Open Access Book Chapter

Published version

Chapter title	Perspectives in Glycomics and Lectin Engineering
Title in English	
Author(s)	Jan Tkac, Tomas Bertok, Jozef Nahalka, Peter Gemeiner
DOI	10.1007/978-1-4939-1292-6_37
Copyright holder	Publisher
License	All rights reserved
Language	English

Original publication

Book title	Lectins. Methods and protocols
Title in English	
Publisher	Springer
Publication date	22-07-2014
ISBN	9781493912919
DOI	10.1007/978-1-4939-1292-6
Editor(s)	Jun Hirabayashi

ERC Grant information

Project title	Electrochemical LEctin and glycan biochips integrated with NAnosti
Project acronym	ELENA
Grant number	311532
ERC principal investigator	Jan Tkac

OAPEN Foundation received funding from the European Research Council (ERC)
under the European Union's Horizon 2020 research and innovation programme (grant agreement No 683680).

Perspectives in Glycomics and Lectin engineering

Jan Tkac[a,b], Tomas Bertok[a], Jozef Nahalka[a,b] and Peter Gemeiner[*,a]

[a]Institute of Chemistry, Center for Glycomics, Slovak Academy of Sciences, Dubravska cesta 9, 845 38 Bratislava, Slovakia;

[b]Institute of Chemistry, Center of excellence for white-green biotechnology, Slovak Academy of Sciences, Trieda Andreja Hlinku 2, SK-94976 Nitra, Slovakia

Summary

This chapter would like to provide a short survey of the most promising concepts applied recently in analysis of glycoproteins based on lectins. The first part describes the most exciting analytical approaches used in the field of glycoprofiling based on integration of nanoparticles, nanowires, nanotubes or nanochannels or using novel transducing platforms allowing to detect very low levels of glycoproteins in a label-free mode of operation. The second part describes application of recombinant lectins containing several tags applied for oriented and ordered immobilisation of lectins. Besides already established concepts of glycoprofiling several novel aspects, which we think will be taken into account for future, more robust glycan analysis are described including modified lectins, peptide lectin aptamers and DNA aptamers with lectin-like specificity introduced by modified nucleotides. The last part of the chapter describes a novel concept of a glycocodon, what can lead to a better understanding of glycan-lectin interaction and for design of novel lectins with unknown specificities and/or better affinities toward glycan target or for rational design of peptide lectin aptamers or DNA aptamers.

Keywords: biosensors, glycomics, lectins, nanoparticles, DNA aptamers, lectin peptide aptamers, recombinant lectins

[*] Corresponding author. Tel.: +421 2 5941 0211, Fax.: +421 2 5941 0222, E-mail: Peter.Gemeiner@savba.sk

1. Introduction

Since the introduction of DNA biochips in 1995 *(1)*, the technology has been intensively applied in assays of genome-wide expression to seek information about possible functions of novel or poorly characterised genes *(2)* and for diagnostic purposes, as well *(3)*. Even though DNA microarray technology has shed light on many physiological functions of genes by determination of expression of gene clusters, there is quite often only a very low correlation between RNA and protein abundance detected in single-cell organisms *(4)* and in higher ones, including humans *(5)*. Since quantitative analysis of proteins is central to proteomics with a focus on design of novel drugs, diagnostics of diseases and their therapeutic applications, protein microarrays were successfully launched to address these issues *(6)* .

Analysis of finely tuned post-translational modifications (PTMs) of proteins is an additional challenge for current analytical technology. Glycosylation is a highly abundant form of PTM of proteins and it is estimated that 70%–80% of human proteins are glycosylated *(7)*. Importance of glycans can be further highlighted by the fact that 70% of all therapeutic proteins are glycosylated *(8)*. Glycan mediated recognition plays an important role in many different cell´s processes such as fertilisation, immune response, differentiation of cells, cell-matrix interaction, cell-cell adhesion etc. *(9, 10)*. Glycans present on the surface of cells are naturally involved in pathological processes including viral and bacterial infections, in neurological disorder and in tumour growth and metastasis *(3, 11–16)*. Thus, better understanding of glycan mediated pathogenesis is essential in order to establish a *"policy"* to develop efficient routes for disease treatment with several recent studies as good examples e.g. "neutralisation" of various forms of viruses *(17,18)* or more efficient vaccines against various diseases *(19, 20)*. A changed glycosylation on a protein backbone can be effectively applied in early stage diagnostics of several diseases, including different forms of cancer with known glycan-based biomarkers *(21–23)*. Moreover, many previously established and even commercially successful strategies used to treat diseases are currently being revisited in light of glycan recognition in order to lower side effects, enhance serum half-life or to decrease cellular toxicity *(3, 24, 25)*. Recently, the first glyco-engineered antibody was approved to the market, what was called by the authors "a triumph for glyco-engineering" *(26)*.

Glycomics focuses on revealing finely tuned reading mechanisms in the cell orchestra based on graded affinity, avidity and multivalency of glycans (i.e. sugar chains covalently attached to proteins and lipids) *(27)*. Glycans are information-rich molecules applicable in coding tools of the cell since they can form enormous number of possible unique sequences from basic building units *(28)*. It is estimated that the size of the cellular glycome can be up to 500,000 glycan modified biomolecules (proteins and lipids) formed from 7,000 unique glycan sequences *(29)*. Thus, it is not a surprise the glycome is sometimes referred to as the ''third alphabet'' in biology, after genetics and proteomics *(30)*. A huge glycan variation can explain human complexity in light of a paradoxically small genome. This glycan complexity together with similar physico-chemical properties of glycans is the main reason why the progress in the field of glycomics has been behind advances in genomics and proteomics *(31)*.

Traditional glycoprofiling protocols rely on glycan release from a biomolecule with a subsequent quantification by an array of techniques including capillary electrophoresis, liquid chromatography and mass spectrometry *(32–34, 30, 35–37)*. There is an alternative way for glycoprofiling by application of lectins, natural glycan recognising proteins *(28, 38, 39)* in combination with various transducing protocols *(12, 40, 41)*. The most powerful glycoprofiling tool relies on lectins arrayed on solid surfaces for direct analysis of glycoproteins, glycolipids, membranes and even glycans on the surface of intact cells *(11, 42, 12, 43)*. Even though lectin microarrays offer high throughput assay protocols with a minute consumption of samples and reagents, there are some drawbacks such as a need to fluorescently label the sample or the lectin, which negatively affects the performance of detection *(11, 12)*, relatively high detection limits and quite narrow working concentration ranges. Thus, the ideal detection platform should be based on protocols without a need to label a glycoprotein or a lectin, in a way similar to natural processes occurring within a cell *(30)*.

Lectins (lat. *legere* = to choose) are proteins able to recognise and reversibly bind to free or bound mono- and oligosaccharides *(44)*. They are not usually catalytically active, do not participate in the immune response of higher organisms and can be found in viruses, bacteria, fungi, plants and animals. They are therefore a relatively heterogeneous group of oligomeric proteins belonging to distinct families with similar sequences and are considered as natural glycocode decipherers *(28)*. Lectins, unlike antibodies, have a low specificity and affinity with K_d ranged from 10^{-3} to 10^{-7} M *(30)* and lectins with a new specificity cannot be raised in a way similar to antibodies.

In the following sections we will focus on ways how to improve glycan detection either by application of novel nanoscale-controlled patterning protocols, nanoengineered devices or by application of novel, recombinant lectins, lectin-like aptamers or lectin peptide aptamers. A final part of this book chapter will focus on a completely novel area in the glycomics – the idea of a glycocodon.

2. Perspectives of Novel Formats of Analysis Applicable in Glycomics

The use of nanotechnology, sophisticated nanoscale patterning protocols and advanced detection platforms can help to overcome the drawbacks of lectin microarray technology allowing it to work in a label-free mode of operation, with a high sensitivity, low detection limits, a wide concentration window and in some cases, real time analysis of a binding event is possible *(9, 45–47, 30, 48–51)*. These devices can differ in their mode of signal transduction compared to traditional methods and will be divided into three categories according to their mode of action. Various traditional analytical techniques based on lectins (i.e. surface plasmon resonance and quartz crystal microbalance) are covered by different chapters accompanied this one within this book and are not discussed here.

2.1. Mechanical Platforms

Microcantilever biochips offer a novel approach for detection of a molecular binding based on a change in mass accumulated on the surface of a cantilever during biorecognition. It is a label-free technique allowing to monitor biospecific interaction in a real time, thus, affinity constants of the interaction can be acquired. When a biorecognition takes place, a particular cantilever bends, what results in the shift of a laser beam angle, allowing for a direct detection of the binding event (**Fig. 1A**).

Fig. 1 near here

The device was prepared with variations in the density and composition of a glycan determinant immobilised via a thiol-gold surface chemistry on a cantilever surface. Namely galactose, trimannose and nonamannose were attached on the surface and probed with two different lectins – cyanovirin A and Concanavalin A (Con A). The later was successfully detected on a surface with optimal glycan composition down to a nanomolar range *(52)*. The sensitivity is not impressive, but comparable to traditional surface plasmon resonance and quartz crystal microbalance lectin-based biosensors. The Seeberger´s group later extended this concept for analysis of several *Escherichia coli* strains on microcantilever biochips functionalised with different mannosides with a specific and reproducible detection with an amount of detectable *E. coli* cells over four orders of magnitude *(53)*.

2.2. Electrical Platforms

Electrical/electrochemical detection is quite often utilised in combination with other techniques in the field of glycomics for some time *(54)*. Electrical platforms of detection of a biorecognition event are primarily based on changes in the electrical signals such as resistance, impedance, capacitance, conductance, potential, and current *(55)*. These analytical techniques are usually non-destructive, extremely sensitive, offering quite a wide concentration working range with a possibility to work in an array format of analysis *(30)*.

2.2.1. Electrochemical impedance spectroscopy (EIS)

The most frequently used label-free electrochemical technique is EIS, which is based on an electric perturbation of a thin layer on the conductive surface by small alternating current amplitude with ability to provide characteristics of this interface utilisable in sensing. EIS results are typically transformed into a complex plane Nyquist plot vectors, which by application of an equivalent circuit can provide information about electron transfer resistance of a soluble redox probe in a direct way (**Fig. 1B**). When a biorecognition took place, an electrode interface is modified and a subtle change in interfacial layer characteristics can be used for detection. EIS investigation is most frequently performed in the presence of a redox

probe with detection of a change of resistance of the interface used for a signal generation. EIS is extensively used as a non-destructive technique for reliable analysis of surface conditions and allows complex biorecognition events to be probed in a simple, sensitive and label-free manner and is being increasingly popular to develop electrochemical lectin-based biosensors for glycan determination *(9, 30)*.

