Capabilities

Project Task Area	Capability	Description
DESIGN	DIVA BioCAD	DIVA is a software platform that brings together a tool-chain that enables the design of (combinatorial) DNA libraries and automates the generation of protocols to construct them. It brings together a tool-chain of DNA component repositories and visualization tools (e.g., ICE/VectorEditor), visual design tools (e.g., DeviceEditor), DNA construction optimization tools (e.g., BOOST, j5), and biosecurity screening tools (e.g., BLiSS), and researcher DNA construction request consolidation and tracking (DIVA). It enables the DIVA DNA construction and DIVA DNA sequence validation capabilities, integrates downstream with the Experiment Data Depot, and importantly captures design intent (of great importance to downstream statistical and mechanistic learning/modeling approaches). Digital DNA sequence sources may be constrained by supported data formats (e.g., FASTA, Genbank, SBOL v1.1 and v2.0 are supported) and sequence length (portions of the tool-chain may not yet well support genome-scale sequences). DIVA is available to ABF researchers, as well as to ABF CRADA projects. Some components (e.g., ICE, VectorEditor, BOOST, j5) are accessible externally to the ABF (potentially through licensing or through commercially-available offerings).
BUILD	Genetic Transformation and Tool Development	The ability to genetically modify microorganisms is foundational to performing rational strain engineering, but most organisms are not transformable when first isolated. We are able to develop initial transformation methodologies and protocols for new organisms or improve upon existing ones. We can further develop high throughput tools for more rapid strain modification in these new organisms, including phage integrase-based genome integration tools for rapid screening of gene expression libraries and CRISPR-based genome editing tools for rapid, targeted gene deletion or gene regulation. Transposon mutant and overexpression libraries can also be created to allow high throughput phenotype screening or selections. Promoter libraries can be created to enable rational selection of gene expression levels for heterologous expression. The ABF has deep expertise in non-model organisms, including work with extremophilic organisms that thrive at

		elevated temperature and extreme pH, including both aerobes and strict anaerobes. We are able to improve transformation efficiency and developed genetic tools for diverse organisms to accelerate rational strain engineering. Development of new microbial platforms or further improving genetic tools in existing platform organisms will provide unique new chassis organisms for advanced bioprocessing to enable a bio-economy.
B U I L D	Fungal Genomics, Synthetic Biology and Bioprocess Development	The ABF houses the filamentous fungal and non-model yeast synthetic biology capability unique within the government laboratories. This effort also comprises integrated fungal genetic tool and medium scale (15-120L) bioprocess development capability. Samples for Testing and Learning by a systems biology approach are generated from large shake flask experimental capabilities or through one of our stirred tank reactors that are highly controlled and have advanced monitoring and sampling capabilities, such as: Transformation systems and parts (promoters, selection markers, etc.) for eight fungal species. Standard modern molecular biology laboratory equipped with all of the thermal cyclers, incubators, electoporators and analytical equipment to generate constructs and verify their insertion sites and DNA sequence. Bioreactors with working volumes from 15-120 L for testing genetically engineered strains, for collecting samples from highly controlled and monitored environments and for testing bioprocesses at an intermediate scale. The limit of filamentous fungal systems generally is the throughput with which genetic changes can be made and tested. Filamentous fungi are less amenable to 96 well culture (1 mL) testing that is translatable back up in scale. Maximum stirred tank reactor volume is 120L. This capability is primarily applicable to one of our three main hosts, Aspergillus pseudoterreus but also is beneficial in working with SNL on the basidiomycete yeast, Rhodosporidium toruloides, and other hosts within current CRADA projects and future ABF users.
B U I L D	DIVA DNA Sequence Validation	Comprises sequence validation of plasmid and/or amplicon DNA using the Illumina MiSeq platform. Customized ABF workflow minimizes library preparation steps and leverages laboratory

