact also the mild fixation could lead to clustering and thus might influence the presented measurements. Consequently peroxisomes were vigorously mixed with a 1 mL pipette in a second step, to break up any appearing clusters. Afterwards a second measurement was done, which is represented by the blue-colored histograms in figure 10.4-C/D. Intriguingly, a shift in size appearances in the mean radius

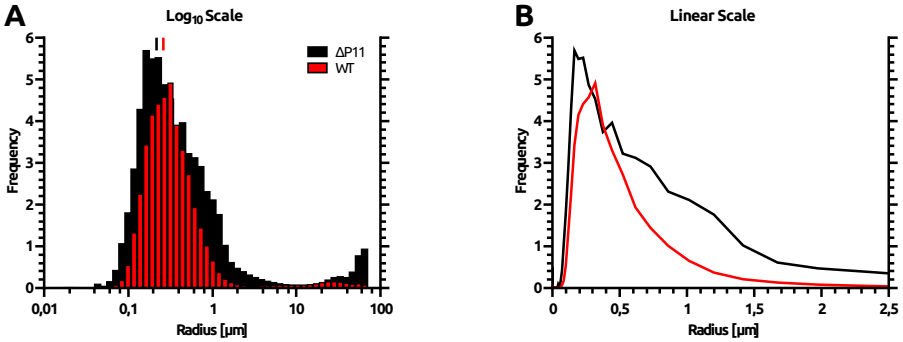

Figure 10.5: ***(A)*** *Hydrodynamic radius distributions for unfixated wildtype and ΔP11 peroxisome solutions, log_{10} scale* ***(B)*** *linear scale; n=30, 30 s; Mean and standard deviation can be found in table 10.1; Data published in [253]*

distributions is observable for both wildtype and ΔPEX11 batch. This can be explained by a depopulating process with regards to fractions representing larger particles (0.5-10 µm), whereas simultaneously fractions of smaller particles (0.1-0.5 µm) are getting more densely populated. This is indicating the presence of larger particles that can be successfully disrupted by vigorous mixing. Furthermore fractions of sizes < 0.1 µm do not drastically increase in occupancy. This might lead to the conclusion that indeed peroxisomes are not disrupted by the treatment which would free a lot of nanometer-sized particles from entrapment. In fact it shows that mildly fixated peroxisome-clusters can survive in buffer without any sorbitol content and that these cluster can be broken up for effective size distribution measurements without interfering with general particle integrity. Mean values of 289 nm and 295 nm can be calculated for wildtype and ΔPEX11 peroxisomes, respectively.
Data that is representing very large particles (> 1 µm) is only shown for completeness, but is generally excluded from the statistical analysis. First because of the sparse appearance of such data points (compare 10.4-C/D

& 10.5-A). Secondly because these fragments most likely represent clustered peroxisomes or contaminants like dust particles that are unrelated to the target of interest. As a next step freshly purified peroxisomes have been used for size

Table 10.1: *Mean and standard deviation of radius frequency distributions on the interval [0.1,1] for experiments presented in figures 10.4 and 10.5 (fixated and fresh peroxisomes)*

Sample	**Mean Radius [nm]**	**StDev [nm]**
Fix. WT -M	355	342
Fix WT +M	289	280
Fix. ΔP11 -M	377	361
Fix ΔP11 +M	295	363
Fresh WT	291	281
Fresh ΔP11	314	305

distribution analysis[1]. The obvious advantage of this approach is of course the circumvention of degradation processes. But unfixated peroxisomes are very sensitive and can break easily (cp. sec. 10.3). Similarly to the fixated batches fresh peroxisomes were centrifuged for 30 s at 500·g and the supernatant was subsequently used for measurements. For all measurements of fresh peroxisomes the original homogenization buffer containing 1.2 M sorbitol was kept to maintain organelle stability. Hence the used viscosity parameter had to be adapted to a value of 1.8 [119]. The obtained results are depicted on a logarithmic and a linear scale in figure 10.5 for both wildtype and ΔPEX11 peroxisomes. In addition all obtained parameters are summarized in table 10.1. The distribution functions are shown in figure 10.5-A/B. These results indicate in general a good agreement of the obtained distributions for unfixated and fixated organelles for both wildtype and ΔPEX11 peroxisomes. 82% and 92% of the data from wildtype and ΔPEX11 peroxisomes fall onto the interval [0.1,1].
The differences between the two unfixated samples are getting more obvious when viewed on a linear scale (fig. 10.5-B). The highest frequency of peroxisome sizes is found at 261 nm for wildtype and at 216 nm for ΔPEX11 peroxisomes, as indicated by the black and red line in figure 10.5-A. But the mean hydrodynamic radius of ΔPEX11 peroxisomes ($\sim$ 314 nm) is slightly larger than for the wildtype case ($\sim$ 291 nm). Thus the largest fraction of contained peroxisomes is in both cases very comparably located at similar radii around 200-300 nm. The related data is listed in table 10.1. Additionally the ΔPEX11 related distribution also shows more densely populated fractions of larger peroxisomes (0.5-1.5 µm) which cannot be seen for the wildtype sample. This is evidence for a sub-population of large peroxisomes that support the so far observed phenotype of this mutant [147]. Due to the aging of cells and organelles and cellular recycling processes, it is not surprising to find a large amount of smaller and most likely younger peroxisomes (cp fig. 2.1).
In general it can be concluded that the presented results verify the phenotype of the ΔPEX11 mutant with regards to peroxisomal sizes, even though the differences are not as pronounced. Additionally, a mostly narrow size distribution for peroxisomes can be verified for the wildtype strain, which is advantagous for SFX experiments and data analysis [155]. Still, large ΔPEX11

[1]Data is published in [253]

peroxisomes could be enriched to obtain a more homogenous distribution of large peroxisomes and hence crystals.

10.3 Verification of Sample Crystallinity by Electron Microscopy

It is already known that wildtype *H. polymorpha* cells contain peroxisomal AOX crystals. Their appearance is correlated with peroxisome proliferation and P_{AOX}-induction, which are induced under methanol conditions [40]. Here we demonstrate that these crystals make up most of the peroxisomal volume in yeast cells and are present within cells when grown under laboratory conditions. Furthermore, it is shown that those crystals are preserved in intact peroxisomes that are the product of the applied purification method (sec. 10.1)[1]. All images were acquired using a *Philips CM12 TEM*, operated at 200 kV, and a cooled slow-scan CCD (*Ultrascan 4000, Gatan*). For experiments the mutant

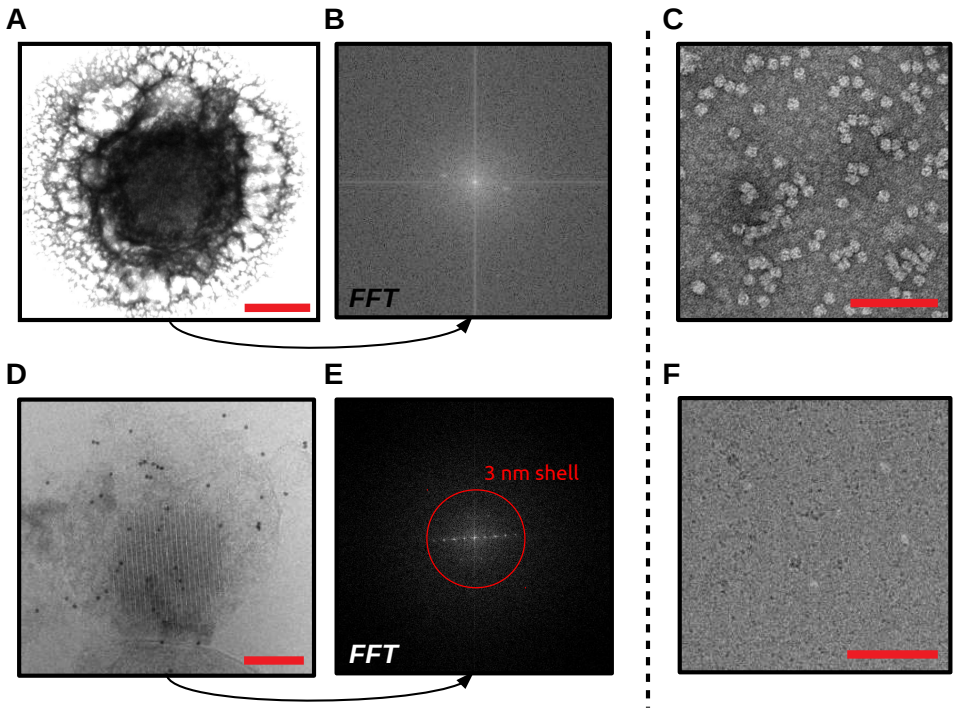

Figure 10.6: *Electron micrographs of purified peroxisome batches; Examples of a peroxisomes in* ***(A)*** *negative stain and* ***(D)*** *cryo- EM with its* ***(B,E)*** *fast fourier transform and* ***(C,F)*** *an example of loose AOX octamers from the same batches; (D-specific) αAOX antibody-gold labelling; Scale bars: (A) 200 nm (C,D,F) 100 nm*

strain *H. polymorpha* PEX14-mCherry was grown for 16 h in complex mineral medium with 0.5% methanol as sole carbon source. Afterwards peroxisomes were obtained in the described manner (Sec. 10.1). For cryo-EM micrographs purified peroxisome or whole cells were subjected to cryo-fixation in liquid ethane by applying the sandwich plunge-freezing method [125]. Subsequently, freeze-substitution was done with 1% osmium tetroxide, 0.5% uranyl acetate and 5% (v/v) distilled water in acetone [167]. After soaking and polymerization

[1] All experiments of section 10.3 were conducted in the labs of Ida v.d. Klei (RUG Groningen, The Netherlands) and supervised by Dr. K. Knoops who also conducted the EM experiments.

with *Epon 812*, thin-sections of about 160-200 nm thickness were cut. For negative-stain micrographs, samples were resuspendend in homogenization buffer and subjected to fixation with 6% glutaraldehyde buffered with 0.1 M Na-Cacodylate. Thin-sections were again obtained by soaking and polymerization with *Epon 812* followed by cutting. Afterwards, staining with 0.5% uranyl acetate was conducted.
Figure 10.6 depicts results of the conducted EM investigation with negative staining (fig. 10.6-A-C) and under cryo-conditions (fig. 10.6-D-F). The three sub-figures A, B, C originate from a first attempt to obtain EM images of peroxisomes by negative staining. Figure 10.6-A was chosen as a representative image of identifiable peroxisomes. The negative staining procedure seems to interfere with peroxisomal integrity as only a few intact peroxisomes could be observed. Additionally, the staining procedure can only barely reveal the intraperoxisomal AOX crystal structure. Hence, it might be suggested that negative staining is detrimental to the fragile structure of purified peroxisomes. Figure 10.6-B represents a fast fourier transform (FFT) of the intraperoxisomal region revealing the occurrence of distinct frequency features. This at least indicates some minor repetitive 2D structure within the image which might be interpreted as crystal. Figure 10.6-C shows another subsection of the prepared sample. A lot of loose AOX octamers can be observed which fits to the earlier mentioned hypothesis of peroxisomes being destroyed during staining. Destruction of the protecting shell obviously leads to a decomposition of AOX crystals. The octameric structure of the observed particles could be verified by the generation of class-sums with the software *e2boxer* and *Relion* (fig. A.10). From these findings it might be concluded that the preparation of purified peroxisomes is too harsh and can lead to degradation. On the other hand the general sample preparation for negative staining can lead to unwanted bias in sample quality and hence might be the source of destroyed peroxisomes. Nonetheless a repetitive structure could be observed and gives evidence for AOX crystallites contained in peroxisomes. To investigate the quality of the preparation procedure and possible bias from negative stain treatment peroxisomes were also investigated under hydrated conditions in cryo-EM.
Figure 10.6-D shows a representative example of a purified peroxisome in cryo-EM. Likewise to the frequency space representation for negative staining

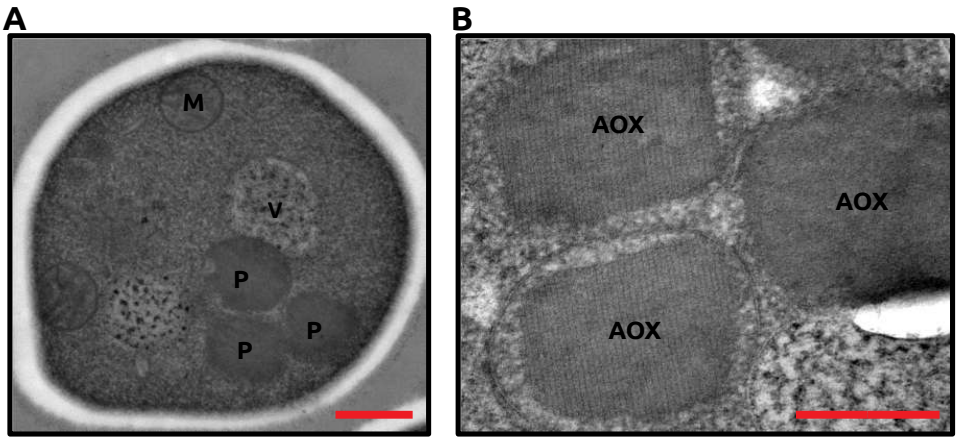

Figure 10.7: ***(A,B)*** *Electron micrographs of cryo-sections of whole H. polymorpha yeast cells; P=Peroxisomes, M=Mitochondria, V=Vacuole, AOX=alcohol oxidase; Scale bars: (A) 500 nm (B) 200 nm*

in sub-figure B, sub-figure E depicts a FFT of the intraperoxisomal region from sub-figure D. This time more distinct frequency features can be obtained, that give strong evidence of a crystalline structure within the peroxisomes. Obviously a better contrast can be obtained under cryo-conditions. Frequency features can be located up to a resolution of 3 nm as indicated by the red circle in figure 10.6-E, which fits to the approximate resolution limit in powder diffraction experiments (cp. sec. 9.1).
Figure 10.6-F represents another section without peroxisomes but with solubilized AOX octamers. Again, this seems to indicate a too harsh treatment of fragile peroxisomes. On the other hand, the emergence of loose AOX octamers for both preparative techniques might indicate that non-crystalline AOX protein is present in the originally purified batch of organelles. Nonetheless, it is possible to conclude that the applied purification techniques for peroxisomes can preserve the crystals within peroxisomes, as long as these remain intact.
For comparison cryo-sections of whole yeast cells were prepared. The preparation has been conducted equally to organelle samples (see above). Figure 10.7-A shows an example of an intact yeast cells containing a visible vacuole, mitochondria and peroxisomes. Those structures are very distinct due to their electron density and have been identified in previous experiments via gold-antibody binding to marker proteins [230]. Additionally, fig. 10.7-B depicts a more detailed view of the electron-dense peroxisomes, revealing the regular structure of AOX crystals, which seem to occupy most of the peroxisomal volume.
The obtained results verify the emergence of AOX crystals within peroxisomes when grown according to the stated protocols. More importantly, it seems legitimate to assume, that AOX crystals occupy most of the peroxisomal volume in intact yeast cells. It is not clear, whether the purification of organelles might introduce too much stress resulting in smaller crystals within intact organelles, or if organelles carrying larger crystals are more prone to destruction.

10.4 Testing Yeast Cell- Beam Injection via Liquid Jets

Different sample delivery methods exists for SFX experiments at FELs (see sec. 1.4). For the injection of whole yeast cells and purified peroxisomes a liquid microjet injection via a GDVN (gas dynamic virtual nozzle) has been chosen [129]. GDVNs allow for the delivery of micrometer-sized samples in a hydrated manner, simultaneously limiting background scattering by the reduced diameter of the liquid jet, which is narrowed down by the peripheral gas flow. Indeed typical GDVNs exhibit a relatively high sample consumption of about 10 µL/min to maintain stable jets at high velocity flow. However sample availability is practically not limited for *in vivo* crystals from yeast cells.
Sample injection has been studied on a test setup under atmospheric pressure conditions. Injection at FELs is typically conducted under high vacuum conditions. Thus vaporization will lead to a slightly shorter jet, but the fast replenishment of liquid solution due to the high flow velocity will counteract on this effect and stabilize the microjet [129, 212, 262]. Other parameters like tube diameter and valve organization were similar to a typical SFX experiment.
Yeast cell integrity is not strongly affected by the ambient ionic strength. Thus yeast cells can easily be resuspended in distilled water, without affecting cell

integrity [18]. Cell suspensions were stored in a pressurized sample reservoir and delivered via 1 mm tubes and a filter set to the fused silica GDVNs. Cell delivery is typically maintained at 100-1000 psi (7-70 bar) to achieve flow rates of 5-50 $\mu L \cdot min^{-1}$. In SFX experiments the sample reservoir is mechanically rotated to avoid settling of cells during typical data collection times of 5-6 hrs [176]. The outer capillary was pressurized by applying molecular nitrogen gas from a gas bottle at a constant flow on the order of 10^3 µg·min^{-1}. Jet characteristics were investigated by using a high-speed camera. Different nozzles were tested for jet stability, length and escape angle.
The resulting test parameters, yielding a stable liquid jet, are listed in table 10.2. Two exemplary jets are shown in fig. 10.8-A and -B, respectively. Most of the tested GDVN nozzles had an inner capillary diameter of 75 µm, achieving flow rates of 8-21 µL·min^{-1} (at 67-380 psi) for a stable jet. Test run numbers 1-9 exemplary show testing of one nozzle with cell suspensions at densities of OD 12- 29, demonstrating that stable jets can be achieved even for dense cell suspensions at low flow rates. The flow of gas in the outer capillary had been kept at a constant value of $2.75 \cdot 10^3$ µg·min^{-1}. Runs 10 and 11 represented those nozzles for which a stable jet could be achieved at the lowest possible flow rate. Run 10 represents a nozzle with an inner capillary diameter (ICD) of 75 µm, while run 11 represents a nozzle with an inner diameter of 50 µm. In run 10 a low flow rate of about 10 µL·min^{-1} could be achieved at low pressures (67 psi), but the decrease of inner capillary diameter in run 11 needed to be

Table 10.2: *Test setup parameters for stable liquid jets at different cell concentrations; ICD=Inner capillary diameter, OD=optical density*

Run	ICD [µm]	Flowrate [µL/min]	Pressure [psi]	Gas Flow [10^3 µg/min]	OD
1	75	21.0	130	2.75	12.1
2	75	15.0	100	2.75	12.1
3	75	12.7	73	2.75	12.1
4	75	21.0	130	2.75	16.1
5	75	15.0	100	2.75	16.1
6	75	12.7	73	2.75	16.1
7	75	21.0	130	2.75	28.7
8	75	15.0	100	2.75	28.7
9	75	12.7	73	2.75	28.7
10	75	9.7	67	2.75	28.7
11	50	11.2	380	2.75	28.7

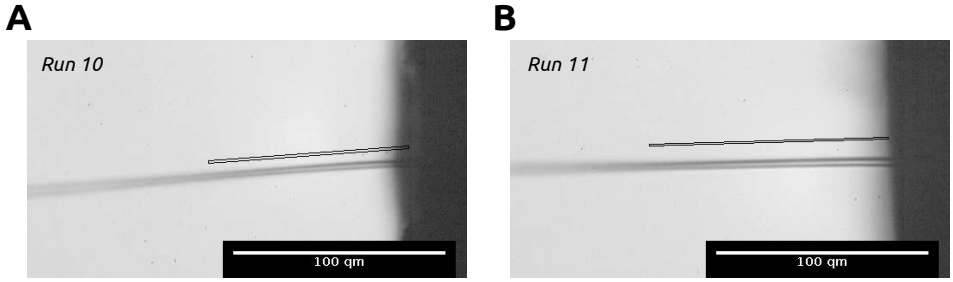

Figure 10.8: *Stable liquid jets for nozzles in* ***(A)*** *run 10 and* ***(B)*** *run 11; Jet breakup was estimated from diffuse regions*

accompanied by a distinct increase in applied pressure (380 psi) to achieve similar flow rates and a stable jet. Liquid jets for runs 10 and 11, with cell suspension at OD 28.7, are shown in fig. 10.8-A and -B, respectively.
In both cases a stable jet exist for about 100 µm as indicated by the drawn lines. The approximate jet diameter is about 3-4 µm (fig. A.11). From this and the wet volume of the cell suspension, it can be approximated that in both cases about 2-3 cells will be present in the jet at any time, limiting the possible hit rate. Unfortunately higher cell densities led to clogging of test nozzles several times (data not shown). The escape angle is for both jets almost 90°, measured in relation to the surface of the capillary tip.
A 10 µm filter has been used to pre-filter the cell suspension for large remains, which could lead to clogging of the nozzle. Cells from a cell suspension at OD 28.7 were collected after successful jetting from run 10. A final OD of 27.5 could be determined. Additionally a microscopic investigation (fig. A.12) revealed intact yeast cells. Thus most of the cells pass through the filter set and tubings and are available for diffraction experiments.

10.5 SFX Experiments with *In Cellulo* Alcohol Oxidase (AOX) Nano-Crystals

Final proof-of-principle SFX experiments[1] were conducted at the CXI beamline (Coherent X-ray Imaging) at LCLS (Linac Coherent Light Source, SLAC Stanford, USA) [233] using the CSPAD detector [173, 225]. Diffraction images were recorded at 120 Hz to match the repetition rate of the FEL, which was running with a pulse length of 30 fs at a photon energy of 7.925 keV (1.56 Å). The FEL had an approximate beam waist FWHM of 100-300 nm.
For injection into the X-ray beam a GDVN (see sec. 10.4) with an inner capillary diameter of 50 µm has been used at a flow rate of 15 $\mu L \cdot min^{-1}$ to obtain a stable jet [129]. The resulting jet had an approximate diameter of 5 µm. The sample reservoir and tubings were maintained at a temperature of 20°C. The sample suspension was constantly agitated within the sample reservoir to avoid sample settling and thus tube- or nozzle clogging.
Whole cells, as well as purified peroxisomes were both considered as potential crystal-stabilizing environments and thus crystal-carriers in typical SFX experiments. A pure batch of carefully purified organelles can theoretically reduce the background scattering by the removal of signal from amorphous molecules and cellular structures. On the contrary, organelles might suffer from severe damage introduced by purification procedures (cp. fig. 10.6-A/C) and can be prone to degradation during prolonged storage. Even more problematic is the storage of organelles in buffers containing sucrose or sorbitol to create an isoosmotic solution. High viscosity solutes have a detrimental effect on the jet stability during SFX experiments. To that effect whole yeast cells offer several advantages as they can be stored for prolonged times at 4°C without degradation, as it was proven by powder diffraction experiments (see fig. A.8). Additionally intact yeast cells exhibit an extreme robustness due to their strong cell wall [58]. Injection of whole cells suspended in distilled water has been demonstrated at high optical densities via GDVNs without cellular degradation or loss of cell material (cp. sec. 10.4). Conclusively the

[1] SFX Experiments were conducted in Oct. 2013 (PCS Proposal LB92) and were published by Jakobi *et al.* [253] in 2016.

usage of whole cell samples was preferred. For diffraction experiments whole ΔPEX11 & ΔPEX5 cells (ch. 2) have been used as promising samples. As it is explained in section 2.2 and schematically depicted in figure 2.1, *H. polymorpha* cells can be genetically modified to yield enlarged peroxisomal or cytosolic crystals constituted of octameric alcohol oxidase (AOX). While a deficiency in the PEX5 protein (ΔPEX5 strain) hinders peroxisomal import of proteins via the PTSI pathway, a deficiency in the protein PEX11 (ΔPEX11 strain) renders a situation in which on average only one enlarged peroxisomes exists per yeast cell. In both cases the formation of crystals is restricted to a single compartment which reduces the risk of simultaneous, multiple and overlapping diffraction patterns. The size distribution of purified organelles, obtained from DLS measurements (sec. 10.2), infers crystal sizes on the order of 200-2000 nm and thus true nanometer-sized crystals.
In total 309,496 frames could be collected during the injection of ΔPEX11 cells and 43,056 frames while injecting ΔPEX5 cells. 1.1% of ΔPEX11 and 0.5% of ΔPEX5 frames could be identified as single-crystal diffraction patterns using *Cheetah* [204]. Figure 10.9-A exemplary shows a single-crystal diffraction pattern obtained from an AOX crystal within a ΔPEX11 cell. The resolution is limited to approximately 30 Å as indicated by the outermost blue circle. Still some crystals showed higher resolution Bragg reflection beyond 30 Å.
Due to the sparsity of Bragg spots, it was not possible to reliably index the diffraction patterns from ΔPEX11 cells, using *CrystFEL* [180]. To still obtain a comparable representation of the diffraction data, a composite powder pattern was created by summation of individual frames, which is shown in figure 10.9-B. Arrows indicate Bragg spots at higher resolution, which extended in very few

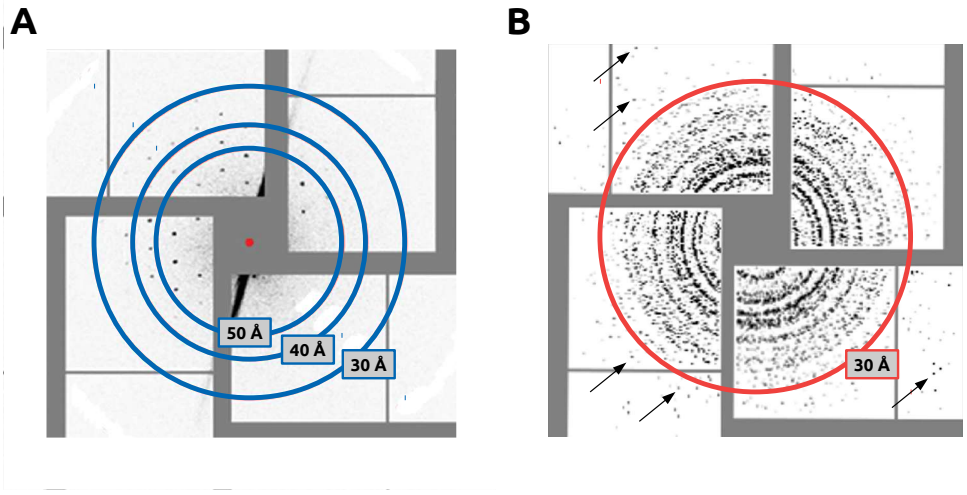

Figure 10.9: ***(A)*** *Single-crystal diffraction pattern from an AOX crystals contained within an intact ΔPEX11 yeast cell* ***(B)*** *Composite powder pattern created by a summation of individual diffraction patterns from ΔPEX11 sample; Blue and red circles indicate the resolution; Data published in [253]*

instances to 6 Å. A crystal solvent content of 63-75% can be inferred from electron microscopy [40] and might be indicative for crystalline disorder and the poor diffraction quality of the crystals, as it has already been suspected in chapter 9.1.
Still, the sparse occurrence of high resolution Bragg spots indicates, that high quality diffraction data can theoretically be obtained from nano-sized crystals, contained within intact yeast cells. The results strongly suggest to

assume a limitation of crystal quality, rather than a limitation of the method. Since in total only 215 frames have been recorded for ΔPEX5 cells that contain single-crystal diffraction, no further analysis was conducted for this sample.

11 | Revealing Properties of the Peroxisomal Crystal Factory

11.1 *In Cellulo* Steady-State Fluorescence Anisotropy as an Analysis Tool

In modern protein crystallographic laboratories a variety of tools exist to investigate the properties of purified protein samples and elucidate their behavior under specified conditions to obtain well-characterized samples. Such tools are an integrative part within a cycle of expression, purification and in-depth analysis, to eventually obtain functional protein in satisfactory amounts and protein crystals for structure determination at the far end. Starting with more generally applicable tools such as dynamic light scattering (DLS) and UV- circular dichroism spectroscopy (CD), but reaching out to isothermal titration calorimetry (ITC), thermal shift assays (Thermofluor) and many more, plenty of methods exist to describe the properties of a protein sample under various conditions [132]. There are likewise easy methods at the stage of crystallographic trials, such as UV-/SHG-imaging, to rate possible successful crystallization conditions well before performing X-ray experiments.
Obviously most projects in protein crystallography can rely on established methods to guarantee at least systematic investigation of the matter. For the novel approach of directed *in vivo* crystallization for SFX experiments researchers need to rethink their approaches since it is dealed with a crystallization environment that is as complex as never before in crystallography. Thus, to understand the behavior of a dynamic and complex system like an eukaryotic cell new methods need to be developed and applied.

Protein concentration is a major factor during *in vitro* nucleation and crystallization and thus inevitably in *in vivo* crystallization. This work will help to elucidate this important system property in yeast *in vivo*/*in cellulo* crystallization.
It is demonstrated that steady-state fluorescence anisotropy (sec. 3.1) in living cells is a suitable, almost non-invasive tool to characterize the cellular environment. Fluorescence anisotropy can be used as a complementary readout to simple fluorescence intensity measurements. With regards to a cellular environment the intensity level can only give information about the general expression level of a probe molecule. On the contrary anisotropy can yield specific information about molecule interactions, concentrations and relative orientation of molecules. The difference in information content between both readouts is illustrated in figure 11.1-A.
This section starts with a brief theoretical discussion to define experimental expectations. Experiments with purified EGFP will demonstrate probe

behaviour *in vitro*, while the subsequent subsections will focus on *in cellulo* EGFP measurements and data analysis. Those chapters will mainly focus on the description of the progression of intraperoxisomal probe protein concentrations and will discuss possible influences on general protein nucleation and crystallization. For all experiments a specifically designed strain of *H. polymorpha* expressing an EGFP probe has been utilized. The EGFP protein carries an N-terminal SKL tag and will be efficiently imported into peroxisomes via the PTSI import pathway (see sec. 2.2). A detailed description of the overall methodology is given in section 7.3. Hereafter an assay is developed that enables to screen for adjuvant system perturbations, that could help to force the peroxisomal environment into conditions that are supportive for crystallization (see subsec. 11.3.2).

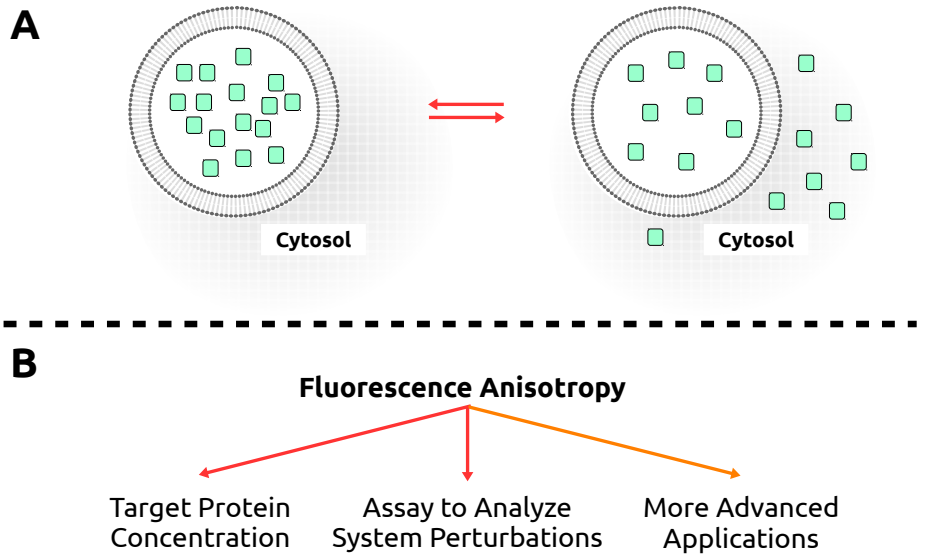

Figure 11.1: ***(A)*** *Schematic depiction of equal fluorescence intensity at different local in cellulo concentrations* ***(B)*** *Utilization of fluorescence anisotropy for different purposes*

In an additional chapter (see ch. 12) a broader application of fluorescence anisotropy in protein crystallography is briefly discussed in a *gedankenexperiment.*

11.1.1 Theoretical Framework for Protein Concentration Analysis by Fluorescence Anisotropy

This section aspires to lay a first theoretical framework for *in vivo* fluorescence anisotropy measurements, that focuses on protein concentration determination. The influence of viscosity and energy transfer on the anisotropy readout at increasing protein concentrations is explored. Subsequently, it is discussed how experimental data needs to be interpreted in order to judge on the development of intraperoxisomal protein concentration levels.
In order to derive a useful descriptive, theoretical model of the peroxisomal system, it is utterly important to reduce the considered complexity. Therefore the peroxisomal system is simply assumed to be of 'liposomic' character with no diffusional limitations for water and small solutes (fig. 11.2). Furthermore an instantaneous probe molecule import into the crystallization compartment

is assumed. Hence the signal contribution of cytosolic probe molecules can be neglected. In fact, it is known from microscopy data (sec. 11.2) that most of the target proteins are indeed located inside of peroxisomes. In addition it is assumed that probe import happens on much faster timescales than bilayer replenishment. This will unavoidably lead to an increase of concentration of the target protein. The probe molecule will be approximated as a hard sphere and concentration limits will be calculated accordingly. The sample is assumed to be homogenous.

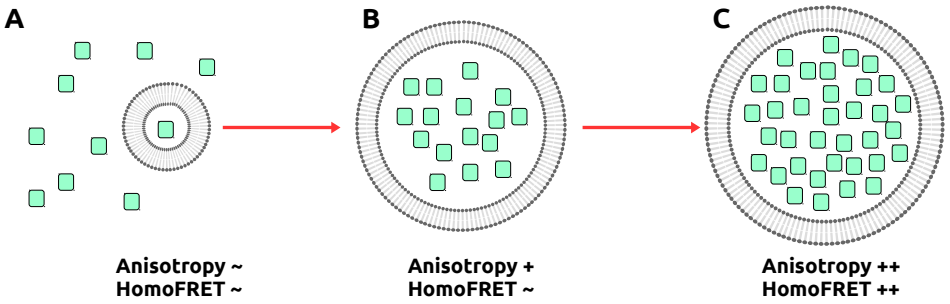

Figure 11.2: *Schematic representation of increasing probe molecule concentrations in an organelle; Concentration is increasing from A to C*

As schematically shown in figure 11.2 higher concentrations of probe molecules in such environment will give rise to a particle-dependent change of viscosity due to steric clashes that will increase the measured anisotropy. Likewise, the average distance between probe molecules will decrease dramatically at higher concentrations and give rise to homoFRET energy transfer, that in turn will decrease anisotropy. These effects will be investigated with regards to the probe volume packing fraction and the expected signal in steady-state measurements. For a realistic description molecule parameters are chosen that resemble the basic properties of EGFP as a probe molecule.

Deriving an Estimation of Steady-State Fluorescence Anisotropy Values for a Probe Molecule at Varying Volume Packing Fractions

From geometry we can derive an estimate of a globluar protein's relation of molecular weight to its radius, by incorporating an average protein density [145]:

$$r = \left(\frac{3M}{4\pi\rho}\right)^{1/3} \tag{11.1}$$

Hereby M denotes the molecular mass of the molecule and ρ is the mass density. The obtained radius usually needs to be extended by an estimate of the protein's hydration shell [144]. Hence a constant factor of 0.2 nm can be added to the calculated value. For the 30 kDa EGFP one can calculate an unhydrated radius of about 2.5 nm.

$$\frac{\eta}{\eta_0} = \left(1 - \frac{\delta}{\delta_{max}}\right)^{-2.5\delta_{max}} \tag{11.2}$$

Further on, the sphere model can be incorporated into the semi-empirical Krieger-Dougherty equation (eq. 11.2), which describes the relation of the concentration of globular hard spheres and visocsity (η) [199]. Logically, it can be easily adopted to describe the relation of medium viscosity and volume

fraction (δ) of probe molecules accumulating in peroxisomes. As an upper limit of volume fraction ($\delta_{\max}$) the *close packing of equal spheres* has been chosen (cp. fig. 11.3), which corresponds to a maximum of $\tilde{7}4\%$ in space occupancy. This is also approximately at the upper limit of protein content in crystals [12]. The lower limit represents free diffusion. Electrostatic interactions between the molecules are completely neglected in this approximation. Under the assumption of a maximum volume packing we can obtain for a 30 kDa globular protein of average mass density a maximum particle density of $1.115{\cdot}10^{25} particles \cdot m^{-3}$, which corresponds to 555 mg$\cdot mL^{-1}$, or 18.57 mM of probe molecule.

