PROJECT FINAL REPORT

Grant Agreement number: 604530

Project acronym: CellulosomePlus

Project title: *Boosting Lignocellulose Biomass Deconstruction with Designer Cellulosomes for Industrial Applications*

Funding Scheme: Collaborative project

Date of latest version of Annex I against which the assessment will be made: 22/12/2017

Period covered: from November 2013 to October 2017

Name, title and organisation of the scientific representative of the project's coordinator[1]: Mariano Carrión Vázquez

Tel: +34 91 5854830

Fax: +34 91 5854754

E-mail: mcarrion@cajal.csic.es

Project website[2] address: www.cellulosomeplus.es

[1] Usually the contact person of the coordinator as specified in Art. 8.1. of the Grant Agreement.

[2] The home page of the website should contain the generic European flag and the FP7 logo which are available in electronic format at the Europa website (logo of the European flag: http://europa.eu/abc/symbols/emblem/index_en.htm logo of the 7th FP: http://ec.europa.eu/research/fp7/index_en.cfm?pg=logos). The area of activity of the project should also be mentioned.

TABLE OF CONTENT

1. Project Context and the Main Objectives

CONTEXT

For the healthy, wealthy and peaceful future of our planet, a reliable supply of sustainable energy is critical. Energy is the world's largest market, with a political and strategic impact that is unmatched by any other sector. Most countries, including European nations, are currently highly dependent on the finite and non-renewable resources of fossil fuels for their energy needs. This allows countries rich in such resources to become major players in world politics, frequently at the expense of countries that lack them.

Lignocellulose, also known as fiber, is formed mainly by the non-edible plant cell wall and constitutes the major source of carbon and energy in the biosphere. This type of biomass is the source for so-called second generation biofuels. However, lignocellulose is a recalcitrant carbohydrate, resistant to degradation. The transition to a more environmentally friendly economy has put the focus of research on enzymes capable of efficiently degrading this feedstock. Bacteria and fungi have evolved complex enzymatic systems enabling their growth on plant material rich in cellulose. Enzymatic conversion of crystalline polysaccharides (saccharification) is also crucial for an environmentally sustainable bioeconomy, but microorganisms that produce these enzymes typically require weeks, months or even years to decompose fallen logs or tilled corn stalks. This process is consistent with Nature's needs. However, for chemical or fuel production from these materials, industry requires affordable enzymatic systems that can do the job in a much shorter scale, of days or even hours.

In the market of second generation biofuels and in general in those of other lignocellulosic derivatives, there is an increasing need for more efficient and cost-effective saccharification processes. Most industrial efforts concentrate on the use of fungal cellulases and fungal-based enzyme cocktails that is becoming the standard product in the second-generation biofuel technology. Bacterial cellulosomes are still restricted to the laboratory level in which researchers concentrate on improving cellulosomes using biophysical and biochemical properties that can be measured and compared by traditional techniques. However, bulk techniques are very limited in their description of the system and there is a clear need to obtain more detailed descriptions from these nanoscale systems including properties than can only be measured at this level (*e.g.* mechanostability or exceedingly high affinity interactions). Thus, the remarkable mechanostability and the extraordinarily high affinity of some of the interactions of the cellulosomal constituents remain to be explained and exploited. In addition to this, current computational multiscale methods allow a first description of this system for verification and prediction purposes.

The problem in which this project focuses is the bottleneck of the second generation of biofuels: the efficient improvement of saccharification through the so-called Designer Cellulosomes (*DCs*).

We propose a hybrid strategy based primarily on rational design (after the knowledge generated by an initial physico-chemical characterization) and using biological evolution (a massive trial and error process filtered by natural selection through eons). However, this approach will be

complemented and improved by high throughput screening, which will explore a wider combinatorial space. Furthermore, a special emphasis is drawn into developing new technology and in combining it with standard methods towards the solution of the problem: efficient improvement of saccharification through *DCs*. In our opinion, this hybrid semi-empirical approach is one of the strengths of this project, which is also reinforced by the prediction and corroborative power of multi-scale modelling.

To reach the ambitious aforementioned challenge, the conrsortium has needed the collaborative work of the following multidisciplinary participant partners:

1 Agencia Estatal Consejo Superior de Investigaciones Científicas (CSIC). Spain. National Research Institute (coordinator)
2 Weizmann Institute (WEIZ). Israel. National Research Institute
3 Ludwig Maximilians Universität (LMU). Germany. Higher Education Institute
4 Polish Academy of Science (IFPAN). Poland. National Research Institute
5 Centre National de la Recherche Scientifique (CNRS). France. National Research Institute
6 University of Limerick (UL). Ireland. Higher Education Institute
7 Designer Energy (DesEn). Israel. SME.
8 Abengoa Bioenergía (ABNT). Spain. Multinational Industry (end user).
9 Biopolis (BIO). Spain. SME (end user)
10 CelDezyner (CD). Israel. SME.

OBJECTIVES

The cellulosome concept

One of the most efficient approaches for degradation of plant cell-wall polysaccharides in Nature is the integration of cellulases and associated enzymes into an extracellular multi-enzymatic complex named the cellulosome. Cellulosome architecture consists of a non-catalytic "scaffoldin" subunit and two complementary recognition modules termed "dockerin" and "cohesin", which together serve to integrate the various enzymes into the complex. The non-catalytic scaffoldin subunit comprises a string of repeating cohesin modules, plus a single Carbohydrate-binding module (CBM). Cellulosomal enzymes contain a catalytic module and an additional dockerin module, which binds tightly to the cohesins of the scaffoldin.

In designer cellulosomes (*DCs*), artificial scaffoldins serve as docking backbones that contain both a cellulose-specific CBM and multiple cohesin modules of divergent specificities. The CBM directs the enzymatic complex to the cellulosic substrate, while the cohesins allow the specific incorporation of matching dockerin-bearing enzymes. The controlled incorporation of cellulases into artificial designer cellulosomes was shown to induce enhanced synergism between cellulases via their targeting to the substrate by the family-3a scaffoldin-borne CBM module, or by the proximity of the cellulases in the complex.

Main Aim: Developing efficient Designer Cellulosomes specifically tailored to provide a high yield of fermentable sugars from industrially suitable biomass (specifically the Organic Fraction of the Municipal Solid Waste, OFMSW) for subsequent production of value-added products like biofuels at a low energy cost.

Main Objectives:

Objective 1: **Characterization of natural cellulosomes and the selected substrate**. CellulosomePlus will produce the basic components of natural cellulosomes as well as other lignocellulosic enzymes and characterize the hydrolysis (by the cellulosome, its components and by the *DCs* to be developed in this project) of the substrate of interest: OFMSW. We will focus on studying the relevant catalytic activities, physicochemical properties (thermodynamic, kinetic and mechanical stabilities), atomic and supramolecular structure, and interactions of the various cellulosomal components. An important (but not indispensable for the main objective of the project) step is to develop a standard enzymatic assay for performance comparisons. The knowledge obtained will provide a deeper understanding of the architecture, nanomechanics and catalytic properties of cellulosomes, and the logic behind their construction.

Objective 2: **Modelling the cellulosome for *in silico* knowledge integration**. Multi-scale modelling (from atomic to supramolecular levels) will provide crucial support for the synthesis, self-assembly and characterization tasks, supplying detailed structural and energetic information that will aid in the design and interpretation of the experiments. Model outputs will provide new information on the conformation-driven and dynamic catalytic properties, together with mechanical stabilities, which will provide new ideas for experiments, in a synergistic feedback loop towards the production of optimized *DCs*.

Objective 3: **Rational design and mass production of *DCs*.** The integration of the acquired knowledge from 1 and 2 into *DCs* and their subsequent screening will provide us with a platform to test the goal of constructing final *DCs* (carrying both cellulosomal and non-cellulosomal components) optimized for the degradation of the selected industrial substrate, and validated at the laboratory-scale. Our final goal is to obtain higher activity as assessed by comparing the rates of hydrolysis and final sugar yield to those of commercial enzymatic cocktails based on grams of enzymes/grams of biomass.

Specific Objectives:

In order to achieve the main goal of the project: generate active Designer Cellulosomes (*DCs*) to efficiently degrade residual biomass into fermentable sugars, a multidisciplinary approach was followed to accomplish it. The close interaction between different research groups from diverse fields and industrial partners ensured rapid development and exploitation of the new technology.

To ensure the main goal, a thorough understanding of the molecular basis of the cellulosome function is needed. Therefore, the work plan comprised eight Research, Technological and Development (RTD) actions corresponding to the work packages 2 to 9. Five of these (WP2-6) constitute the basic experimental research to produce and characterize the key elements of the system (enzymes, scaffoldins and substrate). Computational predictions developed in WP7 provided complementary knowledge of the experimental results. Aforementioned information was integrated to obtain simple systems and test them at the laboratory-scale level, improving their activity (WP8). Finally, all the knowledge generated was unified in WP9 to develop mass production of optimized *DCs* for sustainable energy production.

WorkPlan

WP1: Project management (CSIC): its objective is to coordinate the scientific and logistic aspects of the project.

WP2: Production of substrates and enzymes and assembly of *DCs* (WEIZ, CSIC, ABNT, CNRS, DesEn): The objectives of this WP were to discover, design, optimize, produce and assemble the components necessary for the different studies to be carried out in the other WPs of the project.

WP3: Physicochemical characterization and component improvement (CSIC, WEIZ): The objectives of this WP were to characterize and improve the thermal and pH stability of the selected components and study their mechanical properties, such as mechanical stability by using SMFS.

WP4: Structural characterization (CNRS, CSIC): The objectives of this WP were to characterize the atomic and supramolecular structure of selected components by X-ray diffraction and SAXS.

WP5: Monitoring the enzymatic activity (DesEn, WEIZ, LMU): The objectives of this WP were to develop a standardized activity assay that can be applied to monitor catalytic activity during the project.

WP6: Characterization of interactions (LMU, CSIC, WEIZ): The objectives of this WP were to study the properties of the multiple interacting cellulosome parts using traditional, single-molecule, high-throughput molecular force balances and computational techniques.

WP7: Multi-scale modelling (IFPAN, UL, CNRS, CSIC): The objectives of this WP were to generate atomistic to supramolecular-scale models of the cellulosome structure and properties to help predicting, characterizing, interpreting and manipulating structural, mechanical and catalytic properties across multiple scales of both single protein architectures and multi-protein assemblies.

WP8: Laboratory-scale level integration of *DCs* (WEIZ, CSIC, LMU, CNRS, IFPAN, UL, DesEn, ABNT, BIO, CD): The objectives of this WP were to integrate all the acquired knowledge and the expertise of the industrial partners into optimized *DCs* as relatively simple systems to efficiently degrade lignocellulose at the laboratory-scale level, and to improve their activity.

WP9: Preindustrial-scaling up (ABNT, BIO): The objectives of this WP were to produce and test final *DCs* to efficiently degrade lignocellulose at the preindustrial-scale level.

WP10: Exploitation and Dissemination (CSIC, WEIZ, LMU, IFPAN, UL, CNRS, ABNT, DesEn, BIO, CD). It objective is to disseminate and exploit the results obtained in CellulosomePlus.

2. Description of the main Science & Technical results/foregrounds

The results obtained in each RTD WPs during the project are presented in the following pages:

2.1. Description of the main RTD results in WP2: Production of substrate and enzymes and assembly of *DCs*

The main goals of this WP were to discover, design, optimized, produce and assemble the cellulosomal components necessary for the subsequent studies to be carried out in the other WPs.

Bioinformatics search for novel cellulosome components

In order to locate previously unidentified cellulolytic enzyme candidates, CBMs and cohesin-dockerin pairs, several genomes of cellulosome-producing bacteria were examined by bioinformatics-based approaches studying their cellulosomal system: regarding their numbers, types and modular content of scaffoldin proteins, cohesin modules, dockerin modules, dockerin-containing proteins, glycoside hydrolases, multifunctional proteins and CBMs.

A multi-dimensional perspective on the cellulolytic potential of two fibrolytic ruminococci: *Ruminococcus flavefaciens* and *Ruminococcus albus* were studied. The extreme diversity of enzymes and structural scaffoldins was demonstrated within and also between these species. Additional studies in other ruminocci such as *Ruminococcus champanellensis* and *Ruminococcus bromii* revealed that both appear to be cellulosome-producing bacteria, a new set of cohesin-dockerin interactions were discovered in both species.

On the other hand, bioinformatic analysis of the newly sequenced genome of *Clostridium clariflavum* revealed the multiplicity and complexity of its cellulosomal components, enabled a diverse architectural assembly. The network of cohesin-dockerin pairs provides a thermophilic alternative to those of *C. thermocellum* and a basis for subsequent utilization of the *C. clariflavum* cellulosomal system for biotechnological application. Other clostridia species examined in terms of its cellulosome diversity by *in silico* methods were *Clostridium papyrosolvens* and *Clostridium alkalicellulosi,* as well as the newly sequenced *Pseudobacteroides cellulosolvens* phylogenetically related to the clostridial assemblage.