		automation to enable scale/reduce reagent costs, provides full plasmid coverage, and is integrated with the ICE sequence/strain repository platform. This validates DNA sequences (ensures that they actually are what they purport to be) at effective cost points and turnaround times and prevents incorrect materials from being wastefully investigated or processed further. The process consists of weekly cycles of 384 samples (individually barcoded and then pooled) each. Samples can consist of purified plasmids, PCR amplicons, or (plasmids in) bacterial cultures. There are some constraints around acceptable resuspension buffers, minimum concentrations, minimum volumes, etc. Available internally to ABF researchers, as well as to ABF CRADA projects.
BUILD	DIVA DNA Construction, Synthetic Biology and Bioprocess Development	This capability relates to physical construction of sequence-validated plasmids. The workflow leverages laboratory automation to enable scale/reduce reagent costs, sequence validation provides full plasmid coverage, and is integrated with the DIVA/DeviceEditor/BOOST/j5 design, BLiSS biosecurity screening, and ICE sequence/strain repository platforms. This enables ABF researchers to focus their efforts where they are competitively advantaged (upstream and downstream of DNA construction). ABF operations are made more efficient through the pooling of DNA construction tasks across ABF researchers at scales that significantly benefit from laboratory automation devices. DNA fragments requiring de novo DNA synthesis are constrained by DNA synthesis vendor (e.g., Twist, ThermoFisher) capabilities. Extant physical DNA must be sequence validated before input (as digest substrate or PCR template) into the construction process. Each DNA assembly reaction may contain at most 5 assembly pieces (including the vector backbone); more than 6 assembly pieces requires a hierarchical (multi-step) construction process. PCR reaction size is capped at 5 kb; larger fragments need to be partitioned or a digest strategy may need to be pursued instead. Certain DNA assembly methods have specific requirements (e.g. lack of internal type IIs restriction sites). This capability is available internally to ABF researchers, as well as to ABF CRADA projects.
TEST	Targeted Proteomics	Targeted proteomics assays are applied to specific hypotheses via selected-reaction monitoring (SRM) mass spectrometry methods that accurately quantify a select group of proteins. This approach has been applied to determine pathway bottlenecks, compare enzyme homologs for a specific enzymatic step, and to track dynamic regulation of protein levels engineered microbes. For

		instance, application of targeted proteomics methods helped identify protein-associated bottlenecks in the mevalonate pathway, resulting in over 300% improvement in the final product levels. This technique is also used to quantify protein levels for tool genetic development studies such as engineered promoter and ribosome-binding site (RBS) variants. More recently, targeted proteomic methods have been optimized for greater throughput (over 250 samples/day) as well as used to characterize and quantify stable post-translational modifications for engineered microbes. High-throughput targeted proteomic methods for engineered microbes utilize laboratory automation to reduce sample preparation variability and leverage tight integration with the ABF ICE sequence/strain repository platform and Experiment Data Depot (EDD). The targeted proteomic capability supports the ABF through rapid testing of engineered microbes to reduce the number of DBTL cycles required to successfully produce target molecules. Targeted proteomic assays are limited to quantify fewer than 75 proteins in high-throughput applications, but can accurately quantify hundreds of proteins for lower throughput experiments. The method is also requires that unique peptides for a specific protein can be targeted. Each protein and peptide has different limits of detection, so sensitivity ranges will vary depending on the experiment. The targeted proteomics methods are available to ABF researchers as well as ABF CRADA projects.
T E S T	Targeted Metabolomics	Targeted metabolomic analyses at the pathway and organism level provide functional information for both pathway and host-engineering research. Tools for strain improvement, such as metabolic flux analysis rely heavily on accurate metabolite data and carbon-flux measurements to restrict parameters for model predictions. Monitoring metabolites that are part the engineered pathway as well as central carbon metabolism aids identification of bottlenecks, and helps identify where increasing specific protein levels can yield dramatic improvements to the product titer or where allosteric regulation is limiting flux through the pathway. These assays are conducted by using an ultra-high-performance liquid chromatography quadrupole time-of-flight mass spectrometry (UHPLC-QTOF-MS) system to rapidly acquire high-resolution, high-mass accuracy mass spectra and structural information of metabolites. High-throughput targeted metabolomic methods for engineered microbes utilize laboratory automation to reduce sample preparation variability, high-sensitivity UHPLC-MS detection, and leverage tight integration with the ABF ICE sequence/strain repository platform and Experiment Data Depot (EDD). Rapid quantification of metabolite levels in

