OXYMOD - Optimized oxidative enzyme systems for efficient conversion of lignocellulose to valuable products

1. Connection to Centre for Digital Life Norway

The OXYMOD project will through a transdisciplinary approach define, develop and demonstrate applicability of **new enzyme systems for the efficient biocatalytic conversion of lignocellulose** from abundant Norwegian biomass into valuable products like sugars and aromatic building blocks. OXYMOD will focus on the still largely underexplored and –exploited group of **redox enzymes** and their potential in the depolymerisation of cellulose, hemicellulose and lignin, including aspects such as redox enzyme interplay, co-factors and reaction partners, as well as their interplay with hydrolytic enzymes. OXYMOD will address these enzymes and enzyme systems as they occur and function in, among others, a unique in-house collection of approx. 1000 marine Actinobacteria isolates with genomes recently sequenced by SINTEF. Initial analysis of some of these genomes has already revealed diverse redox enzyme candidates of interest for further processing in the OXYMOD project, thus validating the high relevance of this biodiversity source.

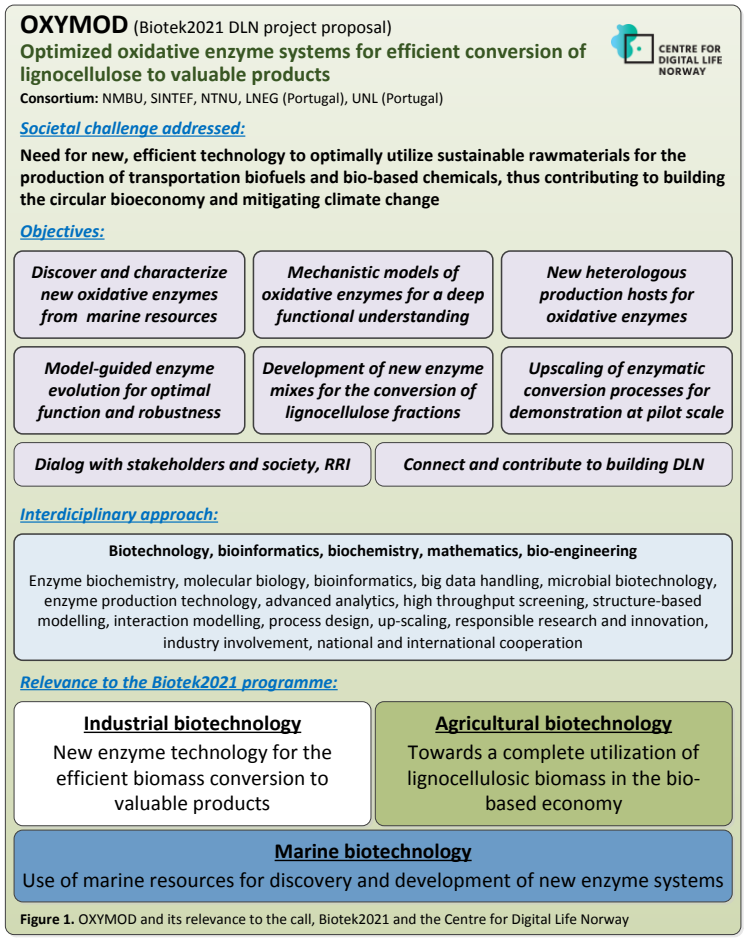


Figure 1. OXYMOD and its relevance to the call, Biotek2021 and the Centre for Digital Life Norway

Redox enzymes require co-factors and redox partners, and there is a considerable degree of cooperativity between different enzyme classes.[1] Understanding and eventually engineering the efficient degradation of lignocellulose by these enzyme systems, a key ability in a future **circular bioeconomy**, requires a transdisciplinary approach far beyond "simple" enzyme discovery.

OXYMOD combines life sciences (enzyme biochemistry, enzyme production technology, microbial biotechnology, high throughput screening, advanced analytics), ICT (bioinformatics, big data handling), mathematical sciences (enzyme systems modelling, process modelling) and engineering (enzyme evolution, synthetic biology) for developing new and optimized biocatalytic systems for industrial application, primarily within the agricultural and forest sectors. It is therefore highly relevant as a project in the Norwegian Centre for Digital Life (DLN).

Besides the enzymes and enzyme systems themselves, additional innovations from OXYMOD concern the generation of **well-defined products streams**, primarily sugars from (hemi-)cellulose and aromatic building blocks from lignin for a variety of downstream applications (e.g. biofuels & bioplastics).

The OXYMOD project is equally relevant for the Biotek2021 thematic focus areas **industrial biotechnology and agricultural biotechnology**. Since the project predominantly explores marine microorganisms and (meta-)genomes as the source of new enzyme systems, it is also relevant for the area of **marine biotechnology**. The project is highly relevant for current efforts to establish a strong bio-based economy in Norway and the use of biotechnology for that purpose.

The national project consortium comprises NMBU, SINTEF and NTNU, which are all already connected to DLN as either nodes (NMBU, SINTEF) or hub partner (NTNU). The proposed project will link the three partners more closely to DLN and add another dimension (enzyme systems) to the DLN Centre. In addition, the OXYMOD project includes **international collaboration** with two strong partners in Portugal, i.e. the National Laboratory of Energy and Geology (LNEG) and the New University of Lisbon, Group of Biological Chemistry (UNL), which complement the competences of the Norwegian project partners and are internationally renowned in the fields of biomass conversion and oxidative enzymes, respectively.

The OXYMOD consortium will readily **connect to the Work Groups (WGs) of DLN** to ensure seamless integration of the project as part of the Centre. The project interacts closely with several relevant commercial stakeholders who will be represented in an OXYMOD **Industrial Reference Group** (→WG2-Innovation & Industry involvement). National leading infrastructure and competence for high throughput screening, bioprocess technology and advanced analysis at NMBU, SINTEF, and NTNU, already linked to the DLN through WG4-Competence & Infrastructure network, will be used in the project. Additional platforms like the Norwegian NMR platform (NNP) and NorBioLab, as well as the FME Bio4Fuels, will be linked to the Centre via OXYMOD and its project partners. The aspects of big data handling and mathematic modelling in the project will benefit from a close link to DLN WG4. **Responsible Research and Innovation (RRI)** is an integral part of the project including one dedicated PhD student and active project participation of a key expert from NTNU in the field, who will, together with the other partners, throughout the project also closely interact with RRI experts at DLN. (WG1-Leadership & Governance & RRI). **Four PhD students and one post-doc** will be employed in the project, and their education will include participation in the PhD school of DLN (→WG3-Training & Recruitment). One representative of the consortium will be responsible for the day-to-day communication with the DLN Centre coordinators, ensuring smooth and efficient communication and flow of information. The project leader will participate in the project leader forum of the Centre and be the primary contact for the DLN Centre leaders.

2. Societal and scientific challenges

2.1 Societal challenges addressed by the project: Challenges connected to the emission of greenhouse gases have high societal focus. There is an **increasing demand for a "green shift"**, moving towards an economy with a decreased dependence on oil, as a result of a more efficient utilization of biomass, e.g. from agriculture and forestry. Today the annual production of biofuels comprises about 24 billion litres of biodiesel and 89 billion litres of bioethanol.[2] These biofuels are so-called 1st generation biofuels produced from oil plants or sugar/starch crops, feedstocks that also can be used as food. Therefore, efforts are made to develop **2nd generation biofuels** based on lignocellulosic biomass, where the production process entails thermal pretreatment, enzymatic saccharification and fermentation, while side-streams such as lignin and fermentation residues are mainly used to generate heat and electricity. Commercial production of cellulosic bioethanol from woody biomass has not yet been established, however, Borregaard of Norway produces 20 million litre of bioethanol from spruce mannan annually as part of its integrated biorefinery. Current systems for lignocellulosic biomass conversion are relatively ineffective and far from approaching complete utilization. Sub-optimal biomass utilization leads to the generation of greenhouse gases such as methane through uncontrolled anaerobic degradation of unutilized and dumped biomass. Furthermore, higher-value conversion routes are desirable, for example to get more value out of lignin than the value created by simply burning it. OXYMOD is based on the idea that **breakthroughs in enzyme technology** can help in meeting these societal challenges.

In the proposed project, we will explore the potential of unique in-house strain collections and –omics data, and the newest of modelling technologies, at the individual enzyme and the enzyme systems level, to develop **new and highly efficient biocatalytic processes for land-originating biomass**. By exploring unique biodiversity resources, i.e. an extensive marine Actinobacteria strain collection at SINTEF/NTNU of which more than 1000 genomes have recently been sequenced by SINTEF, we can adopt a systems approach in enzyme discovery and development for efficient biomass processing. Actinobacteria are known biomass degraders, with complex but manageable enzyme systems. They are genetically tractable, and the comprehensive in-house pan-genomic dataset allows advanced correlation analyses.

