Goal | **Breakthough Capability** | Milestone

On-demand design, generation, and evolution of macromolecules for desired functions.

***De novo* prediction of RNA structure, protein structure, and complexes of DNAs/RNAs and proteins from primary sequence and the ability to make accurate predictions of mutability and effect of mutations from structure.**			
Reliably predict the structure of 300-amino acid proteins and 200-nucleotide RNA domains within 5 Ångstroms from primary sequence.	Reliable *de novo* prediction of RNAs and proteins containing non-canonical structures.	Routine prediction of structures for 500-amino acid proteins and 200-nucleotide RNA domains within 3 Ångstrom.	Routine prediction of structures for 3,000-amino acid proteins, protein-protein and RNA-protein interactions, and protein and RNA-protein complexes.
Improve force-field and backbone-sampling algorithms and capabilities to capture force-fields of post-transcriptionally- and post-translationally-modified nucleosides and amino acids.	Routine redesign of ligand binding sites and/or aptamers for custom ligands with a greater than 50% success rate.	Design proteins that fold correctly 50% of the time and RNA-protein complexes that form correctly 20% of the time. Modeling and design of chromatin states that can be manipulated to change function.	Routine prediction of protein function from structure.

***De novo* design and/or prediction of macromolecular dynamics and dynamic macromolecular structures.**		
Improving computational models of RNA dynamics that can incorporate experimental data.	Incorporating co-transcriptional (for RNA) and co-translational (for protein) processes into design algorithms.	Routine design of enzymes with high activities.
	Design of intrinsic regulatory control into biomolecules.	
	Design of dynamic and responsive protein-RNA nanomachines.	Modeling and design of dynamic RNA nanomachines that can engage with and manipulate the chormatin states of living systems.
	Routine design of large proteins, beta topologies, membrane proteins, and loops.	Modeling and design of dynamic DNA-RNA-protein condensates that can expand beyond the functionality of natural condesates.
	Routine design of protein complexes.	

High-throughput integrated computational, experimental, and evolutionary schemes for refinement of desired biomolecule functions.			
Durable and high-mutation-rate *in vivo* continuous DNA mutagenesis and evolution systems in model organisms.	Durable and high-mutation-rate *in vivo* continuous DNA mutagenesis and evolution systems in non-model organisms.	Full control over all statistical properties of DNA diversification *in vivo*.	*De novo* DNA synthesis *in vivo* with single-cell sequence control.
		Direct sequencing of proteins and carbohydrates.	Ability to select for any function, including those conferred by: A) small molecules, lipids, or carbohydrates, and; B) proteins or nucleic acids.

Special considerations for on-demand design, generation, and evolution of macromolecules that rely on non-canonical/unnatural building blocks.

PCR, reverse transcription, cellular replication, and transcription of fully unnatural nucleotide-containing genes of up to 400 base pairs.			
Identification of "missing" functionality or functionalities in A-T-G-C base pairs.	Improved *in vitro* manipulation of unnatural nucleic acids.	Biosynthesis of unnatural nucleotides.	Establishment of organisms capable of full replication, maintenance, and transcription of a plasmid or artificial chromosome made up entirely of unnatural bases.
	Expansion of unnatural nucleotide toolkit.		

Expanded genetic code systems for translation of >100-amino acid proteins containing fully-unnatural amino acids, and proteins with at least four, distinct unnatural amino acid building blocks.			
Create proteins that are capable of gaining new, therapeutically-useful activities through unnatural amino acids.	Efficient biosynthesis of proteins containing three or more distinct unnatural amino acid building blocks.	Biosynthesis of unnatural amino acids.	Templated biosynthesis and evolution of new polymers with large user-selected sets of unnatural building blocks *in vivo*.

Holistic, integrated design of multi-part genetic systems (i.e., circuits and pathways).

Design of highly-stable, large genetic systems (genomes) with targeted expression levels in a host organism or cell type, incorporating system-wide effects.			
Incorporate gene expression interactions into predictable design of prokaryotic genetic systems.	Incorporate gene expression interactions into predictable design of eukaryotic genetic systems.	Discovery and characterization of mechanistic interactions at the systems-level affecting protein activities inside cells.	Whole-tissue or whole-cell, nucleotide-resolution simulations encompassing several layers of models predicting gene regulatory, metabolic, and system-level behaviors.

Ability to rationally engineer sensor suites, genetic circuits, metabolic pathways, signaling cascades, and cell differentiation pathways.			
Reliable engineering of genetic circuits with more than 10 regulators for sophisticated computations.	Reliable engineering of novel, many-enzyme pathways utilizing combinations of bioprospected enzymes with well-characterized kinetics.	Reliable expression of redesigned synthases to produce secondary metabolites.	Simultaneous, tunable, timed expression of many transcription factors controlling mammalian cell state.
	Five-time improvement and expansion of inducers/promoters for model organisms that respond to environmental inputs and any intracellular metabolite.		
	Utilize machine-learning approaches to use the vast amount of uncurated literature results within pathway design.	Computational design of protein-ligand and RNA-ligand interfaces suitable for engineering protein-based or RNA-based sensors.	

Integrated design of RNA-based regulatory systems for cellular control and information processing.

Porting nucleic acid strand displacement technology into cellular systems with RNA instantiations.			
RNA implementation of strand displacement cascades in bacteria.	RNA implementation of strand displacement cascades in eukaryotic systems.	Engineer computational RNA strand displacement networks in mammalian systems.	Computational design of RNA strand displacement neural networks that process the transcriptome.
	Engineer 'universal' computational strand displacement architectures using strand displacement in bacteria.		Engineer RNA neural networks that dynamically reprogram cell state.

Porting successes in computationally designed bacterial RNA-based genetic regulators into eukaryotic and mammalian systems.			
First generation eukaryotic RNA-based gene regulators that utilize RNA:RNA interactions and/or strand-displacement and achieve 10-fold change in gene expression.	Second generation eukaryotic RNA-based gene regulators that are suitable for computational design to create libraries that are highly-orthogonal and high-performing, achieving 100's-fold change in gene expression.	Expand RNA modification apparatus to modify non-natural RNA alphabets to enhance their functional properties.	Engineering enzymes that can perform non-natural RNA modifications to further expand the chemical repertoire of what is possible and extend RNA ligand recognition, catalysis and genetic control.
Creation of RNA modification machinery that allows programmable site-specific modifications of RNA, focusing on naturally abundant modifications.	Use RNA modifications for programming or fine-tuning RNA functions.		

2 Years | **5 Years** | **10 Years** | **20 Years**