1. Project title

Designing starch – harnessing carbohydrate polymer synthesis in plants

2. Project Acronym

DesignStarch

3. Composition of the consortium

CRP Partner	Last, First Name	Affiliation (Organisation, City – Country)	M/F
1	Ebenhöh, Oliver	Heinrich-Heine-University Düsseldorf, Germany	M
2	Zeeman, Samuel	Eidgenössische Technische Hochschule Zürich, Switzerland	M
3	Field, Robert	John Innes Centre Norwich, UK	M

4. Themes

1	Food security	2	Non-food crops		Adaptation to a changing climate
	Biotic/abiotic stress		Others (not listed in the Call Notice):		

5. Keywords

starch; carbohydrates; biopolymers; synthetic biology; systems biology; biochemistry

6. Composition of the research partner teams

CRP Partner	Organisation	Name of team member: Last, First Name, Title	Expertise / Specialisation
1-1	Heinrich-Heine-University	Ebenhöh, Oliver, Jun.-Prof.	Theoretical Biology, Mathematical modelling
1-2	Heinrich-Heine-University	Postdoc, to be named	
1-3	Heinrich-Heine-University	Schliesky, Simon, Dr.	IT Administration, Data Management
2-1	ETH Zurich	Zeeman, Sam, Prof.	Plant Biochemistry, yeast synthetic biology
2-2	ETH Zurich	Postdoc, to be named	
3-1	John Innes Centre, Norwich	Field, Robert A, Prof.	Glycobiology, Surface biochemistry

3-2	John Innes Centre, Norwich	Postdoc, to be named	

7. Project summary

Starch is a natural product produced by most land plants and algae with remarkable physico-chemical properties. Starch is composed of two polymers of glucose: amylose, a predominantly linear polymer of α-1,4 linked glucose units, and amylopectin, which also contains α-1,6 linkages (branch points) resulting in a tree-like structure. The simple constituents of starch (one type of monomer and two types of linkages) is contrasted by its complex and highly ordered structure, in which crystalline and amorphous layers alternate in a defined and regular fashion. This structure gives starch unique physicochemical properties, which make it an exceptionally tightly packed energy storage that is of such tremendous importance for the human diet and economy as a whole. Despite decades of intense research, it is still not understood how precisely starch granule biogenesis initiates and progresses. A relatively small number of enzymes are involved, but it is unclear how their activities are coordinated in order to ultimately control the structure and properties of starch.

The objective of our project is to gain a profound understanding of the regulation and control of the biophysical and biochemical processes involved in the formation of the complex polymeric structure that is the starch granule. We will apply this understanding to recreate the synthesis of starch *in vitro* and learn to control its physical and chemical properties in a targeted way. By expressing starch synthesising enzymes in yeast, an organism not natively producing starch, we will design starches with desired properties *in vivo*. This will be translated back *in planta* to genetically engineer plants producing starch with desired, pre-defined physico-chemical properties.

To achieve our goal, we will simultaneously follow bottom-up and top-down approaches, complemented by synthetic and theoretical biology activities. We will systematically analyse the components of the starch synthesis systems by characterising all involved enzymes *in vitro* in order to understand their functional specialisations (WP1). These data will inform the mathematical models (WP4) used to predict the combined actions of enzymes *in vitro* (WP1) and the behaviour of engineered pathways *in vivo* (WP2&3). The characterised components (WP1) will be used to engineer starch synthesis in *Saccharomyces cerevisiae* (WP2). Using controllable promoters, we will systematically regulate the levels of individual enzymes, allowing the production of insoluble polymers to be fine-tuned, thereby testing test model predictions (WP4) and forming the basis for the transformation of plants (WP3).

In summary, we will generate a systems-wide understanding of the synthesis of the complex starch polymers and implement design strategies to create starches with desired properties *in vitro* and *in vivo*.

8. Project description

8.1 Project Duration (months): 36 months

8.2 Intended starting date: May 2015

8.3 Objectives of the project (max. 1 A4 page)

The overall aim of the project is to obtain a systems-wide understanding of the biochemical and biophysical processes involved in the synthesis of the complex carbohydrate starch and to use this knowledge to rationally engineer systems to synthesise starch with desired properties. Obtaining such comprehensive insight would not be possible without combining the complementary expertise of our research groups. A central aspect of our research project is the development of validated and predictive mathematical models, which form a comprehensive theoretical framework in which to interpret and integrate all gathered knowledge and data. To achieve our aim, we will follow a unique approach combining biochemistry and molecular biology with modern theoretical and synthetic biology approaches.

Our systems-wide approach necessitates investigation of the starch synthesizing mechanism at multiple levels. Biochemical methods (WP1) will deliver a comprehensive *in vitro* characterization of all enzymes involved in starch synthesis. In addition, enzymes with altered biochemical properties will be produced by site-directed mutations. For substrate specific enzymes in solution, the employed biochemical methods are standard. However, for surface-active and polymer-active enzymes a comprehensive characterization is only possible through the combination of the skills of the Field lab (carbohydrate-coated chip technology), the Zeeman lab (production of enzymes with altered kinetics) and the Ebenhöh lab (theoretical concept development). The resultant dataset will be unique. It will provide the first complete description of enzymes of a pathway involving soluble and insoluble substrates, polymers and monomers.

This knowledge base, which will be made freely accessible (see Data Management Plan), will form the background for predictive mathematical models (WP4) for *in vitro* (WP1), synthetic (WP2), and *in planta* (WP3) systems. Specifically, an extensive library of yeast strains will be produced expressing wild-type and modified starch-synthesizing enzymes in combinations predicted by the models to yield interesting insoluble structures (WP2). The control of gene expression by tunable promoters will confer control over enzyme stoichiometry. These yeast strains will be precisely described by characterizing the insoluble carbohydrate polymers in terms of chain length distributions, branching pattern, crystallinity and particle size distribution. Building on the promising preliminary results of the Zeeman group, we will, for the first time, systematically characterize synthetic starch-like polymer producing systems and interpret and understand their functioning through the application of mathematical models (WP4).

