Applications of a Mass-spectrometric Method for the N-glycosylation Analysis of Immunoglobulin G molecules

Doctoral Thesis

by

DI Johannes Stadlmann

Submitted in partial fulfilment of the requirements for the degree of

„Doktor der Universität für Bodenkultur, Wien“

Carried out under the supervision of

A.o. Univ.-Prof. DI Dr. nat. techn. Friedrich Altmann

at the

Department of Chemistry,

of

The University of Natural Resources and Applied Life Sciences, Vienna.

Vienna, 2009

ABSTRACT

The work presented in this cumulative thesis reflects the efforts in the development of a robust high-throughput analytical platform for the quantitative and site-specific glyco-proteomic characterisation of IgG molecules, and highlights some of its application in the development of various novel, particularly plant-based, expression systems.

Taken together, the herein selected studies shall provide a comparative overview on the N-glycosylation properties of different expression systems and expression strategies, particularly with respect to the ongoing efforts in the development of large-scale production platforms for therapeutically relevant recombinant IgG molecules in plants.

Zusammenfassung

Die hier vorliegende, kumulative Dissertation spiegelt die Entwicklung einer robusten Hochdurchsatz-Analytik-Plattform zur quantitativen und “site”-spezifisch-glyco-proteomischen Charakterisierung von IgG Molekülen wider, und beschreibt einige Anwendungen dieser Methode im Bereich der Entwicklung von neuen, vor allem pflanzlicher, Expressions- Systeme.

In ihrer Gesamtheit sollen die hier zusammengestellten Publikationen eine vergleichende Übersicht der N-glycosylierungs-Eigenschaften verschiedener Expressionssysteme und Expressionsstrategien, besonders im Hinblick auf die aktuellen Entwicklungen in Richtung großtechnischer, pflanzlicher Produktionsplattformen zur Herstellung von therapeutischen, rekombinanten IgG Molekülen, liefern.

We’ll make mistakes and then
Life is the art of learning to live with it
Through time

(Róisín Murphy)

"*Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides*".
Stadlmann, J; Pabst, M; Kolarich, D; Kunert, R and Altmann, F (2008).
Proteomics 8(14): 2858-71.

"*Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern*".
Schahs, M; Strasser, R; Stadlmann, J; Kunert, R; Rademacher, T and Steinkellner, H (2007). Plant Biotechnol J 5(5): 657-63.

"*Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure*".
Strasser, R; Stadlmann, J; Schahs, M; Stiegler, G; Quendler, H; Mach, L; Glossl, J; Weterings, K; Pabst, M and Steinkellner, H (2008). Plant Biotechnol J 6(4): 392-402.

"*Cost-effective production of a vaginal protein microbicide to prevent HIV transmission*".
Ramessar, K; Rademacher, T; Sack, M; Stadlmann, J; Platis, D; Stiegler, G; Labrou, N; Altmann, F; Ma, J; Stoger, E; Capell, T and Christou, P (2008).
Proc Natl Acad Sci U S A 105(10): 3727-32.

"*Recombinant antibody 2G12 produced in maize endosperm efficiently neutralizes HIV-1 and contains predominantly single-GlcNAc N-glycans*".
Rademacher, T; Sack, M; Arcalis, E; Stadlmann, J; Balzer, S; Altmann, F; Quendler, H; Stiegler, G; Kunert, R; Fischer, R and Stoger, E (2008).
Plant Biotechnol J 6(2): 189-201.

"*Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic Arabidopsis seeds*".
Van Droogenbroeck, B; Cao, J; Stadlmann, J; Altmann, F; Colanesi, S; Hillmer, S; Robinson, DG; Van Lerberge, E; Terryn, N; Van Montagu, M; Liang, M; Depicker, A and De Jaeger, G (2007). Proc Natl Acad Sci U S A 104(4): 1430-5.

REFERENCES

APPENDIX

Curriculum vitae

Comprehensive List of Publications

ACKNOWLEDGMENTS

"No man is an island, entire of itself;
Every man is a piece of the continent, a part of the main."
(John Donne; 1572 – 31st March 1631)

I wish to thank A.o. Univ.-Prof. DI Dr. Friedrich Altmann, who excellently supervised the realization of this work. Throughout the years of fruitful collaboration, he has been a continuous source of valuable advice, encouragement, motivation and inspiration; not only in scientific matters.

Mass-spectrometry is a technology-driven field of science and greatly relies on delicate technical equipment. Therefore, I thank DI Dr. Daniel Kolarich for having me introduced to, and familiarized with, the technical aspects in operating the mass-spectrometric equipment.

As analytics is hardly an end in itself, it would not have been without the trustful provision of samples by a large number of collaborators, that this work has gained its eventual scientific significance. Most prominently, however, I would like to thank A.o. Univ.-Prof. Dr. Herta Steinkellner and DI Dr. Richard Strasser (as well as their co-workers at the BOKU), and A.o. Univ.-Prof. Dr. Eva Stöger (as well as her associates at the RWTH in Aachen).

Furthemore, I thank my co-workers in the laboratory, namely Dr. Martin Pabst, Mag. Josephine Grass, Dr. Renaud Leonard, Mag. Richard Fischl, Dr. Chunsheng Jin, and Ing. Thomas Dalik, for having created a delightful and friendly work environment. In particular, I would like to thank Ing. Karin Polacsek, who greatly contributed to this work by excellent technical assistance and her cheerful nature.

I dedicate this work to my family. I express my deepest gratitude to my parents, grand-parents and my brother, for their encouragement, understanding, patience and selfless support during my studies. The biggest debt of gratitude, however, I owe to my *fiancée*, Mlle Lisa Teckentrup, for her loving, caring and unremitting support throughout the years.

The work presented in this thesis was funded by "Pharma-Planta", an EU 6th Framework program of the European Commission (www.pharma-planta.org).

INTRODUCTION

The biotechnological production of protein-based therapeutics is one of the fastest growing sectors of the pharmaceutical industry [1]. The majority of these therapeutic proteins are subject to post-translational modifications (PTMs) of some form, which can affect the therapeutic properties of the eventual biopharmaceutical product [2]. Consequently, the identification and characterization of such PTMs of therapeutic proteins, as well as understanding their biological significance, is of increasing importance to the fields of pharmaceutical research and biotechnological process development.

PTMs result from biochemical reactions which lead to chemical and/or structural alterations of the nascent polypeptide chain, after its translation. PTMs comprise the chemical modification of amino-acids, the formation of disulfide bonds, proteolytic cleavage (e.g. of signal-peptides), and the addition of functional groups to the protein. Many of these modifications are essential for the correct folding, processing and secretion of proteins, their biochemical activity, their in-vivo stability, as well as their immunological properties.[1]

Glycosylation, the covalent attachment of single monosaccharides or large carbohydrate groups to a proteins polypeptide backbone represents one of the most abundant PTMs of eukaryotic proteins; and thus also on therapeutically relevant proteins [2].
With respect to the chemical nature of the linkage between the carbohydrates and the polypeptides amino-acid side-chains two major classes of protein glycosylation are differentiated: O-linked glycosylation and N-linked glycosylation.
In O-linked glycosylation, an initial, single monosaccharide is attached via an α-O-glycosidic linkage to the hydroxyl-group of the amino-acid side-chains of serine, threonin or hydroxyproline. This inital biosynthetic reaction is located to the Golgi-apparatus and can be followed by the subsequent action of various glycosyltransferases, which then lead to a step-wise elongation of O-glycan structures.
By contrast, in N-linked glycosylation a preliminary oligosaccharide-precursor, consisting of 14 monosaccharides, becomes attached via a β-N-glycosidic bond to the amide-group of the amino-acid asparagine when being member of the N-glycosylation amino-acid-“sequon”: asparagine, followed by any amino-acid, but proline, which is then followed by serine, threonine or –rather rarely- cysteine.
The initial transfer-reaction of the oligosaccharide precursor from dolichol-phosphate to the nascent polypeptide takes place in the lumen of the ER and is then followed by the concerted action of glycosyltransferases and glycosidases along the N-glycoproteins secretory pathway from the ER, via the Golgi-apparatus, to its eventual destination. [3-5]
Although the biosynthesis and the transfer of the N-glycosylation specific oligosaccharide-precursor to the nascent polypeptide are both highly conserved throughout the eukaryotic kingdom, later biosynthetic steps -particularly those located to the Golgi-apparatus- are less conserved and sometimes even class- or genus-specific. Intensive research in the fαield of glycobiology, however, provided the detailed characterization of the N-glycosylation patzways of most eukaryotic model species, which have been extensively reviewed in [3-8].

Even though all secreted N-glycoproteins are subject to the same Golgi-residing glycosylation-machinery of a given cell-type, substantially different sets of N-glycan

structures are linked to the various different N-glycoproteins produced by these cells [9]. For example, while the recombinant production of human erythropoietin in CHO cells commonly yields highly complex, tri- and tetra-antennary, N-acetyl-neuraminic acid terminated N-glycans [10], recombinant IgG molecules derived from the same cell line exhibit comparably less complex, bi-antennary structures, which are virtually devoid of N-acetyl-neuraminic acid residues [11].
Similarly, N-glycan structures attached to different N-glycosylation sites of the very same N-glycoprotein were found to differ substantially. For example, the recombinant IgG1 molecule Erbitux bears two N-glycosylation sites, of which one exhibits the IgG-typical, complex type, bi-antennary structures. The second N-glycosylation site, however, exhibits highly complex, tri- and tetra-antennary N-glycans with substantial amounts of terminal N-acetyl-neuraminic acid residues. [11, 12]
The processing of N-glycan structures in a particular expression host thus co-depends on the general structural properties of the respective N-glycoproteins, as well as the spatial conformation of the individual N-glycosylation sites. Additionally, however, numerous studies highlighted the influence of production parameters (e.g. cell culture medium composition, protein production levels) on the N-glycosylation properties of various expression-systems, and confirmed their potential in the generation of N-glycoprotein heterogeneity [10, 13-15].

Functionally, N-glycosylation is crucial to the initial folding-processes of secreted proteins. In the course of protein-folding in ER, N-glycans provide important structural features for the interaction of the nascent polypeptide-chain with the ER-residing chaperones, such as calnexin (CNX) and calreticulin (CRT), as well as for the subsequent protein-trafficking machinery, which catalyses the transport of correctly folded proteins from ER to the Golgi-apparatus, or the degradation of incorrectly folded proteins (e.g. via ER-associated-degradation-pathways; ERAD). [16, 17]
Similarly, N-glycans affect the interactions of N-glycoproteins outside of the cell. As N-glycans greatly contribute to the global, spatial conformations of N-glycoproteins (e.g. Fc-region of IgG molecules; see below) and may also directly serve as structural epitopes, they enable and/or modulate highly complex biochemical processes, such as protein targeting, in vivo stability, or various reactions of the immune system. [3-5]

As a consequence, particularly in the advent of large-scale manufacturing recombinant N-glycoproteins, N-glycan structures were sought to be identical to those found on their respective native counter-parts. Later, extensive studies on the specific impact of the individual N-glycan structures on glycoproteins, however, helped to identify both, minimal structural requirements, as well as particular oligo-saccharide compositions which exhibit beneficial biochemical properties in a number of therapeutic applications of N-glycoproteins (e.g. highly sialylated rhEPO, non-fucosylated IgG, Glucocerebrosidase). Clearly, these findings stimulated research and development in the fields of N-glycoengineering.