2.2 Scientific challenges addressed in the project: The biochemical conversion of lignocellulosic biomass to fuels and other products holds considerable promise. Despite the increased knowledge and building of so-called second generation biofuel plants (see above, 2.1), many challenges remain that hamper wider technology implementation, e.g. in the Nordic countries. Current biorefineries, for example, mostly consider the **lignin fraction** as energy source for combined heat and power. However, lignin could have a much higher value, for example as the only direct **green source of bio-based aromatics**. Hydrolysis lignin (HL), i.e. the lignin-rich fraction remaining after polysaccharide degradation, is therefore an essential stream in a biorefinery, meaning that assessment of the quality and possible valorization routes of HL needs attention. Non-fuel applications of HL, such as conversion of lignin to phenolics, may yield very high market values.

Enzymatic depolymerization of plant biomass is an essential step in the biochemical conversion of biomass and requires the concerted action of enzyme mixtures on the complex substrate of cellulose and various

hemicelluloses and lignin. The enzymes need to deal with the co- and heteropolymeric, partly insoluble and partly crystalline nature of the feedstock. For most feedstocks, including woody biomass, some sort of thermochemical pretreatment is needed to facilitate enzyme action and to reach acceptable levels of enzymatic digestibility. In general, hydrolases are well studied and knowledge of these enzymes is quite abundant. The discovery of **Lytic Polysaccharide Monooxygenases (LPMOs)** as an important tool working in synergy with hydrolases has been a major breakthrough, enabling technological progress, but also raising novel research questions.[1, 3] Presently there is a general lack of understanding concerning redox enzymes and their important roles (together with hydrolytic enzymes, like cellulases and other glycoside hydrolases) in lignocellulose processing, both for cellulose degradation as well as for lignin modification and depolymerisation.[4]

Redox enzymes are particularly challenging to screen for, analyse, and optimize due to their complex mode of action, involving e.g. cofactors. In addition, several challenges lie within the understanding and further optimization of these enzymes and their interactions and in the integration of redox steps into industrial processes.[5] **Cooperativity of different enzymes**, both redox enzymes as well as hydrolases, needs to be better understood to aid optimization, system design and engineering. In addition, production of these enzymes might be challenging, a challenge which needs to be met by using state-of-the-art production systems employing innovative expression vectors and hosts.

To access novel enzyme candidates for application in lignocellulose processing, alternative and relevant biodiversity resources must be employed, such as the present in-house database of approximately 1000 marine Actinobacteria genomes at SINTEF. Actinobacteria have large genomes encoding many potentially relevant enzymes and are known to degrade cellulosic biomass.[6] The consortium has in initial analyses of a few of these genomes identified a large number of redox enzyme candidates of relevant classes, validating the high relevance of this biodiversity source. Targeted harvesting from such biodiversity is not simple, at least not if the ambition is to discover truly novel enzymes and enzyme systems. One way to break new ground is to harness the power of comparative genomics, transcriptomics and correlation analysis, which becomes a major, but potentially highly rewarding effort for such a large dataset. By taking this approach, OXYMOD will touch on systems biology and synthetic biology, where knowledge of enzymes and systems from various organisms is assembled into new combinations for increased system efficiency.

Several challenges and targets of OXYMOD will be met by **building a theoretical foundation of understanding both individual enzyme function and cooperatively working enzyme systems**. Experimental work will interact iteratively with *in silico* approaches used to generate a thorough understanding of how enzymes work and interact with substrates, as well as with each other, directly or via product/substrate relationships.

The OXYMOD project is founded on the following hypotheses and thoughts:

1. There is still considerable potential to improve the efficiency of enzymatic conversion of lignocellulosic biomass, especially for biomass that has not undergone harsh pretreatments and especially for processes where the goal is to preserve and valorize *all* components of the feedstock.
2. Despite recent progress in our knowledge of individual enzymes, there is a large gap in understanding how enzyme *systems* work and can be optimized.
3. Lignin enzymology needs to be developed as a key towards better understanding natural degradation processes and developing novel biorefining routes that preserve feedstock value; the interplay between lignin and polysaccharide conversion is of major interest and requires a systems approach.
4. The focus on redox enzymes requires novel methodological approaches, focusing on monitoring relevant redox species (including reactive oxygen species) and on the interactions between such species and the various enzymes.
5. For characterization and engineering of enzymes for lignocellulose conversion, efficient production systems and functional assays for these enzymes are required and need to be optimized.

3. Status of knowledge

3.1 Redox enzymes of relevance for lignocellulosic biomass degradation: For a long time, combinations of various types of glycoside hydrolases were thought to be the sole responsible for the enzymatic conversion of recalcitrant polysaccharides, as illustrated in Figure 2. The discovery of the LPMOs, in 2010, at NMBU, has changed this paradigm dramatically.

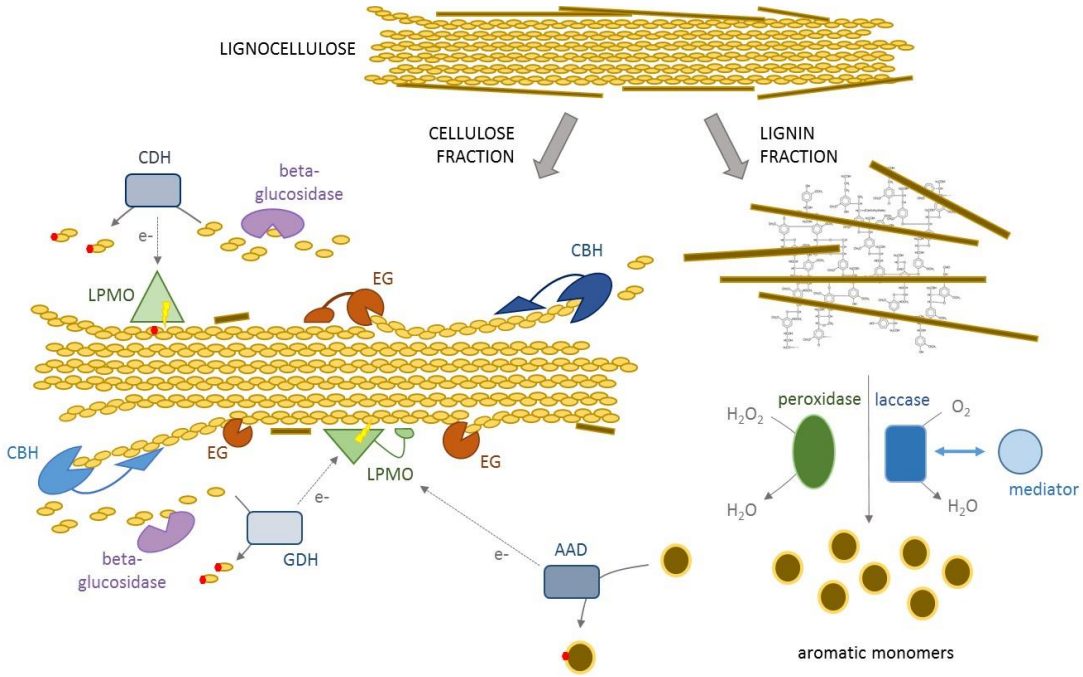


Figure 2. Schematic overview of known steps in enzymatic lignocellulose degradation, showing cellobiose dehydrogenase (CDH), glucose dehydrogenase (GDH), cellobiohydrolases (CBH) endoglucanases (EG), and lytic polysaccharide monooxygenases (LPMO) for cellulose degradation, and peroxidases and laccases, as well as aromatic amino acid decarboxylases (AAD) for lignin processing.

LPMOs, abundantly encoded in microbial genomes, are known to act on cellulose and a variety of hemicelluloses, whereas additional functionalities are likely to be discovered. Cellulose-active LPMOs are already being exploited in industrial biomass conversion processes and are of major importance for the efficiency of current commercial cellulase cocktails.[7] Importantly, LPMOs need copper, molecular oxygen and an external electron donor to function, which raises issues when it comes to design of industrial processes (i.e. access to oxygen and metals, operational redox stability). Studies with light and a variety of enzymatic and non-enzymatic electron donors have shown that LPMO activity can be boosted by better controlling electron supply and has also revealed interactions between seemingly different enzyme sub-systems, involved in biomass processing[1, 8] (e.g. cellobiose dehydrogenase and various CMC-oxidoreductases such as glucose oxidase). Oxygen and/or electron supply also relate to the oxidative stability of the enzyme, which is a major limiting factor for LPMO performance. LPMOs have also been shown to act on co-polymeric substrates[9] and may be pivotal in attacking areas where cellulose and hemicellulose interact, i.e. areas that are difficult to access by regular cellulases or hemicellulases.

