BEILSTEIN SYMPOSIUM

Time-proof Perspectives on Glycoscience

Beilstein Glyco-Bioinformatics Symposium 2019

25–27 June, 2019
Dom Hotel Limburg
Limburg, Germany

Beilstein-Institut and Open Science

The non-profit Beilstein-Institut is one of the most respected organizations in the communication and dissemination of high-quality information in chemistry. Since 1951, when the foundation was established by the Max Planck Society, we have been fulfilling our mission to support the scientific community by providing high-quality information that is essential for research.

Our role has evolved over the years: from the production of the Beilstein Handbook and Database, to being one of the first open access journal publishers in chemistry, to host of interdisciplinary symposia and supporter of open data initiatives. We believe that free access to scientific research results, giving everyone in the world an equal chance to read and reuse experimental findings and data, is the best way to advance science.

Open Science is a new approach to scientific research. It is based on cooperation and uses new ways to disseminate information and broaden knowledge through digital technologies and new collaborative tools. It aims to make the primary outputs of publicly funded research results – publications (open access) and the research data (open data) – publicly accessible in digital format with no or minimal restriction.

The Beilstein-Institut supports open science and makes the results of its projects freely available to the scientific community as open access publications. This is an essential contribution to the foundation's mission to advance the chemical and related sciences. All journal articles, conference proceedings and videos are open access to allow the worldwide, unhindered sharing and exchange of ideas. This allows scientists, students, educators and the public the opportunity to inform themselves of the latest developments in research and to build on these ideas to further advance scientific knowledge.

Our two platinum open access journals, the *Beilstein Journal of Organic Chemistry* and the *Beilstein Journal of Nanotechnology*, which we fully fund, have no fees for authors or readers. Both journals are produced and managed by the Beilstein Editorial Office team, who work together with a global scientific network of experts that are responsible for the peer review. In 2015, the Beilstein Journals were awarded the DOAJ Seal which recognizes the exceptionally high level of publishing standards and best practices adhering to these journals.

An essential prerequisite for open science data is reporting guidelines and technical standards that provide the framework for the exchange of data from one laboratory to another without technical and textual barriers.

The Beilstein-Institut runs two data standards projects: STRENDA which is concerned with the reporting of enzymology data and MIRAGE with the reporting of glycomics experimental results. Both of which are now widely accepted and acknowledged by the scientific community.

The direct interaction and the exchange of thoughts and ideas between scientists are supported by a program of regularly hosted symposia. These international meetings are organized by the Beilstein-Institut and cover a variety of topics ranging from organic chemistry and biochemistry to nanotechnology and open science as well as interdisciplinary meetings on contemporary topics.

The Beilstein-Institut has been hosting symposia since 1988. Each meeting is always an interesting event with an open result: the Beilstein-Institut provides the framework and the lively and intense exchange of thoughts and ideas of the participants turn it into a memorable and lasting experience. The number of participants is usually limited to around 50 and the program is designed specifically to allow sufficient time for discussions. In some ways the talks can be seen as providing a catalyst for these discussions which often go on into the night and have led to subsequent cooperation projects. The resulting exchange between researchers is the underlying goal of the meeting and gives the Beilstein Symposium their unique character.

You will find regularly updated information about our symposia at www.beilstein-symposia.org.

Upcoming symposia in this year are:

Beilstein Enzymology Symposium 2019 *Molecular Functions, Catalysis and Regulation* 10 – 12 September 2019, Rüdesheim, Germany.	Scientific Program: Barbara Bakker, Carsten Kettner, Thomas S. Leyh, Santiago Schnell, Ming-Daw Tsai www.enzymology.beilstein-symposia.org
Beilstein Open Science Symposium 2019 *The What, How and Why of Open Science* 15 – 17 October 2019, Rüdesheim, Germany	Scientific Program: Martin Hicks and Carsten Kettner www.open-science.beilstein-symposia.org

Beilstein Nanotechnology Symposium 2019
MXene at the Frontier of the 2D Materials World

15 – 17 October 2019, Mainz, Germany

Scientific Program:
Yuri Gogotsi, Xinliang Feng, Johanna Rosén

https://www.beilstein-institut.de/en/symposia/nano-2d-materials

Beilstein Nanotechnology Symposium 2019
New Directions for Nanoporous Materials

12 – 14 November 2019, Rüdesheim, Germany

Scientific Program:
Sir Fraser Stoddard, Cafer T. Yavuz

https://www.beilstein-institut.de/en/symposia/nano-porous

Table of Contents

Book of Abstracts

Overview

This symposium will bring together glycochemists and biologists with experts in bioinformatics and computer sciences not only to look back at the roots of glycosciences 30 odd years ago but also to gather the perspectives of present and future applications of glycomics.

Glycomics, in the past considered a subfield of proteomics, has established itself and evolved to become an independent discipline which studies systematically the structures and functions of glycans in a cell or organism.

Consequently, boundaries to proteomics, genomics and metabolomics are crossed towards a systems-wide investigation of the various roles of glycoconjugates in cell-cell communication, recognition processes, signal transduction, molecular trafficking, modulation and regulation of molecular pathways in both eukaryotes and prokaryotes and in many other processes such as diagnosis and therapy of diseases. Glycomics also encompasses further important areas, for example, combining bioinformatics and mathematics approaches to develop methods for data storage, analysis, and modelling functions and relationships.

Under the guidance of the MIRAGE Commission, this conference series also provides a platform to discuss current standards and the needs for additional standards in the glycoscience. The mission of MIRAGE (www.beilstein-mirage.org) is to establish guidelines for the reliable and accurate reporting of glycan data and to interconnect existing and future infrastructure entities to provide the glycoscience community services for the deposition, sharing, analysis and reuse of glycan data.

All participants are encouraged to discuss their latest results, approaches, and methodologies in experimental, theoretic and bioinformatics glycosciences.

Enjoy the Symposium!

Scientific Committee

Kiyoko Aoki-Kinoshita
Soka University
Tokyo, Japan
kkiyoko@soak.ac.jp

Matthew Campbell
Griffith University
Goldcoast, Australia
m.campbell2@griffith.edu.au

Frédérique Lisacek
Swiss Institut of Bioinformatics
Geneva, Switzerland
frederique.lisacek@sib.swiss

Joe Zaia
Boston University
Boston, MA, USA
jzaia@bu.edu

Carsten Kettner
Beilstein-Institut
Frankfurt am Main, Germany
ckettner@beilstein-institut.de

Registration

All participants must be registered to have access to the conference area.

Participants can ask the organizers for a confirmation of the payment of the conference registration fee. Insurance of participants against accidents, sickness, cancellation, theft, property damage or loss is not covered. Participants are advised to take out adequate personal insurance (see also „Liability and Insurance“ in the printed version of the book of abstracts).

Participants are responsible for settling their hotel bills directly with the hotel on departure. The total price for participants staying at the Dom Hotel Limburg is 707 EUR and includes both accommodation for four nights and the conference package that covers lunches, dinners and coffee breaks as well as admits access to the conference room.

> **Participants not staying at the Dom Hotel Limburg are requested to register with the hotel for booking and paying the conference package, i.e. 309 EUR per person.**

Extras, such as drinks, telephone calls etc. are not included in the price.

The Symposium

The symposium will be held from 25 to 27 June, 2019, with the 24th and the 28th for travelling.

The setting and the limited number of participants (max. 60 persons) provide a very convivial atmosphere for the ready exchange of thoughts and ideas.

The scientific program will take place over three days and will

start at 9.00 am on Tuesday, the 25th, and

end in the late afternoon (ca. 5.30 pm) on Thursday, the 27th.

If you wish to extend your stay, please contact the hotel directly.

For the **length of the individual talks**, please refer to the program. Speakers should allow sufficient time for discussion at the end of their talks (e.g. a 35 min slot includes 25 min talk + 10 min for questions). We will have an LCD projector connected to a Windows PC available.

Scientific Program

Monday, 24 June

19.00 Welcome reception

19.30 Dinner

Tuesday, 25 June

09.00	Opening and Introductory Remarks	Carsten Kettner
	Session Chair: Kiyoko Aoki-Kinoshita	
09.15	Structural Investigations of the Roles of Glycans in Infection, Immunity and Cancer	Catherine E. Costello
09.50	Glycome of Stem Cells: From Structural Analysis to Social Implementation	Hiroaki Tateno
10.25	Poster Flash Presentation #1 Posters 1 - 4	F. Bonnardel, F. Schuhmacher, J. Ereño-Orbea, O. Grant
10.45	*Coffee Break and Poster Session*	
11.15	Mining Glycan Microarray Data using Anti-carbohydrate Binding of Antibodies in Human Serum	Richard D. Cummings
11.50	Targeting Siglecs to Suppress Allergies	James C. Paulson
12.25	Software Elevator Talks #1 Software 1 – 3 *chaired by René Ranzinger*	I. Walsh, O. Grant,
12.40	*Lunch*	
	Session Chair: Ulrika Westerlind	
14.00	Giant dsDNA Viruses Taste for Sugars	Cristina de Castro
14.35	Understanding the Molecular Details of Mucin-type O-Gycosylation by Glycosyltransferase Bump-and-hole Engineering	Benjamin Schumann
15.10	Poster Flash Presentation #2 Posters 5 - 8	D. Ungar, K. Paschinger, W.E. Hackett, T. Phung

15.30	*Conference Photo, Tea Break and Poster Session*	
16.00	Glycoprotein Data Visualization in GlyCosmos	Kiyoko Aoki-Kinoshita
16.35	Challenges for the CAZy Database in the Era of High-throughput Sequencing and Functional Screening	Nicolas Terrapon
17.10	Some (in)consistency Issues in Grouping Glyco-entities	Frédérique Lisacek
17.45	Software Elevator Talks #2 Software 4 – 5 *chaired by René Ranzinger*	Y. Akune, N. Karlsson,
17.50	Poster Flash Presentation #3 Posters 9 - 12	N. Karlsson, P. Suchánkova, S. Moon, J. Mariethoz
18.10	*Close*	
19.30	*Dinner*	

Back to Program

Wednesday, 26 June

	Session Chair: Benjamin Schumann	
09.00	Standardization of Glycosaminoglycan (GAG) Sequences binding to Proteins and Creation of a Pipeline for the Curation of GAG-protein Interactions: Application to GAG Interaction Networks	Sylvie Ricard-Blum
09.35	Glycomics and Proteomics of Brain Pathologies	Joe Zaia
10.10	*Coffee Break and Poster Session*	
10.40	Contextualized Functions of Glycans in Human Tissue Formation	Hans H. Wandall
11.15	Challenges in Defining the Structure and Function of Invertebrate *N*-Glycans	Iain B.H. Wilson
11.50	Software Elevator Talks #3 Software 7 – 8 *chaired by René Ranzinger*	S. Chatterjee, J. Mariethoz
12.00	*Lunch*	
13.15	Software Demo Session *chaired by René Ranzinger*	
	Session Chair: Frédérique Lisacek	
14.15	Sugars in the Gas Phase – from Structure to Reaction Mechanism	Kevin Pagel
14.50	Glycomimetic Tools to Understand Decoding of the Glycan Code	Christoph Rademacher
15.25	Dimensions of Glycomimetics	Thisbe K. Lindhorst
16.30	Excursion Meeting in the hotel lobby	
19.30	*Dinner*	

Thursday, 27 June

	Session Chair: Robert Woods	
09.00	Glycans as Biomarkers and Functional Effectors in Diabetes and Cardiovascular Diseases	Gordan Lauc
09.35	Update on the 1st Human Glycoproteomics Initiative (HUPO/HGI) Study: Interlaboratory Evaluation of Software for Intact Glycopeptide Analysis by MS	Nicolle H. Packer
10.10	Glycoscience in the Algae	Robert A. Field
10.45	*Coffee Break*	
11.15	Carbohydrates and Glycomimetics from the Viewpoint of a Computational Chemist	Martin Frank
11.50	From Symbolic Carbohydrate Notations to Atomic Coordinates	Philip Toukach
12.25	*Lunch*	
	Session Chair: Niclas G. Karlsson	
13.30	GlyGen - Computational and Informatics Resources for Glycosciences	René Ranzinger
14.05	Introducing UniCarbKB 2.0 and SPRIT-Gly	Matthew Campbell
14.40	*Tea Break*	
15.10	GlyFinder and GlyProbity: New Online Tools for Locating and Curating Carbohydrate Structures in wwPDB	Rob Woods
15.45	O-Glycologue: a Simulator of the Enzymes of O-linked Glycosylation	Andrew McDonald
16.20	MIRAGE Presentation and Discussion	René Ranzinger, Carsten Kettner
17.30	*Closing Remarks*	Carsten Kettner
19.30	*Dinner*	

Back to Program

List of Posters

The poster presentation includes a short (5 min) oral presentation on Tuesday, 25 June, and the poster sessions during the coffee breaks on Tuesday and Wednesday morning. The posters will be displayed throughout the entire symposium from Tuesday, 25 June, to Thursday, 27 June.