Immunoglobulin Gamma (IgG):

Immunoglobulins (Ig) are multi-functional N-glycoproteins, linking cellular and humoral immunity. This group of molecules is able to specifically bind antigens (e.g. pathogens) and to form immune complexes, which activate various effector mechanisms of the immune system and initiate the removal and/or destruction of these antigens.[18]

In human, most Ig molecules (except for IgM and IgA, which associate in a pentameric and a dimeric conformation, respectively) prevail in a monomeric configuration. Monomeric Ig molecules consist of two identical short polypeptide chains of approx. 25 kDa (i.e. light-chains; LC); and two identical long chains of approx. 50 – 65 kDa (i.e. heavy-chains; HC), each; which are linked via multiple inter-chain disulfide bonds. Both types of the Ig polypeptide chains bear a series of similar, yet not identical, amino-acid sequence motifs of approx. 110 residues, each. These structural motifs, commonly referred to as Ig domains, give rise to individually folded, structurally robust, β-barrel shaped protein domains, which are composed of two disulfide-bonded, anti-parallel β-sheets. [18-20]

The LC polypeptide chains of all Ig molecules consist of two Ig domains: an N-terminal domain, which exhibits a high degree of sequence variability (i.e. variable domain of the LC; VL); and a C-terminal domain, which is formed by either of two highly conserved amino-acid sequence variants (i.e. constant domain of the LC; CL) referred to as κ or λ.
Similarly, Ig HCs made up of one Ig domain with high sequence variability at the N-terminus (i.e. variable domain of the HC; VH), and a limited set of highly conserved amino-acid sequence variants towards the C-termini of the polypeptide-chains. In contrast to the architecture of the Ig LCs, however, the conserved parts of the Ig HCs exhibit three to four constant Ig domains (namely CH1, CH2, CH3, and CH4) and –with the notable exception of IgM and IgE- one structurally unordered “hinge” region, located between the CH1 and the CH2 domain. It is the amino-acid sequence of these highly conserved Ig HC domains which define the individual Ig molecules’ structurally and functionally distinct idiotype: IgG1, IgG2, IgG3, IgG4 [see Fig.1, Fig.2 and Fig.4], IgM, IgA1, IgA2, IgE and IgD. [18, 21]

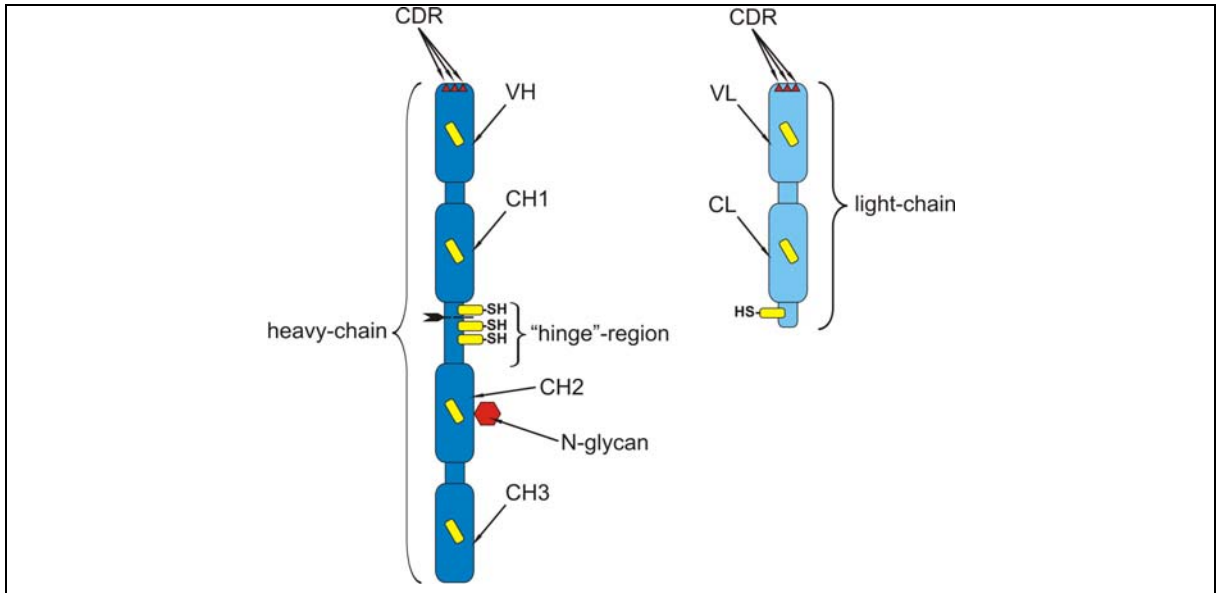

Figure 1. Schematic representation of the Ig domains of the individual polypeptide chains of IgG1 molecule. disulfide bonds are highlighted in yellow; complementarity-determining regions (CDR) are highlighted by red triangles. The papain-cleavage site is indicated by a bold arrow.

```
            1       10        20        30        40        50        60        70        80        90
            |--------+---------+---------+---------+---------+---------+---------+---------+---------|
IGHG1_LCS   ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPS
IGHG2_LCS   ASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVXSSXXGTQTYTCNVDHKPS
IGHG4_LCS   ASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPS
IGHG3_LCS   ASTKGPSVFPLAPCSRSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQXSGLYSLSSVVTVPSSXXGTQTYTCNVNHKPS
Consensus   ASTKGPSVFPLAP.S.STS..TAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQ.SGLYSLSSVVTV.SS..GT.TY.CNV#HKPS

            91      100       110       120       130       140       150       160       170       180
            |--------+---------+---------+---------+---------+---------+---------+---------+---------|
IGHG1_LCS   NTKVDKXVE----------------------------------------------PKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDT
IGHG2_LCS   NTKVDKTVE----------------------------------------------RKCCVE---CPPCPAPP-VAGPSVFLFPPKPKDT
IGHG4_LCS   NTKVDKRVE----------------------------------------------SKYGPP---CPSCPAPEFLGGPSVFLFPPKPKDT
IGHG3_LCS   NTKVDKRVELKTPLGDTTHTCPRCPEPKSCDTPPPCPRCPEPKSCDTPPPCPRCPEPKSCDTPPPCPRCPAPELLGGPSVFLFPPKPKDT
Consensus   NTKVDK.VE..........................................K.......CP.CPAP....GPSVFLFPPKPKDT

            181     190       200       210       220       230       240       250       260       270
            |--------+---------+---------+---------+---------+---------+---------+---------+---------|
IGHG1_LCS   LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
IGHG2_LCS   LMISRTPEVTCVVVDVSHEDPEVQFNWYVDGXEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTK
IGHG4_LCS   LMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVXHQDWLNGKEYKCKVSNKGLPSSIEKTISKAK
IGHG3_LCS   LMISRTPEVTCVVVDVSHEDPEVQFKWYVDGVEVHNAKTKXXEEQXNSTFRVVSVLTVXHQDWLNGKEYKCKVSNKALPAPIEKTISKXK
Consensus   LMISRTPEVTCVVVDVS.EDPEV.F.WYVDG.EVHNAKTK..EEQ.NST%RVVSVLTV.HQDWLNGKEYKCKVSNK.LP..IEKTISK.K

            271     280       290       300       310       320       330       340       350       360
            |--------+---------+---------+---------+---------+---------+---------+---------+---------|
IGHG1_LCS   GQPREPQVYTLPPSRXEXTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHE
IGHG2_LCS   GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHE
IGHG4_LCS   GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSXLTVDKSRWQEGNVFSCSVMHE
IGHG3_LCS   GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAXEWESXGQPXNNYXTTPPXLDSDGSFFLYSXLTVDKSRWQXGNXFSCSVMHE
Consensus   GQPREPQVYTLPPS..E.TKNQVSLTCLVKGFYPSDIA.EWES.GQP.NNY.TTPP.LDSDGSFFLYS.LTVDKSRWQ.GN.FSCSVMHE

            361     370   377
            |--------+-----|
IGHG1_LCS   ALHNHYTQKSLSLSPGK
IGHG2_LCS   ALHNHYTQKSLSLSPGK
IGHG4_LCS   ALHNHYTQKSLSLSLGK
IGHG3_LCS   ALHNXXTQKSLSLSPGK
Consensus   ALHN..TQKSLSLS.GK
```

Figure 2. Amino-acid sequence alignment of the "least common" CH domains of IgG1, IgG2, IgG3 and IgG4. The reference amino-acid sequences for the individual IgG subclasses and their respective allelic variants were obtained from the IMGT-database (imgt.cines.fr). [Stadlmann J, Pabst M, Weber A, Anderle H, Schwarz HP, Altmann F; manuscript submitted]

Although globally stabilized by various inter- and intra-chain disulfide bonds, the formation of the eventual functional three-dimensional structures of Ig molecules is predominantly driven by strong hydrophobic interactions between the individual Ig domains. (This is, for example, highlighted by the observation that also the Ig domains of Fv-fragments, which are devoid of any intermolecular covalent linkage, are able to associate in order to form functionally active, antigen binding molecules.) Within a fully assembled Ig gamma molecule (IgG), the Ig domains of the VL and the VH pair in a non-covalent manner. This close interaction of the variable protein regions of the Ig molecule results in the spatially confined arrangement of their complementarity-determining regions (i.e. CDR) and thus provides the structural basis for the formation of the respective molecule-specific antigen binding "paratope". Similar to the VL and VH domains, the CH1 and the CL (which are additionally interconnected by a disulfide bond) and the two CH3 domains, respectively, engage with each other in a non-covalent manner, providing highly robust structural elements, which further stabilize the IgG molecule. The two spatially opposing CH2 domains, however, do not pair as their hydrophobic patches are shielded by one highly conserved, complex-type, typically bi-antennary and predominantly core-fucosylated N-glycan, each [see Fig.3 and Fig.5]. The interaction of the IgG HCs hinge regions is essentially provided by a number of inter-HC disulfide bonds. [18, 22, 23]

Eventually, the highly ordered association of the various Ig domains of two IgG HCs and two IgG LCs results in the formation of three compact, functionally distinct, globular protein regions of similar size, which are connected through a rather flexible hinge region. As early studies on IgG revealed, these three globular protein regions can be released by partial proteolytic degradation with papain, and give rise to two functionally distinct IgG-derived fractions. One of these fractions exhibits the antigen-binding properties (i.e. Fab-fragment) and consists of the IgG LC domains VL and CL domains, linked –again, via disulfide-bonds- to the VH and the CH1 domains of the IgG HC. The second fraction (i.e. Fc-fragment), however, comprises of the disulfide-linked hinge-regions, the two CH2 and the two rather rigidly associated CH3 domains of the IgG molecule. This highly conserved IgG protein region (i.e. constant region) mediates the molecules effector functions in the immune system. [18] [see Fig.3]

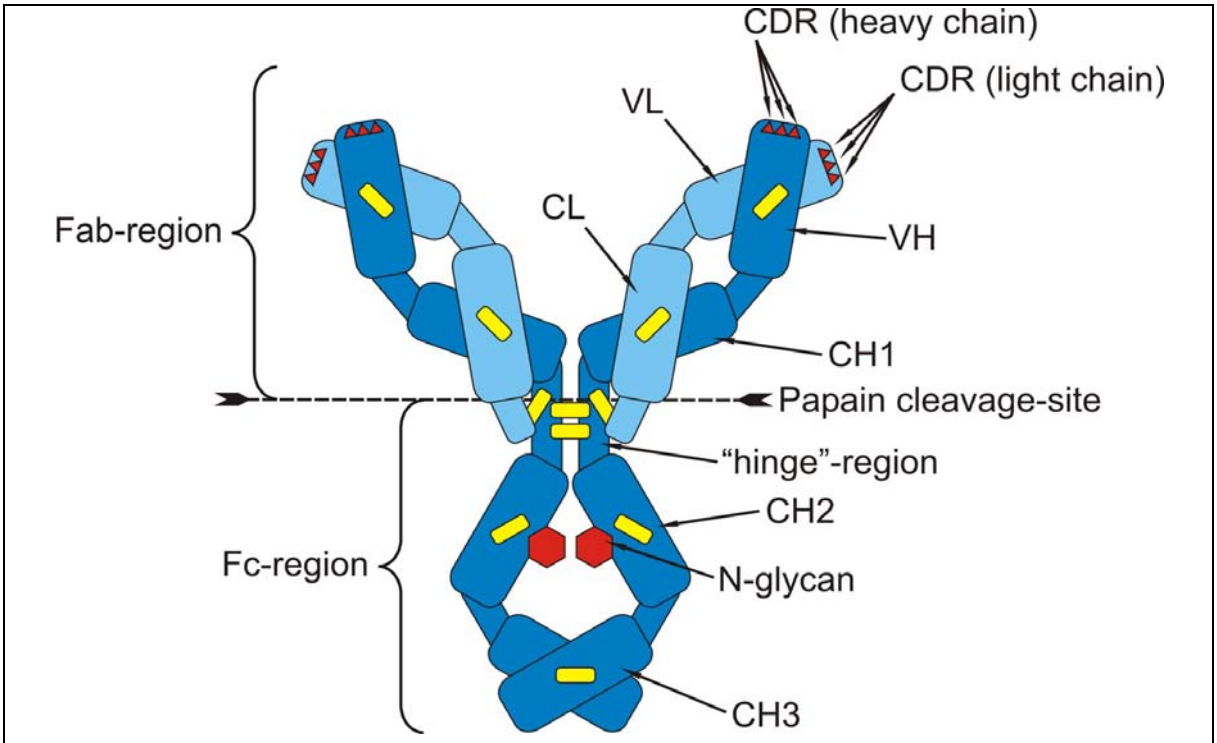

Figure 3. Schematic representation of a fully assembled, monomeric IgG1 molecule. The molecules heavy-chains are indicated in dark-blue; the light-chains are indicated in light-blue; inter-chain disulfide bonds are highlighted in yellow; complementarity-determining regions (CDR) are highlighted by red triangles.