The nature of the interaction among the different cohesin-dockerin pairs allowed to elucidate the organization of the cellulosomal components thereby revealing its architecture. This knowledge procured us with new cohesin and dockerin specificities, providing with new insight into the similarities and diversities evident in cellulosome-producing bacteria that we can then use as building blocks for production of efficient designer cellulosomes. As a result of the search, an in-house database of all the identified cellulosomal components was established.

Producing and assembly of cellulosomal components

The development of new procedures for protein expression based on fusion protein cassettes: where cohesin modules were fused to a CBM while dockerin modules were fused to a xylanase, allowed simple purification procedures, increased the expression level and dramatically improved solubility and stability of the proteins, permitted a rapid screening for matching interactions.

Powerful analytical methodologies have been developed for the routine characterization of the *DCs*, in order to determine functioning of the cohesins and dockerins, the extent of incorporation of the enzyme components, and the purity of the final DC product. To precisely detect interaction of various cohesin and dockerin modules, high throughput methods such as the matching fusion protein system and affinity-based ELISA assay were employed to examine the binding ability of an individual cohesin/dockerin towards a variety of its counterparts.

During this project around a hundred chimeric scaffoldins and enzymes were produced, containing divergent dockerins in order to allow self-assembly together with appropriate scaffoldins for incorporation into designer cellulosomes. Among them, wild-type cellulases, chimaeric cellulases, cohesin and dockerin modules were taken from both thermophilic bacteria (*C. thermocellum, Thermobifida fusca, C. clariflavum* and *Archaeoglobus fulgidus*) and mesophilic bacteria (*R. flavefaciens, R. champanellensis, R. bromii, A. cellulolyticus, P. cellulosolvens C. papyrosolvens* and *C. alkalicellulosi)* and produced for incorporation into designer cellulosome assemblies, matching designer scaffoldins and adaptor scaffoldins.

A growing library of scaffoldins was formed for *DCs* production. These included scaffoldins with a CBM and 1, 3, 4 or 6 cohesins, each cohesin of which was derived from a different cellulosome-producing species with defined specificity. In addition, tetravalent scaffoldins, with a single specificity with and without a CBM, were produced. Finally, in order to create elaborate forms of designer cellulosomes, adaptor scaffoldins that allowed both incorporation of various enzymes and interaction with another chimeric scaffoldin were produced. Using this strategy, the hexavalent scaffoldin can integrate five recombinant enzymes and the adaptor scaffoldin can incorporate three additional enzymes.

A Central Protein facility (CPF) was stablished as service-provider for the expression, purification, characterization, storing and shipping of common enzymes and structural proteins for participants in the project. The purpose of this facility were to ensure homogeneity and fresh proteins, custom production of troublesome proteins and validation of protein activity. The CPF expressed and characterized a library of 30 different known and unknown *C. thermocellum* enzymes. Moreover, the unique High Density Cell Culture (HDCC) approach for production of recombinant proteins resulted in mass quantity production of six different enzymes and structural proteins in quantities ranging 0.2-2 grames.

2.2. Description of the main RTD results in WP3: Physicochemical characterization and component improvement

The main goals of this WP were to characterize and improve the thermal stability of the selected components and study their mechanical properties by AFM-SMFS.

Thermophilic designer cellulosome

Thermostable cellulases offer many advantages in the bioconversion process, which include increase in specific activity, higher levels of stability, inhibition of microbial growth, rise in mass transfer rate due to lower fluid viscosity, and greater flexibility in the bioprocess.

For the construction of cellulase libraries two strategies were followed: i) error prone PCR was used to generate random amino acid substitutions, ii) a semi-rational "consensus approach" based on the substitution of specific amino acids in a particular protein with the most prevalent amino acid present at these positions among the homologous family members.

The endoglucanase Cel8A from *C. thermocellum* was chosen as a model for thermostability enhancement. Approximately 9000 colonies were screened for enzyme activity where 2800 colonies were selected for subsequent screening. Four clones demonstrated significant thermostability compared to wild-type Cel8A enzyme. Four amino acid substitutions were mapped to the thermostable clone where a single mutation was sufficient to produce the thermostable variant.

An analogous procedure was followed to increase the thermostability of the exoglucanase Cel48S from *C. thermocellum* obtained several clones with better thermostability.

In the case of the enzyme β-glucosidase A (BglA) from *C. thermocellum* a random evolution approach was conducted, about 8000 colonies were screened and 40 mutants with increased thermoresistance were selected revealing a thermostable clone, with 50% residual activity after heat-shocked at 68 °C while the wild type BglA showed a residual activity of 9%. This new thermostable clone revealed minor differences in its catalytic efficiency and its maximum velocity.

Another approach to obtain high-temperature active enzymes was to develop an algorithm that combined a structural design based on a physico-chemical energy function with quantitative sequence data from homologous proteins. Following this strategy, the catalytic domain of Cel5A from the cellulolytic mesophilic bacterium *C. papyrosolvens* was redesigned obtaining thermostable mutants that maintained a residual activity of 46% in comparison to the completely inactive wild-type enzyme after 60 minutes of heat challenge.

By incorporating highly thermostable enzymes into our designer cellulosome structures, we intend to address industrial and biotechnological interests. Indeed, highly thermostable enzymes are more appropriate to harsh process conditions, which may reduce the cost of biodegradation of cellulosic waste.

In a first attempt to fabricate a tetravalent designer cellulosome composed by four enzymes thermostable engineered: endoglucanase Cel8A*, exoglucanase Cel48S*, β-glucosidase (BglA*) (all from *C. thermocellum*) together with the engineered non-cellulosomal endoglucanase Cel5A* from *T. fusca*. A novel scaffoldin (Scaf-BTFA), composed mainly of thermophilic modules was designed including a thermostable CBM. In order to integrate the catalytic modules into designer cellulosomes in a control and precise manner, dockerin specificity were switched, creating chimeric cellulases. Moreover, non-cellulosomal enzymes were converted to the cellulosomal mode adding a dockerin in its C- or N-terminus.

A comparison between improved designer cellulosomes to analogous free enzyme system, in terms of activity was performed. At medium temperatures (50 to 60°C), the thermostable mutants were less active than the wild-type enzymes either in the free state or within the designer cellulosome complexes even during a long incubation period (96 h). At elevated temperatures (65 to 75°C) the thermostable mutants either in free state or in designer cellulosomes complexes were more active than their wild-type counterparts. Nevertheless, at those temperatures the designer cellulosomes had no advantage over the free enzymes certainly due to the mesophilic nature of some modules of the chimeric scaffoldin.
Comparing natural cellulosomes to designer cellulosomes and to commercial enzyme preparations it was obtained that natural cellulosomes were 1.6- to 2.6-fold more efficient than

the designer cellulosomes (at 50 and 60°C) respectively while the commercial preparation was 13.5 to 16.5 times more efficient than the designer cellulosomes (at 50 and 60°C), respectively.

A novel scaffoldin (Scaf-GTV), composed entirely of thermophilic modules and consistent with the complementary dockerins of the respective enzymes, was designed. The new scaffoldin allowed the selective integration of the 3 engineered thermostable cellulases: endocelulase (Cel8A*), an exocellulase (Cel48S*) and betaglucosidase (BglA*).

The thermostable designer cellulosomes were also stable for 4 h at 65°C and 3 h at 70°C, as opposed to the conventional designer cellulosomes comprising the corresponding wild-type enzymes. However, the three thermostable enzymes in complex in a chimeric scaffoldin of mainly mesophilic origin were stable only for 48 h at 50°C. Following stability studies of the different complexes, enzymatic activity of both thermostable and conventional (containing wild-type enzymes) designer cellulosomes was examined, and the results were compared to those of the mixtures of the free enzymes (without a scaffoldin) at various times and temperatures. Both conventional designer cellulosomes and thermostable designer cellulosomes were consistently more active than the mixtures of their corresponding free enzymes. In addition, mixtures of the free thermostable enzymes continuously showed higher levels of cellulose hydrolysis than those of the wild-type free enzymes. Moreover, we observed that in each case the thermostable designer cellulosomes was advantageous (1.7 enhancement) in cellulose hydrolysis over that of conventional designer cellulosomes and free enzyme mixtures. For designer cellulosome formats, the use of the thermostabilized scaffoldin proved critical for enhanced enzymatic performance under conditions of high temperatures.

In nature, cellulosomes are found mainly in mesophilic bacteria, as well as in a few thermophilic bacteria possessing an optimal growth temperature not higher than 60˚C, several thermophilic designer cellulosome have been created to date but no designer cellulosome has been reported in the literature to be stable and active at more than 60˚C for 24 h. In an effort to fabricate hyperthermostable enzymatic complexes, designer cellulolosomes were composed by a thermophilic scaffoldin and chimaeric enzymes from the highly thermophilic bacteria *Caldicellulosiruptor bescii*, able to growing at temperatures up to 90°C being the most thermophilic bacterium capable of growing on crystalline cellulose. A library of *C. bescii* cellulases were created, expressed and analyzed its activity against different substrates. Later on, the free enzymes were converted into the cellulosomal mode by linking them to thermostable dockerins of different specificities. Several scaffoldins (monovalent, bivalent and trivalent) were designed and checked its thermostability and activity. The results demonstrated that these complexes were highly thermostable, focusing the efforts on studied the influence of enzymes composition and position.

Examination of cohesin-dockerin stability and specificity

Microarray techniques allow a high-throughput approach for examining cohesin-dockerin interactions. New interactions were discovered using a cellulose-based microarray method, which allowed us to examine every dockerin separately against a large number of cohesins in one reaction, followed by an ELISA method for subsequent confirmation. This approach was used for determining large numbers of cohesins and dockerins interactions. Both modules were selected from the newly described human gut cellulosome-producing bacteria: *R. champanellensis* and *R. bromii*, while additional cohesins were taken from other cellulolytic species to explore the

possibility of cross-species interactions and examined the specificity of cohesin-dockerin interaction. A number of 20 cohesin and 24 dockerin modules from *R. champanellensis* were cloned and overexpressed testing 480 intra-species and 374 inter-species interactions founding 64 interactions positive. *R. champanellensis* dockerins showed three affinity profiles which led us to propose two types of cellulosome architectures, cell-bound and cell-free.

The same approach was followed to reveal the cellulosomal architecture based on the detection of cohesin-dockerin interactions in the alkaliphilic cellulolytic bacterium *C. alkalicellulosi*. With this respect, 21 cohesin and 8 dockerin modules were selected for production and tested for binding interaction, a total of 168 cohesin-dockerin interactions tests were performed. An extensive scaffoldins repertoire was found, providing a variety of sophisticated cellulosomal system that could lead to the development of alkali.stable designer cellulosome with potential applications in different biotechnological fields.

Finally, a new method of covalently locking cohesin-dockerin interactions was developed by integrating the chemistry of *SpyCatcher/SpyTag* approach to target and secure the interaction. The resultant cohesin-conjugated dockerin was resistant to high temperatures, SDS, and urea while high affinity and specificity of the interacting modular components were maintained.

Mechanical hypothesis

Cellulosome is an adhesion system, attached to the cell and able to interact with the substrate (cellulose) thanks to CBM module, the relative movement between substrate and the cell could cause mechanical stress into the portion of scaffoldin located between the two anchoring points. This raises the possibility that the mechanical resistance of cohesin modules in scaffoldins are under selective pressure and therefore that the mechanical properties of these protein structures are key for their activity. If the mechanical stability of the cohesin modules were not enough, these modules would unfold and the enzymes attached to them could be released and they would loss its synergistic action. To avoid this, cohesin modules located in the connecting region between the cell and the substrate should have high mechanostability and resist mechanical stress. Therefore, mechanical properties of the cellulosome should be crucial for its activity.

Previous work, revealed that cohesin modules from the connecting region (attached to the cell and to the substrate) showed extremely high mechanical stability (higher than all other known folded proteins) while the cohesin hanging outside the connecting region showed much lower mechanical stability. This observation agrees with the working hypothesis since selective pressure on mechanical stability is only expected on those modules in the connecting region. High mechanical stability reduces the probability of cohesin mechanical unfolding which is necessary to maintain dockerin interaction surface and therefore keep enzymes bound to the scaffoldin for synergistic high activity of the cellulosome complex.

Nanomechanical studies (AFM-SMSF) require the preparation of fusion polyproteins to assess that the obtained data truly comes from the desired single molecule and not from contaminating molecules or nonspecific interactions. To achieve this, the well stablished heteromeric polyprotein approach was used generating a hybrid fusion protein consisting of the molecule under study flanked by multiple repetitions of a well-known single molecule reporter and cloned them into a specially designed vector that was successfully been already applied.