Recreating starch synthesis from scratch in a non-starch producer and modifying a native starch-producing plant pose quite different challenges. Insight form the *in vitro* and yeast studies will guide our plant transformation strategies, allowing targeted modification of starch biophysical properties (WP3). This will lead to a collection of transformed plants producing modified starches that we will thoroughly characterise in terms of granule size/morphology, chain length distributions, branching pattern, crystallinity and other biophysical parameters to be defined and interpret our observations in the context of the developed mathematical models (WP4).

Throughout the project, the mathematical models for the various levels of organization (*in vitro*, synthetic, in planta) will be continuously used to interpret experimental results as they are obtained, to make new predictions and experimental validation or falsification will directly feed back into model development such that their predictive power is permanently improved and the closest possible link between experimental and theoretical activities are ensured.

8.4 Background (max. 2 A4 pages)

Through extensive biochemical and genetic research, the conserved enzymatic machinery for the biosynthesis of starch in vascular plants has been discovered. The process involves a highly regulated enzyme for ADP-glucose production (ADP-glucose pyrophosphorylase) and three enzyme classes responsible for glucan polymer (amylopectin and amylose) formation and modification – five distinct starch synthases, two distinct branching enzymes and an ISA-type multimeric debranching enzyme. The starch synthase isoforms all have distinct, yet overlapping functions in elongating the chains of amylopectin. The branching enzymes introduce the alpha-1,6-branch points, some of which are selectively removed by debranching reactions. Collectively, these enzymes generate a polymer with a compatible chain-length and branch-point distributions that permit the formation of the crystalline structures that underlie the insoluble nature of starch granules. However, knowing the components and understanding the biosynthetic process represent quite distinct levels of understanding. Recreating starch biosynthesis *in vitro* (WP1) or in a heterologous systems (WP2) and developing theoretical models explaining the emergent properties (WP4) are essential steps towards a full understanding.

An inherent complication of starch enzymology is often ignored – namely that starch is not water soluble, while the enzymes that act upon it are. The biochemistry of starch, and indeed the *in vitro* enzymatic modification of starch, therefore takes place across a solid–liquid interface. This has a profound impact on enzyme action and must be recreated *in vitro* in order to generate biologically meaningful insight. Reflecting this point, the activity of granule-bound starch synthase, for instance, has been shown to be enhanced on crystalline amylopectin compared to soluble substrate[1]. Various technologies have been successfully applied to observe the changes to the surface of particles, including electron microscopy to visualise the products of amylosucrase-catalysed extension of glycogen giving rise to semi-crystalline starch-like material[2], to observe alterations of the starch granule surface or structure during the initial steps of starch breakdown[3], or to directly monitor purified starch granules during alpha-amylase degradation, either using electron microscopy[4] or synchrotron UV fluorescence microscopy[5]. However, such innovative experiments provide information on structures, but not on reaction rates as such, and are thus somewhat remote from the biological system. To understand starch surface biochemistry, we will go beyond these approaches and apply novel technologies to observe the breakdown of the insoluble substrate *per se* in real time.

Previously, quartz crystal microbalance technology has been used to monitor phosphorylase-catalysed amylopectin extension in real time[6] and to provide valuable kinetic information[7]. We previously showed that surface plasmon resonance (SPR) can be used to measure the synthesis of soluble polysaccharides on surfaces in real time[8,9]. In parallel studies, we demonstrated that gold nanoparticles provide an excellent system for displaying carbohydrates on a surface such that they are accessible to carbohydrate-binding proteins[10,11]. Bringing these approaches together, we reported the use of *Arabidopsis thaliana* glucan phosphorylase PHS2 for the enzymatic extension of glucan-coated SPR biosensor surfaces. Whilst PHS2 is not a starch-active enzyme in physiological terms, it is a robust and efficient tool for the generation of alpha-1,4-glucan chains running to >80 residues in a matter of minutes. Together, these approaches produce insoluble glucan architectures that respond to glucan modifying enzymes in a manner reminiscent of the same enzymes acting on starch itself. Further, glyconanoparticles produced by the PHS2-mediated extension of glucan-coated gold nanoparticles produced material that stained strongly with iodine, again reflecting the *de novo* production of starch-like material[12]. These technologies are now ready to be deployed for the kinetic analysis of plant-derived starch-active enzymes such as starch synthases, branching and debranching enzymes.

The majority of non-starch producing organisms (including most prokaryotes, fungi and animals) produce glycogen – a glucan related to amylopectin, but with shorter chains and more branch points. It remains soluble rather than forming the semi-crystalline matrix adopted by amylopectin in starch. *S.*

cerevisiae contains two UDP-glucose-dependent glycogen synthases (with distinct temporal expression patterns) and a branching enzyme. Glycogen molecules are initiated by a self-glucosylating protein, glycogenin, for which there are two homologous genes. No debranching step is required for glycogen production, but a debranching enzyme acts during glycogen mobilisation, together with glycogen phosphorylase. Thus, seven genes are responsible for glycogen production and remobilisation. We are pursuing a synthetic biology approach, progressively engineering *S. cerevisiae* via the introduction of Arabidopsis genes with the aim of enabling it to produce starch. We chose *S. cerevisiae* as it is possible to stably express multiple genes from well-defined chromosomal positions, inserting them step-wise via homologous recombination. Inserted genes are interspersed with essential genes, preventing spontaneous loss, which has formerly been a common problem when engineering yeast. The value of this approach has been demonstrated by engineering the glucosinolate biosynthetic pathway into yeast using Arabidopsis genes[13].

At the time of writing, we have eliminated the two endogenous glycogen synthases, the branching and debranching enzymes and the phosphorylase, leaving a strain incapable of glycogen metabolism. In this quintuple mutant background, we have created 30 strains in which we express a gene encoding an unregulated form of bacterial ADPglucose pyrophosphorylase (for ADPglucose production), all possible combinations of the four soluble starch synthases (SSI-IV) together with a branching enzyme (BE3), and with or without the heteromultimeric debranching enzyme (ISA1/ISA2). We also have numerous strains with mixed combinations of endogenous yeast and introduced plant genes. The introduced genes are under the control of a galactose-inducible promoter such that when grown on glucose, all are repressed, facilitating their handling and sequential manipulation. Through this process, we have created a yeast strain library that we can analyse by iodine staining (see Figure 1), by light and electron microscopy, and by extracting soluble or insoluble glucans so as to analyse their structural parameters using biochemical (e.g. chain length analyses) and biophysical methods (e.g. birefringence of polarised light and small-angle X-ray scattering). This analytical toolbox allows us to assess the crystallinity and molecular arrangement of the constituent polymers and we are already able to produce insoluble glucans in yeast that have some structural characteristics of starch.