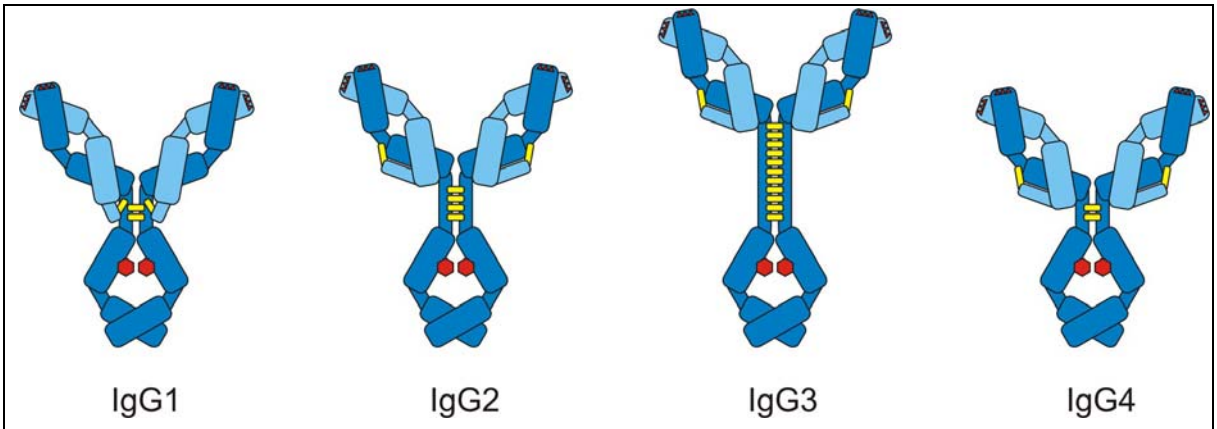

Figure 4. Schematic representation of the fully assembled, monomeric IgG molecules of all four human IgG subclasses. The molecules heavy-chains are indicated in dark-blue; the light-chains are indicated in light-blue; the inter-chain disulfide bonds are highlighted in yellow; complementarity-determining regions (CDR) are highlighted by red triangles.

The principle ligands through which the constant region of IgG molecules interacts with the immune system are the component of complement (i.e. C1 complex) and the members of FcγReceptor super-family. [18, 19, 24-26]
The interplay of the IgG constant region with the C1q (a constituent of the C1 complex), requires the formation of IgG/antigen immune-complexes and activates the classical complement cascade of the innate immune system. A lack of N-linked glycans at the conserved N-glycosylation site of the IgG CH2-domain, however, abolishes this interaction.
The engagement of Fc-fragment with members of the FcγReceptor super-family, on the other hand, initiates various immune-cellular reactions, including antibody-dependant-cellular-cytotoxicity (ADCC), phagocytosis, oxidative burst, and the release of inflammatory components. [18, 19, 24-26]
FcγReceptors are type I trans-membrane glycoproteins, which are expressed on various hematopoietic cells (B-cells, T-cells, macrophages, eosinophils, neutrophils, natural killer-cells, lymphocytes, etc.). The cellular immune responses triggered by this group of proteins can either occur upon activating or inhibitory signals, which are mediated by tyrosine based activation or inhibitory motifs (ITAM and ITIM), respectively, of the various FcγR molecules. In human, the FcγR super-family consists of three large groups: FcγRI (i.e. CD64), FcγRII (i.e. CD32), and FcγRIII (i.e. CD16). Each of these groups is further sub-classified in order to account for multiple FcγR isoforms, alleles and polymorphisms. Most interestingly, each of these receptors was found to exhibit a particular recognition profile for the four human IgG subclasses. However, due to the highly heterogeneous nature of the FcγR super-family members, as well as their apparent different IgG subclass recognition profiles, the data obtained from rather motley experiments can merely provide rough estimates on their individual binding properties and affinities. A more detailed review on the interactions of the respective FcγR molecules with the four human IgG subclasses shall thus be omitted in this work, but may be readily found elsewhere; e.g. [19].
By contrast, it is well established that N-glycosylation of the IgG Fc-region is not only crucial for C1 complex interaction, but also for the efficient activation of all FcγReceptor-mediated response pathways of the immune system [18, 23-25, 27-29]. Thus, the oligosaccharides N-linked to the IgG molecules Fc-regions provide a structural feature, which is essential for its major biological activities in linking humoral and cellular immune mechanisms.

The two N-glycans, common to all IgG molecules, are linked to asparagines residues within the two opposing CH2 domains of the two HCs and are both oriented into the spatial cleft between. Although no direct contact between the carbohydrate groups and the IgG polypeptide-chains could be established, extensive structural investigations of human IgG Fc-region glycoforms yet revealed subtle and reciprocal conformational influences. Most interestingly, however, despite the subtility of the structural alterations, which are imposed onto the IgG polypeptide chain by changes in the IgG N-glycan composition, some were found to significantly impact on the various interactions of the Fc-region. These highly intriguing studies have been expertly reviewed in e.g. [18, 23-25, 27-29].

The oligosaccharides released from human polyclonal IgG typically exhibit complex-type, bi-antennary N-glycan structures. Their structural heterogeneity arises from varying peripheral extensions of a basic “core” hepta-saccharide by α-1,6-linked fucose, β-1,4-linked galactose residues, and/or β-1,4-linked (“bisecting”) N-acetyl-glucosamine. Additionally, a minor

fraction of the galactosylated human IgG N-glycans is terminated by α-2,6-linked N-acetyl-neuraminic acid residues [see Fig.5].[18]

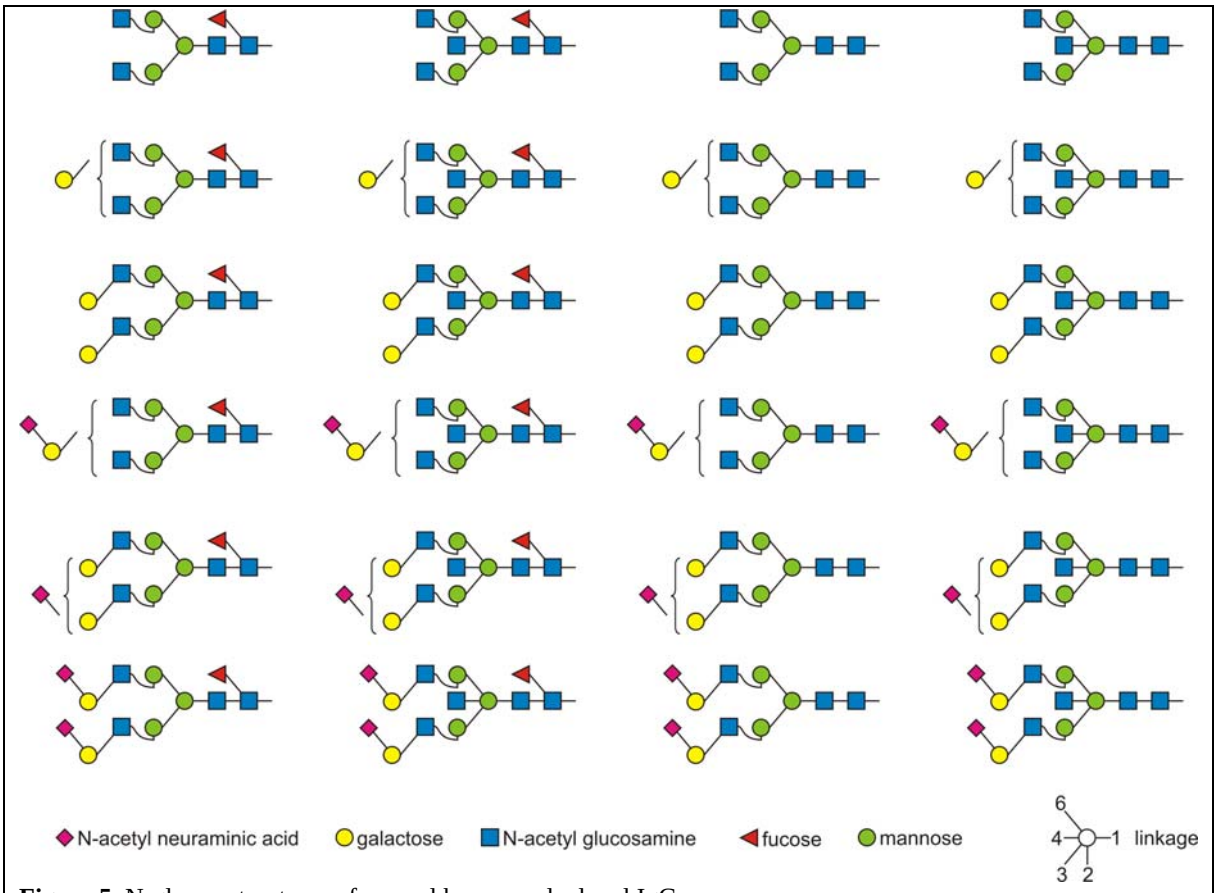

Figure 5. N-glycan structures of normal human polyclonal IgG.

In human, the addition of alpha-1,6-linked ("core") fucose is the most frequently observed extension found on human IgG N-glycans (i.e. > 90% of all glycoforms). Extension by β-1,4-linked galactose-residues, normally present on approx. 70% of all IgG glycoforms, can occur on both N-glycan antennae: the α-1,6-linked and/or the α-1,3-linked GlcNAc residue. The preference for the addition of the first galactose-residue to either arm, which is considered to – at least partially- reflect their respective accessibility for the β-1,4-galactosyltransferase, was reported to differ between molecules of the different IgG subclasses: Whereas IgG1 and IgG4 molecules were reported to exhibit more galactose on the alpha-1,6-arm (approx. 60%), N-glycans released from IgG2 were found to be preferentially galactoslyated on the alpha-1,3-arm. However, the preferential addition of galactose-residues to either antenna of IgG3-Fc-linked N-glycans was found to vary with the respective IgG3 HC-allotype. With respect to the galactosylation it is notable that IgG-"hypogalactosylation" is vaguely associated with a number of inflammatory states, particularly in autoimmune diseases (e.g. rheumatoid arthritis; RA). Bi-secting GlcNAc residues are typically found on approx. 15% of all N-glycans released from polyclonal IgG, and were –most curiously- reported to be also subject to further elongation by β-1,4-linked galactose, yet in infinitesimal amounts [30].[18]

Only a minority of polyclonal IgG Fc-linked N-glycans are mono-sialylated. Nevertheless, these structures exhibit a pronounced bias for the addition of N-acetyl-neuraminic acid to the α-1,3-linked antenna, reflecting, both, the preferential addition of galactose to this very branch in the -usually most abundant- IgG-subclass IgG1, as well as the respective sialyltransferases substrate-preference.[27]

N-glycans are typically linked to asparagine residues being constituent of the general N-glycosylation sequons NXS, NXT or -sometimes- NXC (X being any amino-acid but proline). Due to this rather loose sequence prerequisite, the random amino-acid sequences in the highly variable regions of the IgG VL- and VH-domains may also give rise to potential N-glycosylation sites with in the Fab-region. Studies on the N-glycosylation of Fab-fragments revealed that, indeed, minor fractions (i.e. approx. 15%) of polyclonal IgG also bear N-glycans attached to their variable regions [31]. Interestingly, the N-glycans released from Fab-fragment differ from those N-linked to the IgG CH2-domain. While the N-glycans linked to the IgG Fc-region are factually devoid of di-sialylated, as well as bisected sialylated, such structures are readily detected in N-glycan preparations derived from Fab-fragment. Also, apart from a generally higher degree of galactosylation, the preference of galactose addition to either arm of the “core”-heptasaccharide is not observed on N-glycans released from the variable regions of IgG molecules [Stadlmann J, Pabst M, Weber A, Anderle H, Schwarz HP, Altmann F; manuscript submitted]. These contrasting N-glycan profiles of Fc- and Fab-fragment may be attributed to the steric constraints prevailing in the Fc-region, and the –conversely- greater accessibility of Fab-linked oligosaccharide for glycosyltransferases.

Engineering of IgG and IgG-like molecules:

As was revealed by early studies on IgG, two functionally distinct protein regions can be released by partial proteolytic degradation of IgG. The detailed functional and biochemical characterization of these IgG-derived fragments, in combination with contemporaneous advances in the fields of genetic engineering, paved the way for the recombinant production of functionally optimized therapeutic IgG and IgG-like molecules.[32]
The use of recombinantly produced IgG for the passive immunization of humans requires these molecules to effectively bind the respective antigen, and to be sufficiently integrated into the human immune system. While antigen-binding properties of IgG molecules are stringent criteria in the initial paratope selection and are often even improved in subsequent affinity-maturation processes, the integration of IgG molecules into the human immune system greatly depends on the nature of their constant protein domains. As a consequence, the IgGs derived from non-human B-cells (e.g. mice) need to be adapted to the requirements of the human immune system. Technologies for this purpose comprise the chimerzation and the humanization of IgG molecules. Chimeric IgG molecules are generated by genetically replacing the DNA sequences encoding for the IgG constant domains (i.e. CH1, hinge, CH2, CH3) of the host organism, with those of human origin. In the course of the humanization of IgG molecules, however, only the complementarity-determining regions (i.e. CDR; the minimal structural prerequisite for antigen-binding) of the non-human molecule are incorporated into otherwise entirely human IgG sequences.