With the purpose of testing the mechanical hypothesis based on indirect evidences, the mechanical stability of cohesins from the external and the connecting regions of two primary scaffoldins (those that directly bind enzymes) were studied. Cohesins from two different cellulosomes (CipA from *C. thermocellum* and ScaA from *A. cellulolyticus*) were selected obtaining that the mechanical stability of connecting cohesins in both cases was higher despite sequence identity. Regarding the mechanical unfolding parameters, it is noteworthy to mention that the actual value of this unfolding force depends on the speed at which force is applied. The AFM allows to study the high force regime, although the loading rate that applies to cellulosomes in nature may not be that high. Thus, the mechanical unfolding kinetic parameters were studied and observed at low loading rates (as it could be expected in the natural cellulosomes) and it was found that cohesins form the connecting region still show higher mechanical stability, enable to maintain dockerin interaction surface and keeping enzymes bound to the scaffoldin for synergistic high activity.

Regarding the mechanical stability of cohesins form secondary scaffoldins, cohesin from ScaB and ScaC from *A. cellulolyticus* located in hanging and connecting regions respectively were tested, finding mechanical stability values comparable to those of the connecting region or the hanging region from primary scaffoldins.

To gather further indirect evidence tht the mechanical hypothesis holds, the role of other elements present in natural cellulosomes on the mechanical stability of cohesin modules was studied in order to further characterize the nanomechanical behavior of scaffoldins and test the working hypothesis. We focused on the effects of linker sequences joining cohesins of scaffoldin, the dockerin binding and the presence of multiple cohesin repetitions. The results indicated that these additional elements do not alter the mechanical stability of cohesin modules. Taken together, this indicates that cohesin mechanical stability itself is a main determinant of mechanical properties of a cellulosome.

An alternative mechanism to maintain cellulosome integrity could be the ability of rapid refolding of cohesins after mechanical stress. Therefore, refolding analysis of two different cohesin modules after mechanical unfolding was studied. According to our results, only a fraction of the mechanically unfolded cohesins are expected to refold and it appeared to be independent of the position of the cohesin in the scaffoldin. Considering that cellulose hydrolysis takes long times, this would result in decreased number of functional cellulosomes upon time. Hence, considering the poor refolding capabilities of cohesins, the high mechanical stability of (connecting) cohesins may have been selected to reduce the probability of unfolding, which would not only result in enzyme release but also in damaged scaffoldins.

In order to obtain direct evidence of the mechanical hypothesis of the cellulosome, four different mini-cellulosomes were designed containing three basic elements: one CBM, a single cohesin module and a dockerin module obtained from the primary scaffoldin CipA from *C. thermocellum*. Specific cohesins were selected according to their mechanical stability previously studied by AFM-SMFS and the position of the cohesin varied between or outside the other two elements to study its possible implications in the mini-scaffoldin mechanostability. Mini-cellulosomes were attached to polystyrene beads functionalized with a complementary cohesion type II module, mimicking the presence of a bacterial cell and microcrystalline cellulose was used as substrate. After expressing the mini-cellulosomes recombinantly, corroborated that all their components were functional and checked the proper assembly of the mini-cellulosome complexes on the functionalized microparticles, their activity was tested. The mechanical stress was induced by

high agitation condition. Under these conditions, only the ratio of activity of the mini-celullosome carrying a cohesin of low mechanical stability in the connecting region was lower than the others, indicating that the presence of a cohesin of low mechanical stability in the connecting region interferes with the activity. Thus, placing a low-mechanostability cohesin artificially in a connecting region lowers the enzymatic activity. This effect only appeared with high agitation, under soft conditions this effect was vanished. From the results above we can conclude that there is a relation between the activity of the cellulosome and the mechanical stability of the cohesins located in the connecting region of the scaffoldin. This enables us to provide a few guidelines for the design of mini-cellulosomes according to their mechanical properties: for cellulosomes with two attachment points, the cohesins between these two anchoring points should be of high mechanical stability in order to avoid a loss of activity; for other cases the mechanical stability of cohesins is not critical. These design rules were applied to the design of *DCs*. From our results, we expect that the mechanical stability of cohesins would be a critical parameter to be considered only when cell-bound cellulosomes are used but not for free cellulosomes like in this project. Still we aim to design sturdy cellulosmes that would hold activity when attached to cells in future projects.

2.3. Description of the main RTD results in WP4: Structural characterization

The main goals of this WP were to characterize the atomic and supramolecular structure of candidate components by X-ray diffraction, SAXS (Small-angle X-ray scattering) and RMN.

Current knowledge of the inner workings of the cellulosome is still very limited. The past ten years have seen the development of a ‘dissect and build’ method, combining SAXS to X-ray crystallography, to get a three dimensional vision of cellulosomal complexes. Those techniques allow to understand in atomic detail the importance of these structural arrangements, flexibility and the presence or absence of particular modules/enzymes in these nano-machines. The challenge lie on that cellulosomes are composed by different modules and enzymes, and those enzymes are often multimodular, that contain besides the catalytic domain also CBMs, dockerins and other additional elements. To be able to understand the overall 3D structure of these macromolecular complexes, one needs to understand ‘piece’ by ‘piece’ the structural arrangement in solution, the final goal were to complete the assemblage of a cellulosome that contains the full length scaffoldin associated to more than three enzymes.

Extensive literature survey and screened the 'Protein Data Bank' (PDB) for crystallographic or NMR coordinates were performed in order to develop an interactive table connecting all the structural information available of cellulosomal components. Based on these data and the fact that enzymes from families GH5 and GH9 are very often found as major components in cellulosomal complexes, two different enzymes from GH5 family and two GH9 members belonging to different sub-families were chosen to crystallized. According to data mining and biochemical experiments, these families cover a large panel of substrate specificities, thus we tried to pinpoint specific sequence-patterns that unambiguously define a given substrate specificity of this diverse class of enzymes.

Crystal structures of the first endo-gluco-mannan-β1,4-glucosidase (GH5 from *Zobellia galactanivorans*) and 2 multimodular enzymes from family GH9, namely Cel9E, and Cel9X from *Clostridium cellulolyticum* were solved. These new crystal structures allowed the determination and description of very fine details that are responsible for the variations of substrate specificity within glycosyl hydrolase families and/or sub-families. Traditionally, GH9 family known to be

strict cellulases but identification of Cel9X (actived on xyloglucanan) which is not active on cellulose, opened new panels of substrate specificity for this family. GH5 beta-glucanase was showed active on mixed-linked glucan able to cleave both beta-1,3 and beta-1,4 linkages, however with a net preference for beta-1,3 linkages. Mutagenesis analysis were performed on this, in the aim of determining the key residues that define the substrate specificity. Five mutant proteins were produced and analysed for their enzymatic activities, in conjunction with the 3D crystal structure in complex with a substrate molecule.

Moreover, structures solved were part of the basis for constructing the molecular models of cellulosomes that were studied by SAXS measurements. SAXS studies result in molecular envelopes that allow the building of structural models, within the boundaries of the experimental shapes, thus providing a powerful basis to build up subsequently the large molecular assemblies of cellulosomes.

Cellulosomal complexes, needed a lot of preparation, thus an 'dissect and build' strategy was adopted: in the first place uncomplexed objects and smaller units were studied by SAXS while larger complexes were analyzed in second place. Small angle X-ray scattering curves have provided scattering curves for 4 individual enzymes (one GH8, one GH9, one GH5 and one GH48) and 4 different scaffoldins (CipA from *C. thermocellum*, ScaA and ScaH from *R. champanellensis*) and a designer cellulosome (ScaF-20L). In collaboration with IFPAN, a molecular modelling was developed obtaining several SAXS data of complexes combining enzymes and scaffoldins and a first coarse grain model, based on experimental SAXS scattering curves, for the full length cellulosomal multimodular enzyme Xylanase Z from *C. thermocellum* (XynZ) were analysed and published.

The last technique applied to obtained structural data of cellulosomal components were nuclear magnetic resonance, NMR methodology were applied to study with atomic detail the structure of the region of *C. thermocellum* CipA scaffoldin of higher identity, which comprises cohesin 4 and 5. The cohesin module and the linker between cohesins 4 and 5 isolated were studied. Structural information for the single cohesin shows good agreement with the X-ray structure. Regarding the linker sequence, its NMR analysis revealed that the linker segment shows a significant tendency to adopt extended conformations, it is improbable that it would interact significantly with the cohesin modules and will rather tend to space out the cohesin modules, so as to minimize interactions between them. The other *C. thermocellum* linker segments have identical or similar sequences and will therefore likely adopt similar extended conformers which act to separate the cohesin modules. Finally, preliminary results showed that the construction with two cohesins behave as if the 2 modules were independent.

2.4. Description of the main RTD results in WP5: Monitoring the enzymatic activity

The main goals of this WP were to develop a standardized activity assay to monitored catalytic activity.

Standardized assays for cellulolytic activities

A detailed protocol for low throughput assay for cellulosomal and non-cellulosomal enzymes was established for the first time. The protocol standardized an assay for brief estimation of

endoglucanase, exoglucanase and xylanase activity and it was based on three different substrates: Carboxymethyl cellulose (CMC), amorphous cellulose (PASC) and xylan. Moreover, the protocol allowed to distinguish between endo-acting or exo-acting cellulases in a simple and precise manner. In addition, another low-throughput protocol was provided for the detection of xylanases derived from cellulosome-producing bacteria.
The procedure included predetermined fixed conditions (buffers, temperature and time) which were determined to fit the linear/semi-linear range of the majority of *C. thermocellum* enzymes, thus enabling quantification and comparison of its enzymes. The protocol included an optional internal validation of the semi-linear behavior of the enzymes. To assay enzymes from other cellulolytic organisms an internal validation should be included and assay conditions (temperature and enzyme dosage) should be fitted in order to make sure that the assay is not done in under saturation or undetectable conditions.

Model substrates mentioned above, are highly homogenous substrates, containing only cellulose (or carboxymethyl cellulose) fibers. As opposed to those substrates, lignocellulosic feedstocks have much more complex and heterogeneous composition. Besides cellulose, lignocellulosic feedstocks contain hemicelluloses, lignin and other poly-saccharide such as pectin. Those polymers are integrated into a rigid cell-wall structure, restricting the accessibility of enzymes to the cellulose chains and hindering their degradation. In addition, the complex composition of the lignocellulosic feedstock, as well as the pretreatment step, required for further saccharification process, can increase the formation of different inhibitors. Therefore, to achieve efficient conversion of this complex structure, a synergistic activity of several enzymes, such as cellulases, hemicellulases, esterases, pectinases are required.

Due to the fundamental differences between the model substrates and the lignocellulosic feedstock, a new assay for analyzing enzyme activity on pretreated lignocellulosic substrate provided by Abengoa (ABNT) was developed. The novel protocol was based on the previous mentioned standardized assay. Four different enzymatic preparations were analyzed for their activity: (A) single, purified, recombinant cellulases, (B) mixture of cellulosomal enzymes from *C. thermocellum* (C) native cellulosome from *C. thermocellum* and (D) commercially available cellulosic cocktail. Six different biomasses were screened and the preferred biomass for each application was selected. An optimization step for the enzymatic hydrolysis was performed in order to reach the optimal assay conditions, and the developed assay fits the measurements of all the enzyme combinations. The results reported here indicate that only glucose concentration (and not xylose or cellobiose concentration) should be used in order to analyze the activity. In addition, it was found that the hydrolysis of the lignocellulosic biomass is less efficient while using a single purified enzyme or by un-optimized free enzyme cocktail, compared to the hydrolysis by the native cellulosome or the commercial fungal-based enzyme cocktail. Thus, enzyme combinations of the latter help the cocktail to tackle the difficult-to-degrade plant cell wall, to remove feedback inhibition, and to benefit from the synergistic effect.

Carbohydrate-binding modules are crucial for the activity of cellulases: due to its ability of increase effective enzyme concentrations on the polysaccharide surface, target the catalytic module to the substrate and disrupt the polysaccharide structure. Therefore, a set of robust assays for measurement the binding of CBMs to polysaccharides were established. The first assay was based on titration of enzyme of specific amount of cellulose, which provide a qualitative data on binding, in other words, to test whether a given protein can bind a given insoluble polysaccharide.

By this method binding capacities to several polysaccharides of the family-3a CBM members from *C. thermocellum* were measured.

On the other hand, quantitative assays were established and compared. The first one was based on the qualitative assay protocol (also known as batch assay) while the second assay used Isothermal Titration Calorimetry (ITC). The latter could provide a very accurate data on the affinity of the enzyme to the biomass. Using both methods, binding parameters (maximum absorption and dissociation constant) of three CBM3a-containing proteins in relation to three cellulose-containing substrates (microcrystalline cellulose, phosphoric acid swollen cellulose, and the biomass supplied by ABNT) were measured. Both quantitative assays were compared, revealing similar binding parameters (data ranged in the same order of magnitude). However, ITC-based method was more cumbersome and in some case resulted irrelevant, therefore authors preferred the batch assay.