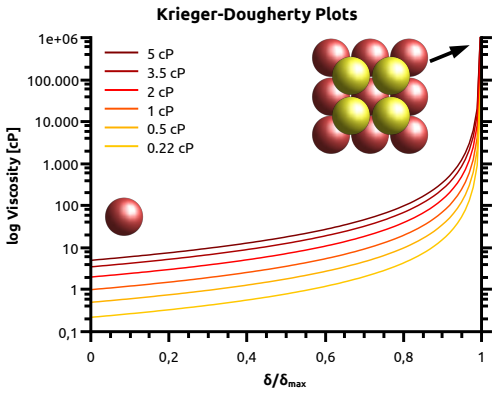

Figure 11.3: KRIEGER-DOUGHERTY *plots at different arbitrary η_0 values, ranging from 0.2 to 5 cP; $^{\delta}/_{\delta_{max}}$ equals a total space occupancy of 74% (close packing of equal spheres) as schematically depicted*

Figure 11.3 is derived from equation 11.2 and shows how viscosity changes upon an increase in concentration. The concentration is expressed as the relative volume packing fraction ($^{\delta}/_{\delta_{max}}$). With the relation of volume fraction to viscosity we can use the STOKES-EINSTEIN-DEBYE equation (eq. 11.3) to build up a relation to the rotational correlation time (θ_r) of the probe molecules:

$$\theta_r = \frac{1}{6D_r} = \frac{8\pi\eta r^3}{k_B T} \tag{11.3}$$

k_B is the BOLTZMANN constant, whereas T describes the temperature in Kelvin. From this it is possible to derive theoretical steady-state fluorescence anisotropy values by using the obtained correlation times in combination with the PERRIN equation (eq. 3.3). By insertion of eq. 11.2 and 11.3 into 3.3 the following dependency of r can be derived:

$$r(\delta) = r_0 \left(1 + \frac{\tau k_B T}{8\pi r^3 \eta_0 \left(1 - \frac{\delta}{\delta_{max}}\right)^{-2.5\delta_{max}}} \right)^{-1} \tag{11.4}$$

Since the KRIEGER-DOUGHERTY equation already accounts for a 3D distribution of molecules, the presented equation is valid for the theoretical peroxisome model. To obtain reasonable theoretical plotting data, it seems necessary to adopt EGFP standard values of different molecule properties from literature and incorporate them into the calculation. As stated above, EGFP weighs about 30 kDa and has a fluorescence lifetime (τ) of about 2.58 ns. Assuming a limiting anisotropy of 0.4 (400 mAns) and an initial viscosity of 0.22 cP

(arbitrary) one can plot the expected anisotropy readout at an increasing viscosity due to higher particle concentration as shown in figure 11.4. Viscosity was set to an arbitrary value to match the experimentally obtained value for EGFP, that is diffusing freely in aqueous solution (cp. fig. 11.9-A).

Table 11.1: *Descriptive plotting parameters; τand R_0 are EGFP-specific values, Protein mass density represents an average value*

Parameter	**Value**	**Ref.**
Temperature	298 K	-
k_B	$1.38{\cdot}10^{-23} J{\cdot}K^{-1}$	[164]
Protein Mass Density	$7.5 \cdot 10^{2} kg \cdot m^{-3}$	[145]
τ(EGFP)	$2.58 \cdot 10^{-9} s$	[65, 111]
R_0 (EGFP)	$4.7 \cdot 10^{-9} m$	[169]

These anisotropy values are fictional in the sense that so far energy transfer between the molecules has not been considered. Table 11.2 holds all values used for the plots depicted in fig. 11.4. As it can be seen the fluorescence lifetime has been varied for comparison of dynamic ranges potentially resulting from this parameter.

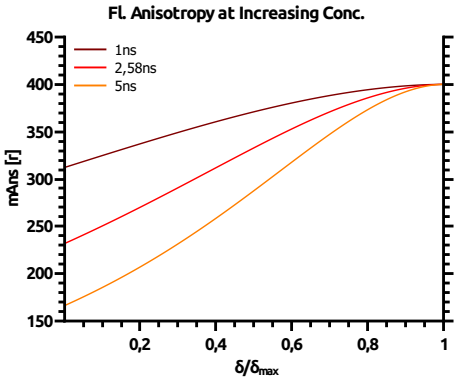

Figure 11.4: *Theoretical steady-state anisotropy of EGFP-like hard spheres at increasing volume packing fractions (δ) and with different fluorescence lifetimes; η_0=0.22 cP*

Deriving an Estimation of homoFRET between Probe Molecules at Varying Volume Packing Fractions

In order to connect the predicted anisotropy with the depolarizing effect arising from energy transfer between neighboring probe molecules (homoFRET), it is necessary to extend the generally applied FRET theory to 3D, as shown in section 3.2. Hence the quantum yield($^{\Phi}/_{\Phi_0}$) of an arbitrary energy donor molecule in a solution can be described as follows (cp. eq. 3.14):

$$\begin{aligned} \frac{\Phi}{\Phi_0} &= 1 - q\sqrt{\pi} \cdot exp\left(q^2\right)\left[1 - \Sigma(q)\right] \\ q &= \frac{\sqrt{\pi} V_F K_{pack}}{2V_M} \cdot \frac{\delta}{\delta_{max}} \end{aligned} \tag{11.5}$$

Hereby q originally describes the ratio of acceptor molecule concentration to a critical concentration C_0 which is calculated from the FÖRSTER radius (see eq. 3.15). For the presented data an isotropic energy transfer and a random orientation of probe molecules is assumed. Hence the more acceptors are in proximity to an excited donor, the stronger the emission will be depolarized and thus reducing the anisotropy readout. Volume packing fraction and concentration are related by $\delta = {}^{V_M N}/_{V}$. V_M is just the volume occupied by one molecule, N is the total number of molecules and V is the considered total volume. Furthermore, we can define $N_{max} = {}^{V \cdot K_{pack}}/_{V_M}$ and hence $\delta_{max} = {}^{N_{max} V_M}/_{V} = K_{pack}$, which yields:

$$\frac{\delta}{\delta_{max}} = \frac{\frac{N}{V} V_M}{K_{pack}} = C \frac{V_M}{K_{pack}} \tag{11.6}$$

K_{pack} is a packing parameter that defines the maximum packing fraction of a certain particle within a volume. N_{max} is the maximum number of particles in V, which occupy the total volume V_{max}. With the new definition of the variable q, equation 11.5 can be utilized to determine the overall homoFRET transfer efficiency at different volume packing fractions, according to [110]:

$$E_{FRET} = 1 - \frac{\Phi}{\Phi_0} \tag{11.7}$$

Considering all of the stated theoretical assumptions, it is possible to obtain an approximation of the energy transfer efficiency between probe molecules in a solute at varying volume fractions. In fact, the transfer efficiency from excited donor molecules is investigated under the assumption that no back-transfer of energy to the original donor is occurring. If the energy transfer rate (k_T) is much faster than the fluorescence lifetime (τ_0), which is likewise faster than the rotational correlation time, energy back-transfer would e.g. be a reasonable explanation for increased anisotropy values.
The plotting results are shown in figure 11.5. Three curves are shown to represent the energy transfer rate between EGFP like molecules with different FÖRSTER radii (R_0). If the FÖRSTER radius equals the reported value of 4.7 nm, transfer efficiency is already equal to 1 at a volume fraction of about 0.1. This corresponds to about 2 mM of protein. If the efficiency is reduced to yield a FÖRSTER radius that is somewhat equal to the molecule radius, an efficiency of 1 is only reached under maximum packing conditions. These results indicate a conclusive approximation of homoFRET in a solution. Therefore the obtained values can be applied as modifying factors in a holistic view of

anisotropy and energy transfer. Within figure 11.5 the horizontal dotted line indicates a complete energy transfer from the donor to surrounding acceptor molecules. An increase of transfer efficiency above 1 simply illustrates the increasing availability of acceptor molecules. In practical terms this can be interpreted as a complete radiation-less acceptance of donor energy, even if an unequal distribution of acceptor molecules is taken as a basis.

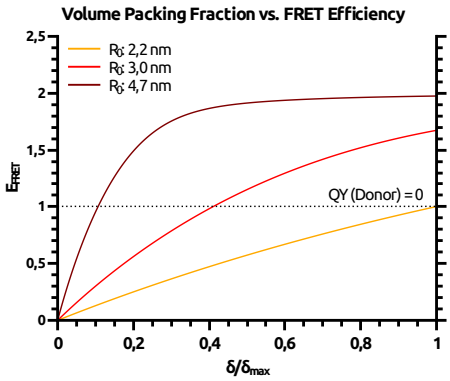

Figure 11.5: *Energy transfer efficiencies (E_{FRET}) in a solution of probe molecules at different volume packing fractions (δ) and with different* FÖRSTER *radii; Molecule Volume & fluorescence lifetime are left unaltered among the plots; The quantum yield (QY) from the donor is approaching 0, when E_{FRET}=1*

Influence of Probe Molecule Dimerization on Anisotropy & Energy Transfer

Most proteins exhibit oligomerization under specific conditions [101]. In case of EGFP it is known that dimerization occurs at higher concentrations with a K_D of 0.11 mM [109]. Conclusively, these dynamics have to be included into the presented formulas, to yield a more realistic behaviour of the model. The volume packing fraction corresponds to particle density and thus likewise to molar concentrations. Therefore, the reported K_D can easily be connected to probe behaviour at different volume packing fractions. From binding kinetics theory we can easily deduce the relation of monomer / dimer concentrations and dissociation constant (K_D):

$$K_D = \frac{[A]^x[B]^y}{[A_xB_y]} = \frac{[A]^2}{[A_1A_1]} \quad \leftrightarrow \quad \Sigma_b : \frac{[AA]}{[A]_T} = \frac{[A]}{K_D + [A]} \tag{11.8}$$

Σ_b denotes the fraction of molecule [A] being part of the dimeric complex [AA]. The related equation yields a hyperbolic curve for the formation of the dimeric fraction. Obviously such curve is only valid for the conditions under which the K_D has been determined and may vary dramatically if a differing pH or varying salt concentrations prevail. The formation of additional, higher oligomers is neglected. The dimer formation model can also be incorporated into the energy transfer efficiency calculations. The energy transfer in a dimeric probe molecule assembly can be calculated according to equation 3.19 in section 3.2:

$$\frac{\Phi}{\Phi_0} = \frac{1 + \tau \cdot k_T}{1 + 2\tau \cdot k_T} \tag{11.9}$$

The fraction Φ/Φ_0 is the quantum yield of the excited donor under dimeric conditions. F_T is the energy transfer rate proportional to $(R_0/R)^6$. In compliance with data from table 11.2 we can derive a quantum yield of a donor in a dimeric complex of about 0.71,when κ^2 equals $^2/_3$.
Figure 11.6-A/B shows the incorporation of the dimerization dynamics into the theoretical anisotropy and energy transfer model, that have been derived above. To emphasize the difference due to dimer formation a logarithmic depiction on the x axis has been chosen. The black curves represent data for a strictly dimeric protein, whereas yellow curves represent strictly monomeric data. Lastly the red curve is giving the theoretical readout under dimer formation conditions and the stated K_D. It should be noted that the direct space between the monomeric and dimeric curves in each plot represent the parameter space of the corresponding output variables.

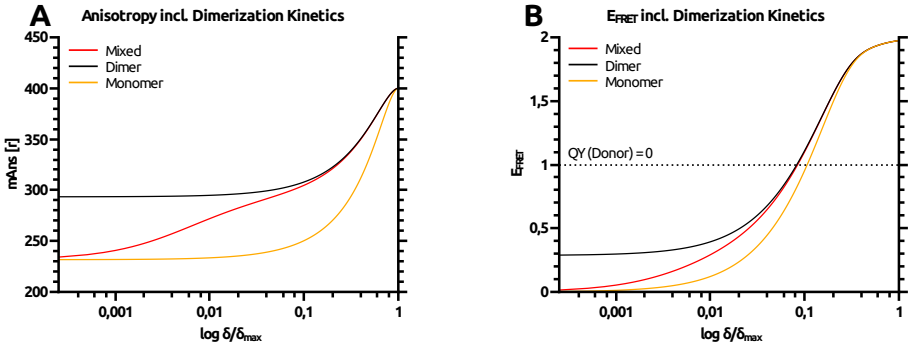

Figure 11.6: ***(A)*** *Steady-state anisotropy (r) and* ***(B)*** *energy transfer efficiency (E_{FRET}) at varying volume packing fractions (δ) in solution, including dimerization kinetics with a K_D=0.11mM; Mixed: Concentration-dependent dimerization*

Theoretical Probe Behaviour at Increasing Volume Packing Fractions Considering FRET & Dimerization

Considering the above derivations one can derive a combined theory of steady-state anisotropy and fluorescence energy transfer at increasing concentrations of intra-organellar probe molecules. The overall depolarization due to energy transfer depends on the orientation of the absorption dipoles of the acceptor molecules to the emission dipole of the donor molecules, which defines the total energy flux across molecules. The probability of a radiation-less energy transfer from a molecule is usually anisotropic and expressed by the orientation factor κ^2 (see sec. 3.1) for the case of FRET. Hence, the factor will ultimately determine the expected maximum depolarization. Here it is assumed that the energy transfer not only is isotropic, but also that it occurs with acceptor molecules in random orientations. Thus, complete energy transfer will lead to an emission signal with vanishing polarization. In such case the anisotropy reduces to zero.
Figure 11.7-A shows how the anisotropy signal develops at increasing volume packing fractions if homoFRET between molecules is considered. The influence of energy transfer on the anisotropy readout, denoted as *homoFRET potential* is plotted in figure 11.7-B accordingly. It simply describes the deviation of the anisotropy signal with homoFRET from the anisotropy signal without homoFRET and might be used as an independent measure of concentration. Data in A has been calculated via a linear combination of energy transfer and anisotropy.

The theoretical development of the anisotropy signal of strictly monomeric (yellow curve), strictly dimeric (black curve), and the dynamic dimer formation (red curve) cases are shown. Additionally, the anisotropy signal of dynamic dimer formation without energy transfer (blue curve) is shown for comparison. Negative anisotropy values are meaningless and originate from the higher potential of energy transfer (cp. fig. 11.5). The emitted fluorescence signal will remain depolarized in these cases.
Evidentially energy transfer has a strong depolarizing effect on the anisotropy readout and causes a steep decrease at increasing volume packing fractions. Consequently, the complete depolarization of the emission signal is reached at a volume fraction of 0.1, equally to plot 11.5 (R_0=4.7nm). Logically this corresponds to a homoFRET potential of 100%. It is interesting to note that the inclusion of the dimer formation does not yield a dramatically different theoretical readout signal. This would of course change in case higher oligomers would be considered, which would lead to a stronger depolarization at low concentrations. A circumstance that is shown in other studies to yield the number of molecules within a cluster [45, 113, 143]. At very low concentrations we can still find a distinct difference if dimer formation is included into calculations. Potential oligomerization enlarges the parameter space as mentioned before. If it is aspired to measure the behaviour of target proteins, especially in the

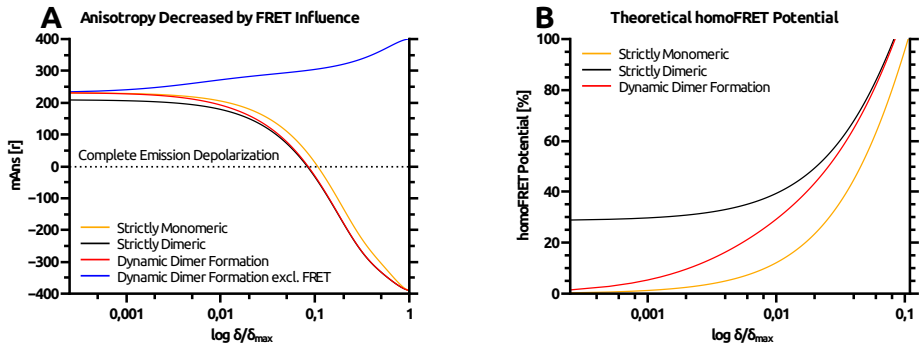

Figure 11.7: ***(A)****Steady-state anisotropy at different volume packing fractions (δ) including the depolarizing effect from energy transfer to neighboring molecules and* ***(B)*** *the arising homoFRET potential;* K_D*=0.11mM,* $\eta_0 = 0.22cP$*,* $\tau = 2.58ns$*,* $R_0 = 4.7nm$

light of *in vivo* crystallization, it seems very advantageous to utilize a methodological combination of anisotropy and FRET. Anisotropy will always reflect the interaction with other proteins or the general viscosity, whereas FRET can directly yield specific information about the proximity of target molecules. This leads to the possibility to differentiate between plausible concentration-related situations. For example, if the general concentration of proteins seem to increase, it could be possible that the concentration of target proteins in fact is decreasing. This is not necessarily captured by a simple fluorescence intensity readout. Furthermore theoretical calculations yield conclusive approximations of the ratio of target to other proteins in an organelle. If compared to experimental data, it is possible to gain detailed insights into a possible *in vivo* crystallization environment. To emphasize this idea, a second plot was created, which is depicted in figure 11.8. Besides the cases of dynamic dimer formation with/without the consideration of FRET, additional curves are shown. They represent the theoretical cases in which the general concentration of protein or solutes is still increasing, but the relative concentration of target proteins is reduced by a factor of 2,4,8 or 16 in the total volume fraction. This corresponds

to a total final space occupancy by target proteins ($\delta_{max,TP}$) of 37%, 18.5%, 9.3% and 4.6%, respectively. In case of 100% target proteins within a volume a space occupancy of 74% is assumed, as discussed above. These considerations are of course only valid, if proteins and solutes are randomly distributed. It is obvious that the parameter space of possible anisotropy values is extended to the whole area between the two extreme cases of full and none energy transfer. Values outside of these boundaries are also possible in case of higher oligomers or complexes with different proteins. In a real world experiment the

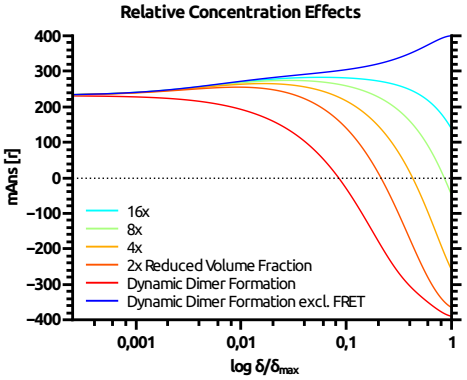

Figure 11.8: *Effect of reduced relative probe molecule concentrations in terms of volume packing fraction; 2x=37% $\delta_{max,TP}$, 4x=18.5% $\delta_{max,TP}$, 8x=9.3% $\delta_{max,TP}$, 16x=4.6% $\delta_{max,TP}$; K_D=0.11mM, $\eta_0 = 0.22cP$, $\tau = 2.58ns$, $R_0 = 4.7nm$*

anisotropy readout will always represent an ensemble, when measured with simple spectroscopic methods. This disadvantage can easily be eliminated by the usage of appropriately equipped microscopes. This is especially important when whole cell populations are investigated which might exhibit pronounced inhomogenities on the cellular level. Another important aspect of *in cellulo* ensemble measurements is the existence of secondary movements of organelles or cells. Fortunately the rotational diffusion of cellular compartments and cells in general can be neglected due to their much larger rotational correlation times in comparison to the fluorescence lifetime of most fluorophores.
The progression of fluorescence intensity is not incorporated into the theoretical considerations. It will show an approximate linear response to higher volume packing fractions.
This section presents a theoretical basis of the behaviour of fluorescent probe molecules, accumulating in 3D-lysosomic cellular structures. Possible dimerization kinetics and the additional import of other proteins are considered and define expectation intervals for anisotropy and energy transfer rates in *in vivo* experiments.

11.1.2 EGFP as a Responsive Probe in Steady-State Anisotropy

GFP and its variants have received a lot of attention as probes in sophisticated anisotropy and other fluorescence-based experiments in the past [55, 86, 143]. This is mostly due to the fluorescent properties that these molecules possess, their photostability, quantum yield and not least because they can be genetically encoded for a minimum invasive approach in biological systems [44].
Here it is shown that the *enhanced* GFP (EGFP) variant can be successfully used as a probe in steady-state anisotropy experiments, as proof is given for a link between rotational diffusion and an emission depolarization. More importantly it is demonstrated that these changes can be captured by the used setup and interpreted in useful ways. Thus, the shown *in vitro* experiments represent a prerequisite for the understanding of the following sections and are a logical continuation of the theoretical considerations in section 11.1.1. In detail this section shows (I) how the fluorescence anisotropy relates to changes in medium viscosity and (II) probe solubility.

Influence of Medium Viscosity

If not stated otherwise, all measurements were conducted equally with a sample volume of 200 μL and a plate reader focus depth of 23 mm at 20°C. For a more detailed description of the experimental setup and sample preparation see sec. ??.

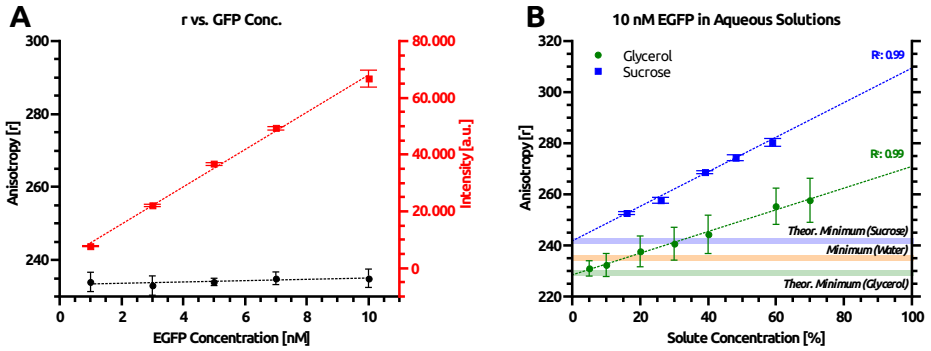

Figure 11.9: ***(A)*** *Nanomolar EGFP concentrations in aqueous buffer* ***(B)*** *10 nM EGFP at increasing solute concentrations; Temp.: 20°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 470 nm (f: 5 nm), $\lambda_{Em.}$: 520 nm (f: 5 nm)*

At first it was of particular interest to observe whether the combination of probe and setup yields the expected results. The setup response to pure EGFP[1] has been studied at a nanomolar range of 1-10 nM under aqueous buffer conditions (50 mM potassium phosphate, pH 7.4). The observed changes in fluorescence intensity and anisotropy are depicted in figure 11.9-A.
The readouts satisfy the expectation. The intensity response is increasing linearly from lower to higher concentrations (red data points). Likewise the measured anisotropy (black data points) exhibit the anticipated behaviour by remaining mainly unaltered. This reflects that at low concentrations no diffusional hindrance is introduced for the probe molecules, nor a spatial proximity is created, that would result in additional changes of the emission polarization (cp. sec. 3.2). In addition the measurements show a low error of 1-5% and hence indicating low systematic and random error during the experiment.

[1] Purified EGFP has been thankfully supplied by Zuzanna Pietras (Svergun group, EMBL Hamburg, Germany).

In order to elucidate the relation of emission polarization and medium viscosity, an *in vitro* experiment has been conducted that utilizes viscous solutes. As stated in section 3.1 higher viscosities can artificially represent binding and clustering events of molecules with regards to the anisotropy readout. Glycerol and sucrose were chosen as highly viscous solutes to maintain a series of different medium viscosities. Both reagents are extensively used for different purposes in biological research and thus were assumed to be mostly non-reactive towards the probe molecule. The probe concentration has been kept at a constant concentration of 10 nM during all measurements, to ensure a sufficient signal strength (data not shown). The result can be found in figure 11.9-B, in which a correlation between fluorescence emission polarization and medium viscosity is successfully described. As the viscosity of the medium increases, it is expected that probe molecules have larger translational and rotational correlation times according to the STOKES-EINSTEIN-DEBYE equation. Hence, the measured fluorescence anisotropy should be increasing when mixed with ever-increasing amounts of a viscous solute. Indeed for both solutes a positive, linear correlation can be found. Starting at a solute concentration of 0%, the aqueous

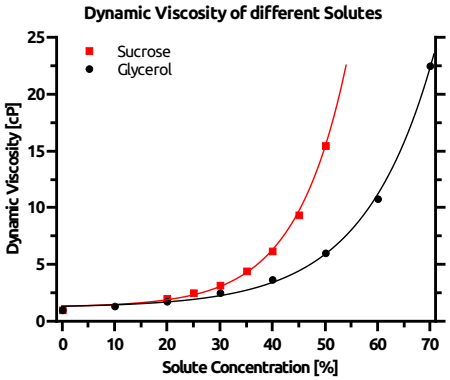

Figure 11.10: *Dynamic viscosities of increasing amounts of sucrose and glycerol in aqueous solution at RT; Data from [10, 97]*

probe buffer is assumed to have a dynamic viscosity of 1 cP ($mPa \cdot s$) at room temperature (20°C). For this probe environment a setup-specific anisotropy value of 234.2 ± 2.65 can be determined (fig. 11.9-A). Figure 11.9-B now clearly shows a linear deviation from this initial anisotropy value as EGFP is mixed in a dilution series, both with sucrose (blue data points) and glycerol (green data points).
For a complete understanding of these measurements it is important to take theoretical values of the dynamic viscosity of aqueous sucrose and glycerol solutions into account (fig. 11.10). From this data it is immediately obvious that the dependency of dynamic viscosity on the solute concentration clearly follows an exponential growth function in both cases. This is in agreement with the plotting of medium viscosity via the KRIEGER-DOUGHERTY equation (eq. 11.2, fig. 11.3). Intriguingly, a rather linear dependency is found between solute concentration and viscosity when monitored via EGFP anisotropy.
This behaviour is best understood by considering the mathematically description of both phenomenons. Anisotropy will respond with a saturating exponential decay curve towards increasing rotational correlation times, whereas medium viscosity shows an opposed exponential growth response towards solute concentrations. The result is a mostly linear response of anisotropy to increasing medium viscosities, as shown in figure 11.4. Such a linear response of EGFP

anisotropy can be independently verified for two different viscosity-increasing solutes. Hence the results prompt, that the probe is responding according to theory.
For both solutes a theoretical maximum fluorescence anisotropy value can inferred from a linear extrapolation to a solute concentration of 100%. The maximum values are approximately r=309 for sucrose and r=271 for glycerol. The maxima are not necessarily equal to the maximum anisotropy of the molecule which is solely determined by the difference in absorption- and emission dipole moments. Due to the higher viscosity of sucrose its maximum value will presumably be closer to the ultimate maximum and may serve as an experimental approximation of it. This also seems plausible as 100% glycerol exists in a liquid state at RT. Likewise the theoretical minimum anisotropy, which represents free diffusion of the particles, can be inferred from the y-intercepts. Apparently these values differ similarly for both solvents. For the obtained minimum values a comparison with a reference value from aqueous buffer conditions yields deviations of about 3%. For sucrose and glycerol one can obtain values of approximately r=242 and r=229, respectively. Besides a difference in y-intercepts the slope of the linear fits in figure 11.9-B deviates from each other.
The anisotropy readout can theoretically be modulated by other factors, like the fluorescence lifetime τof the fluorophore. If the fluorescence lifetime changes, the anisotropy readout will change as well. In case of an effect of a solute onto the fluorescence lifetime, the influence should non-linearly increase with solute concentration. Here such effects can be excluded to explain differences between the solutes, since in both cases the anisotropy response appears strictly linear, even at high solute concentrations. It may further be assumed, that solute-specific properties are responsible for the observed differences. Elastic scattering processes due to increasing turbidity could change the field polarization and hence influence the readout. Still this should also reveal itself in a non-linear dependency. Consequently it may be assumed, that the difference of the slopes arise simply from the fact, that unlike sucrose, glycerol will still show fluidity at a concentration of 100%.
Additional information can be obtained by looking at the excitation- & emission-spectra of EGFP in aqueous buffer and a solute. Such a photospectrometric experiment is shown in figure 11.11-A/B. Spectra were obtained under similar conditions and by the same plate reader setup as the previously discussed anisotropy data. For both types of spectra 1 nM EGFP has been mixed with either buffer, or buffer containing 50% (v/v) glycerol in a total volume of 200 µL. Excitation spectra were obtained with a constant emission readout at 512 nm $\pm 5nm$, while the sample was subject to an excitation scan between 420 to 500 nm. Vice versa the emission spectra were obtained by using a constant excitation at 488 nm $\pm 5nm$ and an emission scan between 512 to 650 nm. Wavelength scanning has been conducted by recording data every 2 nm (Δ_{Scan}) with a bandwidth (f) of 5 nm.
The results do not yield further evidence that the absorption and emission properties of EGFP were altered by the high concentration of solute. This assures the assumption, that the fluorophore is effectively shielded from the surroundings and yields reliable results. More importantly it can be validated that the recorded excitation and emission intensities are very comparable between buffer and glycerol-containing buffer. This suggests that the observed decrease/increase of fluorescence anisotropy in different solutes is not due to

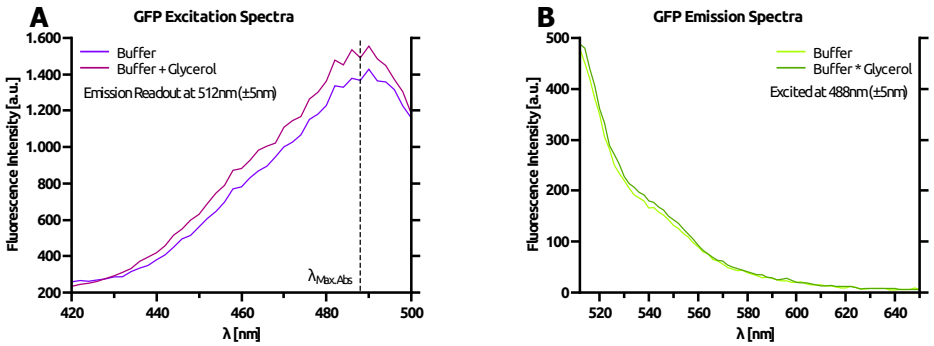

Figure 11.11: ***(A)*** *EGFP Excitation Spectra and* ***(B)*** *EGFP Emission Spectra with 1 nM protein in aqueous buffer with and without 50% (v/v) glycerol; Temp.: 20°C, Vol.: 200 µL, Detector gain: 140, Focus Depth: 23 mm, $\lambda_{Exc.}$: 488 nm (f: 5 nm)/ $\lambda_{Em.}$: 505-650 nm, $\lambda_{Exc.}$: 420-500 nm (f: 5 nm)/ $\lambda_{Em.}$: 512 nm, Δ_{Scan}: 2 nm (f: 5 nm)*

inelastic effects (cp. fig. 11.9-B), but rather elastic light-matter interactions that lead to a depolarization of the emission. Inelastic effects, described by a general extinction, might represent an additional explanation for an altered anisotropy readout at even higher solute concentrations. In general it was found that mixing of the probe molecule with glycerol was more prone to experimental error (cp. error bars in fig. 11.9-B). Most likely this is due to a more effective mixing of molecules with the sucrose solution. Both sucrose and glycerol have a kosmotropic effect in water solutions and hence should stabilize macromolecules. Hence, the larger experimental error for glycerol might be contributed to phase separation processes and incomplete mixing with the aqueous buffer.
For a better understanding it is worth mentioning that at very low solute concentrations both linear responses in figure 11.9-B will indeed be non-linear and approach the minimum value of the aqueous buffer.

Anisotropy Response to Probe Solubility / Dimerization

In a second step it is demonstrated how the anisotropy readout from freely diffusing EGFP molecules responds to changes in probe solubility. This was achieved by adding increasing amounts of sodium chloride to an EGFP-containing solution. EGFP was maintained in 50 mM potassium phosphate buffer at a constant concentration of 5 nM. Sodium chloride affects the availability of water molecules on the protein surface and hence influences protein solubility. Furthermore, the salt will influence the oligomerization state of EGFP by making hydrophobic interactions less favorable according to the HOFMEISTER series. By increasing the salt concentration the hydrophobic interactions are effectively weakened and the EGFP dimer with its approximate K_D of 0.11 mM can dissociate more easily. Additionally, peptides across the EGFP surface can rearrange and influence diffusion [118]. The experiment has to be analyzed very carefully, as both effects, dimer dissociation and change of solubility will occur simultaneously. Nonetheless, the rather high K_D indicates that most of the protein will be exist in a monomeric state at 5 nM.
The results of the experiment are shown in figure 11.12. For every measured time point an equilibrium situation is assumed. Red data points show the progression of fluorescence intensity, whilst black data points represent the change in fluorescence anisotropy. The intensity remains similar for all data points which indicates an intact fluorophore environment. This means, that

EGFP is not denatured, nor precipitated by the high salt concentrations, which would quench the fluorescence.
As it has been suspected beforehand, the fluorescence anisotropy does change upon incubation with different salt concentrations. The effect appears to be negatively correlated with the salt concentration and is best described by a biexponential decay curve. The amplitude contributions of the slow and fast component are 72.3% and 27.7%, respectively. Apparently different effects contribute to the change in anisotropy. If we assume a significant amount of dimeric protein, an induced monomerization would in fact lead to a rather monoexponential decay, because the rotational correlation time of a monomer will be shorter. Nonetheless the concentration of 5 nM does not imply an important role oligomerization kinetics. On the other hand an increased energy transfer between clustered molecules could lead to a decrease in anisotropy (cp. sec. 11.1.1). But bleaching experiments (fig. 11.13) also disprove this explanation. Consequently, it may be suspected, that the general change of

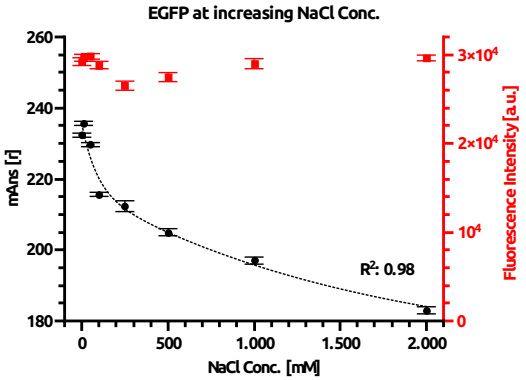

Figure 11.12: *5 nM EGFP at increasing NaCl concentrations; Temp.: 20°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 470 nm (f: 5 nm), $\lambda_{Em.}$: 520 nm (f: 5 nm)*

solubility and the rearrangement of surface groups is committal for the complex readout.
In an additional step a bleaching experiment has been conducted on the same EGFP solutions, to investigate the potential contribution of energy transfer. The results are depicted in figure 11.13-A. A detailed description of the bleaching procedure is given in subsection 11.1.3. The data represent the net change in anisotropy and intensity. The net change was always negative for the intensity. This is indicating a successful bleaching after 60 s, and further after 180 s. The net change in anisotropy is negative for all samples, but no significant difference can be found between the various salt concentrations. Only net changes of about 1-3% were found between bleached and unbleached data among all samples. Conclusively, this indicates that most proteins are neither in a dimeric state, nor clustered.
Indeed, it cannot be excluded that the probability of energy transfer in an EGFP dimer is very low due to the orientation of fluorophores. In such case, further verification e.g. via DLS would be required. Furthermore a change in the fluorescence lifetime can explain a decrease in anisotropy. But it seems very unlikely, since no apparent trend in the anisotropy bleaching data can be found.
The fundamental anisotropy of EGFP is in principle independent of oligomerization states, changes of the radius of gyration or the solubility. Hence, the approximation of the anisotropy to high solute concentrations should yield the

same limiting anisotropy. To determine its value, EGFP samples in 2 M NaCl and without NaCl, both subjected to bleaching, were additionally mixed with glycerol at different concentrations (see fig. 11.13-B). The obtained data is linearly fitted and extrapolated to indicate the maximum anisotropy at 100% glycerol for both cases. Apparently the data suggest a minor difference in the apparent fundamental anisotropy, with and without 2 M NaCl (blue arrows). This can imply slight changes of the fluorophores' environment, that in turn could alter the fluorescence lifetime. Due to the larger error in the different data points, the interpretation remains difficult. Hence even if high salt concentrations would influence the fluorophore itself, the effect still appears very minor and proofs the robustness of the probe molecule. Therefore, changes in anisotropy, which are not accompanied by changes in energy transfer, might be simply ascribed to changes in probe solubility (cp. subsec. 11.1.1). The speed of rotational motion of the probe molecules can be inferred from the slope of the data. For both experiments a linear fit to the ensemble data indicates different rotational correlation times.