		engineered microbes benefit the ABF by reducing DBTL cycle time. The sensitivity of targeted metabolomic assays are limited by the number of cells used to extract the metabolites, the amount of the metabolites inside the cell, and the availability of metabolite standards. Furthermore, metabolites have different physiochemical properties so chromatographic performance and limit of detection will vary depending on the metabolite assayed. Available internally to ABF researchers, as well as to ABF CRADA projects.
T E S T	Riboregulators for Precise Control of Gene Expression	Riboregulators of various classes have been developed to allow for the tuning of gene expression. The two classes we have designed are: cis-repressor/trans-activator pairs, and tuned standalone cis-repressors. The constructs have been designed to be universal irrespective of gene sequence. Cis-repressing riboregulators can be placed in front of genes in the genome of a target organism in their native context to regulate expression of the gene from very low to near native levels. Cis-repressors that completely repress expression have also been developed to work with trans-activator RNA (taRNA) elements. Here, specific taRNA sequences target a cis-repressed gene and activate expression to designed levels. Multiple cognate pairs can transcribed from a vector and be used to control multiple genes in a pathway and tune expression. Libraries of taRNA with differing activation profiles can be selected for a particular phenotype. These RNA constructs have been demonstrated to provide very consistent regulation independent of gene sequence with greater than 12 discrete levels available. These constructs allow for the creation of gene pathways with expression tuned to improve growth and productivity of engineered strains. Cis-repressors must be inserted into the genome using standard techniques for genetic manipulations limiting the diversity of regulatory states. Targeting sequences of taRNA/cisRNA pairs must be validated as neutral and specific to ensure consistent activation, also activation is dependent on the concentration and therefore will vary with vector copy number and promoter. Sequences and help with design and implementation are available without restriction.
T E S T	Microfluidic Screening	Droplet-based microfluidics is used to survey cell-based or cell-free expression libraries of the 104-106 size in picoliter aqueous droplets to identify rare events. The system can identify enzymes with promiscuous activities and can be used for protein engineering. The droplet-based microfluidics complements FACS-based methods, and supports high-throughput protein engineering. A

		fluorescent-based reporter assay is necessary to perform the survey. This capability is available for collaborations with the ABF.
T E S T	Global Metabolomics and Isotopically-Labelled Metabolomics for Metabolic Flux Analysis	Metabolomics is another key aspect of an integrated systems biology approach to understanding engineered and basal strains under different conditions to enable learning that leads to the next set of Designs for DBTL. Global metabolomics is the derivatization GC-MS analysis of the intracellular metabolome. The extracellular metabolites can be analyzed separately to assess substrate utilization and secretion of target molecules as well as side-products. 13C Metabolic Flux Analysis (MFA) involves the use of 13C labeled substrates and subsequent metabolomics and metabolic modeling analysis discussed in the Modeling and Machine Learning Capability to determine fluxes through pathways. LC-MS and NMR methods are techniques complementary to our standard GC-MS metabolomics analysis. They are useful for examining large hydrophilic molecules and can be utilized for verification of structures of complex or novel intermediates or target molecules. There are three GC-MS instruments devoted to metabolomics and one LC-MS. High resolution instruments for isotopically labelled metabolite assignments, enabling 13C-MFA. Multiple NMR instruments of moderate to high fields with diverse probes available. Preference for this activity is ~50 mg of wet cell weight per sample, quadruplicate biological replicates, and sample sets of 48-60 though twice that number can be accommodated.
T E S T	FRET Biosensors	FRET-based biosensors couple ligand-dependent protein conformational changes to a fluorescent signal response. Such biosensors can be used intra- or extracellularly to provide rapid, real-time responses to ligands or metabolites of interest. Unlike typical reporter systems, signals from FRET biosensors are reversible, permitting detection of anabolic and catabolic (conversion) processes and variations in metabolite fluxes. FRET biosensors can be constructed from a variety of metabolite binding proteins. A suite of fluorescent protein scaffolds has been constructed to rapidly identify FRET-responsive ligand binding domains and for optimization of FRET biosensors.

		The biosensors are ratiometric and respond to changes in the emission characteristics of the sensor upon the binding of a ligand. Signals are independent of biosensor concentration. They perform in vitro as well as in vivo. This approach is useful for high-throughput screening and sorting and is adaptable for use in the monitoring of bioreactor titers in real time. A prerequisite is the need a protein assembly that changes conformation upon binding of a desired ligand. Libraries of these exist in nature in the form of solute binding proteins and technology has been recently adapted to domains of certain types of transcription factors. Sensors for a variety of metabolites are available and new sensors directed towards new metabolites could be readily developed through collaboration.
TEST	Experimental Data Depot	EDD is an online tool designed as a repository of experimental data and metadata. EDD can uptake experimental data, provide visualization of these data, and produce downloadable data in several standard output formats. The input of data to EDD is performed through automated data streams: each of these input streams automatically parses the standard outputs of the instruments most commonly used for bioengineering. New input streams can be easily added to adapt to local data production. EDD provides a quick visualization of imported data that allows for a quality check by showing whether the imported data are within the expected range or not. Since data are stored internally in a relational database, all data output is consistent. Outputs can be provided in terms of different standardized files (Systems Biology Markup Language, SBML, or CSV) or through a representational state transfer (RESTful) Application Programming Interface (API). SBML and CSV files can be used in conjunction with libraries such as COBRApy or Scikit-learn to generate actionable results for metabolic engineering. EDD is the only tool that provides a single data repository for all -omics data types which is able to extract data straight from instrument output, visualize this data, and export the data in formats that are readily applicable to modeling tools and libraries. As all data is collected in a standardized form, EDD enables the application of machine learning and mechanistic models to learn from data and guide the metabolic engineering process in a systematic fashion. Code is available to anyone as part of an open source project. The EDD website is internally available to ABF researchers.
TEST	High Throughput	While ‘Design’ and ‘Build’ steps of DBTL cycle for bio-based products show high throughput, the

	Screening platform for enzymes, pathways and strains using 'Smart' Microbial Cell Technology	'Test' step remains a bottleneck due to low efficiency of the current screening technologies. We have developed Smart Microbial Cell technology that utilizes a custom-design sensor-reporter system for a given small molecule of interest (value-added chemical, industrial precursor) in the host strain harboring variations in a single gene, or a pathway, or at the genome level. Based on the production efficiency of enzyme/pathway/strain (determined by genotype), there is a correlated response from the sensor-reporter system (phenotype), offering an easy screening technique using flow cytometry or by simply plating them on a solid phase and visualizing under an illuminator. This capability is unique as it provides the ability to incorporate computational protein design to engineer sensors or intelligent library design for evolution of enzymes. The existing capability enhances the efficiency of the 'Test' step of DBTL cycle for bio-based products. Product strains can be optimized in a high throughput manner. 102-106 variants can be screened in a single day.
T E S T	Biolector and Robolector automated fermentation and sampling platform	The Biolector is a microbioreactor for high-throughput fermentations with the online monitoring of biomass, pH, DO, and fluorescence. The RoboLector combines the high-throughput fermentation and online monitoring capability of the BioLector with the precise and accurate liquid handling of a robotic system. The RoboLector autonomously prepares media compositions, among others, from the design of experiments (DoE) tests, followed by an online-monitored high-throughput fermentation. This combination of Robolector and Biolector provides the capability to automatically do fermentations and sampling, ensuring repeatability and data quality. The constraint on these equipment is on the Biolector, which is limited to 48 fermentations at a single time. These, in turn, enable predictive quantitative modeling of biological processes. Available internally to ABF researchers, as well as to ABF CRADA projects.
T E S T	Biocatalyst Optimization	A microfluidic platform is utilized for rapid enzyme engineering relevant to identify improved and/or missing catalysts in production pathways. This approach is compatible with aerobic/anaerobic enzyme screening (in vivo or cell-free). The system is capable of utilizing TF-sensors or FRET and enzyme-linked sensors. Once candidate genes have been identified, variants are synthesized using rational design/directed evolution approaches. Appropriate sensors are linked to monitor reaction improvements, and in-depth characterization of candidate enzymes is

		carried out in vitro and in target host systems. The pairing of microfluidics and biosensor design/engineering for molecular adaptation is only possible in a few unique research environments. It can be used to eliminate bottlenecks in metabolic engineering. Improved biocatalysts can be incorporated into production strains to improve titers, rates, and yields of bioproducts. This approach is aided by the availability of biosensors and is most successful with target proteins where (i) structures can be inferred from database resources or (ii) functional correlations between active and inactive enzymes can result from the study of evolutionary relationships in sequences of large numbers of homologs. Ready to use through collaboration. The strategy works best if able to be approached with a structural perspective or enzymes from large structural or functional classes.
L E A R N	Regulatory Modeling	The amount of information that can be collected by a multi-omics analysis of fermentation (e.g. combined metabolomics, transcriptomics, and proteomics experiments) or the information that can be collected from sensors in an industrial bioprocess can easily amount to terabytes of data. Accommodating this massive outpouring of information is one of the defining challenges for the computational analysis of biological systems. The goal of Deep Learning is to effectively analyze these data to make meaningful predictions of the behaviors of complex systems and to identify specific actionable interventions for productive alterations to maximize process outputs. Deep Learning is an Artificial Intelligence approach that accomplishes this by mimicking the activity of layers of neurons in a living brain to solve otherwise intractable computational problems. Applications of Deep Learning to sustainable industrial bioprocesses include analysis of bacterial genomes for the rational genomic engineering for improved efficiency in industrial applications or for the rational design of novel biological processes for new bioproducts. Statistical modeling of biological systems using machine learning and artificial intelligence approaches strongly compliments the results of mechanistic biological models. Deep Learning directly benefits ABF by creating a bridge between the laboratory and computational analysis, streamlining the process by which meaningful information can be extracted from biological observations and rapidly turned to actionable information that can help guide continuous improvements to developing bioprocesses or the real-time optimizations of ongoing industrial fermentations. Deep learning requires substantial amounts of raw data for training and validating AI. This restriction

		can be partially relieved by considering other model-generated data type (e.g. metabolic models or gene regulation predictions) as input in addition to direct biological observations.
L E A R N	Pan Genome Analysis	Pan-genomic analysis leverages recent advances in the ability to routinely sequence and annotate whole genomes. No longer restricted to considering one genome at a time, the capability now exists to study multiple genomes from the same taxonomic grouping and observe how genomes change with evolutionary time and with adaptations to new environmental niches. It is possible, for example, to trace the evolution of a single enzyme as structure changes to fit function. However, the role of any genes in living systems is highly pleiotropic, its function dependent upon the interactions and regulation of many other genes. Through pan-genomic analysis we have identified that a microorganism's characteristics is not well predicted by any individual gene or gene function, rather a microorganism's role is best predicted as an emergent property of its genome, i.e. its metabolome and its transportome. Our unique computational approach to modeling a microorganism's transportome, PRTT, has proven to be most predictive of a microorganism's function. This insight had provided specific motivation for protein production and functional characterization. Pan-genomic Analysis leverages previously sequenced annotated microorganism genomes to identify, not only specific gene functions, but also interacting gene networks that are associated with desired biological functions, such as the capacity to consume specific carbon sources. Unique capacity to model secondary metabolome and transportome lead to unique, non-central metabolism targets for genomic manipulation. Pan-genomic Analysis is limited only by the number of available genomes and the knowledge of microorganisms' native habitats.
L E A R N	Metabolic Flux Analysis	Building and employing models for learning form integrated systems biology data is an important capability within the ABF. Pathway genome databases are developed for host strains and used in the building of metabolic models for quantitative analysis of in vivo carbon fluxes in metabolic networks, i.e. intracellular activities of enzymes and pathways. 13C Metabolic Flux Analysis (MFA) involves culturing strains on carbon sources labeled at specific positions with 13C followed by GC-MS metabolomics to determine the labeling pattern of

		metabolites. The labeling data is compared to a computational simulation using a flux model that combines isotopomer and metabolite balancing representing the investigated metabolic network. The flux estimation is based on minimizing the deviation between the measured and the simulated labeling data. In combination with experimentally determined extracellular fluxes, 13C MFA determines the crucial fluxes of substrates through intermediates to target molecules in the metabolic network. This is valuable for understanding bottlenecks and side reactions to maximize efficiency of carbon utilization and hence maximize TRY (titer rate and yield) in the conversion of substrates to target products in our ABF host strains. Genome-scale metabolic models (GEMs) summarize the known metabolic information in a mathematically defined reaction network. A key objective of utilizing GEMs for metabolic engineering is to improve the production of target products in our ABF host strains by providing genetic targets to be deleted, down-regulated or over-expressed. One method employed in the ABF is OptForce, which identifies potential gene targets by classifying metabolic reactions with regard to necessary changes in flux (i.e., increase, decrease, or zero) to meet a pre-specified overproduction target. The efficiency and quality of engineering interventions predicted by OptForce and related methods depends on quality multi-omics, cell mass and other meta-data for the wild-type strain. The modeling capabilities are a tremendous benefit to the ABF with regard to incorporating and interpreting the multi-omics and meta data to provide actionable genetic engineering targets for the next round of DBTL. Development and use of models requires the time of domain experts. The types of isotopically labeled substrates are limited and expensive. Biological replicates within very well designed experiments are crucial to getting statistically robust predictions.
L E A R N	Mechanistic Models of Metabolic Flux	Two-scale 13C Metabolic Flux Analysis (TS-13C MFA) measures intracellular metabolic fluxes by using data obtained from 13C substrate labeling experiments and genome-scale models. Metabolic

		fluxes map how carbon and electrons flow through metabolism to enable cell function and provide the best available insights to guide metabolic engineering. This is the only 13C labeling technique that has been used to produce flux measurements for genome-scale models and improve bioproduct yields. TS-13C MFA provides a mechanistic modeling approach that is proven to guide metabolic engineering efforts to improve bioproducts yields. TS-13C MFA is an accurate but slow capability that may take several months to produce results. Currently we can process 4-5 strains per semester. Code is available as part of the open source jQMM library. Expertise is available internally to ABF researchers, as well as to ABF CRADA projects.
L E A R N	Machine Learning	Often we do not know all the biological mechanisms required to predict the behavior of a bioengineered system through mechanistic modeling (e.g. how does induction time or amount affect flux through a pathway). In those cases, a data-intensive, statistics-based approach can still predict bioengineered systems to the degree required to drive metabolic engineering efforts. We use a variety of machine learning approaches (ranging from scikit learn to deep learning) for this purpose. We have used these techniques to tie proteomic profiles to production and suggest improved proteomics profiles, as well as to predict pathway dynamics. The combination of machine learning techniques with the ability to produce our own data in an automated fashion using the Biolector & Robolector is unique within the national labs. Machine learning provides a systematic method to leverage data stored in the Experiment Data Depot (EDD) to guide metabolic engineering methods without the need for a deep mechanistic understanding. This data-intensive approach is usually limited by the availability of data. Typically, 50-100 conditions (strains) are needed for successful predictions, as well as 3-4 consecutive rounds in which predictions are tested and the new data is used to improve predictions for the next round. This capability is available internally to ABF researchers, as well as to ABF CRADA projects.
L E A R N	Deep Learning	The amount of information that can be collected by a multi-omics analysis of fermentation (e.g. combined metabolomics, transcriptomics, and proteomics experiments) or the information that can be collected from sensors in an industrial bioprocess can easily amount to terabytes of data. Accommodating this massive outpouring of information is one of the defining challenges for the computational analysis of biological systems. The goal of Deep Learning is to effectively analyze

		these data to make meaningful predictions of the behaviors of complex systems and to identify specific actionable interventions for productive alterations to maximize process outputs. Deep Learning is an Artificial Intelligence approach that accomplishes this by mimicking the activity of layers of neurons in a living brain to solve otherwise intractable computational problems. Applications of Deep Learning to sustainable industrial bioprocesses include analysis of bacterial genomes for the rational genomic engineering for improved efficiency in industrial applications or for the rational design of novel biological processes for new bioproducts. Statistical modeling of biological systems using machine learning and artificial intelligence approaches strongly compliments the results of mechanistic biological models. Deep Learning directly benefits ABF by creating a bridge between the laboratory and computational analysis, streamlining the process by which meaningful information can be extracted from biological observations and rapidly turned to actionable information that can help guide continuous improvements to developing bioprocesses or the real-time optimizations of ongoing industrial fermentations. Deep learning requires substantial amounts of raw data for training and validating AI. This restriction can be partially relieved by considering other model-generated data type (e.g. metabolic models or gene regulation predictions) as input in addition to direct biological observations.
T E A / L C A	Techno-Economic Analysis	Techno-economic analysis (TEA) is a powerful tool that quantifies the impact that research and development have on the economics of an integrated process. TEAs are performed to evaluate the potential economic viability of a process. Additionally, TEAs can be used to 1) identify cost drivers of an integrated design and help set specific targets/metrics that must be achieved to support an economically attractive process, 2) outline data gaps that must be addressed to understand the process economics and 3) identify specific uncertainties associated with process design or process performance that may be problematic for process scale-up. Finally, TEA can be integrated with sustainability analyses to help quantify the environmental benefits and burdens associated with the implementation of the researched technologies. The ABF has the capabilities for developing dynamic fermentation models for aerobic production (e.g. Bubble columns) for process design, scale-up/scale-down, monitoring, and control studies. Specific areas of competence include compartmental reactor models to explore mixing effects (e.g. oxygen gradients) and exploration of rate-based phenomena (e.g. mass transfer, cell kinetics). Models are developed in house with

		flexible of tailoring to unique process and performance attributes. Understanding technical feasibility and economic potential is critical throughout the R&D and scale-up in developing new and novel products from biomass. Our analyses span the range of steady state modeling to incorporating dynamic and temporal modeling capabilities.
T E A / L C A	Life-Cycle Analysis – GREET model	The GREET model is used to evaluate environmental impacts of advanced vehicle technologies, alternative transportation fuels, and new bio-derived chemicals. With GREET, one can estimate 1) energy consumption including total energy (energy in non-renewable and renewable sources) and fossil fuels (petroleum, natural gas, and coal), 2) water consumption, 3) greenhouse gases (GHG) emissions, and 4) criteria air pollutant emissions including volatile organic compounds, carbon monoxide, nitrogen oxides, particulate matter with size smaller than 10 microns and with size smaller than 2.5 microns, black carbon, and sulfur oxides. GREET provides a holistic and transparent platform for LCA of conventional and alternative vehicle technologies and fuel production technologies, fossil- and bio-derived chemicals, and materials required for vehicle manufacturing. The GREET model is used by more than 30,000 users around the world that include government agencies, other national labs, universities, and members of the transportation industry. The results of the LCA modeling are used to inform environmental benefits (e.g. reductions in GHG emission, fossil fuel consumption, and water consumption) of bio-derived chemicals and fuels in addition to helping identify primary drivers and technical barriers along the supply chain. The GREET model includes more than 100 fuel production pathways from various energy feedstock sources including conventional fossil resources such as petroleum, natural gas, coal, as well as renewable fuel production pathways including corn, sugarcane, soybeans, cellulosic biomass, etc. Bio-products produced from biochemical, biological, and thermochemical conversion technologies are included. In addition, the GREET model evaluates not only the fuel cycle (GREET 1) impacts but also includes vehicle-cycled modeling (GREET 2) capabilities. GREET has been in public domain and free of charge since 1995. A new version of the GREET model is released almost every year with new information and new pathways. The latest version of the GREET model suite is available at https://greet.es.anl.gov.
S C A L E – U P	Fermentation Scale-up	Fermentation in bioreactors offer environments that are substantially different from shake flask

	Recovery and Purification	environments. Strains that are optimized through the design-build-test-learn cycle will be subjected to these unexpected environments and thereby, most often, lead to unanticipated results. At the ABPDU, we have been resolving this disparity between lab-scale and a commercially relevant scale of biomanufacturing. Once optimized at the bioreactor level, we conduct a fermentation to generate the culture needed to optimize downstream recovery and purification operations. Continuous centrifugation, tangential flow filtration, preparative chromatography, liquid-liquid extraction, distillation, etc. are a few such recovery and purification methods that help realize the product in a specification that are amenable to real-world applications. The ABF has substantial experience in working with multiple hosts in this context, several of which are novel. Worked with over 10 hosts including: Rhodosporidium toruloides[1], Yarrowia lipolytica, Thermoascus aurantiacus[2], Pichia pastoris, Microbial Communities[3], E. Coli, S. cerevisiae, etc. The existing capability can help "Test" optimized strains in a bioreactor and use the data to "Learn" and thereby predict scalability of a biological process. In terms of capability, the ABF has access to 2L, 10L, 19L, 50L, and 300L Fermentation Bioreactors; Disc-Stack Centrifuge (~150L/hour), Decanting centrifuge (~6m3/hour), Tangential Flow Filtration (2.5m2 filter area with ~1L/ min capactiy), Liquid-liquid extraction (2L volume column with ~150 mL/min capacity), Preparative Chromatography GE AKTA equipment (~150 mL/min).
S C A L E – U P	Multiscale Bioreactor Cultivation	The ABF has multiple scales of integrated bioreactor cultivation equipment ranging from the microscale BioLector Pro systems (2 mL cultivation volumes), more than 30 0.5-L fully controlled bioreactors, multiple 2 L, 5 L, and 10 L scales, up to 100, 1,000, and 9,000 L reactors in the Integrated Biorefinery Refinery Facility that opérate on an as-needed basis. These include aerobic and anaerobic processes and are routinely used for multi-omics investigations of non-model strains on feedstocks ranging from clean sugars to biomass hydrolysates to non-standard substrates (e.g., waste gas, lignin-derived liquor, etc.). In situ product recovery is also available for bench scale units. This capability benefits the ABF mission by enabling rapid strain evaluation and multiscale, multi-omics investigations of non-model strains.