In order to establish a medium-to-high throughput assay, the previously developed assays for analysis of cellulolytic activities using insoluble cellulose and lignocellulosic biomass was converted and adjusted to be compatible with liquid handling robotics. The assay involved a saccharification step, conducted in a 96 deep-well plate, following by a sugar concentration measuring step conducted in standard 96-well plates**.** Due to non-consistency with ABNT pretreated biomass, alternatively, pretreated lignocellulosic biomass (from corn stover) which is compatible with the instrumentation, was selected. Finally, an established protocol was developed using a liquid handling robotics-based assay.

A novel AFM-based medium-to high-throughput screening process for the identification of *DCs* with superior catalytic activity were developed. For this purpose, an assembly of large numbers of cellulolytic enzymes complexes on a cellulose chip was required. Cellulolytic activity of enzymes on pretreated biomass substrates was measured on a polymerization-based assay, which relies on monitoring the auto-fluorescence of cellulose and measuring the attenuation of this fluorescent signal as a hydrogel consisting of poly ethylene glycol polymerizes on top of the cellulose in response to glucose produced during saccharification. This one-pot method is label-free, rapid, highly sensitive, and requires only a single pipetting step. This approach was corroborated using model fungal enzymes and recombinant bacterial mini-cellulosomes from *C. thermocellum*. The new strategy demonstrated its ability to differentiate enzyme performance based on differences in thermostability, CBM targeting, and endo/exoglucanase synergy. Based on consortium discussions, it was clear that it was not required testing cellulosomal components in a chip format, because the setup was incompatible with bulk biochemical experiments. Although, on-chip expression provided libraries of cellulosomal proteins allowing screening of cellulolytic activity, though, the requirement of using soluble substrates, incompatible with testing solid biomass, made this platform not suitable for this purpose. Nevertheless, a large-scale screen of cohesin-dockerin binding interactions, for example, would be realistic with this platform, and could provide insight into cellulosomal network topology.

2.5. Description of the main RTD results in WP6: Characterization of interactions

The main goals of this WP were studying the properties of the multiple interacting cellulosome parts using traditional, single-molecule, high-throughput molecular force balances and computational techniques.

Internally controlled method for measuring protein-protein interactions

The study of protein-protein interactions by AFM-SMFS (AFM based single molecule force spectroscopy) has allowed for the estimation of kinetic constants, the mapping of the energy landscape, and the identification of hidden kinetic barriers. However, classical studies on protein-protein interactions by AFM-SMFS require complicated functionalization protocols that in most cases make difficult to ensure single-molecule conditions, and it is also hard to differentiate the bond breakage event from unspecific tip-surface adhesion, or the loss of adhesion of one of the proteins from their respective surface. Thus, an auto-controlled approach for directly identify protein-protein interactions by AFM-SMFS was developed. The strategy was based in the inclusion of single-molecule controls in the proteins of interest, and in the use of a reporter of the interaction ensuring the direct identification of the force peak originated by the rupture of the interaction. This general strategy that does not depend on the nature or the characteristics of the proteins of interest. The approach consisted on fusing 3 repetitions of a single molecule marker to each interaction protein partner (cohesin or dockerin) and a parallel second marker (an elastomeric protein) covalently bound by disulfide bonds to the other two constructs.

The viability of the strategy was first evaluated by the use of an empty model (without interaction proteins) and later assessed by implementing it in a cohesin/dockerin pair from *C. thermocellum* (cohesin 7, from CipA scaffoldin and dockerin from cellulase Cel8A). The mechanical stability of the cohesin-dockerin pair was successfully measured and the AFM recordings presented all the single molecule controls included in the design. When this approach was extrapolated to study other cohesin-dockerin pairs interactions, despite the amount of approaches tested (such as new protein immobilization protocols) the strategy seemed not to be easily to generalize to other cohesion-dockerin interactions. The existence of an alternative protocol created by LMU that was nicely working prompted us leave this internally controlled approach just as a proof of concept.

SMFS measures of cellulosomal components interactions

Previous work showed that cohesin-dockerin interactions, apart from exhibiting high binding specificity and strong affinity, also exhibit high mechanical stability. The lack of fundamental understanding of the mechanism behind dockerin mechanical robustness would undoubtedly hamper development of designer cellulosomes for bioconversion processing. Therefore, identifying dockerin domains that show maximum mechanical stability when bound to cohesins, and incorporate them into bioconversion catalysts which could withstand the harsh high-solids stirred bioreactors used for industrial biomass processing could be crucial.

An intramolecular clasp interaction bridging the N- and C-termini of the dockerin was proposed to be a key factor for determining dockerin stability and affinity. Specifically, mechanical clasp motif of ScaA dockerin (wild-type and several mutants) complexes with ScaBDoc-Coh7 from *R. flavefaciens* was studied. The wild-type complex exhibits high mechanical stability, as has been observed for many other cellulosomal protein modules. In terms of the influence of mutations in clasp motif on the unbinding forces, it is too early to make any confident statement. The low yield of good interactions avoided us to get enough data for all mutants due the lower binding affinities

founded and/or lower expression and purity levels that could contribute lower number of interaction.

Several cohesin-dockerins pairs were studied in terms of mechanical stability to cataloging them, understanding their unbinding behavior using a combination of experimental and computational/simulation-based approaches. Three cohesin-dockerin pairs found from 2 different species of anaerobic bacteria were fully characterized. (1) dockerin S (DocS) from *C. thermocellum* complexed with Coh2 from CipA scaffoldin (2) ScaA dockerin from *R. flavefaciens* complexed with Coh7 from ScaB scaffoldin; (3) XModule-dockerin from *R. flavefaciens* with CohE from Ctta scaffoldin. Additionally, several mutants of complexes (1), and (2) were measured.

In view of the results, the knowledge of the dual binding mode existence in Coh2-DocS interaction was incremented, due to the high level of confidence achieved in the force-based single-molecule assay that can discriminate these two bound states based on differences in unbinding mechanics. In the case of Coh7-Doc ScaB interaction several stabilizing amino acids in the dockerin were identified. In the case of the last system studied (Coh-XDoc), which was stabilized by a so called x-module, unbinding forces surpassed everything known to date in terms of rupture force, achieving unbinding forces near 600-750 pN, the highest of their kind ever reported. All-atom SMD simulations suggested a catch bond-like mechanism as responsible for its remarkable stability. This knowledge in combination with a novel analysis method based on steered molecular dynamics calculations and covariance fluctuation analysis, allowed insight into the molecular mechanisms governing the force propagation. Finally, the development of a streamlined approach to covalently attach mechanically stable receptors (e.g. XDoc3) onto proteins of interest via sortase enzymes, allowed stable anchoring of the cellulosomal constituents, improved pickup efficiency and specificity. This new approach improved sample preparation time, throughput and data quality of SMFS experiments.

Another strategy proposed was to developed of a parallelized one-pot *in vitro* transcription-translation and protein pulldown protocol with drastically reduced sample preparation time. This protocol enabled high-throughput of AFM-SMFS measurements using a single cantilever for measure all the constructs obtaining high data quality and improved relative force comparability. The viability of the strategy was evaluated characterizing the force induced unfolding behaviour of all cohesins from the scaffoldin ScaA from *A. cellulolyticus*, as well as revealed how small changes in sequence can have large effects on force resilience in cohesin domains. The new approach provided an efficient way to test and improve the mechanical integrity of protein domains in general.

These insights into the structure-function relationships of cellulosomal complexes, combined with the ability to engineer, mix, match and assemble synthetic systems from said components will serve as a resource moving forward for developing industrially-relevant designer cellulosomes capable of efficient self-assembly and catalysis.

2.6. Description of the main RTD results in WP7: Multi-scale modeling

The main goals of this WP were to create atomistic to supramolecular-scale models of the cellulosome structure and properties to help in predicting, describing, deducing and manipulating the assembly and catalytic properties of cellulosomal components with the aim of apply this knowledge in the development of designer cellulosomes. Methodologies such as electronic

structure calculations, atomic scale Molecular Dynamics (MD) simulations, free energy calculations, coarse-grained MD equilibrium and non-equilibrium calculations were applied. The information exchange between different consortium-research partners were deeply patent in this work package.

Modelling of *DCs* based on SAXS and crystallography

There is no single method that could yield atomic structures of the full-length cellulosome complexes: they are not directly accessible to X-ray crystallography due to the presence of the disordered and flexible segments (although their constituent domains can be crystallized separately); they are also not accessible to protein NMR because of their large sizes; and their inherent flexibility and the lack of symmetries make them practically inaccessible to cryoEM. In order to outline the representative conformations of cellulosomes, various complementary methods must be combined. In particular, small angle X-ray scattering (SAXS) in solution combined with molecular dynamics or other simulational methods can lead to insights into dynamic properties and conformational heterogeneity of this protein complexes.

Structure-based coarse-grained models are commonly applied to study dynamics of large conformational changes in proteins. In this context, several models of cellulosomal structures were created: i) modeling of the full-length endoglucanase Cel9A of *R. champanellensis*. The Cel9A structures obtained from the simulations were compared to the SAXS data and characterizing the conformational ensemble of Cel9A in solution. The same approach was applied to obtain the modeling of the full-length scaffoldin CipA. However, resulting conformations were not comparable to SAXS data, this discrepancy between the simulations and experiments might be explained by the experimental aggregation that CipA scaffoldin suffered.
Another variant of the coarse-grained structure-based model to explore the conformational variability of one component of the cellulosome, was the multi-domain xylanase Z (called as XynZ) of *C. thermocellum*. Simulation results were on agreement with data from SAXS experiments, thus validating the simulation outcome. Data indicated that in the presence or the absence of the cohesin domain, XynZ enzyme appears to be flexible in the sense that it adopts various compact and extended conformations The XynZ enzyme is held together mainly by the flexible linkers connecting the domains. The role of the disordered linkers seems to be providing the required degree of flexibility while maintaining the integrity of the cellulosome complexes. Taken together, our results present a detailed picture of the conformational ensemble of the multi-domain XynZ module in solution.

In the last period of the project, a SAXS-derived model of a designer cellulosome were built. The *DC* were formed by ScaF 20L scaffoldin in complex with three different substrate-specific cellulases: Cel48S, Cel8A and Cel9R.

Atomistic MD: Self-assembly and binding specificity of cellulosomal components

To understand the mechanisms driving the efficiency of the natural cellulosomes, simulations of Cohesin 2 and 3, and the wild-type and terminal-residue mutants of 46 amino acid linkers of the ScaB scaffoldin of *A. cellulolyticu*s were performed. For the wild type linker, the cohesins approached each other without forming a stable dimer. Conversely, the mutated linker leads to the long lived pairing of the cohesins. This drastic change was achieved by modifications on the structure of the junction between the linker and the cohesins. A sharp U-shape turn disappeared

giving more freedom to the linker to move making possible for the cohesin to come close enough to establish a stable pairing. Thus mutations of the linker drastically affect the overall structure of the system. Taking into account the glycosylation pattern of both the wild type and mutated linker, sugars prevented the linkers from collapsing increasing its apparent stiffness, avoiding to leave the space between the cohesins.
Simulations of Coh1 and Coh2 attached by linkers of varying size and initially paired or apart showed that a longer linker affects the distance of contacts by up to 60% for some contacts or decreases by more than 40% for others. Cohesin paired should be taking into account, because the pairing considerably reduces the flexibility of Coh1 when attached to Coh2 with the short linker, leading to drastic changes in the Coh1contacts.

A comprehensive study of binding specificity of CBM and Expansin to different types of cellulosic material at the atomistic level was performed and results provided a comprehensive overview of the residues that are candidates for site-directed mutagenesis.
The multi domain, catalytic subunit Cel9R, which forms part of the trivalent *DC* were modelled and the affinity of its components to defects on the cellulose layer were explored.

The substrate specificity of a Glycosyl hydrolase from family 5 were studied in collaboration with CNRS. Molecular dynamics simulations of four different substrates (mixed-linked glucans) with the rich-aromatic protein binding site were studied. Results indicated that this GH5 enzyme was able to cleave both beta-1,3 and beta-1,4 linkages, however with a net preference for beta-1,3 linkages where hydrogen bonding and hydrophobic interactions contribute to stable complex.

Finally, thermostability of cohesin-dockerin pairs was analysed by classical MD simulations. One thermophilic and two mesophilic cohesin-dockerin pairs were studied at 300K and 400K. According to the data, mesophilic cohesin-dockerin affinity disrupts with the increment in temperature, being the most affected protein the dockerin. This could be due the loss of a helix in its structure. At 400K, mesophilic Coh-Doc pairs showed a tendency to produce random and short lived hydrogen bonds that consequently destabilise the original hydrogen bonds crucial for its affinity while the absolute free energy of binding is reduced.

Atomistic MD mechanical properties

Molecular dynamics simulations represent a powerful tool to explore with atomic detail the dynamics of a protein not accessible experimentally. In particular, steered molecular dynamics (SMD) simulations allow to directly reproduce the experiments carried out by AFM-SMFS and extract highly valuable information of the mechanical unfolding process supporting experimental data. SMD simulations were concentrated on completing the study of the cohesins of *C. thermocellum* and *C. cellulolyticum* scaffoldins. From a previous SMD study it was obtained that the different cohesins showed a mechanical clamp formed by the first and last β strands of the cohesin module and that this mechanical region is the responsible for resisting the unfolding by force. In addition, dockerin binding seems to have no significant effect on cohesin mechanical stability which indicates that the mechanical clamp is the major determinant of the mechanostability of cohesins.