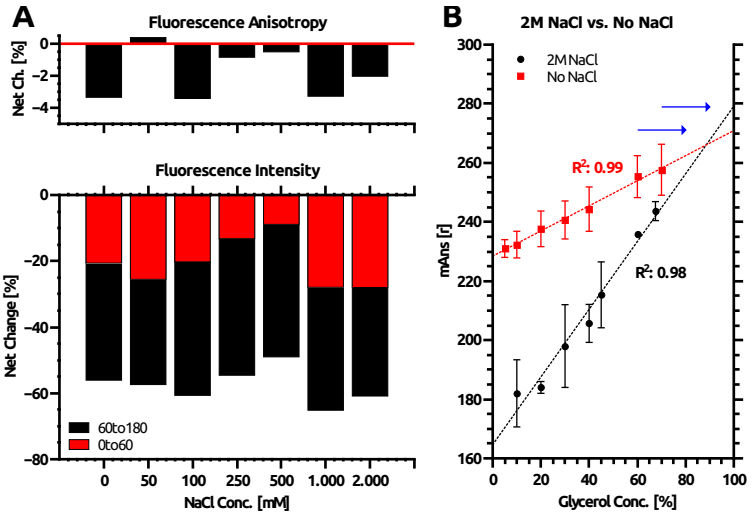

Figure 11.13: ***(A)*** *Total net change of fluorescence anisotropy and proportional change of intensity upon bleaching for 60 and 180 s* ***(B)*** *Bleached samples at different glycerol concentrations to approximate the fundamental EGFP anisotropy; Temp.: 20°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 470 nm (f: 5 nm), $\lambda_{Em.}$: 520 nm (f: 5 nm)*

As mentioned above, dimerization / monomerization processes have been excluded as an explanation for the probe behaviour. Intriguingly, results from Bader *et al.* [143] indicate a low homoFRET probability in EGFP dimers. Therefore, it is still possible that such processes are affecting the readout in this experiment. Of course, without being visible from the rate of energy transfer. In conclusion, it seems plausible to assign the fast and slow decay components of the anisotropy decay (fig. 11.12) to changes in probe solubility and changes in the radius of gyration as mentioned above.
In conclusion, the presented experiments verify the usability of EGFP as a probe for anisotropy measurements, to detect changes in viscosity or oligomerization. The gathered information can be used as a further point of reference for more complex *in vivo* studies.

11.1.3 Measurement of *In Cellulo* Anisotropy Data

This section provides a scientific proof of principle for *in cellulo* anisotropy measurements in the yeast *H. polymorpha*. The strain *H.p.*-12DAC4-EGFPskl has been used for cellular EGFP expression under the control of P_{AOX}. The probe protein carries a SKL-tag and is consequently imported into peroxisomes. For control experiments the strain 12DAC4-PRDX5 (sec. 6.1) has been used, which produces a non-fluorescent, SKL-tagged protein. Both tested cell types are identical in terms of their genome, except for the protein expressed under the control of P_{AOX}.
The yeast *H. polymorpha* possesses a rare methanol utilization pathway often reduced to the strong alcohol oxidase promoter (P_{AOX}), which finds widespread application in heterologous gene expression (see ch. 2). In scientific and industrial applications it is often necessary to screen for the highest level of protein expression under batch or fed-batch growth conditions in complex media. For such screens very often a GFP-fusion tag is used to enable for fast, plate reader- based fluorescence screening [120, 160]. GFP expression is monitored at an emission wavelength of 512 nm. The fluorophore is usually excited at about 488 nm.
Even though this method has been applied to methanol-induced *H. polymorpha* cells, it has been neglected so far that methanol growth conditions lead to a strong and independent induction of FAD producing enzymes, namely riboflavin synthetase, riboflavin kinase and FMN adenylyl transferase, which need to counteract on the strong demand for FAD as a cofactor for the enzyme alcohol oxidase. The activity of this pathway appears independent from the activity of P_{AOX} and the related expression levels, but decreases with extracellular FAD levels [134, 160].
Obtaining high protein yields by exploiting the P_{AOX} promoter requires the passivation of alcohol oxidase expression under methanol growth conditions [73, 239], which would usually not be accompanied by a passivation of the flavin-related enzymes. Under these conditions the GFP readout is strongly overlayed by background fluorescence, possibly originating from a combination of secreted flavins (Riboflavin, FMN, FAD) and vitamins as part of the complex media [252].
Unbiased fluorescence intensity and anisotropy data from living cells can be obtained by extensive washing as shown in this section. In general, data was obtained by conducting the following experimental steps. First, cell samples were produced according to sec. 6.2. Secondly a small fraction of cell material was harvested from the cell culture. Volume fractions were collected in different amounts to reach a cell density of OD 8 in 200 µL of final cell suspension. This was subjected to EGFP-specific intensity, anisotropy and spectral measurements to investigate the signal contribution from other fluorescent species or the cellular environment in general.

EGFP Emission is Overlayed by Other Fluorescent Species

Figure 11.14-A/B depicts excitation- and emissionspectra of EGFP- and non-EGFP expressing (control) yeast cells, before and after extensive washing with ultrapure water. In a direct comparison of the spectra in figure 11.14-B (continuous lines) a considerable conformity of the emission spectra of EGFP- and control cells is apparent. This also holds true for the excitation spectra in figure 11.14-A, and is even more obvious in a comparison of the

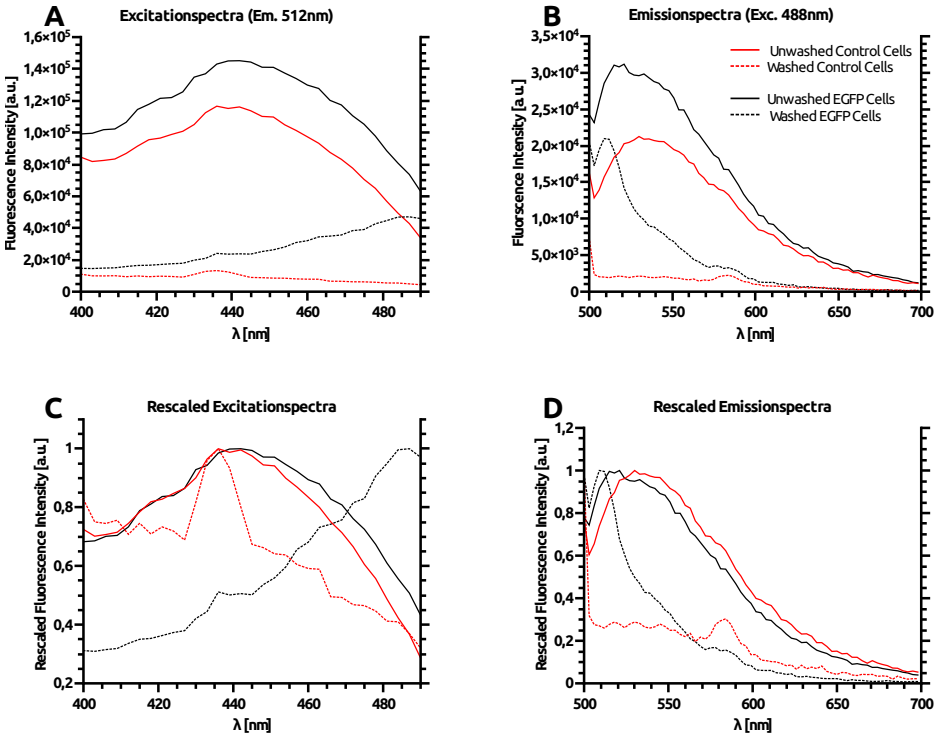

Figure 11.14: ***(A)*** *Excitation- and* ***(B)*** *Emissionspectra of EGFP-expressing & control cells, both washed with ultrapure water & unwashed (growth media)* ***(C)*** *Normalized Excitation- and* ***(D)*** *Emissionspectra; Temp.: 20°C, Vol.: 200 µL, Detector gain: 180, Focus Depth: 23 mm, $\lambda_{Exc.}$: 488 nm (f: 5 nm), $\lambda_{Em.}$: 505-700 nm, Δ_{Scan}: 2 nm (f: 5 nm)*

rescaled spectra shown in figures 11.14-C and -D, respectively. In case of the EGFP-expressing cells, a slight shift of the curves towards the pertinent EGFP emission maximum is revealing some minor contribution of EGFP to the signal, but without yielding conclusive information. Upon washing away the original growth media and resuspension of cells in water, a completely altered situation is making a case for the importance of the right sample treatment: In comparison of the spectra in 11.14-A/B (dashed lines) of the washed control cells, a strong intensity reduction can be found both excitation- and emission-wise. This clearly demonstrates the occurences of media-specific or secreted fluorescent species that interfere with the measurements. In case of the EGFP cells, one can find a strong decrease of signal intensity in those wavelength domains, that are naturally not relevant for light-matter interactions of the EGFP fluorophore. Instead clear EGFP-like excitation- and emission spectra emerge, enabling an unbiased readout of expression levels. At an emission wavelength of 512 nm, a signal reduction of 88% is found for the control cells, whereas for EGFP-expressing cells the signal is reduced by 32% in this experiment. The control cells show a very weak, but still visible background signal. This level of background seems unavoidable, as it probably originates from molecules populating the cells. But the remaining background should in general not be a problem for quantitative measurements as a distinct difference between EGFP- and control cells can be uncovered.
In consequence *in cellulo* measurements, both fluorescence- and anisotropy-wise, require a specific treatment of the cell suspension to yield reliable results. Either extensive preparation via washing of cells, or single-cell analysis (e.g. FACS) in diluted media is required, to exclude foreign fluorescence signals from the analysis.

Fortunately, re-suspension of yeast cells in water is not detrimental for cell integrity and survival. Nonetheless, it may introduce further experimental errors. A distinct analysis of how to circumvent extensive sample washing is presented in subsection 11.3.1. Still, here it will be used to obtain unbiased fluorescence signals from cells.
Figure 11.15-B shows time lapse samples from a cell culture with induced EGFP expression that were subjected to washing. As it can be seen a low background situation will favour correct EGFP fluorescence measurements. The obtained *in cellulo* EGFP spectra (e.g. fig. 11.15-B_{1050}) resemble the *in vitro* spectrum shown in figure 11.2, which is indicative for a mostly unbiased measurement of the protein's fluorescence intensity. Consequently, fluorescence anisotropy measurements will also benefit from this treatment.

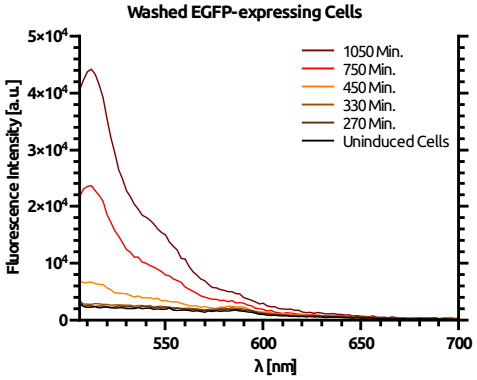

Figure 11.15: *Emissionspectra of EGFP-expressing cells at different time points during growth / P_{AOX} induction; Temp.: 20°C, Vol.: 200 µL, Detector gain: 180, Focus Depth: 23 mm, $\lambda_{Exc.}$: 488 nm (f: 5 nm), $\lambda_{Em.}$: 505-700 nm, Δ_{Scan}: 2 nm (f: 5 nm)*

For comparison an additional emission spectrum for washed cells without EGFP expression is shown in figure 11.15-B (black curve). This is further to evaluate the level of residual fluorescence signal originating from the cellular components. Washed cells still create a minor background signal, that does not change over time (cp. black curve vs. $B_{270/330}$). This background can be considered as a baseline that determines the detection limit. Its absolute value obviously scales with the applied setup gain and the optical density (fig. 11.17).

Procedure to Measure homoFRET Potential & Obtain Unbiased Anisotropy Data

In a next step it has been determined whether not only intensities, but rather correct polarization-related information in the form of fluorescence anisotropy can be obtained from an *in cellulo* sample. Cells were grown until early stationary phase under conditions that induce expression of EGFP and import into the peroxisomal environment as it is described in section 6.1. It was expected that the anisotropy signal is influenced by a homoFRET contribution as described in section 11.1.1 due to a high probe concentration. Hence, a bleaching experiment has been conducted at first, to artificially increase the average distance between functional emitters. This procedure was expected to yield unbiased anisotropy values [165]. The result of this experiment is shown in 11.16-B where decreasing intensity is plotted versus anisotropy. The bleaching setup is schematically shown in 11.16-A and the general approach is described

in section 7.3. Furthermore, 11.16-C depicts intact yeast cells that express EGFP, before and after bleaching procedure. The cells remain intact upon treatment and a decreased EGFP signal can be recognized from peroxisomal structures, verifying a successful experiment. Since bleaching was conducted without specific filter sets, cell samples were cooled during the procedure to reduce molecular damage by deposited heat. Intriguingly, a large homoFRET potential is found for the presented sample (cp. sec. 11.1.1). A systematic analysis and interpretation of this phenomenon for *in cellulo* data is conducted in section 11.1.4.

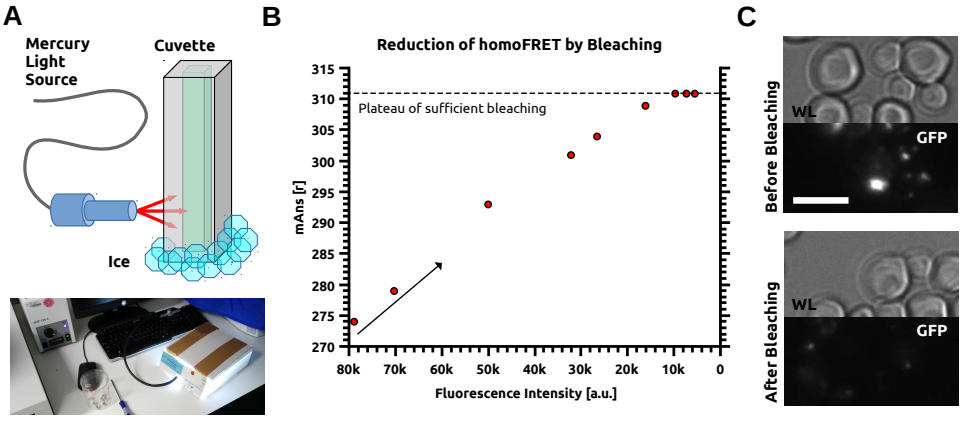

Figure 11.16: ***(A)*** *Bleaching of cell suspension in a cuvette by using a mercury light source* ***(B)*** *Reduction of in cellulo homoFRET between EGFP molecules during successive bleaching; Temp.: 20°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 470 nm (f: 5 nm), $\lambda_{Em.}$: 520 nm (f: 5 nm)* ***(C)*** *Yeast cells before and after bleaching procedure under a microscope; Exposure time: 200 ms, WL: White light illumination, GFP: GFP filter set; Scale bar: 2 µm*

As it can be seen in figure 11.16-B the applied procedure yields expected results. A decrease in general intensity is accompanied by an increase in anisotropy that is caused by a decrease of homoFRET contribution to the measured signal. After several bleaching steps a *plateau of sufficient bleaching* is reached. It exists beyond a point at which the average distance between responsive, functional emitters is much greater than the FÖRSTER distance. Still, the intensity is sufficient to obtain reliable data.

Influence of Scattering and Signal Intensity on the Anisotropy Readout

After bleaching, the cell sample has been used to evaluate the influence of scattering and the sensitivity of the setup. As it can be seen in figure 11.17 the anisotropy readout does depend on the optical density (cell density) and therefore, to some extent on the recorded intensity. The optical density was varied between OD 0 to 8. The obtained anisotropy data is best described by an bi-exponential fit as it becomes increasingly non-linear at lower ODs.
In general the correct anisotropy readout depends on the interplay of sufficient probe signal (intensity) and scattering from the cellular components. The anisotropy decreases linearly with higher optical densities. Since it is known that the anisotropy in general does not depend on the measured intensity (cp. fig. 11.9 & 11.13) the observed dependency might be attributed to a constant depolarization by random scattering from cellular components. Below an intensity threshold of 500-1000 (here OD: 1) the dependency is highly

non-linear. The anisotropy readout seems to be increasingly distorted by the limitations in signal detection. In fact a polarizing effect can be found that leads to an increase in anisotropy. This behaviour might be attributed to an increasing signal contribution of light scattered by the sample holder. This would be in agreement with the increasing error in anisotropy readout at lower intensities, underlining its random character.

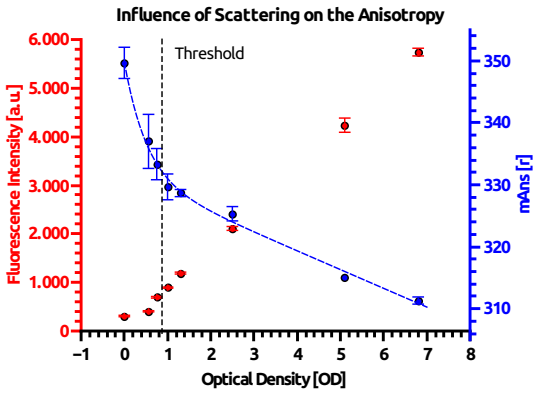

Figure 11.17: *Influence of scattering and fluorescence intensity on the anisotropy readout; Bleached cell samples diluted in ultrapure water; Temp.: 20°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 470 nm (f: 5 nm), $\lambda_{Em.}$: 520 nm (f: 5 nm)*

In conclusion, this points to a detection limit for *in cellulo* anisotropy data at intensity levels between 500-1000. This limit depends on the readout gain of the setup and hence is only valid when measurements are conducted under the same experimental conditions. It may shift to different values when other gain settings are applied. Still, an unexpected high anisotropy readout at 'in comparison' low intensity levels might be a good first hint on incorrect data. In contrast the fluorescence intensity decreases linear with the dilution until an OD of 0.5. Hence, this readout seems to be slightly more robust.
From these findings it can be concluded that the measurement of *in cellulo* fluorescence anisotropy data is feasible as long as the related intensity is maintained above a certain threshold. In case of denser cell suspensions an easy linear correction can account for the depolarizing effects of scattering events caused by cellular components. Correct anisotropy data can be obtained by randomly bleaching a subset of fluorophores in the peroxisomal environment. As a result additional information in the form of a homoFRET potential can be acquired. It can help to understand the local environment of the probe molecule as discussed in section 11.1.4.

11.1.4 Anisotropy-based Analysis of Intraperoxisomal Target Protein Concentrations in Yeast Batch Cultures

In subsection 11.1.3 it is demonstrated how *in cellulo* steady-state fluorescence anisotropy data can be obtained from a yeast cell culture expressing EGFP as a probe molecule. In practice these probe molecules are further imported into peroxisomes (cp. sec. 11.2) via the peroxisome-specific PTSI import pathway (see ch. 2). Accumulated probe molecules can reveal meaningful information about the intraperoxisomal environment. Unraveling the relative protein concentration and target protein proximity facilitates to determine the environment's ability to serve as a site for *in vivo* protein crystallization.
It also has been demonstrated in subsection 11.1.3 that three major readouts emerge under the applied conditions, namely, fluorescence intensity, fluorescence anisotropy and homoFRET potential. Those can be exploited with regards to the intended usage of the cellular compartment. Therefore an exemplary discussion of all tangible readouts is conducted in this section and combined with the knowledge obtained in subsections 11.1.1 and 11.1.2. Possible conclusions and interpretation-wise limitations will also be discussed.
Likewise to subsection 11.1.3 cells were grown under conditions that induce EGFP expression and peroxisome proliferation. Again, a detailed description of this can be found in section 6.1. Other than in the preceding subsection, cell samples were continuously harvested during growth (samples B-G). The optical density and the fluorescence intensity (subsec. 11.1.3) of three separately induced cultures were monitored. The obtained results are shown in figure 11.18-A/B. Both curves, representing the optical density and the fluorescence

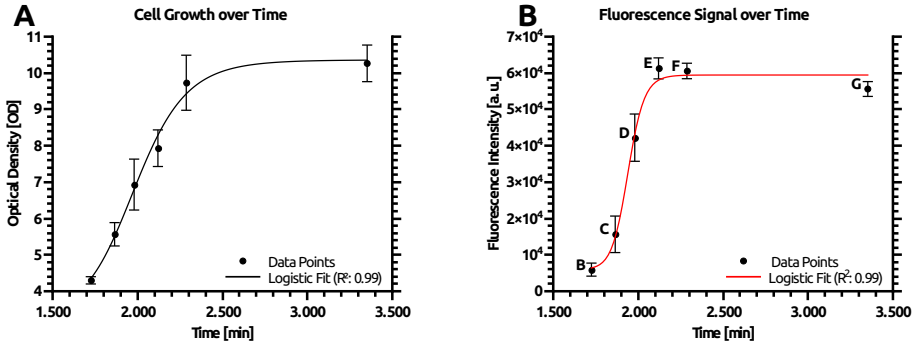

Figure 11.18: ***(A)*** *Yeast growth monitored via optical density (OD) and* ***(B)*** *Development of EGFP fluorescence intensity, both fitted with a logistic growth model; (B-specific) Temp.: 20°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 488 nm (f: 5 nm), $\lambda_{Em.}$: 512 nm (f: 5 nm)*

intensity, respectively, follow a logistic growth as it is expected for a batch culture. For a better understanding, it is worth mentioning that the cultures exhibited a diauxic growth. Hence, cells show two growth phases, here one on glucose without EGFP expression and one on methanol / glycerol with EGFP expression. Only the latter part is of interest. But during glucose depletion the P_{AOX} promoter is derepressed. This leads to minimum level of target protein expression before the logistic growth on methanol/glycerol is initiated. Samples B-G were collected during the second growth phase on methanol/glycerol. Sample B represents data at the very beginning of peroxisome proliferation. On the contrary, samples F and G represent data points in early and late stationary phase, respectively. The increase of fluorescence intensity coincides in a deferred manner with the optical density of the cell suspension. As discussed

in subsection 11.1.3 samples needed to get bleached to obtain unbiased values for anisotropy and a measure of the homoFRET potential. For this reason all three batch cultures were pooled together and subjected to bleaching. Figure 11.19 shows the results of the bleaching experiment. Data points are represented each by a bleaching curve that ends in a specific bleaching plateau. The differences between starting and ending value of the measured anisotropy represents the homoFRET potential as defined earlier. The obvious differences in fluorescence intensities from figure 11.18-B and from the starting points of the bleaching curves in 11.19 is due to setup-specific differences in detection in fluorescence and anisotropy mode.

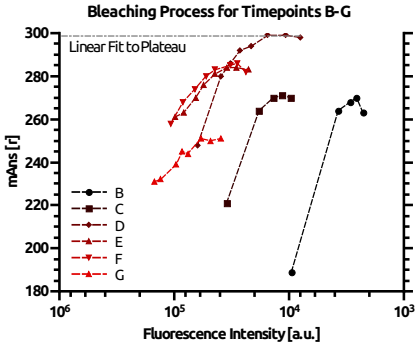

Figure 11.19: *Change of fluorescence anisotropy upon bleaching, in cell samples from different time points during logistic growth; Cells were cooled during bleaching; Temp.:∼15°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 470 nm (f: 5 nm), $\lambda_{Em.}$: 520 nm (f: 5 nm)*

The data can be plotted separately in terms of intensity and anisotropy and the relative changes in both readouts and of each data point can be extracted. This is done for all data points in figure 11.20-A/C. Figure 11.20-B depicts the resulting homoFRET potential of each sample, whereas figure 11.20-D depicts the intensity reduction necessary to reach the *plateau of sufficient bleaching* (see subsec. 11.1.3).

As it can be seen in sub-figure D the reduction in intensity is for all samples in a similar range of 60-70%, but slightly decreasing as indicated by the linear fit. This is in agreement with the assumption that less fluorescent emitters need to be bleached to reach the plateau of sufficient bleaching. In the theoretical peroxisomal system described in subsection 11.1.1, the intensity from equally distributed emitters is a measure of the average particle distance. Logically, this is because a higher intensity is due to more particles. If particles are not

Table 11.2: *Experimental Anisotropy and homoFRET data for time points B-G at ∼15°C; Anisotropy after bleaching (AB) has been determined by a horizontal fit to the plateau of sufficient bleaching; BB = before bleaching*

Time point	**Anisotropy (BB) [mAns]**	**Anisotropy (AB) [mAns]**	**homoFRET Pot. [%]**
B	189 ± 2	263 ± 4	28
C	221 ± 1	270 ± 1	18
D	248 ± 1	298 ± 1	17
E	261 ± 1	283 ± 1	8
F	258 ± 1	282 ± 1	9
G	231 ± 1	251 ± 2	8

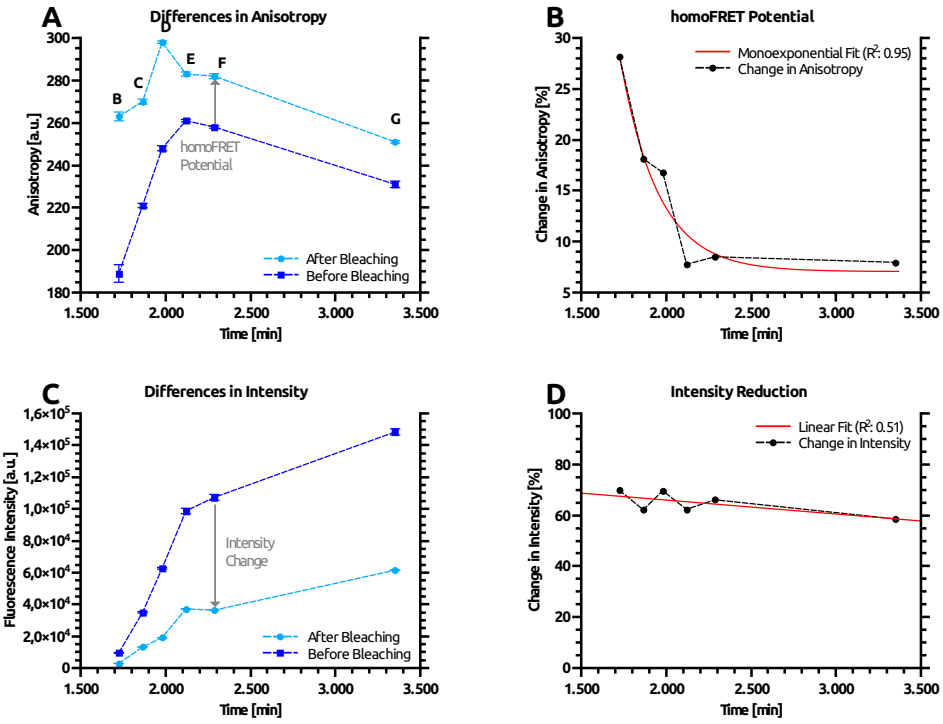

Figure 11.20: ***(A)*** *Change in fluorescence anisotropy upon bleaching in cell samples B-G and the* ***(B)*** *derived homoFRET potential* ***(C)*** *Differences in fluorescence intensity of cell samples B-G, before and after bleaching and a* ***(D)*** *plot of the derived relative changes; Data from fig. 11.19*

equally distributed, the intensity would not yield a mainly linear dependency. Here the intensity change that is needed to reach the *plateau of sufficient bleaching* seems to be negatively correlated with the general fluorescence intensity. Provided that all active emitters are located within peroxisomes, the best explanation for this observation is an increase in the average inter-particle distance of the probe molecules. The homoFRET potential should function as a more sensitive indicator for this hypothesis, since the effect of FRET is non-linearly decreasing with the particle distance. And indeed, a smooth, mono-exponential decay can be inferred from the data in figure 11.20-B, which is in agreement with a constantly increasing inter-particle distance (cp. subsec. 11.1.1).

Quite remarkably the recorded anisotropy values (after photobleaching) give a different picture. Values obtained for freely diffusing EGFP and EGFP at different solute concentrations, as shown in figures 11.9-A/B & 11.13, point towards a hindered rotational motion and a high viscosity, mainly for sample D,E and F. For the other samples at least a moderate hindrance can be inferred. The intraperoxisomal probe molecule movement is comparable to the molecule movement at glycerol concentrations of 20-80% at RT. The cellular environment could lead to a change of the emission polarization. But the recorded data was corrected for such distortion according to figure 11.17. In addition, the photobleached data did not exceed the experimentally determined intensity threshold for an unbiased anisotropy readout.

In a direct trend comparison of anisotropy values and homoFRET potential along data points B-G obvious differences can be seen. While the homoFRET potential is exponentially decreasing, the anisotropy is pursuing an up and down movement. This might be explained by means of the orientation factor

in homoFRET (sec. 3.1) that is interfering with the efficiency of the energy transfer. If certain molecule alignments are favored, e.g. in a crystal, this could lead either to an increase or a decrease in FRET efficiency. This could also explain the extremely high homoFRET potential of about 28% found for sample B. Especially data point D represents interesting data. The anisotropy reaches out to a value of 300 mAns, whereas the homoFRET potential is much lower as for the preceding data points. It seems possible that a far reaching alignment of molecules is observed. A general change in the probe molecule's environment could lead to such a change, too, but seems unlikely during exponential growth of the cells. Nonetheless the cellular system will show constant changes in terms of protein expression, general proliferation and solute concentrations in a batch growth. As it was shown in figure 11.7-A, also an increase in the concentration of other molecules can lead to the observed changes. Hence more complex cellular processes can still prove to be a logical source of explanation for the recorded data.

Comparison of Experimental and Theoretical Data

In a further step data has been recorded at 20°C and was compared to theoretical data from subsection 11.1.1. The results are depicted in figure 11.21 and presented in a plot of anisotropy versus homoFRET potential. The plot has to be interpreted in the following way: The origin of all lines (Ans.: 234, homoFRET: 0%) represents free diffusion of probe molecules at a low concentration. The higher the general concentration of particles, the higher the anisotropy. The specific concentration of probe molecules is represented by the homoFRET potential. If the homoFRET potential is increasing, it indicates an increasing concentration of probe molecules or an increased rate of interaction of such. Obviously the concentration of target proteins can be estimated by the help of the theoretical calculations, without considering the reported intensity values. Arrows 1 and 2 in figure 11.21 indicate an increase in the concentration of other molecules and the concentration of the probe molecule, respectively.

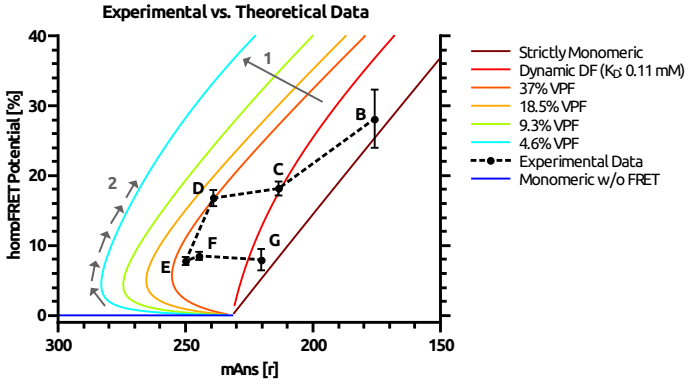

Figure 11.21: *Comparison of experimental and theoretical anisotropy/homoFRET data;* ***(Arrow 1)*** *Conc. increase of other molecules (viscosity)* ***(Arrow 2)*** *Conc. increase of target protein; Unbleached experimental anisotropy data from fig. 11.20; Theoretical data from subsection 11.1.1, VPF data considers dynamic dimer formation*

Intriguingly, the recorded data matches very well into the parametric space of theoretical values, but also reveals a very complex probe behaviour. Sample B represents an early developmental stage during P_{AOX} induction. The corresponding anisotropy and homoFRET data suggests, that the probe molecule

is mostly monomeric at a moderately high concentration. By using the information from the theoretical section an average protein concentration of 0.2-0.3 mM can be inferred. This corresponds to 6-9 mg·mL^{-1} and equals roughly to reported values of the average cytoplasmic protein concentration [189, 236]. Furthermore, the data suggest either a minor dimerization tendency, or the presence of a minor fraction of other proteins. The homoFRET potential is decreasing with later stages C-E. Likewise, the anisotropy is increasing, as it is predicted by the theory. Here an increasing anisotropy is representative of a higher viscosity in certain parts of the parametric space. If it is simultaneously occurring with a decrease in homoFRET potential, this strongly indicates a decrease in relative concentration of probe protein, accompanied by an increase in the concentration of other molecule species.
Data points D and E roughly indicate a 1:1 ratio of target proteins to other molecules, since they coincide with the corresponding line of 37% VPF. Additionally, data point E also represent the lowest observable concentration of probe molecules as indicated by the low homoFRET potential. For data points F and G the homoFRET potential remains mainly unchanged, which is reflecting a stable concentration of probe molecules. Likewise the anisotropy is decreasing, which could indicate an enlargement of the peroxisomal environment, or disposal of other molecule species. Since dimer formation and concentration are related factors it has to be assumed that both play a role in the development of the probe status. It remains difficult to dissect these effects.

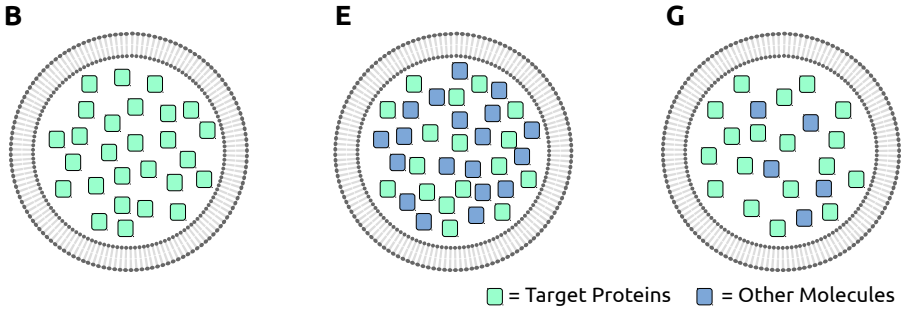

Figure 11.22: *Schematic representation of possible peroxisomal states at time points B, E and G (Compare fig. 11.21);* ***(B)*** *Moderately high target protein conc.* ***(E)*** *Increasing conc. of other molecules (viscosity), depletion of target protein* ***(G)*** *Degradation of other molecules*

In comparison of the experimental anisotropy/homoFRET data to the recorded intensity values (fig. 11.18) it is apparent that the increase of sheer fluorescence intensity shows a strong negative correlation with the decrease in homoFRET potential. This fact supports the interpretation, that the peroxisomal space is enlarging at higher rates than target protein is produced. During the stationary growth phase the concentration of target protein seems to remain stable, whereas other factors are changing and clearly lead to a lower viscosity in the probe environment. The theoretical assumption is, that probe solubility, e.g. influenced by salt concentration, is not tremendously changing in the peroxisomal environment, as cells will tend to maintain homeostasis. However, a direct proof of this assumption cannot be given in this study. Logically, it cannot be excluded, that some of the observed changes can at least partially be contributed to such changes.
The gathered knowledge can be presented in a simple model of the peroxisomal system. Figure 11.22 schematically shows average peroxisomes at time points

B, E and G, that would be in agreement with the above data. Indeed, the presented theory could also be too incomplete, due to unconsidered bias of the readout. But very good evidence exists, that FRET measurements are a reliable tool for sensitive proximity measurements [6, 7, 87]. Furthermore, it is conclusive that such an energy transfer leads to a depolarization and hence influences the anisotropy.
As mentioned before, the results indicate, that under peroxisome-proliferating conditions, the volume increase of peroxisomes exceeds the production of target proteins under the P_{AOX} promoter. This is so far of course only true for the specified conditions. Still, it will be proven in subsection 11.3 for various conditions. In fact, depending on the exact conditions, this could imply an unfavorable situation for crystallization. Considering the phase diagram of crystallography (see fig. 1.8) a rather fast passage through all possible zones could be likely. In such case the concentration required for nucleation would only exist for a very short time or simply never. The constant decrease in concentration would even make it more unlikely that crystals or nuclei could be retained. Possible intracellular precipitants like methanol could render the situation even more complicated and less controllable.