In applications which exclusively require the antigen-binding properties of IgG molecules (e.g. neutralising antibodies), Fc-mediated functions may often be idle or even undesirable. The consequential engineering of IgG-like molecules devoid of the constant protein regions, gave rise to a multitude of exclusively antigen-binding, IgG-derived molecule formats, such as Fab-, Fv-, single chain Fv (scFv)-, or single V-type domain molecules; as well as various multivalent designs thereof [see Fig.6]. Not only do such IgG-derived molecules exhibit certain functional advantages over intact IgG molecules in a number of applications (e.g. increased tissue penetration properties), they are also successfully produced in less complex, more economic expression systems (e.g. bacterial expression systems) [32].
More recently, however, yet another novel format of IgG-like molecules, the so-called Fcab-molecule, was designed. These largely artificial molecules comprise the inter-chain linked hinge-region, two native CH2-domains (bearing an N-glycosylation-site, each) and two engineered CH3-domains. In contrast to typical IgG molecules, where the CDRs are located to the variable domains of the HC and LC (i.e. Fab-fragment), the CDRs of Fcab molecules are incorporated into their CH3 domains [see Fig.6].

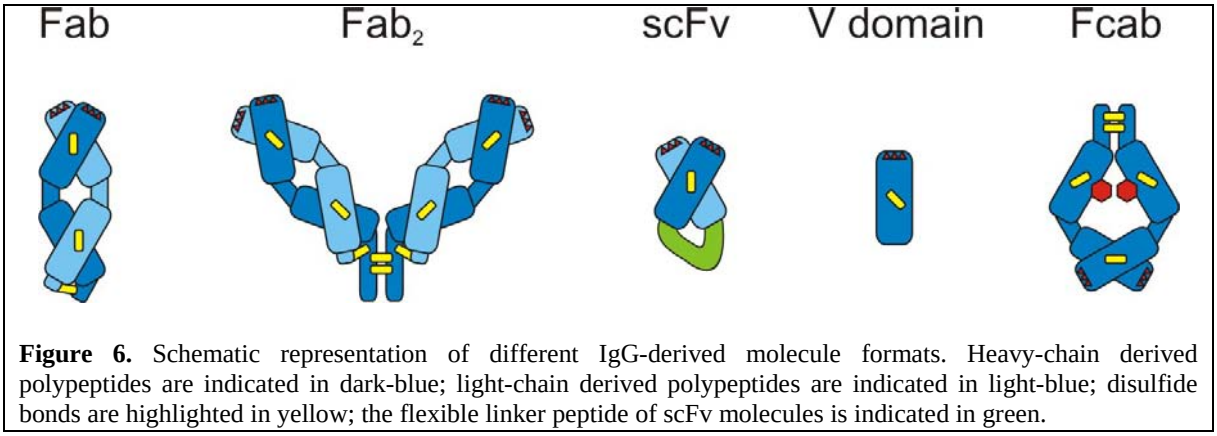

Figure 6. Schematic representation of different IgG-derived molecule formats. Heavy-chain derived polypeptides are indicated in dark-blue; light-chain derived polypeptides are indicated in light-blue; disulfide bonds are highlighted in yellow; the flexible linker peptide of scFv molecules is indicated in green.

Production-platforms for mAbs and recombinant IgG molecules:

The expression of soluble, properly folded, correctly assembled, fully active and intact IgG molecules requires the protein-processing machinery of eukaryotic cells (e.g. folding, disulfide-bond formation, proteolytic processing, glycosylation, secretion).[33]
Initially, two major expression-strategies were used for the production of mono-specific IgG-molecules: hybridoma cell technology, in which B-cells readily producing the respective IgG-molecule (i.e. monoclonal antibody; mAb) are literally fused (i.e. hybridized) with highly productive and immortal myeloma (i.e. B-cell tumor) cells [34]; as well as mammalian recombinant production systems, in which only the genetic information for the production of the respective molecule (i.e. recombinant antibody; rAb) is transferred into the host production cells. Particularly with respect to the production of therapeutic IgG molecules, the recombinant approach clearly provides a greater degree of freedom in engineering both the molecule's primary structures (e.g. chimerization, humanization), and the production cell lines. Additionally, recent advances in the optimization of such large-scale production platforms for recombinant protein expression gave rise to state-of-the-art, serum-free, fed-batch production systems yielding up to 6 g IgG / L culture supernatant [35]. As these yields

of additionally more general and flexible expression platforms exceed those from hybridoma technology (which is per se restricted to IgG production) approx. 30 fold, it is almost needless to mention that current large-scale manufacturing of therapeutic IgG is exclusively based on recombinant cell lines.
To date, host cell lines most frequently used for the production of recombinant IgG molecules comprise of murine myeloma cell lines (e.g. Sp2/0, NSO/1) and murine non-lymphoid cell-lines (e.g. CHO, PC12); with CHO cell lines being the most versatile and thus most widely used system for the production of biopharmaceuticals.

More recently, however, various alternative eukaryotic expression systems have gained considerable interest for the large-scale production of biopharmaceuticals, as they offer a number of advantages in production costs, process scalability and product-safety related issues.[36, 37]
For example, the use of the methylotrophic yeast *Pichia pastoris* as host organism for large-scale recombinant protein production -particularly in high-yielding fed-batch fermentation processes-, offers several benefits over mammalian expression systems, such as comparably simple and thus inexpensive cultivation requirements; rapid growth; a relatively high degree of purity of the secreted recombinant proteins in the culture supernatant; as well as good bio-safety.[38, 39]

Also, insect cell expression systems (e.g. *Spodoptera frugiperda*, *Drosophila melangogaster*) have been successfully employed for the production of recombinant IgG molecules. To this end, insect cells can either be stably transformed or used in transient, baculovirus mediated production schemes. Despite the rather cost and technology-intensive cultivation- and growth-requirements of these cell lines; their general bio-safety properties make these systems of particular interest.[6, 40, 41]

However, cell-based production systems inherently depend on fermenters, as well as skilled personnel to operate them. Alternatively, “molecular pharming” aims at the production of therapeutically relevant recombinant proteins in higher, multi-cellular eukaryotic organisms, such as animals and plants.[42-44]
Although transgenic animals would be the “fermenter-free” alternative to current mammalian cell-culture production strategies, their development is lengthy, and the degree of flexibility as well as process scalability are generally judged to be rather low. Plant-based production schemes, on the other hand, out-compete traditional, cell-based expression systems and transgenic animals by offering high flexibility, excellent scalability, and low production costs. As a consequence, the development of plant-based expression platforms has attracted great interest over the past decade.[43, 45]

Although all these expression systems are generally able to perform all essential protein-processing steps necessary for the secretion of functional full-size IgG molecules, the production of any therapeutic (glyco-) protein in non-human expression systems potentially generates immunogenic structures. Such immunogenic structures are typically associated with different PTM properties of the respective expression system. Since the generic PTM properties (e.g. N-glycosylation, proteolytic processing of signal peptides) of non-human expression systems differ -often substantially- from those of humans, considerable adaptations

of various PTM pathways are crucial. Consequently, many recent efforts in the development of new manufacturing technologies, as well as in the engineering of existing production platforms, particularly focused on product quality by actively controlling the PTM properties of the respective host cells.

Post-translational modifications of IgG molecules:

Despite the uniformity of the genetic information encoding for a recombinantly produced glycoprotein, a number of subsequent enzymatic and non-enzymatic PTMs (which may occur intra- and/or extra-cellularly) give rise to the inherently heterogenous nature of these molecules. Cell culturing conditions, such as temperature, growth rate, or media composition, clearly influence the systems productivity. Interestingly, however, they have also been repeatedly shown to alter the extent of PTMs of the respective recombinant protein produced. Similarly, various down-stream processing steps, product formulations and storage conditions, have been demonstrated to change the chemical composition of biopharmaceuticals. As such variations in the PTM profiles of recombinant proteins may impact on their biological function and/or biocompatibility, the occurrence and variation of PTMs needs to be closely monitored and tightly controlled.[46]

The structural integrity of IgG molecules greatly depends on the correct formation of inter- and intra-chain disulfide bonds. The detection of un-paired, free sulfhydryl groups of cystein residues is thus considered to be indicative for the incomplete formation of disulfide-bonds during protein synthesis, and thus for structural inconsistencies. Furthermore, free sulfhydryl-groups have been reported to catalyze disulfide-bond scrambling, which may – particularly under mild alkaline conditions- result in the formation of multi-molecular aggregates (or even precipitates). [46, 47]

The most C-terminal amino acid all human IgG HCs encode for is lysine. On natural human IgG HCs, however, this amino-acid residue is efficiently removed after protein synthesis by the action of basic carboxypeptidases (e.g. carboxypeptidase B). By contrast, recombinant IgG production platforms (e.g. myeloma cell lines), frequently fail to efficiently remove the C-terminal lysine. Although the biological significance of non-processed C-terminal lysine residues of IgG HCs is still under investigation [48], their contribution to gross product heterogeneity is generally regarded as unfavourable.

The most N-terminal amino-acid of all correctly processed human IgG HCs, on the other hand, is either glutamine or glutamate. In natural human IgG, however, most of these N-terminal amino-acid residues occur in a cyclized form, as so-called pyro-glutamate. Non-human expression systems have been shown to be capable of catalysing identical modifications. However, the degree, to which this enzymatic modification is performed, greatly depends on cell culturing conditions. Furthermore, the degree to which the cyclized N-terminal amino-acid variants occur in the eventual product depends on various down-stream processing steps (e.g. ion-exchange chromatography).
In addition to the enzymatic conversion of glutamine and glutamate to pyro-glutamate, the spontaneous formation of pyro-glutamate, particularly under rather harsh conditions, has been

reported. However, the biological relevance of this modification has also not yet been fully understood.[46, 49]

The non-enzymatic conversion of asparagine-residues to iso-aspartate and aspartate is a common modification of proteins, which is most frequently found at sites in the polypeptide-chain, where asparagines are followed by glycine-residues. Similarly, although at kinetically much lower rates, also glutamine residues may undergo such chemical reactions, which result in the formation of glutamate. This being non-enzymatic reactions, “deamidation” may occur at any stage of the life time of a recombinant IgG product.[46]

Other, structurally more significant changes to the polypeptide chains often arise from proteolytic processing and degradation of IgG molecules. One of the most commonly observed PTM in this respect is the fragmentation of intact IgG molecules by proteolytic degradation. Since the hinge-region of IgG is the least well structured protein-region of these molecules, it does provide the most susceptible site for proteolytic cleavage. Proteases, such as papain, plasmin, cathepsin L, or Lys-C, are well known to actively cleave different peptide bonds in this region. Other types of proteases, including pepsin, are also able to cleave peptide bonds in the CH2 and CH3 domains. As the presence of proteolytic IgG fragments may substantially impact on product efficacy, their occurrence clearly raises a number of safety concerns.[46]

The most common and also functionally most important PTM of IgG molecules, however, is glycosylation. In contrast to rather rare N- and O-linked glycan structures attached to amino-acids constituent of the variable domains, the oligosaccharides linked to the conserved N-glycosylation site at the CH2-domains have been clearly shown to significantly influence the biological properties of IgG molecules. The following section shall thus provide a brief overview on the implications of N-glycosylation for the production of recombinant IgG for therapeutic applications.