Coarse-grain MD

Cellulosomal enzymes degrade cellulose and unraveling their action requires the understanding of various allomorphs of cellulose through theoretical modelling. The coarse-grained model constructed for cellulose-protein complexes were applied to determine amorphous transition states for the room-temperature conversion process between the Iα and Iβ allomorphs and to characterize the interface between the crystalline forms of the allomorphs.
In order to study the properties of other cellulosomal components, experimental and theoretical methods (all-atom and coarse-grained simulations) were combined to assess the effect of a set of point mutations on the highly mechanostable cohesin 7 module from CipA scaffoldin. The proposed theoretical method to determine the effects of single-site mutations on protein structure and stability was corroborated by experimental data. The identification of mutations that may lead to a substantial increase in mechanical and thermodynamic stabilities could guide to obtaining cellulosomal components with improved characteristics.
The role of the linkers was discussed by performing coarse-grained and all-atom simulations of a number of cellulosomal linkers of different lengths and compositions. It was demonstrated that the effective stiffness of the linker peptides, as quantified by the equilibrium fluctuations in the end-to-end distance, depends primarily on the length of the linker and less so on the sequence. Moreover, it was established that the linker stiffness has an influence on the conformations of the folded domains in the mini-cellulosome Cel8A-ScafT of *C. thermocellum*. The short linkers have been found to affect the structure of the domains they are attached to. This phenomenon is expected to affect the efficiency of catalytic sites, but it is hard to predict in what way. Nevertheless, choosing linkers of an appropriate length may be an additional degree of freedom when attempting to construct more efficient designer cellulosomes. Replacing short linkers (containing fewer than ~10 residues) by longer ones should reduce the strain in cellulases and is likely to result in a more effective catalytic action.

The elucidation of the existence of the dual binding, that it the ability that some cohesins and dockerins may bind in two different ways, was made by simulations of mechanical stretching of the complex. The results indicated that each mode of binding leads to two kinds of stretching pathways, which may be mistakenly taken as evidence of dual binding.
The duality of binding would increase the ease with which a cellulosome could access, to the substrate and thus enhance the efficiency of the degradation process.

2.7. Description of the main RTD results in WP8: Laboratory-scale integration of *DCs*

The main goals of this WP were to integrate all the acquired knowledge and the expertise of the industrial partners into optimized *DCs* as relative simple systems to efficiently degrade the selected lignocellulose substrate at the laboratory-scale level, and improve their activity.

Designing the new designer cellulosomes

All the members of the consortium shared its generated knowledge in order to decide on the most suitable designs.

Weizmann Institute wrote a list of components (matching cohesins and dockerins and CBMs) expanded later on in the program. The initial list also included cellulases and xylanases used for

preparation of the hexavalent designer cellulosome. These components (4 xylanases and 2 cellulases) were free enzymes from the aerobic cellulolytic bacteria *T. fusca*, onto which dockerin modules from various cellulosome-producing bacteria were grafted. Additionally, two more *T. fusca* cellulases were included to prepare an octavalent cellulosome via an adaptor scaffoldin. Finally, the best trivalent scaffoldin from the 45 scaffoldins prepared in previously published work (Vazana et al, 2013) was included, along with the three model cellulases from *C. thermocellum* used in that work.

According to the results in mechanical stability, this parameter is not crucial if *DCs* are not going to be anchored to a cell. Nevertheless, the use of high mechanical stability cohesins and the selection of long linkers is preferred. Although initial versions of the *DCs* were not intended to be anchored on a cell or particle, this might be an interesting approach in the future due to the potential advantages of this approach. Focusing on mechano-stable receptor-ligand complexes, it was proposed that Ctta Xmodule-Dockerins from *R. flavefaciens* could be integrated into *DCs* due to the extremely stable mechanics of its interaction (possibly the most stable ever reported). The complex also exhibited high binding affinity based on independent isothermal titration calorimetry experiments performed at LMU. ScaB is an alternative choice, showing high homology with the Ctta XDoc and also high stability mechanics. These receptor-ligands could be used to incorporate enzymes onto scaffolds of designer cellulosomes.

According to simulators partners, new method to address the effects of single-site mutations on mechanical and thermodynamic stabilities were proposed. Using all-atom simulations to predict structural shifts with respect to the native protein and then analyses the mutants using a structure-based coarse-grained model. In addition, the possibility of introducing disulfide bonds into cohesin mechanical clamp to boost its mechanostability was considered. Any substantial effect on the scaffoldin dynamics were found by the use of different linker lengths. However, due to experimental data reported that long linkers promoted higher cellulosomes activities (without affecting mechanical stability) the use of long natural linkers in *DCs* were proposed.

The list of components selected by CNRS in collaboration with Weizmann institute were elaborated on a compromise of the feasibility of the structural investigations and the coherence with respect to the components selected for the other tasks/partners. The mechanical as well as the biochemical stability of the different components were taken into account. Two highly represented families of enzymes, namely GH5 and GH9, were selected to perform 3D crystallographic determination, thus to comprehend the precise role of the residues interacting with the substrate and to understand the determinants of variations of substrate specificity.

CelDezyner gained some insight into the production process of several enzymes and cellulosomal components at high density cell culture. So, this provided an estimation of which proteins, enzymes and modules could be expressed and purified straightforward with high purity, and which could be may be more challenging.

Composition of optimized designer cellulosomes

Four prominent and important cellulases from *C. thermocellum* (two endocellulases, one exocellulase and one xylanase) were selected as well as 3 enzymes from *T. fusca* (one endocellulase, a lytic polysaccharide monooxygenase and a bifunctional xylanase-laccase) and

one betaglucosidase from *C. bescii* for incorporation into designer cellulosomes *via* one primary and one adaptor scaffoldin. Selected enzymes were chosen based on its activity and its themostability.
Designer scaffoldins enable controlled incorporation of cellulases into the designer-cellulosome complex. ScAd·GBcTFAt2 is a chimaeric pentavalent scaffoldin containing divergent cohesins from *A. fulgidus, B. cellulosolvens, C. thermocellum, R. flavefaciens, A. cellulolyticus*, a type II dockerin and a CBM (c) from *C. thermocellum*. In addition, a tetravalent chimeric scaffoldin (T2CAcV), able to interact with the previously mentioned scaffoldin, was developed. This scaffoldin includes four different cohesin from *C. thermocellum*, *C. cellulolyticum*, *A. cellulolyticus* and *C. clariflavum*, together with a family-3a CBM from *C. thermocellum*. Both scaffoldins were able to interact. An assembly of high molecular *DCs* of 10 components was obtained and this was the first time that cellulases/xylanases are assembled together with 3 key accessory enzymes.

All the components were produced in *E.coli* and purified by chromatographic techniques testing their ability for interaction. With this purpose, optimal ratios between i) the tetravalent scaffoldin and its three matching dockerin-containing enzymes, ii) the pentavalent scaffoldin and its five matching enzymes and iii) both scaffoldins were evaluated.
The designer cellulosome complex was formed according to the selected stoichiometric ratios, primarily the complexes between each scaffoldin and its enzymes were formed and secondly the two scaffoldins were allowed to interact.

Enzymatic performance of designer cellulosomes

To test the enzymatic activity of *DCs*, optimized *DCs* were assayed under various conditions, *e.g.*, temperature, pH, mass loading, etc., so as to find the best performance conditions. One of the most significant factors for optimal hydrolysis of biomass by a consortium of enzymes is the pH of the reaction. Since the *DCs* were constructed from enzymes originating from diverse sources, it is safe to assume that each enzyme has its optimal pH. To find the best pH conditions for the *DCs*, a range of pH were empirically tested using two different model substrates, finding that for swollen cellulose the best pH was around 5. A similar approach was made to find the best buffer; several buffers were tested revealing that acetate buffer was the most optimal.
After determining the optimal buffer and pH, optimal temperature for the activity of the *DCs* was evaluated. Best performance temperature of *DCs* and free enzymes was around 60°C.
Comparing *DCs* against free enzymes, results indicated that on both substrates (PASC and MCC) *DCs* showed better performance although in MCC the improvement were more marked. However, when a comparison with a natural cellulosome were made, the outcome showed much higher activity of the natural cellulosome. These results were in agreement with reported literature.
The quantification of releasing sugars from substrate hydrolysis through HPLC allowed to determine that in *DCs* the bulk of sugars obtained after PASC hydrolysis were glucose while in the case of the natural cellulosome high amount of cellobiose was accounted. This is due to the incorporation of a betaglucosidase in the *DC* able to covert cellobiose into glucose.

Once the optimal conditions for hydrolysis on model substrate were settled, the next step was testing the *DC* on pretreated biomasses, the real substrate for cellulosic ethanol production under commercial setting.

The *DCs* was tested on 3 biomasses received from ABNT (wheat straw, corn straw and municipal solid waste) together with several biomasses taken from CelDezyner library of pretreated biomass as a reference.
According to the results, the less crystallinity of the biomass the better the hydrolysis. ABNT biomasses were pretreated by the combination of heat/pressure and the addition of chemical catalyst. Two of the biomasses were derived from the agricultural sector wheat straw and corn straw, the third biomass was derived from a municipal solid waste fraction. Although municipal solid waste showed promising results when hydrolyzed with a natural cellulosome, when DC was applied, the yield was low and almost no hydrolysis was observed. In contrast, the data obtained from corn straw hydrolysis showed that this substrate can be hydrolyze nicely with the DC, better than the free enzyme system (the same enzymes, not associated with any scaffoldin). Although, by using the natural cellulosome better yields were obtained, this could be due to the fact that the native cellulosome contains up to 73 different enzymes, thus making it better equipped to hydrolyze complex biomasses.

2.8. Description of the main RTD results in WP9: Pre industrial scaling up

The main goals of this WP was to produce and test final *DCs* to efficiently degrade lignocellulose at the preindustrial-scale level, as well as cost-benefit and market analysis of *DCs* feasibility.

Mass production of *DCs* and preindustrial scaling up of the saccharification process

In terms of enzymes production, scale production (1 L bioreactor) of ten proteins was performed in order to obtain sufficient material for purification and activity evaluation. The enzymes chosen were: endocellulase Cel8A-b, exocellulase Cel48S, betaglucosidase A, Cel9I-b, and Cel9k-a from *C. thermocellum.* Exocellulase Cel48A and endocellulase Cel5A from *T. fusca.* Moreover, three different chimeric scaffoldins were also produced: B<u>c</u>TFA, 21L and CA<u>c</u>TABF.

1 L scale up was successfully achieved with all of them. The chosen heterologous host for the expression was in the majority of the cases *E. coli*, but *Bacillus subtilis* was used as alternative host for family 9 enzymes of *C. thermocellum* (Cel9I-b and Cel 9k-a). Regardless of efforts to improve purification yields, not all of them were fully purified, exocellulases Cel48A and Cel48S as well as endocellulase Cel8A-b and Cel9k-a were partially purified. However, activity was validated by Biopolis in all the cases. Regarding the scaffoldins, CAcTABF was partially purified while BcTFA and 21L were produced, purified and checked their ability to bind dockerins.

In terms of the saccharification process, which is the procedure of breaking a complex carbohydrate (*e.g*. cellulose) into its monosaccharide components (*e.g*. glucose), an organic fraction from municipal solid waste (MSW) was used as substrate. MSW was subjected to a thermal pretreatment and the organic fiber obtained was used as substrate. The complete composition of the selected raw materials was performed based on two different protocols. The chemical composition content (lignin, hemicellulose, cellulose, pectin, lipids, proteins, starch, ashes and water) of three different samples was obtained.

In order to determine the benchmark of saccharification several ratios enzyme/substrate were analyzed. The effect of enzyme loading (5, 10, 15 and 20 mg/g glucan) and substrate concentration (20% and 25% w/w) on glucose production was tested. Best performance was

obtained using 20% solids and a dosage of enzyme of 20 mg/g glucan, which corresponded to enzymatic hydrolysis yield in the range of 30-35%.

Five cellulase-contained mini-cellulosomes were designed. Each of them consisting in a chimeric scaffoldin: CAcTABF, BcTFA or 21L (TFA) formed by cohesins from *C. thermocellum* (T), *P. cellulosolvens* (B), *A. cellulolyticus* (A) *R. flavefaciens* (F), *C. cellulolyticum* (C) and a family-3a CBM (c) from C. *thermocellum*. The three enzymes attached to these chimeric scaffoldins were taken from the previously 1L scaled up cellulases.