Perrin Plot Analysis

In order to obtain data comparable to experiments conducted with free EGFP, all bleaching samples were subjected to a Perrin plot analysis (see sec. 3.1). Such experiments allow to estimate the molecule's radius at different time points and can help to investigate possible precipitation or clustering. Furthermore, the apparent limiting anisotropy can be determined. In a Perrin plot T/η is plotted against $1/r$. For Perrin plots usually viscosity is altered at a constant temperature. Obvious constraints in accessibility of the cellular system only allowed for the alteration of temperature at a presumably constant viscosity. Thus, samples were measured at varying temperatures of 10-37°C, to not exceed the natural limits of the system. For an *in cellulo* Perrin plot analysis it is required to measure bleached samples to exclude effects from energy transfer.
Figure 11.23 depicts the results of the experiment. Each data point has been measured at 10, 25 and 37°C. A linear fit of the data indicates rotational speed via its slope and fundamental anisotropy via the y-intercept. The related data is shown in table 11.3.
Controversially, the y-intercepts differ for samples B-G. The fundamental anisotropy, should remain similar among all samples. A usual interpretation of such data would either incorporate the measurement at different wavelengths or different temperatures at comparable viscosities. Since samples were treated and measured equally these explanations seem not to apply. The fundamental anisotropy ranges here from about 261 to 319 mAns and is in good agreement with the approximated maximum values of EGFP in sucrose (309 mAns) and glycerol (271 mAns) shown in subsection 11.1.2. A fundamental anisotropy similar to *in vitro* experiments verifies that EGFP seems to be the origin of the detected signals. Smaller fluorescent species would most likely have smaller fluorescence lifetimes and much smaller rotational correlation times. Due to the unknown composition of the peroxisomal interior, it is plausible to assume that small differences in pH, salt & solute concentrations can explain small variations among the fundamental values.
The fitted slopes are a second important readout from the data and are listed

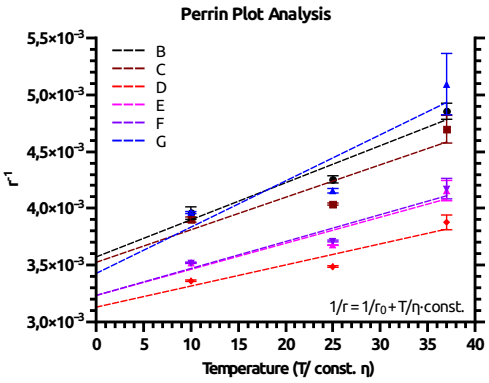

Figure 11.23: PERRIN *plot of samples B-G with data points measured at temperatures between 10 and 37°C.; All samples were allowed to equilibrate with the defined temperature for 10 min.; Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm,* $\lambda_{Exc.}$*: 470 nm (f: 5 nm),* $\lambda_{Em.}$*: 520 nm (f: 5 nm)*

Table 11.3: PERRIN *plot parameters, obtained from fits to the data in figure 11.23*

Sample	**y-Intercept** $[r^{-1}]$	**Slope** $[r^{-1}/(T/\eta)]$	R^2
B	$3.57 \pm 0.23 \cdot 10^{-3}$	$3.27 \pm 0.86 \cdot 10^{-5}$	0.94
C	$3.53 \pm 0.34 \cdot 10^{-3}$	$2.86 \pm 1.30 \cdot 10^{-5}$	0.83
D	$3.13 \pm 0.18 \cdot 10^{-3}$	$1.85 \pm 0.68 \cdot 10^{-5}$	0.88
E	$3.23 \pm 0.22 \cdot 10^{-3}$	$2.30 \pm 0.84 \cdot 10^{-5}$	0.88
F	$3.23 \pm 0.19 \cdot 10^{-3}$	$2.37 \pm 0.73 \cdot 10^{-5}$	0.91
G	$3.83 \pm 0.02 \cdot 10^{-3}$	$4.07 \pm 1.85 \cdot 10^{-5}$	0.83

in table 11.3. A larger slope indicates a faster molecule rotation. Here all slopes are within the same order of magnitude with only small variation. The fastest rotation is found for sample G (late stationary phase) and sample B (induction), whereas the slowest rotation is found for sample D (growth phase). This is in general agreement with the anisotropy values from figure 11.20-A, but not with the homoFRET potentials in figure 11.20-B. Similar to the *in vitro* results in figure 11.13-B the y-intercept and slope are negatively correlated. A PPMCC of -0.98 is indicating a direct relation in the *in cellulo* environment. It is important to mention that a simple change of intraperoxisomal viscosity between different samples would not be accompanied by a change in fundamental anisotropy. The *in vitro* experiments with free EGFP only hint on a very minor influence of extremely high salt concentrations on the fluorescence lifetime. But the obvious correlation of y-intercept and slope definitely indicates a change in protein solubility.

A holistic view of the presented data still supports the model presented in figure 11.22. Presumably, densely packed peroxisomes exist upon P_{AOX} induction with regards to the target protein concentration. This is indicated by the increased energy transfer between probe molecules, which can be perceived as a very robust readout. It is also during induction that the EGFP molecule shows increased rotational motion and a lower fundamental anisotropy. Somewhat contradictory to a densely packed peroxisome, this is indicative for environmental factors that favor a monomeric probe molecule. During late stationary phase, again a faster rotation and a lower fundamental anisotropy are found. It seems that the peroxisomal interior does not hold stable ionic conditions.

Furthermore, the ratio of target protein to other peroxisomal proteins (excluding AOX, CAT, DHAS) seems to decrease as implied by the decrease in energy transfer. No evidence for the clustering or crystallization of a larger molecule fraction can be found, as the rotational diffusion does not change substantially. In total this might lead to the conclusion that peroxisomes render an unfavorable environment for *in vivo* crystallization under the applied conditions.
Intriguingly, the homoFRET contribution seems to be the predominant readout of the anisotropy value (without photobleaching). The unbleached anisotropy data from figure 11.20-A shows a strong negative correlation with the associated homoFRET potential values as it can be expected from the depolarizing influence of random energy transfers. A PPMCC of -0.85 can be calculated for the unbleached data with the homoFRET potential, whereas a value of only -0.10 is found for data after photobleaching. Hence it can be conjectured that the unbleached anisotropy readout might be used as an indicator of probe molecule proximity under different growth conditions.

11.2 Localization & Distribution Analysis of SKL-tagged EGFP via Fluorescence Microscopy

This section presents data obtained via fluorescence microscopy from living wildtype *H. polymorpha* cells and is thought to complement the data from section 11.1. Yeast cells were grown according to section 6.2 and samples were collected similarly to subsection 11.1.4 for all different phases of logistic growth in batch cultivation (see fig. 11.24). A strain expressing EGFPskl has been utilized to efficiently label peroxisomes during peroxisome proliferation. The EGFP protein accumulates within peroxisomes [82, 136] and is used as a marker to determine the approximate size distribution of peroxisomes in each data point. Peroxisomes are also denoted as region of interest (ROI) throughout this section.

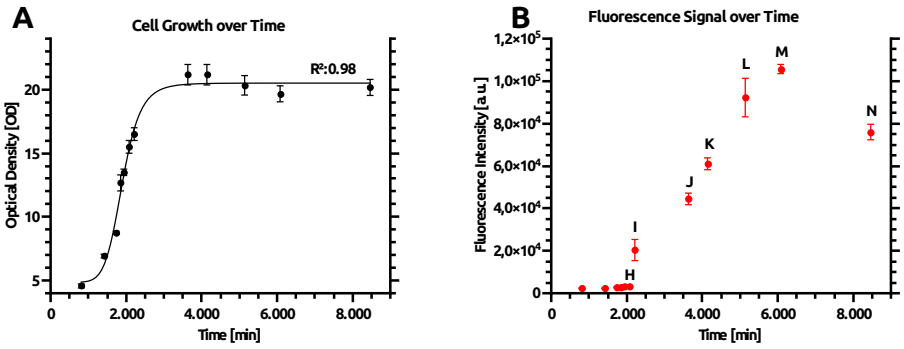

Figure 11.24: ***(A)*** *Yeast growth monitored via optical density (OD) and* ***(B)*** *Development of EGFP fluorescence intensity, both approx. following a logistic growth model; (B-specific) Temp.: 20°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 488 nm (f: 5 nm), $\lambda_{Em.}$: 512 nm (f: 5 nm)*

The applied procedure to obtain such a size distribution is described in section 7.1 and comprises the following steps: (I) recording of z-stacks, (II) bleach correction, (III) deconvolution, (VI) edge detection and (VII) binary image analysis. Data recording & processing had to be done manually since no automatized software had been available. For all processing steps the open-source software suite *Fiji* [178] has been used. Size determination was of course limited by the achievable resolution of the setup. A pixel edge length on the camera corresponded to an approximate distance of 65 nm. Logically, a fluorescence signal confined to such pixel after post-processing represents the minimum area of a detectable peroxisome. When probe molecules accumulate in a spherical volume, the 3D-shape dictates a disparate distribution of signal intensity when reduced to two dimensions. Hence, it can be assumed that peroxisome sizes are systematically underestimated.
Additionally the underlying intensities were extracted from averaged images of the raw data stacks. ROIs were defined by the borders found in volume analysis and intensities were extracted accordingly. The obtained values were weighted with a factor to compensate for different exposure times and differences in stack recording. The exposure times for different data points were chosen to be within the dynamic range of the camera. Weighting factors were calculated from the average intensity/pixel in areas without specimen. Table 11.4 shows a list of exposure times and the applied weighting factors (CC Factors).
The obtained data was subjected to a statistical analysis. Figures 11.25-A/B and -C/D summarize data obtained for different time points for volumes and

Table 11.4: *Exposure times for microscopical image recording and intensity correction factors for samples H-N*

Sample	Exposure Time [ms]	Ratio Int./Px.	CC Factor
H	1000	2115.7	3.2
I	500	2601.2	2.6
J	350	2918.1	2.3
K	250	3559.6	1.9
L	200	3342.2	2.0
M	200	3898.3	1.7
N	200	6690.9	1.0

intensities, respectively. A box plot with [25,75]-percentiles and [5,95]-whiskers is used in figure 11.25-A and -C to visualize the raw data. Normalized frequency distributions of some exemplary time points are shown in figure 11.25-B and -D. Denotations H-N are equivalent to consecutive time points as indicated in figure 11.24 and listed in table 11.5.
The size-related data roughly resembles the DLS data from figure 10.5 in section 10.2. But as mentioned above a slight underestimation of peroxisomal volumes is likely in this case. Rather than an average state, the development of the size distribution during cell growth was of interest at this point.
As it can be seen the data indicates mostly uniformly-shaped size distributions for peroxisomes among different developmental stages during batch growth. Only data point H shows a quite different size distribution. It represents an early developmental stage during peroxisome proliferation. Even though the maximum value of the size distributions is clearly increasing with later developmental stages of the cultures (data not shown) the general shape and median point of the distributions do not follow an obvious trend.

Table 11.5: *Overview of radius and related statistical data of labeled peroxisomes; N(C)= Number cells, N(P)= Number of peroxisomes*

Sample	Time [min]	N (C)	N (P)	Mean [nm]	StDev [nm]	Median [nm]	Max [nm]
H	2080	481	208	83	51	64	408
I	2215	289	427	143	69	137	392
J	3635	246	550	165	85	156	460
K	4140	439	1327	150	87	132	477
L	5140	307	1397	162	99	151	569
M	6080	396	1865	150	84	137	621
N	8465	600	3828	141	83	127	862

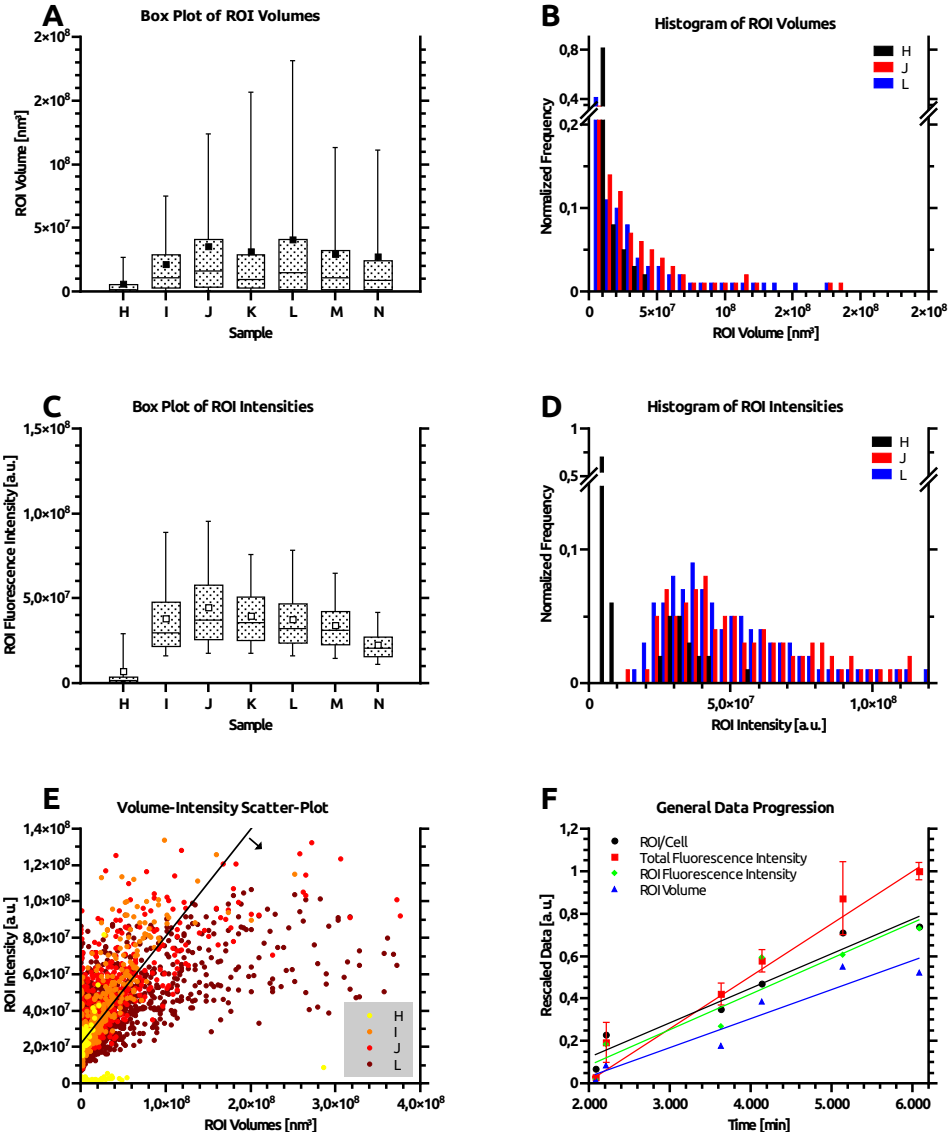

Figure 11.25: ***(A)*** *Box plot of ROI volumes (all time points)* ***(B)*** *Histogram of ROI volumes (time points H,J,L)* ***(C)*** *Box plot ROI intensities (all time points)* ***(D)*** *Histogram of ROI intensites (time points H,J,L)* ***(E)*** *Scatter plot of ROI volumes vs intensities (time points H-J,L)* ***(F)*** *Normalized and averaged data of ROI total volume, ROI total intensity, total ensemble intensity and the number of ROIs per cell (time points H-N); Related data can be found in tables 11.5 & 11.4*

The size distribution gets most expanded for early to mid stationary phase (J-L) and seems to tighten in late stationary phase (N). A trend best followed by looking at the whiskers of the box plots, but which is barely reproduced by the mean and median of the data (cp. table 11.5). In general, the size distribution seems to resemble a Poisson-distribution for most stages.
After the induction of peroxisome proliferation the related size distribution is narrow (H) and most ROIs have apparent volumes of about $6 \cdot 10^6 nm^3$, corresponding to a radius of about $80 \pm 50nm$. It is known that peroxisomes will be generated by *de novo* formation at this stage [196, 208, 213], which is also reflected by the low frequency in peroxisomes/cell (cp. fig. 11.26). Therefore, it is plausible to find mostly small organelles. Due to limitations in detection a broader diversification of this fraction cannot be verified nor

falsified. Short after the induction the volume distribution seems to stabilize. At time points I-N the mean volume fluctuates between $20 - 40 \cdot 10^6 nm^3$, which corresponds to a radius of 140-160 nm and a broad standard deviation of 60-80 nm. Obviously, peroxisomes increase in size short after induction of peroxisome proliferation, as indicated by the jump between data points H and I. Still, except for a minor broadening, the volume distribution remains mostly unaltered during different developmental stages.
Interpretation-wise, compatible distribution patterns emerge for the corrected intensities of ROIs. Likewise to the volume distribution time point H represents an outlier. Here it has a narrow, bimodal distribution pattern (cp. fig. 11.25-D/E). Bimodality might be ascribed to a combination of *de novo* synthesis and fission in the data, as it disappears at later time points. In fact, later time points show a defined intensity minimum for the smallest detectable volume. This is evidential for fission, that only generates ROI with a certain minimum of target proteins. An increase in intensity/ROI is apparent between data points H-J in figure 11.25-C. Afterwards intensity seems to slowly decline, even though target protein production is still ongoing (cp. fig. 11.24-B & 11.25-F). Distributions appear relatively uniform for time points I-M and are best described by a normal distribution which further shows a skewness to the right.

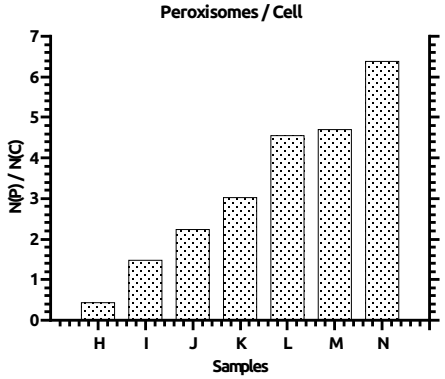

Figure 11.26: *Number of ROIs per cell, calculated from data in table 11.4*

If data for intensities and volumes are summarized for each time point, an easy and appealing comparison of the 'in total' volumes and intensities can be generated. The data is normalized with itself and compared to the general increase in fluorescence, which was obtained from plate reader measurements. The data is shown in figure 11.25-F (blue, green) and its progression in time is linearly approximated. A PPMCC of 0.97 can be calculated, which represents the average correlation between both parameters, ROI intensity and volume increase, respectively. The good correlation is also reflected by the parallelism of the linear fits in figure 11.25-F, and the straight black line in the scatter plot in figure 11.25-E. Obviously the ROI intensities scale with the ROI volumes or vice versa.
Nonetheless, these average values only reflect part of the truth. As cellular development is progressing, a widespread sub-fraction of ROIs is manifesting with deviating volume/intensity ratios. The older the cell culture gets, the more enlarged ROIs can be found without further increase in fluorescence intensity (cp. scatter plot in fig. 11.25-E). Furthermore, sections representing small peroxisomes with relatively high concentrations are getting slighlty depopulated. This reflects a decrease in average target protein concentration

within peroxisomes that fits well to the data obtained in subsection 11.1.4. Intriguingly, the overall fluorescence intensity from the cell interior is increasing at a higher rate and for prolonged times, than the intensity within peroxisomes (red line in fig. 11.25-F). Consequently a higher cytosolic content of target protein at later developmental stages can be found (cp. fig. 11.27). In conclusion, this might render a situation in which peroxisomal uptake of target proteins is not possible anymore, e.g. due to a lack of energization. On the contrary, lipid replenishment and general target protein production still seem to work at this point. It is evidential from the scatter plot in figure 11.25-E and figure 11.25-F, that peroxisomal target protein concentrations is not increasing beyond a vaguely defined limit.

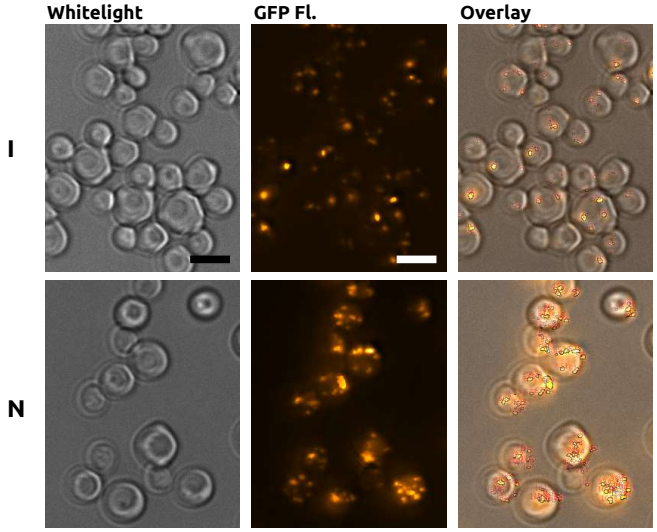

Figure 11.27: *Exemplary images of yeast cells in white light and GFP fluorescence mode for time points I and N; Fluorescence images represent a z-stack average with individual exposure times of 500 ms (I) and 200 ms (N); Overlay includes fitted outlines of ROIs as thin black lines; Scale bar: 2 µm*

In fact, cells counteract on an increase of target protein production mostly by an increase of peroxisomes per cell, and not by an increase in intraperoxisomal concentration. This is obvious from the data in table 11.5 and figure 11.26. It is logical to assume that new peroxisomes form mainly via fission at later stages due to the form of the distribution functions. Constant fission would be consistent with a stable radius distribution and a constant increase in the number of peroxisomes. It is clear that new lipids need to be targeted to those organelles in order to increase their size for a new fission cycle. The increase of cytoplasmic target protein does not seem not to be too pronounced. Hence influences on the ensemble measurements in the preceding subsection seem unlikely.
It is not well understood which processes might trigger the fission of peroxisomes. Defining a dependence of fission on the general intraperoxisomal constituent concentrations lies beyond the scope of this study. But without knowing the exact coherences, the consequences are a limitation in target protein concentration and a decrease in such for later developmental stages, consistent with data obtained in section 11.1 and the model presented in figure 11.22. An increased rate of pexophagy cannot be verified for any time point as the central vacuole is not highlighted by the fluorescence signal of the target protein.

The conclusions from this section are also likely to apply for peroxisomes in ΔPEX11 cells, that constantly increase in size [147]. Due to a lack of fission, peroxisome-specific lipids are redirected to the only available organelle. Data in this section suggest that intraperoxisomal target protein content and peroxisomal volumes are correlated, whereas the general production of target protein is not well correlated. This indicates a volume-dependent limit, rather than a limit due to a lack of energy. Presumably, peroxisomes in ΔPEX11 cells will also be bound to such limits for the target protein concentration. Interestingly ΔPEX11 cells can have more than one peroxisome per cell. In such case limited uptake into the sole peroxisome might trigger generation of new peroxisomes via *de novo* formation.

11.3 Development & Application of an In Vivo Crystallization Condition Screen

The measurement of *in cellulo* fluorescence anisotropy provides a reasonable readout to determine the relative progression of intraperoxisomal concentrations of fluorescent probe proteins. The general feasibility of this approach has been demonstrated by utilizing an EGFPskl construct (*H.p.*-12DAC4-EGFPskl) in sections 11.1 and 11.2.
Key focus of this section is to take advantage of the established methods and strains to determine intraperoxisomal probe molecule concentrations under the influence of cellular growth perturbations. These are thought to lead to possible deviations from the strong correlation of probe molecule concentration and peroxisomal volume. Hence they could render new conditions that lead to successful *in vivo* protein crystallization. Perturbations were introduced as environmental stimuli in form of chemical compounds.
At the beginning of this section a new method is introduced to perform *in situ* bleaching of excreted fluorescent species [134, 252]. It obviates elaborate cell washing (cp. subsec. 11.1.3) and reduces the expenditure of time spend for sample preparation.

11.3.1 *In situ* Quenching of Flavin-like Species Allows for Easy Fluorescence Readout

In this section it is shown how to easily perform *in situ* quenching of fluorescent species in yeast media. As discussed in section 11.1, extracellular fluorescent species are obstructing an unbiased EGFP emission readout from growing yeast cells. The presented method is based on non-radiative deactivation of excited molecules. Reducing agents and halogenic compounds can act as static/dynamic fluorescence quenchers, depending on the conditions in solution. It is consequential to test different molecules, all known for their quenching abilities, for their interaction with the complex media solution and intact yeast cells. This is to overcome the time-consuming and error-prone washing of cells. The intracellular fluorescence emission of EGFP remains unbiased. The presented procedure allows for a reliable and easy EGFP-based expression screening, that can detect even low expression yields at high dilution rates by reducing the background signal. Likewise intracellular anisotropy data can be obtained.

Quenching Ability of Different Compounds Tested on Cell Supernatant

Possible quenching agents were selected according to availability, eco-friendliness and known interactions with biomolecules. An overview of selected agents can be found in table 11.6. The substances TCEP, β-ME and DTT are tested due to their widespread application as reducing agents in biological science. Furthermore ascorbic acid and KBr are also known for their stronger redox potential and thus are tested as possible efficient quenching agents. In fact, ascorbic acid is also chosen as itself is fundamental to life, and hence might not negatively influence the cellular environment. KBr has no known biological function. Therefore an absence of adverse effects on biological material can be suspected.
The quenching effect of these substances can depend on pH and the type of

molecule to be quenched. In general it can be expected that a smaller molecule acts more efficiently than a large molecule, simply due to steric reasons. Moreover, the redox potential has a major effect on the results. It varies among different molecules and experimental conditions. From the known standard redox potentials, one can assume that KBr performs best as a quenching agent, followed by DTT (see table 11.6).

Table 11.6: *Applied quenching agents, Standard redox potentials (E^0) according to [261]*

Compound	**E^0, pH 7 [V]**	**Function**
L-Ascorbic Acid	-0.07	Vitamin, Antioxidant
β-ME	-0.26	Reducing agent; Cleavage of disulfide bonds
TCEP	-0.29	Reducing agent; Cleavage of disulfide bonds
DTT	-0.33	Reducing agent; Cleavage of disulfide bonds
KBr	+1.07	No known function

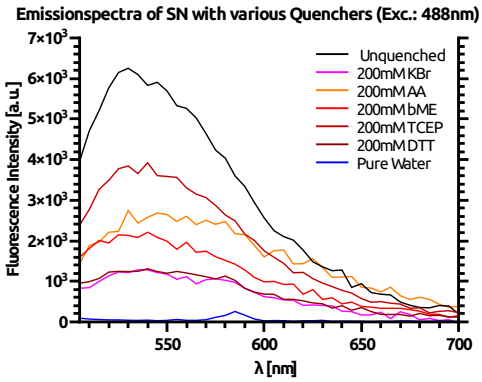

Figure 11.28: *Emission spectra of solutions of quenching agents mixed 1:1 with control cell supernatant; Temp.: 20°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 488 nm (f: 5 nm), $\lambda_{Em.}$: 505-700 nm, Δ_{Scan}: 2 nm (f: 5 nm)*

In figure 11.28 emission spectra are depicted, which were recorded between 510-700 nm, while the excitation wavelength is kept fixed at 488 nm. The spectra represent the corresponding quenching abilities of different quenching agents at equimolar concentrations (final conc.: 200 mM). Solutions were mixed with supernatant from control cells, that were grown to stationary phase, at a ratio of 1:1. Control cells are expressing a non-fluorescent protein under the control of P_{AOX} (cp. subsec. 11.1.3). Furthermore, a spectrum of unquenched supernatant, mixed with water at a ratio of 1:1, and a spectrum of pure water is shown for comparison. The supernatant did not contain cellular debris, but all molecules that cannot be removed by centrifugation. Data is recorded about 2 min after mixing of supernatant and quenching agents. Signals remained stable over time (data not shown). For a more detailed description of the procedure see section 7.3.
The graph in fig. 11.28 clearly shows a reduction of background fluorescence for every substance. The substances' quenching abilities are obviously increasing as follows: *TCEP < Ascorbic Acid < β-ME < DTT, KBr*. Unexpectedly

DTT and KBr show a very similar quenching behavior possibly indicative for a substance-independent limitation. β-ME shows a similarly good quenching ability when its concentration is doubled (data not shown). The observed differences in quenching ability can stem from pH-dependent differences in the redox potentials, that affect the redox state. Indeed, the reducing agents can act as oxidizing agents in the correct pH range, too. The pH of stationary phase yeast media can vary between 4-6. Hence a more detailed investigation would be required to identify the exact quenching mechanism for every substance. Both effects, donation or acceptance of an electron can quench an excited fluorescent molecule. The fluorescence can be quenched via Energy (FÖRSTER)- or Charge (DEXTER)- transfers. In case of heavy atoms the *external heavy atoms effect* leads to fluorescence quenching.

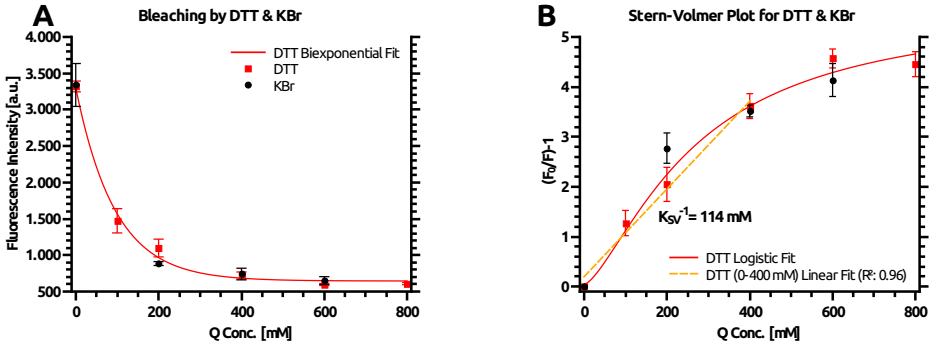

Figure 11.29: ***(A)*** *Bleaching of control cell supernatant by DTT and KBr at conc. from 0-800 mM* ***(B)*** *Stern-Volmer-plot including data from (A);* K_{SV}*: Stern-Volmer quenching constant*

Stern-Volmer kinetics in collisional quenching is described according to the following equation:

$$\frac{F_0}{F} = 1 + K_{SV} \cdot [Q] \tag{11.10}$$

The remaining background signal might originate from non-quenchable molecules in the media. Scattering on the other hand, that arises from instrumental limitations, is an unavoidable source of non-quenchable signal. Nonetheless, the baseline spectrum recorded from pure water indicates no strong signal scattering, arising from the setup itself. Due to the preceding results, the quenching ability of DTT and KBr was tested on supernatant at concentrations between 0-800 mM. The results are depicted in figure 11.29-A. Data was converted into a *Stern-Volmer plot* in figure 11.29-B to make results more tangible. F_0 and F are the fluorescence intensity in the absence and presence of quenching agent, respectively. K_{SV} is the Stern-Volmer quenching constant and [Q] is the applied concentration of quenching agent. In a linear plot the inverse of the slope K_{SV}^{-1} is the concentration at which 50% of the fluorescence intensity is quenched.
As it can be seen in figure 11.29-A the recorded intensity is exponentially decreasing for both, DTT and KBr, but a constant background signal of about 20% remains. The quenching ability of both compounds is comparable among different concentrations. This is especially interesting since DTT is a larger molecule than KBr, and could have theoretically sterical restriction for quenching. In conclusion, this gives evidence that the quenchable fluorescent molecule itself is rather small and/or offers sufficient accessibility. The presented Stern-Volmer analysis (see fig. 11.29-B) reveals a highly non-linear

slope for both compounds. This indicates that multiple fluorescent species are contained in the media. Not all seem accessible for the quenching agents. Apparently this equally applies for DTT, as well as KBr. The data can be mostly linear approximated from 0-400 mM. A linear fit reveals an apparent K_{SV}^{-1} of about 114 ± 10 mM, which can be used as a 'quenching estimate' for the predominant and quenchable fluorescent species in the solution. Thus to remove most of the quenchable signal a considerably higher concentration of quenching agent has to be used. Therefore 200-600 mM of DTT and KBr were tested with regards to their long-term influence on EGFP-expressing cells and their ability to restore an unbiased EGFP emission spectrum.

Long-term Effects of Quenching Agents on Yeast Cell Suspensions

EGFP-expressing cells from early stationary phase were mixed with 200-600 mM (final concentration) of DTT/KBr. The fluorescence emission signals were observed at 512 nm (Exc.: 488 nm) over a prolonged time periode, to estimtate detrimental effects on cellular fitness and fluorescence signal strength. The results of this experiment are depicted in figure 11.30-A.

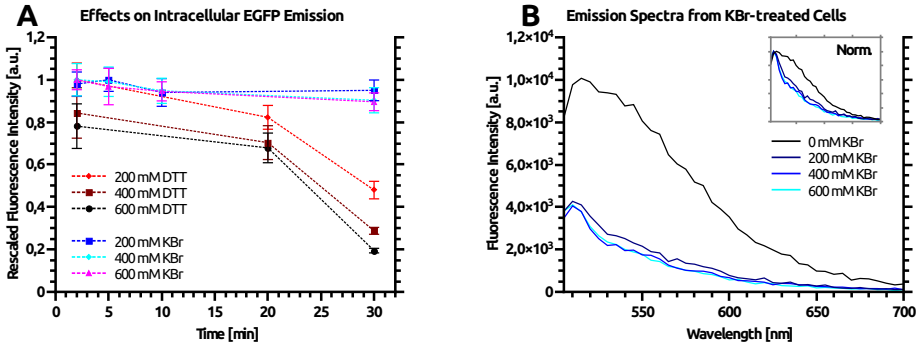

Figure 11.30: ***(A)*** *Long-term observation of intracellular EGFP emission upon cell treatment with high concentrations of DTT and KBr* ***(B)*** *Emission spectra of EGFP-expressing cells after 30 minutes upon incubation with Kbr; Temp.: 20°C, Vol.: 200 µL, Detector gain: 160, Focus Depth: 23 mm, (A-specific) $\lambda_{Exc.}$: 488 nm (f: 5 nm), $\lambda_{Em.}$: 512 nm (f: 5 nm); (B-specific) $\lambda_{Exc.}$: 488 nm (f: 5 nm), $\lambda_{Em.}$: 505-700 nm, Δ_{Scan}: 2 nm (f: 5 nm)*

Intriguingly, DTT seems to increasingly influence the cellular EGFP emission signal already after 20 minutes. 30 minutes after mixing, the cellular EGFP emission signal strength decreased by 50-80%. The effect appears to be more pronounced the higher the concentration is. On the contrary, the EGFP emission signal strength remains stable during long-term observation of KBr-treated cells, regardless of concentration (cp. fig. 11.30-A/B).
Figure 11.30-B shows the change of the emission spectra recorded between 505-700 nm (Exc.: 488 nm), without and with different amounts of KBr after 30 min. As it can be seen, KBr treatment yields a clearly visible EGFP spectrum and removal of background fluorescence. Even though it seems that 200 mM KBr are sufficient to quench most of the accessible background signal, 400 mM KBr will be applied during further experiments. This is, to not be limited in quenching ability due to an insufficient quencher concentration and to allow for a robust assay. Since the EGFP spectrum is very distinct and remains stable over time, it can be assumed that it is possible to conduct mostly unbiased *in cellulo* fluorescence and anisotropy measurements without additional washing of cells.

11.3.2 A Reliable Plate Reader-based Readout of Fluorescence & Anisotropy

To get an overview of cellular development, probe molecule behaviour and the general expression under the P_{AOX} promoter, three major readouts, optical density, fluorescence anisotropy and fluorescence intensity, will be compiled in a screening approach. Subsection 11.3.1 describes an approach to directly obtain meaningful fluorescence-based readout from growing yeast cells, by circumventing the strong bias from excreted fluorescent species. The described procedure will be applied as a general approach to easily quench background fluorescence.
For an accurate compilation of all three readouts the type and magnitude of mutual correlations between the parameters need to be identified. Limitations of the setup could lead to non-linear readout responses and need to be considered. Additionally errors could be introduced by the measured sample itself. For example, the magnitude of scattering of fluorescence signals will depend on the optical density of yeast cell solutions.
It has been shown in section 11.1.3 that the depolarization of *in cellulo* fluorescence signals occurs linearly with the density of a cell suspension at ODs ranging from 0.5 to 8. The gained information can already be used to correct fluorescence anisotropy data obtained at different optical densities. Setup-wise, it has already been shown that a linear response to varying concentrations of fluorescent molecules can be expected (cp. fig. 11.9-A). The (unpolarized) *in cellulo* fluorescence intensity also depends linearly on the optical density of cell solutions (cp. fig. 11.17). Thus fluorescence intensity can easily be normalized to obtain the amount of signal per OD unit (equiv. to signal per cell). Likewise the readout of optical densities by the utilized plate reader is linear in comparison to measurements of diluted samples on a spectrophotometer (see fig. A.13). A detailed description of the plate reader-based assay is given in section 7.3. The section additionally describes data correction and sample treatment.

11.3.3 Introducing System Perturbations for Supportive Crystallization Conditions

As described earlier about 60-70% of the fluorescence signal polarization can be attributed to effects from energy transfer between probe molecules (cp. sec. 11.1). This results in a strong negative correlation of unbleached anisotropy and homoFRET potential. Logically the unbleached anisotropy can be considered as an inverted readout to estimate intraperoxisomal probe molecule concentrations.
To introduce perturbations to the cellular growth different possible external stimuli were chosen based on a literature search. They comprise chemical compounds, that introduce adaptation pressure with regards to pH, osmotic stress, carbon source, oxidative stress, or inhibition of cell division. All applied chemical compounds and the corresponding growth conditions of yeast cells are listed in table 11.7. Carbon sources were altered according to Suppi *et al.* [197][1], to introduce variations in the activation of AOX genes. Methanol, which is not only the main activator of AOX genes, functions also as an intracellular precipitant [34, 53]. Due to the pronounced cellular resistance against high medium concentrations of methanol, its effect on the peroxisomal development has been investigated as well at concentrations which exceed the growth optimum [29].