Tuesday, 25 June

#1	UniLectin3D, a Database of Carbohydrate Binding Proteins with Curated Information on 3D Structures and Interacting Ligands	François Bonnardel
#2	The Raspberry Synthesizer: An Open Source Chemical Laboratory Toolbox where IoT 4.0 meets Chemistry	Frank Schuhmacher
#3	Targeting CD22 with Sialic Acid Derivatives	June Ereño Orbea
#4	Predicting N-Glycan Processing based on Enzyme-glycan Accessibility	Oliver Grant
#5	Modelling Glycan Processing to Probe Golgi-enzyme Organization	Daniel Ungar
#6	Isoabaric and Iosmeric Glycans – m/z alone is not enough	Katharina Paschinger
#7	RAMZIS: Ranking Assessment of m/z Identifications by Similiarity	William E. Hackett
#8	DIALib, an Automated Workflow for Theoretical Ion Library Generation for Peptides and Glycopeptides	Toan Phung
#9	Altered O-linked Glycosylation in Osteoarthrithis Influence Galectin-3 Organisation of the Lubrication of the Cartilage	Niclas G. Karlsson
#10	GlyConvert – Bridging Glycoinformatics and Cheminformatics	Pavla Suchánkova
#11	In-silico Prediction of Chemical Glycosylation	Sooyeon Moon
#12	GlyConnect Compozitor – an Interactive Graph of Glycan	Julien Mariethoz

Back to Program

List of Software Demonstrations

The software demonstrations includes a short (2 min) oral presentation on Tuesday, 25 June, and Wednesday, 26 June, as well as the demo session on Wednesday afternoon (1.45 pm).

Tuesday, 25 June

#1+ #2	Two Software Solutions for Automated Structural Characterization of Glycans Using either HILIC-UPLC-MS or HILIC-UPLC-IMS-MS Glycomics Workflows	Ian Walsh
#3	Anyone Can Display Glycan Symbols in 3D Structures: 3D-SNFG in LiteMol	Oliver Grant
#4	An Update on new Glycan Array Data Software CarbArrayART: ready for Beta Testing	Yukie Akune
#5	The MIRAGE Compatible UniCarb-DR Repository for MS Glycomics	Niclas G. Karlsson
#6	Bioinformatics Tools for Glycoscience	René Ranzinger (to be presented in his talk on Thursday, 27 June)

Wednesday, 26 June

#7	'Plug and Play' with Machine Learning	Sourav Chatterjee
#8	GlyConnect, the Power Engine of Glycomics@ExPASy	Julien Mariethoz
#9	GlyCosmos Portal and GLIC	Kiyoko Aoki-Kinoshita (to be presented in her talk on Tuesday, 25 June)

Back to Program

Abstracts

Tuesday

09.15

Structural Investigations of the Roles of Glycans in Infection, Immunity and Cancer

Catherine E. Costello

Boston University School of Medicine
Department of Biochemistry, Physiology & Biophysics, and Chemistry
Boston, MA, United States of America

Increased understanding of infection, cancer, and the immune system is providing us with powerful antibody-based drugs and diagnostic tools. Infectious agents usually gain entrance to their hosts through the interactions of surface molecules. The immune system is responsible for and exploits the interactions of proteins with one another and with glycans (and other classes). Glycan modifications in cancer affect cell-cell interactions and metastasis and provoke immune responses. In order to explore these phenomena, to investigate how the body can combat infection and/or cancer, and to utilize this knowledge to control disease, we now rely heavily on the tools provided by mass spectrometry (MS) and the Big Data that they yield. This lecture will present MS and tandem MS approaches that we are developing further and using to elucidate the critical players in these pathways.

Back to Program

Tuesday

09.50

Glycome of Stem Cells: From Structural Analysis to Social Implementation

Hiroaki Tateno

National Institute of Advanced Industrial Science and Technology (AIST)
Biotechnology Research Institute for Drug Discovery
Tsukuba, Japan

Human stem cells such as human embryonic stem cells (hESCs), human induced pluripotent stem cells (hiPSCs), and human mesenchymal stem cells (hMSCs) are attractive cell sources for regenerative medicine due to their differentiation potential and proliferation ability. Since glycans are abundant components of the cell surface, reagents that specifically recognize stem cells should be useful tools for the identification, isolation, and manipulation of stem cells. In addition, glycans should be important for stem cell functions. We undertook a global study of the glycome of hESCs, hiPSCs, and hMSCs using lectin microarray, DNA microarray, and mass spectrometry/high-performance liquid chromatography. hiPSCs generated from any somatic cells expressed glycan structures similar to hESCs. Both hiPSCs and hESCs characteristically expressed the three glycan epitopes such as α2-6Sia, α1-2Fuc, and type1 LacNAc. From the glycome analysis of hMSCs with different differentiation potential, α2-6 sialylated N-glycans were found to be closely associated with the differentiation potential of stem cells. Interestingly, rBC2LCN with specificity to Fucα1-2Galβ1-3 motif was found to specifically recognize hESCs/hiPSCs. rBC2LCN is a practical probe for stem cell researches, which can be used for live staining, non-destructive detection, and elimination of tumorigenic hESCs/hiPSCs. These technologies are expected to overcome the tumorigenic risk of hESCs/hiPSCs, which is the major hurdle to apply hESCs/hiPSCs for regenerative medicine. rBC2LCN-based technologies are now commercially available from Japanese company. Recently, we analyzed the glycome of extracellular vesicles (EVs) derived from hESCs, hiPSCs, and hMSCs. We demonstrated for the first time that the characteristic glycan signature of hiPSCs are retained by EVs derived from them. In this presentation, I will talk about the basics and social implementation of the glycome of stem cells.

References

1. Saito et al. *Sci Rep.* **2018** Mar 5;8(1):3997
2. Hasehira et al. *Glycoconj J.* **2017** Dec;34(6):797-806.
3. Tateno et al. *Glycobiology* **2016**, 26(12):1328-1337
4. Hasehira et al. *Mol Cell Proteomics*. **2012** Dec;11(12):1913-23.

Back to Program

Tuesday

Poster
#1

UniLectin3D, a Database of Carbohydrate Binding Proteins with Curated Information on 3D Structures and Interacting Ligands

François Bonnardel[1], Julien Mariethoz[2], Sebastian Salentin, Xavier Robin, Michael Schroeder, Serge Perez[1], Frédérique Lisacek[2], Anne Imberty[1]

[1]CNRS – Cermav, Grenoble, France
[2]Swiss Institute of Bioinformatics, Geneva, Switzerland

Lectins, and related receptors such as adhesins and toxins, are glycan-binding proteins from all origins that decipher the glycocode, i.e. the structural information encoded in the conformation of complex carbohydrates present on the surface of all cells. Lectins are still poorly classified and annotated, but since their functions are based on ligand recognition, their 3D-structures provide a solid foundation for characterization. UniLectin3D is a curated database that classifies lectins on origin and fold, with cross-links to literature, other databases in glycosciences and functional data such as known specificity. The database provides detailed information on lectins, their bound glycan ligands, and features their interactions using the Protein–Ligand Interaction Profiler (PLIP) server. Special care was devoted to the description of the bound glycan ligands with the use of simple graphical representation and numerical format for cross-linking to other databases in glycoscience. We conceived the design of the database architecture and the navigation tools to account for all organisms, as well as to search for oligosaccharide epitopes complexed within specified binding sites. UniLectin3D is accessible at https://www.unilectin.eu/unilectin3D.

Back to Program

Tuesday

Poster
#2

The Raspberry Synthesizer: An Open Source Chemical Laboratory Toolbox where IoT 4.0 meets Chemistry

Frank Schuhmacher

Free University Berlin / Elexon Research

To produce reproducible chemistry is a daunting task for any chemist. Automated chemistry systems are available in various versions. Their functionality often differs significantly due to different application focuses.

The most important task of an automation system is to always do the same with consistent quality and identical results. Another often expressed wish is cheap and usable for all users with an efficient training period.

Which steps are necessary for this and is this possible soon?

One of the most important requirements for complete automation is a comprehensive standardization. Only when a process is completely mastered and describable in all its single steps, IT has technical possibilities for programmatic implementation. The course of a chemical synthesis should be fully conceptualized in an electronic laboratory book before performing. If there are some describing rules, synergy effects for subsequent automation can be achieved. This necessary basics has already been laid in the implementation of the first new oligosaccharide synthesizer generation. Here, the modular structure was consistently implemented for both the hardware and software. The valve & pump controllers, pressure & temperature controls, as well as fraction collectors, were implemented in this first realization. The possibilities of automated purification as well as the separation by photochemical processes should be integrated next into the automation equipment. Online analyses of the current synthesis also require the active feedback of control information that results in a systematic improvement of the local parameter guidance, and the data obtained should be automatically stored in a database in order to be used continuously. To provide a functional data pool from the new syntheses then dynamically proposed to the user.

This makes it possible for new users after a short training period to be able to set up new syntheses without having to deal with the hardware.

In order to make this functionality available to as many labs/users as possible, the possibilities offered by the usage of Markerboards in the realization of hardware modules. This new class of chemical instruments helps in the daily lab work. It can be used directly in individual applications. Expand to any complex system by combination.

Back to Program

Tuesday

Poster
#3

Targeting CD22 with Sialic Acid Derivatives

June Ereño-Orbea[1], Lijuan Pang, Hong Cui, Dorota Borovsky, Corwin Nycholat, Christoph Rademacher[2], Jim Paulson[3] and Jean-Philippe Julien[4]

[1]CICbioGUNE, Derio, Spain
[2]Max Planck Institute of Colloids and Interfaces, Potsdam, Germany
[3]The Scripps Research Institute, La Jolla, CA, United States of America
[4]McGill University, Toronto, Canada

CD22 is a sialic acid-binding immunoglobulin-like lectin (Siglec) that maintains a baseline level of B cell inhibition. Its function and restricted expression in B cells makes CD22 a validated target in therapies against deregulated B cells that cause B cell lymphomas and autoimmune diseases. High-affinity sialic acid based ligands that will compete with natural sialic acid ligands to bind CD22 represent a promising therapeutic tool. Here, we describe the design and synthesis of a sialoside library constructed by chemical modifications on carbon substituents C2, C5 and C9 on the natural Neu5Ac scaffold. Subsequent analysis of binding to human CD22 revealed that addition of non- carbohydrate groups, such as 2,3-dichlorobenzyl, at C2 can improve the affinity towards CD22 from high micromolar to sub-micromolar KD values. Moreover, we have solved the crystal structures of the most N-terminal three Ig domains (d1-d3) of human CD22 in complex with three different sialic acid derivatives. Our results provide a strategy to generate high affinity sialic acid molecules against CD22 that will outcompete the natural ligand for the receptor and modulate its activity.

Back to Program

Tuesday

Poster
#4

Predicting N-glycan Processing Based on Enzyme-glycan Accessibility

Oliver C. Grant, Rob J. Woods

University of Georgia, CCRC, Athens, GA, United States of America

Background:

In this work, computer simulation, glycoproteomics and crystallographic data are combined to show that glycoprotein glycoform distributions depend on the accessibility of N-glycans to the relevant glycosidases in the ER. We illustrate this for three systems: a protein disulfide isomerase percursor named Pdi1p, a hemagglutinin (HA) from influenza A, and the HIV envelope protein.

Methodology:

We leverage the recently solved 3D structure of ER mannosidase I (ERManI) with molecular dynamics simulations of the glycoproteins, where the enzyme's substrate, Man9GN2, is present at each site. We calculate the percentage of simulation time that ERManI is physically able to bind glycan, as it samples different shapes throughout the simulation.

Results:

In the case of Pdi1p, the correlation between accessibility to ERManI and the degree of processing is striking. The modeling also predicted that a domain deletion would expose a glycosylation site on a neighboring domain to processing, which was confirmed experimentally. For influenza HA, we were able to rationalize why certain sites remained as Man9GN2. Further, the modeling work was able to propose a 3D model for how the pulmonary collectin SP-D would bind these Man9GN2 glycans and thus neutralize the virus. When applied to the HIV envelope protein, our modeling approach provides insight into formation of the so-called "high-mannose patch".

Conclusion:

The degree to which site-specific glycan processing can be predicted on the basis of 3D-structure and dynamics is surprising, given the range of other factors involved.

Tuesday

11.15

Mining Glycan Microarray Data using Anti-carbohydrate Binding of Antibodies in Human Serum

Akul Y. Metha and Richard D. Cummings

Harvard Medical School
National Center for Functional Glycomics, Department of Surgery
Beth Israel Deaconess Medical Center,
Boston, MA, United States of America

Glycans are recognized by a dizzying panel of lectins, glycan-binding proteins, and antibodies. The anti-carbohydrate antibodies in human serum (termed the anti-carbohydrate antibody repertoire or ACAR) provide a window of insight into the unique adaptive immune responses of individuals and their potential exposure and susceptibility to glycan immunogens associated with disease. We have generated a large number of glycan microarrays of differing glycan content, including human glycans, animal and insect glycans, and microbial glycans, to explore the ACAR in many individuals. Our studies demonstrate an astonishing complexity in IgG, IgM and IgA antibodies to different glycan antigens in individuals from around the world, as well as those in commercial preparations of IgG, termed IVIG. In addition, we have identified differences and reductions in the ACAR in individuals with primary antibody deficiencies or PADs. A major difficulty in analyzing such data and comparative analyses is the size of the database and the lack of tools for such complex comparisons between individuals across a wide spectrum of glycan antigens.

To aid in this we have developed a tool termed GLAD, which is a web-based tool to visualize, analyze, present and mine glycan microarray data. With GLAD users can input multiple data files and develop comparative analyses on their own computer. GLAD has a number of key attributes, including the ability to visualize the data in grouped bar charts, create specific types of heatmaps, calendar heatmaps, force graphs, and correlation maps and bubble box plots. In addition, the tool allows the simultaneous viewing of the glycan structures in SNFG symbols for each data point. These visualizations and glycan structures can be downloaded by the users as vector SVG images of high quality which can be edited using illustration software for publications.

GLAD also allows the users to normalize, sort and filter the data in order to assist in the mining of the information from the array experiments. GLAD was made using mainly JavaScript based libraries (D3js, LoDash, JQuery, Select2, jStat), therefore it is not server-based but client-based. This means that data loaded into GLAD remains private and never gets uploaded to any server. Analysis can be saved as selections or entire sessions locally to the drive of the user. GLAD is freely available for use on the Web at https://glycotoolkit.com/GLAD/ with all major modern browsers (Edge, Firefox, Chrome, Safari). Online documentation includes screenshots as well as video guides to help users get started.

