Managing Informatics Challenges across Distributed Drug Discovery Projects

Collaboration between drug discovery groups from different sites, organizations and companies has become much more common recently. In some cases, virtual or semi-virtual companies are formed by outsourcing all, or a subset of chemistry and biology services to contract research organizations (CROs). Distributing research in these ways can have profound advantages, but managing distributed research has unique challenges.

Mironid is a biotechnology company with distributed research operations that include core biology and chemistry teams located at two different sites, plus supplementary chemistry, biology and DMPK services contracted through external CROs. These scientists, working in many different locations, collaborate closely to apply an understanding of cell signaling to the design of tests that mirror the behavior of the living cell environment and identify compounds that can better treat target diseases. Informatics technology plays a key role, providing unified data collection and processing to enable seamless collaboration between the groups. We will explore how Mironid manages its data, projects and collaborations.

The Science

Mironid's discovery research focuses on developing novel small molecule drugs to exploit new insights into phosphodiesterase (PDE) enzyme structure and function to target degenerative kidney diseases and chronic inflammatory diseases. Autosomal Dominant Polycystic Kidney Disease (ADPKD) is an orphan genetic indication that affects approximately 250,000 patients in the EU and USA alone and causes major health issues, eventually leading to kidney failure for 50% of patients by the age of 55. Current clinical treatments do not adequately address the underlying cause of cyst formation to stop disease progression.

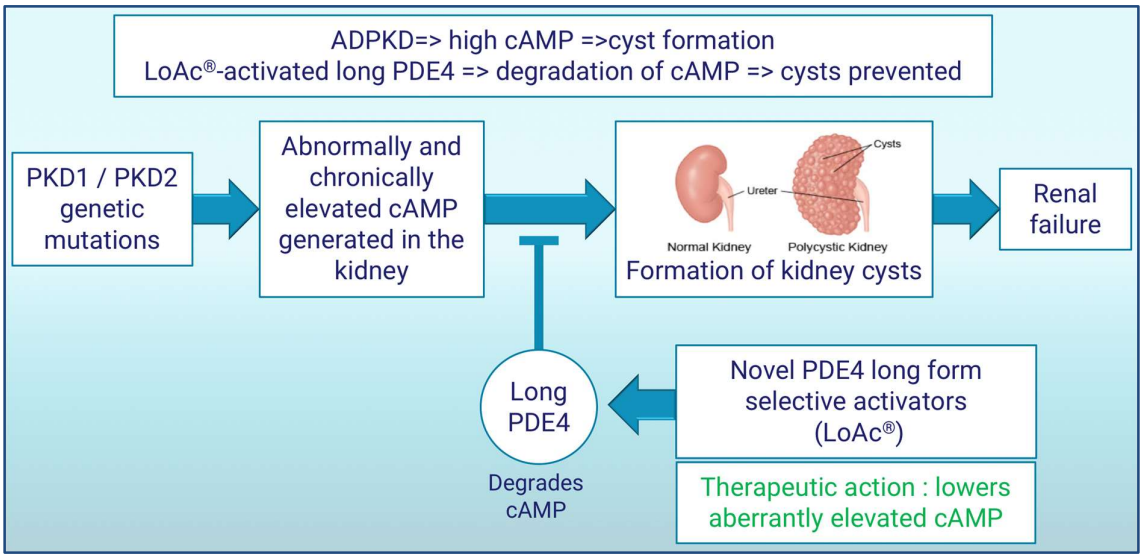

Figure 1 – Metabolic pathway involved in cyst formation in Autosomal Dominant Polycystic Kidney Disease (ADPKD), and the target of therapeutic intervention: Phosphodiesterase 4 (PDE4)

In humans, twenty-four PDE4 isoforms are encoded across four distinct genes (PDE4A-D). Each of these isoforms are characterized by regions of unique sequence, but follow an overarching structure of conserved regulatory domain expression which allows their grouping into 'long' and 'short' forms. While inhibition of PDE4 enzyme activity is a long-established method of influencing cell behavior, the inhibitors have always suffered from side effects caused by a lack of specificity of molecules binding to the catalytic site. Small molecules that *activate* PDE4, on the other hand, had never been reported, until recent disclosure by Mironid [Omar *et al.*, *Proc. Natl. Acad. Sci. U S A.* 2019; 116: 13320–13329].

Mironid has developed insights into allosteric modulation of PDE4 enzymes that allow targeted activation or inhibition of selected subgroups of PDE4 isoforms. Of particular interest are compounds that can selectively activate PDE4 long forms, and thus decrease the levels of the second messenger, cyclic adenosine monophosphate (cAMP) in key subcellular signaling complexes; these are termed **LoAc**® compounds, standing for **Lo**ng form **Ac**tivators. The world's first identification of selective small molecule PDE4 long form activators provides a novel mechanism of action for the development of first in class molecules to treat diseases driven by aberrantly elevated levels of cAMP. These include several rare and orphan diseases, including ADPKD. The activation of PDE4 long isoforms by LoAc® compounds reduces the concentration of cAMP in specific sub-cellular locations and reduces or reverses the formation of cysts in models of ADPKD.

Selecting an Informatics Platform

Informatics plays an important part in the drug discovery process. Proper collection, collation and reporting of results are imperative to show value from a project and provenance to the data. This can become increasingly complicated when compounds and data are being generated on multiple sites and by different organizations.

One of Mironid's first actions was to establish an integrated, structure-searchable drug discovery database to secure its historical data and manage new data going forward. Mironid adopted the informatics platform, **CDD Vault**, as the company's central data repository and informatics framework. This commercial solution is cloud-hosted for a zero footprint system that can provide access to authorized users anywhere in the world from a web browser.

Over the first few months, Mironid's historical compound structures and data were transferred to the new database, in parallel with staff recruitment, setting up new labs and scaling up the synthesis and testing of new molecules. From that point onward, new compounds, protocols and data were added to the database in real time as the standard way of sharing data within the company. This simplified the process of finding and cross-referencing data for particular compounds and interrogating the data by chemical structure, to identify structure-activity relationships and make informed decisions for the rapid optimization of lead molecules.

Managing Data across Distributed Teams

In order to work in the same efficient way with CRO partners, Mironid created isolated project spaces in its CDD Vault and granted access to selected CRO's lead scientists, allowing them to securely upload and view the data they needed. **By recording compound structures and experimental results directly in CDD Vault, they eliminated the security risk of emailing files back and forth, as well as the challenges of keeping data spreadsheets synchronized and current.** Mironid's CDD Vault users have access to their colleagues' latest experimental results in real time, but only the data sets for which they have individually been authorized. **This allows for a close collaboration with external partners, while ensuring that each partner has access to just the subset of data they need.**

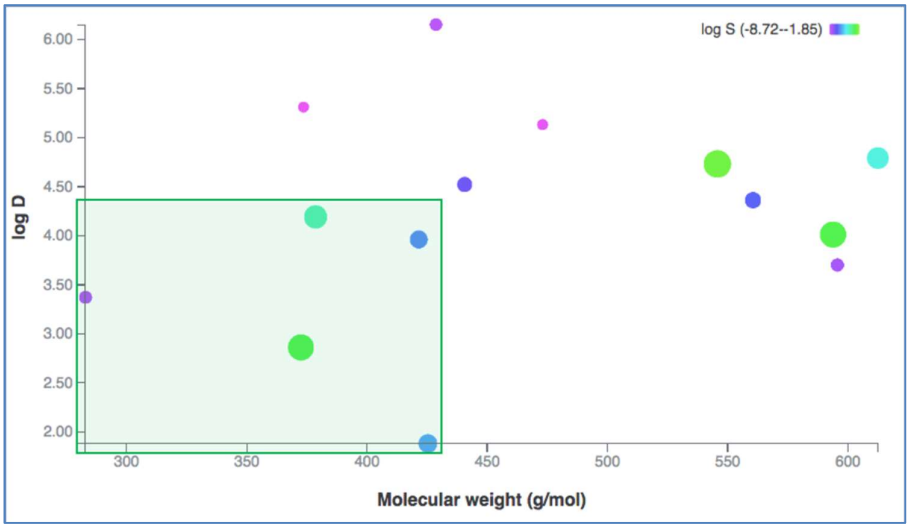

Figure 2 – Hit compounds from the Biomedical Catalyst co-funded project, plotted in CDD Vision by molecular weight and calculated logD. Spot size corresponds to activity and color to predicted solubility. Based on overall properties, the five hit compounds in the green rectangle were deemed most attractive to follow up.

In 2017, Mironid secured an Innovate UK Biomedical Catalyst Award to co-fund the discovery of new LoAc® PDE4 activators for the treatment of ADPKD. For this project, Mironid worked with a network of UK CROs to apply powerful computational modelling, X-ray crystallography and a high-quality library of over 12,000 diverse, drug-like screening molecules to generate a new wave of LoAc® PDE4 allosteric modulators.

Screening data, chemical structures, analytical data, solubility and microsomal stability data were uploaded directly to CDD Vault by CRO partners, while biochemical profiles and cellular assay data were uploaded by Mironid. Data were immediately visible across sites, allowing rapid communication and decision making. In addition to data sharing, searching and retrieval, CDD Vault allowed the project team to readily view and analyze the various hit compounds in scatter graphs to help select the most attractive hits for progression based on their biological activity and physico-chemical properties, such as size, lipophilicity and predicted solubility.

The distributed project team rapidly identified multiple high-quality hit compounds. The hit compounds with the most attractive profiles (Fig 2) were re-synthesized and small arrays of analogs were prepared and tested to determine scope for further optimization within each series. After demonstrating scope for optimization, selected compounds from two distinct chemical series were progressed into cellular disease models and were shown to suppress cyst formation in 3D kidney cell culture, an *in vitro* model of ADPKD.

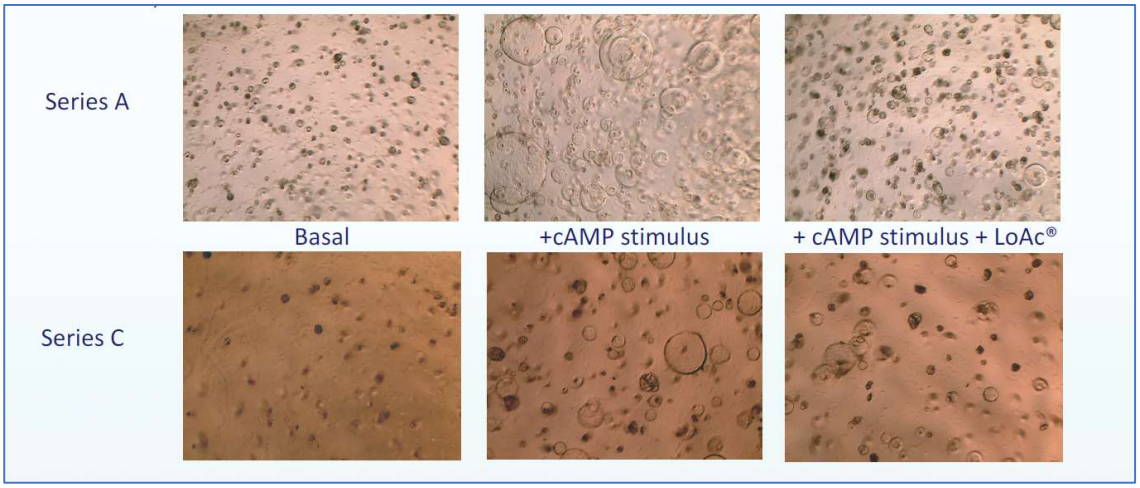

Figure 3 – Hit compounds from Series A and Series C suppressed cyst formation in 3D kidney cell culture, an in vitro model of ADPKD

The Future

By employing a unified informatics platform, the project team was able to coordinate expertise and resources across different sites and use all of the available data in day-to-day decision making. Chemists and biologists were able to work together in a single computational environment, regardless of their physical location.

As the number of compounds evaluated expanded from hundreds into thousands, the CDD Vault database has continued to provide rapid feedback on queries based on single or combined search terms such as substructure, calculated physico-chemical properties, assay type or assay readout.

The Biomedical Catalyst co-funded project yielded a rich resource of diverse, drug-like LoAc® compounds that activate PDE4 long forms, reduce cellular cAMP levels and suppress cyst formation in models of ADPKD. Modern informatics resources fostered these advances by providing infrastructure for researchers from different labs to collaborate in real time and coordinate their best combined efforts. Further optimization of these hit series is in progress within Mironid with the aim of delivering improved therapies for ADPKD patients.

Authors

Julia M. Adam[1]
David R. Adams[1,3]
Keith Bowers[1,2]
David J P Henderson[1,2]
John King-Underwood[4]

[1] Mironid Ltd., BioCity Scotland, Bo'ness Road, Newhouse, North Lanarkshire, ML1 5UH, Scotland, UK
[2] Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 16 Richmond Street, Glasgow, G1 1XQ, Scotland, UK
[3] Institute of Chemical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK
[4] Comp Chem Resource UK, Old Cottage Hospital, Homend, Ledbury, Herefordshire, HR8 1ED