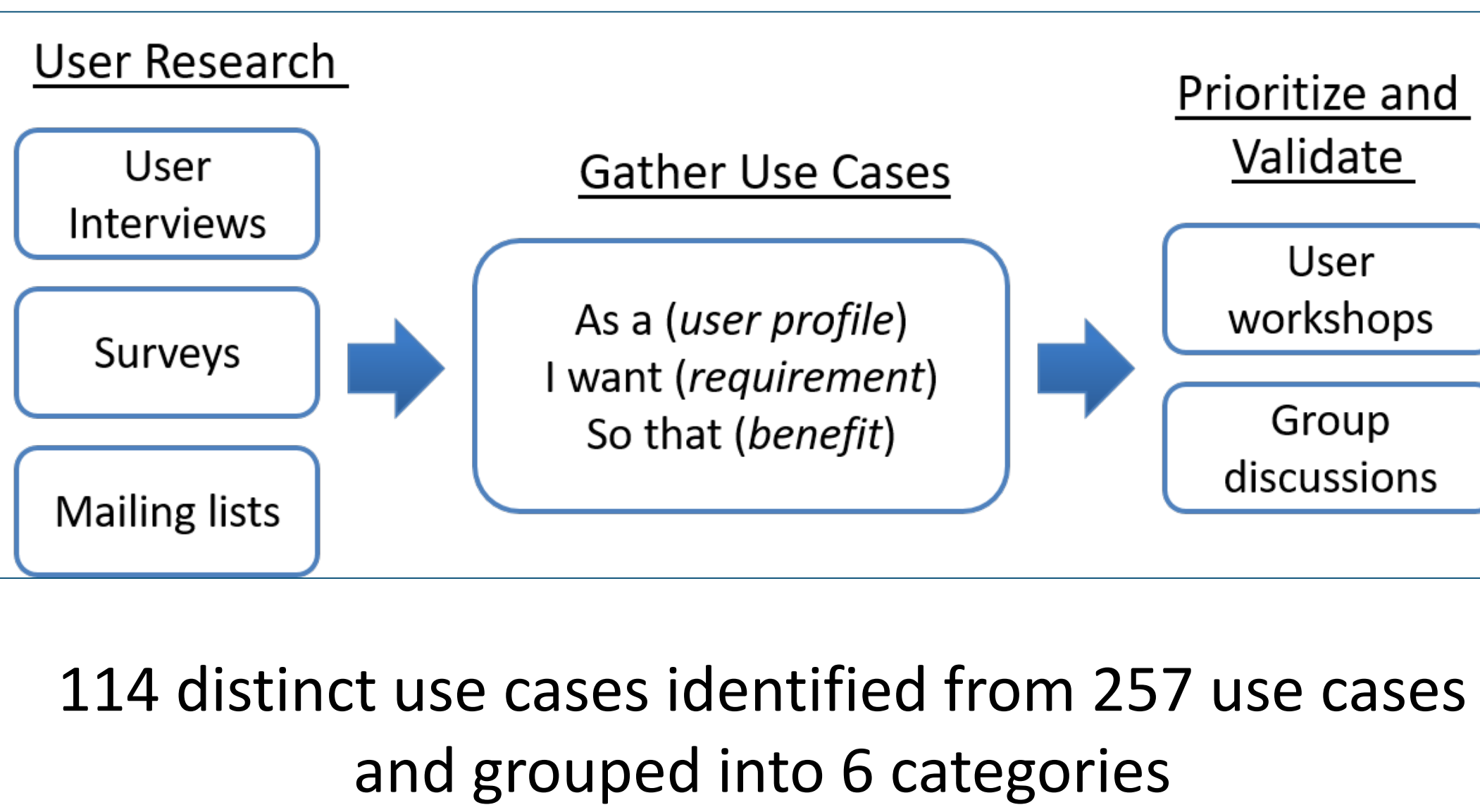




### ABSTRACT

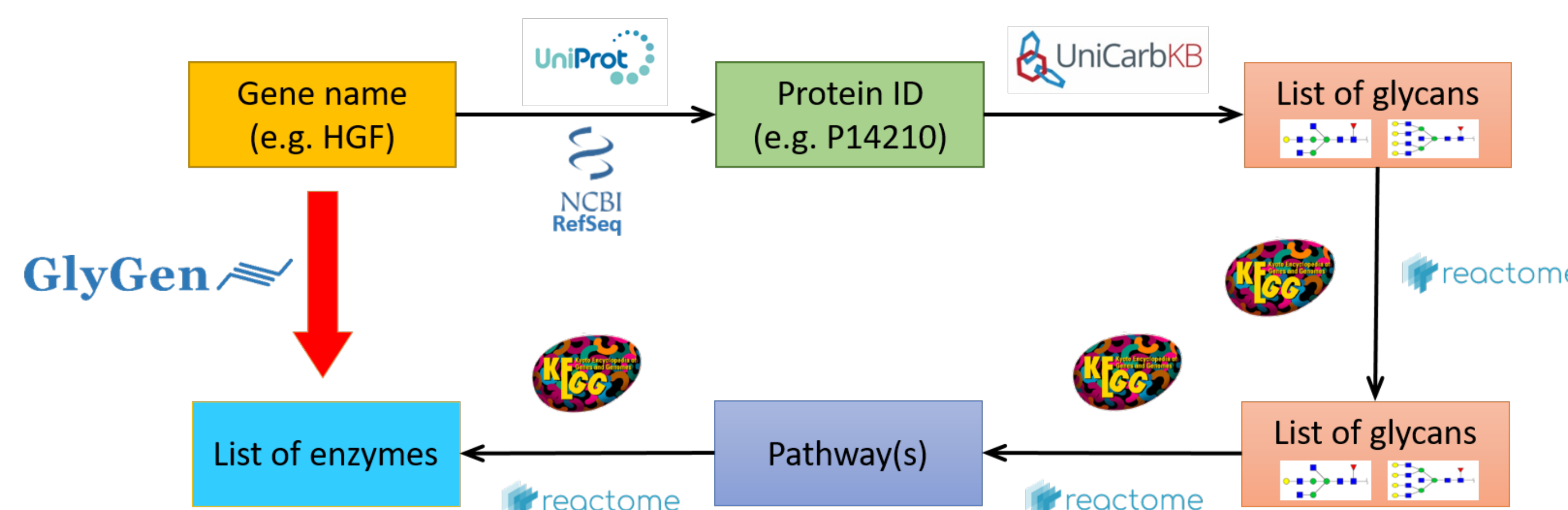
**GlyGen** is an integrated, extendable and cross-disciplinary glycoinformatics resource that will facilitate knowledge discovery in basic and translational glycobiology by integrating multidisciplinary data and knowledge from diverse resources. It will address glycobiology questions that can currently be answered only by extensive literature-based research and manual collection of data from disparate resources. The aims of the GlyGen project includes integrating and exchange of up-to-date glycobiology-related information and data with partnering data sources such as EMBL-EBI, NCBI, UniProt, UniCarbKB, and others; creating an intuitive web portal to search and browse for glycoscience knowledge that will also include off-line data exploration and mining. Furthermore, the GlyGen project includes the development of essential new information resources, namely the Glycan Array Database that will provide key information about the interactions of glycans with other biomolecules and a Glycan Naming Ontology (GNOME) that facilitates interpretation of incomplete structural information in the context of biological functions. Analysis of over 100 use cases contributed by users was used to design a comprehensive data integration framework that allows GlyGen to provide unprecedented support for complex queries spanning diverse data types relevant to glycobiology, extending its scope beyond the mapping of glycan data to genes and proteins. The resource will provide new opportunities for a systems-level understanding of glycobiology in disease and development, even for scientists who do not specialize in glycobiology.