Back to Program

Tuesday
11.50

Targeting Siglecs to Suppress Allergies

James C. Paulson

The Scripps Research Institute
Departments of Molecular Medicine, and Immunology and Microbiology
La Jolla, CA, United States of America

The **s**ialic acid-binding **i**mmuno**g**lobulin-type **lec**tin family of cell adhesion receptors called Siglecs are predominately expressed on white blood cells of the immune system and help the immune system distinguish between self and non-self. Many siglecs carry inhibitory motifs in their cytoplasmic domain acting as checkpoint inhibitors that regulate immune responses. We have developed a liposomal nanoparticle platform that exploits the inhibitory functions of siglecs to suppress antigen mediated immune cell activation.

Key to this platform are synthetic ligands that bind with high avidity and high specificity for a single Siglec. The tolerogenic nanoparticles display an antigen and a ligand of an inhibitory Siglec expressed on the cells that recognize the antigen. Mast cells pre-sensitized with IgE to an allergen normally degranulate upon exposure to antigen and cause allergic symptoms including severe anaphylaxis. However, when exposed to the toleragenic nanoparticles, Siglecs are recruited to the immunological synapse suppressing the immune response and desensitizing the cell to subsequent antigen challenge.

(NIH grants AI050143, AI099141, HL107151).

Back to Program

Tuesday

Software #1+2

Two Software Solutions for Automated Structural Characterization of Glycans using either HILIC-UPLC-MS or HILIC-UPLC-IMS-MS Glycomics Workflows

Ian Walsh[1], Katherine Wongtrakul-Kish[1], Han Wang[2], Lyn Chiin Sim[1], Amelia Mak[1], Tasha Jose[2,3], Christopher H. Taron[3], Pauline M. Rudd[1], Terry Nguyen-Khuong[1]

[1]Bioprocessing Technology Institute, Agency for Science, Technology and Research, Analytics Group, Singapore
[2]Waters Pacific Pte Ltd, Singapore
[3]New England Biolabs, Ipswich, MA, United States of America

Comprehensive glyco-analytics involves extensive characterisation and distinguishing structural differences such as linkages, branching and anomericity of oligosaccharides. Whilst this has been addressed with improvements to analytical technologies to characterise glycosylation, data analysis remains one of the largest bottleneck in this workflow. In this work we present a glycoinformatics pipeline that improves this bottleneck, accurately identifying and quantitating glycans released from glycoproteins. One component of the pipeline, called MAGMap, leverages upon multiple attributes such as Glucose Units, m/z and Collision Cross Section values derived after fluorescently labelled glycans are analysed from a HILIC-UPLC-IMS-MS setup (Waters H-Class UPLC-SYNAPT® G2-S MS). MAGMap utilizes more attributes measured from labelled glycans than conventional methods and thus the precision of the glycomic characterisation was increased using our algorithm (10-20% improvements). Another software, called GlycanAnalyzer, allows for the precise confirmation of glycan structures by analysing data derived from HILIC-UPLC-MS after application of exoglycosidases. Exoglycosidase removal of monosaccharides results in signature peak shifts, in both UPLC and MS, yielding an effective way to pattern match N-glycan structure with high detail. Our laboratory has shown that the pipeline can reduce glycan data analytics from eight to two weeks thus substantially increasing the speed and throughput of traditional glycomic analytical workflows.

Tuesday

Software
#3

Anyone Can Display Glycan Symbols in 3D Structures: 3D-SNFG in LiteMol

Oliver C. Grant, Rob J. Woods

University of Georgia, CCRC, Athens, GA, United States of America

The representation of carbohydrates in 3D space using symbols is a powerful visualization method. This ~30 second software demo will show researchers how to display carbohydrate 3D structures as 3D-SNFG symbols in a web browser using LiteMol. A simple way to use the visualizer is via URL: v.litemol.org/?loadFromCS=5T3X. Any protein databank (PDB) ID can be substituted at the end of the URL. Alternatively, the user may enter a PDB ID or upload a structure from their computer to the visualizer at v.litemol.org.

Availability: LiteMol is available at v.litemol.org and automatically depicts any carbohydrate residues as 3D-SNFG symbols. Web developers can embed LiteMol in a webpage (visit https://github.com/dsehnal/LiteMol).

Back to Program

Tuesday
14.00

Giant dsDNA Viruses Taste for Sugars

Cristina de Castro

University of Napoli
Department of Agriculture
Napoli, Italy

Giant dsDNA viruses are gaining interest in the field of glycobiology because, and differently from other viruses, they are able to biosynthesize rare sugars and to glycosylate their own proteins by an apparent host independent process.

Giant dsDNA viruses consist of a large number of species and this lecture will focus on the recent developments that involve members of *Phycodnaviridae* and *Mimiviridae* families.

Regarding *Phycodnaviridae*, to date information is available for Chloroviruses. Chloroviruses are large (190 nm in diameter) icosahedral, plaque-forming viruses with an internal lipid membrane; they have genomes of 290 to 370 kb that contain up to 400 protein-encoding genes [1]. The prototype chlorovirus, *Paramecium bursaria chlorella virus* (PBCV-1), infects *Chlorella variabilis*, a symbiont of the protozoan *Paramecium bursaria*. The PBCV-1 major capsid protein Vp54 accounts for about 40% of the viral protein, and it is glycosylated by an unusual and complex oligosaccharide [2,3]. The glycobiology of PBCV-1 will be presented as well as information about its antigenic variants and other related chloroviruses [4,5].

As for *Mimiviridae*, these viruses infect *Acanthamoeba sp.* and were initially identified as bacteria because of their large size along with the heavily glycosylated fibrils of the capsid. Within *Mimiviridae*, information is available for *Mimivirus* and *Megavirus* genera, and this lecture will focus on the biosynthetic pathways of activated sugar precursors recently described for both viruses [6,7] and with the structural data available on *Mimivirus*.

References

[1] J.L.Van Etten. *Annu. Rev. Genetics*. **2003**, 37, 153-195.

[2] C. De Castro, A. Molinaro, F. Piacente, J. R. Gurnon, L. Sturiale, A. Palmigiano, R. Lanzetta, M. Parrilli, D. Garozzo, M. Tonetti, J. L.Van Etten. *PNAS*. **2013**, 110, 13956-13960.

[3] C. De Castro, T. Klose, I. Speciale, R. Lanzetta, A. Molinaro, JL Van Etten, MG Rossman. *PNAS*, **2018**, 115, E44-E52.

[4] I. Speciale, G.A. Duncan, L. Unione, I.V. Agarkova, D. Garozzo, J. Jimenez-Barbero, S. Lin, T.L. Lowary, A. Molinaro, E. Noel, M.E. Laugieri, M.G. Tonetti, J.L. Van Etten, C. De Castro. *J. Biol. Chem*. **2019**, 294, 5688-5699.

[5] C. De Castro, I. Speciale, G. Duncan, D.D. Dunigan, I. Agarkova, R. Lanzetta, L. Sturiale, A. Palmigiano, D. Garozzo, A. Molinaro, M. Tonetti, J.L. Van Etten. *Angew. Chem. Int. Ed.* **2016**, 55, 654-658.

[6] F. Piacente, M. Marin, A. Molinaro, C. De Castro, V. Seltzer, A. Salis, G. Damonte, C. Bernardi, J.-M. Claverie, C. Abergel, M. Tonetti. *J. Biol. Chem*. **2012**, 287, 3009-3018.

[7] F. Piacente, C. De Castro, S. Jeudy, A. Molinaro, A. Salis, G. Damonte, C. Bernardi, C. Abergel, M. Tonetti. *J. Biol. Chem.* **2014**, 289, 24428-24439.

Back to Program

Tuesday

14.35

Understanding the Molecular Details of Mucin-type *O*-Glycosylation by Glycosyltransferase Bump-and-hole Engineering

Benjamin Schumann

Imperial College London
The Francis Crick Institute
Department of Chemistry
London, United Kingdom

Mucin-type *O*-GalNAc glycosylation is among the most abundant yet least understood posttranslational modifications: more than 80% of all proteins trafficking through the secretory pathway are predicted to be *O*-GalNAc-glycosylated, and glycans are frequently altered during malignant transformation. Although contributing to the biophysical properties of the glycocalyx, increasing evidence suggests that distinct *O*-glycans are crucial mediators of biological processes. As part of the glyco-code, *O*-glycosylation is encoded by a family of 20 polypeptide GalNAc transferase (GalNAc-T) isoenzymes that introduce the first, Ser/Thr-bound GalNAc residue. Despite partial redundancy, distinct GalNAc-Ts have been associated with disease, suggesting that isoenzyme-specific protein substrates influence important biological processes. However, studying these substrates by glycoproteomics approaches is complicated by the interplay of different isoenzymes with each other and with other glycan-processing enzymes.

Here, we use a chemical biology method termed “bump-and-hole engineering” to dissect the protein substrates of distinct GalNAc-Ts in the native cellular environment. In a structure-guided process, a „hole” is engineered by mutagenesis into the nucleotide-sugar binding site of a GalNAc-T, rendering the enzyme compatible with a chemical functional group (“bump”) in the GalNAc portion of a synthetic UDP-GalNAc derivative. After glycosylation, a traceable chemical handle in the bump allows for the specific detection of glycoproteins by bioorthogonal ligation. Establishment of a GalNAc-T bump-hole pair in the living cell and glycoproteome characterisation reveals details on isoenzyme-specific glycosylation site and glycan structure in a single experiment.

Back to Program

Tuesday

Poster
#5

Modelling Glycan Processing to Probe Golgi-enzyme Organization

Daniel Ungar

University of York, York, United Kingdom

The decoration of proteins by carbohydrates is essential for eukaryotic life, yet heterogeneous due to lacking biosynthetic templates. This complex carbohydrate mixture – the glycan profile – is generated in the compartmentalised Golgi, where level and localization of glycosylation enzymes are key determinants. Here we develop and validate a computational model for glycan biosynthesis to probe how the biosynthetic machinery creates different glycan profiles. We combined stochastic modelling with Bayesian fitting that enables rigorous comparison to experimental data despite starting with uncertain initial parameters. This is an important development in the field of glycan modelling, which revealed biological insights about the glycosylation machinery in altered cellular states. We experimentally validated changes in N-linked glycan-modifying enzymes in cells with perturbed intra-Golgi enzyme sorting and the predicted glycan-branching activity during osteogenesis. Our model can provide detailed information on altered biosynthetic paths, with potential for advancing treatments for glycosylation related diseases and glycan-engineering of cells.

Back to Program

Tuesday

Poster
#6

Isobaric and Isomeric Glycans – m/z alone is not enough

Katharina Paschinger, A. Hykollari, B. Eckmair, S. Yan, J. Vanbeselaere, and Iain B.H. Wilson

Universität für Bodenkultur, Department of Chemistry, Vienna, Austria

Glycan analysis is a challenge and there are many efforts to make it simpler. These days, glycomic approaches are heavily based on mass spectrometry, but relying on the mass to predict the structure will lead to misleading conclusions. Furthermore, many workflows do not rely on fractionating the glycomes before analysis; thus, interesting, low-abundance structures with non-standard ionisation properties are overlooked. Therefore, when we look at unusual glycomes, central aspects of the workflow are pre-fractionation to enrich anionic or more hydrophobic structures, 1D and 2D-HPLC, MS and MS/MS in both positive and negative modes as well as chemical or enzymatic digestions. Thereby, we can resolve glycans of the same mass with the same compositions (in terms of hexoses, N-acetylhexosamines, deoxyhexoses, etc.), but different isomeric structures, or those with very similar masses (more-or-less isobaric), but which have additional building blocks such as sulphate, phosphate or phosphoesters. Various examples from our work on gastropods, slime moulds, nematodes and insects highlight that careful and orthogonal assessment of glycan properties are required in order to solve the structures in an adequate manner.

Acknowledgements

This work was supported by the Austrian Science Fund (FWF; grants P23922, P25058, P26662 and W1224) and the European Union (PITN-GA-2013-608295).

Back to Program

Tuesday

Poster
#7

RAMZIS: Ranking Assessment of m/z Identifications by Similarity

William E. Hackett

Boston University, Department of Bioinformatics,
Boston, MA, United States of America

Glycoproteomics enable deeper understanding of immunology, neurology, cancer, and more, and it can allow for better treatments of diseases. But the field is hindered by a balancing act of sample size, missing values, discovery, time, and cost; current experiments are limited in the appropriate statistical conclusions they can draw and can leave researchers with ambiguous results due to the stochastic nature of the underlying biological processes and the stochastic nature of the data acquisition process. RAMZIS (Ranking Assessment of m/z Identifications by Similarity) is a tool developed to compare glycosylation sites. It performs a data quality assessment, provides a visualized comparison of glycosylation sites via a weighted similarity measure, and ranks the contribution of individual glycosylations via a secondary abstraction. These tools are intended to guide researchers in the design of follow up experiments and draw general conclusions about their experiment and its methodology. Its data quality assessment determines if a researcher should draw any conclusions from a data set, and its visualized comparison allows researchers to determine how likely they are to identify differences in their data or if one group may be a subgroup of the other. The ranking procedure allows for more targeted follow up experiments and guided analysis.