The lignin fraction of the lignocellulosic biomass is essentially decomposed by peroxidases and laccases acting on the same substrate, however, using different reaction mechanisms (Figure 2). Both peroxidases and laccases are found in different variants (see below), and they act synergistically to modify the substrate by oxidation, resulting in depolymerisation of the aromatic structures. **Laccases** are copper-dependent oxidases that attach lignin by oxidation of aromatic moieties.[10] In contrast to peroxidases, laccases utilizes molecular oxygen as electron acceptor and generate water as the sole by-product. In spite of the relatively low redox potential, the activity of laccases towards lignin degradation can be improved in so-called laccase-mediator-systems (LMS).[10b, 11] **Peroxidases** include lignin peroxidase (LiP), manganese peroxidase (MnP), versatile peroxidase (VP), and dye-decolorizing peroxidase (DyP). LiP, MnP, and VP are extracellular fungal peroxidases that belong to the plant and microbial peroxidases superfamily. LiPs are strong oxidants with high-redox potential that oxidize the major non-phenolic structures of lignin. MnP is a Mn-dependent enzyme that catalyzes the oxidation of various phenolic substrates but is not capable of oxidizing the more recalcitrant non-phenolic lignin. VP enzymes combine the catalytic activities of both MnP and LiP that are able to oxidize non-phenolic compounds like LiP. DyPs occur in both fungi and bacteria and are members of a superfamily of heme peroxidases. DyP enzymes oxidize high-redox potential anthraquinone dyes and were recently reported to oxidize lignin model compounds.[12] Even after pretreatment, woody materials contain small amounts of remaining hemicelluloses, and it is well known that hemicellulases, in particular xylanases, can contribute to more efficient cellulose conversion, likely because they remove hemicellulose chains that are tethered with the cellulose, thus blocking enzyme access. In addition, hemicelluloses are cross-linked with lignin, which hinders further access to cellulose. Little is known about these issues for woody biomass, and

studies on the interplay between hemicellulases and cellulases are warranted. Clearly, the redox enzymes described above may also be relevant for tackling such "recalcitrant hemicellulose" and lignin-sugar linkages. Recent data from e.g. fungal secretomes, strongly suggest that the oxidative degradation of polysaccharides and lignin modification are connected, and that much may be gained from understanding how the various enzymes work together, i.e. we need to know how the enzyme *systems* work.

3.2 Technologies for enzyme discovery, production and characterization

In silico approaches for new enzyme discovery: Genomic data can be annotated using standard **bioinformatics tools**, and one may pay special attention to CAZymes, using DbCan and Laccase Engineering Database (LccED).[13] Potential candidate enzymes among "proteins of unknown function" (i. e. proteins not annotated by the above standard methods) may be discovered using a variety of relatively novel approaches that harness the potential of comparative genomics delivered by the pan-genomic Actinobacteria genome data made available for OXYMOD. These methods may in part also be used to sub-classify "known" enzymes and to identify the most interesting candidates:

- The peptide-pattern-recognition algorithm (PPR) developed by Peter Kamp Busk.[14] PPR is an alignment-independent method that could reveal novel sequences of enzymes sharing the same functions with characterized ones even in cases of low sequence identity between them. The method has proven to be successful to identify novel glycoside hydrolases[14] and LPMOs[15] from different fungal genomes.
- Sequence Similarity Network (SSN)[16] provides a method to visualize sequence relationships within groups of sequences based on collections of independent pairwise alignments between sequences. One major advantage of SSNs in comparison with traditional phylogenetic trees is that a very large amount of protein sequences can be visualized using a network which can also reveal the trends in function-related information. Therefore, SSN can be applied on a large dataset such as Actinobacteria genome sequence libraries to discover novel enzyme sequences that are related to the characterized ones. Enzyme Function Initiative (EFI) provides an accessible way to generate SSNs.[17]
- Correlation studies on a very large set of (comparable) genomes, as available in OXYMOD. This implies looking for patterns of co-occurrence, a rather simple but successful example of which appeared in a recent high-impact study of fungal genomes.[1]
- Domain association analysis, i.e. looking for domains of unknown function that are linked to domains with a known and relevant function, e.g. a cellulose-binding domain.[18]

Obviously, comparative transcriptome analysis e.g. of strains grown in the presence and absence of a defined substrate can be applied to further target enzyme selection, as also suggested in OXYMOD.

Homology modelling of proteins provides a preliminary structural impression of new enzyme candidates obtained from data mining. There are tools available that can be used to create homology models of the proteins from amino acids sequences such as PHYRE[19], SWISS-MODEL[20], and I-TASSER.[21] As part of an enzyme candidate pre-screening process, the models can then be used for a preliminary assessment of possible enzyme activity.

Functional screening for new enzymes in (meta-)genomic libraries: Functional screening of (meta)genomic libraries is a powerful tool for discovery of novel enzymes,[22] being independent on sequence homologies to previously known enzymes.[23] Regardless of the enzyme function searched for, large libraries are often required due to a general hit rates of ≤6‰,[24] requiring high throughput screening (HTS) technologies for the primary screening in multi (96, 384 or 1536) well plates. Screening for industrially relevant activity on insoluble biomass is not trivial, requiring e.g. MS-based detection of products, as developed by SINTEF and NMBU for glucose 1-6-mers from cellulase activity and oxidized oligosaccharides from LPMO activity.

Enzyme production systems: Many expression systems are available for production of recombinant enzymes. However, such systems should be controllable, give large expression levels, and allow the use of signal peptides for translocation of the soluble protein product to the periplasm and/or culture medium. One such expression system is the *E. coli*-based VB expression technology platform of Vectron Biosolutions, which is based on the *Pm*/*xylS* expression cassette combined with minimal replicons of the RK2 plasmid. This system has been demonstrated to be suitable for production of a large range of various industrial enzymes, including LPMOs (Aachmann, manuscript in preparation). However, new expression hosts are desired in order to increase the chances of high-level soluble production of enzymes in at least one system. *Streptomyces* spp. represent interesting new hosts for heterologous expression of Actinobacteria enzymes encoding genes, as suggested to be explored in OXYMOD.

Big data management and handling: Good data management, allowing the organization, management, and exchange of interdisciplinary data is crucial in OXYMOD. The FAIRDOM/SEEK platform provides a complete solution to maintain the consistency and integrity of data, integrate data with models, as well as to provide analysis methods within the system.[25] The platform will be used in the project to facilitate the data management for mining of new enzymes and the integration of data and models in continuous feedback between tasks described below. In addition, OXYMOD will implement the use of Synthetic Biology Open Language (SBOL)[26] that has been adapted widely within the synthetic biology community to annotate and exchange data and models among different disciplines.

3.3 Modelling of enzymes and enzyme function

Molecular dynamics and quantum mechanics: Modelling at the atomic and electronic level is becoming a powerful tool for understanding and developing enzymes and enzyme systems. Longer molecular dynamics simulations are easily accessible and are suitable for exploring conformational space and enzyme-ligand interactions. Quantum mechanical methods are increasingly used for studying reactions in which bonds are formed and broken, and are highly suitable for enzymes containing metal centres active in oxidation and reduction, such as LPMOs[27] and laccases.[28] When multiple copper ions are present, the sites are referred to as either coupled or uncoupled, depending on Cu interatomic distances. A distinct difference between the identified mechanisms in the LPMOs and laccases is a reactive superoxide formation during the catalytic cycle of the LPMOs. Variations in effects of backbone structure have been reported.[27, 29] DFT studies treating the Cu sites separately have revealed complicated O_2 reduction mechanisms in laccases.[30] By including two separate QM domains in a QM/MM calculation it was shown that redox potential and acidity constants of either site is considerably affected by the state of the other site.[28] Electron transfer over long distances may be crucial for the interaction between LPMOs and cellobiose dehydrogenases (CDH) and a structural basis for this electron transfer was recently presented.[31] In OXYMOD this range of Cu active sizes will be canvassed and compared for similarities and differences in functional properties. The different backgrounds of the OXYMOD consortium partners will allow us to jointly implement state-of-the art methods to tackle questions concerning reaction mechanisms, mutational effects, ligand binding energies, and so on. Notably tight collaboration between theoreticians and experimentalists doing e.g. EPR (at NMBU), NMR (at NTNU; see below) and ITC (at NMBU) will accelerate our understanding of the enzymes in question.[32] Additionally, interactions between molecular modellers and process modellers, as envisaged in OXYMOD, will eventually allow systems approaches to enzyme functionality.

Experimental studies of enzyme structure and dynamics: NMR spectroscopy is a versatile tool to gather information all the way from structures, ligand interactions and dynamics with atomic level resolution to enzymatic process relevant data. Thus, NMR provides empirical data to guide and validate the modelling at several levels. NMR spectroscopy is a renowned method to solve the 3D structure of proteins in solution, were the sample conditions can be changed easily. This allows real-time interaction studies of protein-ligand complexes, where the interaction area on the protein surface is monitored by NMR techniques. Due to the shared timescale of molecular motion and nuclei spin motions, NMR provides insights into the dynamic processes of molecules such as proteins. All the information provided by NMR can be used to investigate structure-function relationships. Furthermore, NMR is an excellent technique for characterizing enzymatic turnover, as the absolute concentrations of all intermediates can be directly monitored at the same time in the NMR tube. For determining the structures of larger enzymes and enzyme complexes, X-ray crystallography is the method of choice and will be used in OXYMOD, where needed. The enzymes of interest can be crystalized together with their substrates or inhibitors, which in some cases may provide insights into the reaction mechanism.