Clearly one central aim is to see which combination (or combinations) of enzymes permit the biosynthesis of true starch-like polymers. However, this yeast collection is much more valuable than that as it represents the starting point for a deeper and more refined analysis that will take place in the context of the ERA-CAPs network.

A key integrating factor of our ERA-CAPS network is the development of theoretical descriptions and mathematical models, which provide a framework in which to interpret the observed results. So far, the mathematical analysis of complex carbohydrates was restricted to a single enzyme or a very small number of co-operating enzymes [14–16]. However, recent theoretical developments of the Ebenhöh group have provided a novel understanding of how to interpret and understand polymer-active enzymes in the context of statistical thermodynamics [17] and how to systematically describe surface-active enzymes [18]. These key advances pave the way for the development of comprehensive mathematical models describing the synthesis of complex polymeric structures such as starch by a combination of biochemical and biophysical processes. Thus, a central and unique aim of our project is the development of predictive mathematical models, which will promote our basic understanding of the complex interacting processes involved in starch metabolism as well as provide tools to engineer systems producing polymers with desired properties in a targeted way.

8.5 Research plan (max. 5 A4 pages)

The concerted action of many enzymes is required for the biogenesis of the starch granule synthesis; several of these enzymes exist in numerous isoforms. However, apparent differences in enzyme action have been often interpreted in terms of *in vivo* phenotypes when gene expression is repressed (i.e. indirectly, through examining what the remaining enzymes are capable of), rather than through direct consideration of enzyme activity and substrate specificity. The biosensor and nanoparticles technologies that we have developed (Field) provide an opportunity to obtain direct kinetic information about enzymes acting on insoluble starch. This paves the way for quantitative assessment (Field) of starch-active enzymes (from Zeeman), operating either individually or in combination, that will feed into modelling of the action of the starch enzyme system (Ebenhöh). This in turn will complement and underpin the synthetic biology approaches towards *de novo* starch biosynthesis and contribute to understanding starch granule biogenesis *in planta* (Zeeman).

Our pioneering combination of bottom-up and top-down approaches and complementary synthetic and theoretical biology is organised into four work packages (WPs), representing the three levels of complexity (WP1 – *in vitro*, WP2 – synthetic / in yeast, WP3 – *in planta*) and one theoretical work package (WP4) providing the embracing theory.

WP1 Biochemical characterisation of enzymes involved in starch synthesis

Task 1.1. Generation of recombinant forms of all enzymes involved in granule biosynthesis

We will produce, in *E. coli* or *N. benthamiana*, recombinant forms of all 8 of the Arabidopsis enzymes that are throught to be required for starch granule formation (5 starch synthases, 2 branching enzymes and a debranching enzyme; Zeeman and Field). All of the cDNAs for the Arabidopsis proteins, trimmed of their chloroplast transit peptides are already available from our yeast programme (Zeeman). A suite of site-directed mutants will be generated for each in order to address fundamental questions about active site substrate binding and catalysis, and identify remote surface-site substrate binding or regulatory features (e.g. phosphomutant or phosphomimetic variants of known phospho-proteins).

Task 1.2. Generation of defined glucan surfaces

To facilitate kinetic analysis on insoluble glucan substrates, gold surfaces (biosensor chips; nanoparticles) will be produced with appended linear or branched glucans. This will be achieved by (a) coupling fragmented and size-fractionated amylopectin to self-assembled monolayers (SAMs), as we reported previously; or (b) the chemoenzymatic synthesis in solution of defined linear and branched oligosaccharide, followed by their coupling to SAMs. The former glucans are more complex and heterogeneous (in terms of chain length and branching), but natural; the latter are simpler, shorter but homogeneous: both have their merits, depending on the application in question. With these 'primer' glucan surfaces in hand, enzymatic extension may be achieved with PHS2 +/- branching enzyme to access defined amylose-like or amylopectin-like surfaces, respectively. As definition of function of the Arabidopsis starch-active enzymes becomes clear (*vide infra* + WP2/3), these enzymes will be used in place of PHS2 (which is not physiologically involved in starch granule biochemistry) to provide a greater degree of precision in generating mimetics of natural starch surfaces.

Task 1.3. Surface characterisation

Capitalising on protocols that are already in place in the Field lab[12], we will assess all *in vitro* generated starch-like surfaces (a) for their ability to stain with iodine; (b) for their 3D architecture, using either AFM (sensor chips) or TEM (nanoparticles); and (c) for their sensitivity to degradation by starch-active enzymes (glucan chain end-acting alpha-glucosidase and beta-amylase; side-acting

alpha-amylase; debranching isoamylase). Similar approaches will be adopted to characterise surfaces that have been subject to the action of starch synthesis enzymes (*vide infra*).

Task 1.4. Measuring the kinetics of interaction and modification of surfaces

Real-time measurement of enzymes acting on starch mimetic surfaces will be achieved using biosensor technology that relies on surface plasmon resonance (Biacore T-200) or bio-layer interferometry (Forte Bio Octet). We have established that SPR is a powerful tool for the study of surface enzymology[9] and that the immobilisation methodologies translate conveniently to nanoparticle systems[12]. This is important due to the limitation imposed by the 3 channel + control format of the T-200. Translation to nanoparticles, which can be employed in microtitre plates, enables much parallel work in terms of glucan product analysis, but lacks the real-time read-out required for kinetics. Preliminary studies encourage us that glucan immobilisation and enzymatic extension are eminently achievable with the Octet system, which offers excellent scope for multiplexed assays. The Octet's eight independent channels substantially streamlines analyses, which will be invaluable for detailed kinetic analyses. With this suite of technologies in hand, we will perform detailed analysis of the action of recombinant starch active enzymes, chimeras and mutants thereof, along with assessment of the impact of post-translational modification on enzyme activity and selectivity (Starch synthases and branching enzymes are both reported to be phosphorylated in cereals[19,20] and phosphopeptides for SS2 are present in the arabidopsis phosphopeptide databases. Several activities are sensitive to redox potential[21]. Interpretation of the kinetic data will be facilitated by the theoretical concepts for polymer-active enzymes developed in the Ebenhöh lab. The acquired kinetic information will be fed into mathematical models (WP4), which will be compared with multiple enzyme systems constituted *in vitro*.