Back to Program

Tuesday

Poster
#8

DIALib: an Automated Workflow for Theoretical Ion Library Generation for Peptides and Glycopeptides

Toan K. Phung[1], Lucia F. Zacchi[1,2], Nazmi Nazir[1], Aaron Li[1], Benjamin L. Schulz[1,2]

[1]The University of Queensland, School of Chemistry and Molecular Biosciences, St. Lucia, Qld, Australia,

[2]The University of Queensland, ARC Centre for Biopharmaceutical Innovation, St. Lucia, Qld, Australia

With the advent of Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH), in which fragmentation of the precursor ion is independent of its intensity, the amount of data in a single LC-MS experiment has grown exponentially. A common practice for analysis of SWATH data is to use spectral libraries made from data dependent acquisition (DDA) MS experiments. Due to the intensity cutoff for fragmentation during DDA, to the sometimes high complexity of the sample, and to the presence of post-translational modifications, precursor ions can become suboptimally represented in the library. In particular, heterogeneity of glycans is one of the biggest challenges in glycopeptide identification and quantification in SWATH, because standard proteomic softwares are only able to identify a these complex molecules with difficulty.

To get around this problem, it is possible to create a comprehensive theoretical spectral library for quantification of SWATH data to be used with downstream MS analysis software like PeakView (AB SCIEX). However, manually creating theoretical libraries for SWATH is a tedious procedure. For this reason we designed DIALib, a software that assists in the construction of glycopeptide spectral libraries for Y, oxonium, and *b* and/or *y* ions, to be used with PeakView, for DIA quantification. The software allows the selection/combination of type of ion (oxonium, Y, *b*, *y*), post-translational modification (preset or customized), retention time, precursor mass (preset or customized), precursor and fragment ion charge, and transitions for library construction. Using several libraries generated by DIALib, we re-analyzed our published yeast cell wall glycoproteomics dataset and analysed a new serum-derived Immunoglobulin sample.

The new libraries allowed us to identify and quantify glycoform abundance from these samples, recapitulating previously published data obtained using DDA and manually derived libraries. In addition, adaptation of the software functionality for building a library consisting of oxonium ions can quickly provide a preview of whether there are glycosylation differences between two test samples.

Back to Program

Tuesday
16.00

Glycoprotein Data Visualiation in GlyCosmos

Kiyoko Aoki-Kinoshita

Soka University
Faculty of Science and Engineering
Glycan Life Science Integration Center
Tokyo, Japan

We have released the GlyCosmos Glycoscience Web Portal (https://glycosmos.org) on April 1, 2019, as an official portal of the Japanese Society for Carbohydrate Research (JSCR). As such, GlyCosmos will work closely with the Japanese carbohydrate research community to provide up-to-date information about data related to glycans, including glycogenes, glycoproteins, pathways, and diseases.

In this talk, I will present out work on the visualization of the variety of data available in GlyCosmos. For glycoproteins, data from UniProt, GlycoProtDB and GlycoNAVI proteins have been accumulated, and we have developed a user interface to reflect these data from a single page. Glycosylation sites on proteins can be visualized and compared, and relevant data can be accessed by the citations to the original data resources. For pathways, we have developed a tool to visualize the data from Resource Description Framework (RDF) format, including cellular localization and linking to the glycoprotein entry page. Furthermore, we have been working on a visualization tool for the basic known glycosylation pathways, which can reflect glycogene expression values. Moreover, the GlycomeAtlas tool originally developed in RINGS is now also available from GlyCosmos, including new glycomics data from zebrafish.

Data visualization is an important aspect of any bioinformatics resource, determining how much it is used by the community. With the increasing number of useful software tools, it is expected that incorporation into the glycosciences will become easier in the future.

Back to Program

Tuesday

16.35

Challenges for the CAZy Database in the Era of High-throughput Sequencing and Functional Screening

Nicolas Terrapon

Aix-Marseille University
Department of Bioinformatics
Marseille, France

Twenty-eight years ago, the classification of ~300 amino-acid sequences into 35 glycoside hydrolase (GH) families was at the origin of the reference CAZy database. Every day, a handful of CAZy curators semi-manually annotate the newly released protein sequences from Genbank, PDB, as well as JGI fungal genomes. Our duty also involves surveillance of the literature and integration in CAZy of newly discovered families and reported protein activities. As of today, there are more than 160 GH families found across more than 600,000 GH-encoding proteins from GenBank, over 13,000 of which have a published activity which was checked and reported by CAZy curators. Curiously the major challenge that we face is not the volume of sequence to process but the capture of functional information, especially given the increasing pace of biochemical characterizations and the multiplicity of journals where CAZyme biochemistry may be reported, etc.

In this context, I will present several strategies currently discussed in the CAZy team in order to ensure the maintenance of CAZy database in the future, at its level of quality and usefulness for the community.

Back to Program

Tuesday

17.10

Some (in)consistency Issues in Grouping Glyco-entities

Julien Mariethoz, François Bonnardel, Thibault Robin, Oliver Horlacher, Frédérique Lisacek

SIB Swiss Institute of Bioinformatics
Geneva, Switzerland

The overall purpose of our Glycomics@ExPASy initiative (www.expasy.org/glycomics) is to provide web-based resources, which put together reflect a consistent picture of the structural and functional aspects of glycosylation. Two recent inclusions in the collection, i.e., GlyConnect and UniLectin [1,2] exemplify this purpose.

To feed information into each of these two platforms, we process large datasets collected on-line. This larger scale approach imposes grouping and classifying and raises new issues relative to defining appropriate yet not too constraining comparison criteria. We will discuss this thin balance through examples of profiling glycomes, glycoproteins and lectin families. We will also present new tools developed to that end.

References

[1] Alocci D, Mariethoz J, Gastaldello A, Gasteiger E, Karlsson NG, Kolarich D, Packer NH, Lisacek F (2019) GlyConnect: glycoproteomics goes visual, interactive and analytical Software tools and data resources Special issue of *J. Prot. Research* 18(2):664-677

[2] Bonnardel F, Salentin S, Robin X, Perez S, Lisacek F, Imberty A (2019) UniLectin3D, a database of carbohydrate binding proteins with curated information on 3D structures and interacting ligands, *Nucleic Acids Res.* 47 (Database issue): D1236–D1244.

Back to Program

Tuesday

Software
#4

An Update to a Software Tool for Glycan Array Data: CarbArrayART. Ready for Beta Testing

Yukie Akune[1], Sena Arpinar[2], Lisete M. Silva[1], Mark Stoll[1], Angelina S. Palma[3], Yan Liu[1], René Ranzinger[2], and Ten Feizi[1]

[1]Imperial College London, Department of Medicine, Glycosciences Laboratory, London, United Kingdom
[2]University of Georgia, CCRC, Athens, GA, United States of America
[3]NOVA University of Lisbon, UCIBIO-Faculty of Science and Technology, Lisbon, Portugal

We have developed a freely available software, named CarbArrayART (Carbohydrate microarray Analysis and Reporting Tool), for storage, processing and presentation of carbohydrate microarray data. CarbArrayART is based on GRITS Toolbox [Weatherly et al, Glycobiol., 29(6) 452-460, 2019] which was originally developed for processing, interpretation and archiving glycomic mass spectrometry data. GRITS contains functions for storing glycan structures and also metadata, such as project information, sample description and experiment details. We are capitalizing on these functions by developing and implementing CarbArrayART as an extension (plugin) of GRITS Toolbox.

The main features of CarbArrayART are: (i) Storing carbohydrate microarray data from different array formats. Users are able to store the scan data of the binding experiments, e.g. gpr (GenePix) and excel (ProScan) files, and related array-specific metadata, such as glyco-probe lists, array geometry, information on the glycan binding system and experiment protocol information. (ii) Presentation of data with filtering and sorting functions. Tables and charts are generated automatically based on the stored scan data. Filter and sort operations can be applied to the display datasets. (iii) Reporting of data. We have extended the GRITS Toolbox functionality of generating reports from metadata by allowing export conforming to MIRAGE guidelines for glycan microarray-based data [Liu et al, *Glycobiol.*, 27(4) 280-284, **2017**].

Since the first presentation of the CarbArrayART prototype at the 2017 Glyco-Bioinformatics symposium we have improved flexibility of the system by introducing several new input functions for carbohydrate microarray-associated data. These include metadata for carbohydrate-binding samples, microarray experiments and array geometry for multiple glyco-probe layouts. We are presenting the first public version of CarbArrayART which is available to the community for testing.

Back to Program

Tuesday

Software
#5

The MIRAGE Compatible UniCarb-DR Repository for MS Glycomics

Niclas G. Karlsson

University of Gothenburg, Department of Medical Biochemistry and Cell Biology, Gothenburg, Sweden

High throughput glycomics mostly based on mass spectrometry (MS) increasingly depends on databases and software. This requires agreeing on a format for accurately recording of experiments, developing consistent storage modules and granting public access to large glycomic MS data sample sets. The introduction of the MIRAGE reporting standards for glycomics was the first step towards automating glycomic data recording but implementing the MIRAGE guidelines entails not only addressing “What?” but also “How?” to record MS glycomics. Here we present a newly developed web tool for recording MIRAGE records for glycomic MS data available at https://unicarb-dr.biomedicine.gu.se.

This file together with fragmentation data of glycans recorded in Glycoworkbench could then be submitted into the repository UniCarb-DR, as part of a MIRAGE compliant scientific publication. The repository was designed to store submitted MIRAGE information and display fragment spectra and metadata for pre- and post- scientific publishing reviewing.

Back to Program

Tuesday

Poster
#9

Altered *O*-linked Glycosylation in Osteoarthrithis Influence Galectin-3 Organization of the Lubrication of the Cartilage

Sarah A. Flowers, Kristina A. Thomsson, Liagat Ali, Shan Huang, Yolanda Mthembu, Radiosa Gallini, Jan Holgersson, Tannin A. Schmidt, Ola Rolfson, Lena I. Björkman, Martina Sundqvist, Anna Karlsson, Gregoy Jay, Masood Kamali-Moghaddam, Thomas Eisler, Roman Krawetz, Niclas G. Karlsson

University of Gothenburg, Department of Medical Biochemistry and Cell Biology, Gothenburg, Sweden

Synovial fluid lubricin (proteoglycan 4) is a mucin-type O-linked glycosylated (60% of the mass) biological lubricant that can be cross-linked by pentameric synovial galectin-3. Lubricin glycopeptides from patients with Osteoarthritis (OA) and recombinant lubricin (rhPRG4) identified by LC-MS, were compared to previously established 268 O-linked glycosites of lubricin. Dominating short Tn (GalNAca1-) and T (Galb1-3GalNAca1-)glycans with or without sialic acid and sulfate, were found in the lubricin mucin domain. Using glycomics, the level of these truncated and less sialylated glycans was found to be increased in late-stage OA patients. Healthy individuals' lubricin instead contained a significantly higher amount of core-2 (Galb1-3(GlcNAcb1-6)GalNAca1-)structures, while rhPRG4 was totally devoid of core-2. Binding to galectin-3 binding was found to be dependent on core-2, and both the galectin-3 level and galectin-3 interaction with OA lubricin were decreased in OA patients, suggesting a defect crosslinking of lubricating molecules in OA providing novel insights into OA molecular pathology.

Back to Program

Tuesday

Poster
#10

GlyConvert – Bridging Glycoinformatics and Cheminformatics

Pavla Suchánkova[1,2], Julien Mariethoz[3,4], Radka Svobodová[1,2], Frédérique Lisacek[3,4]

[1]Masaryk University Brno, National Centre for Biomolecular Research, Brno-Bohunice, Czech Republic

[2]Masaryk University Brno, CEITEC – Central European Institute of Technology, Brno-Bohunice, Czech Republic

[3]SIB Swiss Institute of Bioinformatics, Geneva, Switzerland

4University of Geneva, Switzerland

Format sharing between biochemistry, chemical biology, carbohydrate chemistry, and glycobiology is a challenge. In fact, the diversity of information sources has created barriers of communication between these fields. While standards such as InChI (International Chemical Identifier), InChIKey and SMILES (Simplified Molecular Input Line Entry Specification) have been widely adopted in chemical biology and cheminformatics, several encoding formats have been proposed as attempts to account for the branched structures of glycans. In the end, toolboxes of converters have been developed over the years in order to cope with this variety.

Admittedly in recent years, GlycoCT [1] has emerged as a structure encoding format shared by a wide range of glyco-databases but GlycoCT has no connection to SMILES or InChi. Over the same period of time, SMILES and InChi have become ever-present in data exchanged between the reference PubChem [2] and ChEBI [3] compound databases and many useful bioinformatics resources spanning protein, enzyme and pathway knowledge. In yet another attempt to connect glycobiology with other fields of Life Sciences, we designed a pipeline to translate GlycoCT into SMILES and InChi thereby bridging glyco- and chemo-informatics with the goal of extending cross-references of our glyco-databases with popular bioinformatics resources.

We also offer the option of exporting glycan structures in those formats via a glycan drawing interface [4].

References

[1] Herget S, Ranzinger R, Maass K, Lieth CW. GlycoCT—a unifying sequence format for carbohydrates. *Carbohydrate Research*. **2008**, 343(12), 2162-2171.

[2] Sunghwan K, Thiessen PA, Bolton EE, et al. PubChem Substance and Compound databases. *Nucleic Acids Research*. **2016**, 44(D1), D1202-D1213.

[3] Hastings J, Owen G, Dekker A, Ennis M, Kale N, Muthukrishnan V, Turner S, Swainston N, Mendes P, Steinbeck C. ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res*. **2016**, 44(D1), D1214–D1219.

