TECHNOLOGY AREAS OF INTEREST
MIT-GSK Gertrude B. Elion Research Fellowship Program for Drug Discovery and Disease

Science that applies nano/microfluidic technologies to biology, increases capacity to collect high content data and advances computational approaches is evolving quickly and offering new ways to incorporate patient information into the early drug discovery process. Connecting patient and disease pathways through state-of-the-art computational modelling and experimental data, now collectible on an unprecedented scale, offers an opportunity to put disease at the core of drug discovery and drive the process for developing new medicines for patients in need.

To further pharmaceutical drug discovery platforms that will enable the next generation of medicines, GlaxoSmithKline (GSK) is pursuing three scientific areas of interest:

- **Computation Modeling and Simulation**
- **Enabling Platforms for Complex Cellular Engineering**
- **Minaturized and Continuous Biology & Chemistry Platforms**

Computation, Modeling and Simulation

- **Physiologically Based Pharmacokinetic (PBPK):** PBPK models try to predict the time course of blood and tissue concentrations within a species to understand safety and efficacy parameters of compounds. Areas of interest include using chemical structure or *in vitro* data to make accurate predictions early in the drug discovery process of (1) *in vivo* clearance and routes of elimination and (2) *in vivo* transporter mediated disposition.

- **Quantitative Systems Pharmacology (QSP) Modeling:** Mechanistic modeling approaches such as QSP and systems biology help understand disease pathways from a quantitative perspective and can be used to validate novel drug targets and to estimate how well preclinical efficacy might translate to the clinic. Areas of interest include approaches that address the following challenges: (1) limited data availability for intracellular and/or tissue concentrations of proteins, (2) understanding interactions between different proteins and (3) differences in protein levels in diseased versus healthy states.

- **Data Visualization and Interaction:** Interest in developing a general and flexible framework to enable visualization and interaction with high dimensional data (e.g., small molecule, 'omics, pathway) from diverse domains (i.e., structural, chemical, biological and clinical) across multiple scales (e.g., atomic, cellular, tissue) for both VR (virtual reality) and AR (augmented reality) to enhance the ability of identifying critical connections within data for improved decision making.

- **Predictive Modeling:** Areas of interest include model scaling, uncertainty quantification and model fusion techniques that robustly couple sparse and variable discovery/development data to predict therapeutic windows of medicines. Specific interests include the following:
    - Data integration: improve ability to perform inference over heterogeneous, noisy and often large datasets with intended gaps
    - Data connectivity: enhance linkages between domain sciences and data science
    - Data exploration/analysis: extract meaningful features and governing equations from data
    - Image analysis challenges: develop scalable video analysis platform for discovery, accelerate workflows, automate interactive visualization of 3D CT (Computed Tomography) data and compute biologically relevant parameters in a robust fashion with relatively little human interaction

Additional areas of interest in computational methods:

- **Computational Toxicology:** Increasing ability to predict off-target binding events and translate these to *in vivo* animal or human safety tolerability outcomes
- **Molecular Recognition:** Quantitatively predict protein-ligand association constants and binding kinetics from first principles at a rate commensurate with synthesis and testing
- **Automated molecular design:** Concurrently use multiple protein-ligand structures for pharmacological and selectivity targets in conjunction with other models (e.g., QSAR) for 'inverse design' of novel selective small molecules
- **Protein design:** Explore *in silico* methods to design proteins with novel activity and enhanced stability within manufacturing conditions; explore *in silico* methods for predicting antibody epitope mapping predictions and subsequent affinity maturation

Enabling Platforms for Complex Cellular Engineering

- **Advanced Cellular Engineering:** Technologies associated with genome editing and stem cell biology to develop *in vitro* models of human disease for use in target identification, target validation, assay development, screening, hit qualification and safety assessment. Technologies should have one or more of the following characteristics:
    - Ability to utilize synthetic gene circuits
    - Ability to modulate target gene expression for activation and repression
    - Ability to conduct the above in induced pluripotent stem cell (iPSC) lines or in cells obtained through iPSC differentiation
    - Can be validated
    - Can be scaled
    - Uses co-culture or 3D models to mimic physiological interaction
    - Can be miniaturized
    - Can measure multiple phenotypes and cellular responses