N-glycosylation engineering of production-platforms for recombinant IgG:

N-glycosylation may directly and/or indirectly impact on various biological properties of IgG molecules. The direct interactions of N-glycans with various lectins or lectin-like proteins can have a pronounced impact on the molecules pharmacokinetics, their potency, and their immunogenic properties. In this respect, particularly GlcNAc- and mannose-terminating N-glycan structures, which have been reported to increase the clearance rate of therapeutic IgG molecules from the serum, are considered to be detrimental [50]. Furthermore, such structures potentially interact with glycan-array binding molecules, such as mannan-binding lectin (MBL), the cellular mannose-receptor (MR) or the dendritic cell specific intercellular adhesion molecule DC-SIGN, and may thus alter the therapeutic properties of IgG molecules. For example, the interaction of terminal GlcNAc-residues of IgG molecules with MBL, a functional homologue of C1q, is able to activate the lectin pathway (rather than the classical pathway) of the complement system via the MBL-associated-serine-protease-complex (MASP1 to MASP3), and may thereby contribute to inflammatory reactions [51]. By contrast, the binding of such IgG molecules to C1q was found to be reduced, and resulted in less

classical CDC reactivity of these molecules, in vitro. More strikingly, however, MR and DC-SIGN are both C-type lectins, which bind to mannose-terminating oligosaccharides and are expressed on macrophages and dendritic cells (N.B.: MR is also expressed on endothelial cells). These receptors are thus hypothesized to catalyse the uptake of IgG with mannose-terminating N-glycan structures by the antigen-presenting dendritic cells, and might thereby provoke autoimmune reactions against IgG molecules [18]. Other adverse reactions directly linked to N-glycan structure may arise from non-human and thus potentially immunogenic oligosaccharide structures. For example, the presence of α-galactose has been reported to induce the generation of anti-α-galactose IgE molecules, which may trigger severe hypersensitity reactions [52].
Nevertheless, also a number of indirect and reciprocal contributions of N-glycans to the quaternary structure of IgG molecules were found to be essential for fundamental effector activities of IgG, and to provide potential for further optimizing such functions. While the mere presence of N-glycans of the IgG Fc-region is fundamental to almost all of the relevant interactions of this protein region with the immune system, also the effects of subtle structural features of the respective oligosaccharides on the binding affinities of IgG molecules to the individual members of the FcγR super family have been demonstrated [53]. Most importantly in this regard, was the observation that IgG molecules (of all subclasses), which are devoid of the otherwise typical α-1,6-linked core-fucose residue on their N-glycans, have a significantly higher binding affinity for the human FcγRIIIa and are thus able to trigger ADCC more efficiently [54]. Inversely, the presence of α-2,6-linked sialic acid residues was reported to lower the binding affinity of IgG molecules for human FcγRIIIa, and thus to reduced ADCC activity [55].

Clearly, such findings stimulated a number of N-glycosylation engineering efforts aiming at the reproducible production of IgG molecules, which exhibit defined, homogenous and non-immunogenic N-glycan structures, providing enhanced and optimized the effector-functions. To this end, a number of general strategies, including the optimization of cell culturing conditions, the generation of novel host cell lines and/or the in vitro (i.e. enzymatic) modification of N-glycan structures, were adopted. Cell-culturing conditions have repeatedly been reported to significantly affect e.g. the degree of galactosylation, sialylation, or the abundance of oligo-mannosidic structures on recombinant IgG molecules from murine production systems [13-15]. Similarly, the use of glycosidase- or glycosyltransferase inhibitors has been demonstrated to provide means for the modulation of the N-glycosylation properties of well-established production cell lines [56]. Both approaches, would thus allow their straightforward integration into currently approved production schemes, without the need to establish and/or approve new production cell lines. The potential application of such approaches, however, is intrinsically limited by the respective host cells native N-glycosylation repertoire, as well as the availability of suitable inhibitor molecules.
By contrast, more profound and more effective adaptations of the N-glycosylation pattern typically requires the generation of novel mutant cell lines by means of genetic engineering, and -economically most importantly- entails the establishment, development and/or approval of the respective new production processes and cell lines. This approach has recently been applied to the re-engineering of the N-glycosylation pathways of currently used and well-established murine cell lines (see below), as well as in the more far-reaching developments of

novel (i.e. hitherto not approved) eukaryotic expression platforms (e.g. yeasts, plants, insect cells).

As technologically relevant examples for contemporary N-glyco engineering efforts, various strategies for the production of non-fucosylated, recombinant IgG molecules shall be briefly presented. Owing to their prominent role in current biopharmaceutical manufacturing processes, the engineering of the respective N-glycosylation pathways of murine cell lines has attracted considerable (commercial) interest. Typically, the N-glycans of recombinant IgG molecules derived from mammalian host cells exhibit a high degree of core-fucosylation (>80%) [57]. Aiming for a decrease in fucosylation (>90%), a number of murine expression systems have been developed, namely the CHO cell line variant Lec13 [54], the rat hybridoma cell line YB2/0 [58], a α-1,6-fucosyltransferase (FUT8) knock-out CHO cell line variant [59], a GDP-mannose 4,6-dehydratase (GMD) knock-out CHO cell line variant [60], as well as a FUT8/GMD double knock-out CHO cell line variant [61], and a β-1,4-N-acetylglucosaminyltransferase III (GnTIII; from rat) expressing CHO line [62, 63]. While the low fucosylation levels of N-glycans exhibited by the Lec13 and the GMD knock-out cell lines are ultimately due to a lack of the donor substrate GDP-fucose, in α-1,6-fucosyltransferase knock-down or knock-out lines (i.e. YB2/0 and FUT8 knock-out CHO line, respectively) the transfer of fucose to the N-glycan structure is (at least partially) impaired [57]. Interestingly, a comparative study showed that only the “true” knock-out lines (FUT8 [57], and GMD [60]) were able to produce substantial amounts of non-fucosylated IgG molecules. The mode of action responsible for the reduction in fucosylation of IgG molecules obtained from GnTIII transformed CHO cells, however, has been attributed to the competitive nature of the GnTIII and the FUT8 reaction. As the addition of β-1,4-N-acetylglucosamine (i.e. bisecting GlcNAc) to the hepta-saccharide core structure forecloses the transfer of α-1,6-linked fucose, the constitutive over-expression of GnTIII in these cell lines yields predominantly bisected, non-fucosylated N-glycan structures [58].
In addition to genetic strategies for the engineering of mammalian N-glycosylation pathways, an alternative, rather curious approach, supplementing a potent α-mannosidase I inhibitor (i.e. kifunensine) to murine cell cultures, was devised [56]. As the inhibition of α-mannosidase I restricts the processing of ER-derived oligo-mannose structures (i.e. Man9), and thus all further structural processing towards complex-type N-glycans (including fucosylation), the IgG molecules derived from such expression systems exhibit oligo-mannose type structures, exclusively; being glycoforms which are commonly associated with the rapid clearance of biopharmaceuticals from the bloodstream [50].
After all, similar N-glyco-engineering approaches also seem feasible in order to generate α-1,3-galactose- or N-glycoly-neuraminic acid-deficient cell lines, and will most certainly enter the stage of large-scale manufacturing of second generation protein-based pharmaceuticals.

Most noteworthy, however, these considerable efforts in engineering the N-glycosylation pathways of mammalian expression systems not only reflect the general attractiveness of glyco-engineered cell lines to the biopharmaceutical industry, but also highlight the increasingly restrictive patent landscape in this field and the consequential interest in other eukaryotic expression systems. Although such alternative non-mammalian expression systems have not yet reached the level of large-scale manufacturing of therapeutic proteins, they still

provide a greater degree of freedom to operate in terms of intellectual property related issues, on top of their beneficial bio safety and production properties.
The most promising future production platforms of this kind currently developed include *Pichia pastoris* [39], a number of insect cell lines (e.g. *Spodoptera frugiperda, Drosophila melangogaster*)[6, 41] and various plant-based systems (e.g. *Physcomitrella patens* [64], *Arabidopsis thaliana* [33], *Nicotiana* ssp. [65], *Medicago* ssp. [66] , *Lemna minor* [67], *Zea mays* [68, 69]). A common obstacle which precludes the application of these systems in the production of biopharmaceuticals arises from their native N-glycosylation pattern. In contrast to the N-glycan structures generated by non-engineered mammalian host cells, which closely resemble those of human, the native N-glycans derived from non-mammalian expression systems typically bear substantial amounts of unfavourable (e.g. oligo-mannose type) or even adverse N-glycan structures (e.g. β-1,2-linked xylose and/or α-1,3-linked fucose). As a consequence, already the early endeavours towards the "humanization" of these expression platforms embraced N-glyco-engineering as an integral part; and have generated substantial insight into the N-glycosylation machinery of the respective systems since.
Today, the N-glycosylation pathways of a number of non-mammalian expression platforms have been successfully "humanized". For example, Pichia pastoris has reportedly been engineered in order to provide sialylated, non-fucosylated, bi-antennary complex-type structures [38, 70]. Also, *Arabidopsis thaliana, Physcomitrella patens* and *Lemna minor* mutants, which completely lack β-1,2- xylosyltranferase and α-1,3- fucosyltransferase [33, 67, 71-73]; as well as *Nicotiana tabaccum* and *Nicotiana benthamiana* plants, which exhibit virtually no β-1,2-xylosyltranferase and α-1,3- fucosyltransferase activity (i.e. RNAi-lines) [65], have recently been generated and are now poised for their full humanization by the addition of terminal sialic acid residues.

Last but not least, also the *in vitro* modification of N-glycan structures of recombinant IgG molecules has been proven feasible. Most impressively, Warnock D., et al. [74] reported the complete *in vitro* galactosylation of 98% of all neutral glycan structures from 1 kg human polyclonal IgG, within 48 hours. Other IgG N-glycan modifications performed *in vitro*, include the addition of β-1,4-N-acetylglucosamine to more than 80% of all N-glycans [75] , and the combined addition of β-1,4-galactose and α-2,6-sialic acid, yielding 90% of sialylated N-glycans [76].

N-glycosylation Analysis of IgG molecules:

A wide range of methods for the analysis of the N-glycosylation of glycoproteins has been developed and was recently reviewed by [77]. Most of these methods and strategies have been also been successfully applied to IgG molecules. The most commonly employed approaches for the N-glycosylation analysis of IgG molecules were recently reviewed by [78].

The multitude of currently employed methods for N-glycosylation analysis can be classified according to their general analytical focus.
A first group of analytical strategies aims at the analysis of the N-glycosylation of intact N-glycoproteins. Analytical approaches of this group include the rather general detection of protein-linked carbohydrate-constituents (e.g. orcin), the comparative assessment of the

electrophoretic mobility of different protein glycoforms (e.g. IEF, upon de-glycosylation with various exo- and endoglycosidases), the identification of N-glycan structure motifs by their (specific) interaction with various lectins or lectin-like proteins (e.g. *Sambuccus nigra* lectin, Concanavalin A, anti-HRP-IgG) or their susceptibility to various glycosidases, the mass-spectrometric (MS) analysis of intact glyco-proteins; as well as various combinations thereof.
Key to all of the analytical techniques of this group, however, is purification of the respective N-glycoprotein and separation of its glycoforms [77]. The purification of IgG molecules is greatly facilitated by the availability and the ease of high-affinity purification methods based on the staphylococcal proteins A and G. The separation of IgG glycoforms, however, necessitates electrophoretic, chromatographic and/or MS techniques. [79]
Generally, electrophoretic methods allow the separation of molecules according to their charge and size. As N-glycans can have a pronounced impact on both of these two, rather global protein parameters, SDS-PAGE, IEF, 2D-Electrophoresis and various capillary-electrophoresis applications are successfully and routinely used for the analysis of various IgG glycoforms, or -at least- groups thereof.
Despite many advantages of electrophoretic methods, chromatographic separation techniques, such as ion-exchange, reversed-phase or affinity chromatography, proved more versatile and more efficient with regard to the separation of IgG glycoforms.
MS (and multi-stage MS even more so) provides means for the separation of IgG glycoforms according to mass. Fundamentally, MS methods depend on the spatial separation of molecules of different masses (e.g. Time-of-flight, Quadrupole) for analysis. Consequently, MS is commonly used for the detection and analysis of IgG glycoforms, also of intact IgG molecules.
However, the comprehensive and detailed characterization of N-glycosylation by analysing intact N-glycoproteins is intrinsically limited to proteins bearing only one N-linked oligo-saccharide. Although this limitation might not apply to the majority of recombinant IgG molecules, there are notable exceptions of therapeutically relevant IgG molecules which bear multiple N-glycoslyation sites and are thus not sufficiently characterized by such methods (e.g. Erbitux). Furthermore, the degree of information on the precise N-glycan structures and structural isomers obtained by these methods is usually very low [78].