WP2 Engineering yeast strains to synthetically produce insoluble carbohydrate polymers

Task 2.1. Using established yeast strains to inform a model on starch biosynthesis

Each of the conserved enzymes in starch synthesis is thought to have a distinct function. This is exemplified by the starch synthase gene family, of which there are 5 members (SSI-SSIV and GBSS). SSI - SSIV are involved in amylopectin synthesis. Studies performed mostly on mutant and transgenic plants have led to the idea that SSI, SSII and SSIII elongate short, medium and long chains of amylopectin respectively, while SSIV is involved in the poorly-characterised process of starch granule initiation. GBSS (for Granule Bound Starch Synthase) is unique in that it is trafficked to the starch granule (by a novel protein, PTST, for Protein Transport to STarch, recently characterised in the Zeeman laboratory) and exclusively found there, where it synthesises the amylose fraction of starch.

The role of each of the soluble starch synthases (SSI – SSIV) towards amylopectin synthesis will be analysed in yeast lines expressing a single starch synthase, together with a single branching enzyme (BE3). Glucans will be extracted and analysed for their total molecular weight, chain-length distribution and branch point distribution (determined by chain length analysis after pre-treatment with exoamylase to specifically remove the external chains). These data will form the baseline from which to analyse further lines expressing two, three or all four starch synthases. The data will also be compared with those derived from the actions of the recombinant purified Arabidopsis enzymes tested *in vitro* (WP1). Together, these empirical data will inform the initial mathematical models (WP4) for amylopectin biosynthesis. Further, we will analyse the extracted glucans from the comparable yeast strains also expressing the ISA-type debranching enzyme, which is thought to tailor nascent amylopectin molecules to promote their crystallisation. We will define which starch synthase combinations create glucan structures that can be selectively debranched by isoamylases such that they

will crystallise. The mathematical models developed in parallel will ultimately allow us to define the structural preconditions for this debranching step.

Task 2.2. Creating new yeast strains to understanding the determinants of starch enzyme specificities

All starch synthase proteins contain a conserved glucosyl transferase domain. In addition, each has unique domains that are likely to be critical in determining their specific sub-functions or for their regulation. For example, SSII has an amino-terminal domain which is subject to phosphorylation; SSIII has a set of three carbohydrate binding modules in addition to the substrate binding residues of the active site; SSII, SSIV and GBSS have predicted coiled coil domains, while SSIII and SSIV have predicted 14-3-3 binding sites, all of which could mediate protein-protein interactions. Our yeast-based method is perfect to explore these potentially important protein features. Likewise, branching and debranching enzymes have domains outside the active site that are required for substrate interactions and likely modify the substrate specificities and kinetic characteristics of the enzymes. Beyond this, there is evidence from cereal species that the starch biosynthetic protein complexes exist including both synthase and branching enzyme isoforms, together with other proteins.

New yeast lines will be created in which a given Arabidopsis enzyme is replaced by versions altered through site-directed mutagenesis, through truncation, or through domain-swap experiments. For example, the SSIII will be replaced by truncated or site-directed versions altering the amino-terminal carbohydrate binding modules, or altering the substrate binding residues of the active site. Alternatively the glucosyl transferase domain will be replaced by that of another enzyme (e.g. SSI or a glycogen synthase). Initially these replacements will be done in a yeast line expressing only SSIII and BE3. Subsequently, when comparisons of the altered versions with the wild type enzyme provide interesting and informative changes in function, the altered versions will be expressed in combination with other starch biosynthetic enzymes in yeast and in appropriate plant backgrounds (see below).

Task 2.3. Quantitative protein expression in yeast.

We will investigate the extent to which relative enzyme amounts affect the structure of the resultant glucans. To control the expression levels of individual genes we will specifically replace galactose-inducible promoters of a single gene with a promoter that can be controlled to modulate gene expression levels (e.g. Cu^{2+}- or Ca^{2+}-inducible systems). Thus, by placing SSIII under the control of a Cu^{2+}-inducible promoter, its expression can be modulated in relation to BE3 by controlling the level of Cu^{2+} in the medium. Alternatively, BE3 expression could be modulated in relation to ISA-type debranching enzyme. Clearly a great many combinations are possible. The choice of which proteins to vary in abundance will derive partly from the initial results in WP1 and other tasks in WP2 (see above) as well as an idea of the absolute protein levels for each protein *in planta* and when expressed under the gal-promoter in yeast. These data will be valuable as they will reveal whether amylopectin structure generation during starch biosynthesis is a predominantly a competition for substrates amongst the enzymes present or a more orchestrated processes where a combination of compatible enzyme specificities determine structure.

WP3 Generation of modified Arabidopsis plants producing starch with altered characteristics

Task 3.1. Expression of variant enzymes in Arabidopsis mutants.

The development of practical, realistic *in-vitro* and heterologous expression systems provide superb tools for rapid hypothesis and mathematical model development and iterative testing. Ultimately, however, it is necessary to return to a plant system, where starch metabolism proceeds in its true biological context. We will use Arabidopsis for three key reasons; firstly, it remains the best model system for higher plants, including all important starch crops. Secondly, we are experts in Arabidopsis starch biosynthesis. Thirdly, in the course of our research we have established or obtained an extensive

collection of over 50 Arabidopsis knock-out mutant combinations lacking one or more starch metabolic genes (i.e. starch synthases, branching enzymes, debranching enzymes and combinations thereof[22,23]). This mutant collection represents an extremely valuable tool in which to assess the impact of alterations in kinetic characteristics or relative gene expression levels that are revealed by the other layers of this interdisciplinary research program. Using these mutant lines as backgrounds for genetic transformation, constructs will be made using gateway multisite cloning to express variant forms of starch biosynthetic enzymes, under the control of their own promoter or a foreign promoter. The impact of these changes on starch granule synthesis in these plant systems can be rapidly evaluated using molecular, biochemical and microscopic methods in which we have proven expertise. This work will build on our existing knowledgebase and will provide a vitally important comparison with the *in-vitro* and yeast-based systems. We have to recognise that the plant system may have additional complexities, the importance of which have not yet been fully realised (i.e. the contribution of starch catabolising enzymes in determining structure, or the presence of other, as-yet unidentified proteins in plants which are important for starch granule biosynthesis).