[1] And others [21, 108, 142, 200, 239]

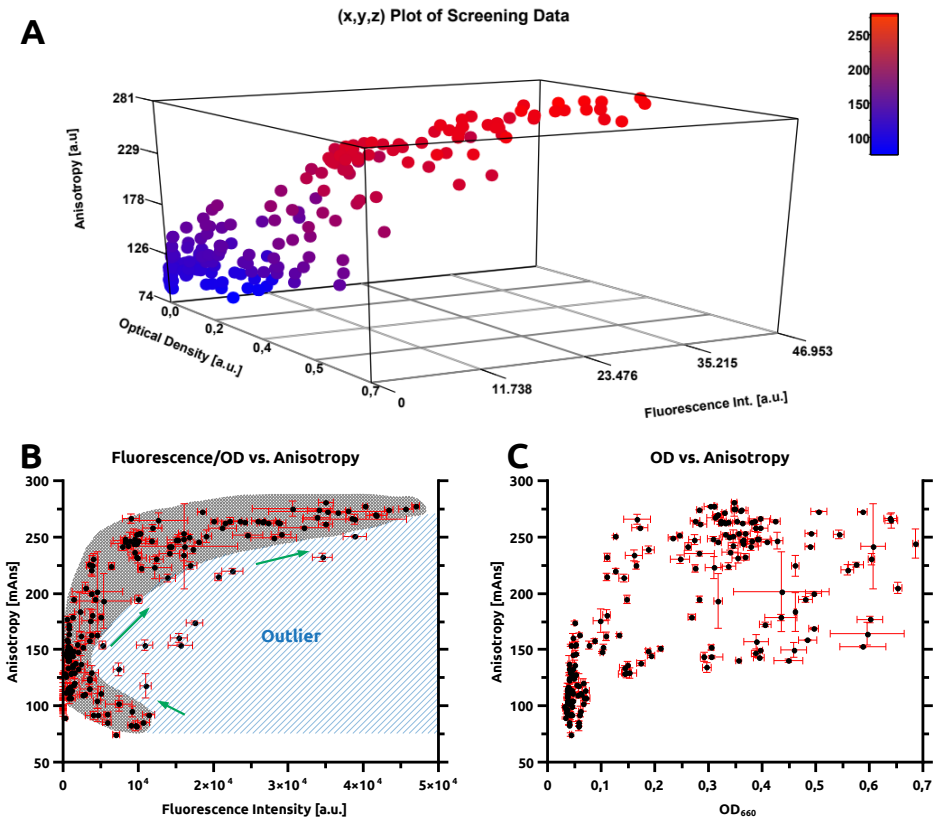

Figure 11.31: *Representation of compiled data from plate reader screening;* ***(A)*** *(x,y,z) plot showing correlation of OD, fluorescence intensity and -anisotropy* ***(B)*** *plot of fluorescence anisotropy versus -intensity, black hatched area shows typical correlations, green arrows indicate development over time* ***(C)*** *Plot of fluorescence anisotropy versus OD; Data recorded according to subsec. 7.3*

Furthermore the pH of the growth medium has been altered. Such changes may lead to a shift of intracellular pH according to Arroyo-López *et al.* [142] and Peña *et al.* [238]. Oxidative stress has been exerted by exposure to high concentrations of DMSO, which can lead to growth inhibition and a change of the proteom in related yeast species [16, 195]. Likewise the anti-cancer drug *nocodazole* has been used to partially inhibit cell division [22, 152]. As stated by Brotherton [11] and Hohmann [83], yeast cells can adapt to osmolarity by an adjustment of the tugor pressure. This is necessarily accompanied by changes in gene transcription e.g. regulated by MAP kinase pathways, and might also effect AOX genes.

Except for the perturbing chemicals, cells were grown under typical P_{AOX}-inducing conditions. Usually cells have been investigated after 24 hrs of allowed batch growth. For some conditions cells were investigated at different time points from 4- 72 hrs after exposure to a perturbing environment. A detailed description of cell growth and screening is given in section 7.3. One data point represents a minimum of 3 yeast cell cultures grown under identical conditions. For all measurement the yeast strain *H.p.*-12DAC4-EGFPskl has been utilized (cp. sec. 11.1).

Corrected data, that has been obtained for various conditions and time points is summarized in figure 11.31A-C. Plot A shows the 3D correlation of optical density, fluorescence/OD and anisotropy. Plot B and C show a 2D plot of fluorescence intensity or optical density versus anisotropy, respectively. As

it can be seen, especially fluorescence/OD and anisotropy show an obvious non-linear correlation, whereas optical density and anisotropy are somewhat uncorrelated.
The intraperoxisomal anisotropy almost always follows a predetermined developmental path upon increasing fluorescence intensity per OD unit (fig. 11.31-B, black hatched area). This behaviour is basically reconfirming the interconnection of target protein expression (fluorescence intensity) and peroxisomal development, which has already been described earlier. Unfortunately, the combined results also indicate a mostly unbiased peroxisomal development, while cellular growth is impeded for many conditions. Nonetheless a few outliers can be found, that show higher fluorescence intensities while simultaneously exhibiting a stronger signal depolarization by homoFRET (fig. 11.31-B, blue hatched area). They represent cell samples that are exclusively grown at pH 7-8, stabilized by 50 mM potassium phosphate. Another interesting observation is a small bulge towards higher intensities, visible for extremely low anisotropy values. This bulge corresponds to an early expression phase of probe molecules. Following the theory developed in section 11.1, this phase holds the highest concentration of probe molecules. Areas in figure 11.31-B were drawn manually, based on the typical relation of anisotropy and intensity found for unbiased conditions (fig. 11.32). Remarkably, cellular growth and P_{AOX}-induction seem not to be strongly correlated (cp. fig. 11.31-C).
For a correct interpretation of the shown figures it is necessary to remember that a lower anisotropy value reflects a higher degree of energy transfer and thus of probe proximity/ concentration. Ultimately it is of course a convolution of energy transfer and viscosity. Data on a horizontal line (at a constant anisotropy value) corresponds either to comparable concentration levels of probe molecules or to a certain extent to comparable viscosities. Thus, while a higher fluorescence intensity would usually be interpreted as higher concentration, here it is rather reflective of non-tangible changes in the peroxisomal interior. For example, if a certain anisotropy value is reached at different fluorescence intensities this is most likely indicative for differing peroxisomal volumes that lead to the same relative concentration. Likewise the same intensity at differing anisotropy values indicates different concentrations. Hence it is logical to assume a less steeper slope for enlarged peroxisomes in diagrams as in figures 11.31-B or 11.32.
Data recorded for yeast cells grown at different medium pHs is shown in a separate plot in figure 11.32. Different time marks emphasize the progres-

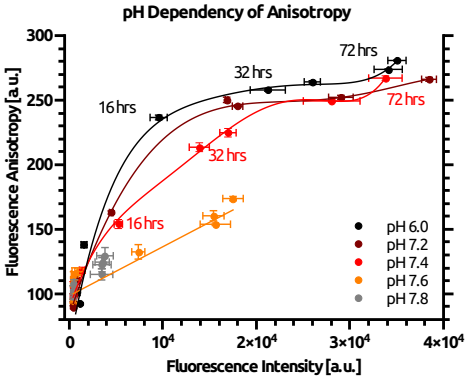

Figure 11.32: *Screening data from yeast cells grown in batch mode at media pHs of 6-7.8, plot of anisotropy versus fluorescence; Data recorded according to subsec. 7.3*

sion of data points during cellular growth. From this a clear dependency of anisotropy/fluorescence on the pH can be identified. The higher the pH the more flat is the increase in anisotropy at higher probe molecule expression levels. The steepest increase is found for standard conditions, the growth at pH 6 (black curve) and flattest for cells grown at pH 7.6 (orange line). Cells incubated in media at pH 7.8 were limited in growth and did not yield distinct data (gray data points). Data is interpolated either with a 5-order polynomial (pH 6, 7.2, 7.4) or with a linear function (pH 7.6).

This section mainly demonstrates how to simultaneously obtain three major readouts from growing yeast cell cultures, that were made comparable over a wide dynamic range via data correction. The developed screening approach has been utilized to probe peroxisomal development and P_{AOX}-induction during growth under the influence of different stress inducers. The aim has been to identify possible systematic perturbations that could help to crystallize recombinantly expressed proteins. The acquired data mainly indicates, that almost none of the tested conditions lead to a distortion in the balance of P_{AOX}-induced expression and peroxisomal development. Although for high, but not low pHs a change in peroxisomal development could be detected. This change can at best be interpreted as the occurrence of enlarged peroxisomes e.g. by a restriction of fission. Still, further evidence is required to create a more accurate description of the observed phenomenon. Even though this finding proves the applicability of the screening approach it has to be questioned, whether such changes can lead to an increased propensity of recombinant proteins to crystallize. Unfortunately, no other test condition exhibited strongly deviating dynamics.

Table 11.7: *Test conditions for perturbation screening; If not stated otherwise, all conditions contained 0.5% (v/v) methanol, const. Glyc. indicates a glycerol concentration of 0.2% (v/v); Results are shown in fig. 11.31-A*

Nb.	Stimulus	Compound	Conc.
1	Carbon Source	Xylose (w/o Glyc.)	0.1% (v/v)
2	Carbon Source	Xylose (w/o Glyc.)	0.2% (v/v)
3	Carbon Source	Ribose (w/o Glyc.)	0.1% (v/v)
4	Carbon Source	Ribose (w/o Glyc.)	0.2% (v/v)
5	Carbon Source	Oleic Acid	0.2% (v/v)
6	Carbon Source	Oleic Acid	0.1% (v/v)
7	Carbon Source	Oleic Acid	0.05% (v/v)
8	Carbon Source	only Glycerol	0.5% (v/v)
9	Carbon Source	only Glycerol	0.25% (v/v)
10	Carbon Source	only Glycerol	0.1% (v/v)
11	Carbon Source	Glycerol	0.5% (v/v)
12	Carbon Source	Glycerol	0.25% (v/v)
13	Carbon Source	Glycerol	0.1% (v/v)
14	Precipitant	Methanol (const. Glyc.)	1.0% (v/v)
15	Precipitant	Methanol (const. Glyc.)	1.5% (v/v)
16	Precipitant	Methanol (const. Glyc.)	2.0% (v/v)
17	Precipitant	Methanol (const. Glyc.)	2.5% (v/v)
18	Precipitant	Methanol (const. Glyc.)	3.0% (v/v)
19	Osmolarity	Sorbitol	0.5 M
20	Osmolarity	Sorbitol	1.0 M
21	Osmolarity	Sorbitol	1.5 M
22	Osmolarity	NaCl	0.2 M
23	Osmolarity	NaCl	0.5 M
24	Osmolarity	NaCl	1.0 M
25	Osmolarity	NaCl	1.5 M
26	Oxidative Stress	DMSO	4%
27	Oxidative Stress	DMSO	6%
28	Oxidative Stress	DMSO	8%
29	Oxidative Stress	DMSO	10%
30	Oxidative Stress	DMSO	12%
31	Growth Inhibition	Nocodazole	0.05µM
32	Growth Inhibition	Nocodazole	0.1µM
33	Growth Inhibition	Nocodazole	0.25µM
34	Growth Inhibition	Nocodazole	0.50µM
35	Growth Inhibition	Nocodazole	1.00µM
36	Growth Inhibition	Nocodazole	3.00µM
37	pH	w/o PP Buffer	
38	pH	PP Buffer	pH 4.5
39	pH	PP Buffer	pH 5.0
40	pH	PP Buffer	pH 5.5
41	pH	PP Buffer	pH 6.0
42	pH	PP Buffer	pH 6.5
43	pH	PP Buffer	pH 7.0
44	pH	PP Buffer	pH 7.2
45	pH	PP Buffer	pH 7.4
46	pH	PP Buffer	pH 7.5
47	pH	PP Buffer	pH 7.6
48	pH	PP Buffer	pH 7.8
49	pH	PP Buffer	pH 8.0

12 | Implications from Fluorescence Anisotropy for Protein Crystallography

In protein crystallography it is assumed that concentration is one of the main driving forces in protein oligomerization, nucleation and crystallization. It has been proven in section 11.1, that the fluorescence anisotropy of a probe molecule can help to understand this important parameter in an *in vivo* crystallization environment. It has been demonstrated that the movement of fluorescent units and their proximity can yield specific information about their behaviour in terms of particle density and clustering. It is logical to look into further possible applications of this technique in crystallography.
The big advantage of fluorescence anisotropy is truly its influenceability by depolarizing effects. The depolarizing effect of FRET (as it has been applied throughout this thesis) is rather caused by the random orientation of acceptor molecules and not by the energy transfer itself. But the probability of energy transfer strongly depends on the orientation of donor and potential acceptor molecules towards each other. More specifically, the orientation of the donor emission transition moment and the acceptor absorption transition moment. In homoFRET the angle between these dipole moments is the same for donor and acceptor. Besides the orientation, other properties like spectral overlap, fluorescence lifetime and quantum efficiency play a role in determining the rate of energy transfer between donor and acceptor.
In FRET theory (see sec. 3.1) the magnitude of energy transfer between two dipoles is summarized by the FÖRSTER radius. It includes the orientation factor κ^2, whose variations may be used to discern protein clusters and crystals, based on the random versus non-random orientation of molecules.
In case of clustering/crystallization of fluorescent molecule, we will find two opposing effects upon examination via fluorescence anisotropy. First, an increase of anisotropy due to an increase of the rotational correlation time, second a decrease of anisotropy due to the influence of energy transfers. The latter effect will vary with the orientation of acceptor molecules. Hence if both values could be determined from an integrated theory, clusters and crystal should yield different values.

12.1 Short-Range Order Permits Variations of the Orientation Factor κ^2

The anisotropic orientation factor κ^2 (sec. 3.1) is strongly influencing the rate of energy transfer in FRET [87, 110]. It depends on the relative orientations of emission- and absorption transition moments in a molecular arrangement. Thus in a non-random situation of short- or long- range order, energy transfer between adjacent molecules can heavily increase or decrease. If molecules are just constraint in their translational and rotational freedom without exhibiting some kind of order, only an averaged orientation factor will contribute to the rate of energy transfer.

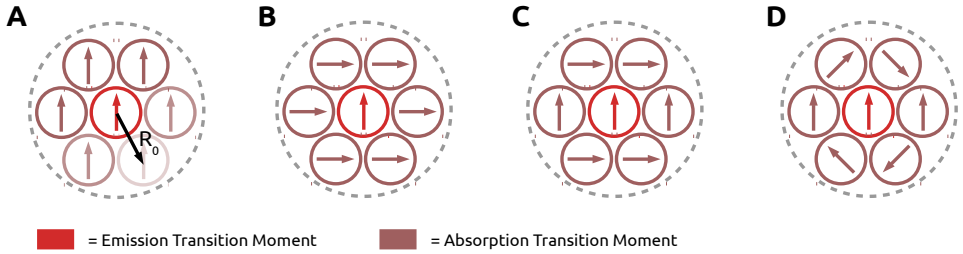

Figure 12.1: *Different molecular arrangement of donor and acceptor molecules. Orientation of emission- and absorption transition moments are indicated by arrows for donor and acceptor, respectively;* R_0*:* FÖRSTER *distance*

The average value of κ^2 can be approximated for different two-dimensional molecular arrangements (see fig. 12.1), that exhibit short-range order. For this the absorbing molecules (receiver) are assumed to be located within normal FÖRSTER distance (R_0) of the emitter. This allows to estimate differences in energy transfer efficiency in comparison to completely randomly oriented receivers. In the case of completely random orientations the 2D-κ^2 averages to a value of 1.25. Figure 12.2 depicts plots of the 2D-κ^2 of the molecular arrangements A-D (fig. 12.1), when different degrees of rotational freedoms are allowed for the receiving molecules.

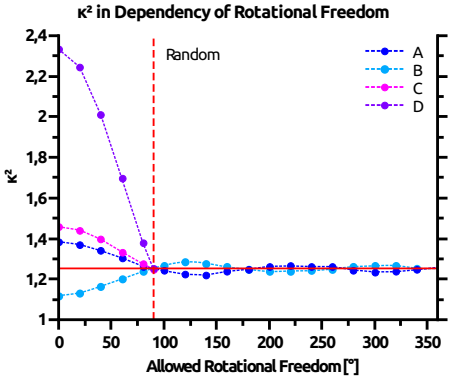

Figure 12.2: *Variation of a theoretical 2D-*κ^2 *in molecular arrangements A-D (fig. 12.1); Each data point results from* $5 \cdot 10^4$ *averaged calculations of uniformly distributed angles in the denoted range; E.g.: 50°= Allow angles between 0°and 50°; The average 2D-*κ^2 *reduces to 1.25 (random) when 90°or more degrees are allowed*

Those are arranged in every case around a central emitter at fixed angles of 30, 90, 150, 210, 270 and 330 degrees. If translational freedom is kept restricted while rotational freedom is completely allowed, the 2D-κ^2 again averages out

to a value of about 1.25. Every data point in figure 12.2 was created from averaging $5 \cdot 10^4$ calculations of random parameters from the denoted range of rotational freedom angles[1]. The corresponding values were calculated according to [198]:

$$\kappa^2 = (cos\theta_R - 3cos\theta_D cos\theta_A)^2 \tag{12.1}$$

Here θ_R describes the angle between the emission transition moment of the donor and the absorption transition moment of the acceptor molecule. θ_D and θ_A describe the angle between the donor emission- and the acceptor absorption transition moment, respectively, and the connecting direction vector between both. A more detailed description can be found in section 3.1.
It can be assumed that any sort of real rotational freedom would most likely be more restrictive and should perhaps be approximated by a GAUSSIAN distribution. It is apparent from figure 12.2 that short-range order can drive the orientation factor away from its average value. Depending on the exact molecular arrangement energy transfer might heavily increase or decrease, given that the overall distance between molecules is low.
Ordered molecular arrangement can be found in biological structures such as protein crystals. It is therefore plausible to assume that protein crystals can give rise to an increased or decreased rate of energy transfer, which might render them distinguishable from protein clusters with random molecule orientations. A prerequisite of this is of course the presence of a fluorophore that allows for such detection.

12.2 FRET Orientation Factor in a $P2_12_12_1$- EGFP Crystal

Section 12.1 pinpoints to the variations among a theoretical 2D-κ^2 factor in molecular arrangements of donor and acceptor. But real protein crystals are seldom two-dimensional. Therefore it needs to be investigated, whether 3D crystals have non-random κ^2 factors. It is broadly accepted that in case of random 3D orientations the κ^2 factor will average out to $^2/_3$ [110, 198].

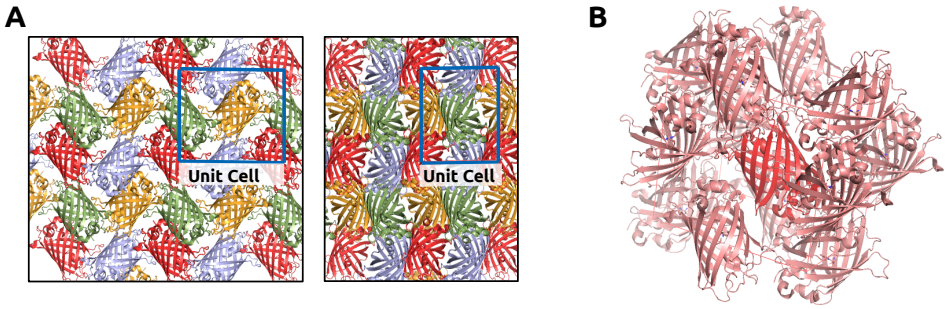

Figure 12.3: ***(A)*** *Depiction of an EGFP crystal (PDB: 4EUL) with its unit cell* ***(B)*** *A randomly chosen donor molecule with possible acceptor molecules in* FÖRSTER *distance*

The $P2_12_12_1$ space group is one of the most abundant space groups among protein crystals [47]. Therefore it is used as an example to approximate the average orientation factor seen by an excited donor molecule in such a crystal. For consistency the crystallographic data of an EGFP crystal has been used (PDB: 4EUL). Its structure is shown in figure 12.3-A. The absorption- and

[1] Calculated in *LibreOffice Calc* (Vers. 5.4.6.2) and *Matlab* (2011b)

Table 12.1: *Extracted angles θ_D, θ_R, θ_A and calculated κ^2 values for symmetry mates 1-12; Values correspond to the molecular arrangement shown in figure 12.3-B*

Sym.-Mate	**D [°]**	**R [°]**	**A [°]**	**κ^2**
1	79	90	146	0.233
2	35	97	112	0.647
3	46	152	113	0.003
4	118	90	99	0.052
5	68	97	145	0.647
6	34	90	101	0.233
7	42	152	128	0.249
8	81	90	62	0.052
9	15	97	111	0.844
10	67	152	134	0.003
11	69	97	165	0.844
12	52	152	138	0.249
Avg.				**0.338**

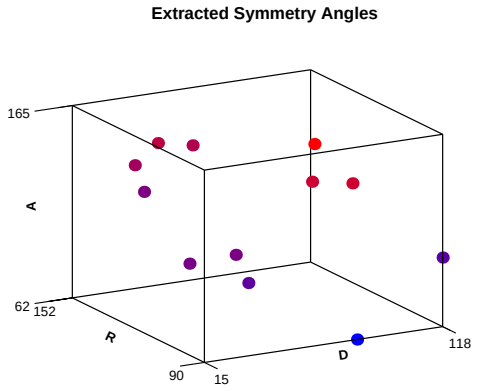

Figure 12.4: *3D representation of extracted angles θ_D, θ_R, θ_A; Values correspond to the molecular arrangement shown in figure 12.3-B*

emission transition moments were approximated as vectors along the long molecular axis of EGFP, which seems legitimate due to the fixed geometrical relation of the molecules. The relative orientations were calculated by extracting transformation matrices from *PyMol* [234] and decomposing the contained rotation matrices. For the approximation only molecules in the range of the EGFP-FÖRSTER distance (R_0= 4.7 nm) were considered (see fig. 12.3-B). The size of the unit cell (a,b,c) is $\sim 51 \cdot 62 \cdot 70$Å. Therefore, equally oriented molecules were not considered in the calculation. Due to the regular pattern within a crystal, the contribution of molecules outside of the unit cell should yield non-deviating values with descending influence on the final result. The same applies for a cluster of proteins with true random orientations. In total 12 symmetry mates (light red, fig. 12.3-B) of a central donor molecule (red) were included in the calculation. The extracted angles θ_D, θ_R, θ_A and the corresponding theoretical κ^2 value are listed in table 12.1. The angles are also shown in figure 12.4 to visualize their replication of a 3D relation.
Likewise to the 2D case (sec. 12.1), the anisotropic response of the orientation factor favours some orientations over others. As it can be seen the value of

the orientation factors varies a lot between different symmetry mates. The highest value found is 0.844 and the lowest is 0.003. The calculated average κ^2 factor for the investigated $P2_12_12_1$ crystal has a value of 0.338, which is considerably less than the average value of $^2/_3$ for random orientations. For another theoretically excited molecule within the crystal arrangement a value of 0.306 can be calculated.
In conclusion, the average energy transfer rate within an EGFP crystal will be less than in a cluster with random orientations. This seems to be backed up by data from Inoue *et al.* [75][84], who found polarized light irradiating from GFP crystals. But this would indeed require some odd molecule arrangement, that is required to disallow most of the energy transfer to yield polarized emission. Due to the proximity of molecules in a crystal, the energy transfer efficiency will be 1 regardless of the difference between a random and a non-random molecular arrangement. Hence, to successfully measure such differences other properties of the homoFRET energy transfer would require fine-tuning.

12.3 Prerequisites for the Differentiation of Protein Crystals & Clusters

Most importantly the homoFRET energy transfer rate has to be dramatically reduced for an investigated conglomerate, in order to make the κ^2 differences visible. At best the transfer efficiency should not exceed ≈ 1 for a volume packing fraction (VPF) of 1 (see subsec. 11.1.1). Otherwise it is not possible to detect any differences in an average κ^2 arising from the molecule arrangement. As mentioned in subsection 11.1.1, this could be achieved by using fluorophores with an increased STOKES shift [192], or by a decreased quantum yield according to (cp. eq. 3.5):

$$R_0 = 0.211 \left(\kappa^2 n^{-4} Q_D J(\lambda)\right)^{1/6} \tag{12.2}$$

Q_D is the quantum yield of the donor (in absence of acceptor), $J(\lambda)$ is the overlap integral of donor and acceptor spectra and n is the refractive index. The latter one is usually assumed to be 1.4 for proteins in aqueous solutions [110]. It has to be determined to what extend the anisotropy of n in a crystal would influence the anticipated measurements. Further on, a reduction in energy transfer rate could also be achieved, if fluorophores are restricted in proximity due to the size of the investigated molecules. In case of the EGFP's R_0 this would require an expansion of the molecule to a radius of about 4.7 nm, which in turn equals the R_0 (cp. fig. 11.5). Another possibility is the reduction of the labelling rate.
For a *gedankenexperiment* we may assume an EGFP-like fluorescent particle with a FÖRSTER distance that is reduced to R_0=2.23 nm. Furthermore size, fluorescence lifetime and quantum yield are assumed to be identical (Compare subsec. 11.1.1). Under the assumption that κ^2 again equals $^2/_3$, the measured fluorescence anisotropy from excited molecules should reduce to 0, when the particle density reaches a VPF of 1 (See table 12.2). If the orientation factor now reduces to 0.338, similar to the example presented in figure 12.3-B, the energy transfer efficiency will drop to 0.77 and yield a steady-state anisotropy of about 90 (assuming a limiting anisotropy of 400 mAns). The difference in depolarization and hence in energy transfer can be detected and should enable to differentiate a molecule cluster from a molecule crystal. Indeed, another assumption was made throughout this section, to derive such obvious results:

Table 12.2: *Theoretical values for homoFRET transfer efficiencies and fluorescence anisotropy in an EGFP-like molecule conglomerate for different κ^2 factors; Other molecule properties are identical to those in subsec. 11.1.1*

κ^2	mAns [r]	E_{FRET}
0.667 (Cluster)	0	1
0.476 (Cluster)	46	0.89
0.338 ($P2_12_12_1$ Crystal)	90	0.77

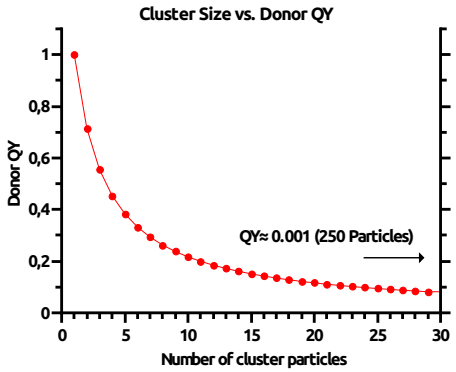

Figure 12.5: *Plot of cluster size versus the quantum yield of an excited donor molecule within the cluster; Calculated via eq. 11.9*

It has been assumed, that the average particle density in protein crystals and protein clusters (precipitate/aggregates) is almost similar. This is not necessarily true and would require further clarification.
Fig. 12.5 schematically depicts the theoretical quantum yield of an excited molecule in a protein conglomerate of increasing size. The quantum yield has been calculated via equation 11.9. It is immediately apparent that it is decreasing rather slowly due to possible energy back-transfers to the donor molecule. But it reaches a value of below 0.001 for about 250 particles. The decay of a molecule's quantum yield will depend on its position within the cluster. This is due to the finite size of the protein conglomerate. Likewise, it can be assumed that the quantum yield of an excited molecule will increase, the more molecules are initially excited within a cluster. For a crystal system with its repeating structure the maximum signal depolarization also depends on the distance of excited molecules and the number and orientation of available acceptor molecules. Not only is the energy transfer itself determined by the relative acceptor orientations, but also the degree of depolarization due to this orientation. For a crystal system it might be expected that even less depolarization occurs, due to the non-random orientations of molecules.
It can be concluded that a finite number of a few hundred particles will render a situation in which statistically all energy will dissipate from its original molecule and $E_{FRET} \approx 1$ can be assumed.

12.4 Application of Anisotropy-based Crystal Identification

The simple *gedankenexperiment*, which has been presented in the previous sections, relies on the comprehensive availability of fluorophores. Therefore, the

presented idea is difficult to extend to proteins without an intrinsic fluorescent center. But this drawback might be circumvented by several approaches.
First, it is of course possible to create a fusion protein from a target- and a fluorescent protein [100]. Such a fusion protein can presumably quash any tendency of a target protein to crystallize. But it may also serve the opposing side and be the cause of crystallization. This is indeed suggested by several studies, using GFP- [161], MBP- [259], thioredoxin [128]- or lysozyme-fusion tags [231]. A review of fusion tags used in crystallograpy is provided in [92]. Therefore, an utilization of the proposed approach might be extended to non-fluorescent target proteins via the usage of a fluorescent fusion protein, that provide a fluorescent center. In the light of the anticipated measurements, it is necessary to provide a rigidly attached fluorescent center. Any introduced rotational freedom will reduce the influence of the molecule orientation on the energy transfer efficiency (cp. fig. 12.2).
An alternative approach to fluorescent fusion proteins is embodied by a certain class of biarsenical, organic dyes [117]. They stand out due to their special properties: First, they are easily-attachable to a tetracysteine-tag on a protein and hence are almost non-invasive. Second, they can penetrate biological membranes, which renders them also interesting for *in vivo* crystallography. Lastly, their fluorescence strongly depends on the binding to the tetracysteine-tag. Unbound dye will not be fluorescent. A formidable approach to reduce background signal. The second generation dye *CrAsH* [103] is a good representative of this class of dyes. Obviously, a dynamic binding scheme involves the risk of incomplete binding and requires special attention.

It is worth mentioning that the presented approach would also benefit from the usage of more elaborate optical techniques. In conclusion, a fluorescence anisotropy approach may allow to distinguish between protein clusters in form of precipitate / aggregates and protein crystals under certain conditions. It would expand the toolbox of early crystallographic characterization.

13 | *In Vitro* Production of Protein-Nanofibers for Diffraction Experiments at XFELs

The following chapter comprises an experimental approach in fiber X-ray diffraction[1]. It is shown how the formation of amyloid protein fibers composed of the human protein superoxide dismutase 1 (SOD1) is optimized in order to obtain a suitable sample for single-fiber structural characterization at a XFEL. SOD1 is a copper and zinc- containing 32 kDa homodimeric enzyme. The 16 kDa monomers contain each one Cu and Zn ion. The enzyme is relevant in the removal of reactive oxygen species in human neuronal cells. It is a known actuator of neurodegenerative diseases e.g. familial amyotrophic lateral sclerosis (fALS) and is ultimatively causing motor neuron death due to misfolding and fiber-like aggregation. More than 110 known mutations within the 153 aa-long protein lead to a form of fALS, but the mechanism that causes neurodegeneration is not yet fully understood. It has been proposed that mainly a combination of oxidation and a loss of metal ions lead to the formation of amyloid structures [85, 126, 187, 210].
As shown by Sunde *et al.* [54] , amyloid structures exhibit a diffraction response to X-ray radiation due to their repetitive structure of amino acids, folded in β-spines. To understand the emergence of neurodegeneration in fALS, XFELs offer an excellent opportunity to yield structural information at an atomic resolution to understand the mechanism of action in different mutated forms of SOD1. Here it is described how to induce, analyze and optimize fiber formation under laboratory conditions for such studies.

13.1 Optimization of SOD1-Amyloid Fiber Formation

The fiber formation is supposedly triggered by depriving Zn and Co ions from the protein and induced conformational changes due to oxidation [131]. For experiments the chemically modified and ultra stable 2ME-SOD1 has been used, in which a β-ME molecule is covalently linked to Cys^{111}. It can be converted into the wildtype form of the protein by incubation with a strong reducing agent [115, 174]. Lyophilized 2ME-SOD1 protein and an initial protocol for fiber formation was supplied by collaborators[2]. The initial protocol suggested fiber formation at conditions that comprises a pH range of 3-5, addition of reducing agents, EDTA, and sodium chloride.

[1] Project has been initiated by Leonard Chavas (Former researcher at CFEL, Hamburg) under proposal nb. 2015A8003

[2] Protein was produced by Ube Industries Ltd. and supplied by Noriko Fujiwara, Hyogo College of Medicine, Nishinomiya, Japan

13.1.1 Investigation of Fiber Formation via Thioflavin T Assay

To verify the fiber formation hypothesis for SOD1 all related chemical compounds were investigated in terms of their influence on the protein. ThT has been used as a reliable fluorescent indicator for protein fiber formation [114, 185, 206, 210]. First, lyophilized SOD1 protein has been resuspended in distilled water at 100 mg·mL^{-1}. The solution was filtered with a sterile 0.22 µm filter. Afterwards a variety of solutions were prepared to cover the following conditions: distilled water, unbuffered acidic pH, and low (50 mM) and high salt (500 mM) sodium acetate buffer at pH 3. Protein-containing solutions were freshly prepared by mixing 0.96 µL from 100 mg·mL^{-1} stock into a final volume of 150 µL of buffer or aqueous solution, containing 40 µM of ThT. Protein behaviour was monitored in terms of ThT fluorescence at $\lambda_{Em.}$: 485 nm in a plate reader (sec. 7.2). Constant agitation of the sample solution and heating to 37°C was applied. The results are depicted in figure 13.1-A.

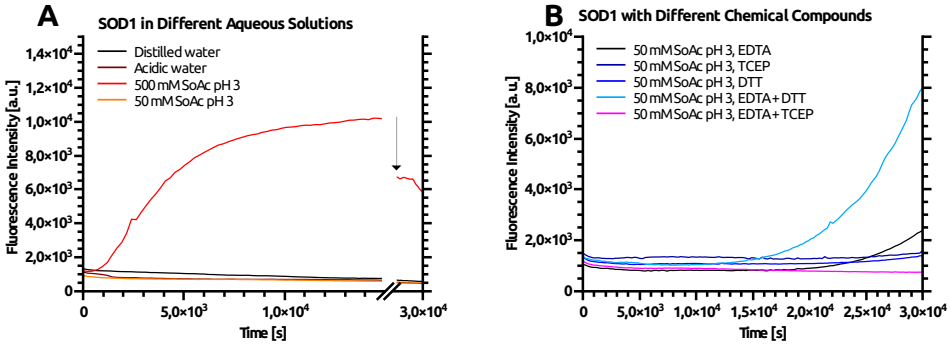

Figure 13.1: *ThT assay with SOD1 protein under presumed fiber formation conditions* ***(A)*** *at different buffer conditions and* ***(B)*** *with different additional compounds at low salt conditions; Temp.: 37°C, Vol.: 150 µL, Detector gain: 160, Focus Depth: 23 mm,* $\lambda_{Exc.}$*: 442 nm (f: 5 nm),* $\lambda_{Em.}$*: 485 nm (f: 5 nm)*

As it can be inferred from the trend of the curves, high salt concentrations at an acidic pH already lead to a strong destabilization of the SOD1 structure causing the formation of protein conglomerates that interact with ThT. On the contrary no change in fluorescence intensity is observed for protein in low salt buffer at the same pH. Additionally protein mixed with water at a low pH and in water at neutral pH does also not show any change in ThT fluorescence. These result indicate that high salt concentrations may lead to protein precipitation/ fiber formation, but an acidic pH alone does not. The protein was apparently able to withstand acidic conditions over a timespan of 8 hrs without exhibiting a ThT-related increases of fluorescence. A drop in fluorescence for high salt conditions upon a prolonged incubation can be attributed to intense evaporation (black arrow).
A second experiment has been conducted to verify the influence of further compounds, that have previously been identified as actuators in fiber formation. 20 µM protein were mixed with 50 mM sodium acetate at pH 3, 40 µM ThT and either 10 mM EDTA, 100 mM DTT, 100mM TCEP, or both EDTA and TCEP/DTT. As depicted in figure 13.1-B the obtained results indicate foremost an influence on fiber formation processes by EDTA and DTT, but not by TCEP. When SOD1 is soley incubated with EDTA (in buffer solution) a slight increase of fluorescence can be found after several hours (black curve). If additionally DTT is present, the increase in fluorescence is

more rapid and distinct (cyan curve) and reaches similar intensity levels as for high salt conditions (fig. 13.1-A). Intriguingly, DTT or TCEP alone have no effect on the occurence of fluorescent signal (dark blue/ blue curve). This indicates that the removal of metal ions from the protein via a strong chelator could be the first necessary step to achieve amyloid formation, whereas DTT seems to play a crucial role in the progression of the process. Mixing of TCEP and EDTA does not result in an increase of ThT fluorescence (magenta curve). At an acidic pH range, TCEP will maintain its reducing power, whereas DTT will basically act as an oxidizing agent. Therefore it might indeed be concluded that oxidation via DTT plays a crucial role in the fiber formation process. This is also strongly indicated by results obtained from protein solutions at various pHs, containing EDTA and DTT (see fig. A.14).
In a third experiment is has been elucidated whether the mechanism of precipitate/fiber formation is comparable among high salt and low salt conditions. Since an increase of fluorescence has been observed in 500 mM sodium acetate buffer without addition of any other compounds, one may conclude that the destabilization of the protein by a high salt buffer is already sufficient to trigger fiber formation. To elucidate the matter, protein has been exposed to high salt buffer conditions mixed with different additional compounds. The results are presented in figure 13.2.