### Collection and Prioritization of User Input



### Complex User Queries translated into simple 'Quick Search'

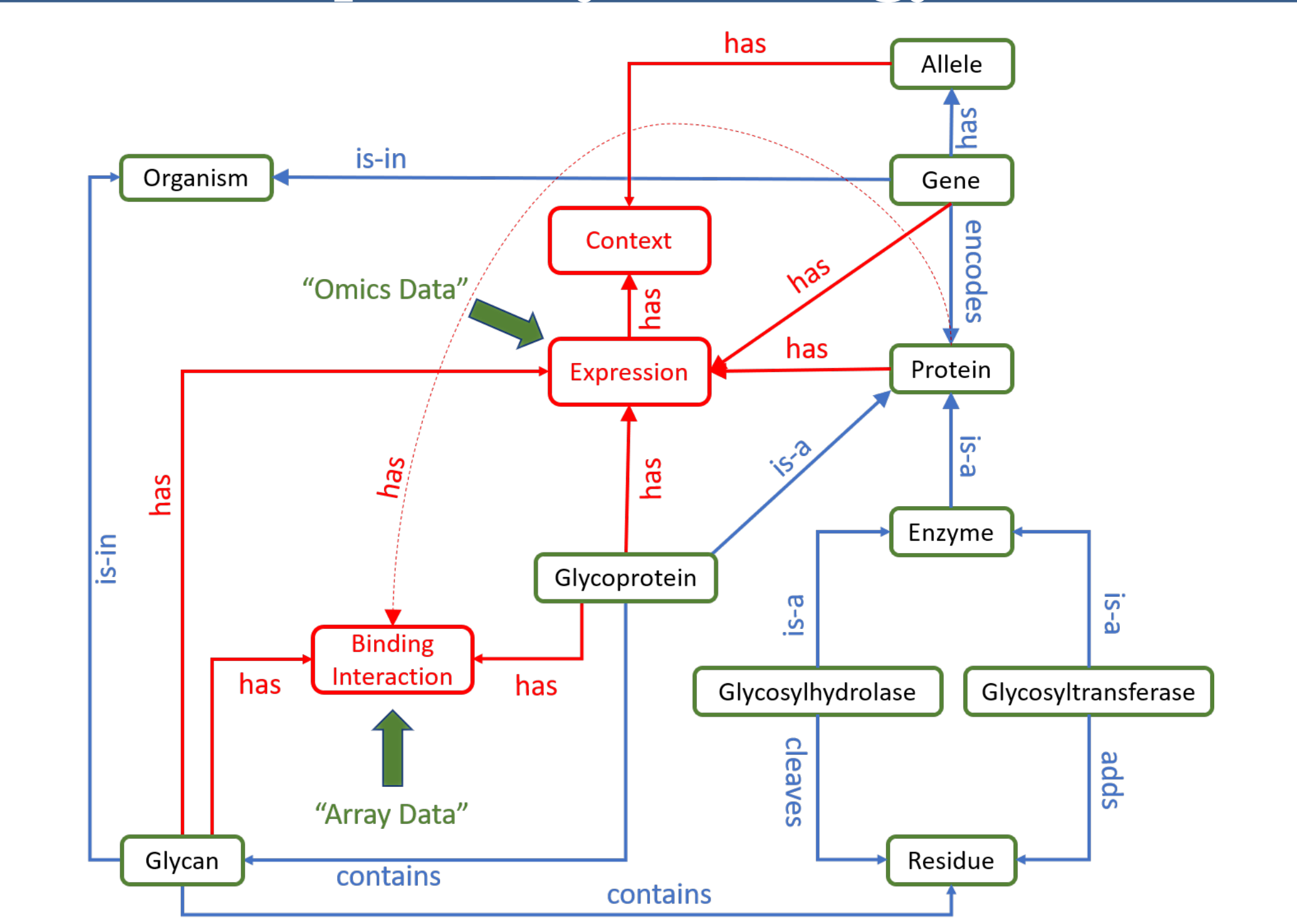
What are the **enzymes** involved in the biosynthesis of the glycans on [protein X](#)?



### Glycosylation Details for Protein P14210 (HGF-Hepatocyte Growth Factor)

Source	GlycoCan Accession	Type	Enzyme	Image of Glycan Structure
UniCarbKB	G015432X	N-linked	Asn653	
UniCarbKB	G117689DH	N-linked	Asn653	