[4] Alocci D, Suchánková P, Costa R, Hory N, Mariethoz J, Svobodová R, Toukach P, Lisacek F. SugarSketcher: Quick and Intuitive Online Glycan Drawing. *Molecules*. **2018**, 23(12):3206

Back to Program

Tuesday

Poster
#11

In-silico Understanding of Glycosylation through Automation and Machine Learning

Sooyeon Moon[1,2], Sourav Chatterjee[1], Peter. H. Seeberger[1,2], Kerry Gilmore[1]

[1]Max Planck Institute of Colloids and Interfaces, Department of Biomolecular Systems, Potsdam, Germany

[2]Freie Universität Berlin, Institute of Chemistry and Biochemistry, Berlin, Germany

Predicting stereo-chemical outcome of chemical glycosylation is one of the most challenging problem in organic chemistry due to sheer number of factors dictating the reaction outcome. Gaining a detail mechanistic understanding of such complicated reaction is often a formidable challenge. Recently we have performed around

300 reproducible experiments on model glycosylation reaction on a fully automated platform. We have intercepted several factors affecting glycosylation such as temperature, reaction stoichiometry, equivalents, and influence of donor, acceptor, activator and solvents without human intervention [1]. These data presented itself as a perfect opportunity for application of machine learning algorithms in order to predict stereochemical selectivity of model glycosylation reactions.

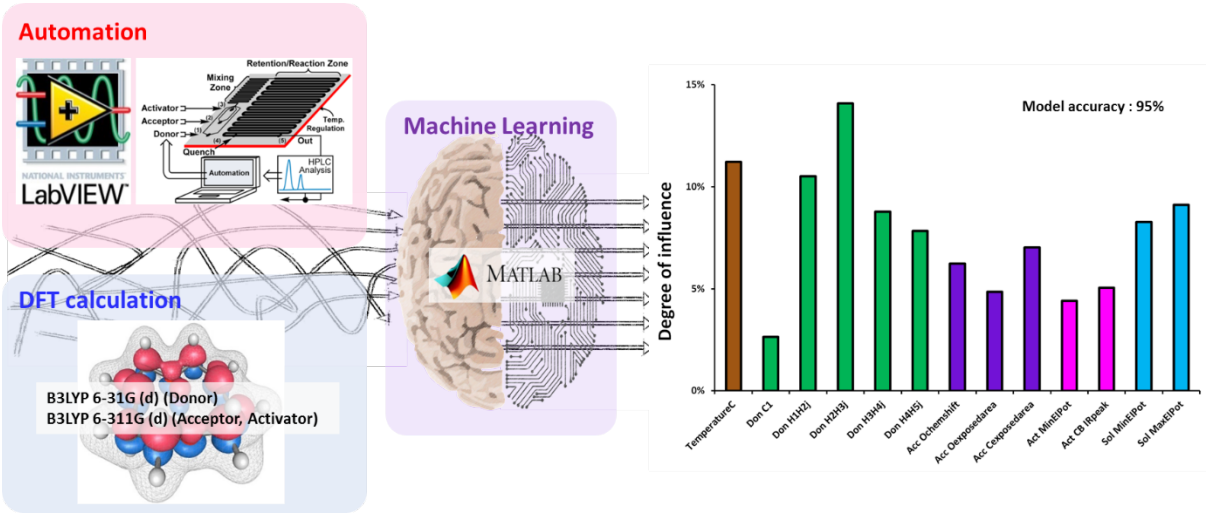


Figure 1. In-silico understanding and prediction of chemical glycosylation through Machine Learning.

With an empirical understanding of glycosylation mechanism, descriptors to quantify the molecular property of donor, acceptor, activator and solvent were calculated with density functional theory (DFT) and cheminformatics calculations. For a successful numeric description of glycosylation reaction, the quantification of the reactivity of the coupling partners along with steric effects and nucleophilicity becomes utmost important.

It was revealed that reactivity of donor could be numerically quantified by using the 13C NMR chemical shift at the anomeric position [2]. Also, the Hn-Hn+1 (n = 1 to 4) j coupling constant were considered for quantifying donor's stereochemical properties. Similarly, for acceptor, nucleophilicity is characterized by 18O NMR chemical shift and for quantification of steric effects, O and αC exposed surface area were calculated. IR absorption peak of O-H bond and minimum value of the electrostatic potential (MinElPot) of the conjugate base were chosen as activator descriptors, and for solvent, MinElPot and maximum value of the electrostatic potential (MaxElPot) were chosen as input parameters.

As stereo-selective outcome of glycosylation is a continuous numerical parameter, machine-learning regression in the form of Random Forest algorithm was chosen instead of classification models. This regression methodology, being inherently quantitative in nature, gave the advantage of not only prediction of stereo- selectivity but also the required reaction conditions for obtaining the same.

References

[1] Chatterjee, S.; Moon, S.; Hentschel, F.; Gilmore, K.; Seeberger, P. H. *J. Am. Chem. Soc.* **2018**, 140, 11942–11953.

[2] Zhang, Z.; Ollmann, I. R.; Ye, X. S.; Wischnat, R.; Baasov, T.; Wong, C.H. *J. Am. Chem.* Soc. **1999**, 121, 734–753.

Back to Program

Tuesday

Poster #12

GlyConnect Compozitor – an Interactive Graph of Glycan Compositions

Thibault Robin[1,2,3,4], Julien Mariethoz[2,3], Frédérique Lisacek[2,3,5]

[1]SIB Swiss Institute of Bioinformatics, CALIPHO Group, Geneva, Switzerland
[2]SIB Swiss Institute of Bioinformatics, Proteome Informatics Group Geneva, Switzerland
[3]University of Geneva, Computer Science Department, Geneva, Switzerland
[4]University of Geneva, Department of Microbiology and Molecular Medicine, Geneva, Switzerland
[5]University of Geneva, Section of Biology, Geneva, Switzerland

GlyConnect Compozitor is the latest visualization tool added to the GlyConnect environment. The GlyConnect database contains a wealth of information about protein glycosylation, reporting numerous experimentally detected glycosylation sites. For each glycosylation site, the reported glycan compositions are provided along with the corresponding glycan structures that were resolved. With the recent rise of mass spectrometry in proteomics-based approaches, glycan compositions are increasingly identified at specific sites without a full resolution of the associated structure. Consequently, there is a growing need to map the diversity of glycan compositions.

GlyConnect Compozitor aims to address this need by providing an interactive graph of the observed glycan composition for a given protein, tissue or cell line. Note that the GlyConnect tissue and cell line controlled vocabularies have been substantially revised to provide the user with an accurate selection of terms. The different compositions are linked together by their differences in terms of monosaccharides, leading to a graph describing compositional dependency. Optionally, a single glycosylation site can be selected to display the relative compositions and easily determine whether they are site-specific. A mode supporting the comparison between two different criteria is also under development, allowing for instance to compare the composition graph between two proteins or glycosylation sites.

GlyConnect Compozitor is designed as a single page application using the D3 JavaScript library. A prototype version is openly available on https://dev.glyconnect.org/compozitor.

Back to Program

Wednesday

09.00

Standardization of Glycosaminoglycan (GAG) Sequences Binding to Proteins and Creation of a Pipeline for the Curation of GAG Protein Interactions: Application to GAG Interaction Networks

Sylvie Ricard-Blum

University Lyon 1
Institute of Molecular and Supramolecular Chemistry and Biochemistry
UMR 5246 CNRS
Villeurbanne, France

Mammalian glycosaminoglycans (GAGs) are complex polysaccharides comprising heparan sulfate, heparin, dermatan sulfate, chondroitin sulfate, keratan sulfate and hyaluronan. GAG-protein interactions reported in the literature are curated by MatrixDB database (http://matrixdb.univ-lyon1.fr/), which belongs to the IMEX consortium (http://www.imexconsortium.org/) and follows its curation rules. However, a standard nomenclature and a machine-readable format of GAGs for curation together with bioinformatics tools for mining their interaction data are lacking. We have built an automated pipeline to (i) standardize the format of GAG sequences interacting with proteins manually curated from the literature, (ii) translate them into the machine-readable GlycoCT format and SNFG (Symbol Nomenclature For Glycans) images and (iii) convert them into a format processed by a builder generating 3D models based on a repertoire of conformations validated by data from GAG-protein co-crystals.

We have curated GAG sequences binding to proteins, translated them into the GlycoCT and SNFG formats using the pipeline and cross-referenced the GAG entries of MatrixDB with the Chemical Entities of Biological Interest (ChEBI, https://www.ebi.ac.uk/chebi/) and the glycan repository GlyTouCan (https://glytoucan.org/). We have also cross-referenced the GAG entries of GlyTouCan with MatrixDB, and ChEBI entries with GlyTouCan to increase the interoperability of the major databases for GAGs.

Furthermore, we have developed and integrated into the pipeline a converter (CT23D), which automatically translates the GlycoCT code of a GAG sequence into the input file required to construct a 3D model (https://github.com/OlivierClerc/convert-glycoct-inp). The 3D models, used to display the GAG binding sites, can be built on MatrixDB website using a version of the POLYS glycan builder we have developed for GAGs (http://glycan-builder.cermav.cnrs.fr/gag/), downloaded as pdb files and visualized. The ultimate goal is to integrate biochemical and structural features of GAGs contributing to protein binding into the interaction networks they establish either as free chains or as chains attached to proteins to form proteoglycans.

Back to Program

Wednesday
09.35

Glycomics and Proteomics of Brain Pathologies

Joe Zaia

Boston University
Department of Biochemistry
Boston, MA, United States of America

Dysregulation of the cellular microenvironment occurs in cancers, neurodevelopmental and neuropsychiatric diseases. Known as the matrisome, the set of extracellular matrix and cell surface molecules control the availability of growth factors to cellular receptors and the mechanical-physical properties of the microenvironment. In brain, region-specific regulation of matrisome molecule glycosylation controls the neuronal microenvironment and becomes dysregulated in neuropsychiatric diseases. In neurodegeneration, proteoglycans bind to and play roles in the aggregation of proteins including Aβ, tau, prion protein, and α-synuclein. In bipolar disorder and schizophrenia patient brains, alterations in chondroitin sulfate (CS) sulfation patterns are observed.

We developed a workflow for quantitative mapping of alterations in matrisome structure at the glycomics and proteomics levels from brain tissue slides. Our highly successful workflow for profiling glycosaminoglycans (GAGs) including heparan sulfate (HS) and chondroitin sulfate (CS), *N*-glycans, and proteins provides a readout of GAG quantities, domain structures, and non-reducing end structures using simple enzyme digestions with minimal need for workup. The final proteomics of tryptic peptides identifies ~1200 proteins from the 10 nL tissue volume, providing deeper coverage than can be obtained from an MS imaging approach. I will show results from our studies of brain aging and pathologies including drug abuse, Schizophrenia, Parkinson's disease and glioma.

Back to Program

Wednesday

10.40

Contextualized Functions of Glycans in Human Tissue Formation

Hans H. Wandall

University of Copenhagen
Department of Cellular and Molecular Medicine, Center for Glycomics
Copenhagen, Denmark

More than 200 different types of post-translational modifications (PTMs) finetune the structure and function of proteins in human tissue. The most diverse and abundant subtype of PTM is believed to be glycosylation. Glycans exhibit a large structural diversity with cell-type specificities that underlie defined biological functions. However, our functional understanding of the glycome is limited, partially due to the lack of simple model systems that allow for open-ended, unbiased screens of glycan function in human tissues. We here present the first human organotypic platform to systematically interrogate glycan functions in tissue formation. Using CRISPR-Cas9 and a 3D organotypic model of human skin, we have generated a human tissue library with truncation of the key glycan structures, thus providing a platform contextualized to a human setting with broad discovery potential. The library demonstrates distinct phenotypes associated with loss of individual glycosylation pathways, including the effect of complex *N*-glycans on wound healing, *O*-GalNAc glycans on cell-cell adhesion, differential roles for *O*-Fucose and *O*-Glucose glycosylation in NOTCH signaling and glycosphingolipids in EGF signaling and skin barrier formation. The platform can help define the roles of the glycome in epithelial homeostasis, epithelial transformation, cell-cell and cell-matrix adhesion, signaling and host pathogen interactions, enabling glycobiology to move beyond the constrains of 2D cell culture assays, printed glycan arrays, and animal models.

Back to Program

Wednesday

11.15

Challenges in Defining the Structure and Function of Invertebrate *N*-Glycans

Iain B.H. Wilson, S. Yan, J. Vanabeselaere, A. Hykollari, B. Eckmair and K. Paschinger

Universität für Bodenkultur
Department für Chemie
Wien, Austria

Despite years of research, the glycomes of invertebrates continue to surprise. Gone are the days when it could be said that “simple” organisms have “simple” glycomes. Typically, any biological sample will yield a complex mixture of 100 *N*-glycans or more, whereby we do reach a limit in terms of detection of low-abundance structures, whose analyses do not reach certain minimum standards, e.g., in terms of MS/MS. Another aspect is the sheer amount of data being generated - for previously unknown glycomes there are no databases and probably no software is currently able to accurately predict structures; thus, manual interpretation is absolutely necessary in order to guide the experimental steps for structural verification.

Then, there is the whole question of function - what do these various glycans do? As between species there are variations in structures or abundance of their glycoconjugates, we can hypothesise that speciation correlates with glycomic alterations. Species which have evolved special ecological niches, whether they are free-living in terrestrial or aquatic environments or they are parasites which infect hosts and are transmitted through other invertebrates. Thus, the final glycomic result is the product of evolution and of selective pressure that certain protein-carbohydrate recognition events either take place or are prevented. The pressure in science to include functional aspects means that structural glycomic analysis must be accompanied by data of biological relevance, e.g., recognition by proteins of the immune system in a glycan array format. This is the direction in which we are going as shown by our recent data on the *N*-glycomes of honeybee royal jelly and canine heartworm.