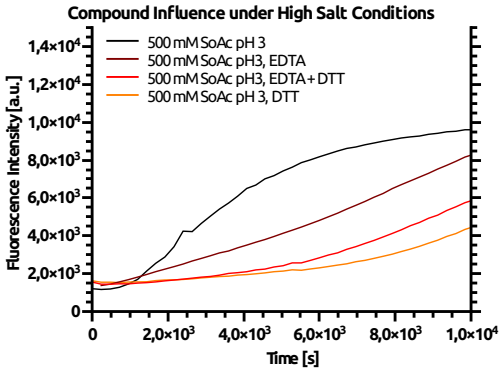

Figure 13.2: *ThT assay with SOD1 protein at high salt conditions; Temp.: 37°C, Vol.: 150 µL, Detector gain: 160, Focus Depth: 23 mm, $\lambda_{Exc.}$: 442 nm (f: 5 nm), $\lambda_{Em.}$: 485 nm (f: 5 nm)*

The black curve represents the previously obtained results with 20 µM protein in 500 mM sodium acetate buffer at pH 3. If EDTA (dark red curve) or DTT (orange curve) were added to the solution a dramatic change in formation kinetics could be observed. Both compounds slowed down the increase in fluorescence signal, indicating an individual influence on the precipitate formation in both cases. Interestingly a mixture containing both EDTA and DTT seems to exhibit some intermediate kinetics (red curve).
Obviously the results hint on two distinct mechanisms that lead to precipitate/fiber formation susceptible to ThT binding. Thereby one is solely induced by EDTA and DTT, whereas another one is induced by a general destabilization of the protein at high salt conditions. Although an increase in fluorescence indicates the formation of precipitate in both cases, the underlying structures do not necessarily need to exhibit an amyloid shape. It might be suspected that high salt precipitation does not lead to fiber formation, but only to the formation of protein aggregates. Samples have been further investigated via dynamic light scattering (DLS).

13.1.2 Investigation of SOD1 Fiber Formation via DLS

Measurements were conducted in order to elucidate the particle size distribution under conditions that favour fiber formation (cp. subsec. 13.1.1). DLS has been performed on a *SpectroSize 600* (Xtal Concepts, Germany) in 5 µL droplets covered by immersion oil and maintained at 20°C. Unfortunately, agitation or stirring could not be applied throughout these measurements. Radius frequency distributions were calculated every 5 min for 16 hrs. For size calculations an aqueous viscosity parameter of 1.006 $mPa \cdot s$ has been used. The results are depicted in figure 13.3-A-E. Samples contained 2 mg·mL^{-1} protein in 50 mM sodium acetate, 100 mM DTT and 10 mM EDTA were prepared at pH 3 (A) and pH 4 (D). Samples depicted in figure 13.3-B/C contained 3.5 and 5 mg·mL^{-1} (both at pH 3), respectively. Sample E contained only EDTA and no DTT at a pH of 3. Samples A-C exhibit a strong separation into two fractions at hydrodynamic radii of 2-3 nm and at around 100 nm. The acquired fraction signal intensity representing smaller sizes seems to follow

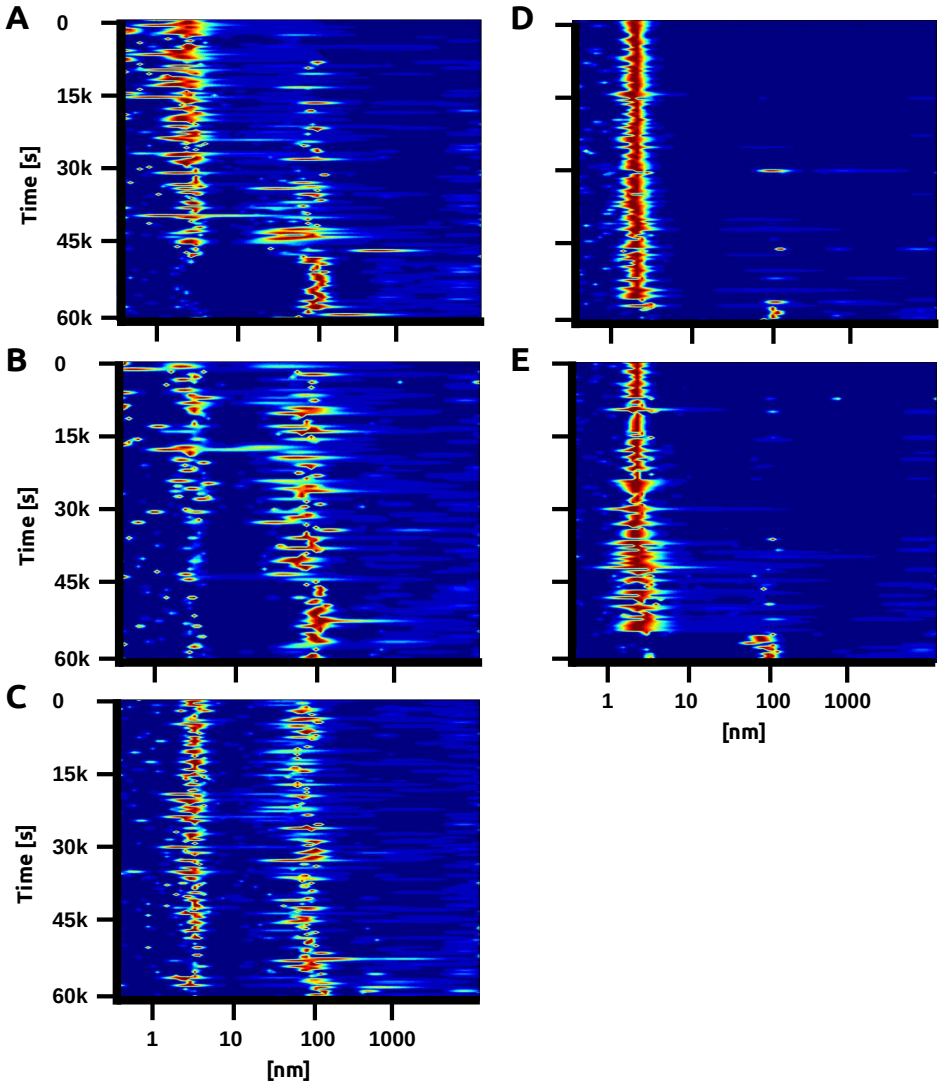

Figure 13.3: *Radius frequency distributions over time; 20 µM SOD1 in 50 mM SoAc buffer containing 100 mM DTT, 10 mM EDTA* ***(A)*** *pH3, 2 mg·mL^{-1}* ***(B)*** *pH3, 3.5 mg·mL^{-1}* ***(C)*** *pH3, 5 mg·mL^{-1}* ***(D)*** *pH 4, 2 mg·mL^{-1}* ***(E)*** *pH 3, no DTT, 2 mg·mL^{-1}; 20°C, 16 hrs; Red indicates high, blue indicates low counting frequencies*

a moderate concentration dependency when compared among samples A-C. It disappears during prolonged incubation for the lowest investigated concentration, but remains visible for higher concentrations. This fraction certainly is related to the homodimeric or an oligomeric form of SOD1, as it also appears in a control experiment at pH 4 (fig. 13.3-D) that did not yield any ThT response (cp. subsec. 13.1.1).
The fraction representing larger particles appears almost instantaneously in all cases A-C. The time of its appearance seem to be correlated with the concentration, as it is not present at the beginning of measurements with the lowest concentration. Instead the frequency distribution appears more frayed at early time points. Intriguingly, it is difficult to identify intermediate states, indicating a less gradual, but foremost fast and spontaneous process. However minor fractions of possible intermediate states can be identified from figure 13.3-A, which disappear with prolonged incubation times.
The identification of larger fragments appears to be difficult due to sedimentation processes within the droplet. A second control measurement lacking DTT, indicates slow kinetics as it has been observed previously (fig. 13.3-E). Only after an incubation of about 12 hrs the occurence of a larger particle fraction is visually identifiable from the frequency distributions. For measurements at pH 4 (fig. 13.3-D) a similar but less distinct radius distribution reveals even slower protein dynamics. However this conditions did not yield any fluorescence response in the ThT assay.
In conclusion the obtained results indicate a fast and instantanous formation of larger particles from small particles under conditions that already have been identified in a ThT assay to be in favour of fiber formation processes. In fact, for lower concentrations the small particle fraction vanishes over time, which is indicative for a non-equilibrium process, that consumes small particles due to energetic reasons. Apparently the observation of even larger particles is limited due to sedimentation and a lack of agitation. For conditions without DTT, a particle fraction corresponding to larger sizes only appears after several hours of incubation. It is noteworthy, that both, the lack of agitation and the change in incubation temperature, might change the aggregation kinetics or even hinders the assembly of larger particles on a larger scale.

13.1.3 Identification of Particle Species by Electron Microscopy

Fiber samples were investigated by electron microscopy to clarify particle shapes and to create visible proof of fiber formation by SOD1. Samples were prepared similarly to subsection 13.1.1 by incubation at 37°C in 50 mM sodium acetate buffer at different pH. Furthermore 100 mM DTT and/or 10 mM EDTA were added. Incubation was conducted for 3 days with additional sample agitation. The obtained results are depicted in figure 13.4-A-C.
Subfigure A represents conditions that already showed a positive response in both, the ThT assay and during DLS measurements, containing 100 mM DTT and 10 mM EDTA at pH 3. The electron micrographs reveal mostly small elongated particles at lengths of about 80-150 nm. Some of the particles are clustered or appear to be assembled into larger fiber-like structures, which show a strong curvature. These particles represent most likely amyloid protofibrils [187, 210]. A distinct difference is found for sample incubated at pH 4 as depicted in figure 13.4-B. Almost no fiber-like structures are visible in the obtained micrographs and a lot of globular particles with a diameter below 10 nm are found. This is in agreement with findings from subsections 13.1.1 and

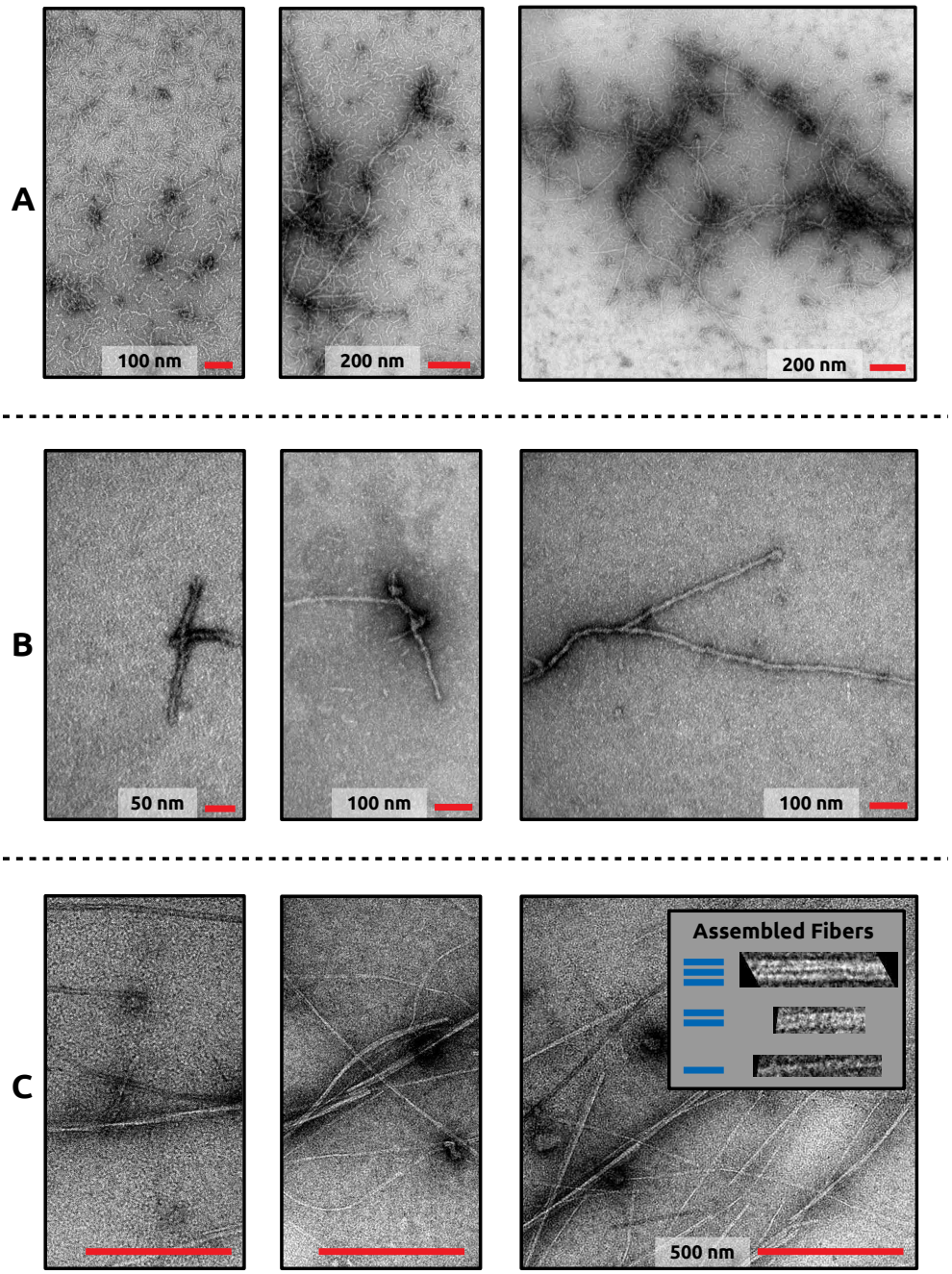

Figure 13.4: *Electron micrographs of SOD1 protofibrils & fibers; Samples obtained after 3 days of agitated incubation at 37°C of 20 µM SOD1 in 50 mM SoAc buffer* ***(A)*** *pH 3, 100mM DTT, 10 mM EDTA* ***(B)*** *pH 4, 100mM DTT, 10 mM EDTA* ***(C)*** *pH 3, 10 mM EDTA; Sizes as indicated*

13.1.2, in which neither fiber formation, nor the occurence of a distinct fraction of larger particle could be verified. At last figure 13.4-C depicts micrographs recorded for sample incubated with EDTA, but without DTT at pH 3. As it was previously found, this condition may lead to a response in the ThT assay after a longer lag phase. After 3 days of agitated incubation this condition apparently yields amyloid fibers at low background [153]. Most of the protein seems to be incorporated into the fiber structures. The fibers themselve show a clear tendency to form higher oligomers, as indicated by the image inlet.

Mono- di- and tri-mers arranged in a parallel arrangement can be identified. Hence the fibers seem not to build up twisted structures.

It is known from the literature that mature amyloid fibrils are composed of several protofibrils, each structured in a cross-βmotif. Amyloid fibers usually have diameters of around 10 nm, span several hundred nanometers in length and their occurrence is related to a complicated kinetic theory. Intriguingly, one protein can assemble into several structurally distinct fibrils, depending on the ambient conditions. Protein fibers can emerge due to *nucleation-dependent polymerization* (NDP), or by *linear polymerization*, that does not require nucleation events [153, 156].
Results obtained from ThT assays indicate a nucleation-dependent polymerization for SOD1 fiber formation. A short lag phase followed by a logistic signal increase is evidential from measurements shown in figures 13.1-A/B and 13.2. On the contrary, results obtained from DLS measurements (Fig. 13.3-A-C) indicate the almost instantanous occurrence of larger particles, which would in fact plead for an isodesmic (linear) polymerization process. But due to the slightly different experimental conditions in comparison to the conducted ThT assays, a direct correlation of the data seems difficult.
Still, it is known from other proteins to fibrilate in a process termed *assembly via oligomeric intermediates* [156]. Such process would start via the rapid formation of spherical oligomers or protofibrils that turn into mature fibers after prolonged incubation. The formation of such pre-fibrillar aggregates seems not to be limited by nucleation and indeed would show characteristics of an isodesmic polymerization, as it was found for the formation of larger particles by SOD1 in DLS measurements. Conclusively the formation of mature fibers could follow a different kinetic model for SOD1, as indicated by results obtained from ThT assays.
In general it is presumed that potofibrils are not an off-pathway species, but rather pre-fibrillar aggregates. But it is also known that protofibrils and fibers may form under different conditions for some proteins [153, 156]. For SOD1 it appears to be the case, that protofibrils are a direct precursor for mature fibers. As evidential from figure 13.4-A a lot of protofibrils are formed by SOD1, that eventually assemble into larger fiber structures (cp. fig. 13.4A-C). The lack of DTT under certain conditions seems to rapidly decrease the occurrence of initial protofibrils, but leads to well assembled fibers after prolonged incubation. The general decrease in nucleation seeds for fiber maturation seem to lead to fully assembled fibers in the end.

13.2 Serial-Femtosecond Fiber Diffraction Experiments with SOD1 Amyloids

Fiber diffraction experiments were carried out at beamline BL2 at SACLA (SPring-8 Angstrom Compact FEL, Japan). The beamline was equipped with an octal multiport CCD detector for data acquisition at 30 Hz [211]. A vacuum chamber, including a multi-nozzle cartridge for GDVNs, has been provided by the Chapman workgroup (CFEL, Hamburg) for sample delivery [226]. The FEL beam waist was about 2 µm in size. Furthermore, the FEL was maintained with a photon energy of 7.0 keV (1.77 Å) and a pulse length of 50 fs. The setup provided the possibility to introduce tilt to the flow angle of the liquid jet. This special property of the setup theoretically allows for better structural analysis

for non-randomly oriented samples like protein fibers. The injection of buffers and fiber-containing solutions has been tested similarly to section 10.4 prior to XFEL experiments (data not shown). Stable jets could be maintained at flow rates of 12-15 µL·min^{-1}. Samples were loaded into agitated and Pelletier-cooled sample reservoirs similar to yeast samples (sec. 10.5) to avoid sample settling. 10 mL of fiber solution at 2 $mg \cdot mL^{-1}$ were prepared on site at SPring-8. Fiber formation was verified by incubation with ThT.
The beamtime was intended to provide the possibility to optimise sample delivery of protein fibers to an ultra-intense X-ray beam by GDVNs. The applied nozzle technology does help to flow-align fibrillar structures in a low-background situation and hence allows for an easier and faster data acquisition. The utilization of femtosecond, coherent X-rays allows to overcome sample size limitations and circumvent radiation damage. Therefore, it was intended to obtain high quality diffraction data from SOD1 nano-fibers to possibly identify the molecular basis of amyloid formation by SOD1. Filamentous protein structures are a direct link between currently investigated protein nano- and micro crystals [163, 172, 212, 221, 253], and single particles [170]. Hence, their investigation can also be considered as a logical step in experiments with coherent X-rays to provide data at the limits of Bragg diffraction and diffuse scattering similar to experiments with imperfect crystals [250].
Unfortunately, no fiber diffraction data could be acquired during $3 \cdot 12$ hrs due to alignment problems between the experimental vacuum chamber and the X-ray beam. It was not possible to conduct any follow-up X-ray experiments.

V | Conclusions

The general objectives of this thesis have been explained in the motivation part and are further introduced in the chapters thereafter. It has been a major goal of the presented research project to establish a new general supply chain for nano-sized macromolecule crystals and related structures as promising samples which can be utilized in serial femtosecond X-ray crystallography (SFX) at free-electron laser (FEL) facilities around the world. Yeast *in vivo* crystallization has been primarily investigated as a tool to deliver crystals composed of nano-sized cellular machines for structural investigation, but also non-crystalline fiber structures have been brought up as possible intermediate samples to converge with other research approaches on the far goal of single-particle imaging.

Yeast as Potential *In Vivo* Crystal Factories & Carriers

As explained in section 2.1, the heterologous expression of proteins is well-established in yeast-based systems. Although natural *in vivo* crystallization has been observed not only in yeast, but also more specifically in peroxisomes across different species (cp. sec. 2.3), yeast peroxisomes have not yet been considered as general factories and carriers of crystalline material. Logically, this promising site of versatile character has been explored with regards to its capabilities in order to fulfill the necessary requirements for crystallization from a biological and physical standpoint.
At first, it has been tested whether high-quality diffraction data can be obtained not only from *in vivo*-grown, but rather *in cellulo*-situated nano-crystals. To accomplish this goal, the natural tendency of *H. polymorpha* cells to form peroxisomal AOX crystals has been exploited in powder diffraction (sec. 9.1) and SFX experiments (sec. 10.5). It is demonstrated that the dense cellular environment is not obstructing the acquisition of diffraction data, both in air at synchrotrons (cp. fig. 9.1) and under vacuum conditions at FELs (cp. fig. 10.9). Still, a strong limitation of detectable diffraction at higher Bragg angles has challenged the idea of high-quality data acquisition from crystals which are embedded in a thick layer of biological material. Fortunately, a dilution series of crystalline phase in amorphous phase yielded evidence that crystalline material can be easily detected *in cellulo*, even at cellular volume fractions on the order of one-tenth of a percent (subsec. 9.1.2). This finding is also supported by the poor diffraction quality of *in vitro*-grown AOX crystals, indicating a limitation due to crystalline disorder (sec. 8.2) rather than by X-ray attenuation. The available diffraction data was indexed to yield information for a general crystallographic classification (sec. 9.1). It was found that AOX crystals are organized in a rather large unit cell and that the total number of diffracting units is thus even more constrained in a nano-sized crystal. Therefore, it seems plausible to exclude disadvantageous SNR as prevalent limitations for yeast *in cellulo* diffraction experiments in the first place. Still, powder diffraction experiments indicated a rather disadvantagous SNR during experiments. Yet, we could refute the importance of this finding for SFX experiments in which diffraction beyond 10Å has been observed [253]. Further experiments also proved convenient handling of yeast cells for beam injection via GDVNs (sec. 10.4). Additionally, parameters for powder diffraction experiments in transmission geometry have been optimized for *in cellulo* AOX crystals.
It is obvious that the tremendous amount of standardized tools in protein crystallography can hardly be applied to the approach of direct *in cellulo* crystallization. This renders a difficult situation for preliminary sample assessment

with effortful experiments at FELs in view. Therefore, the Williamson-Hall plot analysis has been explored as a first investigative step to judge on the quality of obtained crystalline material. As shown in subsection 9.1.3, the method can easily be applied at modern synchrotron beamlines and gives a readout about crystal domain size and strain. It is also superior in comparison to an investigation via electron microscopy, because sample processing is minimized (cp. sec. 10.3). This enables the optimization of diffractive properties of crystalline material, as it is exemplarily shown for AOX crystals in figures 9.5 and 9.6 with regards to cellular growth time and incubation temperature. It is demonstrated that the size of AOX crystal domains is indeed increasing over time. Nevertheless, crystal strain is also increasing. Intriguingly, the comparison of incubation temperatures suggest larger crystal domain sizes at higher temperatures. Even though this finding finding seems contradictory to thermodynamics at first, retrograde solubility is a commonly observed phenomenon [17]. It implies the shift of a reaction's equilibrium free energy with the ambient temperature according to $\delta\Delta G^0/T/\delta T$ and is usually referred to as Le Chatelier principle. Among other proteins such behaviour is known from concanavalin A [223]. The obtained results seem to indicate the first observation of enhanced *in vivo* crystallization at higher temperatures. But further experimental evidence is required for a clarification. It fits into the concept that an increased AOX crystal domain sizes at higher temperatures go hand in hand with an increased number of crystallographic defects. This is opposed to the observed trend of strain reduction with incubation time.
As a logical next step, genetic engineering of *H. polymorpha* yeast cells has been conducted. Early in, the production of diffraction-quality *in vivo* crystals formed by heterologously expressed proteins in peroxisomes has been defined as the next achievable milestone of the project. For this purpose various promising protein candidates have been selected either due to reported crystallization success in other biological systems, or based on availability (cp. sec. 2.3, sec. 8.3 & 9.2). The DNA sequence of proteins was manipulated to yield proteins complemented with a PTSI peroxisomal targeting sequence (cp. sec. 2.2). Unfortunately, the succeeding *H.p.*-specific engineering process turned out to be error-prone, but a successful creation of multiple integration yeast clones was finally achieved. Genomic integration and expression in *H.p.* was exemplary verified for the protein *N.c.* Hex-1 (sec. 8.3 & 8.4). Additionally, genomic integration strains were created, that carry the sequence of the proteins PRDX5, Esat-6 and Cathepsin B. Further strains were obtained, that facilitate the expression of EGFP and lysozyme. All clones are based on a triple-deletion strain, which is genetically modified to carry gene deletions with regards to the three major peroxisomal matrix proteins AOX, DHAS and CAT. Further measures were taken to expand the experimental efforts to a cross-validation approach, utilizing the yeast *S. cerevisiae* to serve the same purpose of *in vivo* crystallization. It might be hypothesized that both yeast systems are highly similar, but they are not too closely related on the phylogenetic tree. While peroxisomal targeting is conserved among both species, maximum expression rates had to be achieved differently, either by methanol- or oleic acid induction. Since crystal formation of AOX is a natural process in *H.p.* cells, it was attempted to accomplish artifical crystal formation in *S.c.* peroxisomes. Additionally, heterologous expression and peroxisomal targeting of Hex-1 was tested in *S.c.*, since experiments by Würtz *et al.* [140] suggested possible success. In a collaborative attempt it was possible to verify peroxisomal targeting and strong

expression for both *N.c.* Hex-1 and *H.p.* AOX protein in *S.c.* cells. Although several growth conditions have been tested with all available expression strains from both species (cp. tab. 9.5), crystallization of further proteins could not be achieved. Varied conditions ranged from different growth media over incubation times to different growth temperatures.

Interpreting Intraperoxisomal Probe Molecule Behaviour & Concentration Measurements

So far it seemed to be a profound hypothesis that yeast *in vivo* crystallization is feasible due to the strong protein expression via the utilized genetic promoters (see sec. 2.1). Therefore, nucleation and crystallization in yeast peroxisomes was assumed to be initiated by a steady accumulation of proteins in solution, which at some point would energetically favor the formation of stable nuclei and crystal growth (cp. sec. 1.5). The failure of generating further *in vivo* crystals strongly suggested to reassess this assumption.
To successfully investigate the matter, a novel approach based on the utilization of intraperoxisomal EGFP probes in *H. polymorpha* has been developed (sec. 11.1). The fluorescent probe molecules are imported into peroxisomes of a triple-deletion *H.p.* strain via the PTSI pathway and model the principle behaviour of a globular protein in such an environment during peroxisomal biogenesis. The probe molecules enable steady-state fluorescence anisotropy measurements, that are convolved with a potential energy transfer between probe molecules. The readout can be interpreted in terms of relative probe concentration, molecular interactions and viscosity of the environment.
A theoretical model is initially derived in subsection 11.1.1 to describe the behaviour of the probe in a structure with lysosomic characteristics. The change in fluorescence anisotropy and energy transfer is modeled in relation to the molecular movement (tumbling) and probe proximity. The mathematical considerations assumed the applicability of the Perrin- and SED equation (cp. eq. 3.3). Energy transfer was modeled as an isotropic process. It has been additionally assumed that Förster's original solution for energy transfer in solution is valid (eq. 3.14). Properties of the EGFP probe molecule were modeled according to values from the literature (cp. tab. 11.2). At last, dimerization kinetics were incorporated into the model for comparison (eq. 11.8).
Further on, a scheme has been worked out to obtain experimental *in cellulo* data to which the model can be compared (subsec. 11.1.4). Additionally, interpretation of data is cross-validated with a statistical analysis of fluorescence microscopy data (sec. 11.2). The comparison of experimental and theoretical data yields new and interesting insights into peroxisomal dynamics.
Although the obtained results indicate a net increase of the total cellular probe concentration, the average intraperoxisomal concentration is simultaneously decreasing (see fig. 11.21). It can be suggested that organelle-specific processes prevent the development of higher concentrations. Firstly, it appears that a constant replenishment of bilayer phospholipids leads to a constant increase in the average size of peroxisomes regardless of organellar fission. Evidence for a non-vesicular phospholipid transport between ER and peroxisomes has been brought up by Raychaudhuri & Prinz [137] and might explain the general size increase of peroxisomes. This would also favor a situation in which both old and new organelles are targeted by phospholipid supplies. Since the general production of probe molecules continues into late stationary phase, an energetic

limitation on a cellular level is not sufficient to explain non-increasing concentrations. This is also true for the peroxisomal level, since findings by Palmieri *et al.* [76] indicate the presence of the Pex3p/Pex19p-dependent ATP-transporter Ant1p in *S.c.* peroxisomal membranes. Likewise, the PTSI pathway is known to depend on ubiquitination and ATP hydrolysis processes on the cytosolic side of the membrane [208, 258]. Additionally, measurements by Ozalp *et al.* [157] show a high concentration of ATP in starved yeast cells. If uptake is theoretically not limited due to a lack of energy sources, it might either be regulated by unknown processes, or protein expression and lipid replenishment are geared to each other. In such a case, the usage of different expression promoters or multiple genomic integration sites might increase protein production and intraorganellar concentrations of such [184]. However, expression trials of Hex-1 and AOX in *S.c.* effectuated by the P_{PGK} promoter seem not to induce crystallization, although peroxisome proliferation is not induced. In general, the cellular system of *H. polymorpha* counteracts on the increasing expression of peroxisome-targeting probe molecules by organellar fission under typical methanol conditions. Unfortunately, only weak promoters are available for conditions which do not induce peroxisome proliferation.
While it seems obvious that ΔPEX11 strains could render a diverging situation, the presented evidence and other studies can be used to falsify this assumption. As pointed out by Van Dijken *et al.* [20] and later by Roggenkamp *et al.* [32], *H.p.* peroxisomes tend to increase in size depending on the specific conditions and demand for intraorganellar space. Furthermore, Krikken *et al.* [147] identified unnaturally enlarged peroxisomes in Pex11p deletion strains, which indicates a direct redirection of bilayer supplies to the sole left organelle. In conclusion, a steep increase in concentration seems only feasible upon advanced engineering of mechanisms involved in membrane regulation and not just organellar fission. The presented findings also explain the appearance of cytosolic AOX crystals in ΔPEX5 cells (sec. 10.5), which was also formerly reported by v.d. Klei *et al.* [38]. It seems obvious to assume that AOX crystals do not appear due to extremely high protein concentrations but rather due to an endogenous tendency for critical nucleation under natural conditions. On the contrary, the lack of peroxisomal AOX crystals upon heterologous expression in *S.c.* cells seems contradictory. However, degradation and host-specific matrix protein expression render a completely different situation. It is known that peroxisomes exhibit increased densities among organelles [71]. Since this is true for peroxisomes of different species with and without crystals, a direct correlation of crystalline material with the increased density is not self-evident. Intriguingly, *in vitro* crystallization trials of *H.p.* AOX have only been successful in solutions with high viscosities, which in turn appear to resemble the dense peroxisomal lumen (sec. 8.2).
The recent success of the Sf9/Baculovirus system in generating *in vivo*-grown crystals (cp. sec. 2.3) seems to be related to the ability of the virus-infected system to divert cellular resources solely to protein expression, actively breaching the cell's capability to maintain homeostasis. In contrast, yeast cells will follow their generic metabolic and genetic pathways to oppose any 'unnatural' changes to their interior and may require far more genetic engineering to be positioned as competitors in the field (sec. 2.2). Such engineering attempts could also implement tools from optogenetics. Following the ideas of Rost *et al.* [241], molecular cloning was conducted to create a Pex3p-based fusion protein comprising Archaerhodopsin-3 and pHIurion (data not shown). This could be

used upon successful expression and targeting to function as a light-triggered switch to (de)acidify the peroxisomal lumen. Similar efforts are required to fine-tune intraperoxisomal conditions for crystallization. Future attempts for yeast *in vivo* crystallization may also comprise targeting of different organelles to cover different naturally accessible conditions. Despite the required genetic fine-tuning of the yeast system, it has proven outstanding qualities as a general crystal carrier.
The presented approach of fluorescence anisotropy can be improved by time-resolved measurements on the ns scale. Especially the implementation of time-resolved fluorescence anisotropy imaging microscopy (TR-FAIM) [219] could yield more information about protein behaviour in an *in vivo* crystallization environment not only in yeast-, but also in insect- and mammalian cells. Even the investigation of nucleation and crystallization dynamics seems feasible, depending on the design of the experiment and the resolution of the setup.

Potential Approach to Discern Protein Clusters & Crystals

This general idea is closely related to the experiments conducted by Jullien [31] who reported the utilization of steady-state fluorescence anisotropy to observe the *in vitro* pre-nucleation phase of ribonuclease A. Although the requirements for such measurements already seem strictly defined, the idea might be applicable in a broader sense. Furthermore, the readout might provide greater detail than e.g. DLS measurements. Steady-state or time-resolved fluorescence anisotropy measurements were already applied to observe the formation process of general protein clusters/oligomers [143]. However, most published studies only consider an isotropic resonance energy transfer between fluorescent probe molecules. Taking the FRET orientation factor κ^2 (sec. 3.2) into account may provide the basis of a novel method to differentiate between protein clusters and crystals. This possibility is briefly discussed in a *gedankenexperiment* in chapter 12. For a proof-of-claim, the rate of energy transfer is described for four fixed 2D molecular arrangements. It clearly deviates from the classical assumption, which is only valid for completely random orientations. Afterwards, the principle is extended to the 3D case by taking the example of a $P2_12_12_1$ protein crystal. Most importantly, ideas are brought up to extend such measurements to non-fluorescent proteins. The idea could be advanced into an applicable method, especially relevant for *in vivo* crystallization.

A Plate Reader-based Assay to Investigate System Perturbations

Following up on the possibility to relate intraperoxisomal probe concentration and fluorescence anisotropy, a plate reader-based assay has been developed to make use of this relationship (subsec. 11.3.3). Here, external chemical stimuli (perturbations of the milieu) were introduced in order to test the response of the *H.p.* peroxisomal system. As reported for many cases (see subsec. 11.3.3), external stimuli can influence the intracellular environment and thus could change it to yield favorable conditions for *in vivo* crystallization. Again, a *H.p.* triple-deletion strain, that is expressing an EGFP probe has been utilized in order to catch transitions within peroxisomes. But in order to provide an easily-applicable assay for screening, an unbiased readout of the EGFP fluorescence signal had to be achieved. To circumvent elaborate sample

treatment, it was necessary to develop an approach to reduce background fluorescence in the growth media. Most of it originates from Flavin-like molecules which are secreted by yeast cells during growth [134]. Different fluorescence quenchers were tested for this purpose. KBr was found to be not interfering with cellular integrity and EGFP fluorescence but to be highly effective as a quenching agent. Considering further dependencies of intensity, optical density and anisotropy, it is possible to reliably generate comparable readouts. System perturbations were introduced by means of pH, oxidative stress, osmolarity, carbon source, intracellular precipitants and growth inhibitors. However, the results for most conditions reveal a similar correlation among fluorescence intensity and anisotropy. This implies an unbiased development of the peroxisomal system with regards to promoter activation and peroxisome proliferation. Intriguingly, cellular growth seems not strongly correlated with peroxisomal development. Only an increasingly basic pH of the growth media led to an alternate relationship (subsec. 11.3.3). Still, none of the tested conditions was found to induce higher peroxisomal probe concentrations, a prerequisite for crystallization.

In Vitro Production of SOD1 Nanofibers

Protein nanofibers are the connecting link between classical X-ray crystallography and single-particle imaging as they only contain 1D translational symmetry. This thesis presents an approach to reliably produce SOD1 amyloid fibers which can be used for structural investigation at FELs. Fiber formation has been investigated by means of DLS, a Thioflavin T assay and electron microscopy. Strong indication is given that SOD1 amyloids form by *assembly via oligomeric intermediates*, comprising the formation of pre-fibrillar structures via *isodesmic oligomerization* and a subsequent rather nucleation-dependent formation of full amyloid-like fibers (cp. subsec. 13.1.2 & 13.1.3). Unfortunately, technical difficulties prevented the acquisition of diffraction data from single flow-aligned nano-fibers. Still, fibrillar structures remain a hot topic in the FEL community, although fiber X-ray diffraction from multiple, aligned fibers has been used before in structural investigation [138]. As of recently, studies have been published which demonstrate successful single-fiber diffraction and data analysis at a FEL X-ray source [266, 270].