The highest degree of information on N-glycan structures, particularly on structural isomers, however, is typically provided by a second group of analytical methods, which focuses on the analysis of either chemically or enzymatically released N-glycans. The analytical techniques most commonly used for this purpose include liquid chromatography, capillary electrophoresis, MS, and combinations thereof. [80]

Here fore, N-glycans can either be chemically released by hydrazinolysis, or enzymatically using peptide-N4-(N-acetyl-beta-glucosaminyl) asparagine amidases, such as PNGase F or PNGase A. The use of hydrazine is a non-selective, method for the release of all peptide-bound glycans (i.e. N- and O-linked glycans). This rather harsh method, however, requires water-free samples and leads to the inevitable loss of N-acetyl groups (e.g. of N-acetylglucosamine-residues). By contrast, alternative enzymatic reactions, which perform under significantly milder conditions, are exclusively able to release N-linked glycans and exhibit some substantial substrate specificities. Most notably, PNGase F, a bacterial enzyme, which typically releases N-glycans with high yields, does not release N-glycans bearing α-1,3-

linked core-fucose residues (i.e. it does not cleave a number of plant- or insect-specfic complex-type N-glycans). PNGase A, on the other hand, an enzyme derived from almonds, which is able to cleave these N-glycan structures, is most active at the glycopeptide level and thus requires prior protease digestion of the glycoprotein of interest. [78]

However released, free N-glycans are commonly subjected to chromatographic, MS, and electrophoretic analysis, as well as various combinations thereof. For this, the reducing ends of free glycans are typically derivatized with fluorescent aromatic amines by reductive amination, which facilitates sensitive detection, quantification, and -in some cases- chromatographic separation [81]. Despite the multitude of such labels reportedly used for this purpose, mainly three of them predominate in the analysis of IgG N-glycans, namely 2-AB (2-aminobenzamide), 2-AP (aka PA; i.e. 2-aminopyridine) and 2-AA (2-aminobenzoic acid), and became generally accepted standard methods for the quantitative and structural analysis of N-glycan [80].
The analytical methods of this group, which are most widely used for IgG N-gylcosylation analysis, thus, include normal-phase chromatography of 2-AB or 2-AA derivatized glycans (providing separation primarily according to oligosaccharide size and sometimes also the discrimination of constitutional isomers; e.g., β-galactose either linked to GlcNAc-residue linked to the the α-3- or the α-6-arm of the "heptasaccharide" core-structure) [80, 82, 83]; reversed-phase chromatography of 2-AP labelled glycans (with advantageous wide spacing of nonfucosylated, fucosylated, and bisected N-glycans; yet also substantial overlapping of sialylated structures) [84]. Additionally, several multi-dimensional approaches, which combine either weak anion-exchange chromatography [85] or normal-phase chromatography with reversed-phase chromatography [31, 86], are frequently used; particularly for the quantitative and structural assessment of sialylated N-glycans.

Chromatographic methods applied to free (i.e. unlabelled) glycans ultimately depend on detection systems other than UV-absorbance or fluorescence detection. Although the pioneering efforts in analysing native N-glycans using High-pH-Anion-Exchange-Chromatography in combination with Pulsed Amperometric Detection (i.e. HPAEC-PAD) were soon outmatched by almost contemporaneous advances in the reductive amination department, the more recent availability of more affordable soft-ionization MS equipment revived efforts in this field.

Nowadays, MS is widely used for the rapid and highly sensitive detection, as well as the compositional characterization, of N-glycans. To this end, underivatized, reduced, fluorescently labelled or permethylated N-glycans are typically analysed by single-stage MS, particularly MALDI-TOF MS [80]. Notably, due to the incidence of negatively charged monosaccharide constituents (e.g. sialic acids) the equal detection of neutral and charged oligo-saccharides by MS (in either ion-mode) requires special sample treatment, including permethylation of N-glycans [87] and esterification of sialic acids [88].
However, as single-stage MS is not able to separate or to discriminate between isobaric structural isomers of N-glycans, and recently developed multi-stage MS approach only partially compensate for [89-91], the full structural elucidation of such N-glycan structures yet greatly relies on other separation techniques, prior to MS analysis [92]. Thus, the combination of chromatographic methods for fluorescently labelled N-glycans, either directly (i.e. ESI) or

indirectly (i.e. MALDI) interfaced with MS [12, 30, 83, 93-96], greatly enriched the N-glycosylation analysis tool-box.
Additionally, in contrast to UV or fluorescence detection systems, MS is also amenable to free glycans, and thus eludes the use of rather harsh labelling reactions, which potentially generate structural artefacts (e.g. loss of α-2,3-linked sialic acid residues, or α-1,3-linked fucose residues) and typically require subsequent steps for the removal of excess reagents [80]. Although the panel of separation techniques for free N-glycans has been stated to be rather restricted [78], it yet comprises the recently developed approach of porous graphitized carbon chromatography (PGCC) of reduced N-glycans [11, 97, 98]. Particularly, the recently developed hyphenation of PGCC and MS (i.e. PGCC-ESI-MS) [11, 97-100] proved to be a non-discriminative method, able to separate and to detect both, neutral and sialylated oligosaccharides, and to allow the full structural assignment of IgG-derived N-glycans based on retention time and mass [11].

Nevertheless, methods exclusively based on the enzymatic or chemical release of all N-glycans in a given sample intrinsically fail to provide site-specific information on the N-glycosylation profiles of proteins with multiple N-glycosylation sites, and are clearly also more prone to errors arising from contaminant glycoproteins. Although these seemingly minor limitations may be negligible in the analysis of highly purified glycoproteins (e.g. biopharmaceutical products), more heterogeneous samples (e.g. rAbs prior to affinity-purification steps or other substantial down-stream processing) clearly pose a challenge to exclusively N-glycan-based analytical approaches and limits their use along research and production pipe-lines. In addition to these limitations, such methods also inherently fail to provide information on non-glycosylated or de-glycosylated protein fractions, as well as on N-glycan structures which are either not released by the commonly employed peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidases or not covered by standard separation or detection methods (e.g. single GlcNAc residues resulting from endo-glycosidase activity).

A viable strategy in order to evade a number of limitations arising from both, the analysis of intact glyco-proteins and/or the analysis of released N-glycans, is the analysis of glyco-peptides, employing a so-called glyco-proteomic approach. In contrast to other techniques, this analytical strategy allows the site-specific characterization of protein N-glycosylation. Also, this approach is able to provide detailed information on both, N-glycan and O-glycan structures, their precise location on the respective glyco-protein, the detection (and distinction) of non- or de-glycosylated protein fractions, and allows the unambiguous discrimination of contaminant glyco-peptides. In addition, the glyco-proteomic characterization concomitantly provides highly detailed information on the polypeptides primary structure.[11, 68, 69, 71, 101-103]
To this end, the glyco-proteomic approach makes use of well-established proteomic methods, typically including the electrophoretic separation of glyco-proteins, their *in situ* degradation by proteases with defined endo-peptidase activities, and the MS analysis of the resulting glyco-peptides.
Limitations to this approach, however, arise from the prerequisite need of the respective glyco-proteins amino-acid sequence (which -in the post-genomic era- is a minor limitation indeed!), a considerable lack of discriminative power for isobaric N-glycan structure isomers,

and –more generally- the potential pitfalls in quantitative MS of glyco-conjugates (e.g. equal detection of neutral and charged glycan structures).

AIM OF STUDY

"*The mission of the Pharma-Planta Consortium is to develop efficient and safe strategies for the production of clinical-grade protein pharmaceuticals in plants, and to define procedures and methods for the production of these proteins in compliance with all appropriate regulations.*" (Mission Statement of the Pharma-Planta Consortium; www.pharma-planta.org)

Although the initial aim of this doctoral thesis was the analysis and *in vitro* enzymatic modification of the N-glycosylation on therapeutic antibodies (and if possible also other pharmacologically relevant glycoproteins), contemporaneous achievements of other groups in this field [74-76], soon rendered further efforts towards the production of antibodies with a pre-defined and homogeneous N-glycosylation by post-harvest glycan-modifications (using fucosidase, GlcNAc-transferase I or II, galactosyltransferase or sialyltransferase) obsolete.
Furthermore and even more importantly, other, more potent strategies for the *in vivo* production of such glyco-modified antibodies were explored and successfully developed by other members of the Pharma-Planta Consortium, namely by the Steinkellner-Strasser group at the University of Natural Resources and Applied Life Sciences, Vienna (i.e. three floors above).

In the light of these developments, glycosylation analysis of recombinantly produced IgG molecules gained significance for the members of the Pharma-Planta Consortium glycoproteins. To this end, a robust and sensitive methodology for the quantitative N-glycosylation analysis, also applicable to samples from early stage strain development, should be established. Above all, the envisioned methodology should deliver comprehensive information regarding antibody processing, including underglycosylation, de-glycosylation as well as unexpected deviations from the expected amino acid sequence.

These requirements eventually directed this work towards an in-depth study of (glyco-)peptide analysis using RPC-ESI-MS. The validity of the quantitative information obtained for IgG1 molecules by this approach could be confirmed by an independent method.
Another line of development focused on the primary structure of IgG molecules and led to the idea that the data generated during glyco-proteomic analysis contained valuable information allowing also the relative quantification of IgG subclasses.

The developed methodology was extensively applied to numerous samples derived mainly from animal cell- or plant-based expression systems and was finally employed to define the glycosylation of the anti-inflammatory fraction of human polyclonal IgG.

Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides.

Stadlmann, J; Pabst, M; Kolarich, D; Kunert, R and Altmann, F (2008). Proteomics 8(14): 2858-71.

In order to assess the N-glycosylation properties of currently approved expression platforms for the large-scale production of therapeutic rAbs, the N-glycosylation profiles of six commercially available therapeutic IgG molecules were analysed. The product panel analysed, comprised of six IgG1 molecules, of which three were recombinantly expressed in CHO cell cultures (i.e. Xolair, Herecptin and Rituxan), and three were derived from murine myeloma cell cultures (i.e. Erbitux, Remicade, Zenapax).
The high degree of purity and homogeneity of the respective samples allowed the concomitant analysis of PNGase released 2-AB labelled by NP-HPLC, which provided commonly accepted quantitative data for a comparative evaluation of the glyco-proteomic approach.
In addition to the results already presented in [11], the following section provides a brief summary on the therapeutic application of the respective rAbs, and recapitulates additional results obtained from the glyco-proteomic characterisation of these products (e.g. N-terminal peptide integrity, C-terminal peptide processing). Unless stated otherwise, all information on the amino-acid sequences of the respective IgG molecules were obtained from DrugBank (www.drugbank.ca).

Omalizumab (Xolair; DrugBank acc. no.: BTD00081)

Xolair (Genentech / Novartis) is a humanized IgG1 κ monoclonal antibody, recombinantly expressed by CHO cell suspension culture and indicated for the treatment of asthma caused by allergies. The rAb selectively binds human IgE and thereby prevents their binding to the high-affinity IgE receptor FcεRI on the surface of mast cells and basophils. The reduction of IgE binding to these cells eventually limits their release of mediators of the allergic response.

Xolair HC:
EVQLVESGGGLVQPGGSLRLSCAVSGYSITSGYSWNWIRQAPGKGLEWVASITYDGSTNYADS
VKGRFTISRDDSKNTFYLQMNSLRAEDTAVYYCARGSHYFGHWHFAVWGQGTLVTVSSGPSVF
PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVP
SSSLGTQTYICNVNHKPSNTKVDKKAEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLM
ISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLN
GKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA
VEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL
SLSPGK

Xolair LC:
DIQLTQSPSSLSASVGDRVTITCRASQSVDYDGDSYMNWYQQKPGKAPKLLIYAASYLESGVP
SRFSGSGSGTDFTLTISSLQPEDFATYYCQQSHEDPYTFGQGTKVEIKRTVAAPSVFIFPPSD
EQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKAD
YEKHKVYACEVTHQGLSSPVTKSFNR

The expected N-terminal peptide (E^{1}VQLVESGGGLVQPGGSLR^{19}, $[M+H]^{+}$= 1882.003, pI= 4.53) of the rAbs heavy-chain was detected. Nevertheless, a (slightly later eluting) small portion of the peptide was also detected to carry an N-terminal pyrrolidone carboxylic acid (PyroE, -18 amu). As it is unclear to which extent the PyroE-modification alters the ionization properties of the analyte, a quantification of this N-terminal modification was not performed.

The expected C-terminal peptide (S^{440}LSLSPGK^{447}, $[M+H]^{+}$= 788.451, pI=8.47) was not detected. Instead, exclusively the C-terminally processed version of the peptide (S^{440}LSLSPG^{446}, $[M+H]^{+}$= 660.356, pI=5.24) was detected.

Trastuzumab (Herceptin; DrugBank acc. no.: BTD00072)

Herceptin is a humanized IgG1 κ monoclonal antibody produced in CHO cells and is used in the treatment of HER2-positive metatsatic breast cancer.
The rAb selectively binds to the extracellular domain of the human HER2 (or c-erbB2) proto-oncogene, an EGF receptor-like protein, which was found to be over-expressed in 25-30 % of primary breast cancers. Binding HER2 positive cells induced antibody mediated (complement mediated) killing of these cells.

