



## The Structural Genomics Consortium

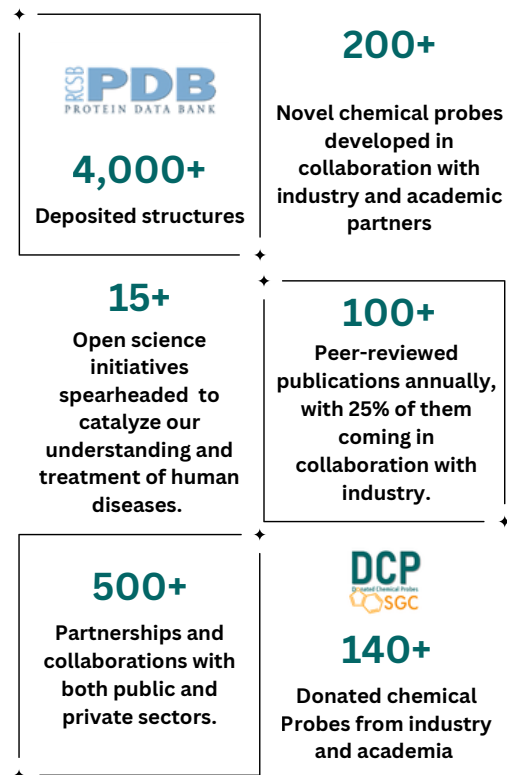
A global public-private partnership that supports the discovery of new medicines through open access research.

## About SGC

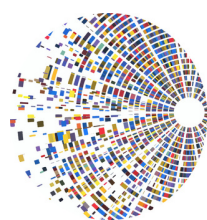
The Structural Genomics Consortium (SGC) is an international public-private partnership dedicated to understanding the functions of all human proteins. Founded in 2003, the SGC operates with laboratories in Canada, the United States, Germany, the United Kingdom, and Brazil, supporting over 250 scientists working across structural biology, chemical biology, data science, and machine learning.

Over the past two decades, SGC has developed openly available scientific tools that have been used to understand fundamental human biology and accelerate early-stage drug discovery. These include more than 4,000 publicly deposited protein structures, more than 200 high-quality chemical probes, and widely used datasets that form the foundation for new biological insights and therapeutic hypotheses. SGC-developed probes and data have been adopted globally by academic and industry groups, cited in thousands of studies, and used to support multiple clinical programs.

The SGC is widely recognized for its open-science model, sharing data, reagents, and workflows without restriction. This commitment has enabled broad collaboration, minimized duplication of effort, and strengthened the reproducibility and quality of research worldwide. Today, SGC continues to lead the shift toward open, data-driven drug discovery and serves as the coordinating organization behind Target 2035.



## Target 2035: Our flagship initiative



**TARGET  
2035**

Target 2035 is a global, open-science initiative with an ambitious mission: to develop pharmacological modulators for most human proteins by 2035.

Now transitioning out of its stealth mode, Target 2035 is prioritizing the development of advanced computational tools to accelerate probe development, from hit finding to hit optimization. A key focus of this new phase is generating freely accessible, large-scale, high-quality protein-ligand binding datasets to train and benchmark next-generation AI models.

**The SGC leads and coordinates Target 2035, mobilizing a global team of scientists from academia, industry, technology companies, government agencies, and non-profit foundations towards its mission.**

### The unfulfilled promise of the human genome

Most of the human proteome – all proteins encoded by our genes – is still poorly understood. While many have been linked to disease, only a small fraction has approved medicines or the high-quality chemical tools necessary for their interrogation. This area, often referred to as the “dark proteome”, represents a major untapped opportunity to decipher human biology and discover new treatments.

The first step in developing a chemical probe is identifying small molecules (“hits”) that bind to a protein. Historically, this process has been expensive, and difficult to scale. Artificial intelligence (AI) has the potential to significantly accelerate and reduce the cost of this process, but only if large, high-quality datasets are available to train robust models. These datasets do not currently exist at the scale or quality required.

SGC projects aligned with the mission of Target 2035 are supported by Innovative Health Initiative (IHI) under Horizon Europe, the Bill & Melinda Gates Foundation, the Michael J. Fox Foundation, the U.S. National Institutes of Health (NIH), the Canada Foundation for Innovation (CFI), and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).



### Ultra-large protein-ligand datasets

Support robust, generalizable machine learning models



### Open, Findable, Accessible, Interoperable and Reusable (FAIR)

Data infrastructure for integration, benchmarking and reuse.



### A global community

A multidisciplinary community spanning experimental and computational sciences.

### Core activities of our next phase include:

- Generating high-quality protein-ligand binding data across thousands of proteins using screening platforms such as Enantioselective Affinity Selection Mass Spectrometry (E-ASMS) and DNA-Encoded Libraries (DEL).
- Building shared infrastructure – including the [AIRCHECK](#) platform- to house and share high-quality, interoperable data.
- Hosting open benchmarking challenges, including machine-learning evaluation campaigns, to engage the global machine learning community.
- Getting shared assay and data standards across global partners to enable reproducible, scalable experimentation.

### The data backbone of Target 2035



## AIRCHECK

AIRCHECK is a cloud-based open platform developed by SGC to serve as an open-access hub for AI and machine learning model training and validation. The platform integrates data generation, storage, dissemination, and model development to facilitate the prediction, testing, and refinement of small-molecule protein binders.

Designed to serve as the “Protein Data Bank for ligands,” AIRCHECK enables researchers to train models, compare performance, and openly deposit new methods.

## How to get involved

The SGC offers multiple ways for the scientific community to contribute to and benefit from the Target 2035 mission.

### Contribute Proteins

#### Join the [Protein Contribution Network](#)

and submit purified, high-quality proteins to be screened for small molecule binders. Allow the screening data to be shared, and in turn receive validated hits for your protein of interest.



### Participate in Open Benchmarking Challenges

Target 2035 collaborates with BEACON, the new benchmarking super-consortium comprising CACHE, DREAM, CASP, OpenADMET, to provide a platform for computational scientists to benchmark hit-finding algorithms in real world settings, with experimental testing of model predictions.



### Join [MAINFRAME](#)

A new international network of machine learning researchers, computational chemists, and data scientists. Membership is free and provides access to curated datasets for model testing and benchmarking.



### Be part of the mission

SGC welcomes new partners from industry, technology companies, and research foundations. Joining the SGC offers the opportunity to influence and govern the overall direction of the consortium and its projects, including Target 2035, and collaborate with the SGC scientists on your priority projects.