3.4 Protein engineering:

Protein engineering, either by rational design or by screening-based approaches, is a valuable tool for improving enzyme properties.[33] With the increasing number of structure and sequence databases, many combinatorial approaches for generation of sensible diversity have been developed. Combinatorial active-site saturation test (CAST), iterative saturation mutagenesis (ISM), and iterative CASTing are methods of step-wise randomization of residues in the substrate-binding pocket.[34] Focused database-driven mutagenesis strategies have been shown an effective way to create smaller and "smarter" libraries. Many available computational tools, either as web-based services or desktop software, that can be used to analyse amino acid sequences or to visualize protein structures and to predict the potentially relevant residues for mutation. For example, B-FITTER for improving of thermostability or Hotspot-Wizard for identifying residues to improve substrate activity or selectivity. Obviously, the very large collection of

Actinobacteria sequences, allowing a range of clever comparative analyses, will provide a good starting point for designing smart libraries.

4. Scientific objectives

Primary objective: Define, develop and apply innovative enzyme systems that fully harness the potential of redox enzymes for the efficient conversion of lignocellulose from Norwegian biomass into valuable products such as sugars and aromatic building blocks.

Sub-goals:

- Discover and characterize redox enzymes from a proprietary strain collection of marine Actinobacteria and derived comprehensive genome databases using an advanced bioinformatics toolbox and new functional assays in high throughput screening.
- Establish mechanistic models of selected enzymes and describe reaction mechanisms at atomic scale.
- Develop efficient heterologous production systems for redox enzymes, including using a *Streptomyces* production host.
- In-depth experimental studies of the functional interplay between redox enzymes acting on lignin and/or polysaccharides and classical hydrolases, accompanied by system modelling efforts to better understand limiting factors.
- Perform enzyme evolution for functional optimization of relevant enzymes and enzyme systems.
- Model and combine multiple enzyme functions in new enzyme mixes for conversion of lignocellulose and isolated fractions thereof (cellulose, lignin), with extensive experimental verification.
- Perform upscaling of new enzymatic conversion processes to small pilot scale.
- A continuous dialog with relevant stakeholders and society on technology development for the bio-based economy, its chances and limitations.
- Closely connect to and contribute to a positive development of the Centre for Digital Life Norway as a lighthouse of Norwegian biotechnology R&D.

Main deliverables, impact and achievements of the project:

- ✓ A comprehensive database of classified enzyme candidates from genomes of marine Actinobacteria as a source for further development within and beyond the OXYMOD project
- ✓ A dedicated LPMO web resource
- ✓ Optimized, robust enzyme systems for efficient conversion of wood-derived lignocellulose and fractions thereof into fermentable sugars, high quality lignin fractions, and/or aromatic chemical building blocks
- ✓ New and optimized heterologous production systems for redox enzymes, including simultaneous production of multiple enzymes
- ✓ In depth, atom-scale functional understanding of selected redox enzymes in the LPMO, laccase, and peroxidase families
- ✓ Consolidated collaboration between leading biotech research organizations in Norway and Portugal
- ✓ New enzymatic processes, readily adaptable by the Norwegian biorefinery industry
- ✓ Consolidate involvement of the participating institutions in the Centre for Digital Life Norway

5. Transdisciplinary approach

In OXYMOD, the convergence of computational chemists, biophysicists and biochemists will provide a solid foundation for accelerated optimization of enzyme systems for efficient lignocellulose conversion. In an ongoing strategic project, SINTEF is applying QM and MD methods to in-house discovered laccases, whereas NMBU is conducting state-of-the-art QM/MM calculations on LPMOs. A large ensemble of various Cu, Fe, Mn, and flavin -containing redox enzymes will be studied both experimentally and computationally. This provides an excellent basis for developing tailored force field parameters for the non-standard residues. Furthermore, while atomistic and electronic computational methods may be used by biochemists, physicists and chemists alike, the project consortium brings together theoreticians with different backgrounds. This ensures transfer of methods between the various disciplines, as well as the possibility of new insight due to changing viewpoints. This is also true from the reaction chemistry point of view, where the project includes chemists with long experience in heterogeneous and homogeneous catalysis.

The application of newly developed bioinformatics tools for sequence data mining, such as sequence similarity networks (SSNs) and peptide pattern recognition (PPR), will help explore and discover new carbohydrate active enzymes in the project. Protein structure modelling provides an additional analysis angle to evaluate functions of discovered enzymes, which is also important for the engineering of single enzymes and enzyme systems. Enzyme production technology of Vectron Biosolutions will play an important role also in

screening and engineering effort as it is essential to have the enzymes expressed in a good amount to detect the activity. Information on structure, interactions, and dynamics of selected enzymes will be provided by NMR spectroscopy. The NMR data will be used as input for the computational modelling aimed at functional understanding and for suggesting modifications of the selected enzymes. These data will be the foundation for mathematical descriptions of enzyme process models.

Through existing collaborations and interactions with the Industrial Reference Group (IRG) the project also has a clear industrial component. Enzyme system development will be driven by realistic questions coming from realistic settings in industry conveyed to the project by industrial process engineers. The involvement of RRI expertise in the form of a direct project involvement in OXYMOD will secure a close dialog of the project activities with relevant stakeholders and society.

6. The project plan, project management, organisation and cooperation

6.1. Overall concept of the project plan: OXYMOD aims for the discovery, development and process implementation of new **redox enzyme systems in lignocellulosic biomass conversion**, with the aim of complete biomass utilization (cellulose, hemicellulose, lignin) to generate sugars and value-added chemicals. Though the main focus area is on new redox enzyme systems, combinations of these with other enzyme classes, like diverse glycoside hydrolases will be included. Enzyme systems will be modelled, experimentally tested and optimized in repetitive cycles in order to finally integrate them into scalable biomass conversion processes of industrial relevance. The project builds on the partners' complementary infrastructure platforms and competence, as well as existing comprehensive in-house biodiversity resources such as strain collections, DNA sequence databases, and metagenomics libraries (Figure 2).

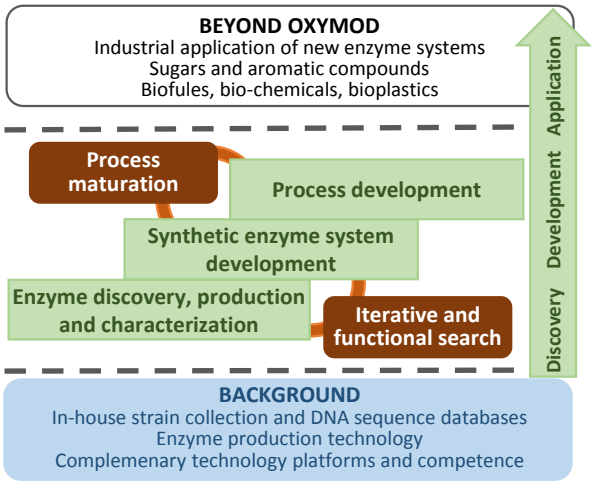


Figure 2. Integrated transdisiplinary concept of OXYMOD.

The project focuses on **lignocellulosic biomass** (cellulose, hemicellulose, and lignin) from abundant Norwegian resources. Initially, enzyme systems will be tested on synthetic and more simple substrates, but the overall focus will be on natural substrates such as wood chips, saw dust, hydrolysis lignin, Kraft lignin, black liquor, as well as other industrially relevant substrates. Biomass pretreatments, such as steam explosion [35] will be carried out when appropriate.

The main **biodiversity source** for enzyme discovery is a collection of more than 1000 marine Actinobacteria strains at SINTEF/NTNU with recently determined genome sequences. This strain collection originates from the Trondheim fjord, receiving large amounts of forestry and agriculture influx. The existing genome sequence pool, comparative transcriptome analysis of selected strains, and high throughput functional screening will in parallel be applied in OXYMOD to discover new redox enzymes. Initial sequence analysis of selected marine Actinobacteria genomes has already revealed a large number of redox enzyme candidates of all relevant classes, proving these strains as a suitable source for the discovery of such enzymes.