Initial efforts to detect glycoproteins by EIS were launched by the group of Prof. Joshi with sialic acid binding *Sambucus nigra* agglutinin (SNA) and a galactose binding peanut agglutinin covalently immobilised on printed circuit board electrodes *(56)*. The assays were really quick with a response time of 80 s and with sensitivity of glycoprotein detection down to 10 pg mL^{-1} (e.g. 150 fM), while using a cost-effective electrode material *(56)*. A group of Prof. Oliveira put a substantial effort to use lectin modified surfaces with EIS detection for discrimination between healthy human samples and samples from patients infected by a mosquito-borne Dengue virus (breakbone fever) with a high mortality rate *(9)*. Their device with two different lectins immobilised on gold nanoparticles offered a detection limit in the low nM range *(57)*. Another EIS-based biosensor was built on a surface of the silicon chip with an array of gold electrodes interfaced with nanoporous alumina membrane with high density of nanowells *(58)*. The biosensor offered a high reliability of assays and a good agreement with enzyme-linked lectin assays (ELLA). The detection limit of a biosensor for its analyte was 5 orders of magnitude lower compared to ELLA (i.e. 20 fM *vs.* 4.6 nM). An assay time for the biosensor of 15 min was much shorter compared to 4 h needed for ELLA. Moreover, a minute amount of sample (10 μL) was sufficient for the analysis by the biosensor *(58)*.

In our recent work we focused on the development of ultrasensitive impedimetric lectin biosensors with detection limits down to a single-molecule level based on controlled architecture at the nanoscale *(59–61)*. In the first study the biosensor was able to detect a glycoprotein in a concentration window spanning 7 orders of magnitude with a detection limit for the glycoprotein down to 0.3 fM, what was the lowest glycoprotein concentration detected *(59)*. In the following study an incorporation of gold nanoparticles offered even lower and unprecedented detection limit of 0.5 aM with quite a wide dynamic concentration range covered *(61)*. In our last study the EIS-biosensors were constructed with three different lectins to be able to detect changes on immunoglobulins with progression of a rheumatoid arthritis in humans. The biosensor with improved antifouling properties offered a detection limit in the fM range and worked properly even with 1,000x diluted human plasma. The biosensor performance was directly compared to the state-of-the-art glycoprofiling tool based on fluorescent lectin microarrays with a detection limit in the nM level *(60)*. Moreover, a sandwich configuration offered a detection limit down to aM concentration *(60)*. A detection limit down to fM range for analysis of alpha-fetoprotein (a biomarker for hepatocellular carcinoma) was recently observed on a device modified by arrays of single-walled carbon nanotubes and wheat-germ agglutinin with EIS as a transducing mechanism *(62)*.

2.2.2. *Nanotube field effect transistor (NTFET) sensors*

In NTFETs, semiconducting nanotubes or nanowires act as a channel between two metal electrodes (source and drain) while the two electrodes are held at a constant bias voltage using a so called gate electrode (**Fig. 2A**) *(49, 30)*. When the device with an immobilised biorecognition element is exposed to the sample containing its binding partner, a change of the device conductivity can be applied for quantification of the analyte. The application of the FET devices in the field of glycobiology was pioneered by Star´s group *(63, 64)*. In the initial study carbon nanotubes were employed as a channel when glycoconjugate was immobilised on a surface of the device and an analyte lectin down to 2 nM concentration could be detected *(63)*. A forthcoming study confirmed that a carbon nanotube biosensor for detection of a lectin outperformed a device based on graphene *(64)*. However, semiconducting carbon nanotubes with a high purity are required to achieve better signal quality as a further research goal. Silicon nanowires were applied as a FET channel more effectively compared to carbon nanotubes and graphene, since a detection limit for a lectin down to 100 fg mL^{-1} (≈ fM level) was achieved *(65)*. Even though such a remarkable concentration of lectin with a glycan modified FET device was detected, analysis of glycoproteins on a lectin immobilised surface can be more problematic since the device is able to detect changes in a close proximity to the surface and a biorecognition lectin-glycoprotein can be too far from the surface to be detected. The solution, however, can lie in an application of ultra-diluted (100x or 1,000x) phosphate buffers allowing to detect biorecognition event at distances 7.5–23.9 nm from the surface, but for analysis of protein levels in serum, serum has to be desalted prior detection *(66)*.

Fig. 2 near here

Another interesting approach applied in glycoassays was based on immobilisation of mannose inside a nanochannel and changes in the nanochannel conductance were after binding of Con A detected in a concentration window from 10 nM to 1,500 nM *(67)*. The question how sensitive analysis of glycoproteins with lectins immobilised within a nanochannel can be, has to be still answered.

2.3. Optical Platforms

There are two different optical sensing mechanisms applied in label-free glycoanalysis. The first is based on an intrinsic fluorescence of carbon nanotubes and the second one on a localised surface plasmon resonance detected on gold nanoislands. Both concepts have advantage since there is no necessity for an electronic interfacing, what is a problematic aspect of FET devices, and the nanoscale sensors require only a minute amount of sample for analysis.

2.3.1. Quenching of an intrinsic carbon nanotube fluorescence

This platform of detection employs fluorescent carbon nanotubes with a flexible NTA-nickel tether attached, modulating fluorescent intensity of carbon nanotubes on one side and being applied as a coupling agent for His_6-tagged lectins. When the glycoprotein interacts with an immobilised lectin, a nickel ion moves away from the carbon nanotube surface, partially restoring a quenched fluorescence of carbon nanotubes (**Fig. 2B**). An increase in the fluorescence output can be applied not only for quantification of a glycoprotein level, but for monitoring of the interaction in a real time, providing kinetic and affinity constants, as well. The absolute detection limit of the device for the glycoprotein was not that impressive (2 μg i.e. 670 nM), but authors believe the device has a room for improvement (i.e. by using high quality nanotube sensors) *(47, 68)*. In a recent study authors extended this initial study for glycoprofiling of different forms of IgGs *(69)*.

2.3.2. A localised surface plasmon resonance

Noble metal nanostructures exhibiting a localised surface plasmon resonance, sensitive to changes in the refractive index near the nanostructures, can be integrated into a biosensor device. There is only one report on application of such a device in glycoassays (**Fig. 3**). In this case mannose was immobilised on the surface of Au nanoislands and sensitivity towards a lectin was probed under stationary or flow conditions. Moreover, kinetic parameters of lectin interaction were obtained in an agreement with traditional techniques. Mannose-coated transducers offered an excellent selectivity towards Con A down to concentration of 5 nM in the presence of a large excess of bovine serum albumin (BSA) *(70)*.

Fig. 3 near here

In summary, it can be concluded that from all novel nanoengineered devices offering label-free mode of detection EIS-based biosensors have a great potential for glycan analysis since they can clearly outperform the state-of-the-art tool in a glycoprofiling, lectin microarrays, in terms of a detection limit achieved and a dynamic concentration range of analysis offered. EIS lectin biosensors were successfully applied in analysis of complex samples such as human serum even at dilution of 1,000x. Only such sensitive devices can really detect ultralow concentration of disease markers directly in human serum, a feature important for early stage prognosis of a particular disease. Moreover, biosensor devices with a detection limit down to single molecule level have a potential to be applied for identification of novel biomarkers, which can be present in human body liquids at concentrations not detectable by other analytical platforms of detection involving lectins. Other analytical tools based on mechanical, FET and optical signal transduction mechanisms have to prove their analytical potential in glycoprofiling with lectins immobilised on surfaces of such devices.

3. Perspectives in Lectin Engineering

The glycan binding sites of lectins are usually a shallow groove or a pocket present at the protein surface, or at the interface of oligomers *(7)*. Four main amino acids are part of an affinity site including asparagine, aspartic acid, glycine (arginine in Con A) and an aromatic residue for interaction with glycan via hydrogen bonds and hydrophobic interactions *(71)*. Ionic interactions are especially involved in recognition of negatively charged glycans containing sialic acids. Lectin-monosaccharide binding is relatively weak, this is why several approaches were applied to enhance practical utility of lectins in glycoprofiling *(7)*.

Recombinant DNA technology for producing lectins was traditionally applied to establish primary structure; to study genetics, evolution and biosynthesis; to elucidate the role of amino acids in recognition; to produce lectins with altered specificity and/or affinity; and to study their function in the organism of origin *(72)*. Novel trend is to apply this technology for producing lectins to be utilised for construction of various lectin-based biodevices. Recombinant lectin technology can significantly reduce drawbacks of traditional lectin isolation such as a long processing time, often quite a low yield, batch-to-batch variation of the product quality depending on the source, with presence of various contaminants or different lectin isoforms *(73, 11)*. Moreover, recombinant technology offers to produce lectins either without any glycosylation, which can in many cases complicate glycoprofiling, by expression in prokaryotic hosts and to introduce various tags (His_6-tag, glutathione transferase), which can be effectively utilised not only for one-step purification process, but more importantly for an oriented immobilisation of lectins on various surfaces *(38)*. Although lectin peptide aptamers have not been produced yet, it is a question of time, when such artificial glycan binding proteins emerge as an efficient tool in the area of glycobiology. It is estimated that another player in the area of glycoprofiling will make a substantial fingerprint lectin aptamers based on expanded genetic alphabet by introduction of modified nucleotides. Other concepts based on modified lectins in glycoprofiling with added value are finally described.

3.1. Oriented Immobilisation of Recombinant Lectins

A controlled immobilisation of lectins on a diverse range of surfaces can have a detrimental effect on the sensitivity of assays, since lectins can be attached in a way a biorecognition site is directly exposed to the solution phase for an efficient biorecognition. As a result almost 100% of immobilised lectin molecules can have a proper orientation with an increased chance for catching its analyte (**Fig. 4A**). Moreover a presence of a linker, which attaches tag to the protein backbone can significantly lower possible interaction of the protein with the surface, what can eventually lead to a denaturation of a protein *(74)*.

Fig. 4 near here

In a pilot study seven bacterial lectins, having a His_6-GST tag, were expressed in *E. coli* and subsequently applied for construction of a complete recombinant lectin microarray, which was utilised to probe differences between several tumour cells (ACHN, TK10, SK-MEL-5 and M14 cancer cell lines) *(75)*. For that purpose isolated membrane micelle from the tumour cell lines were employed. Such a procedure avoids use of proteases, which can change composition of samples containing glycoproteins, and at the same time there is no need to work with whole cells, allowing to work with a small spot sizes. The results showed distinct variations between tumour cell lines expressing different glycan moieties. In order to have a control spot on a lectin microarray a mutated form of one lectin was introduced, allowing to quantify specificity of interaction. Moreover, it was found out that in the presence of monosaccharides during a lectin printing process better resolved spot morphology and lectin activity was achieved *(75)*. In a next study of the Mahal´s group, the effect of oriented immobilisation of lectins on the sensitivity of lectin microarray assays was quantified. Oriented immobilisation of lectins offered a detection limit of approx. 12 ng mL^{-1} (ca. 640 pM protein) *(76)*, a significantly lower level compared to a detection limit of 10 µg mL^{-1} achieved on a lectin microarray with a random immobilisation *(33)*. In a next study the group developed oriented immobilisation of recombinant lectins in a single step deposition of lectins together with a glutathione to an activated chip surface. Such an approach simplifies an overall immobilisation process because the surface does not need to be modified by glutathione prior lectin immobilisation *(77)*.