Another perspective is also opened up by manipulating glycosylation pathways by examining mutant glycomes, as we have done in *Caenorhabditis* - which reveals that glycan biosynthesis is carefully balanced and removing enzymes from the Golgi can have a major impact on the glycome, while not being incompatible with viability in a laboratory setting.

Acknowledgements

This work was supported by the Austrian Science Fund (FWF; grants P23922, P25058, P26662 and W1224) and the European Union (PITN-GA-2013-608295).

Back to Program

Wednesday

Software
#6

Bioinformatics Tools for Glycoscience

René Ranzinger

University of Georgia
Complex Carbohydrate Research Centre
Athens, GA, United States of America

To advance our understanding of the roles that glycans play in development and disease bioinformatics tools have become more and more important. As throughput of analytical data production increase software tools assisting in the interpretation of the data are essential to keep up with the amount of data produced. In addition, as more knowledge about glycans, proteins and der interactions are discovered, databases are crucial for storing and accessing this information.

GRITS Toolbox

GRITS Toolbox is a software developed for the interpretation of glycomics mass spectrometric (MS) data. The program allows for (semi-)automated annotation of MS with glycan structures from database. The software supports multiple different MS methods (MS profile, MS/MS, TIM, LC-MS/MS) and glycan preparation methods (native and per-methylated glycans, labels, etc.). The program is freely available at http://www.grits-toolbox.org.

GRITS Toolbox DANGO

DANGO is and extension of our GRITS Toolbox software that allows the (semi-)automated annotation of intact glycosphingolipid MS data. Similar to GRITS Toolbox the software supports different MS methods and sample preparation methods. DANGO and sample preparation protocols for intact glycolipid MS analysis are freely available on our webpage (http://www.ms-dango.org).

GRITS Toolbox GlycanBuilder

The GlycanBuilder extension integrates a carbohydrate structure builder into the GRITS Toolbox software. The drawing tool is based on the widely used GlycoWorkbench application but supports the newest notation of the SNFG cartoon representation for glycan structures as

well as additional monosaccharides. The extension can also be used to create new glycan databases for GRITS Toolbox and DANGO to be used for the interpretation of MS data.

GlyGen

The GlyGen project is a data integration project aiming to integrate data from the glycomics and glyco-proteomics domain into a single data resource. This integrated and standardized resource allows to perform queries across multiple domains and answering questions that otherwise cannot be answered or only be answered with considerable effort. The data and a portal to query the information is freely available at https://www.glygen.org.

Back to Program

Wednesday

Software
#7

‘Plug and Play’ with Machine Learning

Sourav Chatterjee[1], Sooyeon Moon[1,2], Peter H. Seeberger[1,2], Kerry Gilmore[1]

[1]Max Planck Institute of Colloids and Interfaces, Department of Biomolecular Systems, Potsdam, Germany

[2]Freie Universität Berlin, Institute of Chemistry and Biochemistry, Berlin, Germany

Application of machine learning in scientific research have become ubiquitous at present, due to availability of ‘big data’, either from literature or from automated laboratories. Machine learning algorithms can be applied in various facets of research such as prediction of crystals form failed reactions [1] or prediction of C-N cross-coupling [2]. Predicting selectivity and yields of chemical reactions require regression based machine learning algorithms instead of classification. However, this leads to the daunting challenge of numerically characterizing the model chemical reaction by means of chemical descriptors. Here we propose a software capable of automated chemical descriptor search with machine learning. Instead of hand picking mechanistically plausible descriptors having inherent biases, we chose descriptors, numerically quantifying important properties of individual molecules such as donors, acceptors, activators and solvents taking part in a model chemical reaction such as glycosylation.

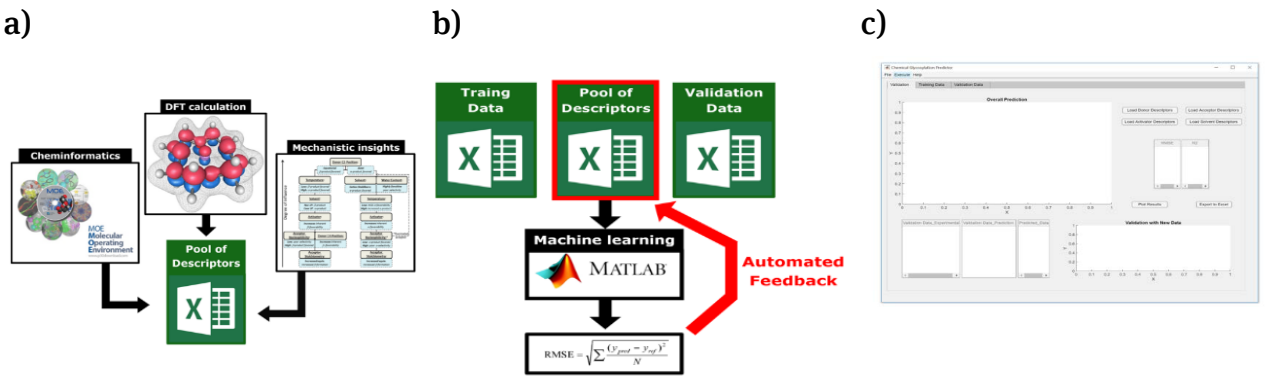


Figure 1. (a) Generation of descriptor pool, (b) automated descriptor search, (c) GUI of chemical glycosylation predictor software.

A number of numeric descriptors including Quantum mechanical (QM), Qualitative structure activity relationship (QSAR) and thermodynamic were calculated using DFT and Cheminformatics approach, for individual reactants, to numerically represent electronic, steric and thermodynamic properties.

These descriptors formed the descriptor pool (Figure 1a). Software was developed in MATLAB to feed these descriptors, using a user-friendly GUI (Figure 1c). The software performed a sequential descriptor search using Random Forest algorithm. Each descriptor was evaluated based on its individual and collective performance, quantified by root mean square error (RMSE) and R2, in the overall training set and validation benchmarking data (Figure 1b). We have around 300 data points in our model chemical glycosylation, which forms the training set. Hence, we restricted ourselves for a maximum of 4 descriptors per donor, acceptor, activator and solvent generating 16 descriptors in total to avoid any overfitting to get statistically meaningful results.

This software gives access to innovative machine learning algorithms without the need to have any programming and mathematical skills, and enables the 'plug and play' approach for applying machine learning to chemistry data. The software can be used to perform automated descriptor search, deploying different machine learning algorithms such as support vector machines, Gaussian process regression and Random forest. The software also allows easy visualization of data, model accuracies and data exporting in Excel for further processing.

References

[1] Raccuglia, P.; Elbert, K. C.; Adler, P. D.; Falk, C.; Wenny, M. B.; Mollo, A.; Zeller, M.; Friedler, S. A.; Schrier, J.; Norquist, A. J. *Nature*, **2016**, 533, 73–76.

[2] Ahneman, D. T.; Estrada, J. G.; Lin, S.; Dreher, S. D.; Doyle, A. G. *Science* **2018**, 360, 186–190.

Back to Program

Wednesday

Software
#8

GlyConnect, the Power Engine of Glycomics@ExPASy

Julien Mariethoz, Thibault Robin, Frédérique Lisacek

SIB Swiss Institute of Bioinformatics, Proteome Informatics Group, Geneva, Switzerland

Glycomics@ExPASy the glycomics tab of the Swiss Institute of Bioinformatics server (www.expasy.org/glycomics) was created in 2016 to centralise web-based glycoinformatics resources developed within an international network of glycoscientists. The philosophy of this toolbox is to be {glycoscientist AND protein scientist}–friendly with the aim of popularising (a) the use of bioinformatics in glycobiology and (b) the relation between glycobiology and protein-oriented bioinformatics resources [1].

Glyconnect was designed to (i) centralise curated data collected from selected items listed in the Glycomics@ExPASy database section as well as supplementary material of large-scale glycoproteomics experiments and (ii) integrate several tools listed in the Glycomics@ExPASy tool section [2]. GlyConnect also cross-references reciprocally with several external databases such as GlyTouCan, UniProt, PDB, GeneCards and neXtprot that hardly need introduction in the audience of this meeting. Recently, we added the GlyConnect Compozitor (see related poster) to manage the growing set of glycan compositions accumulated in glycoproteomics data. We also revised the tissue controlled vocabulary inherited from GlycoSuiteDB (core data of GlyConnect) and moved to Gene Ontology (subcellular location) and the UBERON ontology of anatomy for greater precision and flexibility.

In this demo, we will show the impact of increasing the database by 35% in glycoproteins and refining site-specific information, on query results. We will highlight the advantages of improved connectivity within and across GlyConnect as exemplified among others by communication with our UniLectin sister project.

References

[1] Mariethoz J, Alocci D, Gastaldello A, Horlacher O, Gasteiger E, Rojas-Macias MA, Karlsson NG, Packer NH, Lisacek F (2018) Glycomics@ExPASy: Bridging the gap, *Mol Cell Proteomics* 17: 2164–2176

[2] Alocci D, Mariethoz J, Gastaldello A, Gasteiger E, Karlsson NG, Kolarich D, Packer NH, Lisacek F (2019) GlyConnect: glycoproteomics goes visual, interactive and analytical, *J. Prot. Research* 18(2):664-677

Back to Program

Wednesday

Software #9

GlyCosmos Portal and GLIC

Kiyoko Aoki-Kinoshita

Soka University, Faculty of Science and Engineering, Tokyo, Japan

After acknowledgement from the Japanese Society for Carbohydrate Research as their official portal, the GlyCosmos Portal (https://glycosmos.org) was released on April 1, 2019. This portal aims to serve as a gateway for glycan-related omics data. It serves as a culmination of the large amounts of glycan-related database development that has taken place in Japan for over 12 years; from GlyCosmos, users can browse through glycan related genes, proteins, diseases, pathways and glycomics data, all through a user-friendly Web interface.

This demo will invite users to go through the portal and provide feedback to the developers.

Much glycan analysis software has also been developed over the past decade, and the glycoinformatics consortium (GLIC; https://glic.glycoinfo.org) has been attempting to organize and coordinate among the developers. The GLIC website also provides a form by which glycobiologists can contact developers for assistance in analytical problems. Thus, the GLIC site will also be presented for those interested in participating in GLIC and/or in finding GLIC members for collaborations.

Back to Program

Wednesday

14.15

Sugars in the Gas Phase – from Structure to Reaction Mechanism

E. Mucha[1], M. Marianski[1], M. Lettow[1], G. von Helden[1], G. Meijer[1], Peter H. Seeberger[2,3], Kevin Pagel[1,3]

[1]Fritz Haber Institute of the Max Planck Society, Berlin, Germany

[2]Max Planck Institute for Colloids and Interfaces, Potsdam, Germany

[3]Freie Universität Berlin, Institute of Chemistry and Biochemistry, Berlin, Germany

Oligosaccharides or glycans are essential in nature and central participants in virtually every biological process. The extensive structural diversity enables glycans to encode rich information in biological functions; however, it also creates major challenges in almost all aspects of the glycosciences. The structural analysis of glycans is often challenging due to the coexistence of multiple isomers and is further complicated by poorly understood rearrangement reactions during mass spectrometric analysis. Also the synthetic formation of glycosidic bonds during glycan assembly is mechanistically still not fully understood. This is largely a result of highly-reactive, but short-lived oxocarbenium ion intermediates, which are difficult to study using established techniques. However, the structure of these intermediates dictates the stereochemistry of the resulting glycosidic bond, the control of which is absolutely crucial for a successful synthesis.

Here we show, how ion mobility-mass spectrometry and gas-phase IR spectroscopy can be used to unravel the structure, fragmentation pathways and reaction mechanisms of glycans. First, a series of synthetically derived standards were analyzed. A separation based on size and shape by ion mobility spectrometry enabled a distinction for most isomers, although in some cases the resolution was not sufficient for an unambiguous identification [1]. Cold-ion gas-phase IR spectroscopy, on the other hand, yielded highly diagnostic absorption patterns with a variety of well-resolved bands, which enables the identification of any conceivable isomer [2]. This remarkable resolution was further used to study the detailed mechanism of fucose migration – an erratic rearrangement reaction that often leads to erroneous assignments in the tandem MS analysis of glycans [3].

Our results showed that fucose migration already occurs at low energy conditions and does not require dissociation, which has fundamental consequences for the interpretation of glycan MS data.

Finally, we were recently able to determine the first high-resolution structure of short-lived oxocarbenium ion intermediates occurring during glycan synthesis [4]. Comparison between experimental and theoretical cold ion IR spectra revealed detailed structural insights and showed that the ring conformation is different for individual monosaccharides, which in turn determines the stereochemical outcome of the glycosylation reaction.

References

[1] J. Hofmann et al. *Nature* **2015**, 526, 241.

[2] E. Mucha et al. *Angew. Chem. Int. Ed.* **2017**, 56, 11248.

[3] E. Mucha et al., *Angew. Chem. Int. Ed.* **2018**, 57, 7440.

[4] E. Mucha and M. Marianski et al., *Nature Commun.* **2018**, 9, 4147.

Back to Program

Wednesday

14.50

Glycomimetic Tools to Understand Decoding of the Glycan Code

Dongyoon Kim[1], Felix Fuchsberger[1,2], Christoph Rademacher[1,2]

[1]Max Planck Institute of Colloids and Interfaces
Department of Biomolecular Systems
Potsdam, Germany

[2]Freie Universität Berlin, Institute of Chemistry and Biochemistry,
Berlin, Germany

Every living cell is decorated with a dense fur of glycans. Multicellular organisms make use of this matrix to encode for specific information such as cellular identity, metabolic and activation status, and circadian clock. Cell surface glycans, present on glycoproteins and glycolipids modulate protein localization, their interaction partners and receptor life cycles. Since glycans are secondary gene products, their formation as well as their breakdown is determined by many factors, not limited to gene expression of potentially over 200 enzymes and transporters. Regulatory circuit arose on many different time scales: very quick alteration of the glycocalyx can be introduced using secreted hydrolases and endo- and exocytosis of glycoproteins. Long-term remodeling of glycans can result from gene expression and histone modifications.