Saccharification reactions were performed in a total volume of 60 mL including 20% of substrate and a dosage of mini-cellulosomes previously determined (20 mg/g glucan). Additionally, betaglucosaidase (BglA) activity was included as free enzyme. Reactions were carried out at 50 °C for 72 hours, released glucose was determined by HPLC. According to the results, four out of five mini-cellulosomes had comparable saccharification activity to a commercial cocktail in terms of glucose release. Two cellulosomes achieved better saccharification yields than the enzymatic cocktail: one composed by CAcTABF scaffoldin containing Cel48A, Cel8A-b and Cel 9k-a cellulases while the other mini-cellulosome was composed by BcTFA scaffoldin with Cel5A, Cel 9l-b and Cel9k-a cellulases attached to. In view of these results, an optimal mini-cellulosome, combining best performance scaffoldin and cellulases, was designed for pre-industrial scale performance in batch stirred reactor. Saccharification conditions and substrate/mini-cellulosome ratio were similar as mentioned above, though the reaction was performed in 1 L volume. Optimal mini-cellulosome achieved comparable glucose released to the commercial enzymatic preparation used as a control. These results probed that mini-cellulosomes are a realistic alternative to free enzymes in bioethanol production industry.

Market analyses

The global bioethanol market was valued at USD 52.66 Billion in 2016 and is projected to reach USD 68.95 Billion by 2022. The market is driven by the increased demand for bioethanol from various end-use industry segments, such as transportation, pharmaceuticals, cosmetics, alcoholic beverages, and others. In 2016, the transportation industry led in terms of volume the global bioethanol market. The same year, the world production of bioethanol remained at levels of 117.7 million m^3. The world leaders in the production of bioethanol are the USA (59.5 million m^3), followed by Brazil (27.8 million m^3). In the European Union (EU), with the implementation of the "Renewable Energies Directive", that prescribes a mandatory part of 10% renewable energies in the transport sector by 2020, the consumption of bioethanol is expected to rise.
Lignocellulosic ethanol is produced by agricultural biomass including corn cobs, perennial grasses, wheat straws and forest biomass such as wood chips. The availability of feedstock is a major growth driver for cellulosic ethanol market. However, challenges exist in harvesting, storage and transportation and the fluctuation of raw material prices can develop as an obstacle in the bioethanol market growth. Lately, Municipal Solid Waste (MSW) that holds a great potential as biomass source, has been proposed as an alternative raw material for the bioethanol production. Using this type of waste as biomass feedstock provides around a 70% of potential biodegradable materials, which aligns with the E.U. Directive 2003/30/EC. On the other hand, the economic benefits of using MSW as fuel, such as obtaining large biomass quantities at low cost are obvious. Regarding the enzymes, some authors pointed out that the cellulase cost was still as high as up $0.69/gal ethanol or even $1.47/gal ethanol by quoting the current market prices of cellulase products in others industries. Cost reduction on these enzymes usage has been the central effort

in the commercialization of bioethanol production from lignocellulose biomass. Commercially available cellulases are dominated by fungal enzymes, mainly from *Trichoderma* species. Alternatively, other bacterial systems (e.g. cellulosome) have been investigated intensively. However, so far, no cellulosome or cellulosome-derived system is yet commercially available. In terms of the cellulosome product, a pricing advantage is considered due to the high yields of the saccharification process and to the reusability of the enzymatic complex.

Considering direct competitors, Novozymes has launched Cellic® CTec3 product, the third generation of a cellulase cocktail claimed to be 1.5-fold better than the previous generations decreasing in five times the enzyme dosage. This new product could make the cost of cellulosic biofuels to around $2.0/gal of ethanol, which is competitive with production of corn ethanol and gasoline. Companies are investing heavily for improving their distribution channel by strengthening their research & development, marketing & sales, and supply chain & logistics activities. Apart from Novozymes, industry players in enzymes market include: Enmex, Roche, DuPont, Advanced Enzymes, BASF SE, Specialty Enzymes and Biotechnologies (SEB), Associated Bristish Foods plc, MAPS Enzymes Limited, Epygen Labs FZ LLC and Megazyme among others.

Regarding the marketing options for the cellulosome, the main applications are related with technical industries (detergent production, textiles, paper, organic synthesis or biofuels), food industries, food processing and cosmetics. Concerning the global bioethanol market, a segmentation in the application field allows us to stablish different channels through which enzyme complexes may reach the market. In addition, with the rise in population and disposable income, the number of transportation vehicles is also increasing, therefore, such high demand for bioethanol from the transportation industry is creating several growth opportunities. Additionally, there are secondary channels within each of the industrial areas where to use either free enzymes or in complexes. Such fields above mentioned (food industry, food processing and cosmetics) have implicit activities where enzymes can be used to extract the interesting substances: juices, milk, starch, oils, dyes and antioxidants.

Concerning enzymes formulations, it is necessary to study which form (powder, liquid solution, etc.) best suits the application. However, not only is crucial the delivery mode but also the presence of stabilizers and preservatives. Commercial cellulase preparations (e.g. Crystalzyme, Celluclast, Cellulase AP30K, Cellubrix, etc.) are mainly water-soluble and liquid form, however powder forms of some cellulases are also available. Moreover, it should be taken into account the significant impact on the quality, stability and nutritional functions of the final products obtained by the use of enzymes. Therefore, in order to obtain the maximum effectiveness of an enzymatic cocktail an optimization process is required.

Life cycle assestment

Life Cycle Assessment (LCA) methodology is the analysis tool to quantitatively compare the environmental impacts associated with a product or process taking into account all the steps in its life cycle, from its production to its use. Some studies have already addressed the LCA of bioethanol production from lignocellulosic biomass (wheat straw and corn stover) concluding that when compared to gasoline, bioethanol use produces a shift in the environmental problems, reducing both the fossil fuel dependency and greenhouse gas emissions, which are the major associated concerns.

From all the involved steps in bioethanol production, enhanced saccharification (the target of *CellulosomePlus* consortium) would lower the amount of feedstock needed (thus reducing those impacts related with biomass production), and allow milder or shorter pre-treatments (with similar productions) reducing those impacts where pre-treatment is a major factor (*i.e.*, acidification and eutrophication).
Simplified LCA of the production of 1 L ethanol from urban waste consisting of a residue of municipal solid waste was detailed. LCA analysis included the following environmental impact indicators: Global warming potential, decrease in resources (fossil depletion), acidification and eutrophication of water and marine, ecotoxicity and land use.
ISO 14040 LCA framework, SimaPro software and ReCiPe Midpoint (the most recent and harmonized indicator approach available) were employed. In order to carry out the inventory, the parameterization of the process scaled to 300L was taken into account as a starting point for the assessment of energy consumption on an industrial scale. Logically, the scaling to a process plant makes it possible to reduce the consumption associated with each stage within the limits of the system by several orders of magnitude Inventory for cellulosome production included consumption and emission values related to the transport of raw materials, through information obtained from literature and data from the Ecoinvent 3 database. The rest of the inventory has been based on the characteristics of the currents referred to in the material balances, except for the hypotheses indicated above.

The simplified analysis carried out shows the environmental interest of the process, although further improvements are needed both in the biotechnological phases (mainly the hydrolysis stage) and the recovery phase, and an escalation of the process to industrial levels, which allows energy consumption to be modulated. The information limitations applicable to this scale prevent the scope of this LCA from being broadened, although a more favorable assessment, with minor impacts, is envisaged when obtaining the final product, up to the end of its life, including the substitution of petroleum derivatives with high environmental impact.

Product cost

The cost of obtaining ethanol from the organic fraction of municipal solid waste (MSW), in which enzymatic hydrolysis is conducted using designer cellulosomes was studied.

An evaluation of the product cost, understood as the financial costs of manufacturing process, was carried out. Due to the fact that no standardized methodology exists for this analysis, thus, an approximation of the profitability and economic feasibility of the process, projected to an industrial scale were made. The final purpose of this analysis was to validate whether the product can be obtained at a manufacturing cost that does not exceed of the current cost of the benchmark.

The economic feasibility of an ethanol production plant from MSW has been estimated, assuming a capacity of 6,000 tons of processed waste per year and based on reasonable hypotheses regarding the results obtained and the potential of the process. The process would be transferable and extendable to other geographic areas MSW with similar characteristics, which is considered a route with high potential for indirect exploitation of the results. A first estimate for the cost of the product is obtained in the range of 0.77 €/kg of product, competitive enough in relation to current state of the application.

2.9. CONCLUSIONS

The main conclusions of the results are the following:

-The cellulosome of 9 bacterial species were analyzed creating a complete database of cellulosomal components.

-The thermal stability of key cellulosomal enzymes was increased by random and semi-rational mutagenesis, achieving better hydrolysis yields.

-The mechanical hypothesis of the cellulosome was validated and it should be taken into account in order to develop *DCs* bound to cells.

-The development of a unique high density cell culture facilitated the production of cellulosomal components in large quantities. In addition, several standardized assays for monitoring the enzymatic activity were performed, accomplishing a medium throughput assay based on liquid-handling robotics.

-The 3D structures of three glycoside hydrolases were solved and the SAXS data contributed greatly to the development of computational models.

-Different approaches to characterize interactions between cellulosomal components by AFM-SFMS methods were developed, reducing sample preparation time and increasing the data quality and comparability. These methods allowed to discover the strongest protein-protein interaction in nature.

-Several models of multi-domain enzymes and *DCs* were built presenting a detailed picture of the cellulosomal assemblies. Moreover, relevant data in terms of thermostability and substrate specificity of key cellulosomal elements were obtained.

-An optimized *DC* formed by ten components was developed for the first time, mixing cellulase/xylanase activities with three accessory enzymes.

- The complexation of enzymes into optimized *DCs* increased the hydrolysis on model substrates and pretreated lignocellulosic biomass.

-1L scaled up of ten selected cellulosomal components were accomplished.

-The optimized *DC* achieved comparable saccharification activity to a commercial cocktail using Municipal Solid Waste as substrate

3. Potential impact & main dissemination activities and exploitation of results

3.1. IMPACT

SPECIFIC IMPACTS

The *CellulosomePlus* project has impacted many areas. Here we summarize the impact on the **Second Generation biofuel market,** on the **Biotech industry,** on **research,** on **employment (societal),** and on **the partners** themselves.

Impact on the Second Generation biofuel market

The general impact on the market of 2nd Generation Biofuels of the *DC*s area estimated taking into account the following aspects: final cost of the *DCs* cocktail, efficiency of the process, bioethanol yield, environmental impact and final cost of the produced bioethanol.

The main specific conclusions related to the product cost can be summarized as follows:

-Five cellulosomes have been designed. All of them have comparable activity on saccharification to commercial benchmark cellulases at laboratory scale. One optimized synthetic cellulosome has been tested in bioreactor and it has worked as the enzymatic commercial preparation used as a control. Therefore, cellulosomes are a realistic alternative to free enzymes in bioethanol production industry.

-The economic feasibility of an ethanol production plant from OFMSW has been estimated, assuming a capacity of 6,000 tons of processed waste per year and based on reasonable hypotheses regarding the results obtained and the potential of the process.

-The process would be transferable and extendable to other geographic areas OFMSW with similar characteristics, which is considered a route with high potential for indirect exploitation of the results.

-A first estimate for the cost of the product is obtained in the range of 0.77 €/kg of product, competitive enough in relation to current state of the application.

Taking in account the efficiency of the process the pre-industrial results from the industrial partner Biopolis indicate that the *DC*s tested are highly active on OFUSW, comparable to the commercial cocktails. Thus, opening the door to their direct application to the 2nd generation Biofuels industry.

On the other side, the knowledge gained and contacts made in the supportive, competitive environment of *CellulosomePlus* have moved our industrial positioning 2 points in the TRL scale, from TRL3 to TRL5, focussing more on direct industry needs and the final market.

Impact on the biotech industry

-We have discovered and tested a **new industrial parameter** of importance in enzymatic processes: protein mechanostability. In addition to well-known parameters like thermostability, this discovery is expected to guide future protein engineering towards the improvement of enzymatic activity of enzymatic cascades.

-In the future paradigm of Consolidating Bioprocesing, when using *DCs* bound to cells the mechanostability of scaffoldin (the scaffolding protein of *DC*s) is expected to become a critical parameter since it is affecting enzymatic activity.

-We have developed new low and medium throughput enzymatic assays for cellulases activities.

Research impact

-The mechanical hypothesis of the Cellulosome has been tested and validated. It remains to be investigated the source of mechanical stress in the system in order to introduce new improvements in the future.

-We have stablished an internally-controlled method to directly measure protein-protein interactions (the so-called “molecular staple”).This represents just a proof of concept but new developments should be implemented in the future in order to make this strategy a general one for measuring interactions.

-We have discovered the strongest protein-protein interactions found in biology: the type III cohesin-dockerin interaction from *Ruminococcus flavefaciens*, where we measured rupture forces in the range of 600-750 pN.

-We have discovered and described 9 new bacterial cellulosomes including the most complex cellulosome ever (which has 160 enzymes).

-We have developed new atomistic and coarse grain models for cellulosomal components and cellulose.

Societal impact: employment

-*CellulosomePlus* consortium has created around 62 new jobs during the four-year period of the project

-As a result of inner collaboration in the *CellulosomePlus* consortium the new doctoral and postdoctoral researchers contracted have been exposed a highly multidisciplinary environment acquiring a unique profile, which makes them highly competitive and versatile scientists in academia and industry.

Impact on Training

-*CellulosomePlus* has created 62 new positions of Master Thesis students, PhD students and Post-Doctoral researchers.
-As a result of the *CellulosomePlus* activities a number of PhD students and postdoctoral researchers have been trained in the different disciplines involved in the project.

-As a result of inner collaboration in the *CellulosomePlus* consortium the new doctoral and postdoctoral researchers contracted have been exposed a highly multidisciplinary environment acquiring unique profiles, which make them highly competitive scientists in academia and industry

Impact on partners

-Biopolis has been recently acquired by the American Group Archer Daniels Midland (ADM), one of the largest US producers of agricultural goods. As a result, Biopolis (former SME) has become a large company.

-The group of Damien Thompson (Partner 6) got a new position and moved from his prior position at the Tyndall National Institute at the University-College Cork to his current position at the University of Limerick.

-The group of Mariano Carrión-Vázquez (Partner 1) was promoted to a higher position in the CSIC ranking at the Cajal Institute-CSIC.

-*CellulosomePlus* funding has allowed all partners to increase their size particularly in the number of PostDocs contracted and their productivity with 34 (+10 in progress) scientific publications in high-impact journals, most of them in collaboration with other *CellulosomePlus* partners.

-The CellulosomePlus consortium has proved to be highly successful in terms of networking. Its partners have carried out more than 20 networking activities to tackle different technical aspects of the project, to discuss specific results or to traine consortium students. As a result, it has been published 17 jointed papers (5 of them in preparation) and it has been increased the network of experts to collaborate in future projects.

POTENTIAL IMPACT

The *CellulosomePlus* consortium has already developed the first prototypes of these novel enzymatic complexes, and when successfully validated at a larger scale, will be commercially attractive biocatalysts. Furthermore, since there is freedom to operate, it is also expected that in the near future the industrial property rights will be protected through patent application when all technological information has been gathered.

CellulosomePlus has delivered the proposed pre-industrial results, which we expect that will impact industry and society in turn. We expect that the use of self-assembled *DC*s as bio-inspired nanocatalysts produced by our consortium will reduce significantly both the cost of the saccharification step and the environmental impact of the whole process. These *DC*s should be marketable by European biotechnology industries working in the transport-related sector in addition to other chemical industries to process urban waste and residues from the agro-food, paper, and forestry based industries. This should reduce Europe's reliance on oil, strengthen SMEs from the EU, stimulate job creation and reduce the environmental impact of the second generation biofuel sector.

Our end-user, Biopolis, an active member of the European BioBased Industry Consortium (full member since 2015), estimates a return on investment of 100k/year to benefit all contributing partners, if any of these enzymatic complexes is successfully launched to the market, for research/laboratory scale applications. In case any of the products reach full market scale for bulk applications, such as enzymatic hydrolysis of lignocellulosic materials—sugar platform, or improved fermentation through consolidated bioprocessing, a 0.1% share of the global biofuel enzymes market is estimated as 300k/year (>USD 1.500 million, >USD 300 million for cellulases in 2016, growing at >7% CAGR).

The major expected impact of *CellulosomePlus* is that the rational design of *DC*s will enable fast industrialization of materials of very high activity and selectivity, and minimum energy use in their preparation and during the work cycle. *DC*s have the potential of improving the performance of existing industrial processes including energy production and to lead to exploitation of renewable, efficient and inexpensive sources for biotechnological processes.

The results obtained by *CellulosomePlus* are expected to significantly advance European cutting edge research in bio-molecular science and technology. The impact of *CellulosomePlus* is indeed envisaged as to be long-term not only by providing new tools for scientists, but also by making possible the rational design of tailored *DC*s based on their true bio-molecular mechanisms.

Furthermore, the introduction of new nanotechnology concepts and parameters into the established industry will promote transformation of the industry to high value-added production. The education level of the work force is expected to increase significantly and together with more challenging tasks in high-technology industry will increase employee job satisfaction.

Finally, regarding the environmental impact, *CellulosomePlus* will certainly have positive environmental impact both in terms of new and precise methods for analysis and for the development of novel biological catalysts. Furthermore, the analysis of biological nanosystems by experiment and modelling should allow unprecedented insights into the true biomolecular mechanism. It therefore provides new means to control and analyze biological (*e.g.* enzymatic)

processes. In addition, as the *CellulosomePlus* systems developed by this consortium, the *DC*s, spatiotemporally control cascade reactions, this will greatly increase the selectivity due to the nanoscale dimensions as well as the efficiency of reactions due to the reduction in the entropy of the system. This will also result in new and more efficient industrial enzymes. The *CellulosomePlus* technology platform will make it possible for the industry to use less raw materials, produce less waste, and apply a more environmentally friendly production technology.

Summarizing, the *CellulosomePlus* Project has taken the promising technologies in development from the partners involved in this consortium formed by a research-industry complementation in order to provide an increase in the Technological Readiness Level toward a pre-commercial development. It will enable the European industries to better bridge the 'innovation gap' and the "valley of death" between technology development and commercialization. Both very pertinently apply to the field of biorefineries, which requires significant investments, typically beyond the financial reach of individual private companies. Public intervention is required to foster industry leadership and to promote long-term industry commitment in research and innovation related to bio-based industries.

3.2. MAIN DISSEMINATION ACTIVITIES AND EXPLOITATION OF RESULTS

The consortium has continually disseminated the technical and scientific results obtained during the development of the project through Dissemination & Communication activities. These activities aim to i) support the EC efforts to promote EU programmes, ii) to comply with the contractual obligations to disseminate & communicate the outcomes of the project funded, and iii) to benefit the whole consortium and the project partners.

CellulosomePlus has performed a substantial number of activities regarding the dissemination and exploitation of the results obtained in this project. Following we describe the details of these activities:

3.2.1. DISSEMINATION ACTIVITIES

PRESS RELEASES

During the project the partners have made an effort to disseminate project´s results through press releases. Here we list them:

1. *Article in BiofuelDigest website. Published on March 12th 2014.* Partner: **UL**. http://www.biofuelsdigest.com/bdigest/2014/03/12/university-of-limerick-leads-eu-project-to-improve-biomass-processing/
2. *Article in Física Hoy (only in Spanish). Published on 10th May 2014*Partner: **CSIC**. http://greenenergyinvesting.net/innovations-in-biofuel-energy-solutions
3. *Article in Green Energy Investing. Published on 12th March 2014.* Partner: **UL**. http://greenenergyinvesting.net/innovations-in-biofuel-energy-solutions
4. *Article in University of Limerick website in News 'section. Published on March 10th 2014.* Partner: **UL**.

   http://www.ul.ie/news-centre/news/biofuel-energy-solutions-investigated-by-ul-researchers

5. Article in Madri+d website (only in Spanish). Published on 8th May 2015. Partner: **CSIC.**

http://www.madrimasd.org/informacionidi/noticias/noticia.asp?id=60305&origen=notiweb_suplemento&dia_suplemento=jueves&seccion=cooperarparacompetir

6. *Article in Cantabria INFONET (only in Spanish). Published on8th May 2015* Partner: **CSIC**.
https://pas2.cantabria.es/infonet/noticia.asp?id=506901&palabra=cellulosomeplus
7. Article. CellulosomePlus Partner: CSIC. Magazine AENOR. Title of the article: *Las normas acercanla innovación al Mercado H2020*. (Only available in Spanish) Febrero 2015 - Nº 301
http://www.aenor.es/revista/307/index.html
8. Article in Weizmann website. Published on July 2016. Partner: **WEIZ**
http://www.weizmann.ac.il/resdev/WeizmannDirect/news/July/2016/whats-new-in-science
9. Press release. American Institute of Physics Media press release, September 13, 2017, written by Julia Majors. Partner: IFPAN

https://publishing.aip.org/publishing/journal-highlights/getting-point-mutations-re-engineering-biofuel-producing-bacterial

10. Press release. Getting to the point (mutations) in re-engineering biofuel-producing bacterial enzymes. Sep 13, 2017. Partner: IFPAN

Republished by https://eurekalert.org/pub_releases/2017-09/aiop-gtt091317.php

Additionally, The consortium has also **advertised its activities** in the following **WEBSITES:**

1. CellulosomePlus Consortium website. http://cellulosomeplus.eu/en/
2. CellulosomePlus Partner: CSIC. http://carrionvazquez-lab.org/es/page.cfm?id=69&title=proyectos-activos#.VS_rPNysXTo

3. CellulosomePlus Partner: BIOPOLIS. http://www.biopolis.es/en/company/R-and-D-track-record.php

4. CellulosomePlus Partner: UL. *University of Limerick website in Blog 'section. Published on March 10th 2014.* Partner
http://universityoflimerickresearch.com/dev1/blog/this%20is%20strange

5. CellulosomePlus Partner: WEIZ. http://www.weizmann.ac.il/pages/bacteria-biofuel

6. CellulosomePlus Partner: IFPAN.
http://info.ifpan.edu.pl/ACTIVITY/Press_release_launch_of_CellulosomePlus_draft_V1.pdf

PUBLICATIONS, CONGRESSES, CONFERENCES, MEETINGS, WORKSHOPS, SEMINARS & PhD THESES

The *CellulosomePlus* consortium has made a great effort along the whole project to effectively disseminate and communicate the main technical and scientific outcomes of the project. It has been published 34 articles (plus nine in different phases of publication) in high impact generalist and specialized journals (*Proc. Natl. Acad. Sci., ACS Nano, Adv. Mat., Angew Chem, J. Chem. Phys., Cellulose, Journal: Int Ed Engl., Biotechnol. Biofuels,).* The number of communication actions taken during this initial period were 88 in total (57 attendances to congresses and conferences, 13 workshops and 15 seminars). The results of the project were presented in the most representative congresses and conferences in the specialized fields of biophysics,

nanomechanics, computational science and the biorefineries field. 3 PhD Theses were also defended as a result of CellulosomePlus activities.

Here we provide the **list of publications** derived from this project:

1. Arfi Y, Shamshoum M, Rogachev I, Peleg Y, **Bayer EA**. Integration of bacterial lytic polysaccharide monooxygenases into designer cellulosomes promotes enhanced cellulose degradation. Proc Natl Acad Sci U S A. 2014 Jun 24;111(25):9109-14. doi: 10.1073/pnas.1404148111. Epub 2014 Jun 9.

2. Artzi L, Morag E, Barak Y, Lamed R, **Bayer EA**. *Clostridium clariflavum*: Key Cellulosome Players Are Revealed by Proteomic Analysis. MBio. 2015 May 19;6(3):e00411-15. doi: 10.1128/mBio.00411-15.

3. Chwastyk M, Bernaola AP, **Cieplak M**. Statistical radii associated with amino acids to determine the contact map: fixing the structure of a type I cohesin domain in the Clostridium thermocellum cellulosome. Phys Biol. 2015 May 27;12(4):046002. doi: 10.1088/1478-3975/12/4/046002.

4. Chwastyk M, Vera AM, Galera-Prat A, Gunnoo M, **Thompson D**, **Carrión-Vázquez M**, **Cieplak M**. Non-local effects of point mutations on the stability of a protein module. J Chem Phys. 2017 Sep 14;147(10):105101. doi: 10.1063/1.4999703.

5. Dassa B, Borovok I, Lombard V, Henrissat B, Lamed R, Bayer EA, **Moraïs S**. Pan-Cellulosomics of Mesophilic Clostridia: Variations on a Theme. Microorganisms. 2017 Nov 18;5(4). pii: E74. doi: 10.3390/microorganisms5040074.

6. Dassa B, Borovok I, Ruimy-Israeli V, Lamed R, Flint HJ, Duncan SH, Henrissat B, Coutinho P, Morrison M, Mosoni P, Yeoman CJ, White BA, **Bayer EA**. Rumen cellulosomics: divergent fiber-degrading strategies revealed by comparative genome-wide analysis of six ruminococcal strains. PLoS One. 2014 Jul 3;9(7):e99221. doi: 10.1371/journal.pone.0099221. eCollection 2014.

7. Davidi L, Moraïs S, Artzi L, Knop D, Hadar Y, Arfi Y, **Bayer EA**. Toward combined delignification and saccharification of wheat straw by a laccase-containing designer cellulosome. Proc Natl Acad Sci U S A. 2016 Sep 27;113(39):10854-9. doi: 10.1073/pnas.1608012113. Epub 2016 Sep 12.

8. Durner E, Wolfgang O, Nash M, **Gaub H**. Post-Translational Sortase-Mediated Attachment of High-StrengthForce Spectroscopy Handles. ACS Omega 2017, 2, 3064−3069. doi: 10.1021/acsomega.7b00478

9. Galanopoulou AP, Moraïs S, Georgoulis A, Morag E, **Bayer EA**, Hatzinikolaou DG. Insights into the functionality and stability of designer cellulosomes at elevated temperatures. Appl Microbiol Biotechnol. 2016 Oct;100(20):8731-43. doi: 10.1007/s00253-016-7594-5. Epub 2016 May 21.

10. Gunnoo M, Cazade PA, Galera-Prat A, Nash MA, **Czjzek M**, **Cieplak M**, **Alvarez B, Aguilar M, Karpol A**, **Gaub H, Carrión-Vázquez M, Bayer EA**, **Thompson D**. Nanoscale Engineering of Designer Cellulosomes. Adv Mater. 2016 Jul;28(27):5619-47. doi: 10.1002/adma.201503948. Epub 2016 Jan 7.

11. Hamberg Y, Ruimy-Israeli V, Dassa B, Barak Y, Lamed R, Cameron K, Fontes CM, **Bayer EA**, Fried DB. Elaborate cellulosome architecture of *Acetivibrio cellulolyticus* revealed by selective screening of cohesin-dockerin interactions. PeerJ. 2014 Oct 30;2:e636. doi: 10.7717/peerj.636. eCollection 2014.

12. Jobst MA, Milles LF, Schoeler C, Ott W, Fried DB, **Bayer EA**, **Gaub HE**, Nash MA. Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy. Elife. 2015 Oct 31;4. pii: e10319. doi: 10.7554/eLife.10319..

13. Malinowska KH, Rind T, Verdorfer T, **Gaub HE**, Nash MA. Quantifying Synergy, Thermostability, and Targeting of Cellulolytic Enzymes and Cellulosomes with Polymerization-Based Amplification. Anal Chem. 2015 Jul 21;87(14):7133-40. doi: 10.1021/acs.analchem.5b00936. Epub 2015 Jun 26.

14. Moraïs S, Stern J, Kahn A, Galanopoulou AP, Yoav S, Shamshoum M, Smith MA, Hatzinikolaou DG, Arnold FH, **Bayer EA**. Enhancement of cellulosome-mediated deconstruction of cellulose by improving enzyme thermostability. Biotechnol Biofuels. 2016 Aug 4;9:164. doi: 10.1186/s13068-016-0577-z. eCollection 2016.

15. Muñoz-Gutiérrez I, Ortiz de Ora L, Rozman Grinberg I, Garty Y, **Bayer EA**, Shoham Y, Lamed R, Borovok I. Decoding Biomass-Sensing Regulons of Clostridium thermocellum Alternative Sigma-I Factors in a Heterologous Bacillus subtilis Host System. PLoS One. 2016 Jan 5;11(1):e0146316. doi: 10.1371/journal.pone.0146316. eCollection 2016.

16. Nawrocki G, Cazade PA, **Thompson D**, **Cieplak M**. Peptide Recognition Capabilities of Cellulose in Molecular Dynamics Simulations. Journal of Physical Chemistry C. J. Phys. Chem. C, 2015, 119 (43), pp 24404–24416. doi: 10.1021/acs.jpcc.5b07118.

17. Vera A. & **Carrión-Vázquez M**. Direct Identification of Protein-Protein Interactions by Single-Molecule Force Spectroscopy. Angew Chem Int Ed Engl. 2016 Nov 2;55(45):13970-13973. doi: 10.1002/anie.201605284.

18. Oroz J, Bruix M, Laurents DV, Galera-Prat A, Schönfelder J, Cañada FJ, **Carrión-Vázquez M**. The Y9P Variant of the Titin I27 Module: Structural Determinants of Its Revisited Nanomechanics. Structure. 2016 Apr 5;24(4):606-616. doi: 10.1016/j.str.2016.02.016. Epub 2016 Mar 24.

19. Poma AB, Chwastyk M, **Cieplak M**. Coarse-grained model of the native cellulose $$\hbox {I}\alpha$$ I α and the transformation pathways to the $$\hbox {I}\beta$$ I β allomorph. Cellulose. Vol. 23/Issue 3. Springer Netherlands. doi: 10.1007/s10570-016-0903-4.

20. Poma AB, Chwastyk M, **Cieplak M**. Polysaccharide-Protein Complexes in a Coarse-Grained Model. J Phys Chem B. 2015 Sep 10;119(36):12028-41. doi: 10.1021/acs.jpcb.5b06141. Epub 2015 Aug 26.

21. Rozman Grinberg I, Yin G, Borovok I, Berg Miller ME, Yeoman CJ, Dassa B, Yu Z, Mizrahi I, Flint HJ, **Bayer EA**, White BA, Lamed R. Functional phylotyping approach for assessing intraspecific diversity of Ruminococcus albus within the rumen microbiome.FEMS Microbiol Lett. 2015 Jan;362(3):1-10. doi: 10.1093/femsle/fnu047.

22. Różycki B, Cazade PA, O'Mahony S, **Thompson D**, **Cieplak M**. The length but not the sequence of peptide linker modules exerts the primary influence on the conformations of protein domains in cellulosome multi-enzyme complexes. Phys Chem Chem Phys. 2017 Aug 16;19(32):21414-21425. doi: 10.1039/c7cp04114d.

23. Różycki B, **Cieplak M**, **Czjzek M**. Large conformational fluctuations of the multi-domain xylanase Z of Clostridium thermocellum. J Struct Biol. 2015 Jul;191(1):68-75. doi: 10.1016/j.jsb.2015.05.004. Epub 2015 May 23.

24. Różycki B, **Cieplak M**. Stiffness of the C-terminal disordered linker affects the geometry of the active site in endoglucanase Cel8A. Mol Biosyst. 2016 Nov 15;12(12):3589-3599.

25. Schoeler C, Bernardi RC, Malinowska KH, Durner E, Ott W, **Bayer EA**, Schulten K, Nash MA, **Gaub HE**. Mapping Mechanical Force Propagation through Biomolecular Complexes. Nano Lett. 2015 Nov 11;15(11):7370-6. doi: 10.1021/acs.nanolett.5b02727. Epub 2015 Aug 19.

26. Schoeler C, Malinowska KH, Bernardi RC, Milles LF, Jobst MA, Durner E, Ott W, Fried DB, **Bayer EA**, Schulten K, **Gaub HE**, Nash MA. Ultrastable cellulosome-adhesion complex tightens under load. Nat Commun. 2014 Dec 8;5:5635. doi: 10.1038/ncomms6635.

27. Smith SP, **Bayer EA**, **Czjzek M**. Continually emerging mechanistic complexity of the multi-enzyme cellulosome complex. Curr Opin Struct Biol. 2017 Jun;44:151-160. doi: 10.1016/j.sbi.2017.03.009. Epub 2017 Apr 6. Review.

28. Stern J, Anbar M, Moraïs S, Lamed R, **Bayer EA**. Insights into enhanced thermostability of a cellulosomal enzyme. Carbohydr Res. 2014 May 7;389:78-84. doi: 10.1016/j.carres.2014.01.014. Epub 2014 Jan 27.

29. Stern J, Moraïs S, Lamed R, **Bayer EA**. Adaptor Scaffoldins: An Original Strategy for Extended Designer Cellulosomes, Inspired from Nature. MBio. 2016 Apr 5;7(2):e00083. doi: 10.1128/mBio.00083-16.

30. Vera AM, **Carrión-Vázquez M**. Direct Identification of Protein-Protein Interactions by Single-Molecule Force Spectroscopy. Angew Chem Int Ed Engl. 2016 Nov 2;55(45):13970-13973. doi: 10.1002/anie.201605284. Epub 2016 Oct 13.

31. Verdorfer T, Bernardi RC, Meinhold A, Ott W, Luthey-Schulten Z, Nash MA, **Gaub HE**. Combining in Vitro and in Silico Single-Molecule Force Spectroscopy to Characterize and Tune Cellulosomal Scaffoldin Mechanics. J Am Chem Soc. 2017 Nov 3. doi: 10.1021/jacs.7b07574. [Epub ahead of print]

32. Wasserberg D, Cabanas-Danés J, Prangsma J, O'Mahony S, Cazade PA, Tromp E, Blum C, **Thompson D**, Huskens J, Subramaniam V, Jonkheijm P. Controlling Protein Surface Orientation by Strategic Placement of Oligo-Histidine Tags. ACS Nano. 2017 Sep 26;11(9):9068-9083. doi: 10.1021/acsnano.7b03717. Epub 2017 Sep 13.

33. Wojciechowski M, **Cieplak M**. Dual binding mode in cohesin-dockerin complexes as assessed through stretching studies. J Chem Phys. 2016 Oct 7;145(13):134102.

34. Wojciechowski M, **Thompson D**, **Cieplak M**. Mechanostability of cohesin-dockerin complexes in a structure-based model: anisotropy and lack of universality in the force profiles. J Chem Phys. 2014 Dec 28;141(24):245103. doi: 10.1063/1.4904726.

35. Wołek K, **Gómez-Sicilia À**, **Cieplak M**. Determination of contact maps in proteins: A combination of structural and chemical approaches. J Chem Phys. 2015 Dec 28;143(24):243105. doi: 10.1063/1.4929599.

Publications in progress

1. J. Dorival, S. Ruppert, M. Chapelais, J. Dabin, M. Gunnoo, A. Galera-Prat, T. Barbeyron, A. Labourel, G. Michel, **D. Thompson**, **M. Czjzek**, S. Genicot. Structural insights into the substrate specificity of a marine glycoside hydrolase family 5 mixed-linkage-glucanase from Zobellia galactanivorans. FEBS Journal in preparation

2. Dorival J, Różycki B, Gunnoo M, Dabin J, Ruppert S, Cazade PA, Galera-Prat A, **Cieplak M, Thompson D**, **Czjzek M**. Small angle X-ray scattering and molecular modelling of designer nanosomes built from the ScaF-20L and cipA. JBC in preparation.

3. Dorival J, Dabin J, Fierrobe H-P, Ravachol J, Jeudy A, **Czjzek M**. Exploring the determinants for substrate specificity in the 3D structures of variant GH9 enzymes. JBC in preparation

4. Vera A., Galera-Prat A., Laurents DV, **Bayer ED, Carrión-Vázquez M**. Direct test of the dual binding mode in dockerin-cohesin complexes. In preparation.

5. Galera-Prat A, Laurents DV, **Carrión-Vázquez M**, Pantoja-Uceda D. Solution conformation of a cohesin module and its scaffoldin linker from a prototypical cellulosome. Accepted in Archives in Biochemistry and Biophysics.

6. Galera-Prat A, Sarah Moraïs S, Yael Vazana Y, **Bayer ED**, **Carrión-Vázquez M**. Activity tests strongly support the mechanical hypothesis of the cellulosome. In preparation to PNAS.

7. Galera-Prat A, Sarah Moraïs S, Yael Vazana Y, **Bayer ED**, **Carrión-Vázquez M**. The cohesin module is a major determinant of cellulosome mechanical stability. Accepted in JBC.

8. Barruetabeña, Alonso-Lerma B, Galera-Prat A, Joudeh N, Barandiaran L, Aldazabal, Garcia-Parra, Marie Fertin P, Manteca A, Rodriguez-Couto S, Arbulu M, De Sancho D, **Carrión-Vázquez M**, Raul Perez-Jimenez. Improving the catalytic efficiency of a cellulase by ancestral sequence reconstruction. Submitted to Nature Structure and Molecular Biology.
9. Fernández-Ramírez MC, Hervás R, Galera-Prat A, Laurents DV, **Carrión-Vázquez M**. A novel strategy for accurate and efficient nanomechanical analysis of intrinsically disordered proteins. Submitted to Biophysical journal.
10. M. Wojciechowski, B. Rozycki, Pham D. Q. Huy, M. S. Li, **E. A. Bayer**, and **M. Cieplak**, Dual binding in cohesin-dockerin complexes: the energy landscape and the role of short, terminal segments of the dockerin module (submitted to Scientific Reports)

Book chapters

1. Stern J, Artzi L, Moraïs S, Fontes CMGA, **Bayer EA**. Carbohydrate Depolymerization by Intricate Cellulosomal Systems. Methods Mol Biol. 2017;1588:93-116. doi: 10.1007/978-1-4939-6899-2_8.
2. Artzi, L., Dassa, B., Borovok, I., Shamshoum, M., Lamed, R., and **Bayer, E**. **A.** Cellulosomics of the cellulolytic thermophile *Clostridium clariflavum*. Biotechnol. Biofuels. 2014; 7:100.

3.2.2. EXPLOTATION OF RESULTS

ANALYSIS OF IPR.

The possible outcome of a patent or a commercial product from the *CellulosomePlus* project due to the research done within the framework of the project is highly likely. To make sure that the patent or the commercial product does not infringe other patents, we have recently performed a "Freedom To Operate" (FTO) search through publically available patent databases (WIPO, USPTO, LENS). The search period was restricted to patents that were issued between our last FTO search to the current date (02/ Nov /2016 to 01/Nov/2017) – a period of 12 months. Our final conclusion from the FTO search was that we did not found any patent that we may infringe within the Cellulosome Plus project.