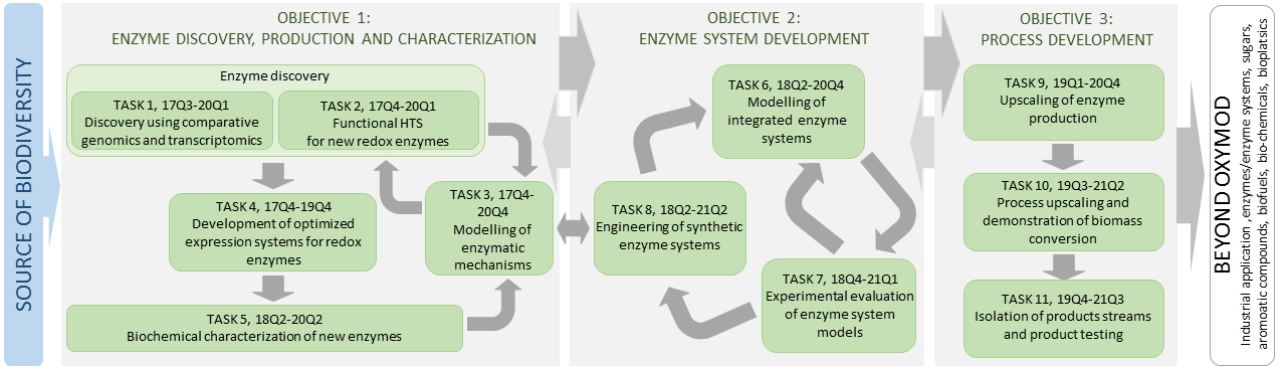
For efficient and environmentally benign degradation and utilization of the biomass, **enzymes and in particular cooperative enzyme systems** are needed. For **development of functional enzyme systems**, consisting of enzymes, co-factors, and redox mediators, modelling is crucial in order to utilize the full potential of the system. Members of each enzyme class will be subjected to detailed studies at both atomistic and electronic computational level, with the aim to develop unified methods that can be applied to all enzymes systems relevant for OXYMOD. A wide array of molecular modelling methods including molecular dynamics (MD) simulations, quantum mechanical/ molecular mechanical (QM/MM) calculations, pure QM calculations, and molecular docking calculations, will be used. Based on modelling, enzymes and enzyme systems will be engineered for optimization. The interaction of experimental and theoretical approaches will promote incremental development of the enzymatic systems.

Expected **end products** from cellulose and hemicellulose degradation are sugars and sugar fractions that can be used in downstream fermentation processes. Lignin degradation will produce aromatic monomers which are platform chemicals and precursors for fine chemical synthesis and biopolymers. End products are planned to be evaluated and used e.g. as sugars for fermentation processes for production of e.g. biofuels (e.g. in projects linked to the FME Bio4Fuels) Aromatic monomers from lignin degradation are also targets for new

biopolymers. Notably, the enzymes developed in OXYMOD may find a wide variety of application and may thus be end products themselves, one example being the use of laccases and peroxidases in bleaching and bioremediation processes.

6.2. Workplan

The proposed project has three main objectives; **Enzyme discovery, production and characterization**, **Enzyme system development,** and **Process development**, approached by 11 Tasks with their corresponding Milestones and Deliverables, as given in Figure 3.



MILESTONES

MS1.1. Sequence mining tools optimized for LMPOs, laccases and peroxidases and first > five candidates in expression pipeline [18Q2]
MS1.2. Paper on comparative genomics on Actinomycete strains [19Q1]
MS1.3. Comparative transcriptomics of Actinomycetes initiated [18Q4]
MS1.4./MS2.5. Extended enzyme candidate list [19Q3]
MS2.1. Library construction for functional screening [18Q1]
MS2.2. HTS methodology for functional screening for redox enzymes [18Q2]
MS2.3. HTS of enzymes (fosmids) and enzyme libraries, with different lignocellulosic substrates, [18Q3]
MS2.4 / MS8.1. First enzyme mutant with interesting properties [19Q1]
MS3.1. Establishment of molecular dynamics simulation methodology [18Q2]
MS3.2. Enzyme mechanism modelling methods estanblished; publication [19Q1]
MS4.1. Optimized expression systems for redox enzymes [18Q4]
MS4.2. > 20 Enzymes expressed in >0.1 gram amounts (19Q2)
MS5.1. Extensive enzyme characterization initiated [18Q4
MS5.2. >4 selected enzyme candidates for functional evaluation [19Q3]
MS6.1. First partial models for enzyme systems [19Q3]
MS6.2. First testable integrated model [20Q1]
MS7.1. Assembly and evaluation of modeled enzyme systems [19Q3]
MS7.2. Synthetic biology approach for enzyme system engineering [20Q1]
MS8.2. Optimization of individual enzyme components; > 2 papers [20Q2]
MS8.3. Positive trends upon iterative optimization of enzyme system composition [20Q3]
MS8.4. Prediction based screening for identification of optimal enzyme combinations [20Q4]
MS9.1. Gram-scale production for > 2(redox) enzymes [20Q1]
MS10.1. First process run in > 100 g scale [20Q1]
MS10.2. Demonstration of novel lignocellulosic biomass conversion process [21Q1]
MS11.1. Development of protocols for isolation of product streams [19Q4]
MS11.2. Demonstration of purified end product in > 100 g scale [21Q3]
MS0.1. Four PhDs [21Q4]
MS0.2. First joint paper with RRI [19Q4]
MS0.3. > 15 sceintific publications [21Q4]

DELIVERABLES

D1. Gererated enzyme candidate list
D2. Developed HTS methodology for (redox) enzyme functional screening
D3. Modelling methodology for enzymatic mechanisms
D4. Optimized expression systems for redox enzyme production
D5. Characterization of selected enzyme candidates
D6. Optimized modelling method for integrated enzyme systems
D7. Designed and optimized synthetic enzyme systems
D8. Upscaled process using developed enzyme system
D9. Demonstration of lignocellulosic biomass conversion
D10. Protocol for isolation of product streams

Figure 3. OXYMOD work plan overview. Milestones in bold face are part of the contract with the Research Council

Objective 1: Enzyme discovery, production and characterization

Task 1: Enzyme discovery. Homology-based data mining for sequences of new carbohydrate-active enzymes (CAZymes), and in particular LPMOs, laccases and peroxidases will be performed based on in-house curated actinomycete –omics databases using tools such as SSNs and PPR.[14, 16] Comparative genomics analysis will be applied in order to identify proteins/domains of unknown function possibly involved in biomass conversion by carrying out (1) extensive, big-data correlation analyses (co-occurrence patterns) and (2) domain association analysis (e.g. looking for domains of unknown function that are coupled to a known cellulose-binding domain). Selected isolates will be cultivated in the presence and absence of lignocellulosic substrates, and transcriptome analysis performed using RNAseq in order to identify genes that are upregulated by the substrate. These approaches will yield new enzyme candidates, which will be analyzed *in silico* prior to experimental verification.

Task 2: Functional high throughput screening for new redox enzymes. In parallel and complementary to the bioinformatics-based approaches presented in Task 1, fosmid libraries will be used in function-based screening, using pan-genomic DNA of selected Actinobacteria. Functional HTS assays for different redox enzyme classes of interest (LPMOs, laccases, peroxidases) will be adopted from the literature and the different partner groups (e.g. for LMPOs), or be developed in OXYMOD. Functional screening will involve direct spectrophotometric detection, coupled assays, and MS-based detection using Agilent RapidFire high throughput MS at SINTEF. Lignocellulose derived substrates of different quality will be used and the assays adjusted accordingly. Interesting fosmids ("hits") will be sequenced, and candidate genes will be subjected to experimental verification. Assay methods developed in this task are also relevant and will be adapted further in the approaches for enzyme and enzyme system optimization in Task 8.

Task 3: Modelling of enzymatic mechanisms of enzyme target classes. A significant challenge when studying molecular systems is to sample the plethora of possible conformational space sufficiently before e.g. QM or QM/MM simulations. To improve such sampling, we will apply replica exchange methodology and accelerated MD simulations, using state-of-the-art software as Amber, NAMD, and GROMACS and computational resources available at the NOTUR supercomputer centres. Molecular dynamics simulation

protocols will be developed for the three oxidative enzyme classes mentioned above, and when necessary the transition metal active sites will be parameterized using tools as Parmed (Amber package) and protocols developed by Seminaro.[36] Structures from MD trajectories will provide starting geometries for QM/MM and QM cluster calculations. Using QM methods, various reaction pathways can be tested, important active site amino acid residues can be identified, and their roles can be elucidated. This is information that is crucial for comparative studies and for translation of mechanistic determinants between systems. When building enzyme-substrate models, we will as far as possible utilize experimental data from e.g. NMR, mutagenesis studies, and activity assays, to guide model construction. For many of the enzymes involved in lignocellulose decomposition, the active site or electron transfer centers are partly or completely buried within the enzyme. We need to understand how electrons are transferred between active sites, and how this occurs. There are several possible options that can be tested with QM calculations, including tunneling effects, hydrogen transfer, or protein assisted electron transfer. With knowledge of electron transfer processes, potential transfer pathways relayed by amino acids can be identified, and care must be taken to conserve these when performing enzyme engineering in Task 8.