A | Appendix

Table A.1: *Williamson-Hall plot data, Comparison of H.p. wildtype- (WT) and ΔPEX11-grown AOX crystals; Fitted and corrected FWHM (2Θ) of incorporated hkl reflections;* $B = {}^{FWHM}/_{2}$

Cell line	**Time [hrs]**	**hkl Refl.**	**FWHM** $[10^{-3}\ °]$	$\boldsymbol{sin\Theta}$ $[10^{-3}]$	$\boldsymbol{B \cdot cos\Theta}$ $[10^{-3}]$
WT	6	(1 1 0)	8.398	3.026	4.199
		(2 0 0)	7.586	4.227	3.793
		(3 1 0)	8.625	6.699	4.312
		(3 2 1)	9.639	7.930	4.819
		(4 0 0)	8.271	8.464	4.135
		(4 2 0)	6.241	9.464	3.120
WT	24	(1 1 0)	4.929	3.008	2.464
		(2 0 0)	4.950	4.232	2.475
		(3 1 0)	5.384	6.694	2.692
		(3 2 1)	5.581	7.922	2.791
		(4 0 0)	6.337	8.462	3.169
		(4 2 0)	6.000	9.463	3.000
ΔP11	6	(2 0 0)	7.932	4.229	3.966
		(3 1 0)	7.361	6.704	3.680
		(3 2 1)	4.936	7.948	2.468
		(4 0 0)	7.963	8.466	3.981
ΔP11	24	(1 1 0)	2.847	3.010	1.423
		(2 0 0)	3.444	4.236	1.722
		(3 1 0)	4.466	6.700	2.233
		(3 2 1)	4.716	7.927	2.358
		(4 0 0)	5.230	8.470	2.615
		(4 2 0)	5.587	9.469	2.794

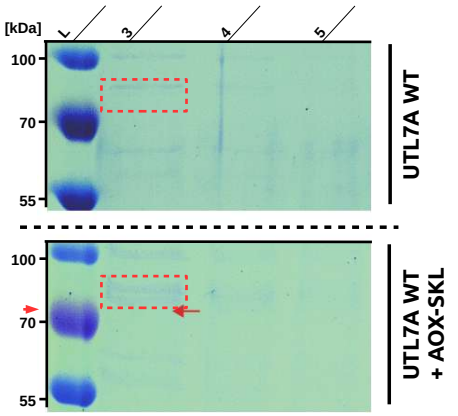

Figure A.1: *Coomassie-blue stained SGS-PAGE gel depicting* ***(Lower)*** *quantitative expression of AOX in S.c. UTL7A mutants, transformed with the pMSC13 vector (P_{FOX3} promoter) and* ***(Upper)*** *a S.c. UTL7A wildtype control*

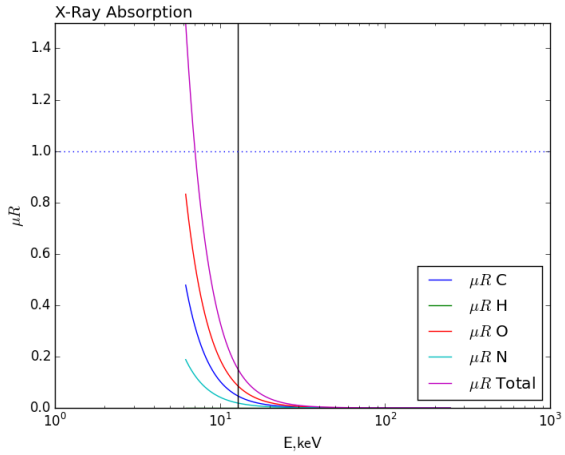

Figure A.2: *Theoretical X-ray absorption of the elemental composition $CH_{1.8}O_{0.5}N_{0.2}$ and its constituents at a density of $1.09 g/cm^3$ and a beam energy of 12.8 keV (black line); Data calculated as stated in subsec. 9.1.1*

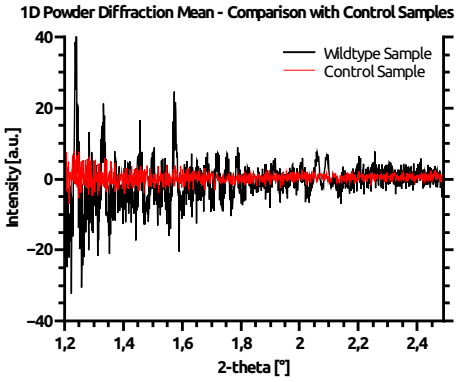

Figure A.3: *Post-processed 1D powder pattern from H.p. wildtype (black) and 12DAC4 control (red) yeast cells 24 hrs after induction of crystal formation*

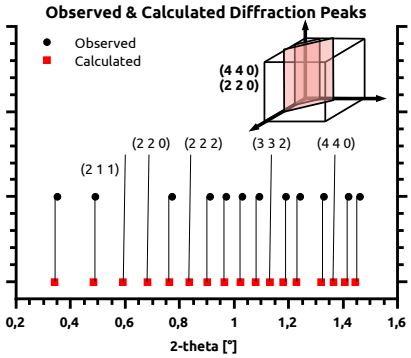

Figure A.4: *Observed versus calculated hkl reflections for a body(I)-centered cubic lattice of AOX in vivo crystals (Image Inlet) Schematic indication of anisotropic disorder in AOX crystals due to missing hkl reflections*

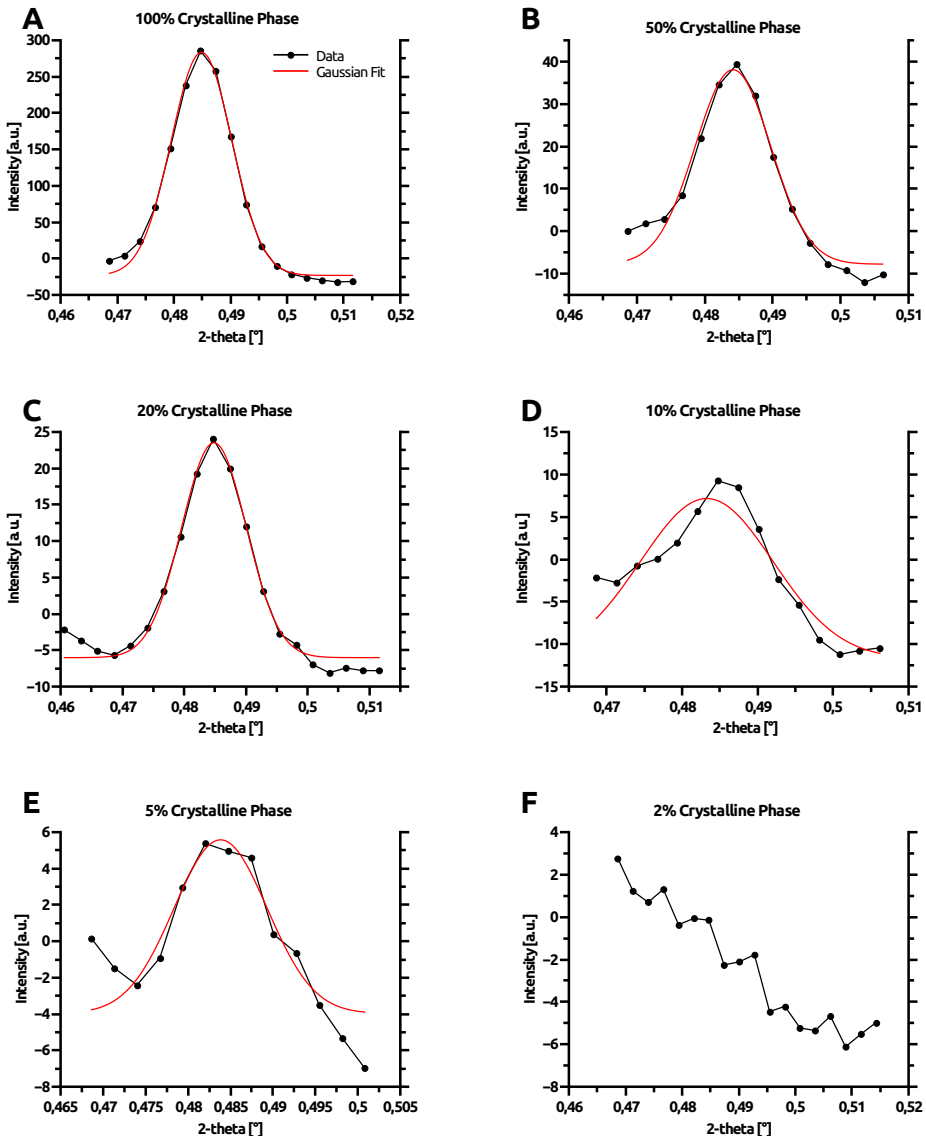

Figure A.5: *100% peak intensity/shape and manual fitting at different volume packing fractions of crystalline phase; Data from 1D powder patterns of in vivo AOX crystals*

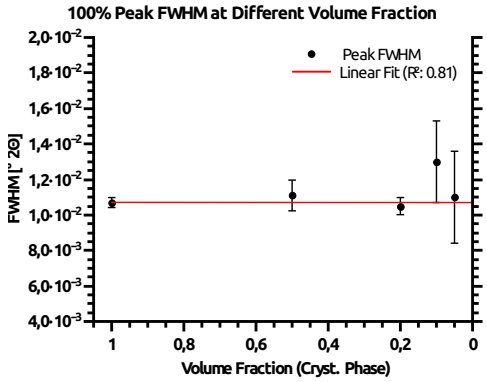

Figure A.6: *100% peak FWHM at different volume packing fractions of crystalline phase, Indication of reliable peak fitting at lower scattering intensities; Data from 1D powder patterns of in vivo AOX crystals*

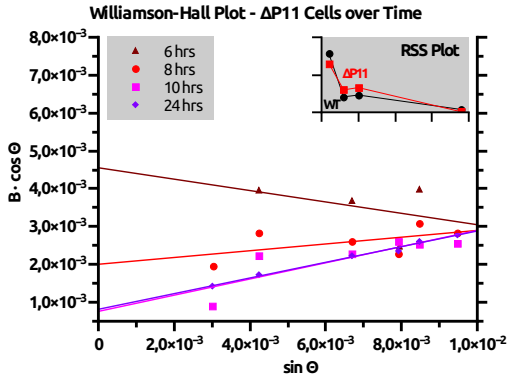

Figure A.7: *Time-resolved WH plot analysis of ΔPEX11 cells containing AOX crystals (Image Inlet) Residual sum of squares (RSS) for data points and linear fit model for different time points, fit accuracy is increasing over time*

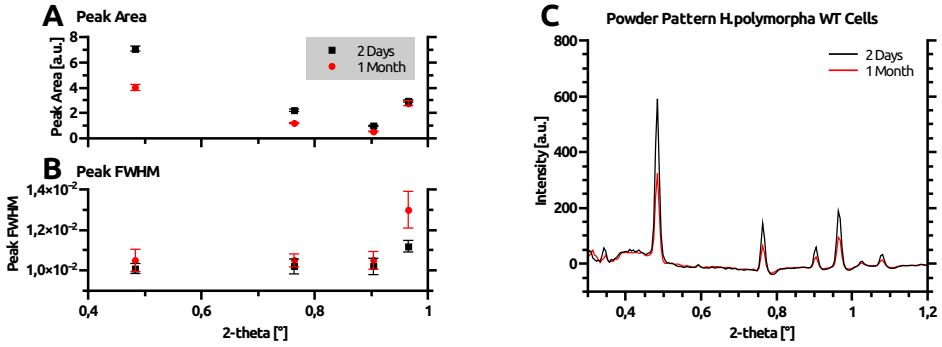

Figure A.8: *Comparison of **(A)** peak intensity and **(B)** FWHM during prolonged storage of in vivo AOX crystals in yeast cells stored at 4°C **(C)** Corresponding 1D powder patterns*

Table A.2: *Williamson-Hall plot data, Comparison of H.p. wildtype-grown AOX crystals at different growth temperatures; Fitted and corrected FWHM (2Θ) of incorporated hkl reflections;* $B = {}^{FWHM}/_{2}$

Growth Temp. [°C]	**hkl Refl.**	**FWHM [10^{-3} °]**	$\boldsymbol{sin\Theta}$ **[10^{-3}]**	$\boldsymbol{B \cdot cos\Theta}$ **[10^{-3}]**
25	(1 1 0)	4.584	3.029	2.292
	(2 0 0)	5.069	4.212	2.534
	(3 1 0)	4.719	6.664	2.360
	(3 2 1)	4.163	7.888	2.081
	(4 0 0)	5.470	8.419	2.735
	(4 2 0)	4.796	9.413	2.398
37	(1 1 0)	4.275	2.978	2.138
	(2 0 0)	4.103	4.216	2.051
	(3 1 0)	4.273	6.663	2.136
	(3 2 1)	4.122	7.884	2.061
	(4 0 0)	5.434	8.421	2.717
	(4 2 0)	6.296	9.416	3.148
45	(1 1 0)	3.987	2.961	1.993
	(2 0 0)	4.213	4.214	2.107
	(3 1 0)	4.991	6.661	2.495
	(3 2 1)	4.123	7.883	2.061
	(4 0 0)	5.931	8.417	2.966
	(4 2 0)	7.586	9.406	3.793

Table A.3: *DLS data analysis, CONTIN analysis results of purified peroxisomes from H.p. wildtype and ΔPEX11 cells, indicating different peak fractions*

Peak	**Mean Radius [nm]**	**PDI [%]**	**Amplitude**
WT	**- mixing**		
1	239.96	13.5	0.217
2	799.20	22.3	0.311
3	4467.97	22.1	0.071
WT	**+ mixing**		
1	217.66	27.9	0.465
2	908.64	35.5	0.163
ΔPEX11	**- mixing**		
1	149.45	10.1	0.010
2	581.92	29.1	0.430
3	2585.51	37.8	0.170
ΔPEX11	**+ mixing**		
1	150.03	10.5	0.084
2	377.88	31.0	0.524
3	2463.96	27.6	0.132

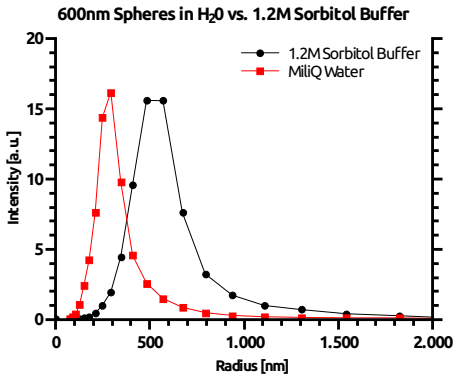

Figure A.9: *Control DLS measurements with Duke Standard 600 nm NIST Traceable Polymer Microspheres (Thermo Scientific, USA) in water and 1.2 M sorbitol with an aqueous viscosity parameter*

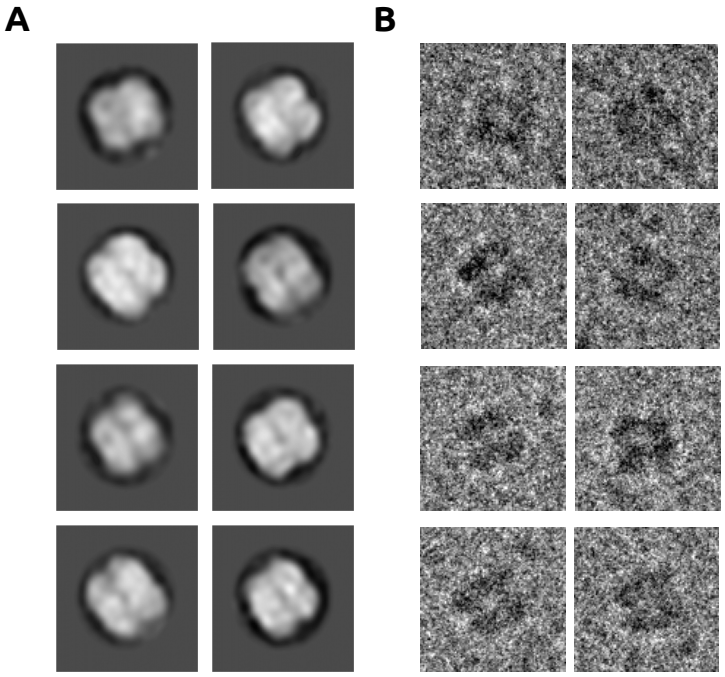

Figure A.10: ***(A)*** *Class-sums of particles in negative stain-EM possibly representing loose AOX octamers* ***(B)*** *Cryo-EM, single images of particles possibly representing AOX octamers*

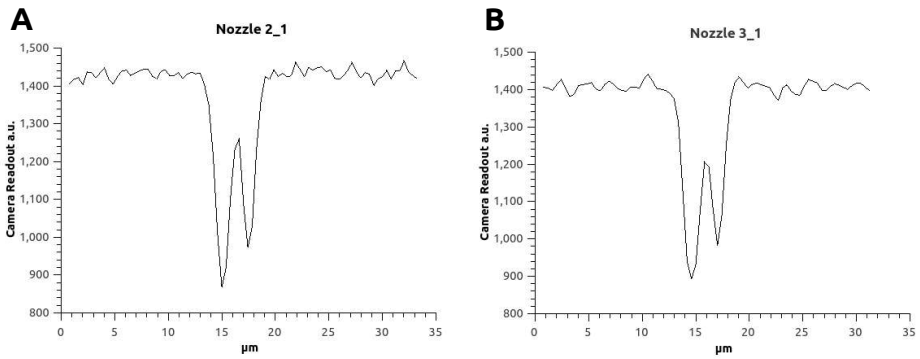

Figure A.11: *GDVN jet diameter approximation from linear plots in vertical direction of the liquid flow* ***(A,B)*** *Representation of jets from two different nozzles as stated in sec. 10.4*

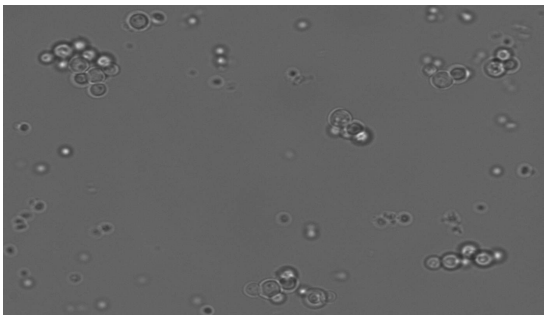

Figure A.12: *Verification of yeast cell integrity after jetting in a typical GDVN setup as described in sec. 10.4*

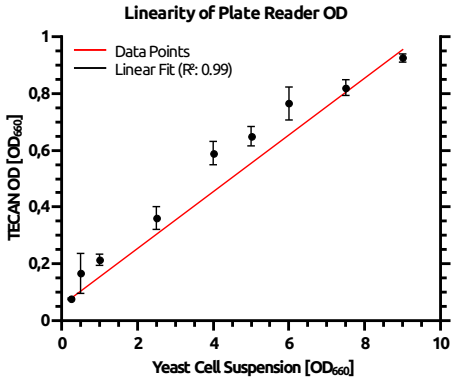

Figure A.13: *Linear dependency of plate-reader based and undiluted OD measurements (y-axis) in comparison to diluted OD measurements on a spectrophotometer (x-axis)*

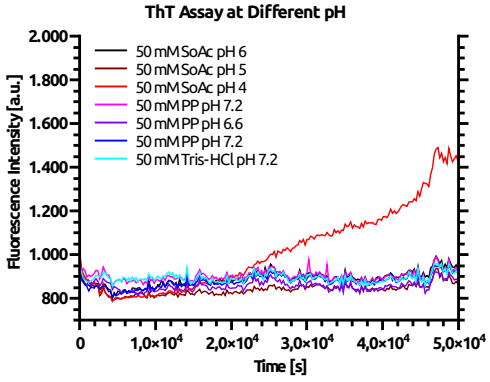

Figure A.14: *ThT assay, control measurements to investigate the pH influence on the SOD1 fiber formation process; Experimental conditions are specified in subsec. 13.1.1*

Table A.4: *DNA Primers in 5'-3' direction used for SLiCE cloning of pHIPZ4/pHIPX4 cloning of target protein sequences, spacer peptides and SKL tag*

Name	Description	5'-3' Sequence
Pex11Verify-FWD	Verif. of Pex11p Deletion	GTCCAATCCGCGTTCTCCTC
Pex11Verify-RV	Verif. of Pex11p Deletion	GCGACTGATTCGGCAAGATG
ShuttleSeq-FWD	Sequencing primer	GACATCCCCAGTTTTTGCCC
ShuttleSeq-RV	Sequencing primer	GTAGGAAGGCTGGATGTCCA
ShuttleSeq2-Fwd	Sequencing primer	CAGCTGGATGGTGTCATGTG
Spacer-B1	Spacer Cloning	GACATCAATCTAAAGTACAAAAACGGTATGGACGAATTGTACAAGTCGAAGCTGTAATCTAGAGGATCGATCCCCGGGCC
Spacer-B2	Spacer Cloning	GGCCCGGGGATCGATCCTCTAGATTACAGCTTCGACTTGTACAATTCGTCCATACCGTTTTTGTACTTTAGATTGATGTC
Spacer-B3	Spacer Cloning	GACATCAATCTAAAGTACAAAAACAAGCTTGGTATGGACGAATTGTACAAGTCGAAGCTGTAATCTAGAGGATCGATCCCCGGGCC
Spacer-B4	Spacer Cloning	GGCCCGGGGATCGATCCTCTAGATTACAGCTTCGACTTGTACAATTCGTCCATACCAAGCTTGTTTTTGTACTTTAGATTGATGTC
Hex1-FWD	Hex-1 SLiCE	TCAATCTAAAGTACAAAAACGATCTATGGGCTACTACGAC
Hex1-RV	Hex-1 SLiCE	TTGTACAATTCGTCCATACCGAGGCGGGAACCGTGGACGA
PRDX5-FWD	PRDX5 SLiCE	TCAATCTAAAGTACAAAAACATGTCAAGAAATTTTCAAAC
PRDX5-RV	PRDX5 SLiCE	TTGTACAATTCGTCCATACCCAACTTTGCTAGAACCCTGT
CatB-FWD	Cathepsin B SLiCE	GTGGTGACATCAATCTAAAGTACAAAAACATGGCCCTCGTTGCTGAAGAC
CatB-RV	Cathepsin B SLiCE	CAGCTTCGACTTGTACAATTCGTCCATACCCAATTTGGACGCCGTGTTGG
Esat6-FWD	Esat-6 SLiCE	GTGGTGACATCAATCTAAAGTACAAAAACATGGCTCAGATCAAGCTGACCC
Esat6-RV	Esat-6 SLiCE	CAGCTTCGACTTGTACAATTCGTCCATACCACCGGAGATCTGGGAAGC
Lsy-FWD	T4 Lysozyme SLiCE	GTGGTGACATCAATCTAAAGTACAAAAACATGAATATATTTGAAATGTTACG
Lys-RV	T4 Lysozyme SLiCE	CAGCTTCGACTTGTACAATTCGTCCATACCTAGATTTTTATACGCGTCC
Luc-FWD	Luciferase SLiCE	GTGGTGGACATCAATCTAAAGTACAAAAACATGGAAGACGCCAAAAACAT
Luc-RV	Luciferase SLiCE	CAGCTTCGACTTGTACAATTCGTCCATACCCAATTTGGACTTTCCGCCCT
MBP-FWD	MBP SLiCE	GTGGTGGACATCAATCTAAAGTACAAAAACATGGAAACTGAAGAAGGTAA
MBP-RV	MBP SLiCE	CAGCTTCGACTTGTACAATTCGTCCATACCACCGCTGCCGGTACCAGTCT
DAO-FWD	DAO SLiCE	GTGGTGGACATCAATCTAAAGTACAAAAACATGACTGATTCCAAATACGT
DAO-RV	DAO SLiCE	CAGCTTCGACTTGTACAATTCGTCCATACCTTTGGCAGCTTCAAAATACA
GenAOX-FWD	Genome Primer P_{AOX}	TCGACGCGGAGAACGATCTC
AOX-RV	Primer P_{AOX}	GATGTCACCACCGTGCACTG

References

1. Bravais, A. Mémoire sur les systémes formées par des points distribués réguliérement sur un plan ou dans l'espace. *J. Éc. Polytech.* **19,** 1–128 (1850).
2. Gibbs, J. W. *Elements of Vector Analysis Arranged for the Use of Students in Physics* 92 (New Haven Printed by Tuttle, Morehouse & Taylor, 1884).
3. Röntgen, W. C. Über eine neue Art von Strahlen. *Annalen der Physik* **300,** 1–11 (1898).
4. Bragg, W. H. The Reflection of X-rays by Crystals. (II.) *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **89,** 246–248 (1913).
5. Friedrich, W., Knipping, P. & Laue, M. Interferenzerscheinungen bei Röntgenstrahlen. *Annalen der Physik* **346,** 971–988 (1913).
6. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Die Naturwissenschaften* **33,** 166–175 (1948).
7. Förster, T. Experimentelle und theoretische Untersuchung des zwischenmolekularen Übergangs von Elektronenanregungsenergie. *Z.Naturforsch.* **4A,** 321–327 (1949).
8. Haddad, S. A. & Lindegren, C. C. A Method for Determining the Weight of an Individual Yeast Cell*. *Applied and Environmental Microbiology* **1,** 153–156 (1953).
9. Williamson, G. K. & Hall, W. H. X-ray line broadening from filed aluminium and wolfram. *Acta Metallurgica* **1,** 22–31 (1953).
10. GPD. *Physical Properties of Glycerine and Its Solutions* 27 (Glycerine Producers' Association, New York, 1963).
11. Brotherton, J. Lack of Swelling and Shrinking of Pityrosporum ovale in Media of Different Osmotic Pressures and its Relationship with Survival in the Relatively Dry Conditions of the Scalp. *J Gen Microbiol* **48,** 305–308 (1967).
12. Matthews, B. Solvent content of protein crystals. *Journal of Molecular Biology* **33,** 491–497 (1968).
13. Fahimi, H. D. Cytochemical localization of peroxidatic activity of catalase in rat hepatic microbodies (peroxisomes). *Journal of Cell Biology* **43,** 275–288 (1969).
14. Frederick, S. E. & Newcomb, E. H. Cytochemical localization of catalase in leaf microbodies (peroxisomes). *Journal of Cell Biology* **43,** 343–353 (1969).

15. Craver, F. & Knox, R. Theory of polarization quenching by excitation transfer II. Anisotropy and second-neighbour considerations. *Mol. Phys.* **22,** 385–402 (1971).

16. Yee, B., Tsuyumu, S. & Adams, B. G. Biological effects of dimethyl sulfoxide on yeast. *Biochemical and Biophysical Research Communications* **49,** 1336–1342 (1972).

17. Lilje, K. C. & Macomber, R. S. Recrystallization: unexpected behavior. An undergraduate organic demonstration or experiment. *L. Chem. Educ.* **50,** 567–568 (1973).

18. McGinnis, M. R., Padhye, a. a. & Ajello, L. Storage of stock cultures of filamentous fungi, yeasts, and some aerobic actinomycetes in sterile distilled water. *Applied microbiology* **28,** 218–222 (1974).

19. Reynolds, R. J. The Lorentz factor for basal relfections from micaceous minerals in oriented powder aggregates. *American Mineralogist* **61,** 484–491 (1976).

20. Van Dijken, L. P., Otto, R. & Harder, W. Growth of Hansenula polymorpha in a methanol-limited chemostat - Physiological responses due to the involvement of methanol oxidase as a key enzyme in methanol metabolism. *Archives of Microbiology* **111,** 137–144 (1976).

21. Eggeling, L. & Sahm, H. Derepression and partial insensitivity to carbon catabolite repression of the methanol dissimilating enzymes in Hansenula polymorpha. *European Journal of Applied Microbiology and Biotechnology* **5,** 197–202 (1978).

22. Künkel, W. Effects of the antimicrotubular cancerostatic drug nocodazole on the yeast Saccharomyces cerevisiae. *Zeitschrift für allgemeine Mikrobiologie* **20,** 315–324 (1980).

23. Cromer, D. T. & Liberman, D. A. Anomalous dispersion calculations near to and on the long-wavelength side of an absorption edge. *Acta Crystallographica Section A* **37,** 267–268 (1981).

24. Veenhuis, M., Harder, W., van Dijken, J. P. & Mayer, F. Substructure of crystalline peroxisomes in methanol-grown Hansenula polymorpha: evidence for an in vivo crystal of alcohol oxidase. *Molecular and cellular biology* **1,** 949–57 (1981).

25. Kingsman, S. M., Kingsman, A. J., Dobson, M. J., Mellor, J. & Roberts, N. A. Heterologous Gene Expression in Saccharomyces cerevisiae. *Biotechnology and Genetic Engineering Reviews* **3,** 377–416 (1985).

26. Reynolds, R. C. The Lorentz-Polarization Factor and Preferred Orientation in Oriented Clay Aggregates. *Clays and Clay Minerals* **34,** 359–367 (1986).

27. Huang, T. C., Hart, M., Parrish, W. & Masciocchi, N. Line-broadening analysis of synchrotron x-ray diffraction data. *Journal of Applied Physics* **61,** 2813–2816 (1987).

28. Herskowitz, I. Life cycle of the budding yeast Saccharomyces cerevisiae. *Microbiol Rev* **52,** 536–553 (1988).

29. Giuseppin, M. L. Mathematical modeling of growth and production of alcohol oxidase by Hansenula polymorpha grown on methanol–glucose mixtures. *Biotechnology and Bioengineering* **33,** 524–535 (1989).

30. Höfte, H & Whiteley, H. R. Insecticidal crystal proteins of Bacillus thuringiensis. *Microbiological reviews* **53,** 242–255 (1989).

31. Jullien, M. Early steps in the crystallization process of proteins: An approach by fluorescence anisotropy. *FEBS Letters* **253,** 38–42 (1989).

32. Roggenkamp, R, Didion, T & Kowallik, K. V. Formation of irregular giant peroxisomes by overproduction of the crystalloid core protein methanol oxidase in the methylotrophic yeast Hansenula polymorpha. *Molecular and Cellular Biology* **9,** 988–994 (1989).

33. Van der Klei, I. J. *et al.* Use of electron microscopy in the examination of lattice defects in crystals of alcohol oxidase. *FEBS Letters* **244,** 213–216 (1989).

34. McPherson, A. Current approaches to macromolecular crystallization. *European Journal of Biochemistry* **189,** 1–23 (1990).

35. Van der Klei, I. J., Bystrykh, L. V. & Harder, W. Alcohol oxidase from Hansenula polymorpha CBS 4732. *Methods in enzymology* **188,** 420–427 (1990).

36. Van der Klei, I. J., Rytka, J., Kunau, W. H. & Veenhuis, M. Growth of catalase A and catalase T deficient mutant strains of Saccharomyces cerevisiae on ethanol and oleic acid - Growth profiles and catalase activities in relation to microbody proliferation. *Archives of Microbiology* **153,** 513–517 (1990).

37. Gellissen, G. *et al.* Heterologous gene expression in Hansenula polymorpha: Efficient secretion of glucoamylase. *Bio/Technology* **9,** 291–295 (1991).

38. V.d. Klei, I., Sulter, G. J., Harder, W. I. M. & Veenhuis, M. Assembly of Alcohol Oxidase in the Cytosol of a Peroxisome-Deficient Mutant of Hamenula poZymorpha- Properties of the Protein and Architecture of the Crystals. *Yeast* **7,** 15–24 (1991).

39. Zaar, K., Volkl, A. & Fahimi, H. D. Purification of marginal plates from bovine renal peroxisomes: Identification with L-α-hydroxyacid oxidase B. *Journal of Cell Biology* **113,** 113–121 (1991).

40. Vonck, J. & Van Bruggen, E. F. J. Architecture of peroxisomal alcohol oxidase crystals from the methylotrophic yeast Hansenula polymorpha as deduced by electron microscopy. *J Bacteriol* **174,** 5391–5399 (1992).

41. Einerhand, A. W. C., Kos, W. T., Distel, B. & Tabak, H. F. Characterization of a transcriptional control element involved in proliferation of peroxisomes in yeast in response to oleate. *European Journal of Biochemistry* **214,** 323–331 (1993).

42. George, A & Wilson, W. W. Predicting protein crystallization from a dilute solution property. *Acta Crystallogr D Biol Crystallogr* **50,** 361–365 (1994).

43. McNew, J. A. & Goodman, J. M. An oligomeric protein is imported into peroxisomes in vivo. *Journal of Cell Biology* **127,** 1245–1257 (1994).

44. Chalfie, M. GREEN FLUORESCENT PROTEIN. *Photochemistry and Photobiology* **62,** 651–656 (1995).

45. Runnels, L. W. & Scarlata, S. F. Theory and application of fluorescence homotransfer to melittin oligomerization. *Biophysical journal* **69,** 1569–1583 (1995).

46. Walton, P. a., Hill, P. E. & Subramani, S. Import of stably folded proteins into peroxisomes. *Molecular Biology of the Cell* **6,** 675–683 (1995).

47. Wukovitz, S. W. & Yeates, T. O. Why protein crystals favour some space-groups over others. *Nature Structural Biology* **2,** 1062–1067 (1995).

48. Evers, M. E., Titorenko, V., Harder, W., Van Der Klei, I. & Veenhuis, M. Flavin adenine dinucleotide binding is the crucial step in alcohol oxidase assembly in the yeast Hansenula polymorpha. *Yeast* **12,** 917–923 (1996).

49. Fan, G. Y. *et al.* In vivo calcineurin crystals formed using the baculovirus expression system. *Microsc Res Tech* **34,** 77–86 (1996).

50. Langford, J. I. & Louer, D. Powder diffraction. *Rep. Prog. Phys.* **59,** 131–234 (1996).

51. Wille, K. *Introduction to insertion devices* in *CAS - CERN Accelerator School : Synchrotron Radiation and Free Electron Lasers* (1996), 61–75. doi:10.5170/CERN-1998-004.61.

52. Gellissen, G. & Hollenberg, C. P. *Application of yeasts in gene expression studies: A comparison of Saccharomyces cerevisiae, Hansenula polymorpha and Kluyveromyces lactis - A review* in *Gene* **190** (1997), 87–97. doi:10.1016/S0378-1119(97)00020-6.

53. Kozak, M, Kozian, E, Grzonka, Z & Jaskólski, M. Crystallization and preliminary crystallographic studies of a new crystal form of papain from Carica papaya. *Acta Biochim Pol* **44,** 601–605 (1997).

54. Sunde, M *et al.* Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J Mol Biol* **273,** 729–739 (1997).

55. Swaminathan, R, Hoang, C. P. & Verkman, A. S. Photobleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. *Biophys J* **72,** 1900–7 (1997).

56. Ten Wolde, P. R. & Frenkel, D. Enhancement of protein crystal nucleation by critical density fluctuations. *Science* **277,** 1975–1978 (1997).

57. Dodson, G. & Steiner, D. *The role of assembly in insulin's biosynthesis* 1998. doi:10.1016/S0959-440X(98)80037-7.

58. Lipke, P. N. & Ovalle, R. *Cell wall architecture in yeast: New structure and new challenges* 1998.

59. Müntz, K. Deposition of storage proteins. *Plant Molecular Biology* **38,** 77–99 (1998).

60. Subramani, S. Components involved in peroxisome import, biogenesis, proliferation, turnover, and movement. *Physiol Rev* **78,** 171–188 (1998).

61. Walker, G. M. *Yeast Physiology and Biotechnology* 362. doi:978-0-471-96446-9 (Wiley, 1998).

62. Macheroux, P. in *Flavoprotein Protocols. Methods in Molecular Biology* 1–7 (1999). doi:10.1385/1-59259-266-X:1.

63. Mayer, A. F. *et al.* An expression system matures: A highly efficient and cost-effective process for phytase production by recombinant strains of Hansenula polymorpha. *Biotechnology and Bioengineering* **63,** 373–381 (1999).

64. Neal, B. L., Asthagiri, D., Velev, O. D., Lenhoff, A. M. & Kaler, E. W. Why is the osmotic second virial coefficient related to protein crystallization? *Journal of Crystal Growth* **196,** 377–387 (1999).

65. Pepperkok, R., Squire, A., Geley, S. & Bastiaens, P. I. H. Simultaneous detection of multiple green fluorescent proteins in live cells by fluorescence lifetime imaging microscopy. *Current Biology* **9,** 269–274 (1999).

66. Perkins, D. N., Pappin, D. J. C., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20,** 3551–3567 (1999).

67. Von Stockar, U. & Liu, J. S. Does microbial life always feed on negative entropy? Thermodynamic analysis of microbial growth. *Biochim Biophys Acta* **1412,** 191–211 (1999).

68. Galkin, O. & Vekilov, P. G. Control of protein crystal nucleation around the metastable liquid – liquid phase boundary. *PNAS* **97,** 6277–6281 (2000).

69. Haas, C. The Interface between a Protein Crystal and an Aqueous Solution and Its Effects on Nucleation and Crystal Growth. *The Journal of Physical Chemistry B* **104,** 368–377 (2000).

70. Heinze, M., Reichelt, R., Kleff, S. & Eising, R. High resolution scanning electron microscopy of protein inclusions (cores) purified from peroxisomes of sunflower (Helianthus annuus L.) cotyledons. *Crystal Research and Technology* **35,** 877–886 (2000).

71. Lodish, H. *et al. Molecular Cell Biology (4th edition)* 1184. doi:10.1016/S1470-8175(01)00023-6 (W.H. Freeman, New York, 2000).

72. Tenney, K. *et al.* Hex-1, a gene unique to filamentous fungi, encodes the major protein of the Woronin body and functions as a plug for septal pores. *Fungal Genetics and Biology* **31,** 205–217 (2000).

73. Van Dijk, R., Faber, K. N., Kiel, J. A.K. W., Veenhuis, M. & Van Der Klei, I. The methylotrophic yeast Hansenula polymorpha: A versatile cell factory. *Enzyme Microb Technol* **26,** 793–800 (2000).

74. Curtis, R. A., Blanch, H. W. & Prausnitz, J. M. Calculation of phase diagrams for aqueous protein solutions. *Journal of Physical Chemistry B* **105,** 2445–2452 (2001).

