Designing a Bioremediator: Mechanistic Models Guide Cellular and Molecular Specialization

Marco Zaccaria¹, William Dawson², Viviana Cristiglio³, Massimo Reverberi⁴, Laura E. Ratcliff⁵, Takahito Nakajima², Luigi Genovese⁶, and Babak Momeni¹

¹ Department of Biology, Boston College, Chestnut Hill MA, United States.
² RIKEN Center for Computational Science, Kobe, Japan.
³ Institut Laue-Langevin, Grenoble, France.
⁴ Department of Environmental & Evolutionary Biology, Sapienza Università di Roma, Rome, Italy.
⁵ Department of Materials, Imperial College, London, United Kingdom.
⁶ CEA, INAC-MEM, L-Sim, Université Grenoble-Alpes, Grenoble, France.

Corresponding Author:
Babak Momeni¹
140 Commonwealth Ave, Chestnut Hill, MA 02467 USA
Email address: momeni@bc.edu
Designing a Bioremediator: Mechanistic Models Guide
Cellular and Molecular Specialization

Marco Zaccaria¹, William Dawson², Viviana Cristiglio³, Massimo Reverberi⁴, Laura E. Ratcliff⁵, Takahito Nakajima², Luigi Genovese⁶ and Babak Momeni¹

¹ Department of Biology, Boston College, Chestnut Hill MA, United States.
² RIKEN Center for Computational Science, Kobe, Japan.
³ Institut Laue-Langevin, Grenoble, France.
⁴ Department of Environmental and Evolutionary Biology, “Sapienza” Università di Roma, Rome, Italy.
⁵ Department of Materials, Imperial College, London, United Kingdom.
⁶ CEA, INAC-MEM, L-Sim, Université Grenoble-Alpes, Grenoble, France.

ABSTRACT
Rational, mechanistic design can substantially improve the performance of bioremediators for applications including waste treatment and food safety. We highlight how such improvement can be informed at the cellular level by theoretical observations especially in the context of phenotype plasticity, cell signaling, and community assembly. At the molecular level, we suggest enzyme design using techniques such as Small Angle Neutron Scattering and Density Functional Theory. To provide an example of how these techniques could be synergistically combined, we present the case-study of the interaction of the enzyme laccase with the food pollutant aflatoxin B₁. In designing bioremediators, we encourage interdisciplinary, mechanistic research to transition from an observation-oriented approach to a principle-based one.

1. Introduction
To decontaminate an environment from pollutants, chemical and physical means of remediation often rely on compounds that prove unsafe in the long run. The alternative is to employ biological agents, referred to as bioremediation. Bioremediation has been attempted on several targets, with varying degrees of success; in particular, hydrocarbons, pesticides, explosives, radioactive compounds, heavy metals, and metalloids [1–4 and references therein]. The main asset of a successful bioremediation system is arguably long-term sustainability. However, the intricacies of biological systems make bioremediation an inherently complex feat. Bioremediation could span beyond the competences and customary research approaches of biologists. In what follows, we outline some strategies to guide the design of efficient bioremediators and for that, we make a case for an interdisciplinary approach. We will point out specific tools and techniques not consistently adopted in bioremediation, yet well established in other areas of physics, chemistry, and biology.
Depending on the focal biological entity, the conventional bioremediation approaches can be divided into two main methods: cellular and molecular.

**The cellular approach:** In these instances, bioremediation is performed by populations or communities of different cells or multicellular organisms. To this day, bacteria and fungi have been the main candidates to perform cellular bioremediation. Bacteria can make reliable bioremediators when the target pollutant is exploited for the bioremediator’s specific energetic metabolism [5]. When the target molecule is abundant in an environment to which the bioremediator is adapted, a bacterial bioremediator is likely the best option. Fungi, on the contrary, are often metabolic generalists with higher energetic demands. Even as eclectic scavengers, they will suffer competition with bacteria when placed outside of their customary niche, which reduces their applicability range to specific environments. Conversely, whenever the target pollutant is not readily bioavailable, either physically or biochemically, hyphal growth and an extensive enzymatic inventory allow fungi to affect pollutants more efficiently than bacteria [5].