Task 4: Development and optimization of expression systems for redox enzymes. Vectron Biosolutions' *Pm/xylS* based expression platform is highly suited for this project due to the ability to induce expression using benzoate derivatives and possibility for fine-tuning the expression cassette by using different *Pm*, *xylS* and 5'-UTR variants for efficient production of enzymes. The induction of other commonly used expression systems, such as the araBAD operon and the pRSET system, is dependent on the carbohydrate metabolism of the cell, that negatively affects both yield and control of the system. The work will focus on optimizing expression of active enzymes by screening for the best combinations of *Pm* promoter variants, *xylS*, 5'-UTR, plasmid copy number and signal peptide. Cultivation conditions will be optimized for improved yields and solubility in the best enzyme producing strains, including fine-tuning of growth medium, concentration of inducer, time of induction, in addition to time of harvesting.

Actinobacteria derived enzymes are probably best produced heterologously in a closely related expression host of the same phylum. With the ERASysAPP project SYSTERACT and the Biotek2021-DLN project INBioPharm, SINTEF is currently coordinating two projects centered around Systems Biology of *Streptomyces coelicolor* and its use as a host for heterologous production of new bioactive compounds. This provides an excellent basis for exploring this organism as a heterologous production host for enzymes in OXYMOD. Protein expression in actinomycetes is quite well established in the field.

Task 5: Biochemical characterization of new enzymes. Enzyme candidates derived from activities in Tasks 1 and 2 will be ranked according to relevance with respect to the OXYMOD objectives and a subset taken further for in-depth verification of activity and full biochemical characterization. Established and newly developed enzyme production systems will be applied to produce sufficient amounts of enzymes at SINTEF, LNEG and UNL. Enzyme characterization will include determination of activity and selectivity on different substrates including catalytic constants, pH and temperature optima, multimeric state, co-factor content, need for reaction partners, need for and sensitivity to ROS, and general robustness (temperature/ chemical/inhibitor/solvent tolerance). Best candidates will be selected for in depth functional analysis by NMR, structure determination, as well as for further development as parts of new enzyme systems according to Objective 2. Data generated in Task 5 will be input for enzyme system modelling in Task 6.

Objective 2: Enzyme system development

Task 6: Modelling of integrated enzyme systems. Understanding how enzymes with different functionalities interact with each other and cooperatively when acting on lignocellulosic substrates can provide insight and guidance for the development of optimized enzyme cocktails for industrial scale processes. Redox enzymes often require other protein partners and co-factors to perform optimally. LPMOs, for example, can receive electrons from cellobiose dehydrogenases, and synergize with glycoside hydrolases. Peroxidases and laccases can work synergistically in the degradation of lignin to smaller aromatic monomers. Insights into enzyme-enzyme interactions, enzyme stability, enzyme cooperativity, co-factors and mediators, substrate specificity and turnover are needed. Biochemical data for individual enzymes and experimental data, e.g. from NMR, will be used as the basis for computational modeling of enzyme mixes, analogous to a cellular metabolic system that can be described using a Systems Biology approach. Since the number of different conversions and metabolites will be limited (starting with 2-3 and incrementally increasing complexity up to ~15-20 different enzymes), it will be possible to use methods like molecular dynamics to generate complete system

models that can be optimized in feedback loops with experimental evaluation in Task 7 and optimization of individual components in the mixture and the ratios between these in Task 8.

Task 7: Experimental evaluation of enzyme system models. Predictions from the enzyme system models generated in Task 6 will be used to assemble enzyme mixes that will be tested for performance in lignocellulosic substrate conversion. This will generate feedback to Task 6 for model verification or further refinement to improve predictability of the models. Multiple Systems Biology cycles of prediction, experimental evaluation and model refinement with successive increase of complexity will be applied in order to ultimately yield a reliable final predictive model for lignocellulose degradation based on the native enzymes. In a subsequent stage, the experimental evaluation tools will also be applied for the evaluation of the engineered synthetic enzyme systems to be developed in an integrated feedback loop between Task 6 and Task 8.

Task 8: Engineering of synthetic enzyme systems: Results from atom-scale modelling of individual enzymes in Task 3 and enzyme systems modeling in Task 6 will guide approaches for optimization of both individual enzymes (enzyme engineering and evolution) and enzyme mixtures (synthetic enzyme systems; Synthetic Biology). Tools such as the docking programs AutoDock, HADDOCK, and prediction servers like Hotspot-Wizard can be used for designing enzyme mutants or for devising directed evolution strategies based on clever and small libraries. Mutant libraries will be screened using (adapted) high-throughput functional assays from Task 2. In case of enzyme systems involving multiple enzymes (systems modelled in Task 6, plus enzymes designed/evolved in Task 8), the optimal combinations of variants from different enzymes will be identified through prediction-based screening. Optimized enzymes systems will give circular feedback to improving models in Task 6 via experimental evaluation in Task 7.

Objective 3: Process Development

***Task 9. Upscaling of enzyme production*.** For selected enzyme systems generated under Objective 2, upscaling of production for process implementation will be performed using optimized enzyme production technology from Task 4. Scale-up of enzyme production from the shake-flask format will primarily be done in 3-liter high cell-density cultivations (HCDC), including optimization of HCDC conditions for maximum enzyme yield and solubility. Further upscaling up to the 300 L scale will be considered where necessary. Fermentations will be performed using the fermentation platforms at SINTEF and NMBU, including multiple 1-3 litre bioreactors, as well as 30 and 50 litre reactors of the NorBioLab infrastructure, and a 300 litre pilot plant bioreactor at SINTEF.

***Task 10. Process upscaling and demonstration of lignocellulose and lignin conversion*.** Optimized enzyme systems from Objective 2 will be used in attempts to stepwise upscale lignocellulose conversion processes from micro-reactor scale up to a 50 L high dry matter pilot scale. Demonstration will be targeted for one complete lignocellulose biomass system and one selected lignin degradation system. Analysis of enzyme system performance at the 1-5 mL microreactor and 50-100 mL shake-flask (formats for testing conditions, pH, temperature, etc.), as well as 0,5-3 L fermentor scale will generate input data for process modelling and model optimization prior to further upscaling to pilot-scale (30-50 L reactor). At all stages, results will feedback to models of Task 6 to evaluate robustness of the model also during upscaling. Any new finding will be used to further refine the respective optimized enzyme systems models. Results from the upscaling studies will be used for exploitation in cooperation with relevant commercial partners beyond the scope of OXYMOD.

***Task 11. Isolation of product streams and product testing*.** During upscaling activities, protocols for the isolation of product streams from enzymatic lignocellulose and lignin conversion will be developed, focusing on scalable chromatographic methods for fractionation of sugars and aromatic monomers. Mono-sugars and sugar fractions will be used to perform fermentations at SINTEF for the production of different end products in the frame of other at that time ongoing projects at SINTEF, e.g. in connection to the FME Bio4Fuels. Data from such evaluations can give valuable feedback to OXYMOD for final enzyme system adaptations. If applicable, aromatic monomers produced from the lignin fraction will be channeled into relevant ongoing project activities, such as projects related to bioplastics production.

6.3. The OXYMOD consortium: The project consortium consists of partners covering all relevant expertise within the project (see Table below). The project has direct involvement of industry by Vectron Biosolutions, as a partner for enzyme production. In addition, other relevant industries are connected by participation in an Industrial Reference Group (IRG, see below). International collaboration includes two Portuguese partners

with high expertise in enzyme modelling and enzyme production and (redox) enzyme characterization, respectively.

Project relevant competence	Partner								
	Norway							Portugal	
	NMBU (Eijsink)	NMBU (Røhr)	NTNU (Aachmann)	NTNU (Myskja)	SINTEF Biotech.	SINTEF Nanotech.	Vectron Biosolutions	LNEG	UNL
Biodiversity resources	√				√				
Bioinformatics and sequence mining	√				√				
Enzyme modelling (stability, cofactors, cooperat.)		√	√			√			√
Structure/function modelling					√	√			√
Atom-scale modelling (mol. dynam., quant. mechan.)		√				√			√
Functional HTS screening					√				
Enzyme production technologies					√		√	√	
Enzyme/Enzyme system characterization	√	√	√		√			√	√
Enzyme/Enzyme system engineering	√	√	√		√				
Enzyme system and process integration					√			√	
Process engineering					√				
Project Management	√								
Dissemination	√	√	√	√	√	√	√	√	√
RRI				√					

6.4. Project management: The Project leader Vincent Eijsink (NMBU) will be responsible for the administrative, technical and strategic management of the project. A consortium agreement will be signed to regulate confidentiality, dissemination and intellectual property. The nature of the project and the consortium are such that project results will be owned by the research partners, which will define and bring in their own background to the project. The project will start with a kick-off meeting to which all involved researchers, as well as company representatives in the IRG (see below) will be invited. Bi-annual project meetings will follow in order to exchange and evaluate results and to plan further research. On each second project meeting, the IRG will participate and be encouraged to provide input to the research plan for the following year. All partners in the project will have one representative (PI) in a management team that will be responsible for the "daily" management of the project; this team will meet at least monthly by phone/Skype. A communication plan will be set up, also covering the RRI aspects as described in section 9. Already confirmed companies being represented in the OXYMOD IRG are Borregaard, ArcticZymes, and Norske Skog. Further relevant companies (e.g. from the groups of primary producers, biorefinery companies, enzyme companies, end users) will be invited once the OXYMOD project is granted.