How can such a specially and temporally regulated, complex and stochastic system encode reliable for information that can successfully be decoded by other cells through the use of lectin receptors? To address this intriguing question, we have developed glycomimetic tools, small molecules that specifically bind to selected mammalian lectins [1-5].

These tool compounds are used in combination with nano- and microparticles to ask very fundamental questions such as: How often does a protein-carbohydrate interaction on a cell surface lead to a productive, biological response? How do lectins and glycans find each other on two-dimensional surfaces? How much information can be transferred through such a communication channel and how does this compare to glycan-independent pathways of cell-cell communication?

References

[1] Aretz, J., Wamhoff, E.C., Hanske, J., Heymann, D., and Rademacher, C. (2014). Computational and experimental prediction of human C-type lectin receptor druggability. *Front Immunol.* 5, 323.

[2] Aretz, J., Baukmann, H., Shanina, E., Hanske, J., Wawrzinek, R., Zapol'skii, V.A., Seeberger, P.H., Kaufmann, D.E., and Rademacher, C. (2017). Identification of Multiple Druggable Secondary Sites by Fragment Screening against DC-SIGN. *Angew. Chem. Int. Ed. Engl.* 56, 7292-7296.

[3] Hanske, J., Aleksic, S., Ballaschk, M., Jurk, M., Shanina, E., Beerbaum, M., Schmieder, P., Keller, B.G., and Rademacher, C. (2016). Intradomain Allosteric Network Modulates Calcium Affinity of the C-Type Lectin Receptor Langerin. *J. Am .Chem .Soc.* 138, 12176-12186.

[4] Aretz, J., Anumala, U.R., Fuchsberger, F.F., Molavi, N., Ziebart, N., Zhang, H., Nazare, M., and Rademacher, C. (2018). Allosteric Inhibition of a Mammalian Lectin. *J Am Chem .Soc.* 140, 14915-14925.

[5] Wamhoff, E.-C., Schulze, J., Bellmann, L., Bachem, G., Fuchsberger, F.F., Rademacher, J., Hermann, M., Del Frari, B., van Dalen, R., Hartmann, D., van Sorge, N., Seitz, O., Stoitzner, P., Rademacher, C. (2018). A specific, glycomimetic Langerin ligand for human Langerhans cell targeting. *bioRxiv*. doi:10.1101/286021

Back to Program

Wednesday

15.25

Dimensions of Glycomimetics

Thisbe K. Lindhorst

Christiana Albertina University of Kiel
Otto Diels Institute of Organic Chemistry
Kiel, Germany

Through about 20 years, my group have designed, synthesized and used various functional glycomimetics to explore carbohydrate-protein interactions.

I will explain the dimensions of our glycomimetic design, their structure and special features and will ask the question if these synthetic non natural glycoconjugates are amenable to a systematic glycobioinformatic analysis.

See also our most recent paper: *Org. Biomol. Chem.*, **2019**, Advance Article, https://doi.org/10.1039/C9OB00124G

Back to Program

Thursday

09.00

Glycans as Biomarkers and Functional Effectors in Diabetes and Cardiovascular Diseases

Gordan Lauc[1,2]

[1]University of Zagreb, Faculty of Pharmacy and Biochemistry, Zagreb, Croatia

[2]Genos Glycoscience Research Laboratory, Zagreb, Croatia

The majority of proteins that evolved the after appearance of multicellular life are glycosylated and glycans significantly affect structure and function of these proteins. However, due to structural complexity of glycans and the absence of a direct genetic template, the analysis of protein glycosylation is much more complicated than the analysis of DNA or proteins. Consequently, the knowledge about the importance of individual variation in glycans for both normal physiological processes and diseases is still limited.

In the last few years it is becoming increasingly clear that variations in a DNA sequence are only a beginning of the understanding of complex human diseases. Genetic polymorphisms have to be put in the context of complex biology of life and a more elaborate approach that combines different 'omics phenotypes is needed to understand disease mechanisms and perform patient stratification that transcends genomics. Glycomics, as by far the most complex posttranslational modification, has an immense potential in this respect, which is only beginning to be investigated.

By generating glycomic data for over 80,000 individuals from some of the best characterized clinical and epidemiological cohorts we enabled glycomics to meet other 'omics. The analysis of this rich gold mind is painting a picture of a very complex genetic and epigenetic regulation of glycosylation that fine tunes protein activity in multiple biological systems and, if altered, contributes to development of different complex diseases. In particular, the evidence is accumulating that in cardiometabolic diseases changes in glycosylation are not only biomarkers, but functional effectors that actively participate in disease development.

Back to Program

Thursday

09.35

Update on the 1st Human Glycoproteomics Initiative (HUPO/HGI) Study: Interlaboratory Evaluation of Software for Intact Glycopeptide Analysis by MS

Morten Thaysen-Andersen[1], Daniel Kolarich[2], Rebeca Kawahara Sakuma[1], Hannes Hinneburg[1], Kai-Hooi Khoo[3], Katalin Medzihradszky[4], Joseph Zaia[5], Goran Larson[6], Stuart Haslam[7], Giuseppe Palmisano[8], Jong Shin Yoo[9], and Nicolle H. Packer[1,2]

[1]Department of Molecular Sciences, Macquarie University, Sydney, NSW, Australia

[2]Institute for Glycomics, Griffith University, Southport, QLD, Australia

[3]Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

[4]Mass Spectrometry Facility, The University of California, San Francisco, CA, United States of America

[5]School of Medicine, Boston University, Boston, MA, United States of America

[6]Department of Clinical Chemistry, Gothenburg University, Gothenburg, Sweden

[7]Department of Life Sciences, Imperial College, London, United Kingdom

[8]Department of Parasitology, University of Sao Paulo, Sao Paulo, Brazil

[9]Korea Basic Science Institute and Graduate School of Analytical Science and Technology, Chungnam National University, Daejon, Korea

Analytical advances in mass spectrometry have now facilitated LC-MS/MS-based glycoproteomics studies that are identifying hundreds and even thousands of unique intact glycopeptides from a single experiment. However, significant bottlenecks still exist in the accurate annotation of the large volumes of MS/MS spectral data and in the confident identification of the corresponding intact glycopeptides. Efficient software for automated glycopeptide identification is essential and glycoproteomics has recently seen the development, by both commercial and academic developers, of various software tools for

automated or semi-automated annotation and identification of intact glycopeptides from MS/MS spectral data.

This is an update on the results from the interlaboratory study that was initiated in September 2017 by the Human Glycoproteomics Initiative (HGI) that is part of the Human Proteome Organisation (HUPO). This study evaluates the performance of the currently available glycoproteomics informatics capabilities for intact glycopeptide identification in complex human samples. The 22 study participants include both expert users (13) and software developers (9) across both industry and research in the glycoproteomics community.

All participants were supplied the same two high resolution LC-MS/MS datasets of intact *N*- and *O*-glycopeptides acquired from analysis of digested human serum glycoproteins using complementary fragmentation modes (HCD, EThcD, ETciD and CID) as obtained by collaboration with Thermo Scientific. The participants provided reports of identified intact *N*- and *O*-glycopeptides, annotated spectra and details of the informatics approaches they used.

Preliminary study of the results from results provide interesting insights into the diversity of approaches being applied to glycopeptide analysis of complex samples, the group-to-group (vari)ability to accurately identify intact glycopeptides and the performance of the available glycoproteomics software. This documenting of the current state of glycoproteomics analytical software is vital to improve further informatics development and to stimulate the application of system-wide glycopeptide analysis by more researchers.

Back to Program

Thursday
10.10

Glycoscience in the Algae

Robert A. Field

John Innes Centre
Department of Biological Chemistry
Norwhich, United Kingdom

The algae are an enormous, diverse collection of organisms, ranging from single celled, micron diameter microorganisms to seaweeds that are many metres in length. While there is a general appreciation that macroalgae are sources of polysaccharides - alginates and fucoidans, for instance - the glycobiology of the algae in general remains largely unexplored.

This presentation will highlight aspects of the glycobiology of the freshwater microalgae *Euglena gracilis*, which have been enabled by transcriptomics and follow up biochemical studies. Further studies on the glycoscience of the harmful-algal-bloom causing *Prymnesium parvum* and its' lytic virus will also be presented.

References

[1] NDP-β-L-rhamnose biosynthesis across the algal taxonomic groups: an evolutionary perspective. *J. Biol. Chem.*, **2019**, under revision.

[2] Unravelling the subtleties of the specificity of β-1,3-glucan-utilising phosphorylases: GH94 vs GH 149 vs GH161. *J. Biol. Chem.*, **2019**, accepted doi: 10.1074/jbc.RA119.007712 .

[3] Identification of de novo KDN biosynthesis in the toxic haptophyte, *Prymnesium parvum*, and implications for widespread sialic acid biosynthesis amongst microalgae. *J. Biol. Chem.*, **2018**, 293, 16277-16290.

[4] CuAAC click chemistry for the enhanced detection of novel alkyne-based natural product toxins. *Chem. Commun.*, **2018**, 54, 1223-1227.

[5] Insights into toxic *Prymnesium parvum* blooms: the role of sugars and algal viruses. *Biochem. Soc. Trans.*, **2018**, 46, 413-421.

[6] Identification of *Euglena gracilis* β-1,3-glucan phosphorylase and establishment of a new glycosyl hydrolase family GH 149. *J. Biol. Chem.*, **2018**, 293, 2865-2876.

[7] Exploring the glycans of *Euglena gracilis*. *Biology*, **2017**, 6, 45 doi: 10.3390/biology6040045.

[8] Sugar nucleotide profiling, *Methods in Enzymology*, **2017**, 597, 209-239.

[9] Isolation and characterization of a double stranded DNA megavirus infecting the toxin-producing haptophyte, *Prymnesium parvum*. *Viruses*, **2017**, 9, article 40 doi: 10.3390/v9030040.

[10] Fluorescent mannosides serve as acceptor substrates for glycosyltransferase and sugar-1-phosphate transferase activities in Euglena gracilis membranes. *Carbohydr. Res.*, **2017**, 438, 26-38.

[11] Gene Discovery for Synthetic Biology: Exploring the Novel Natural Product Biosynthetic Capacity of Eukaryotic Microalgae. *Methods in Enzymology*, **2016**, 576, 99-120.

[12] Euglena in time: evolution, control of central metabolic processes and multi-domain proteins in carbohydrate and natural product biochemistry. *Perspectives in Science*, **2015**, 6, 84-93.

[13] The transcriptome of Euglena gracilis reveals unexpected metabolic capabilities for carbohydrate and natural product biochemistry. *Molecular Biosystems*, **2015**, 11, 2808-2820.

Back to Program

Thursday

11.15

Carbohydrates and Glycomimetics from the Viewpoint of a Computational Chemist

Martin Frank

BIOGNOS AB
Gothenborg
Sweden

Advances in the functional understanding of carbohydrate–protein interactions have enabled the development of a new class of small-molecule drugs, known as glycomimetics [1]. The computational design of glycomimetic inhibitors or carbohydrate-based therapeutics requires the application of methods from chemoinformatics, bioinformatics and molecular docking [2]. For decades, the rigid 'lock-and-key'-principle had been the foundation for structure-based drug design. However, it has recently become evident that, when carbohydrates meet their macromolecular receptor, a successful binding event involves frequently significant conformational dynamics. With the advance of computer technology it has become feasible to simulate such motions - even for large molecular systems - routinely on the microsecond timescale. This helps to interpret existing experimental results better and can give valuable new ideas for planning the next round of experiments.

Detection and use of transient dynamic binding pockets has led to the successful design of high-affinity glycomimetic inhibitors [3] and MD simulations have also been instrumental for understanding the dynamics of the IgG1 Fc-domain which will help to design improved antibody therapeutics [4]. Here I will discuss the current challenges, prospects and limitations of computational carbohydrate-based inhibitor design and present in-house workflows that will help to speed-up the development of glycotherapeutics.

Additionally a novel encoding format for carbohydrate building blocks will be presented that will standardize the storage of glycomimetic molecules in databases and will facilitate the mining for carbohydrate-like scaffolds in commercial compound libraries.

References

[1] Ernst B, Magnani JL (2009) From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov* 8:661–677. doi: 10.1038/nrd2852

[2] Frank M (2014) Computational Docking as a Tool for the Rational Design of Carbohydrate-Based Drugs. In: Rademacher C, Seeberger PH (eds) *Topics in Medicinal Chemistry: Carbohydrates as Drugs*. Springer,

[3] Prescher, H., Frank, M., Gütgemann, S., Kuhfeldt, E., Schweizer, A., Nitschke, L., Watzl, C. & Brossmer, R. Design, Synthesis, and Biological Evaluation of Small, High-Affinity Siglec-7 Ligands: Toward Novel Inhibitors of Cancer Immune Evasion. *J. Med. Chem.* 60, 941–956 (2017).

[4] Frank, M., Walker, R. C., Lanzilotta, W. N., Prestegard, J. H. & Barb, A. W. Immunoglobulin G1 Fc Domain Motions: Implications for Fc Engineering. *J. Mol. Biol.* 426, 1799–1811 (2014).

Back to Program

Thursday

11.50

From Symbolic Carbohydrate Notations to Atomic Coordinates

I. Yu. Chernyshov, Philip V. Toukach

Russian Academiy of Sciences
Zelinsky Institute of Organic Chemistry
Moscow, Russia

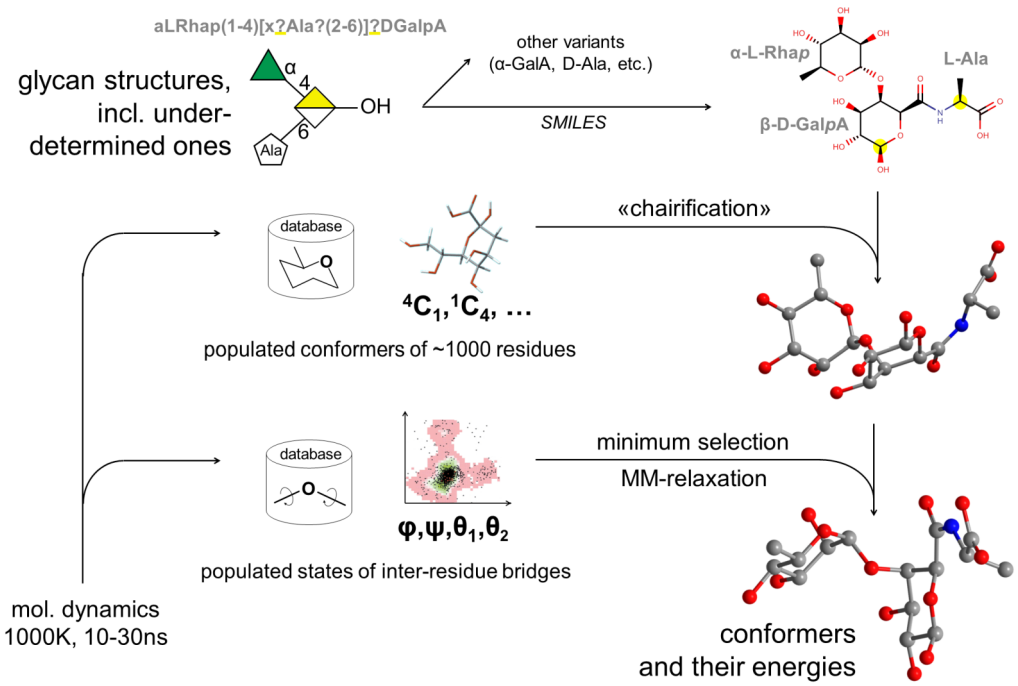
Molecular modeling is an inevitable step in rational design of pharmaceuticals with desired activity and in docking of candidate structures to biomacromolecules. It always starts with a formal description of a primary structure. In glycomics, the structure of biological glycans and their derivatives is traditionally published in symbolic notations dealing with monosaccharides and other residues rather than with atoms. However, none of the existing carbohydrate notations is supported by molecular modeling software, and these notations have to be converted to cheminformatic formats prior to modeling. Many published primary structures of bioglycans have underdetermined moieties. A few solutions allow conversion of selected symbolic notations to PDB coordinates but they can process only fully defined structures comprised of a limited set of residues.

To solve these problems and to eliminate the restrictions of glycan modeling, we developed a tool for translation of a symbolic carbohydrate notation (CSDB Linear [1]) to a universal chemical notation (SMILES) and for generation of optimized atomic coordinates. Our approach is called REStLESS, which stands for REsidues as SMILES, LinkagEs as SMARTS [2]. It is implemented at the platform of Carbohydrate Structure Database (CSDB [3]) and utilizes its data on ~20,000 natural carbohydrates and on ~1,000 monosaccharide residues and other building blocks of bioglycans. REStLESS is able to generate SMILES and atomic coordinates for glycoconjugates containing residues undefined in CSDB, if they are specified in SMILES format.

The underlying algorithm generates SMILES codes of bioglycans by retrieving SMILES of prototype residues from a service database, adjusting them to actual anomeric, absolute and ring-size configurations, and combining them by virtual condensation reactions in SMARTS format. This approach supports underdetermined structures typical for glycomics.

For undefined residues and aglycons, SMILES codes must be specified in the symbolic notation. At the second stage, the obtained SMILES codes are expanded to all possible stereomers and converted to atomic coordinates. The coordinates are optimized using predominant conformers of residues and conformation maps of inter-residue bridges pre-calculated by high-temperature molecular dynamics. The resulting set of conformers of the full structure can be used as validated initial geometries for subsequent resource-intensive calculations.

The conformation maps of di- and trisaccharide fragments are calculated automatically using REStLESS and molecular dynamics, and low-energy conformers are stored. The perspective usage of these data and tool are predictions of nuclear Overhauser effects for any given oligosaccharide or glycoconjugate and usage of atomic coordinates for large-scale virtual screening. Up to now, such applications were hardly possible due to a high computer cost of preliminary conformational studies of each involved molecular structure. The general route of data processing is summarized in the figure.



Our tool has a web-interface and is freely available at http://csdb.glycoscience.ru/csdb2atoms.html. The work on linking the symbolic and atomic notations was funded by Russian Science Foundation grant 18 04 00094. The conformational studies were funded by Russian Foundation for Basic Research grant 18 14 00098.

References

[1] Toukach Ph.V. Bacterial carbohydrate structure database 3: principles and realization – *J. Chem. Inf. Model* **2011**, 51(1): 159–170. doi: 10.1021/ci100150d.

[2] Chernyshov I.Yu., Toukach Ph.V. REStLESS: automated translation of glycan sequences from residue-based notation to SMILES and atomic coordinates - *Bioinformatics* **2018**, 34(15): 2679-2681. doi: 10.1093/bioinformatics/bty168.

[3] Toukach Ph.V., Egorova K.S. Carbohydrate structure database merged from bacterial, archaeal, plant and fungal parts - *Nucleic Acids Res* **2016**, 44(D1): D1229–D1236. doi: 10.1093/nar/gkv840.

Back to Program

Thursday

13.30

GlyGen – Computational and Informatics Resources for Glycosciences

René Ranzinger

University of Georgia
Complex Carbohydrate Research Center (CCRC)
Athens, GA, United States of America

Advancing our understanding of the roles that glycosylation plays in development and disease is hindered by the diversity of the data that must be integrated to gain insight into these complex phenomena. GlyGen is a new initiative supported by the NIH Common Fund with the goal of democratizing glycoscience by implementing a comprehensive data repository that integrates diverse types of data, including glycan structures, glycan biosynthesis enzymes, glycoproteins, and three-dimensional glycoprotein structures along with genomic and proteomic knowledge.

As part of GlyGen we have established collaborations with database providers from different domains (including but not limited to EBI, NCBI, GlyTouCan and UniCarbKB) in order to populate the repository with data. All information from these resources are standardized and crosslinked with dataset from the other resources to allow queries across multiple domains. To provide the community with an easy way to access the information, an intuitive, web based interface (http://glygen.org/) has been developed to visually represent the data and the connections between the individual datasets. In addition to the browser based interface, we are also developing RESTful webservice based APIs and an SPARQL endpoint, allowing programmatic access to the integrated datasets.

Our goal is to provide both trained and aspiring glycoscientists an easy way to access the complex information involving glycans and proteins in biology. One aspect we have implemented in this context is a query interface which suggest specific questions that a user is likely to have, such as “What are the enzymes involved in the biosynthesis of glycan X in humans?”

By specifying only one or few pieces of information (e.g. glycan X in the example) a complex query across multiple datasets and domains is performed and the result returned to the user. These questions are based on a collection of use cases that we compiled using mailing lists and community meetings.

Back to Program

Thursday
14.05

Introducing UniCarbKB 2.0 and SPRIT-Gly

Taherzadeh G[1], Brian Fochtman[2], Rahi Navelkar[2], Jodie L. Abrahams[3], Raja Mazumder[2], Yaoqi Zhou[1,3], Matthew P. Campbell[3]

[1]Griffith University, School of Information and Communication Technology, Gold Coast, Australia

[2]George Washington University, Department of Biochemistry and Molecular Medicine, Washington, DC, United States of America

[3]Griffith University, Institute for Glycomics, Gold Coast, Australia

UniCarbKB (http://unicarbkb.org) is a collaborative, international open science effort that aims to integrate and curate structural, experimental and functional glycoproteomics information from many sources. Since its launch in 2012 UniCarbKB has grown spanning 'large scale' glycoproteomics data collections to cell line glycomic profiling, however, over time the underlying architecture and interface has become less intuitive.

To improve data representation and discovery we have started to rebuild UniCarbKB from the ground up to provide a lighter, faster and simpler resource. In collaboration with the GlyGen initiative (http://glygen.org) UniCarbKB 2.0 has been extensively curated including revisions to the legacy GlycoSuiteDB data collections and now provides access to more recent glycoproteomic data collections. These efforts alone have doubled the number of curated glycoproteins in the knowledgebase and describe newly reported glycosylation sites not listed in other resources.

Here, we will debut new analyses over our integrated data collections, highlighting the breadth of knowledge that can be acquired from UniCarbKB 2.0 together with a wide-range of data services, including the adoption of Semantic Technologies and the availability of a unique and novel web services. In addition, cross-references between UniCarbKB and GlyGen enable users to explore structures, pathways and networks. Current efforts aim to increase the coverage and quality of data generated through this collaboration and examples will be provided in this presentation.

The continuing growth of UniCarbKB presents a number of opportunities to develop data mining tools that harness the curated knowledge available. In this presentation, we'll discuss how we used UniCarbKB to deliver a machine learning platform, called SPRINT-Gly, to improve the prediction and validation of *N*-/*O*-linked glycosylation sites using deep learning neural networks and Support Vector Machine (SVM) classifiers. To train SPRINT-Gly we constructed the largest known dataset of human and mouse glycosylation sites sourced from UniCarbKB and other open-access resources including GlycoProtDB, UniPep, UniProtKB and dbPTM.

Here we present an overview of SPRINT-Gly including the machine learning approaches (Tensorflow and LIBSVM), and how we integrated glycosylation site feature representations, protein sequence-based features, structural properties (generated by SPIDER 3.0), and intrinsic disorder information (determined by SPOT-disorder) to develop a unique approach for the prediction of potential glycosylation sites.

Back to Program

Thursday

15.10

GlyFinder and GlyProbity: New Online Tools for Locating and Curating Carbohydrate Structures in wwPDB

Robert Woods

University of Georgia
Complex Carbohydrate Research Center (CCRC)
Athens, GA, United States of America

The World Wide Protein Data Bank (wwPDB) contains more than 140,000 3D structures of biomolecules, many of which are directly relevant to human health and disease. Approximately 20% of these structures contain carbohydrates as ligands or as post-translational modifications. While numerous tools exist to curate protein 3D structural data, no such tools have been adopted by the PDB as part of the validation checks performed upon coordinate deposition. This oversight has resulted in a large number of errors and inconsistencies in annotation and structure in the carbohydrate structural data. Here we report on a joint project with the developers of wwPDB to create and implement tools to address these issues as part of a broader carbohydrate remediation initiative at the wwPDB.

At the present time there are three serious problems that hinder the utilization of carbohydrate data stored in the wwPDB:

1) There are an unacceptably high proportion of errors in the deposited coordinates.
2) No convenient interface exists for searching for carbohydrate structures in the PDB.
3) There is no quality rigorous assessment or curation of the deposited carbohydrate coordinates

We have created an online search interface, “GlyFinder” implemented at GLYCAM-Web that greatly simplifies the task of locating relevant carbohydrate containing structures. We are also generating an online tool called “GlyProbity" for checking the accuracy and internal consistency of 3D structures of carbohydrates, and will implement this tool for the data remediation.

Taken together, these aims should significantly impact the development of glycomimetic therapeutics, as well as the generation of structure/function relationships in glycobiology, and will be essential for achieving interoperability with additional databases or data mining services in the future.

We report on the current status and capabilities of the GlyFinder and GlyProbity tools.

Back to Program

Thursday

15.45

O-Glycologue: a Simulator of the Enzymes of *O*-linked Glycosylation

Andrew McDonald

Trinity College Dublin
School of Biochemistry and Immunology
Dublin, Ireland

O-Glycologue is a formal-language based simulator of the biosynthetic enzymes of *O*-linked glycosylation. Glycan acceptor substrates are entered in their native format, or else imported as GlycoCT condensed, or IUPAC condensed names. The substrates are passed to the enzymes, which are modelled as regular-expression string substitutions. The resulting networks of reactions can be exported as SBML, while the effects of knocking out different sets of enzyme activities can be compared. The enzymes used to produce a given substrate can be predicted, and the method is applied to a set of human gastric mucin *O*-glycans. Adaptation of the method to other systems of glycosylation is also demonstrated.

Back to Program

Thursday

16.20

MIRAGE Presentation and Open Discussion on Data and Tools

René Ranzinger, Carsten Kettner and
the MIRAGE Commission

The co-ordinating team and the MIRAGE Commission

MIRAGE stands for “Minimum Information Required for A Glycomics Experiment”. The aim of the MIRAGE project is to improve the quality of glycomics data in the scientific literature. Researchers seeking to understand the biochemical structure–function relationships of carbohydrates require detailed descriptions of the assay conditions and the experimental results. Currently, these data are insufficiently reported in the literature.

Under the auspices of the Beilstein-Institut, in 2011 the MIRAGE project started to define requirements for the publication of data in glycomics. The members of the MIRAGE Commission meets at least once a year for a few days to discuss publication guidelines and strategies for co-operations and publications. Since the definition of reporting guidelines cannot be a top-down approach the Commission relies on comments and suggestions from the glycomics community to gain widest acceptance and support to make the proposition of guidelines a movement from the grass roots .

Here, the latest developments will be presented and discussed with the wider audience at the symposium.

Back to Program