WP4 Development of Mathematical models of starch synthesis as theoretical frameworks and predictive tools

As has become clear in the recent decades, starch metabolism is not simply understandable by a straight-forward interpretation of single knock-out mutants, but rather represents a complex system with properties that go beyond direct intuition. We will therefore investigate the systemic properties with mathematical models.

Task 4.1. Theoretical description of single enzymes kinetics

Any mathematical model describing the dynamics of a metabolic system is based on expressions describing the kinetics of the single components, i.e. enzymes, involved in the pathway. In this project, we face the challenge that we will describe a system in which biochemical (enzymatic) and biophysical (crystallisation) processes cooperate. The major processes (e.g. chain elongation together with branching and debranching reactions) occur in solution and/or on the granule surface and biosynthesis of nascent material proceeds with crystallisation of mature polymers. Most of the involved enzymes act on polymers and are not specific to one or a small number of substrates.

To serve as building blocks for more complex models (Tasks 4.2 and 4.3) we will first develop models describing the observed behaviour of single enzymes expressed under controlled conditions *in vitro*, using in particular data resulting from WP1 and WP2 (Task 2.2). Some of these carbohydrate-active enzymes will act in solution, while others are surface-active. Model development will draw greatly on the experience of the Ebenhöh group who recently developed novel theoretical concepts to describe polymer-active enzymes and surface-active enzymes [17,18]. With methods from statistical thermodynamics and stochastic simulations, we could successfully describe the *in vitro* behaviour of several glucanotransferases [17,24]. For each of the enzymes involved in starch synthesis, corresponding descriptions will be formulated and parameterised by the *in vitro* data from WP1.

Task 4.2. Theoretical advancement to include biophysical processes and simulation of simple reconstituted pathways

The models and concepts developed so far can describe enzymatic reactions on polymers and on granule surfaces. However, so far no conceptual framework exists, by which biophysical processes such as crystallisation can be included. An important theoretical goal of this project is therefore to develop models that can provide a comprehensive description of the formation of a semi-crystalline carbohydrate polymer. A major challenge in the development of theories and models is the vast complexity of starch. A granule cannot be described in every molecular detail. Rather, important

characteristic quantities need to be identified which a) provide a sufficiently detailed description of a starch granule, b) can be measured either directly or indirectly through biochemical, biophysical or imaging technologies and c) are valuable parameters to modify for technological applications. Such quantities include but are not limited to the number of crystalline and amorphous layers, thickness of layers, branching patterns, crystalline arrangement, amylose content and chain lengths. Once this conceptual framework is established, simple reconstituted systems consisting of a relatively small number of enzymes will be modelled, where models are tested against and parameters are derived from the results from simple *in vitro* (WP1) and yeast (WP2) systems.

Task 4.3: Modelling complex starch synthesis pathways and in planta systems

The ultimate goal is to translate the knowledge gained from the *in vitro* and yeast systems back to the plant system. However, as stated above (WP3), the various native processes acting on starch, such as enzymes involved in starch degradation, are likely to interfere with the engineered starch synthesis pathways. The results of Tasks 4.1 and 4.2 enable us to construct mathematical models of higher complexity involving the description of the action of a complete set of starch synthesising enzymes, to include biophysical processes, and also to include key enzymes involved in degradation for the *in planta* systems. We will begin by mimicking the inducible yeast systems and reproduce the observed results (WP2, Task 2.3) to understand how quantitative changes in enzyme levels alter the physico-chemical properties of the produced starch granules. Once a realistic model is established and has been validated against the synthetic yeast system, it will serve as a starting point for the development of more complex models mimicking starch synthesis in the Arabidopsis background.

8.6 Complementarity of the teams and transnational added value (max. 1 A4 page)

An interdisciplinary and multi-faceted project such as this requires the expertise of specialists in various fields that span the physical and life sciences, integrating theoretical and experimental approaches. No single research group in Europe, or anywhere else in the world, could tackle on their own the scientific questions we are addressing here. Our consortium has been carefully selected to ensure a maximal synergy between the partners by identifying the complementary skills and experience needed to successfully complete the project: we need to be able to express, purify and isolate enzymes, analyse their kinetic and thermodynamic properties, both in solution and on surfaces, engineer cassettes of enzymes into *S. cerevisiae*, transform *A. thaliana* plants, characterise structural and physico-chemical properties of soluble and insoluble carbohydrates, develop computational models and unifying theories, place our finding in a sound conceptual framework that allows us to make novel predictions and in future to engineer designer starches.

Sam Zeeman is a renowned expert in starch metabolism, and has established an extensive and invaluable library of Arabidopsis mutants. His research uses both targeted reverse-genetic and molecular biological approaches to investigate gene function together with unbiased classical genetics for novel gene discovery. His work has resulted in conceptual advances and patents applications for novel methods to modify starch structure. His research team has recently demonstrated the feasibility of engineering starch synthesising pathways in yeast, opening up a whole new set of possible approaches such as those described here. Rob Field has made innovative advances in analysing kinetic properties of surface-active enzymes and their substrate specificities employing his recently developed carbohydrate-coated sensor chips and gold nanoparticles. This builds on a more extensive physical sciences programme aimed at developing new tools for glycoscience research, part of which has recently led to the spin-out of Iceni Diagnostics to commercialise carbohydrate-based surfaces for

medical and agricultural applications. Field has extensive experience of working in multidisciplinary teams, with academia and industry, including the UK Glycoarrays consortium, The European Glycoscience Forum steering group and the US Consortium for Functional Glycomics. Oliver Ebenhöh has a long-standing expertise in theoretical biology and mathematical modelling. He has profound experience in developing mathematical models of plant energy metabolism, which he is now applying in a broad context as Junior Professor for Quantitative and Theoretical Biology within the Cluster of Excellence in Plant Science (CEPLAS), funded by the German Research Foundation. He has been instrumental in developing novel theoretical concepts to support and advance our understanding of polymer biochemistry and the action of surface-active enzymes, a unique expertise without which the theoretical activities of this project would not be possible. As coordinator of the Marie-Curie Initial Training Network 'AccliPhot' (Acclimation of Photosynthesis), in which 9 academic and 3 industrial partners collaborate to optimise industrial algal biotechnology, he has proven experience and skills in program management and leading interdisciplinary research consortia.