**The molecular approach:** Bioremediation can employ subcellular elements, such as enzymes, to target select molecules. Fungal enzymes have been the most studied in bioremediation applications. The energetic metabolism of fungi relies on efficient scavenging in disparate environmental conditions, and thus revolves around enzymes that display broad substrate affinity. Such versatility makes them more suitable to tackle pollutants of synthetic origin (e.g. pesticides), natural compounds of not high occurrence (e.g. metals and metalloids), and recalcitrant organic molecules in general (e.g. polychlorinated compounds).

In both the cellular and molecular approaches, we endorse the development of a formalized, mechanism-oriented *modus operandi* centered around iterative improvement of bioremediation performance. In the former, we will highlight recent theoretical developments on microbial population dynamics. In the latter, we will discuss how enzymes with bioremediation potential, such as laccase [6], could be optimized. To show the practical implementation of our discussions, we will highlight a specific case that our laboratory is currently working on: contamination of food commodities by mycotoxins (Box 1).

**BOX 1. Bioremediation for decontaminating food commodities.** Mycotoxins are among the most dangerous carcinogenic natural compounds [49]. Among mycotoxins, especially concerning are the Aflatoxins, which are synthesized by the genus Aspergillus [50]. Aflatoxins are a classic topic in bioremediation research. They are aromatic compounds and, incidentally, practical molecules to experiment on, thanks to their environmental stability and the natural fluorescence attributed to the lactone ring in their structure, which is also the main effector in the toxicity [51]. Therefore, it is possible to correlate loss of fluorescence to loss of toxicity. They stand as our case study for two main reasons: 1) as hydrocarbons, they are a category theoretically amenable to bacterial bioremediation and exploitable as a carbon source; 2) as aromatic compounds, they are vulnerable to the action of well-characterized ligninolytic enzymes present in fungal species. Yet, in spite of this theoretical disposition (as previously highlighted), successful bioremediation has yet to be proven feasible in most contexts. Aflatoxins thus stand as a relevant example of pollutants that will require substantial methodological development in the bioremediation field to be efficiently detoxified.
2. The Cellular Approach

In cellular bioremediation, ideally select species/organisms are applied to remove or deactivate pollutants in a context that guarantees the long-term sustainability of the process of interest. The variables at play are numerous, and the know-how to factor them all requires a strong theoretical background, especially when the employed bioremediators are targeted to enhance their traits of interest. Artificial Selection (AS) in this context is a powerful technique towards designing more efficient bioremediators. Bacteria, in particular, represent an apt target, thanks to their fast generation-turnover and suitability for molecular manipulations.

We highlight three examples of cross-disciplinary efforts that can guide the design of efficient bioremediators. These examples also emphasize the importance of collaboration between theorists and experimentalists.

2.1. Theoretical Insight about Variation and Plasticity Informs Efficient Artificial Selection

When employing cells for bioremediation, genotypic and phenotypic variability [7,8] are assets that can be exploited towards a desired outcome. When a bioremediator is applied in the form of a single population, measures to promote long-term survivability of a species in a new environment will benefit from bet-hedging. This is even more so for populations selected to display a fast growth rate, which, because of that very trait, would be more vulnerable to extinction through demographic fluctuations.

An insightful, theoretical paper by Kussel and Leibler [9] helps formalize intrapopulation variation and correlate environmental fluctuations to contextually favorable adaptation mechanisms [10,11] and thereby to develop informed AS experiments. Specifically, phenotypic plasticity is classified as either responsive (R, triggered by sensing) or stochastic (S, autonomous). The model formalizes how, in an unpredictable environment, “R” individuals who develop sensing elements are favored, whereas in predictably varying contexts the “S” individuals who spare the cost of developing a sensor have the advantage. In our specific case of aflatoxin bioremediators, we observe species of the genus Rhodococcus achieving degradation through an uncharacterized metabolic process. Applying the model to our own experimental system, one can infer what intervals of environmental variability of toxin exposure are favorable to solicit Rhodococcus degradation; at the same time, this can reveal the underlying mechanism of response (responsive versus stochastic) and its implications.