Herceptin HC:
EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYTRYADS
VKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGTLVTVSSASTKGP
SVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVV
TVPSSSLGTQTYICNVNHKPSNTKVDKKVEPPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPK
DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQ
DWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYP
SDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT
QKSLSLSPGK

Herceptin LC:
DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGVPSRFS
GSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPPSDEQLK
SGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
KVYACEVTHQGLSSPVTKSFNRGEC

The expected N-terminal peptide (E^{1}VQLVESGGGLVQPGGSLR^{19}, $[M+H]^{+}$= 1882.003, pI= 4.53) of the rAbs heavy-chain was detected. Nevertheless, a (slightly later eluting) small portion of the peptide was also detected to carry an N-terminal pyrrolidone carboxylic acid (PyroE, -18 amu). A quantification of this N-terminal modification was not performed (since it is unclear to which extent the modification alters the ionization properties of the analyte).
Of the expected C-terminal peptide (S^{440}LSLSPGK^{447}, $[M+H]^{+}$= 788.451, pI=8.47) only trace amounts were detected. The major peptide form, however, exhibited the C-terminal K-truncation (S^{440}LSLSPG^{446}, $[M+H]^{+}$= 660.356, pI=5.24)

Cetuximab (Erbitux; DrugBank acc. no.: BTD00071)

Cetuximab is a humanized monoclonal IgG1 κ antibody produced in (not further described) murine myeloma cell culture. Its medial indication is for treatment of metastatic colorectal cancer. The rAb specifically bind to the epidermal growth factor receptor (EGFr, HER1, c-ErbB-1) on both normal and tumor cells, yet over-expressed in many colorectal cancer cells. The competitive binding of cetuximab to EGFr results in the inhibition of cell growth, the induction of apoptosis and a reduction of metastatic spread.

Erbitux HC:
QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTDYNTPF
TSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVTVSAASTKGPS
VFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVT
VPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSPKSCDKTHTCPPCPAPELLGGPSVFLFPPKP
KDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLH
QDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFY
PSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHY
TQKSLSLSPGK

Erbitux LC:
DILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFS
GSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKRTVAAPSVFIFPPSDEQLK
SGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
KVYACEVTHQGLSSPVTKSFNRGA

The expected N-terminal peptide ($Q^{1}VQLK^{5}$, $[M+H]^{+}$= 615.382, pI =8.75) was not detected at all. Instead, exclusively the N-terminally processed variant, carrying a pyrrolidone carboxylic acid ($PyroQ^{1}VQLK^{5}$, $[M+H]^{+}$= 598.356) was detected.
Surprisingly, the expected C-terminal peptide ($S^{440}LSLSPGK^{447}$, $[M+H]^{+}$= 788.451, pI=8.47) was detected in substantial amounts. The peptide variant exhibiting the C-terminal K-truncation ($S^{440}LSLSPG^{446}$, $[M+H]^{+}$= 660.356, pI=5.24) was also detected, with slightly higher abundance.
Erbitux bears an additional N-glycosylation site in the VH region of the rAbs heavy-chain, which was, in perfect agreement with the work recently published by [12] also found to be N-glycosylated. In contrast to the glycans detected at the N-glycosylation site located to the rAbs Fc-region, however, the N-glycan structures detected on the glycopeptide of the VH-region were bi- and tri-antennary, core-fuosylated structures, bearing substantial amounts of alpha-galactose residues and N-glycolyl-neuraminic acid (Ng).
Interestingly, the VL region of the molecules light-chain contained an N-glycosylation sequon (NGS), which was however not glycosylated (similar to what was observed with Remicade).

Infliximab (Remicade, DrugBank acc. no.: BTD00004)

Remicade is a humanized monoclonal IgG1 antibody produced by a recombinant murine (mouse) myeloma cell line (SP-210) cultured by continuous perfusion. The rAb finds its application in the treatments of Crohn's disease, psoriasis, rheumatoid arthitis and ankylosing

spondylitis. The antibody competitively binds to tumor necrosis factor α (TNF-α), thereby reducing synovitis and joint erosion in collagen-induced arthritis, eventually allowing the affected joints to heal.

For Remicade, amino-acid sequence data are only available for the heavy- and light-chain variable regions (Sequence No. 5 and 3, respectively, both from US patent 5656272). Due to rather a clandestine deposition of the rAbs primary structure information, the most probable amino-acid sequences were first confirmed by MS/MS experiments and tryptic peptide mass fingerprinting.
Interestingly, the VL region of the molecules light-chain contained an N-glycosylation sequon (NGS), which was however not glycosylated (similar to what was observed with Erbitux).

Remicade VH:
EVKLEESGGGLVQPGGSMKLSCVASGFIFSNHWMNWVRQSPEKGLEWVAEIRSKSINSATHYA
ESVKGRFTISRDDSKSAVYLQMNSLRTEDTGVYYCSRNYYGSTYDYGQGTTLTVS[**X**ASTKGP
SVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVV
TVPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSPKSCDKTHTCPPCPAPELLGGPSVFLFPPK
PKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVL
HQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGF
YPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNH
YTQKSLSLSPGK]

Remicade VL:
DILLTQSPAILSVSPGERVSFSCRASQFVGSSIHWYQQRTNGSPRLLIKYASESMSGIPSRFS
GSGSGTDFTLSINTVESEDIADYYCQQSHSWPFTFGSGTNLEVK[TVAAPSVFIFPPSDEQLK
SGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKH
KVYACEVTHQGLSSPVTKSFNRGEC]

The expected N-terminal peptide ($E^{1}VK^{3}$, $[M+H]^{+}$= 375.224, pI =6.10) was not detected. The lack of detection of this peptide was most probably due to its poor retention-properties on the RP-C18 material. The C-terminal peptide $S^{440}LSLSPGK^{447}$ ($[M+H]^{+}$= 788.451, pI=8.47), however, was detected in substantial amounts. The peptide variant exhibiting the C-terminal K-truncation ($S^{440}LSLSPG^{446}$, $[M+H]^{+}$= 660.356, pI=5.24) was also detected, yet with slightly lower abundance.

Dalizumab (Zenapax, DrugBank acc. no.: BTD00007)

Zenapax is a humanized monoclonal IgG1 molecule produced in murine (mouse) myeloma cell (GS-NS0/1) culture. The molecule specifically binds the Tac subunit of the high-affinity IL-2 receptor complex, which is primarily expressed on activated, but not resting, lymphocytes. The consequential inhibition of IL-2 binding of these cells was found to reduce the risk of renal transplant rejection.

Zenapax HC

QVQLVQSGAEVKKPGSSVKVSCKASGYTFTSYRMHWVRQAPGQGLEWIGYINPSTGYTEYNQK
FKDKATITADESTNTAYMELSSLRSEDTAVYYCARGGGVFDYWGQGTTLTVSSGPSVFPLAPS
SKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLG
TQTYICNVNHKPSNTKVDKKAEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTP
EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK
CKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWES
NGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG
K

Zenapax LC
DIQMTQSPSTLSASVGDRVTITCSASSSISYMHWYQQKPGKAPKLLIYTTSNLASGVPARFSG
SGSGTEFTLTISSLQPDDFATYYCHQRSTYPLTFGSGTKVEVKRTVAAPSVFIFPPSDEQLKS
GTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHK
VYACEVTHQGLSSPVTKSFNR

Low amounts of the expected N-terminal peptide (Q^{1}VQLVQSGAEVK^{12}, $[M+H]^{+}$= 1285.711, pI =6.00) were detected. The major peptide variant detected represented the N-terminally processed variant, carrying a pyrrolidone carboxylic acid (PyroQ^{1}VQLVQSGAEVK^{12}, $[M+H]^{+}$= 1268.6840).
The expected C-terminal peptide (S^{440}LSLSPGK^{447}, $[M+H]^{+}$= 788.451, pI=8.47) was detected in substantial amounts. However, the peptide form exhibiting the C-terminal K-truncation (S^{440}LSLSPG^{446}, $[M+H]^{+}$= 660.356, pI=5.24) was also detected, with significantly higher abundance.

Rituximab (Rituxan or MabThera, DrugBank acc. no.: BTD00014)

Rituxan is a humanized monoclonal IgG1 antibody produced by CHO suspension culture in a nutrient medium containing gentamycin. The rAb is purified from cell culture supernatant by means of affinity and ion exchange chromatography. It binds the CD20 antigen predominantly expressed on mature B lymphocytes and more than 90% of non-Hodgkin's lymphoma B-cells, and induces their selective lysis.

Rituximab HC
QAYLQQSGAELVRPGASVKMSCKASGYTFTSYNMHWVKQTPRQGLEWIGAIYPGNGDTSYNQK
FKGKATLTVDKSSSTAYMQLSSLTSEDSAVYFCARVVYYSNSYWYFDVWGTGTTVTVSGPSVF
PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVP
SSSLGTQTYICNVNHKPSNTKVDKKAEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLM
ISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLN
GKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA
VEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL
SLSPGK

Rituximab LC
QIVLSQSPAILSASPGEKVTMTCRASSSVSYMHWYQQKPGSSPKPWIYAPSNLASGVPARFSG
SGSGTSYSLTISRVEAEDAATYYCQQWSFNPPTFGAGTKLELKRTVAAPSVFIFPPSDEQLKS

GTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHK
VYACEVTHQGLSSPVTKSFNR

Of the expected C-terminal peptide ($S^{440}LSLSPGK^{447}$, $[M+H]^+$= 788.451, pI=8.47) only trace amounts were detected. The major peptide form, however, exhibited the C-terminal K-truncation ($S^{440}LSLSPG^{446}$, $[M+H]^+$= 660.356, pI=5.24)

Conclusions

All six antibodies analysed were found to exhibit very simple CH2-linked N-glycans, generally consisting of non-sialylated, fucosylated, bi-antennary structures, with a low degree of galactosylation. In addition, minor amounts of oligo-mannosidic and hybrid-type N-glycan structures were detected. However, despite the over-all similarity of the N-glycosylation profiles, subtle differences were observed. In comparison to the CHO cell culture derived products, myeloma cell culture derived rAbs exhibited a slightly higher amount of hybrid N-glycan structures (e.g. MA(a3,4)F or Man5GnF; ranging from 4.1-10%, compared to 1.3-2% in CHOs).
Similarly, subtle differences in the polypeptide processing properties of the two groups of mammalian expression systems (i.e. murine myeloma and CHO cells) were detected. Most noteworthy, in contrast to CHO cell derived products, myeloma cell derived products seemingly contain larger fractions of molecules still bearing the C-terminal K-residue. This finding was consistent with the observations of others [104].

DOI 10.1002/pmic.200700968

RESEARCH ARTICLE

Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides

Johannes Stadlmann[1], Martin Pabst[1]*, Daniel Kolarich[1]**, Renate Kunert[2] and Friedrich Altmann[1]*

[1] Department of Chemistry, University of Natural Resources and Applied Life Sciences (BOKU), Vienna, Austria
[2] Department of Biotechnology, University of Natural Resources and Applied Life Sciences (BOKU), Vienna, Austria

Two LC-ESI-MS methods for the analysis of antibody glycosylation are presented. In the first approach, tryptic glycopeptides are separated by RP chromatography and analyzed by ESI-MS. This "glycopeptide strategy" allows a protein- and subclass-specific quantitation of both neutral and sialylated glycan structures. Additional information about under- or deglycosylation and the protein backbone, *e.g.*, termini, can be extracted from the same data. In the second LC-ESI-MS method, released oligosaccharides are separated on porous graphitic carbon (PGC). A complete structural assignment of neutral and sialylated oligosaccharides occurring on antibodies is thereby achieved in one chromatographic run. The two methods were applied to polyclonal human IgG, to commercial mAb expressed in CHO cells (Rituximab, Xolair, and Herceptin), in SP2/0 (Erbitux and Remicade) or NS0 cells (Zenapax) and the anti-HIV antibody 4E10 produced either in CHO cells or in a human cell line. Both methods require comparably little sample preparation and can be applied to SDS-PAGE bands. They both outperform non-MS methods in terms of reliability of peak assignment and MALDI-MS of underivatized glycans with regard to the recording of sialylated structures. Regarding fast and yet detailed structural assignment, LC-MS on graphitic carbon supersedes all other current methods.

Received: October 11, 2007
Revised: February 25, 2008
Accepted: April 2, 2008

Keywords:
Antibody / Glycopeptides / Glycoproteomics / *N*-glycan / Structural analysis

1 Introduction

Antibodies are the largest group of recombinantly produced glycoproteins in terms of the number of different molecules as well as the amount produced. Analysis of their glycosylation status constitutes an integral part of regulatory necessities.

Correspondence: Dr. Friedrich Altmann, Department of Chemistry, University of Natural Resources and Applied Life Sciences (BOKU), Muthgasse 18, A-1190 Wien, Austria
E-mail: friedrich.altmann@boku.ac.at
Fax: +43-1-36006-6059

Abbreviations: BPI, base peak intensity; **GlcNAc,** *N*-acetylglucosamine; **NP,** normal phase; **PGC,** porous graphitic carbon; **PGCC,** porous graphitic carbon chromatography

Apart from regulatory issues, two currently emerging developments in the research area require (high-throughput) analysis of antibody glycosylation: (i) the recently discovered effects of glycosylation on the biological efficacy of antibodies, in particular on antibody-dependent cellular cytotoxicity (ADCC) [1] and (ii) the development of novel expression systems (other than mammalian cells), *e.g.*, yeasts or green plants [2–6].