8.7 Plan for use and dissemination of knowledge (max. 0.5 A4 page)

The consortium follows the general principles of open science and open source. This means that, as a general rule, all results will be made accessible to all consortium members immediately after generation and made freely available for reuse under the creative commons license through dissemination over the consortium website and open access articles. For details on the data storage, organization and dissemination see the separate Data Management Plan.

At the same time the consortium respects the interests of partners generating results with potential commercial interest. Such results can be expected for example from the generation of yeast strains or modified plants producing starch with physico-chemical properties of industrial importance, novel enzymes for starch modification and sensor chip design and applications. It is the responsibility of the partner producing the result to identify and communicate the interest and intention to reserve the intellectual property rights and delay public dissemination of the result. In this case, the corresponding protocols to generate such strains will be made available to the consortium members only and a patent will be filed by the consortium member who has generated the result as soon as possible.

Details on ownership and joined ownership follow the general principles laid out in Annex II to the Call Notice (IPR conditions) and will be detailed in a Consortium Agreement, which will be agreed on by the partners latest 3 months after project start.

8.8 Coordination with outside groups (max. 0.5 A4 page)

Technologies in synthetic biology are developing very rapidly, and we will adopt emerging methodologies in our project as and when it will speed up our work without compromising the continuity of our analyses. For example, through our interactions with Dr. Chris French and colleagues at the university of Edinburgh, we will follow the YeastFab programme, which is involved in making many standard 'parts' and developing techniques for rapid pathway assembly in yeast. In particular, their plan to convert all yeast promoters into standard modular parts which can be reassembled with genes of interest using the rapid and efficient 'yeast golden gate' assembly system. Such technologies, combined with gene synthesis technologies will make it very easy to assemble multi-gene pathways quickly and integrate them into yeast. We will also benefit from collaboration with Prof Gebhard Schertler (Paul Scherrer Institute, ETH Domain) for X-ray scattering analysis of starch polymers using the Swiss Light Source (SLS) synchrotron.

8.9 Economic, societal and/or environmental relevance (if appropriate) (max. 0.5 A4 page)

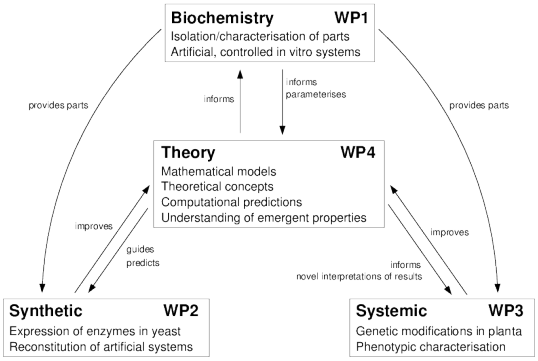
Understanding starch biosynthesis is a major challenge, but a very important one. Academically, starch biosynthesis represents a prototype complex system. We believe that we understand most components, yet we cannot explain or predict the system's behavior. Thus, understanding this complex pathway, in which biochemical and biophysical processes in solution and on reaction surfaces interact, will be a milestone in understanding emerging properties of even more complex biological systems, such as cell wall or organelle formation. The industrial and economic importance of starch cannot be over-estimated, given its position as the main calorific input for mankind and a bulk commodity for the chemical industry. Our project will set the basis from which to harness starch biosynthesis *in vivo* and control starch properties in a systematic fashion. With the knowledge obtained through our collaborative effort it will become possible to move the generation of diverse starch functionality from the chemical plant into the crop plant or algal workhorse. Thus, we see our research as providing the foundation for a new branch of the biotechnology industry, which is based on synthetic biology underpinned by sound theoretical concepts and experimental validation.

In addition, it is likely that there will be other discoveries made during the project. For example, we can envisage starch surface sensors being commercially useful for screening prospective starch processing enzymes, or novel enzymes for starch modification (glycosyltransferases, esterases, lipases). Therefore, consideration will be given to protecting IP associated with sensor chip design and application.

Ultimately, all data will be made freely available over the consortium website and as open access papers. Computer models will be made freely available as documented and version controlled open source software. For details on data sharing see the separate Data Management Plan.

8.10 Project management and reporting (max. 1.5 A4 page)

The project is coordinated by Oliver Ebenhöh at HHU Dusseldorf. Scientifically, the project is divided into four workpackages (WP1-4). Oliver Ebenhöh and Rob Field will lead one WP each (WP4 and WP1, respectively), Sam Zeeman will lead WP2 and WP3.



Because of the interdisciplinarity of our collaborative research project, a regular communication is key to success. We will implement the following means to ensure efficient communication between the partners.

As detailed in our Data Management Plan, all project partners will regularly submit their results to the central Data Management Unit, also located in the project coordinator's group at Düsseldorf. This ensures that all partners always have access to all data and results produced within the consortium. To maintain a close exchange and personal communication, we will hold monthly seminars distributed over the Internet. Sam Zeeman and Oliver Ebenhöh have made excellent experience with such seminars, which are now a regular means of communication between PIs and fellows of the Marie-Curie Initial Training Network AccliPhot. Oliver Ebenhöh, as the coordinator of AccliPhot, was instrumental to establish this successful monthly webinar series and we will adopt this scheme to exploit modern technology to ensure a regular exchange between the partners. On these seminars the PIs and researchers will report on their progress to the partners in presentations. Additional written reports will be provided by the PIs to the project coordinator every 12 months and yearly consortium

meetings in person will be held, starting with a kickoff meeting at project start (May 2015) in Dusseldorf. Subsequent yearly meetings will be held in Zurich (2016), Norwich (2017) and the final meeting again in Dusseldorf (April 2018). These meetings will be used to summarise and critically assess the yearly progress, and agree on the written reports of the consortium. Research visits of the employed researchers to partner labs will complement the scientific exchange. These visits are necessary because many results will be obtained by a collaborative effort. The visits further provide additional training to the researchers giving them the chance to learn techniques that the partners are specialised in.

For external quality control, we will identify two experts (one theoretician and one synthetic biologist) to serve as an external supervisory board (SAB) before the start of the project. Members of the SAB will sign a confidentiality agreement and be granted access to all internal data, which is not yet released, to all internal reports, and be invited to join the webinars.

The duration of the tasks as defined above is indicated in the table below. Here, the dark shaded areas indicate the core activity in a particular task. Naturally, the success of the tasks is not always predictable, so we plan for a degree of flexibility to obtain missing results where we experienced unexpected difficulties or start early on tasks when the first results of previous tasks have been produced (light shaded areas). In general, the tasks are defined such that they do not depend on the full completion of other tasks but can start with partial datasets.

Task\Timeline (months)	**1-3**	**4-6**	**7-9**	**10-12**	**13-15**	**16-18**	**19-21**	**22-24**	**25-27**	**29-30**	**31-33**	**34-36**
WP1 Biochemical characterisation of enzymes involved in starch synthesis												
Task 1.1. Generation of recombinant forms of all enzymes involved in granule biosynthesis												
Task 1.2. Generation of defined glucan surfaces												
Task 1.3. Surface characterisation												
Task 1.4. Measuring the kinetics of interaction and modification of surfaces												
WP2 Engineering yeast strains to synthetically produce insoluble carbohydrate polymers												
Task 2.1. Using established yeast strains to inform a model on starch biosynthesis												
Task 2.2. Creating new yeast strains to understanding starch enzyme specificities												
Task 2.3. Quantitative protein expression in yeast												
WP3 Generation of modified Arabidopsis plants producing starch with altered characteristics												
Task 3.1. Expression of variant enzymes in Arabidopsis mutants												
WP4 Development of Mathematical models of starch synthesis as theoretical frameworks and predictive tools												
Task 4.1. Theoretical description of single enzymes kinetics												
Task 4.2. Theoretical advancement, biophysical processes + simulation of reconstituted pathways												
Task 4.3: Modelling complex starch synthesis pathways and in planta systems												

8.11 Legal requirements

☐ Yes **X** No (if "no" explain the current status)

Not applicable

8.12 References (max. 2 A4 pages)

1. Edwards, A. *et al.* Specificity of starch synthase isoforms from potato. *Eur J Biochem* **266,** 724–736 (1999).

2. Putaux, J.-L., Potocki-Véronèse, G., Remaud-Simeon, M. & Buleon, A. Alpha-D-glucan-based dendritic nanoparticles prepared by in vitro enzymatic chain extension of glycogen. *Biomacromolecules* **7,** 1720–1728 (2006).

3. Sun, Z. & Henson, C. A. Degradation of Native Starch Granules by Barley alpha-Glucosidases. *Plant Physiol* **94,** 320–327 (1990).

4. V. Planchot, P. Colonna, D. J. Gallant and B. Bouchet. No Title. *J. Cereal Sci.* **21,** 163–171 (1995).

5. Tawil, G. *et al.* In situ tracking of enzymatic breakdown of starch granules by synchrotron UV fluorescence microscopy. *Anal Chem* **83,** 989–993 (2011).

6. Murakawa, A., Mori, T. & Okahata, Y. Direct Monitoring of Both Phosphorolysis and Elongation of Amylopectin Catalyzed by Phosphorylase on a 27-MHz Quartz-crystal Microbalance. *Chem. Lett.* **36,** 312–313 (2007).

7. Nishino, H., Murakawa, A., Mori, T. & Okahata, Y. Kinetic studies of AMP-dependent phosphorolysis of amylopectin catalyzed by phosphorylase b on a 27 MHz quartz-crystal microbalance. *J Am Chem Soc* **126,** 14752–14757 (2004).

8. C. Cle, C. Martin, R. A. Field, P. K. and S. B. No Title. *Biocatal. Biotransform.* **28,** 64–71 (2010).

9. Clé, C. *et al.* Detection of transglucosidase-catalyzed polysaccharide synthesis on a surface in real time using surface plasmon resonance spectroscopy. *J Am Chem Soc* **130,** 15234–15235 (2008).

10. Schofield, C. L., Field, R. A. & Russell, D. A. Glyconanoparticles for the colorimetric detection of cholera toxin. *Anal Chem* **79,** 1356–1361 (2007).

11. Schofield, C. L. *et al.* Colorimetric detection of Ricinus communis Agglutinin 120 using optimally presented carbohydrate-stabilised gold nanoparticles. *Analyst* **133,** 626–634 (2008).

12. O'Neill, E. C. *et al.* Sugar-coated sensor chip and nanoparticle surfaces for the in vitro enzymatic synthesis of starch-like materials. *Chem. Sci.* **5,** 341–350 (2014).

13. Mikkelsen, M. D. *et al.* Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab Eng* **14,** 104–111 (2012).

14. Nakatani, H. Monte {C}arlo simulation of 4-α-glucanotransferase reaction. *Biopolymers* **50,** 145–151 (1999).

15. Nakatani, H. Monte {C}arlo simulation of multiple attack mechanism of β-amylase-catalyzed reaction. *Biopolymers* **42,** 831–836 (1997).

16. Nakatani, H. Monte {C}arlo simulation of hyaluronidase reaction involving hydrolysis, transglycosylation and condensation. *Biochem. J.* **365,** 701–705 (2002).

17. Kartal, O., Mahlow, S., Skupin, A. & Ebenhöh, O. Carbohydrate-active enzymes exemplify entropic principles in metabolism. *Mol Syst Biol* **7,** 542 (2011).

18. Kartal, O. & Ebenhöh, O. A generic rate law for surface-active enzymes. *FEBS Lett* **587,** 2882–2890 (2013).

19. Tetlow, I. J. *et al.* Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. *Plant Physiol* **146,** 1878–1891 (2008).

20. Hennen-Bierwagen, T. A. *et al.* Starch biosynthetic enzymes from developing maize endosperm associate in multisubunit complexes. *Plant Physiol* **146,** 1892–1908 (2008).

21. Glaring, M. A., Skryhan, K., Kötting, O., Zeeman, S. C. & Blennow, A. Comprehensive survey of redox sensitive starch metabolising enzymes in Arabidopsis thaliana. *Plant Physiol Biochem* **58,** 89–97 (2012).

22. Streb, S. *et al.* Starch granule biosynthesis in Arabidopsis is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase. *Plant Cell* **20,** 3448–3466 (2008).

23. Pfister, B. *et al.* Genetic evidence that chain length and branch point distributions are linked determinants of starch granule formation in Arabidopsis. *Plant Physiol* (2014). doi:10.1104/pp.114.241455

24. Ruzanski, C. *et al.* A bacterial glucanotransferase can replace the complex maltose metabolism required for starch-to-sucrose conversion in leaves at night. *J Biol Chem* (2013). doi:10.1074/jbc.M113.497867

9. Breakdown of costs

Table 9.1 Requested costs (overview)

CRP Partner No.	Partner Affiliation	Country	Personnel costs (total) (k€)	Consumables (k€)	Travel (k€)	Equipment (k€)	Other (k€)	Total requested (k€)
P1	Heinrich-Heine-University	Germany	190.8	5.2	4.8	4.2	41	246
P2	ETH Zurich	Switzerland	253.7	60	4.8	0	13	331.5
P3	John Innes Centre	United Kingdom	161.6	64.3	4.8	0	247.1	477.8
Totals			606.1	129.5	14.4	4.2	301.1	1055.3

Please give a justification for all requested costs in the different categories (per Partner):

Breakdown of Personnel costs

CRP Partner No.	Type of personnel [postdoc/ PhD/ technical assistant/ other]	Person months	Person costs/month (k€)	Personnel costs (total) (k€)
P1	Postdoc	36	5.3	190.8
P2	Postdoc	36	7.05	253.7
P3	Principal Investigator (Rob Field)	2.88	12.6	36.4
P3	Postdoc	28.8	4.3	125.3

…				
P4				

Please justify requested personnel category and persons months in each case.

A breakdown of requested costs for standard consumables is not generally required. Non-standards items, especially if costly, should be mentioned and justified.

For a breakdown of requested travel costs, please list the total amount for conference visits, the amount for visiting collaboration partners, and any other project-related travel costs.

Requested pieces of equipment should be listed and justified.

"Other costs" should be specified and justified. Overhead costs should be listed in brackets under "Other costs" unless specified otherwise by the respective National financial regulations.

P1: Personnel costs are requested for 36 months of a postdoc, to be named. The postdoc, who will have training in physics/biophysics and experience in theoretical approaches to biology, will develop mathematical models for enzyme kinetics for polymer-active and surface-active enzymes and perform simulations of the reconstituted *in vitro* and yeast systems. We request €4250 for consumables, which includes the purchase of software and database licences (e.g. KEGG), as well as a contribution towards recruitment costs and towards transferable skills training for the postdoc and €2250 for publication costs. A further €4800 is requested for attendance at project meetings by the PI and postdoc to liaise with partners and present results at international conferences (e.g. International Conference on Systems Biology). We request equipment for €4200, which include a high-performance PC to carry out simulations (€3000) and a network attached storage drive (€1200) to secure and make available the data resulting from the project (see Data Management Plan). Other costs (€40700) are requested as contribution to overheads (20% of incurred costs) as specified by the regulations of the German Research Foundation.

P2: Personnel costs are requested for 36 months of a postdoc, to be named. The postdoc will have training in biochemistry and molecular biology techniques and experience working with microbial cultures and plants. The postdoc will perform cloning and mutagenesis of starch biosynthetic enzymes for expression in *E. coli*, *S. cerevisiae* and *A. thaliana*, purification of recombinant proteins as well as biochemical and biophysical analyses of soluble and insoluble glucan structures. We request €20000 per year for consumables which will cover the costs for basic laboratory consumables and labware, growth facilities and media, the extensive kits and reagents needed for cloning, sequencing and gene synthesis activities, columns for HPLC and a contribution towards the HPLC maintenance costs. A further €4800 is requested for attendance at project meetings by the PI and postdoc to liaise with partners and present results at international conferences. No additional equipment is required for the project, but a further €13000 is requested to cover the costs of work done in the context

of external collaborations (e.g. SLS beamline) and for open-access publication costs. Indirect costs (overheads) are not listed here, but will be dealt with as specified by the regulations of the Swiss National Science Foundation.

P3: Personnel costs are requested for 36 months of an unnamed postdoc, to be named. The postdoctoral research assistant (PDRA), who will have skills in carbohydrate (bio)chemistry, will perform kinetic analysis of Arabidopsis starch-active enzymes by biosensor and nanoparticle approaches. We also request a contribution towards 10% of the costs of Professor Rob Field. He will manage the work to be done at JIC, supervise the PDRA, liaise with the other partners and contribute to manuscript preparation and dissemination. We also request €43200 for general laboratory consumables e.g. enzymes, laboratory plastics, €7680 for use of the JIC mass spectrometry service and €9600 for use of the JIC NMR facility to characterise synthetic and biologically-derived glycan samples. We also request €3840 under Consumables as a contribution towards recruitment costs for the PDRA and a contribution towards the costs of transferable skills training for the PDRA. These costs are all based on costs for similar current projects in our laboratory. A further €4800 is requested for attendance at project meetings by the PI and postdoc to liaise with partners and present results. €247113 is requested for Indirect costs, Estates costs, contribution towards depreciation, pool staff and infrastructure technicians. These are all calculated on the basis of staff time on the project in accordance with BBSRC guidelines, at rates approved by BBSRC. All figures given are for 80% of Full Economic Cost, as requested from BBSRC, and costed at £1=€1.2 as stated in the BBSRC guidelines for this call.

Table 9.2 Own contribution

CRP Partner No.	Person months	Type [postdoc/ PhD/ technical assistant/ other]	Personnel costs (total) (k€)	Consumables (k€)	Travel (k€)	Equipment (k€)	Other (k€)	Total contributed (k€)
P1								
P2								
P3	0.72 7.2	Principal Investigator (Rob Field) Postdoc. to be named	40.4	16.1	1.2	0	61.8	119.5

Totals			40.4	16.1	1.2	0	61.8	119.5

Please give a justification of the own contribution in the different categories.

10. Funding from other sources for all or part of this work

This research program is based on techniques and materials developed in each of the partner's labs. The yeast and Arabidopsis strains that will be used are being produced in the Zeeman lab by a PhD student funded through a Swiss National Science Foundation (SNF) grant. However, none of the experiments proposed here are part of that SNF proposal, which is multi-faceted and also focusses on other experimental approaches. Rather, the SNF-funded research has enabled this extension to a transnational and interdisciplinary project, which will create a much greater overall impact.