Another group developed an oriented immobilisation of recombinant lectins produced with a fused Fc-fragment. Such a fragment has an affinity towards protein G (expressed in *Streptococcus* sp., much like a protein A) or a carbohydrate moiety of Fc fragment has an affinity for boronate derivatives (**Fig. 4B**). Thus, a surface modified by a boronate derivatives or a protein G was effectively applied for oriented immobilisation of a recombinant lectin via a fused Fc fragment *(78)*. Although boronate immobilisation approach showed the highest sensitivity of detection, presence of boronate functionalities on the chip surface induced non-specific interactions with glycoproteins and thus dextran blocking was introduced to minimise unwanted glycoprotein interactions. Additional drawback of such approach can be expected by introduction of Fc fragment having glycan entities, what can interact with glycan-binding proteins, which might be present in complex samples.

3.2. Perspectives for Peptide Lectin Aptamers

An alternative to production of mutated forms of lectins or glycosidases *(7)* for subsequent application in glycoprofiling in a future might be a preparation of novel forms of glycan-recognising proteins termed here as peptide lectin aptamers (PLA). Such proteins will be peptide aptamers with a lectin-like affinity to recognise various forms of glycans. A term peptide aptamer was coined by Colas et al. in 1996 *(79)* and is defined as a combinatorial protein molecule having a variable peptide sequence, with an affinity for a given target protein, displayed on an inert, constant scaffold protein *(80)*. Construction of various

bioanalytical devices such as peptide lectin aptamer microarrays can benefit from such biorecognition elements since it would be possible to generate „army of terracotta soldiers“ looking at the molecular level almost identical besides distinct „facial“ feature of each entity provided by a unique peptide sequence. Even though lectin peptide aptamers have not been prepared yet, in our opinion, it is only a question of time, when such recognition elements will be prepared.

A beneficial feature of such a protein will be high solubility, a small, uniform size of a scaffold protein with an extended chemical and thermal stability and a possibility to express such proteins in prokaryotic expression systems, what is a cost-effective process *(81)*. Moreover, when a small PLA will be immobilised on the surface of various bioanalytical devices, a higher density of biorecognition element can enhance sensitivity of detection, while suppressing non-specific interactions and lowering background signals *(80)*. Peptide aptamers are produced by protein engineering from high-complexity combinatorial libraries with appropriate isolation/selection methods *(82, 80, 81)*. Thus, a need to have knowledge of the protein structure and the mechanism behind binding is not necessary. There are however some requirements for the scaffold protein to posses such as lack of a biological activity and ability to accommodate a wide range of peptides without changing a 3-D structure *(83)*.

Currently there are over 50 proteins described as potential affinity scaffolds, but only quite a few of them reached stage behind a proof of the concept phase *(81)*. Scaffold proteins were constructed from a diverse range of proteins differing in origin, size, structure, engineering protocols, mode of interaction and applicability and typically have from 58 up to 166 amino acids (**Fig. 5**) *(81, 84)*. For example peptide aptamers based on a Stefin A protein (a cysteine protease inhibitor) are working well in an immobilised state on gold and modified gold surfaces with K_d of a peptide aptamer for its analyte down to nM range *(85, 86)*. Moreover, the scaffold based on Stefin A can accommodate and tolerate more than one peptide insert, what can dramatically widen practical application of such peptide aptamers *(87)*. It is possible that in case of lectin peptide aptamers a restricted range of amino acids enriched in four amino acids involved in glycan recognition could be possible, a concept which was successfully implemented for recognition of a maltose binding protein in the past *(88)*.

3.3. Perspectives for Novel Lectin-like Aptamers

The name aptamer coming from the Latin expression “aptus” (to fit) and the Greek word “meros” (part) was coined in 1990 by Ellington and Szostak in order to introduce artificial RNA molecules binding to a small organic dye *(89)*. Aptamers are single-stranded oligonucleotides selectively binding small molecules, macromolecules or whole cells, generally with a size of 15–60 NTs (i.e. 5–20 kDa) (**Fig. 6**) *(80)*. Additional advantages of aptamers include relative simplicity of a chemical modification (introduction of a biotin or a fluorescent label), simple regeneration/reusability, and stability at a high temperature and/or at a high salt concentration *(90)*. Due to a small size of aptamers, they can be effectively

attached to the interfaces with high densities, a feature important for construction of various robust and sensitive bioanalytical devices *(91, 92)*. Either DNA or RNA has to be chosen for preparation of aptamers keeping in mind a final application. For example RNA is structurally more flexible compared to DNA and thus, theoretically such aptamers can be raised against a wider range of analytes *(81)*. Contrary, a major limitation of using RNA is their susceptibility to chemical and/or enzymatic degradation. Furthermore, selection of RNA aptamers is a time-consuming process requiring additional enzymatic steps. Modifications of the DNA or RNA backbone or introduction of modified nucleotides can produce aptamers more resistant to degradation *(93)*.

Fig. 5 near here

When aptamers interacts with its analyte, usually a conformational change occurs creating a specific binding site for the target. Aptamers for proteins generally exhibit quite a high affinity in nM or sub nM level due to presence of large complex areas with structures rich in hydrogen-bond donors and acceptors *(80)*. A relatively high affinity of aptamer makes such oligonucleotides an attractive alternative to lectins or antibodies as detection reagents for carbohydrate antigens. In order to increase palette of analytes being recognised by aptamers, modified nucleotides were introduced.

A rational approach for preparation of aptamers with a high affinity binding of glycoproteins by an extending library of nucleotides modified by incorporation of a boronic acid moiety was recently introduced by Wang´s lab *(94)*. The study showed that affinity with K_d of 6–17 nM for fibrinogen using boronate modified DNA aptamers was higher compared to the affinity with K_d of 64–122 nM for the same analyte using DNA aptamers with natural pool of nucleotides. The fact that for the interaction between a glycoprotein and boronate modified DNA aptamer it is important an interaction between boronate moiety and glycan of fibrinogen was confirmed by analysis of a deglycosylated fibrinogen with a decreased affinity (K_d of 87–390 nM) *(94)*.

An interesting approach for preparation of a wider library of nucleotides was recently introduced by incorporation of six new 5-position modified dUTP derivatives with 5 derivatives containing an aromatic ring *(95)*. DNA aptamers based on an extended pool of nucleotides were able to bind a necrosis factor receptor superfamily member 9 (TNFRSF9) with a high affinity of K_d=4–6 nM for the first time. Interestingly two new derivatives either containing indole derivative or a benzene ring were the best TNFRSF9 binders. This fact is quite interesting since TNFRSF9 is a glycoprotein *(96)* and we can only speculate that these two aromatic derivatives of dUTP were involved in recognition of TNFRSF9 via a glycan interaction. In a recent and similar study a derivative of imidazole (7-(2-thienyl)imidazo[4,5-*b*]pyridine, Ds) containing nucleotides was applied for generation of novel DNA aptamers binding to two glycoproteins vascular endothelial cell growth factor-165 (VEGF-165) and interferon-γ (IFN-γ) with an enhanced affinity *(97)*. The study revealed K_d down to 0.65 pM

with DNA aptamers based on Ds-nucleotides, while the best K_d of 57 pM for the DNA aptamers containing natural nucleotides was found for VEGF-165 *(97)*. Similarly, DNA aptamers based on Ds-nucleotides offered much lower K_d of 0.038 nM compared to DNA aptamers based on natural nucleotides with K_d of 9.1 nM for IFN-γ *(97)*. Here we can again only speculate if the role of Ds nucleotides in enhanced affinity for two glycoproteins is in interaction of Ds modified nucleotides with the glycan moiety of glycoproteins.

Fig. 6 near here

3.4. Other Novel Forms of Lectins

There are several very interesting strategies how to enhance analytical applicability of lectins by their simple modifications, which can dramatically influence the field of glycoprofiling in a future.

The first study focused on application of multimers of eight different lectins prepared by incubation of biotinylated lectins with streptavidin. A wheat germ agglutinin (WGA) multimers integrated into lectin microarrays showed 4–40 times better sensitivity in analysis of glycans in human plasma and much better performance in glycoprofiling of samples from people having pancreatic cancer compared to utilisation of WGA lectin monomer *(98)*. Authors of the study suggested that such lectin multimers with an enhanced affinity towards glycans can broaden the range of glycans, which can be detected. Moreover, according to authors lectin multimers might provide a fundamentally new biorecognition information not achievable by lectin monomers *(98)*. The second study described attachment of a boronate functionality to two different lectins in order to enhance affinity 2 to 60-fold for a particular glycan binding *(99)*. Such modified lectins were tested in a whole cell lysate with an excellent specificity for analysis of 295 *N*-linked glycopeptides. These results revealed that application of boronate modified lectins can facilitate identification of glycans present on the surface of low-abundant glycoproteins *(99)*. The third study indicated that by preparation of a lectin mutant with artificially introduced cysteine into lectin *Galanthus nivalis* agglutinin it was possible to prepare lectin dimmers via a disulfide linkage between two lectin mutants *(100)*. Agglutination activity of a lectin dimmer increased 16-fold compared to a lectin monomer and interestingly a transformation monomer/dimmer can be redox-switchable by addition of mild oxidation or reducing agents *(100)*.

It can be summed up there are very exciting concepts already introduced in the field of lectin glycoengineering such as integration of deglycosylated forms of recombinant lectins into lectin microarrays with enhanced sensitivity of analysis and with lower detection limits achieved. It is only a question of time, when a wider range of recombinant lectins with purification/immobilisation tags will be applied for oriented immobilisation of lectins combined with various transducers or devices. Application of lectins modified by boronate derivatives or

in a form of multimers/dimmers is a promising way for analysis of low abundant glycoproteins and possibly for analysis of glycoproteins, which cannot be detected by unmodified lectins.

We propose a future application of lectin peptide aptamers, what can even further enhance overall order of an immobilisation process compared to immobilisation of recombinant lectins with different tags for construction of various devices applicable in glycoprofiling. Although there is one rational approach for designing DNA aptamers with enhanced affinity towards glycoproteins by introduction of a boronate moiety into nucleotides there are two other reports focused on enhanced nucleotide alphabet created for generation of novel high affinity DNA aptamers against important targets. Interestingly in these two studies described especially glycoproteins were the main targets for such novel DNA aptamers based on an extended nucleotide alphabet, since such glycoproteins could not be recognised by "natural" DNA aptamers consisted of only natural nucleotides. Moreover, nucleotides in such novel DNA aptamers were modified mainly by aromatic amino acids, which are usually involved in glycan recognition by lectins. Thus, it is necessary to prove in a future if such modified nucleotides are really involved in glycan recognition.

4. A Glycocodon Hypothesis

A codon ($n_1n_2n_3$) is a sequence of three DNA or RNA nucleotides that corresponds to a specific amino acid or stop signal during protein synthesis, and the full set of codons is called a genetic code. The current state-of-the-art knowledge about the origin of the genetic code still remains as one of unsolved problems, and enormous number of theories can be divided to RNA world theories, protein world theories, co-evolution theories and stereochemical theories. Integration of these theories leads us to the conclusion that a system of four codons ("gnc", n= a - adenine, g - guanine, c - cytosine, u - uracil) and four amino acids (G - glycine, A - alanine, V - valine, D – aspartic acid) could be the original genetic code *(101–105)*. Research on a selection of particular RNA sequences with an amino acid binding activity, and a relation of those activities to the genetic code, have revealed an evidence that there is a highly robust connection between the genetic code and RNA-amino acid binding affinity *(106)*. It seems that the main part of the genetic code is influenced by a stereochemical prebiotic selection during the first polymerisation of G, A, V and D amino acids and g, c, u and a nucleic acids, however, only first two nucleotides of codons (n_1n_2) are directly related to amino acid stereoselectivity *(107)*.

Recently, a similar evolution process was proposed for "the glycocode" *(108)*. The bioinformatics quantification of "GAVD-dipeptides" in monosaccharide specific proteins revealed that the amino acid triplets, the glycocodons ($aa_1aa_2aa_3$), can be deduced for each glycan letter (monosaccharide). The glycocodons are composed from one polar amino acid, interacting with sugar –OH groups, and one specific dipeptide, usually detecting C-C hydrophobic patch (see Glc, Gal and GlcNAc binding, **Fig. 7a** and **7c**). **Figure 7** depicts a quantification spectra of "GAVD-dipeptides" in glucose (Glc), galacose (Gal), mannose (Man), fucose (Fuc), *N*-acetylglucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc)

specific proteins. In the case of Glc, Gal and GalNAc the maximal values of incidence of "GAVD-dipeptides" were taken for coding; AA for Glc, GA plus AG for Gal, and DD plus VV for GalNAc. In the case of Man and GlcNAc GD plus DG and GV plus VG from "GAVD-dipeptide pool" were taken for coding, because maximal values have been already taken by previous monosaccharides. During evolution, the GAVD-glycocodons were transformed to novel glycocodons by a positive selection for the increased diversity and functionality of a "sugar-protein language" that can be made with a larger amino acid alphabet. Nevertheless, evolution process holds hydropathic similarity; amino acids in the glycocodons are substituted by amino acids with similar polar properties, what minimises errors in established sugar-protein interactions. The bioinformatics quantification of dipeptides composed from all 20 amino acids revealed that GA plus AG for Gal were substituted mainly with SW and WS, AA for Glc can be substituted with MF, GD plus DG for Man can be substituted with AY plus YA, GV plus VG for GlcNAc can be substituted with SF plus FS, and DD plus VV for GalNAc can be substituted with QD plus LF. AV plus VA from a "GAVD-dipeptide pool" were selected for NAc-group sensing, in the case of GalNAc, they were transformed to MS plus SM and IT plus TI dipeptides. **Figure 7d** shows how the GalNAc glycocodons are used during *N*-glycosylation by bacterial oligosaccharyltransferase (*Campylobacter lari*, PGlB). PGlB accepts different oligosaccharides from a lipid carrier requiring an acetamido group at the C2 carbon of the first monosaccharide (GalNAc is the best), or even a "monosaccharide" *N*-acetylgalactosamine-diphospho-undecaprenyl is a good substrate *(109, 110)*. PGlB connects the C1 carbon in the first saccharide moiety (*N*-acetylgalactosamine) with the amide nitrogen of the acceptor (sequon) asparagine. DQNATF peptide has been recognised as an optimal acceptor sequence for PGlB *(111)*. According to the glycocodon theory, DQ dipeptide plus the acceptor asparagine makes the glycocodon for GalNAc. The next ATF sequence of the sequon peptide is inserted into the catalytic centre in such a way that the KTI and HLF glycocodons are formed. This process shows a basic difference between the codons and the glycocodons. When the codons are read from the nucleotide sequence, they are read in succession and do not overlap with one another. Contrary, the glycocodons are used in a way of a "key-lock" principle – three different protein chains can make the glycocodon in 3D space and the glycocodons frequently overlap (**Fig.7**).

Fig. 7 near here

It should be emphasised that the glycocodons were theoretically deduced by a bioinformatics study and it will be necessary to perform a study in the laboratory to establish the strongest correlation between the monosaccharides and the glycocodons and to determine the shortest peptides for the recognition of the specific monosaccharide. However, the glycocodon theory represents a tool showing how the peptide lectin aptamers or novel DNA aptamers based on nucleotide derivatives should be organised and programmed.

5. Conclusions

This chapter described various tools, which have been recently applied in order to extend analytical usefulness of lectin based devices in glycoprofiling. A positive aspect of recent effort in the field is utilisation of a great potential nanotechnology can bring into quite a complex and challenging analysis of glycans. Such approaches proved analysis of glycans by different biosensors can be extremely sensitive with a concentration range spanning few orders of magnitude, what are features essential in analysis of low-abundant glycoproteins. Control of immobilisation of lectins is other important issue, which was successfully addressed in a pilot study showing that attachment of recombinant lectins on surfaces via various tags present in recombinant lectin improved sensitivity of glycan analysis. The book chapter described also future prospect of peptide lectin aptamers in order to increase sensitivity and stability of analysis, while suppressing non-specific interactions. Additional issue to focus on in a future is investigation if modified nucleotides can be successfully applied in preparation of novel DNA aptamers targeting glycoproteins with high affinity and selectivity. The final part of the chapter describe the concept of a glycocodon, what can lead to a better understanding of glycan-lectin interaction and for design of novel lectins with unknown specificities and/or better affinities toward glycan target or for rational design of peptide lectin aptamers or DNA aptamers.

Acknowledgement

The financial support from the Slovak research and development agency APVV 0282-11 and from VEGA 2/0127/10 is acknowledged. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement n. 311532. This contribution is the result of the project implementation: Centre of excellence for white-green biotechnology, ITMS 26220120054, supported by the Research & Development Operational Programme funded by the ERDF.

References

1. Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467-470.

2. Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci* 95, 14863-14868.

3. Alizadeh, A.A., Eisen, M.B., Davis, R.E., Ma, C., Lossos, I.S., Rosenwald, A., Boldrick, J.C., Sabet, H., Tran, T., Yu, X., Powell, J.I., Yang, L., Marti, G.E., Moore, T., Hudson, J., Lu, L., Lewis, D.B., Tibshirani, R., Sherlock, G., Chan, W.C., Greiner, T.C., Weisenburger, D.D., Armitage, J.O., Warnke, R., Levy, R., Wilson, W., Grever, M.R., Byrd, J.C., Botstein, D., Brown, P.O. and Staudt, L.M. (2000) Distinct types of diffuse large b-cell lymphoma identified by gene expression profiling. *Nature* 403, 503-511.

4. Gygi, S.P., Rochon, Y., Franza, B.R. and Aebersold, R. (1999) Correlation between protein and mrna abundance in yeast. *Mol Cel Biol* 19, 1720-1730.

5. Gry, M., Rimini, R., Stromberg, S., Asplund, A., Ponten, F., Uhlen, M. and Nilsson, P. (2009) Correlations between rna and protein expression profiles in 23 human cell lines. *BMC Genom* 10, 365.

6. Lee, J.-R., Magee, D.M., Gaster, R.S., LaBaer, J. and Wang, S.X. (2013) Emerging protein array technologies for proteomics. *Expert Rev Proteom* 10, 65-75.

7. Arnaud, J., Audfray, A. and Imberty, A. (2013) Binding sugars: From natural lectins to synthetic receptors and engineered neolectins. *Chem Soc Rev* 42, 4798-4813.

8. Baker, J.L., Çelik, E. and DeLisa, M.P. (2013) Expanding the glycoengineering toolbox: The rise of bacterial n-linked protein glycosylation. *Trends Biotechnol* 31, 313-323.

9. Bertók, T., Katrlík, J., Gemeiner, P. and Tkac, J. (2013) Electrochemical lectin based biosensors as a label-free tool in glycomics. *Microchim Acta* 180, 1-13.

10. Varki, A. (2009) Essentials of glycobiology. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

11. Gemeiner, P., Mislovicová, D., Tkác, J., Svitel, J., Pätoprsty, V., Hrabárová, E., Kogan, G. and Kozár, T. (2009) Lectinomics:: Ii. A highway to biomedical/clinical diagnostics. *Biotechnol Adv* 27, 1-15.

12. Katrlík, J., Švitel, J., Gemeiner, P., Kožár, T. and Tkac, J. (2010) Glycan and lectin microarrays for glycomics and medicinal applications. *Med Res Rev* 30, 394-418.

13. Krishnamoorthy, L., Bess, J.W., Preston, A.B., Nagashima, K. and Mahal, L.K. (2009) Hiv-1 and microvesicles from t cells share a common glycome, arguing for a common origin. *Nat Chem Biol* 5, 244-250.

14. Schauer, R. and Kamerling, J.P. (2011) The chemistry and biology of trypanosomal trans-sialidases: Virulence factors in chagas disease and sleeping sickness. *ChemBioChem* 12, 2246-2264.

15. Song, X., Lasanajak, Y., Xia, B., Heimburg-Molinaro, J., Rhea, J.M., Ju, H., Zhao, C., Molinaro, R.J., Cummings, R.D. and Smith, D.F. (2011) Shotgun glycomics: A microarray strategy for functional glycomics. *Nat Meth* 8, 85-90.

16. Vaishnava, S., Yamamoto, M., Severson, K.M., Ruhn, K.A., Yu, X., Koren, O., Ley, R., Wakeland, E.K. and Hooper, L.V. (2011) The antibacterial lectin regiiiγ promotes the spatial segregation of microbiota and host in the intestine. *Science* 334, 255-258.

17. Burton, D.R., Poignard, P., Stanfield, R.L. and Wilson, I.A. (2012) Broadly neutralizing antibodies present new prospects to counter highly antigenically diverse viruses. *Science* 337, 183-186.

18. Pejchal, R., Doores, K.J., Walker, L.M., Khayat, R., Huang, P.-S., Wang, S.-K., Stanfield, R.L., Julien, J.-P., Ramos, A., Crispin, M., Depetris, R., Katpally, U., Marozsan, A., Cupo, A., Maloveste, S., Liu, Y., McBride, R., Ito, Y., Sanders, R.W., Ogohara, C., Paulson, J.C., Feizi, T., Scanlan, C.N., Wong, C.-H., Moore, J.P., Olson, W.C., Ward, A.B., Poignard, P., Schief, W.R., Burton, D.R. and Wilson, I.A. (2011) A potent and broad neutralizing antibody recognizes and penetrates the hiv glycan shield. *Science* 334, 1097-1103.

19. Kim, J.-H., Resende, R., Wennekes, T., Chen, H.-M., Bance, N., Buchini, S., Watts, A.G., Pilling, P., Streltsov, V.A., Petric, M., Liggins, R., Barrett, S., McKimm-Breschkin, J.L., Niikura, M. and Withers, S.G. (2013) Mechanism-based covalent neuraminidase inhibitors with broad-spectrum influenza antiviral activity. *Science* 340, 71-75.

20. Klein, F., Halper-Stromberg, A., Horwitz, J.A., Gruell, H., Scheid, J.F., Bournazos, S., Mouquet, H., Spatz, L.A., Diskin, R., Abadir, A., Zang, T., Dorner, M., Billerbeck, E., Labitt, R.N., Gaebler, C., Marcovecchio, P.M., Incesu, R.-B., Eisenreich, T.R., Bieniasz, P.D., Seaman, M.S., Bjorkman, P.J., Ravetch, J.V., Ploss, A. and Nussenzweig, M.C. (2012) Hiv therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature* 492, 118-122.

21. Chandler, K.B. and Goldman, R. (2013) Glycoprotein disease markers and single protein-omics. *Mol Cel Proteom* 12, 836-845.

22. Ferens-Sieczkowska, M., Kowalska, B. and Kratz, E.M. (2013) Seminal plasma glycoproteins in male infertility and prostate diseases: Is there a chance for glyco-biomarkers? *Biomarkers* 18, 10-22.

23. Gilgunn, S., Conroy, P.J., Saldova, R., Rudd, P.M. and O'Kennedy, R.J. (2013) Aberrant psa glycosylation[mdash]a sweet predictor of prostate cancer. *Nat Rev Urol* 10, 99-107.

24. Schmaltz, R.M., Hanson, S.R. and Wong, C.-H. (2011) Enzymes in the synthesis of glycoconjugates. *Chem Rev* 111, 4259-4307.

25. van Bueren, J.J.L., Rispens, T., Verploegen, S., van der Palen-Merkus, T., Stapel, S., Workman, L.J., James, H., van Berkel, P.H.C., van de Winkel, J.G.J., Platts-Mills, T.A.E. and Parren, P.W.H.I. (2011) Anti-galactose-[alpha]-1,3-galactose ige from allergic patients does not bind [alpha]-galactosylated glycans on intact therapeutic antibody fc domains. *Nat Biotech* 29, 574-576.

26. Beck, A. and Reichert, J.M. (2012) Marketing approval of mogamulizumab: A triumph for glyco-engineering. *mAbs* 4, 419-425.

27. Raman, R., Raguram, S., Venkataraman, G., Paulson, J.C. and Sasisekharan, R. (2005) Glycomics: An integrated systems approach to structure-function relationships of glycans. *Nat Meth* 2, 817-824.

28. Gabius, H.-J., André, S., Jiménez-Barbero, J., Romero, A. and Solís, D. (2011) From lectin structure to functional glycomics: Principles of the sugar code. *Trends Biochem Sci* 36, 298-313.

29. Cummings, R.D. (2009) The repertoire of glycan determinants in the human glycome. *Mol BioSystems* 5, 1087-1104.

30. Reuel, N.F., Mu, B., Zhang, J., Hinckley, A. and Strano, M.S. (2012) Nanoengineered glycan sensors enabling native glycoprofiling for medicinal applications: Towards profiling glycoproteins without labeling or liberation steps. *Chem Soc Rev* 41, 5744-5779.

31. Bertozzi, C.R., Kiessling and L., L. (2001) Chemical glycobiology. *Science* 291, 2357-2364.

32. Furukawa, J.-i., Fujitani, N. and Shinohara, Y. (2013) Recent advances in cellular glycomic analyses. *Biomolecules* 3, 198-225.

33. Rakus, J.F. and Mahal, L.K. (2011) New technologies for glycomic analysis: Toward a systematic understanding of the glycome. *Annu Rev Anal Chem* 4, 367-392.

34. Smith, D.F. and Cummings, R.D. (2013) Application of microarrays to deciphering the structure and function of the human glycome. *Mol Cel Proteom* 12, 902-912.

35. Alley, W.R., Mann, B.F. and Novotny, M.V. (2013) High-sensitivity analytical approaches for the structural characterization of glycoproteins. *Chem Rev* 113, 2668-2732.

36. Lazar, I.M., Lee, W. and Lazar, A.C. (2013) Glycoproteomics on the rise: Established methods, advanced techniques, sophisticated biological applications. *Electrophoresis* 34, 113-125.

37. Novotny, M., Alley, W., Jr. and Mann, B. (2013) Analytical glycobiology at high sensitivity: Current approaches and directions. *Glycoconj J* 30, 89-117.

38. Oliveira, C., Teixeira, J.A. and Domingues, L. (2013) Recombinant lectins: An array of tailor-made glycan-interaction biosynthetic tools. *Crit Rev Biotechnol* 33, 66-80.

39. Murphy, P., André, S. and Gabius, H.-J. (2013) The third dimension of reading the sugar code by lectins: Design of glycoclusters with cyclic scaffolds as tools with the aim to define correlations between spatial presentation and activity. *Molecules* 18, 4026-4053.

40. Mislovičová, D., Katrlík, J., Paulovičová, E., Gemeiner, P. and Tkac, J. (2012) Comparison of three distinct ella protocols for determination of apparent affinity constants between con a and glycoproteins. *Colloids Surf B: Biointerf* 94, 163-169.

41. Mislovičová, D., Gemeiner, P., Kozarova, A. and Kožár, T. (2009) Lectinomics i. Relevance of exogenous plant lectins in biomedical diagnostics. *Biologia* 64, 1-19.

42. Hirabayashi, J., Yamada, M., Kuno, A. and Tateno, H. (2013) Lectin microarrays: Concept, principle and applications. *Chem Soc Rev* 42, 4443-4458.

43. Krishnamoorthy, L. and Mahal, L.K. (2009) Glycomic analysis: An array of technologies. *ACS Chem Biol* 4, 715-732.

44. Lis, H. and Sharon, N. (1998) Lectins: Carbohydrate-specific proteins that mediate cellular recognition. *Chem Rev* 98, 637-674.

45. Cunningham, S., Gerlach, J.Q., Kane, M. and Joshi, L. (2010) Glyco-biosensors: Recent advances and applications for the detection of free and bound carbohydrates. *Analyst* 135, 2471-2480.

46. Gerlach, J.Q., Cunningham, S., Kane, M. and Joshi, L. (2010) Glycobiomimics and glycobiosensors. *Biochem Soc Trans* 38, 1333-1336.

47. Reuel, N.F., Ahn, J.-H., Kim, J.-H., Zhang, J., Boghossian, A.A., Mahal, L.K. and Strano, M.S. (2011) Transduction of glycan–lectin binding using near-infrared fluorescent single-walled carbon nanotubes for glycan profiling. *J Am Chem Soc* 133, 17923-17933.

48. Sanchez-Pomales , G. and Zangmeister, R.A. (2011) Recent advances in electrochemical glycobiosensing. *Int J Electrochem* 2011.

49. Tkac, J. and Davis, J.J. (2009) Label-free field effect protein sensing In: Davis JJ (ed) Engineering the bioelectronic interface: Applications to analyte biosensing and protein detection Royal Society of Chemistry, Cambridge, pp 193-224. doi:10.1039/9781847559777

50. Reichardt, N.C., Martin-Lomas, M. and Penades, S. (2013) Glyconanotechnology. *Chem Soc Rev* 42, 4358-4376.

51. Zeng, X., Andrade, C.A.S., Oliveira, M.D.L. and Sun, X.-L. (2012) Carbohydrate-protein interactions and their biosensing applications. *Anal Bioanal Chem* 402, 3161-3176.

52. Gruber, K., Horlacher, T., Castelli, R., Mader, A., Seeberger, P.H. and Hermann, B.A. (2011) Cantilever array sensors detect specific carbohydrate-protein interactions with picomolar sensitivity. *ACS Nano* 5, 3670-3678.

53. Mader, A., Gruber, K., Castelli, R., Hermann, B.A., Seeberger, P.H., Radler, J.O. and Leisner, M. (2012) Discrimination of escherichia coli strains using glycan cantilever array sensors. *Nano Lett* 12, 420-423.

54. Jelinek, R. and Kolusheva, S. (2004) Carbohydrate biosensors. *Chem Rev* 104, 5987-6015.

55. Luo, X. and Davis, J.J. (2013) Electrical biosensors and the label free detection of protein disease biomarkers. *Chem Soc Rev* 42, 5944-5962.

56. La Belle, J.T., Gerlach, J.Q., Svarovsky, S. and Joshi, L. (2007) Label-free impedimetric detection of glycan−lectin interactions. *Anal Chem* 79, 6959-6964.

57. Oliveira, M.D.L., Correia, M.T.S., Coelho, L.C.B.B. and Diniz, F.B. (2008) Electrochemical evaluation of lectin–sugar interaction on gold electrode modified with colloidal gold and polyvinyl butyral. *Colloids Surf B: Biointerf* 66, 13-19.

58. Nagaraj, V.J., Aithal, S., Eaton, S., Bothara, M., Wiktor, P. and Prasad, S. (2010) Nanomonitor: A miniature electronic biosensor for glycan biomarker detection. *Nanomedicine* 5, 369-378.

59. Bertok, T., Gemeiner, P., Mikula, M., Gemeiner, P. and Tkac, J. (2013) Ultrasensitive impedimetric lectin based biosensor for glycoproteins containing sialic acid. *Microchim Acta* 180, 151-159.

60. Bertok, T., Klukova, L., Sediva, A., Kasák, P., Semak, V., Micusik, M., Omastova, M., Chovanová, L., Vlček, M., Imrich, R., Vikartovska, A. and Tkac, J. (2013) Ultrasensitive impedimetric lectin biosensors with efficient antifouling properties applied in glycoprofiling of human serum samples. *Anal Chem* 85, 7324-7332.

61. Bertok, T., Sediva, A., Katrlik, J., Gemeiner, P., Mikula, M., Nosko, M. and Tkac, J. (2013) Label-free detection of glycoproteins by the lectin biosensor down to attomolar level using gold nanoparticles. *Talanta* 108, 11-18.

62. Yang, H., Li, Z., Wei, X., Huang, R., Qi, H., Gao, Q., Li, C. and Zhang, C. (2013) Detection and discrimination of alpha-fetoprotein with a label-free electrochemical impedance spectroscopy biosensor array based on lectin functionalized carbon nanotubes. *Talanta* 111, 62-68.

63. Vedala, H., Chen, Y., Cecioni, S., Imberty, A., Vidal, S. and Star, A. (2011) Nanoelectronic detection of lectin-carbohydrate interactions using carbon nanotubes. *Nano Lett* 11, 170-175.

64. Chen, Y.N., Vedala, H., Kotchey, G.P., Audfray, A., Cecioni, S., Imberty, A., Vidal, S. and Star, A. (2012) Electronic detection of lectins using carbohydrate-functionalized nanostructures: Graphene versus carbon nanotubes. *ACS Nano* 6, 760-770.

65. Zhang, G.J., Huang, M.J., Ang, J.J., Yao, Q.F. and Ning, Y. (2013) Label-free detection of carbohydrate-protein interactions using nanoscale field-effect transistor biosensors. *Anal Chem* 85, 4392-4397.

66. Huang, Y.-W., Wu, C.-S., Chuang, C.-K., Pang, S.-T., Pan, T.-M., Yang, Y.-S. and Ko, F.-H. (2013) Real-time and label-free detection of the prostate-specific antigen in human serum by a polycrystalline silicon nanowire field-effect transistor biosensor. *Anal Chem* 85, 7912-7918.

67. Ali, M., Nasir, S., Ramirez, P., Cervera, J., Mafe, S. and Ensinger, W. (2013) Carbohydrate-mediated biomolecular recognition and gating of synthetic ion channels. *J Phys Chem C* 117, 18234-18242.

68. Kruss, S., Hilmer, A.J., Zhang, J., Reuel, N.F., Mu, B. and Strano, M.S. (2013) Carbon nanotubes as optical biomedical sensors. *Advanced Drug Delivery Reviews* DOI: 10.1016/j.addr.2013.07.015.

69. Reuel, N.F., Grassbaugh, B., Kruss, S., Mundy, J.Z., Opel, C., Ogunniyi, A.O., Egodage, K., Wahl, R., Helk, B., Zhang, J., Kalcioglu, Z.I., Tvrdy, K., Bellisario, D.O., Mu, B., Blake, S.S., Van Vliet, K.J., Love, J.C., Wittrup, K.D. and Strano, M.S. (2013) Emergent properties of nanosensor arrays: Applications for monitoring igg affinity distributions, weakly affined hypermannosylation, and colony selection for biomanufacturing. *ACS Nano* 7, 7472-7482.

70. Bellapadrona, G., Tesler, A.B., Grunstein, D., Hossain, L.H., Kikkeri, R., Seeberger, P.H., Vaskevich, A. and Rubinstein, I. (2012) Optimization of localized surface plasmon resonance transducers for studying carbohydrate-protein interactions. *Anal Chem* 84, 232-240.

71. Jin, S., Cheng, Y., Reid, S., Li, M. and Wang, B. (2010) Carbohydrate recognition by boronolectins, small molecules, and lectins. *Med Res Rev* 30, 171-257.

72. Streicher, H. and Sharon, N. (2003) Recombinant plant lectins and their mutants. In: Yuan CL, Reiko TL (eds) Methods in enzymology, vol Volume 363. Academic Press, pp 47-77. doi:http://dx.doi.org/10.1016/S0076-6879(03)01043-7

73. Geisler, C. and Jarvis, D.L. (2011) Letter to the glyco-forum: Effective glycoanalysis with maackia amurensis lectins requires a clear understanding of their binding specificities. *Glycobiol* 21, 988-993.

74. Alava, T., Mann, J.A., Théodore, C., Benitez, J.J., Dichtel, W.R., Parpia, J.M. and Craighead, H.G. (2013) Control of the graphene–protein interface is required to preserve adsorbed protein function. *Anal Chem* 85, 2754-2759.

75. Hsu, K.-L., Gildersleeve, J.C. and Mahal, L.K. (2008) A simple strategy for the creation of a recombinant lectin microarray. *Mol BioSystems* 4, 654-662.

76. Propheter, D.C., Hsu, K.-L. and Mahal, L.K. (2010) Fabrication of an oriented lectin microarray. *ChemBioChem* 11, 1203-1207.

77. Propheter, D.C. and Mahal, L.K. (2011) Orientation of gst-tagged lectins via in situ surface modification to create an expanded lectin microarray for glycomic analysis. *Mole BioSystems* 7, 2114-2117.

78. Chen, M.-L., Adak, A.K., Yeh, N.-C., Yang, W.-B., Chuang, Y.-J., Wong, C.-H., Hwang, K.-C., Hwu, J.-R.R., Hsieh, S.-L. and Lin, C.-C. (2008) Fabrication of an oriented fc-fused lectin microarray through boronate formation. *Angew Chem Int Ed* 47, 8627-8630.

79. Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J. and Brent, R. (1996) Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature* 380, 548-550.

80. Mascini, M., Palchetti, I. and Tombelli, S. (2012) Nucleic acid and peptide aptamers: Fundamentals and bioanalytical aspects. *Angew Chem Int Ed* 51, 1316-1332.

81. Ruigrok, V.J.B., Levisson, M., Eppink, M.H.M., Smidt, H. and van der Oost, J. (2011) Alternative affinity tools: More attractive than antibodies? *Biochem J* 436, 1-13.

82. Ståhl, S., Kronqvist, N., Jonsson, A. and Löfblom, J. (2013) Affinity proteins and their generation. *J Chem Technol Biotechnol* 88, 25-38.

83. Woodman, R., Yeh, J.T.H., Laurenson, S. and Ferrigno, P.K. (2005) Design and validation of a neutral protein scaffold for the presentation of peptide aptamers. *J Mol Biol* 352, 1118-1133.

84. Gebauer, M. and Skerra, A. (2009) Engineered protein scaffolds as next-generation antibody therapeutics. *Curr Opin Chem Biol* 13, 245-255.

85. Davis, J.J., Tkac, J., Humphreys, R., Buxton, A.T., Lee, T.A. and Ko Ferrigno, P. (2009) Peptide aptamers in label-free protein detection: 2. Chemical optimization and detection of distinct protein isoforms. *Anal Chem* 81, 3314-3320.

86. Davis, J.J., Tkac, J., Laurenson, S. and Ferrigno, P.K. (2007) Peptide aptamers in label-free protein detection: 1. Characterization of the immobilized scaffold. *Anal Chem* 79, 1089-1096.

87. Stadler, L.K.J., Hoffmann, T., Tomlinson, D.C., Song, Q.F., Lee, T., Busby, M., Nyathi, Y., Gendra, E., Tiede, C., Flanagan, K., Cockell, S.J., Wipat, A., Harwood, C., Wagner, S.D., Knowles, M.A., Davis, J.J., Keegan, N. and Ferrigno, P.K. (2011) Structurefunction studies of an engineered scaffold protein derived from stefin a. Ii: Development and applications of the sqt variant. *Protein Eng Design Selection* 24, 751-763.

88. Koide, A., Gilbreth, R.N., Esaki, K., Tereshko, V. and Koide, S. (2007) High-affinity single-domain binding proteins with a binary-code interface. *Proc Natl Acad Sci* 104, 6632-6637.

89. Ellington, A.D. and Szostak, J.W. (1990) In vitro selection of rna molecules that bind specific ligands. *Nature* 346, 818-822.

90. Keefe, A.D., Pai, S. and Ellington, A. (2010) Aptamers as therapeutics. *Nat Rev Drug Discov* 9, 537-550.

91. Iliuk, A.B., Hu, L. and Tao, W.A. (2011) Aptamer in bioanalytical applications. *Anal Chem* 83, 4440-4452.

92. Liu, J., Cao, Z. and Lu, Y. (2009) Functional nucleic acid sensors. *Chemi Rev* 109, 1948-1998.

93. Tolle, F. and Mayer, G. (2013) Dressed for success - applying chemistry to modulate aptamer functionality. *Chem Sci* 4, 60-67.

94. Li, M., Lin, N., Huang, Z., Du, L., Altier, C., Fang, H. and Wang, B. (2008) Selecting aptamers for a glycoprotein through the incorporation of the boronic acid moiety. *J Am Chem Soc* 130, 12636-12638.

95. Vaught, J.D., Bock, C., Carter, J., Fitzwater, T., Otis, M., Schneider, D., Rolando, J., Waugh, S., Wilcox, S.K. and Eaton, B.E. (2010) Expanding the chemistry of DNA for in vitro selection. *J Am Chem Soc* 132, 4141-4151.

96. Eckstrum, K. and Bany, B. (2011) Tumor necrosis factor receptor subfamily 9 (tnfrsf9) gene is expressed in distinct cell populations in mouse uterus and conceptus during implantation period of pregnancy. *Cell Tissue Res* 344, 567-576.

97. Kimoto, M., Yamashige, R., Matsunaga, K.-i., Yokoyama, S. and Hirao, I. (2013) Generation of high-affinity DNA aptamers using an expanded genetic alphabet. *Nat Biotech* 31, 453-457.

98. Cao, Z., Partyka, K., McDonald, M., Brouhard, E., Hincapie, M., Brand, R.E., Hancock, W.S. and Haab, B.B. (2013) Modulation of glycan detection on specific glycoproteins by lectin multimerization. *Anal Chem* 85, 1689-1698.

99. Lu, Y.-W., Chien, C.-W., Lin, P.-C., Huang, L.-D., Chen, C.-Y., Wu, S.-W., Han, C.-L., Khoo, K.-H., Lin, C.-C. and Chen, Y.-J. (2013) Bad-lectins: Boronic acid-decorated lectins with enhanced binding affinity for the selective enrichment of glycoproteins. *Anal Chem* 85, 8268-8276.

100. McDonald, R.E., Hughes, D.J. and Davis, B.G. (2004) Modular control of lectin function: Redox-switchable agglutination. *Angew Chem Int Ed* 43, 3025-3029.

101. Eigen, M. and Schuster, P. (1978) The hypercycle. *Naturwissenschaften* 65, 341-369.

102. Ikehara, K., Omori, Y., Arai, R. and Hirose, A. (2002) A novel theory on the origin of the genetic code: A gnc-sns hypothesis. *J Mol Evol* 54, 530-538.

103. Ikehara, K. (2005) Possible steps to the emergence of life: The [GADV]-protein world hypothesis. *Chem Rec* 5, 107-118.

104. Di Giulio, M. (2008) An extension of the coevolution theory of the origin of the genetic code. *Biol Direct* 3, 1-21.

105. Lehmann, J., Cibils, M. and Libchaber, A. (2009) Emergence of a code in the polymerization of amino acids along RNA templates. *Plos One* 4.

106. Yarus, M., Widmann, J. and Knight, R. (2009) RNA–amino acid binding: A stereochemical era for the genetic code. *J Mol Evol* 69, 406-429.

107. Nahalka, J. (2011) Quantification of peptide bond types in human proteome indicates how DNA codons were assembled at prebiotic conditions. *J Proteom Bioinf* 4, 153-159.

108. Nahalka, J. (2012) Glycocodon theory—the first table of glycocodons. *J Theor Biol* 307, 193-204.

109. Wacker, M., Feldman, M.F., Callewaert, N., Kowarik, M., Clarke, B.R., Pohl, N.L., Hernandez, M., Vines, E.D., Valvano, M.A., Whitfield, C. and Aebi, M. (2006) Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc Natl Acad Sci* 103, 7088-7093.

110. Li, L., Woodward, R., Ding, Y., Liu, X.-w., Yi, W., Bhatt, V.S., Chen, M., Zhang, L.-w. and Wang, P.G. (2010) Overexpression and topology of bacterial oligosaccharyltransferase pglb. *Biochem Biophys Res Commun* 394, 1069-1074.

111. Schwarz, F., Huang, W., Li, C., Schulz, B.L., Lizak, C., Palumbo, A., Numao, S., Neri, D., Aebi, M. and Wang, L.-X. (2010) A combined method for producing homogeneous glycoproteins with eukaryotic n-glycosylation. *Nat Chem Biol* 6, 264-266.

Figure caption

Figure 1: A) A mechanical platform of detection based on an array of microcantilever biochips with a cantilever bent after a biorecognition took place. **B)** An electronic detection platform of analysis based on an electrochemical impedance spectroscopy (EIS) with an increase in overall resistance of a soluble redox probe, represented by a red arrow, to the interface after a biomolecular interaction.

Figure 2: A) A field-effect transistor (FET) sensing based on a changed conductivity of a single-walled carbon nanotube (SWCNT) positioned in between a source and a drain. **B)** An optical detection platform based on quenching of an intrinsic fluorescence of a SWCNT by Ni-tether employed for lectin immobilisation via His_6 tag. A fluorescence of a SWCNT is partly restored after biorecognition since Ni-tether is pushed away from the SWCNT surface.

Figure 3: A localised surface plasmon resonance (LSPR) employed for a label-free recognition based on a shift of reflected light as in case of traditional SPR technique.

Figure 4: A) A schematic representation of possibilities to control uniformity of glycan binding proteins with a random amine coupling (upper image), oriented immobilisation via introduced purification tag (image in the middle) and immobilisation of uniform lectin peptide aptamers differing only in peptide insert providing a biorecognition element; GBS – glycan binding site. **B)** Various ways for immobilisation of Fc-fused lectin on interfaces based on boronate affinity towards glycans present in Fc fragment (left), on affinity of protein G towards Fc fragment (middle) or a random amine coupling (right).

Figure 5: Peptide aptamers based on an affibody (58 AA, PDB code 1LP1, on left) or an DARPin (166 AA, PDB code 2BKK, on right) scaffold in complex with its analyte. A peptide aptamer is in both cases at the bottom part of a figure, while its analyte is above the peptide aptamer.

Figure 6: An RNA aptamer (purple) bound to its analyte peptide (white-magenta chain) (a PDB code 1EXY).

Figure 7: A glycocodon theory - G (glycine), A (alanine), V (valine), and D (aspartic acid) are elementary amino acids and the first primordial interactions between GAVD-peptides and sugars were evolutionary conserved and used in the glycocodons; the full set of glycocodons is proposed to call - the glycocode *(108)*. **A)** A distribution of GAVD-dipeptides in galactose

(Gal) and glucose (Glc) specific proteins *(108)*. In human Galectin-3 (a yellow protein structure, 1KJL), Gal is sensed by two overlapping glycocodons – NWG plus WGR that are derived from ancient GA – AG specific dipeptides; in perchloric acid-soluble protein from Pseudomonas syringae (a blue protein structure, 3K0T), Glc is sensed by the RAA glycocodon – ancient AA specific dipeptide can be today transformed to MF. **B)** A distribution of GAVD-dipeptides in mannose (Man) and fucose (Fuc) specific proteins *(108)*. In bacteriocin from Pseudomonas sp. complexed with Met-mannose (a cyan protein structure, 3M7J), Man is sensed by two overlapping glycocodons – QGD plus DGN – ancient GD - DG specific dipeptides can be transformed today to AY – YA dipeptides; in a PA-IIL lectin from P. aeruginosa (an orange protein structure, 2JDK), Fuc is sensed by the specific glycocodon NSS and by two GTQ and GTD glycocodons derived from ancient GA dipeptide specific for Gal (Fuc is actually 6-deoxy-L-galactose). **C)** A distribution of GAVD-dipeptides in N-acetylglucosamine (GlcNAc) and glucose (Glc) specific proteins *(108)*. In human L-ficolin (an assembly of two monomers is shown, 2J3O), GlcNAc is sensed by two overlapping glycocodons – RVG plus VGE – ancient VG specific dipeptide can be transformed today to FS dipeptide. **D)** A distribution of GAVD-dipeptides in N-acetylgalactosamine (GalNAc) and galactose (Gal) specific proteins *(108)*. In oligosaccharyltransferase from Campylobacter lari (3RCE), GalNAc (a first moiety of the oligosaccharide) is sensed by two overlapping glycocodons – EMI plus ITE derived from ancient VV and VA dipeptides specific for GalNAc. GalNAc is transferred to asparagines of the DQN glycocodon - DQ dipeptide is derived from ancient DD dipeptide specific for GalNAc.

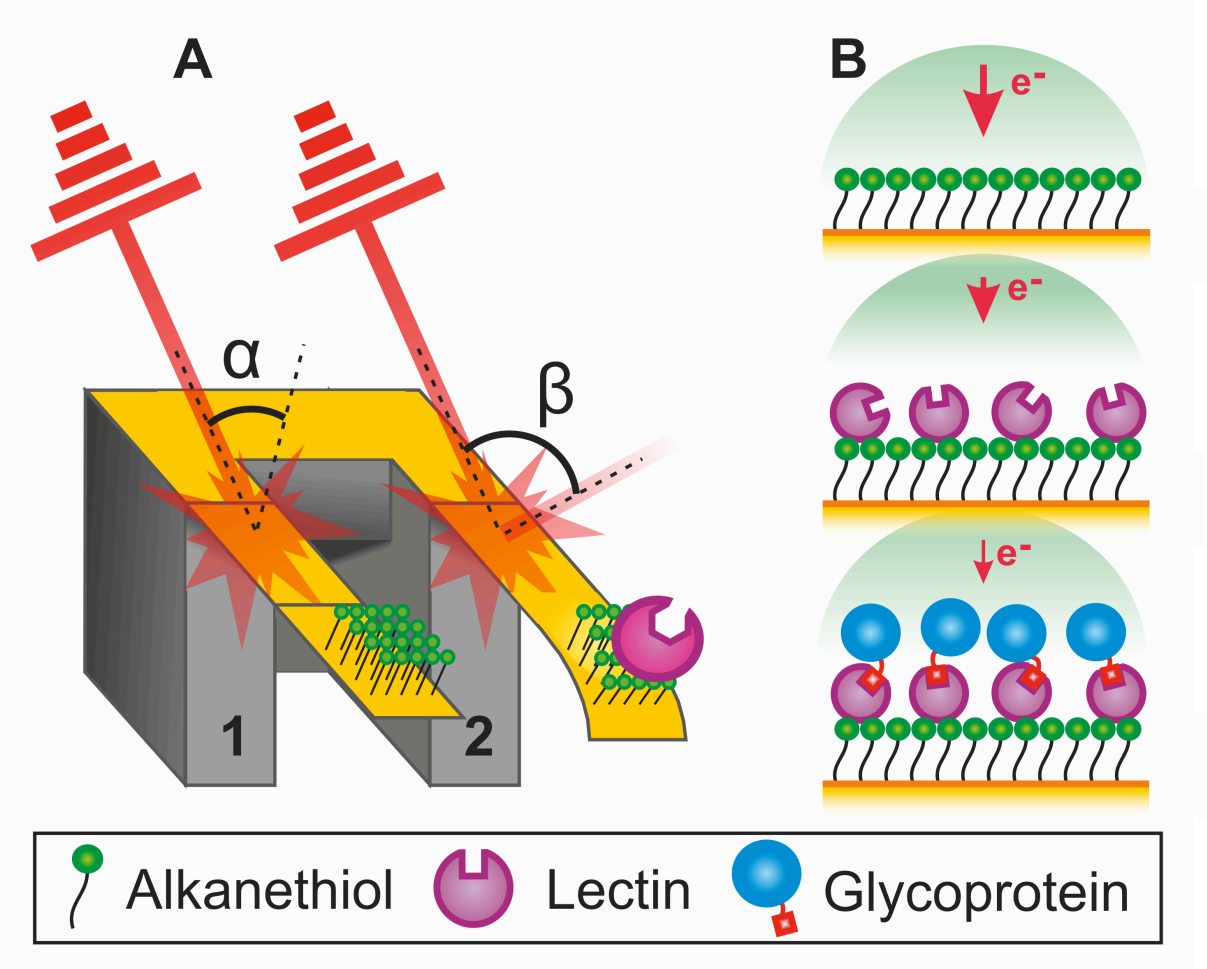

Figure 1

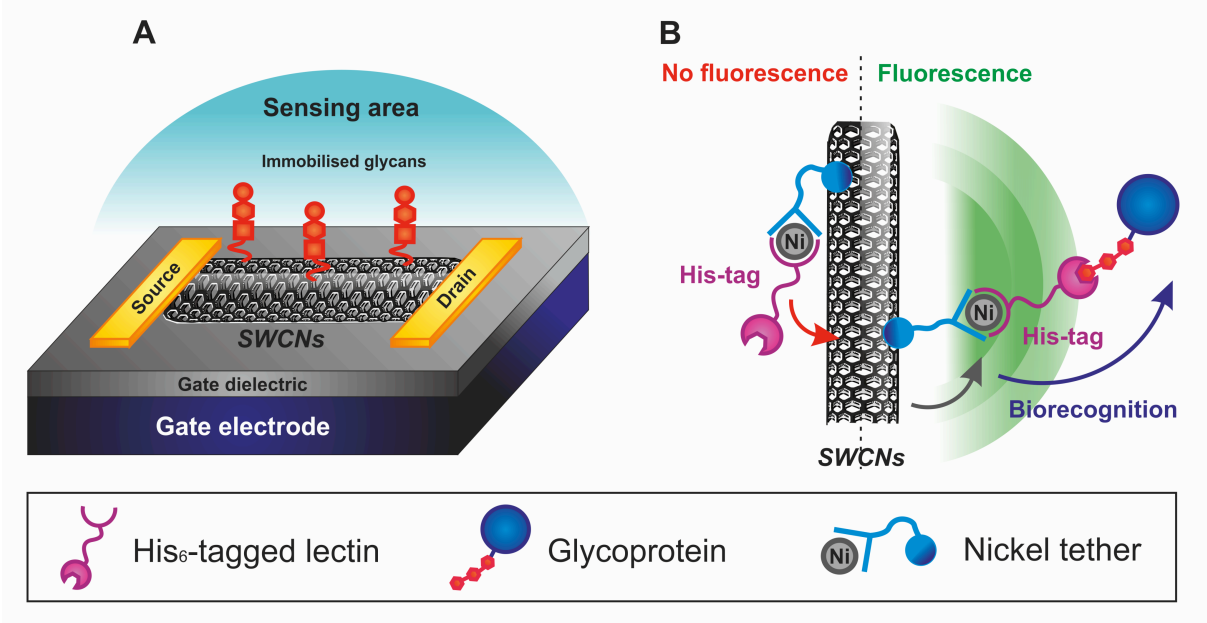

Figure 2

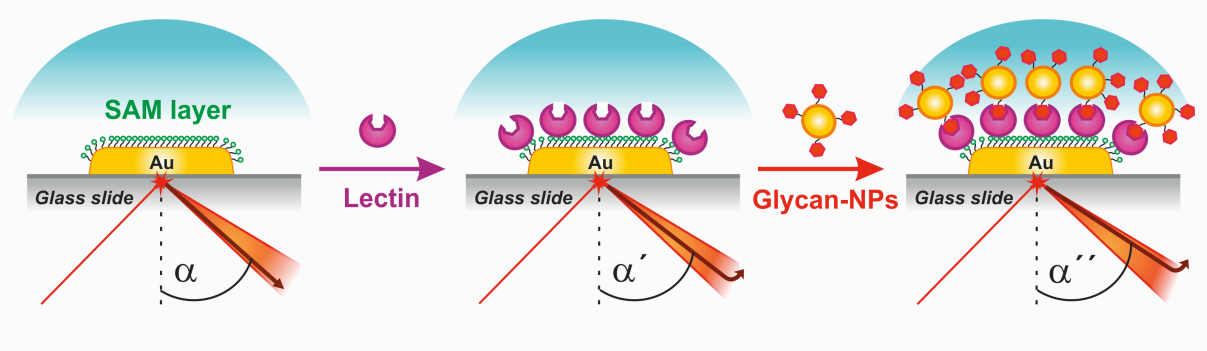

Figure 3

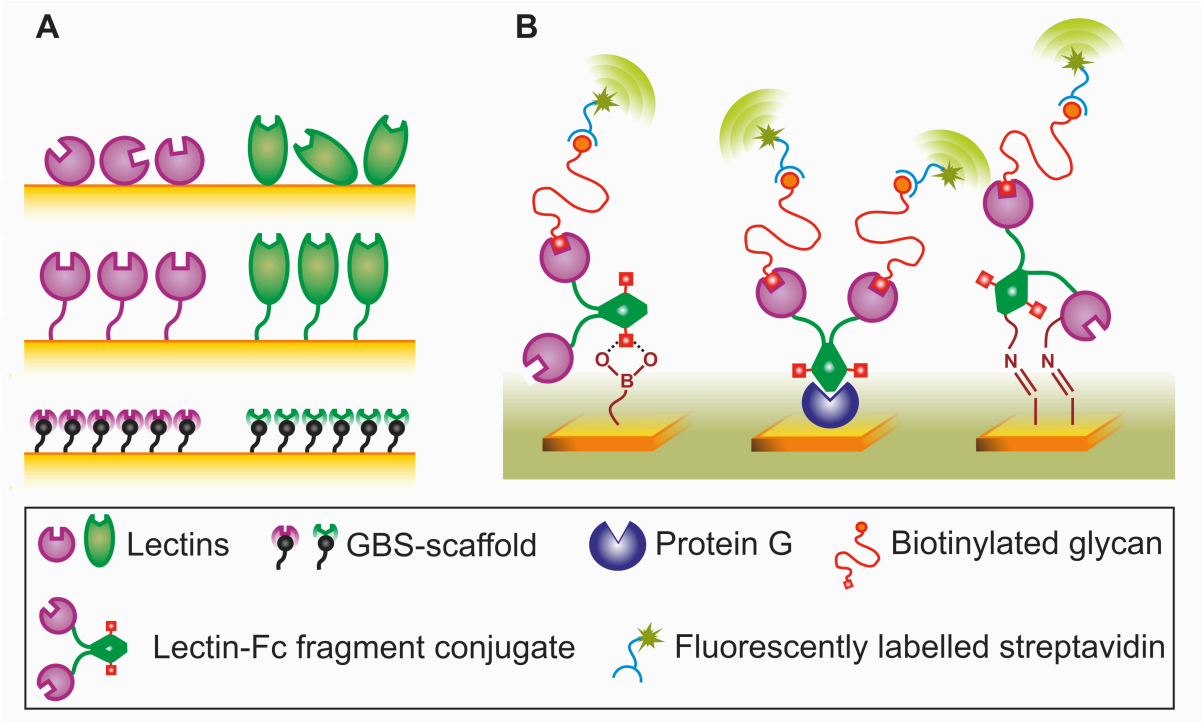

Figure 4

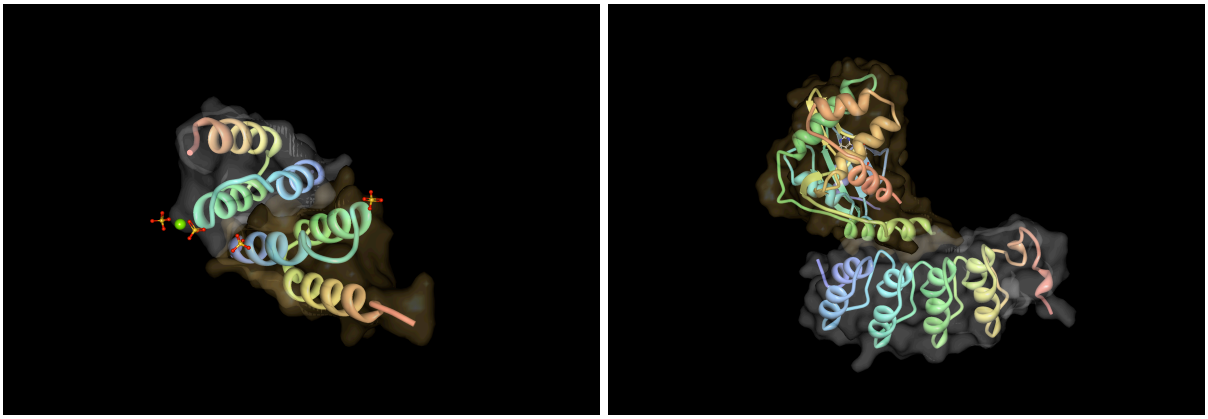

Figure 5

Figure 6

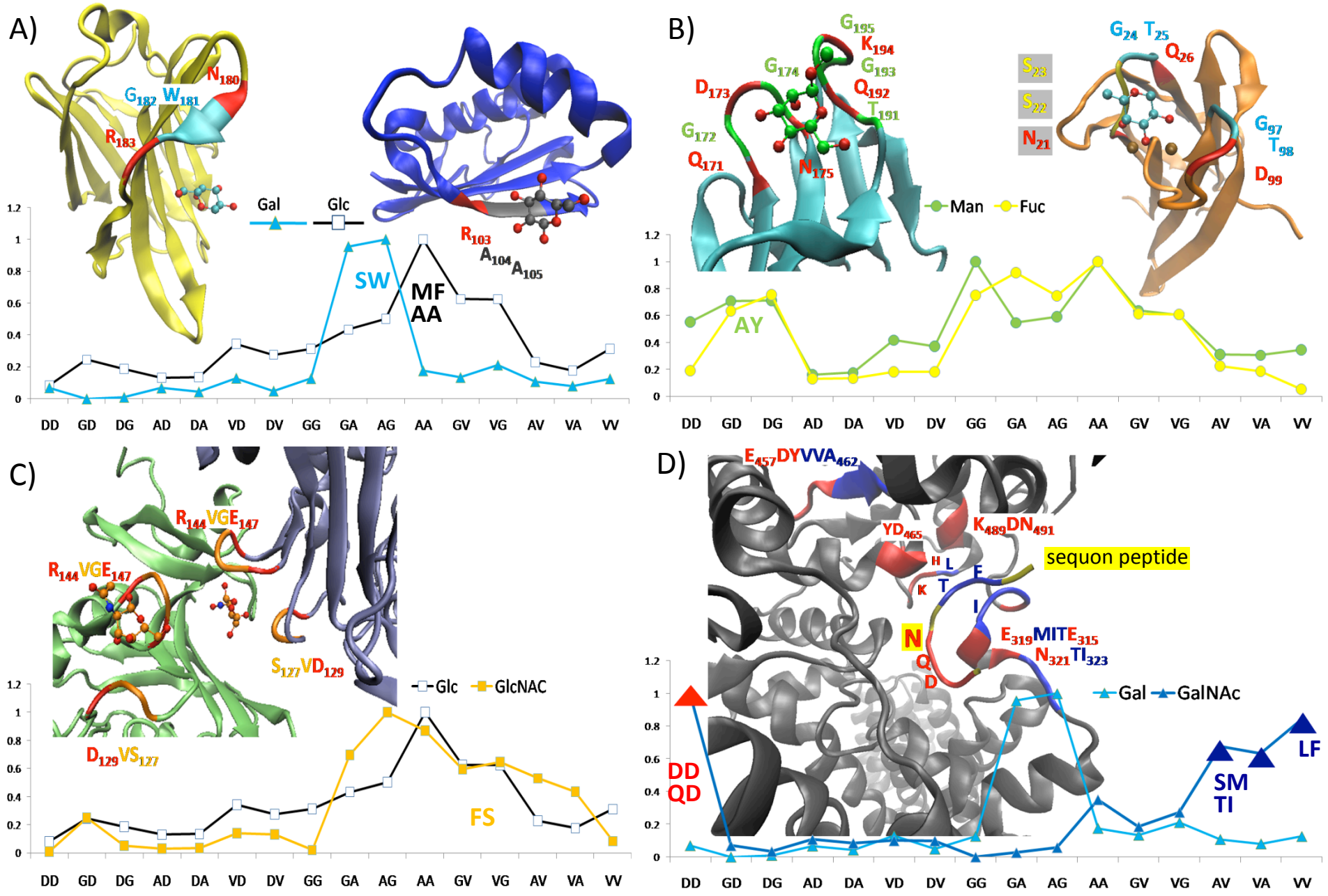
A)
N180
G182 W181
R183
R103
A104 A105
Gal
Glc
SW
MF
AA
DD GD DG AD DA VD DV GG GA AG AA GV VG AV VA VV
B)
G195
K194
G193
G174
D173
Q192
T191
G172
Q171
N175
S23
S22
N21
G24 T25
Q26
G97
T98
D99
Man
Fuc
AY
C)
R144VGE147
S127VD129
D129VS127
Glc
GlcNAC
FS
D)
E457DYVVA462
YD465
K489DN491
sequon peptide
H L
K T
F
I
N
Q
D
E319MITE315
N321TI323
Gal
GalNAc
DD
QD
SM
TI
LF

Figure 7