

## CELLULOSOME STRUCTURE DETERMINATION BY ATOMISTIC SIMULATIONS COMBINED WITH EXPERIMENTAL ASSAYS

**Allocation:** Illinois/680 Knh

**PI:** Isaac Cann<sup>1</sup>

**Co-PI:** Rafael C. Bernardi<sup>1</sup>

**Collaborators:** Klaus Schulten<sup>1</sup>, Edward Bayer<sup>2</sup>, Hermann Gaub<sup>3</sup>, Michael Nash<sup>4</sup>

<sup>1</sup>University of Illinois at Urbana-Champaign

<sup>2</sup>Weizmann Institute

<sup>3</sup>Ludwig Maximilian University of Munich

<sup>4</sup>Universit of Basel

### EXECUTIVE SUMMARY

Cellulosomes, which deconstruct cellulose and hemicellulose, are present in many anaerobic bacteria. Cellulosomal cohesion-dockerin interactions enable the incorporation of catalytic cellulases and hemicellulases onto the cellulosome scaffoldin. Their very efficient mechanism of degrading plant cell-wall biomass makes cellulosomes of interest to the second-generation biofuel industry, which aims to produce ethanol from agricultural waste. Furthermore, the recent discovery of cellulosomal bacteria in the lower gut of humans is paradigm-shifting as it has allowed demonstration of the capacity to degrade both hemicellulose and cellulose, at least, in the gut of some humans. Employing molecular dynamics simulations to complement single-molecule and biochemistry experiments, we characterized cellulosome's components, showing that even a single mutation can cause a large change in cellulosome structural stability.

### RESEARCH CHALLENGE

Symbiotic bacteria greatly influence human health and play a significant role in pathogenesis, disease predisposition, physical fitness, and dietary responsiveness [1]. Moreover, bacteria play a key role in the second-generation biofuel industry where their cellulolytic enzymes are used for plant cell-wall degradation [2]. Here, we investigate key processes underlying bacterial activity, namely plant fiber metabolism. Specifically, we seek the structure and function of cellulosomes, the highly cooperative macromolecular complex that is central for this metabolic process in some bacteria [3].

Cellulosomes are multi-enzyme complexes that enable deconstruction of cellulose and hemicellulose in anaerobic cellulosome-containing bacteria. Integration of cellulosomal components occurs via highly ordered protein-protein interactions among three major components. In cellulosome assembly, a large noncatalytic polypeptide called the scaffoldin, embedded with various cohesins, anchors dockerin-containing enzymes through cohesion-dockerin interactions (see Fig. 1). Specificity of the cohesion-dockerin interaction allows incorporation of different catalytic cellulases and hemicellulases onto the scaffoldin, which may or may not be bound to another domain tethered to the

cell wall [4]. Cellulosome assembly promotes the exploitation of enzyme synergism because of spatial proximity and enzyme-substrate targeting.

Using stochastic search algorithms connected to MD tools, we are building the first comprehensive structure of a cellulosome including enzymatic domains. We expect that a complete model of a cellulosome's structure will shed light on the mechanism that allows these enzymatic complexes to be highly efficient.

### METHODS & CODES

Combined with biochemical and single-molecule experiments, we employ molecular dynamics (MD), steered MD (SMD), and generalized simulated annealing (GSA) simulations on Blue Waters utilizing QwikMD [5], a graphical interface connecting Visual MD (VMD) and Nanoscale MD (NAMD). Using stochastic search algorithms coupled to NAMD we can generate thousands of different structure conformations for the cellulosome [6]. GSA [7] analysis shows that the different linkers between cohesin and dockerin (a carbohydrate-binding molecule) in cellulose-integrating protein A (CipA) scaffoldin assume a number of stable conformations. Small angle X-ray scattering analysis has previously shown that three conformations are observed for one of the cellulosomal linkers. GSAFold can predict these three conformations and all the other conformations for CipA. To perform this analysis, 20,000 conformations were obtained per linker and clustered. Combined, these linker conformations would give us  $10^{43}$  CipA conformations. From clustering, we reduce this number to 3,888 structures that were obtained and subjected to a cluster analysis that gave rise to the five most significant structures.

### RESULTS & IMPACT

Following well-established protocols for large macromolecular systems, and using one of the CipA conformations that we obtained using GSAFold, we built a first model of an entire cellulosome structure. MD simulations are now being employed to study the quaternary structure stability. Also, combining biochemical and single-molecule experiments with MD and SMD simulations, we investigated a series of cellulosomal cohesins from *Acetivibrio cellulolyticus*. We revealed that these cellulosomal components withstand different amounts of force depending on their position

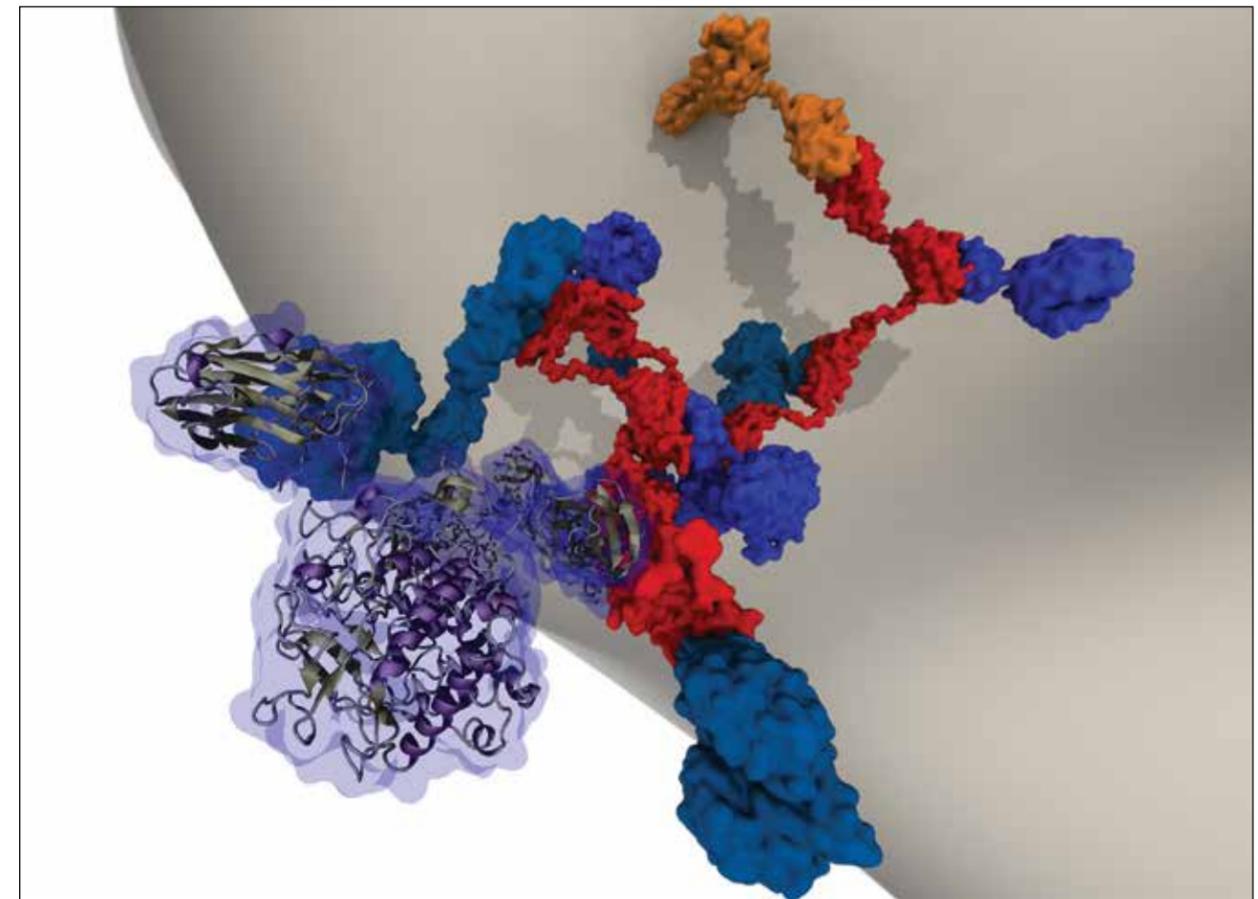


Figure 1: Illustration of a complete cellulosomal structure over a bacterial cell surface. The cohesin-based scaffoldin (red) can be attached by a specific cohesion-dockerin interaction to a cell-anchoring domain (in orange). Another cohesion-dockerin interaction is responsible for attaching the enzymatic domains (blue) to the scaffoldin.

in the protein network. In this study, we combined one-step *in vitro* expression and specific covalent pulldown of protein constructs to assess the mechanical stability of highly related proteins in a parallel single molecular force spectroscopy assay. Using SMD simulations, we reproduced the experimental results and identified important amino acids. In addition, we used the simulations performed on Blue Waters to suggest mutations that were experimentally performed by site-directed mutagenesis, engineering proteins to pin down single crucial amino acids promoting force resilience.

### WHY BLUE WATERS

Investigating the structure and functional processes of large enzymatic complex machineries such as the cellulosomes is only possible on petascale computing resources like Blue Waters. Structures obtained using enhanced sampling techniques such as GSA are only reliable if thousands of conformations (models) are predicted. Employing GSA for the numerous linkers of the cellulosome is a well-suited task for the large-scale parallel architecture of Blue Waters.

### PUBLICATIONS AND DATA SETS

Schoeler, C., et al., Ultrastable cellulosome-adhesion complex tightens under load. *Nat. Commun.*, 5 (2014), p. 5635.

Schoeler, C., et al., Mapping mechanical force propagation through biomolecular complexes. *Nano Letters*, 15 (2015), p. 7370.

Bernardi, R.C., M.C.R. Melo, and K. Schulten, Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim. Biophys. Acta-Gen. Subj.*, 1850 (2015), pp. 872-877