2.2. Cell Signaling Allows Taking Advantage of Existing Cell Machinery

Systems biologists have made consistent progress in identifying strategies to improve cellular adaptation. An insightful publication by James Ferrell correlates adaptation to “conventional” cell-signaling, and highlights mechanisms through which cellular signals underlie adaptation to specific stimuli [12]. One of the established ways of adaptive response to a signal is called negative feedback. Negative feedback requires at least two proteins (e.g. A and B) involved in a negative loop (A induces B; B inhibits A). The adaptive mechanism in this scheme ideally reinstates pre-stimulus conditions regardless of external input. In physical systems, negative feedback is
especially relevant in cases where the performance of the hardware components cannot be further improved [13]. In a biological system, the “hardware components” are intrinsically less amenable to modular modification; therefore, refinement of negative loop feedback control becomes central. Theoretical research has pointed out how the best way to achieve negative feedback-mediated adaptation is to induce ultra-sensitivity of protein B to protein A [14], which is in turn obtainable by keeping the interactions with B always close to saturation [15].

Bacterial chemotaxis is an example of cellular negative feedback loop. For the bioremediation of environments where bacterial mixing is critical yet hard to achieve, bacterial chemotaxis may play an essential role. Such is the case for the decontamination of drinking water supplies in groundwater basins, which poses a challenge both for the abundance of recalcitrant pollutants and the difficulty to obtain homogenous dispersion of the bioremediator [16].

One of the most studied communication mechanisms is the acyl-homoserine lactone (AHL) quorum-sensing system. A tool to identify orthogonal AHL-mediated signals in silico has been recently developed and experimentally validated with E. coli co-cultures that simultaneously employed up to three independent communication channels [17]. This is a relevant step forward towards methodological assembly of microbial consortia and, consequentially, towards the design of a multifunctional biological system.

2.3. Community Assembly Enables Synergistic Division of Labor

Assembly of a complex, multifunctional biological system is a way to circumvent the limits of single, monoclonal populations, usually unfit at multitasking [18,19]. The process of community assembly around a desired function is the most consistent way to achieve a functional combination of metabolic activities [20], but hardly a straightforward one. Several theoretical studies have identified potential strategies for constructing communities, including incorporating cooperation among intended members [21] or constructing larger communities from smaller coexisting sets [22]. Systems biologists have hatched interesting observations of powerful simplicity [23] that can guide the development of bioremediators. Coordinated performance of engineered consortia also relies on fine understanding of species-species and species-environment interactions [24,25], including metabolic influences and cellular communication.

A remarkable advance has been the development of two-strain microbial consortia that, through modular pathway reconfiguration, can realize any of the six different social interaction modes [26]. The authors posit how modular pathway reconfiguration can be achieved through modification of native gene clusters, and accordingly altered the nisin and lactococcin A (lcnA) pathway of Lactococcus lactis, another quorum-sensing mechanism. This is a notable progress in rational ecosystem design with implications of relevance to the engineering communities aimed at specific functions, bioremediation included.

Minimal Interspecies Interaction Adjustment (MIIA) has recently been formalized [27]. MIIA shows that pairwise interactions are minimally affected by the introduction of other members. Moreover, the model highlights how interactions are strongly altered when perturbed by a small
number of species, but gradually resemble the original pairwise outcome when the number of perturbing species increases. We believe this latter trait, if proved consistent across different experimental cases, surprisingly implies that community assembly could become more and more modular as the number of interacting species increases.

3. The Molecular Approach

Rather than cells as the unit of selection, the molecular machinery of biological systems can also be the target of design for efficient bioremediation. In particular, identifying and enhancing enzymatic activities towards specific applications can have a great impact on bioremediation. One of the most rewarding techniques is arguably Directed Enzyme Evolution (DEE). DEE relies on the generation of a library of random mutants empirically screened for their efficacy at a specific function of interest. Alternatively, a mechanistic, de novo enzymatic design (DED) of the active site can be attempted on the basis of preexistent, specific knowledge of the catalyzed reaction, its molecular mechanisms and accessory interactions [28].

We highlight two techniques, Small Angle Neutron Scattering (SANS) and Density Functional Theory (DFT), that—Independently and in combination—can help invigorate rational enzymatic design. In the following, we will illustrate how SANS and DFT could inform enzyme design in a practical example of removing the pollutant Aflatoxin B1. We are working on employing these methods to improve fungal oxidases, including multicopper oxidase laccase, as a promising category of enzymes for bioremediation. (Box 2).

3.1. Small Angle Neutron Scattering Reveals Enzyme-Substrate Interaction

A mechanistic investigation of the molecular structure of the bioremediation machinery can yield important information. An established technique for the extraction of low-resolution structural information of biological samples is SANS [29–32]. In conjunction with Small Angle X-rays Scattering (SAXS), SANS has been widely used to characterize macromolecular structures and their interactions. SANS uses low-energy thermal neutrons with wavelength and energy ranges which can resolve information on the nanometer to micrometer length scales, making it a well-suited technique for the study of the mesoscopic structure of proteins, enzymes, and complex macromolecules in a variety of phases. In a SANS experiment, information about the spatial arrangement of the secondary structure of the assembly is encoded in two factors: the particle form factor P(Q) and the structure factor S(Q), Q being the neutron momentum transfer [33]. These functions can be calculated using analytical models and numerical tools, such as ab initio or Monte Carlo methods. P(Q) reveals observables such as the size of the constituent, gyration radius, and molecular mass, whereas S(Q) reveals the particles’ shapes and interactions, the aggregation states, and the distribution of particles in solution.
Compared to SAXS, SANS offers additional advantages. First, neutrons can distinguish different isotopes in a macromolecular structure, in particular hydrogen and deuterium in biological samples. Using partially or entirely deuterated samples, atomic isotopic substitution and contrast variations can be used to label particular regions of complex structures [34,35]. Second, the penetrative and non-destructive scattering of neutrons within the sample enables real-time monitoring of kinetic processes, time-resolved studies, and longer exposure times. With recent technological advances, it is now possible to work online with time-resolved structural data SANS experiments coupled with a variety of light scattering techniques, such as fluorescence, UV or chromatography [36,37]. This combination enables monitoring of the kinetics of the process while providing access to complementary structural information. This simultaneous probing is important for a reliable interpretation of a biophysical process. Even though the data extracted from a SANS experiment is specific to the low-resolution nanometric scale, careful interpretation of SANS data can offer the necessary insight for rational, molecular design.

In relation to our case-study, it is known that laccase is able to perform degradation of aflatoxins [38,39]. One could begin by determining the atomic positions of the enzyme and toxin. Such a mechanistic description can be validated or refuted by analyzing SANS/SAXS data, which provide insights on the actual conformation that the enzyme will have during the biological process. Various secondary structure conformations of laccases can be compared with the experimental curves to confirm the overall structure and identify the mechanistic model to be employed for the rational approach. To perform such a comparison, it is fundamental to have theoretical analysis of the Small Angle Scattering spectra, with tools like e.g. the SASSIE suite [40].

BOX 2. Laccase: a multicopper oxidase as the ideal training ground for QM-informed bioremediation. Laccase activity allows fungal species to degrade lignin. Lignin is among the most complex whilst abundant natural occurring polymers, and it constitutes about 30% of the organic carbon in the biosphere [52]. For fungal species, lignin degradation at the same time facilitates plant invasion and grants access to a ubiquitous carbon source, over which competition is limited. To this end, fungi have evolved very sophisticated forms of task-specific laccases.

Laccases are ancient, interkingdom enzymes with a conserved active site, while the rest of their structure is highly variable [53]. The fungal isoforms dislodge the most resistant entities in the lignin molecule, the aromatic moieties, thanks to the interactor-mediated action of a large and flexible active site. In the never ending “arms race” between plants and their fungal pathogens, the former resist the attack by evolving reshuffled lignin structures (lignin is multiform: e.g. lignans), and the latter respond by always selecting novel laccase isoforms to keep up [54,55]. Laccase has thus proved to be remarkably versatile also to the point of function-reassignment; a well-documented example is repurposing laccase in Vitis vinifera’s defense against its natural enemy, Botrytis cinerea [56].

From an evolutionary standpoint, the observed plasticity, along with an undisputable potential for industrial applications, encourages laccase employment as the workhorse through which a fully mechanistic, QM-based approach to the formalization of enzymatic dynamics can enable meaningful and impactful steps. It is of great interest to develop a method to exploit the natural “malleability”, as testified by millions of years of host-pathogen coevolution, of the laccase off-active site structure to tailor high-affinity isoforms towards specific pollutants. In particular, laccase ability to break aromatic moieties makes it promising against the most recalcitrant hydrocarbons. On laccase and its potential applications, existing work [48, 57-61] can be used as a solid ground for further developments.
3.2. Quantum Mechanics Modeling Offers Mechanistic Insight into Enzymatic Activity

In combination with high quality experimental data, *in silico* models can provide valuable insights into the properties of different structures. Among the various mechanistic modeling techniques, Density Functional Theory (DFT) [41,42] is a quantum mechanical approach that has been successful in predicting properties of systems with up to a few hundred atoms. DFT has been, for more than twenty years, the workhorse method for simulations within the solid-state physics and chemistry fields. DFT is promising for applications in biology; however, its application has long been limited to mixed quantum-classical approaches [43] due to its high computational cost. In recent years, thanks to the development of linear scaling (LS) approaches [40,44], its scope has been expanded to larger systems with thousands or even tens of thousands of atoms. Such LS methods open up the possibility not just of treating larger systems, but also new types of material systems and calculations [45], including biological [46]. This capability is now extending the range of possible applications to new fields which are focused on larger systems. Similar to how DFT spread in the Quantum Chemistry community [47], we are entering a second era of DFT calculations, where large-scale quantum mechanical treatments are being rapidly adopted in other fields.

In this context, it is of great interest to investigate how information derived from DFT can be employed in the understanding of the quantities relevant to bioremediation. The current state of the art in the modeling of proteins relies on classical force field based methods of varying accuracy to predict active sites and to simulate the equations of motion. Mechanistic models based on DFT, however, are able to go beyond classical models and extract specific physical observables that represent the distribution of electrons in the system. Such a rational understanding brings physics, chemistry, mathematics, and biology together organically and offers a novel perspective for enzyme engineering. Such an approach will enable the mechanistic design of enzymes, and can be combined with a mixed -omic/heterologous expression approach for high-throughput enzyme engineering.

Going back to our case-study of aflatoxin bioremediation, after validating the enzyme model, one may proceed with quantum mechanical modeling. The motivation for such an investigation is that information about the distribution of electrons will indicate the degree of oxidation of the toxin [62]. Quantum mechanical methods can directly compute the charges on atoms in a system in order to quantify the degree of oxidation for a given set of atomic positions. In addition, while the classical modeling of docking can predict the binding energy, quantum mechanical modeling is able to provide a more detailed picture of binding positions. Identifying the toxin-enzyme orientation configuration that oxidizes the toxin the most provides a hint that efficiency can be improved by redesigning the enzyme to make the more oxidized orientation also a favorable binding position. This principle of engineering the substrate arrangement in the active site is in agreement with a similar finding for laccase made by Monza *et al* [48] using a mixed quantum-classical approach. From these calculations, one can begin to develop a mechanistic understanding of the efficiency limitations of this enzyme, and how it can be rationally engineered and improved.
5. Final Thoughts

Cellular and molecular approaches offer a promising perspective for achieving efficient bioremediation. The distinction between these approaches lies in the biological entity that is the subject of the design. In the cellular approach, one begins with cells that—on their own or in conjunction with other cells/species—have the bioremediation capability. The constraints of cells limit the range of possibilities, but also offer resilience and tolerance to changes. The path forward involves improving this capability, with synthetic biology and natural and artificial selection being the main tools. In the molecular approach, the starting point is the molecular machinery that achieves the function of interest. The path forward in this case involves evolving or designing novel molecules of interest.

In the cellular approach, general procedures for natural or artificial selection exist; recent work in community ecology has also boosted the prospect of assembling bioremediation communities of interest. However, developing a systematic roadmap to compile a community with a particular function still remains challenging. In the molecular approach, abundant information about a given system can be extracted, but the question of how to go from insight to novel designs (sometimes called an inverse problem) remains an open challenge. The ultimate solution will combine cellular and molecular approaches to overcome the limitations of each.

**Fig 1.** Comparison between cellular and molecular approaches. Solid black lines represent established methodologies, whereas the dotted red lines designate open questions.
Acknowledgements

This work was supported by the Next-Generation Supercomputer project (the K computer) and the FLAGSHIP2020 project (Supercomputer Fugaku) within the priority study5 (Development of new fundamental technologies for high-efficiency energy creation, conversion/storage and use) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. L.G., N.T., and W.D. also gratefully acknowledge the joint CEA–RIKEN collaboration action. Work in the Momeni Lab is supported by a start-up fund from Boston College and an Award for Excellence in Biomedical Research from the Smith Family Foundation. LER acknowledges an EPSRC Early Career Research Fellowship (EP/P033253/1).

6. References


31. Svergun DI, Koch MHJ: Small-angle scattering studies of biological macromolecules in solution. Reports Prog Phys 2003,