6.5. Organization and cooperation:

The project leader. Prof. Vincent Eijsink (NMBU) is 54 years old and has broad experience in relevant fields, documented by > 240 published papers in international journals of good standard, including recent publications in Science, PNAS, and Journal of Biological Chemistry (h-factor is 52). He is a regular speaker at international conferences. Eijsink has ample experience from participating in and leading research projects, including large collaborative projects and projects in collaboration with industry (see CV for details).

National project partners:

Norwegian University of Life Sciences (NMBU). The work at NMBU will primarily be carried out in the Protein Engineering and Proteomics group at the Department of Chemistry, Biotechnology and Food Science. NMBU has access to all necessary equipment, either in-house or by access to Norwegian national infrastructure platforms such as the biorefining platform (= the NorBioLab national facility, which includes fermentation facilities for enzyme production and reactors for high dry matter processes). NMBU has an advanced instrument park for analysis of all thinkable forms of carbohydrates, oligosaccharides, lignin components as well as, as of late 2016, insoluble biopolymers (SEC-MALLS system run in ionic liquids). In addition to the project leader, Vincent Eijsink, the project team at NMBU will consist of the post-doctoral researcher funded by the project, Dr Åsmund K Røhr, recipient of a FRIPRO "talented reserachers" grant and one of Norway leading experts is MD and QM simulations of proteins, and female upcoming post-doctoral researcher talents Aniko Varnai and Zarah Forsberg.

SINTEF Materials and Chemistry (SINTEF). The **Department of Biotechnology and Nanomedicine** has for many years involved in industrial and publicly financed projects covering diverse applications of microbiology, molecular biology, bioprocess technology, high throughput screening and advanced MS-based analytics. Based on their state-of-the-art infrastructure platforms, the department held key roles in the recent large-scale Biotek2021 enzyme projects NorZymeD and MarPol. SINTEF Biotek is also currently responsible for the DLN project INBioPharm, led by Sr. res. Alexander Wentzel. SINTEF's HTS laboratory, comprehensive MS analytics facilities, molecular biology infrastructure, and its fermentation platform with bioreactors ranging

from micro scale to 300 L pilot scale will play central roles in OXYMOD. Key scientists: Giang-Son Nguyen, Geir Klinkenberg, Anna Lewin, Alexander Wentzel and Research Manager Håvard Sletta. The **Department of Materials and Nanotechnology**. The Department of Materials and Nanotechnology participates in the present project through its research group for nano- and hybrid materials. The group will contribute with competence in atom-scale modelling (atomistic and electronic) of enzyme function. Organic/inorganic macromolecules is a main competence in the group, which lends relevance to OXYMOD as well. Key scientists: Francesca Lønstad Bleken and Ole Swang. Stein Tore Johansen of the **Department for Oil and Gas Process Technology** will contribute to the project with process modelling and engineering competence.

Norwegian University of Science and Technology (NTNU). The **Department of Biotechnology** with NOBIPOL holds strong expertise in structure-function studies of polysaccharide active enzymes and proteins and in polysaccharide engineering. The department has state-of-the-art equipment for molecular biology, polysaccharides characterization and NMR spectrometers (800-400 MHz) and is part of the Norwegian NMR Platform (NNP). Prof. Finn L. Aachmann and the biopolymer NMR group have a broad experience in NMR spectroscopy-based structure determination of proteins and especially on LPMOs. The group have extensive expertise with molecular biology, isotopic labelling of proteins, protein purification as well as a solid background in physical-chemical characterisation of proteins and protein interactions. Besides Aachmann is currently leading the NT-faculty NMR center at NTNU and had a key role in building up the NMR structure determination facilities and he is in the board of the graduate school BioCat. The **Department of Philosophy and Religious Studies** has wide experience in research on bioethics, ethics of technology and political philosophy. Over the last decade, they have built competence on applied ethics in research connected to plant biotechnology, nanotechnology and medical ethics, and are among the leading research groups within Responsible Research and Innovation (RRI) in Norway. Notably, this Department allocates a full PhD student to the OXYMOD project. Key person: Prof. Bjørn K. Myskja.

Vectron Biosolutions AS. Vectron is a privately owned company that was founded in 2008 as a start-up from NTNU focusing on commercializing an expression vector technology platform originally developed in the research group of Prof. Svein Valla. The company has focused on marketing its technologies to the biopharmaceutical industry. This industry is constantly in need of improved expression technologies; in particular, because of the recent patent cliff that results in numerous biopharmaceutical drugs losing their patent protection allowing for the emergence of bio-similars and subsequent increased competition. Vectron recently had a significant market breakthrough with a deal with the major pharmaceutical company Boehringer Ingelheim RCV. In addition to patents on the Pm/xylS technology, the company has other proprietary expression technologies and non-patented know-how related to bacterial gene expression. Key scientists: CEO Dr. Trond Erik Vee Aune, CSO Dr. Jostein Malmo, Dr. Anne Krog, McS Maëliss Lemoine and MSc Pieter van Kuilenburg.

International collaboration:

National Laboratory of Energy and Geology (LNEG) is the largest state laboratory of the Ministry of Economy focused on R&D activity in close cooperation with private sector in the area of Energy and Geology. LNEG participates in a wide range of international projects convenes us a key partner role for European cooperation. The Unit of Bioenergy of LNEG is the largest R&D unit of LNEG with almost 70 members, being 35 PhD staff. The unit has currently four R&D Programs, focused on Biomass deconstruction, Cell Factories and Enzymes, Bioprocess Engineering, and Microalgae and Anaerobic Digestion for Bioenergy. Besides it hosts the only National Research Infrastructure for Biomass and Bioenergy, the Unit of Bioenergy has an accredited analytical Laboratory for Biofuels and Environment laboratory under ISO 17025, with a wide range of standard in-house methods and specific equipment for physical and chemical characterisation of biofuels and raw materials in compliance with European directives. LNEG holds the “HR Excellence in Research” Logo from the European Commission. Key scientists: Susana Marques, Luis Alves, Susana Alves, Florbela Carvalheiro, and Francisco Gírio.

New University of Lisbon, Group of Biological Chemistry (UNL). The UCIBIO-REQUIMTE FCT/UNL team is a leading group in the metalloproteins field. Main interest focuses on structure/function studies of metalloenzymes of denitrification and in the characterization of respiratory switches (nitrate versus sulfate) in sulfate reducing bacteria. Major contributions have been given on the characterization of the structure, reactivity, reaction mechanisms, and modeling of numberous proteins. Given examples are respiratory NaR, periplasmatic NaR, CcNiR, Cd1NiR, CuNiR enzymes and several hemic proteins, including di-heme peroxidades and KatG (enzymatic systems related to the interests of the project) and novel nitrite reductases

(dependent on molybdenum, playing a central role in the formation of signalling NO in mammals under hypoxia, and in ROS and RNS chemistry/biology). To achieve the main interests described, NMR, EPR and (Bio)Electrochemistry techniques are available. Protein-Protein interactions and Transient Electron-transfer Complexes have been studied, using NMR tools and a dedicated algorithm developed by the group (BIGGER). Key scientists: Isabel Moura and José J. G. Moura.
Importantly, several consortium partners have extended relevant international networks as documented by the publication lists in the attached CVs.

6.6 Budget: The project's active period is 4.42 years (53 months) with a total budget of 31 597 kNOK. 20 000 kNOK are applied for from the Norwegian Research Council. The 11 597 kNOK own contribution includes three PhD students provided by NTNU (2) and NMBU (1). The project will also have one post-doctoral researcher. A detailed budget is given in the electronic Grant application form.

7. Strategic aspects

Partners NMNU, NTNU and SINTEF are all connected to DLN, being nodes (SINTEF, NMBU) or Center (NTNU), and interactions are described in section 1. Partners NMNU, NTNU and SINTEF have a pronounced focus on bioeconomy, biorefinery and green biotechnology in their central strategies. NMBU, SINTEF and NTNU are all involved in several running projects (Foods of Norway SFI, BioMim, OXYTRAIN (Marie Curie ITN), ERA-IB OXYPOL, NorZymeD, H2020 Metafluidics, BIA 'XPress' (Vectron partner), and the FME Bio4Fuels. The partners are hence involved in many activities on the utilization of biomass for production of bioenergy (biogas, biofuels etc), biochemicals as well as other products. All partners are also part of the NorBioLab collaboration for infrastructure in biorefining. The partners have close interactions and collaborations with Norwegian industry active within the bioeconomy and focusing on utilization of Norwegian resources.