### Data Model for Visualization of Complex Glycobiology Data

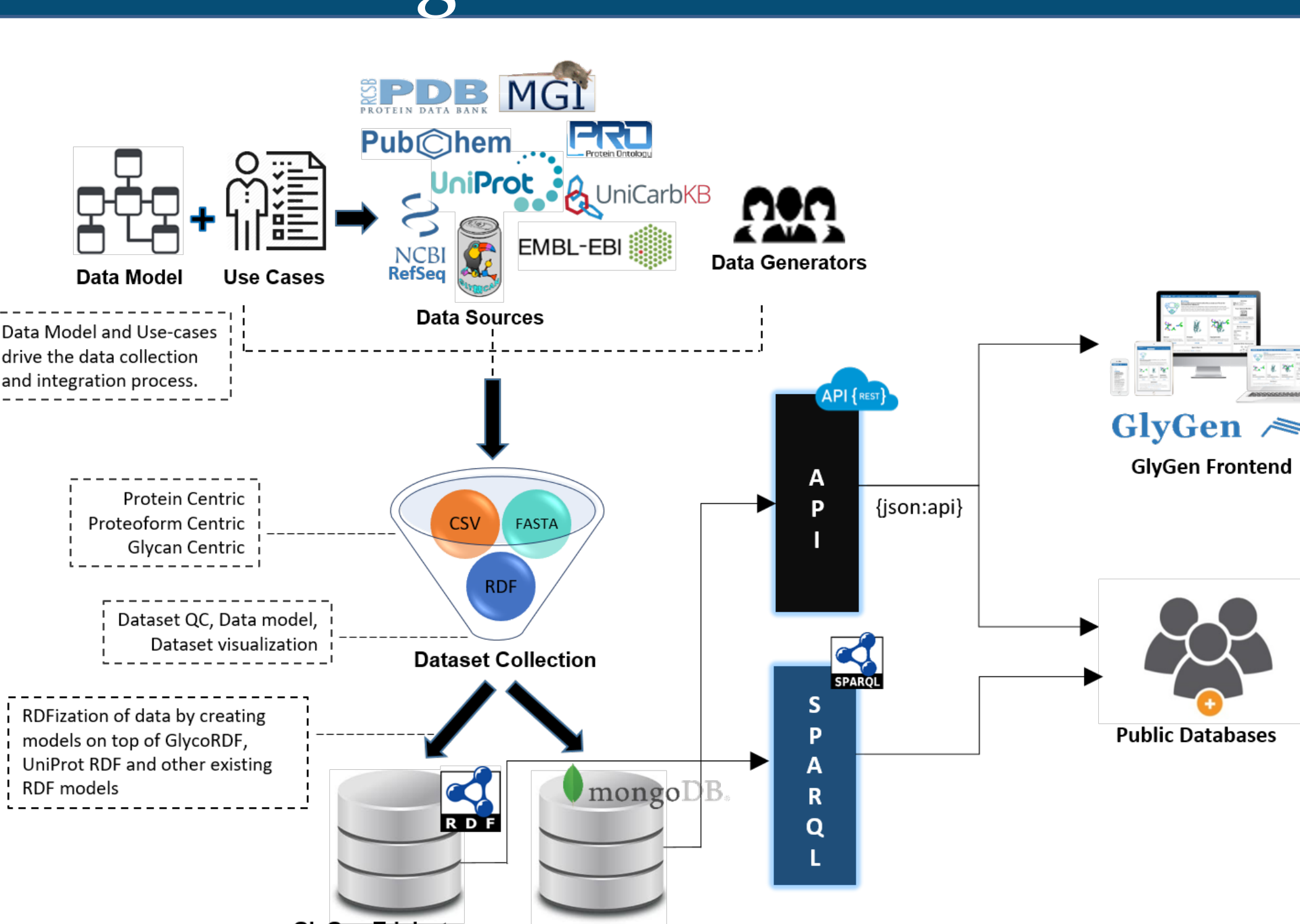


### GlyGen Portal and Dataset collection - glygen.org

### Glycan Detail Page for Glycan G17689DH (Bi-antennary N-Glycan Complex)

Source	Protein Name	UniProtKB Accession	Position
UniCarbKB	Hepatocyte growth factor	P14210-1	653
UniCarbKB	Hepatocyte growth factor	P14210-1	294
UniCarbKB	Hepatocyte growth factor	P14210-1	566

### GlyGen Data Collection and Integration Workflow



### Metadata captured in Readme using BioCompute Objects Schema

### Acknowledgement

NIH Grant - U01 GM125267-01, Radoslav Goldman, Judith Blake, Michael Pierce, William York, Raja Mazumder, Nathan Edwards, Rene Ranzinger, Robel Kahsay, Maria Martin, Kiyoko Aoki-Kinoshita, Matthew Campbell, Evan Bolton, Darren Natale, Karen Ross, Rob Woods, Ten Feizi, Jeff Gildersleeve, Richard Cummings, Brian Fochtman, Nicki Packer, Tatiana Williamson, Leyla Garcia, Preethi Vasudev, Rupali Mahadik.