- **Innovation in Data Analytics:** *In silico* models to design relevant *in vitro* models for human disease; enhance disease-relevance of complex *in vitro* models by incorporating *in silico* models into the design cycle

Miniaturized and Continuous Biology & Chemistry Platforms

- **Rapid chemical tool discovery platform:** Creative approaches to accelerate the tool discovery process could include rapid chemical synthesis platforms, cell engineering, chemical biology, cell-permeable libraries, miniaturised screening and/or computational approaches

- **Experimental analysis and computational model to predict *ex vivo* human tissue viability and function:** Maintaining tissue viability and function outside the human body poses a major barrier to incorporating tissue slice assays into early drug discovery campaigns. GSK is interested in strategies to maintain viability and function of human organotypic slice cultures for time periods that allow for compound testing.

- **Closed-loop design-make-test:** Advances towards a miniaturized multiplex Artificial Intelligence platform that could achieve the following:
    - Automate iterative synthetic design of drug molecules
    - Carry out automated chemical synthesis
    - Conduct rapid tests using multiplexed, miniaturized biology
    - Automate information analysis

- **Delivery and Sampling Device for small molecule screening platforms:** Develop technology to deliver reagents and sample live cells while in culture, ideally with on-board quantification of (multiplexed) extracted cytosolic content in a miniaturized or nano-screening setting, offering the ability to (1) measure cellular responses as a function of time at unprecedented levels of precision and in a non-destructive manner and (2) introduce, in a selective fashion, reagents, compounds, sRNA and/or virus to isolated cell populations and quantify response at cellular level in a multiplexed format.

- **Platform for High Throughput Identification of On/Off Target Effects:** Develop an ultra-fast, highly sensitive detection platform that can identify binding events, i.e., the next generation ASMS [Affinity Selection Mass Spectrometer, aka size-exclusion chromatography mass spectrometer (SEC-MS)] platform. Detection could be single targets for the first iteration, with a longer-term goal of developing a multiplex platform to interrogate multiple (10's) targets in the same system. Desired characteristics include the ability to:
    - Operate at speeds of 1 kHz or greater
    - Test millions of compounds
    - Require ultra low protein and compound concentrations
    - Operate within a wide dynamic range, i.e., detect binding events from double digit micro molar down to nanomolar
    - Assign ligand identity unambiguously once a binding event is detected
    - Assign target ID as well as compound ID for multiplexed targets

Sensors and (Bio)sensors for Miniaturized Devices: Miniaturized devices to collect multi-parametric data for drug discovery biological assay readouts such as high content imaging, detecting protein or cytokines with fluorescently labeled antibodies or transcriptional measures. Multiplexing in miniaturized devices is challenging for some of these readouts due to spectral overlap and limitations of current instrumentation

and processing steps required to run experiments and gather data. Recent developments with miniaturized sensors have shown that extremely small and, in some cases, inexpensive detection systems can detect proteins or other metabolic indicators. Key questions to address include the following:

- What methods best apply to drug screening assays?
- What are the categories of detected biomolecules? Do we need antibodies/ligands?
- What detection/signal indicators can be used, and how can those signals be captured in an automated way? Are they visual, electronic, other?
- Can label-free optic methods detect cellular/pathway changes? What options can be developed for typical plate reader instruments?

- **Biopharm Molecular Discovery:** Interest in technologies to facilitate identification and development of biologic therapies including the following:
  - Platforms/technologies for high content functional assays of large panels of B cells, particularly emphasizing ways to link genotypes to recovery of clones with desirable phenotypic properties
  - Technologies that expand the target landscape for antibodies (e.g., alternative dosing routes such as oral or inhaled) or new target binding parameters (e.g., intracellular targets, targets behind the blood brain barrier)
  - Replace or augment assay systems with *in silico* technologies. Can antibody sequence alone predict developability? Can antibodies be designed *in silico* based on sequence and binding profiles of entire B-cell repertoires?
  - Technologies to generate ion channels at high copy number (>10^5) in cells to facilitate selections and subsequent assays
  - Innovations to enable screening large numbers of molecules efficiently across the range of parameters thought to predict developability
  - Way(s) to mimic UF/DF (Ultrafiltration/Diafiltration) at nanoliter scale to enable screening antibodies for their ability to withstand purification procedures that concentrate antibody solutions during commercial manufacture