8. Relevance to industry and potential for innovations

The applicants' current research on enzyme technology, fermentation and biomass conversion is characterized by close contacts between research and industry, in a variety of projects. In fact, already today, Norwegian companies such as Borregaard and Cambi use knowledge generated by the applicants in their biomass processing operations. The present project will communicate to a large network of users via its own industrial reference group, by connecting to DL dissemination platforms, and by connecting to the Bio4Fuels FME. There will thus be very short lines between the research and its industrial implementation. Strong national competence in enzyme technology will be of great value to any existing or future industry focusing on improved resource utilization and increased value generation from biomass.

9. Responsible research and innovation, communication and dialogue with society

Research projects like OXYMOD hold the potential to generate a range of new applications and products. It is challenging to foresee the consequence and impact of the activity of the project, as it is not obvious what is at stake at an early phase. This is why we will integrate in the project the four dimensions of RRI as defined by Owen et al. (2013): anticipation, inclusion, reflexivity, responsiveness. The first stage will consists in mapping the research within the project, in order to anticipate further research trajectories, potential applications and societal ramifications. This assessment will be an integrated part of the consortium meetings, which will have an outreach to the general public and where we aim at gathering external feedback. The outcomes will be used to reflect on the project in a societal context where the ultimate goals and their perceived benefits and negative side-effects are also evaluated. The second stage will be to engage the project partners and stakeholders in a discussion on research objectives, the choices made during research and how the long term results of the project can contribute to the common good. As the range of potential applications will be open to alterations, the project objectives will be fine-tuned to ensure that end-products are developed with and for benefit of society. There are two main focus areas within RRI in the project that we will emphasize. The first focus area is that the overarching aim of the project is the industrial utilization of lignocellulose. Traditionally, this kind of activity has been evaluated from the economic point of view, where the most significant aspects are the materials costs, process yields, plant operating costs and the capital cost of the production facility. However, as a part of RRI in this project the evaluation will also include questions related to the project's contribution to a turn towards a biobased economy. What are the conditions for the industrial development of this technology? What regulatory frames are necessary to make it competitive, and what are the main obstacles? How may the utilization of forest resources be in conflict with other interests, such as preservation of green recreational areas and wildlife diversity. The second focus area is the societal perception of new bio-based product alternatives in the market. This will include environmental impact resulting from the longevity of the product and end-of-life and an overall balance of

positive and negative impacts resulting from the new products and processes. The ambition of the project is to become a driving force in sustainable innovation and value creation. The RRI component in the project is implemented by involvement of a dedicated PhD student and this student's supervisor, Bjørn K. Myskja (NTNU, see partner description) as active partners in the project.

10. Other relevant aspects

10.1. Environmental impact: Apart from producing big data-based advanced research in Norway, the societal relevance of this project concerns its impact on making our economy greener and more sustainable. The most direct benefit lies in development of more efficient methods for enzymatic depolymerization of biomass, a key step in biorefining and biomass valorization. The project will provide efficient and environmentally friendly approaches for utilization of side streams, hence contributing to approaching a more complete utilization of resources. By more efficient processing of the biomass, fewer resources are going to waste, and also less greenhouse gases might be generated in anaerobic processing of dumped side stream biomass. Hence, the proposed project will have significant positive environmental impacts. We do not envisage any negative environmental impact of the proposed project per se but we are, of course, fully aware of the potential negative impacts of wrong bioresource management.

10.2. Ethical perspectives: The project does not include any research on humans or animals, and there will be no release of genetically modified organisms. Hence, the project does not subject the environment or humans any risks. All screening activities and data collection will follow established ethical guidelines. The project will utilize land based biomass, however, primarily side streams normally being unutilized. Therefore the proposed project will contribute to a more complete utilization of resources, being in-line with ethical perspectives concerning resource exploitation. Also, all laboratories, equipment and protocols used are in compliance with national laws and regulations and adhere to good laboratory practice to protect personnel in the lab and to prevent spread of potential dangerous microbes.

10.3. Gender issues (Recruitment of women, gender balance and gender perspectives): The applicant's institutions are aware of the importance of gender balance in research projects, and the Research Council's effort for recruitment of women. The project partners all have a high share of female employees, and partners will continue actively to work towards gender balance when recruiting new employees, like the for the PhD and post doc positions at NTNU and NMBU.

REFERENCES

[1] D. Kracher, et al., *Science* **2016**, *352*, 1098.
[2] M. X. Guo, et al., *Renew Sust Energ Rev* **2015**, *42*, 712.
[3] aS. J. Horn, et al., *Biotechnol Biofuels* **2012**, *5*, 45; bG. Vaaje-Kolstad, et al., *Science* **2010**, *330*, 219.
[4] J. Zhang, et al., *Proc Natl Acad Sci U S A* **2016**, *113*, 10968.
[5] G. Muller, et al., *Biotechnol Biofuels* **2015**, *8*, 187.
[6] A. Saini, et al., *Enzyme Res* **2015**, *2015*, 279381.
[7] K. S. Johansen, *Biochem Soc Trans* **2016**, *44*, 143.
[8] D. Cannella, et al., *Nat Commun* **2016**, *7*, 11134.
[9] Y. Kojima, et al., *Appl Environ Microbiol* **2016**, AEM.01768.
[10] aG. Guerriero, et al., *Eng Life Sci* **2016**, *16*, 1; bA. I. Canas, S. Camarero, *Biotechnol Adv* **2010**, *28*, 694.
[11] O. V. Morozova, et al., *Appl Biochem Microbiol* **2007**, *43*, 523.
[12] C. Liers, et al., *Appl Microbiol Biotechnol* **2010**, *85*, 1869.
[13] D. Sirim, et al., *Database (Oxford)* **2011**, *2011*, bar006.
[14] P. K. Busk, L. Lange, *Appl Environ Microbiol* **2013**, *79*, 3380.
[15] P. K. Busk, L. Lange, *BMC Genomics* **2015**, *16*, 368.
[16] H. J. Atkinson, et al., *PLoS One* **2009**, *4*, e4345.
[17] J. A. Gerlt, et al., *Biochim Biophys Acta* **2015**, *1854*, 1019.
[18] V. V. Vu, et al., *Proc Natl Acad Sci U S A* **2014**, *111*, 13822.
[19] L. A. Kelley, et al., *Nat Protoc* **2015**, *10*, 845.
[20] M. Biasini, et al., *Nucleic Acids Res* **2014**, *42*, W252.
[21] A. Roy, et al., *Nat Protoc* **2010**, *5*, 725.
[22] M. Martin, et al., *Microbiol Res* **2016**, *186-187*, 52.
[23] F. Boehmwald, et al., in *Biotechnology of Extremophiles:* (Ed.: P. H. Rampelotto), Springer International Publishing, **2016**, pp. 321.
[24] aC. J. Duan, et al., *J Appl Microbiol* **2009**, *107*, 245; bG. Bastien, et al., *Biotechnol Biofuels* **2013**, *6*, 78.
[25] K. Wolstencroft, et al., *BMC Syst Biol* **2015**, *9*, 33.
[26] N. Roehner, et al., *ACS Synth Biol* **2016**, *5*, 498.
[27] S. Kim, et al., *Proc Natl Acad Sci U S A* **2014**, *111*, 149.
[28] M. R. Blomberg, et al., *Chem Rev* **2014**, *114*, 3601.
[29] C. H. Kjaergaard, et al., *J Am Chem Soc* **2015**, *137*, 8783.
[30] D. E. Heppner, et al., *J Am Chem Soc* **2014**, *136*, 17788.
[31] T. C. Tan, et al., *Nat Commun* **2015**, *6*, 7542.
[32] aF. L. Aachmann, et al., *Proc Natl Acad Sci U S A* **2012**, *109*, 18779; bA. S. Borisova, et al., *J Biol Chem* **2015**, *290*, 22955; cG. Courtade, et al., *Proc Natl Acad Sci U S A* **2016**, *113*, 5922.
[33] aU. T. Bornscheuer, et al., *Nature* **2012**, *485*, 185; bV. G. Eijsink, et al., *J Biotechnol* **2004**, *113*, 105; cV. G. Eijsink, et al., *Biomol Eng* **2005**, *22*, 21.
[34] aT. R. Manfred, *Angew Chem Int Ed* **2011**, *50*, 138; bM. T. Reetz, et al., *J Am Chem Soc* **2010**, *132*, 9144.
[35] aT. Pielhop, et al., *Green Chem.* **2016**, *18*, 5239; bV. Vivekanand, et al., *Bioresour Technol* **2013**, *127*, 343.
[36] J. M. Seminario, *Int J Quantum Chem* **1996**, *60*, 1271.