Plasma-derived human antibodies carry a limited number of diantennary complex-type *N*-glycans: mostly with core α1,6-fucose, none- to two-galactose residues, maybe one,

* Both these authors contributed equally.
** Current address: Core of Biomolecular Frontiers, Macquarie University, NSW 2109 Sydney, Australia

rarely two α2,6-linked sialic acids and a small amount of bisecting *N*-acetylglucosamine (GlcNAc), as has been extensively demonstrated (because of the exceptional importance of this serum protein) [7–13]. Recombinantly produced mAbs, however, usually exhibit a slightly simpler spectrum, lacking bisecting GlcNAc and/or nonfucosylated species [14].

Various methods for the analysis of glycoproteins have recently been reviewed by Geyer and Geyer [15]. The first group of such methods commonly applied to antibodies solely employs chromatography of released *N*-glycans. The pioneering efforts using gelfiltration [16] have been superseded by (i) normal-phase (NP) HPLC on amide-silica of glycans derivatized either with 2-aminobenzamide (AB) or anthranilic acid, providing separation of glycans primarily according to size, and sometimes also discrimination of constitutional isomers, *e.g.*, species with one galactose on either the 3- or the 6-arm [7, 13, 17]; (ii) RP HPLC in phosphate buffer of 2-aminopyridine derivatized glycans, which yields an advantageous wide spacing of nonfucosylated, fucosylated, and bisected *N*-glycans [8, 11]. It should, however, be noted that the hardly avoidable racemization products of the fucosylated species elute close to the nonfucosylated glycans (Altmann, unpublished results); (iii) Anion-exchange HPLC on amine-silica of anthranilic acid labeled glycans [18]; and (iv) Mixed-mode HPLC on amine-silica [19]. In addition to these methods, multidimensional separation techniques, combining anion-exchange fractionation and subsequent RP-HPLC of 2-aminopyridine labeled glycans, have been used in order to facilitate the analysis of neutral and charged IgG glycans [11].

Two other powerful separation approaches are high-pH anion exchange chromatography of free oligosaccharides [20–22] and CE of fluorescently labeled glycans [9, 23], which both require specialized and intricate equipment.

The second large set of methods frequently applied to IgG glycans uses MS, and here in particular MALDI-TOF MS. Released glycans are predominantly analyzed either underivatized [14] or permethylated [13]. A salient advantage of MALDI-TOF MS is that the results can be directly translated into glycan composition. Although it is often only a small step from glycan composition to structural assignment, particularly for products from a well characterized biological systems, *e.g.*, CHO cells or green plants [2, 14], structural isomers cannot be discriminated by single-stage MS alone. Modern MALDI-TOF/TOF instruments may deliver more information [15, 24, 25], but another drawback of MALDI-TOF-MS arises from the fact that sialylated glycans require special sample preparation in order to be detectable with a sensitivity comparable to that of neutral glycans [15]. For this purpose permethylation, which is stated to ensure an equal detection sensitivity of neutral and sialylated *N*-glycans [13, 26], and esterification of sialic acids, which removes their negative charge and increases glycan stability [27] have been used. If only sialylated glycans are of interest, these can be measured very sensitively in the negative mode with the matrix trihydroxyacetophenone [9, 28]. Two recent papers present new tricks for this old problem, *i.e.*, negative-ion mode MALDI with (i) native glycans and 2-anthranilic acid as the matrix [18] and (ii) 2-anthranilic acid-labeled glycans with DHB as the matrix [19].

In an impressive multilaboratory work most of the methods mentioned above have been thoroughly compared with respect to the analysis of IgG *N*-glycans [13]. Unfortunately, two LC-ESI-MS methods have not received the attention they deserve in our view.

The first, RP-LC-ESI-MS of IgG glycopeptides, was pioneered already in the last millennium [29, 30] and shall herein be termed RPC-MS. While some researchers have preferred MALDI-MS of offline isolated glycopeptides [13, 31], we and others have found direct LC-ESI-MS to be a highly useful approach for analysis of Fc-region [6, 30–33] or variable region glycosylation [34]. Like MALDI-TOF MS it is not able to resolve isomers. However, as the ionization is predominantly dictated by the peptide portion, absence or presence of charges on the glycan moiety can be expected to matter little, or not at all. A thorough investigation on the occurrence of in-source fragmentation and on the detector response of neutral *versus* sialylated glycans has, however, not yet been made.

The second method, LC of glycans on porous graphitic carbon (PGC) followed by ESI-MS detection (PGC-LC-ESI-MS or PGC chromatography (PGCC)-MS), has also been used by two groups taking part in the multilaboratory test [13]. However, even though the qualification of PGCC-MS for quantitation of neutral glycans was demonstrated, no data on the method's ability to detect sialylated glycans were shown. Moreover, the recently demonstrated ability of PGCC-MS to separate and identify structural isomers [35] has not yet been applied to antibody glycans, where structural variety is introduced, *e.g.*, by sialic acid in α2,3- or α2,6-linkage, being either attached to galactose on the 3- or the 6-arm. The distinction of these structural isomers becomes especially important in the case of nonhuman expression hosts, as may be exemplified by digalactosylated glycans, which are usually assumed to carry two terminal β1,4-linked galactose residues. In mouse cells, however, the glycans attached to recombinantly expressed antibodies may also receive α1,3-linked galactose, giving rise to a glycan with one terminal GlcNAc and a Galα1,3Gal disaccharide on the other arm (*i.e.*, $GnA^{a3\text{-}4}$, which is isobaric with A^4A^4, see structures **12** and **40** in Fig. 1).

In this paper, we demonstrate that neutral and sialylated *N*-glycans of antibodies can be simultaneously measured by LC-ESI-MS. Tryptic glycopeptides are used for the profiling of *N*-glycans – sub-class and site-specific if required – by RPC-MS, while PGCC-MS of free oligosaccharides delivers a complete qualitative structural assignment of antibody glycosylation.

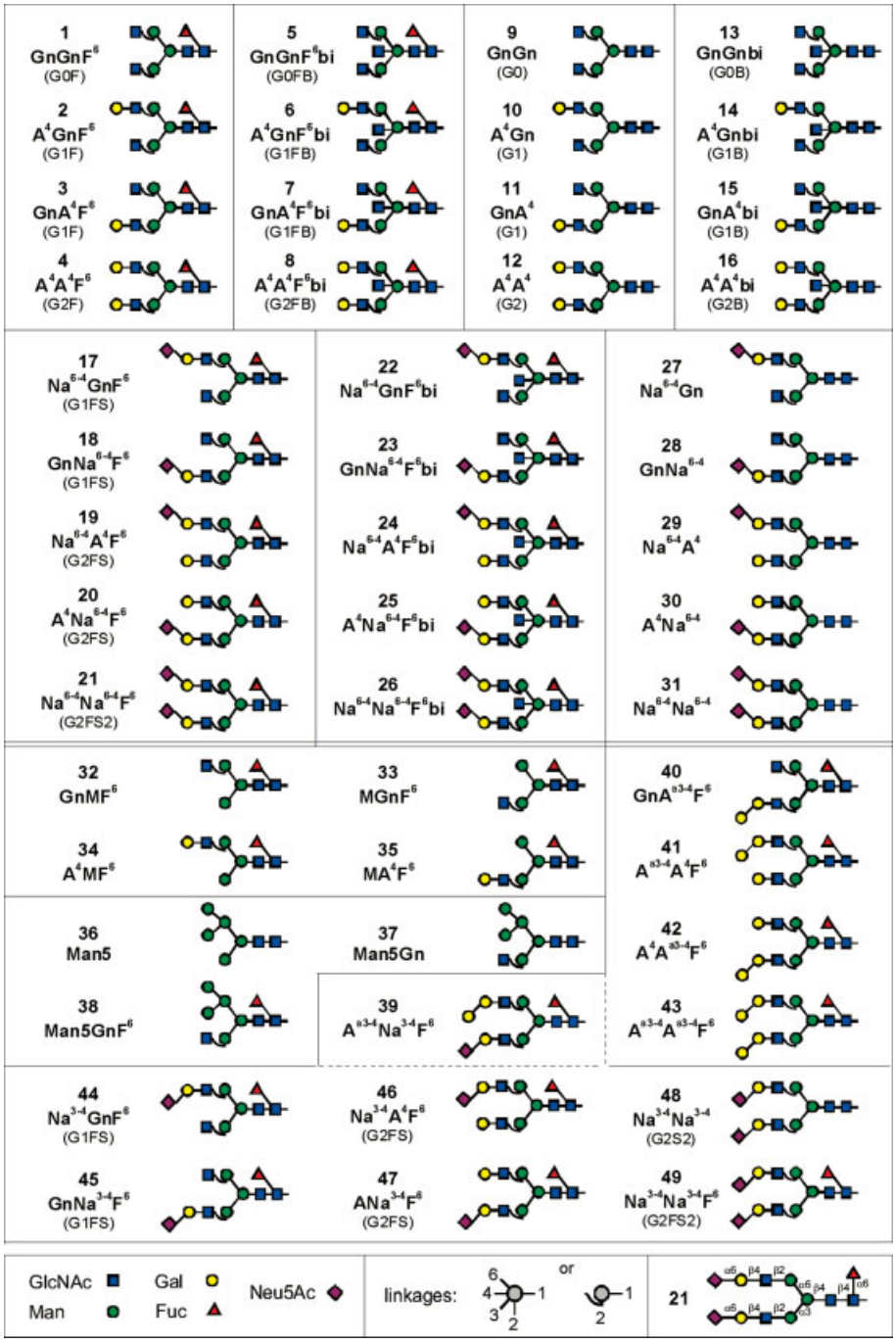

Figure 1. Structures of the *N*-glycans dealt with in this study. The monosaccharide symbols are used as suggested by the Consortium of Functional Glycomics (CFG, www.functionalglycomics.com). Glycans are named according to the proglycan system with full designation of linkages as indicated at the bottom (www.proglycan.com) [35]. The lower lines – where present – show a different type of annotation that counts the number of galactose, fucose, and sialic acid residues [7]. For convenience, the glycans also have consecutive numbers. In the bottom right corner, glycan 21 is shown in true CFG style. Structures **1–31** are the combinations of the usual structural elements fucosylation, galactosylation, bisection, and sialylation found on human IgG [7–13]. Some of these structures may be hardly detectable or absent at all such as, *e.g.*, NaNaFbi. The truncated glycans 32–35 occur in small quantities in some antibodies and/or as artifacts of MS analysis. Glycans 36–49 were found in certain monoclonals. Some very minor structures having been found elsewhere are not shown here, *e.g.*, triantennary glycans [13] or glycans with galactose linked to bisecting GlcNAc [51].

2 Material and methods

2.1 Antibodies

The protein reference preparation CRM470 (International Federation of Clinical Chemistry and Laboratory Medicine, Milano, Italy) was used to obtain multidonor, polyclonal human IgGs. In order to avoid a Protein-A/G-derived bias for or against particular IgG-subclasses, the IgG heavy chain band of the CRM470 sample was directly excised from an SDS-PAGE gel for the glycopeptide work. For oligosaccharide preparations, IgG was isolated using protein A [5]. Guinea pig IgG was prepared by SDS-PAGE from serum of laboratory animals (Medical University Vienna). The monoclonals rituximab, trastuzumab, and daclizumab (trade names Rituxan, Herceptin, and Zenapax, respectively) were from Roche (Basel, Switzerland), omalizub (Xolair) from Novartis (Basel, Switzerland), Cetuximab (Erbitux) from Bristol-

Myers Squibb (New York, USA), and Infliximab (Remicade) from Johnson & Johnson (New Brunswick, NJ, USA). The drugs were kindly provided by Dr. Gottfried Himmler (f-star Vienna, Austria) or, in the case of Zenapax, Dr. Anton Stift (Medical University, Vienna).

The anti-HIV monoclonal 4E10 was produced in CHO cells (ATCC 9096) in protein-free medium under GMP conditions by Polymun Scientific (Vienna, Austria). 4E10 was additionally produced in a human Burkitt lymphoma cell line.

2.2 Glycopeptide and oligosaccharide preparation

Antibody (1–2 μg) was subjected to standard SDS-PAGE under reducing conditions. Heavy chain bands were excised rather generously to include variants slightly differing in size due to glycosylation or polypeptide chain length. *S*-carbamidomethylation, trypsin digestion, and extraction were performed by routine methods [36].

The standard approach for oligosaccharide preparation started with the glycopeptide mixture as prepared by the above procedure – however, from samples containing 3 μg or more of antibody previously purified by protein A affinity chromatography. The dried samples were redissolved in 10 μL of 25 mM ammonium acetate buffer pH 5.0 and heated for 5 min to 95°C to inactivate residual trypsin. Then 0.3 mU of PNGase A (Proglycan, Vienna, Austria) were added. After incubation at 37° for 12 h the liquid was passed through a spec PT C18 cartridge (Varian, