



ChEMBL

ChEMBL



CHEMISTRY MEETS BIOLOGY

ChEMBL is a unique public knowledge base of chemical compounds and small molecules with their biological targets. It pulls together high-value information on compounds and their effects on biological systems from the available academic literature in a structured database.

EMBL-EBI dynamically links information from the academic literature in ChEMBL to chemical patent documents recorded in the SureChEMBL database. SureChEMBL provides live free access to chemical data extracted from the patent literature. Data in patents are important to drug discovery researchers because new discoveries often appear in patents 2-3 years ahead of the published scientific papers.

The ChEMBL database was originally developed as a commercial product, becoming an open and freely available EMBL-EBI service in 2008 with funding from the Wellcome Trust. In 2013, this was extended to include patent information held in the SureChEMBL database.

FUNDING

ChEMBL is primarily funded by



with supplementary funding from



IMPACT

ChEMBL's focus on enabling all aspects of discovery, is utilised by academics and industries of all sizes, strengthening innovation from new research, and the discovery of new treatments and drugs benefiting human health and agriculture.

In the recent Strategic vision for UK e-infrastructure report⁰⁰², Professor Dominic Tildesley of Unilever identifies the ChEMBL database as key in their product development of anti-perspirants. They used the database, to identify active components for anti-perspirants and the ChEMBL data to build a model of their inhibition activity.

Below are quotes from four external users of the ChEMBL database which demonstrates how ChEMBL improves R&D, increases productivity and performance and underpins scientific investment.

USER IMPACTS

Syngenta (Industry: AgroChemicals)

Syngenta is a leading agriculture company and employs 28,000 people in over 90 countries. 5,000 are in R&D and apply world class science to provide innovative crop solutions to transform how crops are grown.



"It has been estimated that without crop protection compounds (pesticides, fungicides, weed killers, etc.) 40% of the world's food would not exist. Our scientists use ChEMBL to support projects in our research towards innovative new products, and ChEMBL has links between both chemistry and biology data which makes it searchable in ways that the underlying literature would not be. People at the EMBL-EBI do a fantastic job in making a vast amount of data of different types openly available to researchers, and without the EMBL-EBI resources in general I'm sure life science research would be greatly hindered."

– Mark Forster, Syngenta

DATA STORAGE

13.5
MILLION
RECORDS



ChEMBL contains information on more than 1.4 million compounds and 13.5 million records of their effects on biological systems.

SureChEMBL provides a live, updated daily, view of chemical patents, with approximately 50,000 new documents added per month.

UNIQUE RESOURCE

70%



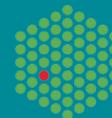
If the ChEMBL resource did not exist 70% of its users stated that they wouldn't have been able to obtain the data from anywhere else.⁰⁰¹

PUBLICATIONS

1,125

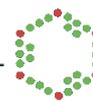


ChEMBL reference papers have been cited 1,125 times⁰⁰³ with use of the database directly cited in 196 publications. Covering areas as diverse as colon cancer, drug design & development, virtual screening, experimental modelling & complimentary biology databases.



ChEMBL

ChEMBL



USERS IMPACTS Cont.

The University of Sheffield Information School**Placed 1st for the impact of research by the 2014 UK Research Excellence Framework.**

“ChEMBL provides a unique resource in terms of publicly available data about compounds and their properties on a scale that is not available elsewhere. This facilitates the process of developing new methods for virtual screening and allows different research groups to compare methods more easily than would be possible without this data, which helps to advance the field more quickly. Without ChEMBL, the methods in general use for drug discovery might be poorer and have reduced success rates for virtual screening. This would lead to higher costs because of having to do more testing; fewer successful drug developments, and potentially an impact on human health as a result.” - *Professor Val Gillet, Professor of Chemoinformatics, Head of School, University of Sheffield Information School*

MRC Technology (Charity)**An independent life science medical research charity working to bridge the gap between basic research and commercial application.**

“ChEMBL database has proven to be invaluable to the computational group at MRCT. Whilst commercial sensitivity prevents the identification of specific targets and programmes active within MRCT, there are many cases of analyses covering a wide range of protein classes that have been moved forward with valuable data based on mining the ChEMBL data sources.” - *Dr Andy Merritt, MRCT*

Cambridge MedChem Consulting (Drug Discovery)**Cambridge MedChem Consulting is a micro-company based in the UK providing a range of consultancy services in drug discovery and medicinal chemistry.**

“ChEMBL acts as a standard source so everyone is working from the same base. There are an increasing number of independent consultants and small companies like MedChem. Without ChEMBL there would be some things I would just not be able to do. For example, going through the original literature would be prohibitively time-consuming (months/years). It is pretty unique as a data source, and the web services and API allow you to build custom tools and very efficient pipelines for access. It is an invaluable resource and the EMBL-EBI staff are helpful, enthusiastic and keen to promote it” - *Dr Chris Swain, Cambridge MedChem Consulting*

SUPPORTING PUBLICATIONS

- 001 **The value and impact of The EMBL-European Bioinformatics Institute**
Impact Survey conducted in 2015 with 4185 respondents of which 771 classified themselves as ChEMBL users.
Charles Beagrie Ltd
- 002 **A Strategic Vision for UK e-Infrastructure professor Dominic Tildesley, Unilever PLC**
A independant report commissioned by UK Gov BIS Minister for Universities and Science. A roadmap for the development and use of advanced computing, data and networks
Professor Dominic Tildesley, Unilever PLC
- 003 **Google Scholar**

USER IMPACT

68%



Of surveyed ChEMBL users said it would have a major or severe impact on their work or study if they could not access EMBL-EBI services and resources.⁰⁰¹

BRAIN GAIN

94%



94% of ChEMBL users were involved in research, spending an average of 34 hours a week on research activities, of which 60% was spent working with data.⁰⁰¹

RESEARCH EFFICIENCY

46%

MORE EFFICIENT



The efficiency value placed on our services represents a direct worth of between £5,382 to £26,000 per respondent with the overall average that EMBL-EBI services allowed users to be 46% more efficient in their work¹.



All interviews and subsequent economic analysis were undertaken by Charles Beagrie Ltd. on behalf of EMBL-EBI.