75. Inoue, S., Hole, W. & Goda, M. Gree Fluorescent Protein: Enhanced Optical Signals from Native Crystals. *Biol. Bull* **201,** 231–233 (2001).

76. Palmieri, L. *et al.* Identification and functional reconstitution of the yeast peroxisomal adenine nucleotide transporter. *EMBO Journal* **20,** 5049–5059 (2001).

77. Pellegrini, C. *Design considerations for a SASE X-ray FEL* in *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **475** (2001), 1–12. doi:10.1016/S0168-9002(01)01527-3.

78. Valeur, B. *Molecular Fluorescence Principles and Applications* 84–86. doi:10.1002/3527600248 (2001).

79. Wilson, E. *An introduction to Particle Accelerators* 268 (Oxford University Press, 2001).

80. Bodenstaff, E. R., Hoedemaeker, F. J., Kuil, M. E., de Vrind, H. P. M. & Abrahams, J. P. The prospects of protein nanocrystallography. *Acta Crystallogr D Biol Crystallogr* **58,** 1901–1906 (2002).

81. Erdmann, R. & Gould, S. J. *Visualization and purification of yeast peroxisomes* 2002. doi:10.1016/S0076-6879(02)51858-9.

82. Faber, K. N. *et al.* Import of assembled PTS1 proteins into peroxisomes of the yeast Hansenula polymorpha: Yes and No! *Biochim Biophys Acta.* **1591,** 157–162 (2002).

83. Hohmann, S. Osmotic Stress Signaling and Osmoadaptation in Yeasts. *Microbiol Mol Biol Rev* **66,** 300–372 (2002).

84. Inoue, S., Shimomura, O., Goda, M., Shribak, M. & Tran, P. T. Fluorescence polarization of green fluorescence protein. *Proceedings of the National Academy of Sciences* **99,** 4272–4277 (2002).

85. Rakhit, R. *et al.* Oxidation-induced misfolding and aggregation of superoxide dismutase and its implications for amyotrophic lateral sclerosis. *J Biol Chem* **277,** 47551–47556 (2002).

86. Zacharias, D. A., Violin, J. D., Newton, A. C. & Tsien, R. Y. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296,** 913–916 (2002).

87. Berney, C. & Danuser, G. FRET or no FRET: A quantitative comparison. *Biophys J* **84,** 3992–4010 (2003).

88. Koenderink, G. H. *et al.* On the validity of Stokes–Einstein–Debye relations for rotational diffusion in colloidal suspensions. *Faraday Discussions* **123,** 335–354 (2003).

89. Pellegrini, C. & Stöhr, J. X-ray free-electron lasers - Principles, properties and applications. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **500,** 33–40 (2003).

90. Peskir, G. On the diffusion coefficient: The Einstein relation and beyond. *Stochastic Models* **19,** 383–405 (2003).

91. Saldin, E. L., Schneidmiller, E. A. & Yurkov, M. V. Statistical properties of radiation from SASE FEL driven by short electron bunches. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **507,** 101–105 (2003).

92. Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P. & Kobe, B. Crystal structures of fusion proteins with large-affinity tags. *Protein Science* **12,** 1313–1322 (2003).

93. Yuan, P. *et al.* A HEX-1 crystal lattice required for Woronin body function in Neurospora crassa. *Nat Struct Biol* **10,** 264–70 (2003).

94. Doye, J. P. K., Louis, A. a. & Vendruscolo, M. Inhibition of protein crystallization by evolutionary negative design. *Phys Biol* **1,** P9–13 (2004).

95. Harju, S., Fedosyuk, H. & Peterson, K. R. Rapid isolation of yeast genomic DNA: Bust n' Grab. *BMC Biotechnology* **4,** 8 (2004).

96. Komduur, J. a., Bellu, A. R., Knoops, K., van der Klei, I. J. & Veenhuis, M. Cold-inducible selective degradation of peroxisomes in Hansenula polymorpha. *FEMS Yeast Res* **5,** 281–5 (2004).

97. Asadi, M. *Beet-Sugar Handbook* 866. doi:10.1002/9780471790990 (Wiley, 2005).

98. Le Bail, A. Whole powder pattern decomposition methods and applications: A retrospection. *Powder Diffraction* **20,** 316–326 (2005).

99. Nunes, C., Mahendrasingam, A. & Suryanarayanan, R. Quantification of crystallinity in substantially amorphous materials by synchrotron X-ray powder diffractometry. *Pharmaceutical Research* **22,** 1942–1953 (2005).

100. Snapp, E. in *Current Protocols in Cell Biology* (2005). doi:10.1002/0471143030.cb2104s27.

101. Tong, Y., Hughes, D., Placanica, L. & Buck, M. When monomers are preferred: A strategy for the identification and disruption of weakly oligomerized proteins. *Structure* **13,** 7–15 (2005).

102. Brocard, C. & Hartig, A. Peroxisome targeting signal 1: is it really a simple tripeptide? *Biochim Biophys Acta* **1763,** 1565–73 (2006).

103. Cao, H., Chen, B., Squier, T. C. & Mayer, M. U. CrAsH: a biarsenical multi-use affinity probe with low non-specific fluorescence. *Chemical Communications,* 2601 (2006).

104. Distel, B. & Kragt, A. Purification of yeast peroxisomes. *Methods Mol Biol* **313,** 21–6 (2006).

105. Doye, J. P. K. & Poon, W. C. K. Protein crystallization in vivo. *Current Opinion in Colloid and Interface Science* **11,** 40–46 (2006).

106. Gozzo, F. *et al.* The instrumental resolution function of synchrotron radiation powder diffractometers in the presence of focusing optics. *Journal of Applied Crystallography* **39,** 347–357 (2006).

107. Hartner, F. S. & Glieder, A. Regulation of methanol utilisation pathway genes in yeasts. *Microbial cell factories* **5,** 39 (2006).

108. Hartner, F. S. *et al.* Regulation of methanol utilisation pathway genes in yeasts. *Microbial Cell Factories* **5.** doi:10.1186/1475-2859-5-39 (2006).

109. Jackson, S. E., Craggs, T. D. & Huang, J. R. *Understanding the folding of GFP using biophysical techniques* 2006. doi:10.1586/14789450.3.5.545.

110. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy* 954. doi:10.1007/978-0-387-46312-4. arXiv: arXiv:1011.1669v3 (2006).

111. Liu, Y., Kim, H. R. & Heikal, A. A. Structural basis of fluorescence fluctuation dynamics of green fluorescent proteins in acidic environments. *Journal of Physical Chemistry B* **110,** 24138–24146 (2006).

112. Rehani, B. R., Joshi, P. B., Lad, K. N. & Pratap, A. Crystallite size estimation of elemental and composite silver nano-powders using XRD principles. *Indian Journal of Pure and Applied Physics* **44,** 157–161 (2006).

113. Bader, A. N., Hofman, E. G., van Bergen En Henegouwen, P. M. P. & Gerritsen, H. C. Imaging of protein cluster sizes by means of confocal time-gated fluorescence anisotropy microscopy. *Opt Express* **15,** 6934–45 (2007).

114. Banci, L. *et al.* Metal-free superoxide dismutase forms soluble oligomers under physiological conditions: a possible general mechanism for familial ALS. *PNAS* **104,** 11263–11267 (2007).

115. Fujiwara, N. *et al.* Oxidative modification to cysteine sulfonic acid of Cys111 in human copper-zinc superoxide dismutase. *J Biol Chem* **282,** 35933–35944 (2007).

116. Georgieva, D. G., Kuil, M. E., Oosterkamp, T. H., Zandbergen, H. W. & Abrahams, J. P. Heterogeneous nucleation of three-dimensional protein nanocrystals. *Acta Crystallogr D Biol Crystallogr* **63,** 564–70 (2007).

117. Hearps, A. C. *et al.* The Biarsenical Dye Lumio™ Exhibits a Reduced Ability to Specifically Detect Tetracysteine-Containing Proteins Within Live Cells. *Journal of Fluorescence* **17,** 593–597 (2007).

118. Ishiii, M, Kunimura, J., Jeng, H., Penna, T. & O, C. Evaluation of the pH- and thermal stability of the recombinant green fluorescent protein (GFP) in the presence of sodium chloride. *Appl Biochem Biotechnol* **137,** 555–571 (2007).

119. LIU, Y., SHI, M., CAO, R., ZHANG, Y. & HU, Y. Densities and Viscosities of the Quaternary System Mannitol-Sorbitol-D-Glucose-H2O and Its Ternary Subsystems at 298.15K. *Chinese Journal of Chemical Engineering* **15,** 703–709 (2007).

120. Newstead, S., Kim, H., von Heijne, G., Iwata, S. & Drew, D. High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae. *PNAS* **104,** 13936–13941 (2007).

121. Platta, H. W. & Erdmann, R. *The peroxisomal protein import machinery* 2007. doi:10.1016/j.febslet.2007.04.001.

122. Sear, R. P. Nucleation: Theory and applications to protein solutions and colloidal suspensions. *Journal of Physics Condensed Matter* **19,** 1–28 (2007).

123. Stanley, W. a., Fodor, K., Marti-Renom, M. a., Schliebs, W. & Wilmanns, M. Protein translocation into peroxisomes by ring-shaped import receptors. *FEBS Letters* **581,** 4795–802 (2007).

124. Astier, J. P. & Veesler, S. *Using temperature to crystallize proteins: A mini-review* in *Crystal Growth and Design* **8** (2008), 4215–4219. doi:10.1021/cg800665b.

125. Baba, M. Electron microscopy in yeast. *Methods in enzymology* **451,** 133–49 (2008).

126. Banci, L. *et al.* SOD1 and amyotrophic lateral sclerosis: Mutations and oligomerization. *PLoS ONE* **3.** doi:10.1371/journal.pone.0001677 (2008).

127. Chayen, N. E. & Saridakis, E. Protein crystallization : from purified protein to diffraction-quality crystal. *Nat Methods* **5,** 147–153 (2008).

128. Corsini, L., Hothorn, M., Scheffzek, K., Sattler, M. & Stier, G. Thioredoxin as a fusion tag for carrier-driven crystallization. *Protein Science* **17,** 2070–2079 (2008).

129. DePonte, D. P. *et al.* Gas dynamic virtual nozzle for generation of microscopic droplet streams. *Journal of Physics D: Applied Physics* **41.** doi:10.1088/0022-3727/41/19/195505. arXiv: 0803.4181 (2008).

130. Dinnebier, R. E. & Billinge, S. J. L. *Powder diffraction: Theory and practice* 582. doi:10.1039/9781847558237 (2008).

131. Furukawa, Y., Kaneko, K., Yamanaka, K., O'Halloran, T. V. & Nukina, N. Complete loss of post-translational modifications triggers fibrillar aggregation of SOD1 in the familial form of amyotrophic lateral sclerosis. *Journal of Biological Chemistry* **283,** 24167–24176 (2008).

132. Gräslund, S. *et al.* Protein production and purification. *Nature methods* **5,** 135–146 (2008).

133. Hinterberger, F. *Physik der Teilchenbeschleuniger und Ionenoptik* 417. doi:10.1007/978-3-540-75282-0 (Springer-Verlag Berlin Heidelberg, 2008).

134. Marx, H., Mattanovich, D. & Sauer, M. Overexpression of the riboflavin biosynthetic pathway in Pichia pastoris. *Microbial Cell Factories* **7,** 1–11 (2008).

135. Mittemeijer, E. J. & Welzel, U. The "state of the art" of the diffraction analysis of crystallite size and lattice strain. *Zeitschrift fur Kristallographie* **223,** 552–560 (2008).

136. Nagotu, S., Saraya, R., Otzen, M., Veenhuis, M. & van der Klei, I. J. Peroxisome proliferation in Hansenula polymorpha requires Dnm1p which mediates fission but not de novo formation. *Biochim Biophys Acta* **1783,** 760–769 (2008).

137. Raychaudhuri, S. & Prinz, W. a. Nonvesicular phospholipid transfer between peroxisomes and the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 15785–15790 (2008).

138. Tsuruta, H. & Irving, T. C. *Experimental approaches for solution X-ray scattering and fiber diffraction* 2008. doi:10.1016/j.sbi.2008.08.002. arXiv: NIHMS150003.

139. Van Zutphen, T., van der Klei, I. J. & Kiel, J. a.K. W. Pexophagy in Hansenula polymorpha. *Methods Enzymol* **451,** 197–215 (2008).

140. Würtz, C., Schliebs, W., Erdmann, R. & Rottensteiner, H. Dynamin-like protein-dependent formation of Woronin bodies in Saccharomyces cerevisiae upon heterologous expression of a single protein. *FEBS Journal* **275,** 2932–2941 (2008).

141. Zhao, Y. & Zhang, J. Microstrain and grain-size analysis from diffraction peak width and graphical derivation of high-pressure thermomechanics. *Journal of Applied Crystallography* **41,** 1095–1108 (2008).

142. Arroyo-López, F. N., Orlić, S., Querol, A. & Barrio, E. Effects of temperature, pH and sugar concentration on the growth parameters of Saccharomyces cerevisiae, S. kudriavzevii and their interspecific hybrid. *International Journal of Food Microbiology* **131,** 120–127 (2009).

143. Bader, A. N., Hofman, E. G., Voortman, J., Van Bergen En Henegouwen, P. M. P. & Gerritsen, H. C. Homo-FRET imaging enables quantification of protein cluster sizes with subcellular resolution. *Biophys J* **97,** 2613–2622 (2009).

144. Erickson, H. P. *Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy* 2009. doi:10.1007/s12575-009-9008-x. arXiv: arXiv:1011.1669v3.

145. Fischer, H., Polikarpov, I. & Craievich, A. F. Average protein density is a molecular-weight-dependent function. *Protein Science* **13,** 2825–2828 (2009).

146. Ishchuk, O. P., Voronovsky, A. Y., Abbas, C. A. & Sibirny, A. A. Construction of Hansenula polymorpha strains with improved thermotolerance. *Biotechnology and Bioengineering* **104,** 911–919 (2009).

147. Krikken, A. M., Veenhuis, M. & van der Klei, I. J. Hansenula polymorpha pex11 cells are affected in peroxisome retention. *FEBS J* **276,** 1429–39 (2009).

148. Lemaire, K *et al.* Insulin crystallization depends on zinc transporter ZnT8 expression, but is not required for normal glucose homeostasis in mice. *Proc Natl Acad Sci U S A* **106,** 14872–14877 (2009).

149. Satyanarayana, T. & Kunze, G. *Yeast Biotechnology: Diversity and applications* 744. doi:10.1007/978-1-4020-8292-4. arXiv: arXiv:1011.1669v3 (Springer, 2009).

150. Biancalana, M. & Koide, S. *Molecular mechanism of Thioflavin-T binding to amyloid fibrils* 2010. doi:10.1016/j.bbapap.2010.04.001. arXiv: NIHMS150003.

151. Emma, P. *et al.* First lasing and operation of an ångstrom-wavelength free-electron laser. *Nature Photonics* **4,** 641–647 (2010).

152. Endo, K., Mizuguchi, M., Harata, A., Itoh, G. & Tanaka, K. Nocodazole induces mitotic cell death with apoptotic-like features in Saccharomyces cerevisiae. *FEBS Letters* **584,** 2387–2392 (2010).

153. Greenwald, J. & Riek, R. Biology of amyloid: Structure, function, and regulation. *Structure* **18,** 1244–1260 (2010).

154. Groenning, M. *Binding mode of Thioflavin T and other molecular probes in the context of amyloid fibrils-current status* 2010. doi:10.1007/s12154-009-0027-5.

155. Kirian, R. A. *et al.* Femtosecond protein nanocrystallography — data analysis methods. *Opt Express* **18,** 5713–5723 (2010).

156. Kumar, S. & Udgaonkar, J. B. *Mechanisms of amyloid fibril formation by proteins* 2010. doi:10.1016/j.jmb.2008.11.033.

157. Ozalp, V. C., Pedersen, T. R., Nielsen, L. J. & Olsen, L. F. Time-resolved measurements of intracellular ATP in the yeast Saccharomyces cerevisiae using a new type of nanobiosensor. *The Journal of biological chemistry* **285,** 37579–37588 (2010).

158. Robinson, I., Gruebel, G. & Mochrie, S. *X-ray beams with high coherence* 2010. doi:10.1088/1367-2630/12/3/035002.

159. Rupp, B. *Biomolecular Crystallography - Principles, Practice, and Application to Structural Biology* 809 (Taylor & Francis Ltd., 2010).

160. Scheidle, M. *et al.* High-throughput screening of Hansenula polymorpha clones in the batch compared with the controlled-release fed-batch mode on a small scale. *FEMS Yeast Research* **10,** 83–92 (2010).

161. Suzuki, N. *et al.* Crystallization of small proteins assisted by green fluorescent protein. *Acta Crystallographica Section D: Biological Crystallography* **66,** 1059–1066 (2010).

162. Wolf, J., Schliebs, W. & Erdmann, R. *Peroxisomes as dynamic organelles: Peroxisomal matrix protein import* 2010. doi:10.1111/j.1742-4658.2010.07739.x.

163. Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* **470,** 73–78 (2011).

164. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics, Vol. I: The New Millennium Edition: Mainly Mechanics, Radiation, and Heat* 560. doi:10.1119/1.1972241 (2011).

165. Ganguly, S., Clayton, A. H. & Chattopadhyay, A. Organization of higher-order oligomers of the serotonin1A receptor explored utilizing homo-FRET in live cells. *Biophysical Journal* **100,** 361–368 (2011).

166. Hasegawa, H. *et al.* In vivo crystallization of human IgG in the endoplasmic reticulum of engineered chinese hamster ovary (CHO) cells. *Journal of Biological Chemistry* **286,** 19917–19931 (2011).

167. Mcdonald, K. L. & Webb, R. I. Freeze substitution in 3 hours or less. *Journal of Microscopy* **243,** 227–233 (2011).

168. Neiman, A. M. Sporulation in the budding yeast Saccharomyces cerevisiae. *Genetics* **189,** 737–765 (2011).

169. Sahoo, H. *Förster resonance energy transfer - A spectroscopic nanoruler: Principle and applications* 2011. doi:10.1016/j.jphotochemrev.2011.05.001.

170. Seibert, M. M. *et al.* Single mimivirus particles intercepted and imaged with an X-ray laser. *Nature* **470,** 78–82 (2011).

171. Von Dreele, R. B., Suchomel, M. R. & Toby, B. H. *CL Algorithm - Python Script* 2011.

172. Boutet, S. *et al.* High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science* **337,** 362–364 (2012).

173. Hart, P. *et al. The Cornell-SLAC pixel array detector at LCLS* in *IEEE Nuclear Science Symposium Conference Record* (2012), 538–541. doi:10.1109/NSSMIC.2012.6551166.

174. Ihara, K. *et al.* Structural switching of Cu,Zn-superoxide dismutases at loop VI: insights from the crystal structure of 2-mercaptoethanol-modified enzyme. *Bioscience Reports* **32,** 539–548 (2012).

175. Ishikawa, T. *et al.* A compact X-ray free-electron laser emitting in the sub-ångström region. *Nature Photonics* **6,** 540–544 (2012).

176. Lomb, L. *et al.* An anti-settling sample delivery instrument for serial femtosecond crystallography. *Journal of Applied Crystallography* **45,** 674–678 (2012).

177. Saraya, R. *et al.* Novel genetic tools for Hansenula polymorpha. *FEMS Yeast Res* **12,** 271–8 (2012).

178. Schindelin, J. *et al. Fiji: An open-source platform for biological-image analysis* 2012. doi:10.1038/nmeth.2019. arXiv: 1081-8693.

179. Schmit, J. D. & Dill, K. Growth rates of protein crystals. *J Am Chem Soc* **134,** 3934–3937 (2012).

180. White, T. a. *et al.* CrystFEL: A software suite for snapshot serial crystallography. *Journal of Applied Crystallography* **45,** 335–341 (2012).

181. Williams, C., Bener Aksam, E., Gunkel, K., Veenhuis, M. & van der Klei, I. J. The relevance of the non-canonical PTS1 of peroxisomal catalase. *Biochim Biophys Acta* **1823,** 1133–1141 (2012).

182. Zhang, Y., Werling, U. & Edelmann, W. SLiCE: A novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res* **40,** 1–10 (2012).

183. Authier, A. *Early Days of X-ray Crystallography* 464. doi:10.1093/acprof:oso/9780199659845.001.0001 (International Union of Crystallography, Oxford University Press, 2013).

184. Aw, R. & Polizzi, K. M. Can too many copies spoil the broth? *Microb Cell Fact* **12,** 128 (2013).

185. Chan, F. T. S., Pinotsi, D., Kaminski Schierle, G. S. & Kaminski, C. F. in *Bio-nanoimaging: Protein Misfolding and Aggregation* 147–155 (2013). doi:10.1016/B978-0-12-394431-3.00013-4.

186. Iwase, T. *et al.* A simple assay for measuring catalase activity: a visual approach. *Sci Rep* **3,** 3081 (2013).

187. Leal, S. S., Cardoso, I., Valentine, J. S. & Gomes, C. M. Calcium ions promote superoxide dismutase 1 (SOD1) aggregation into non-fibrillar amyloid: A link to toxic effects of calcium overload in amyotrophic lateral sclerosis (ALS)? *J Biol Chem* **288,** 25219–25228 (2013).

188. Liu, W. *et al.* Serial femtosecond crystallography of G protein-coupled receptors. *Science (New York, N.Y.)* **342,** 1521–4 (2013).

189. Milo, R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* **35,** 1050–1055 (2013).

190. Müller, U. *Symmetry Relationships between Crystal Structures* 2013.

191. Nielsen, J., Larsson, C., van Maris, A. & Pronk, J. *Metabolic engineering of yeast for production of fuels and chemicals* 2013. doi:10.1016/j.copbio.2013.03.023.

192. Piatkevich, K. D. *et al.* Extended stokes shift in fluorescent proteins: Chromophore-protein interactions in a near-infrared TagRFP675 variant. *Scientific Reports* **3.** doi:10.1038/srep01847 (2013).

193. Ravin, N. *et al.* Genome sequence and analysis of methylotrophic yeast Hansenula polymorpha DL1. *BMC Genomics* **14.** doi:10.1186/1471-2164-14-837 (2013).

194. Redecke, L. *et al.* Natively inhibited Trypanosoma brucei cathepsin B structure determined by using an X-ray laser. *Science* **339,** 227–230 (2013).

195. Sadowska-Bartosz, I., Pączka, A., Mołoń, M. & Bartosz, G. Dimethyl sulfoxide induces oxidative stress in the yeast Saccharomyces cerevisiae. *FEMS yeast research* **13,** 820–30 (2013).

196. Smith, J. J. & Aitchison, J. D. Peroxisomes take shape. *Nat Rev Mol Cell Biol* **14,** 803–17 (2013).

197. Suppi, S., Michelson, T., Viigand, K. & Alamäe, T. Repression vs. activation of MOX, FMD, MPP1 and MAL1 promoters by sugars in Hansenula polymorpha: The outcome depends on cell's ability to phosphorylate sugar. *FEMS Yeast Research* **13,** 219–232 (2013).

198. Van der Meer, B. W., van der Meer, D. M. & Vogel, S. S. in *FRET - Förster Resonance Energy Transfer: From Theory to Applications* 63–104 (2013). doi:10.1002/9783527656028.ch04.

199. Willenbacher, N & Georgieva, K. in *Prod Des Eng Formul Gels Pastes* 7–49 (Wiley-VCH, 2013). doi:https://doi.org/10.1002/9783527654741.ch1.

200. Adam Kawałek and Ida J. van der Klei. At neutral pH the chronological lifespan of Hansenula polymorpha increases upon enhancing the carbon source concentrations. *Microbial Cell* **1,** 163–209 (2014).

201. Ahmad, M., Hirz, M., Pichler, H. & Schwab, H. *Protein expression in Pichia pastoris: Recent achievements and perspectives for heterologous protein production* 2014. doi:10.1007/s00253-014-5732-5.

202. Axford, D., Ji, X., Stuart, D. I. & Sutton, G. In cellulo structure determination of a novel cypovirus polyhedrin. *Acta Crystallogr D Biol Crystallogr* **70,** 1435–1441 (2014).

203. Barends, T. R. M. *et al.* De novo protein crystal structure determination from X-ray free-electron laser data. *Nature* **505,** 244–7 (2014).

204. Barty, A. *et al.* Cheetah: Software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *Journal of Applied Crystallography* **47,** 1118–1131 (2014).

205. Fischer-Friedrich, E., Hyman, A. A., Jülicher, F., Müller, D. J. & Helenius, J. Quantification of surface tension and internal pressure generated by single mitotic cells. *Scientific Reports* **4,** 4–11 (2014).

206. Freire, S., De Araujo, M. H., Al-Soufi, W. & Novo, M. Photophysical study of Thioflavin T as fluorescence marker of amyloid fibrils. *Dyes and Pigments* **110,** 97–105 (2014).

207. Gati, C. *et al.* Serial crystallography on in vivo grown microcrystals using synchrotron radiation. *IUCrJ* **1,** 87–94 (2014).

208. Hettema, E. H., Erdmann, R., van der Klei, I. & Veenhuis, M. Evolving models for peroxisome biogenesis. *Curr Opin Cell Biol* **29,** 25–30 (2014).

209. Heymann, M. *et al.* Room-temperature serial crystallography using a kinetically optimized microfluidic device for protein crystallization and on-chip X-ray diffraction. *IUCrJ* **1,** 349–360 (2014).

210. Ivanova, M. I. *et al.* Aggregation-triggering segments of SOD1 fibril formation support a common pathway for familial and sporadic ALS. *PNAS* **111,** 197–201 (2014).

211. Kameshima, T. *et al.* Development of an X-ray pixel detector with multiport charge-coupled device for X-ray free-electron laser experiments. *Review of Scientific Instruments* **85.** doi:10.1063/1.4867668 (2014).

212. Kupitz, C. *et al.* Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature* **513,** 261–265 (2014).

213. Mukherji, S. & O'Shea, E. K. Mechanisms of organelle biogenesis govern stochastic fluctuations in organelle abundance. *eLife* **3,** e02678 (2014).

214. Nielsen, K. H. Protein expression-yeast. *Methods in Enzymology* **536,** 133–147 (2014).

215. Perry, S. L. *et al.* In situ serial Laue diffraction on a microfluidic crystallization device. *Journal of Applied Crystallography* **47,** 1975–1982 (2014).

216. Prabhu, Y. T. & Rao, K. V. X-Ray Analysis by Williamson-Hall and Size-Strain Plot Methods of ZnO Nanoparticles with Fuel Variation. *World Journal of Nano Science and Engineering* **4,** 21–28 (2014).

217. Sawaya, M. R. *et al.* Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam. *PNAS* **111,** 12769–74 (2014).

218. Sugahara, M. *et al.* Grease matrix as a versatile carrier of proteins for serial crystallography. *Nat Methods* **12,** 61–63 (2014).

219. Suhling, K., Levitt, J. & Chung, P.-H. Time-resolved fluorescence anisotropy imaging. *Meth Mol Biol* **1076,** 503–519 (2014).

220. Veenhuis, M. & van der Klei, I. J. *A critical reflection on the principles of peroxisome formation in yeast* 2014. doi:10.3389/fphys.2014.00110.

221. Weierstall, U. *et al.* Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nature communications* **5,** 3309 (2014).

222. Weinhandl, K., Winkler, M., Glieder, A. & Camattari, A. Carbon source dependent promoters in yeasts. *Microbial cell factories* **13,** 1–17 (2014).

223. Wilson, W. W. & Delucas, L. J. Applications of the second virial coefficient: Protein crystallization and solubility. *Acta Crystallographica Section F:Structural Biology Communications* **70,** 543–554 (2014).

224. Yoshikawa, H. Y. *et al.* Laser ablation for protein crystal nucleation and seeding. *Chemical Soc Rev* **43,** 2147–58 (2014).

225. Blaj, G. *et al.* X-ray detectors at the Linac Coherent Light Source. *Journal of Synchrotron Radiation* **22,** 577–583 (2015).

226. Chavas, L. M., Gumprecht, L. & Chapman, H. N. Possibilities for serial femtosecond crystallography sample delivery at future light sources. *Structural Dynamics* **2.** doi:10.1063/1.4921220 (2015).

227. Duszenko, M. *et al.* In vivo protein crystallization in combination with highly brilliant radiation sources offers novel opportunities for the structural analysis of post-translationally modified eukaryotic proteins. *Acta Crystallographica Section:F Structural Biology Communications* **71,** 929–937 (2015).

228. Ginn, H. M. *et al.* Structure of CPV17 polyhedrin determined by the improved analysis of serial femtosecond crystallographic data. *Nat Commun* **6,** 6435 (2015).

229. Kim, P. K. & Hettema, E. H. *Multiple pathways for protein transport to peroxisomes* 2015. doi:10.1016/j.jmb.2015.02.005.

230. Knoops, K., De Boer, R., Kram, A. & Van Der Klei, I. J. Yeast pex1 cells contain peroxisomal ghosts that import matrix proteins upon reintroduction of Pex1. *Journal of Cell Biology* **211,** 955–962 (2015).

231. Kobe, B., Ve, T. & Williams, S. J. Fusion-protein-assisted protein crystallization. *Acta Crystallographica Section:F Structural Biology Communications* **71,** 861–869 (2015).

232. Levantino, M., Yorke, B. A., Monteiro, D. C. F., Cammarata, M. & Pearson, A. R. *Using synchrotrons and XFELs for time-resolved X-ray crystallography and solution scattering experiments on biomolecules* 2015. doi:10.1016/j.sbi.2015.07.017.

233. Liang, M. *et al.* The Coherent X-ray Imaging instrument at the Linac Coherent Light Source. *Journal of Synchrotron Radiation* **22,** 514–519 (2015).

234. LLC., S. *The PyMOL Molecular Graphics System, Version 1.8* 2015.

235. McPherson, A. & DeLucas, L. J. Microgravity protein crystallization. *npj Microgravity* **1,** 1–15 (2015).

236. Milo, R. & Phillips, R. *Cell Biology by the Numbers* 356 (Taylor & Francis Ltd, 2015).

237. Neutze, R., Brändén, G. & Schertler, G. F. Membrane protein structural biology using X-ray free electron lasers. *Curr Opin Struct Biol* **33,** 115–125 (2015).

238. Peña, A., Sánchez, N., Álvarez, H., Calahorra, M. & Ramírez, J. Effects of high medium pH on growth, metabolism and transport in Saccharomyces cerevisiae. *FEMS yeast research* **15.** doi:10.1093/femsyr/fou005 (2015).

239. Peng, B., Williams, T. C., Henry, M., Nielsen, L. K. & Vickers, C. E. Controlling heterologous gene expression in yeast cell factories on different carbon substrates and across the diauxic shift: a comparison of yeast promoter activities. *Microbial Cell Factories* **14,** 91 (2015).

240. Roedig, P *et al.* A micro-patterned silicon chip as sample holder for macromolecular crystallography experiments with minimal background scattering. *Sci Rep* **5,** 10451 (2015).

241. Rost, B. R. *et al.* Optogenetic acidification of synaptic vesicles and lysosomes. *Nat Neurosci* **18.** doi:10.1038/nn.4161 (2015).

242. Sauter, A., Roosen-runge, F., Zhang, F. & Lotze, G. On the question of two-step nucleation in protein crystallization †. *Faraday Discussions* **179,** 41–58 (2015).

243. Schlichting, I. Serial femtosecond crystallography: the first five years. *IUCrJ* **2,** 246–255 (2015).

244. Schönherr, R. *et al.* Real-time investigation of dynamic protein crystallization in living cells. *Struct Dyn* **2,** 1–17 (2015).

245. Su, X. D. *et al. Protein crystallography from the perspective of technology developments* 2015. doi:10.1080/0889311X.2014.973868. arXiv: 15334406.

246. Thandavan, T. M. K., Gani, S. M. A., Wong, C. S. & Nor, R. M. Evaluation of Williamson–Hall Strain and Stress Distribution in ZnO Nanowires Prepared Using Aliphatic Alcohol. *Journal of Nondestructive Evaluation* **34,** 1–9 (2015).

247. Tsutsui, H. *et al.* A diffraction-quality protein crystal processed as an autophagic cargo. *Molecular Cell* **58,** 186–193 (2015).

248. Wedekind, J. *et al.* Optimization of crystal nucleation close to a metastable fluid-fluid phase transition. *Scientific Reports* **5,** 1–7 (2015).

249. Aroyo, M. I. *International Tables for Crystallography, Volume A, Space-group symmetry* doi:10.1107/97809553602060000114 (Wiley, 2016).

250. Ayyer, K. *et al.* Macromolecular diffractive imaging using imperfect crystals. *Nature* **530,** 202–206 (2016).

251. Colletier, J.-P. *et al.* Serial Femtosecond Crystallography and Ultrafast Absorption Spectroscopy of the Photoswitchable Fluorescent Protein IrisFP. *J Phys Chem Lett,* 882–887 (2016).

252. Galbán, J., Sanz-Vicente, I., Navarro, J. & De Marcos, S. *The intrinsic fluorescence of FAD and its application in analytical chemistry: A review* 2016. doi:10.1088/2050-6120/4/4/042005.

253. Jakobi, A. J. *et al.* In cellulo serial crystallography of alcohol oxidase crystals inside yeast cells. *IUCrJ* **3,** 1–8 (2016).

254. Martin-Garcia, J. M., Conrad, C. E., Coe, J., Roy-Chowdhury, S. & Fromme, P. Serial femtosecond crystallography: A revolution in structural biology. *Archives of Biochemistry and Biophysics* **602,** 32–47 (2016).

255. Nass, K. *et al.* Protein structure determination by single-wavelength anomalous diffraction phasing of X-ray free-electron laser data. *IUCrJ* **3,** 180–191 (2016).

256. Pellegrini, C. *X-ray free-electron lasers: From dreams to reality* in *Physica Scripta* **2016** (2016). doi:10.1088/1402-4896/aa5281.

257. Pellegrini, C., Marinelli, A. & Reiche, S. The physics of x-ray free-electron lasers. *Reviews of Modern Physics* **88.** doi:10.1103/RevModPhys.88.015006 (2016).

258. Sibirny, A. A. *Yeast peroxisomes: Structure, functions and biotechnological opportunities* 2016. doi:10.1093/femsyr/fow038.

259. Waugh, D. S. *Crystal structures of MBP fusion proteins* 2016. doi:10.1002/pro.2863.

260. Decking, W. & Weise, H. *Commissioning of the European XFEL accelerator* in *IPAC 2017: The eighth International Particle Accelerator Conference* (2017), 1–6. doi:10.18429/JACOW-IPAC2017-MOXAA1.

261. Haynes W. M. *CRC Handbook of Chemistry and Physics, 97th Edition* 2560. doi:10.1136/oem.53.7.504 (CRC Press, 2017).

262. Li, X. *et al.* Diffraction data of core-shell nanoparticles from an X-ray free electron laser. *Scientific Data* **4,** 170048 (2017).

263. Mansour, S. F., El-Dek, S. I. & Ahmed, M. K. Physico-mechanical and morphological features of zirconia substituted hydroxyapatite nano crystals. *Scientific Reports* **7.** doi:10.1038/srep43202 (2017).

264. Meents, A. *et al.* Pink-beam serial crystallography. *Nature Communications* **8.** doi:10.1038/s41467-017-01417-3 (2017).

265. Oberthuer, D. *et al.* Double-flow focused liquid injector for efficient serial femtosecond crystallography. *Scientific Reports* **7.** doi:10.1038/srep44628 (2017).

266. Popp, D. *et al.* Flow-aligned, single-shot fiber diffraction using a femtosecond X-ray free-electron laser. *Cytoskeleton* **74,** 472–481 (2017).

267. Roedig, P. *et al.* High-speed fixed-target serial virus crystallography. *Nat Methods* **14,** 805–810 (2017).

268. Standfuss, J. & Spence, J. Serial crystallography at synchrotrons and X-ray lasers. *IUCrJ* **4,** 100–101 (2017).

269. Weinert, T. *et al.* Serial millisecond crystallography for routine room-temperature structure determination at synchrotrons. *Nature Communications* **8.** doi:10.1038/s41467-017-00630-4 (2017).

270. Wojtas, D. H. *et al.* Analysis of XFEL serial diffraction data from individual crystalline fibrils. *IUCrJ* **4,** 795–811 (2017).

271. Berman, H. M. *et al. Protein Data Bank, Table of Released Entries* doi:doi:10.1093/nar/28.1.235.

List of Related Publications

Arjen J. Jakobi, Daniel M. Passon, Kèvin Knoops, Francesco Stellato, Mengning Liang, Thomas A. White, Thomas Seine, Marc Messerschmidt, Henry N. Chapman, Matthias Wilmanns (2016) ***In cellulo* serial crystallography of alcohol oxidase crystals inside yeast cells**, IUCrJ. 2016 3(Pt 2): 88–95.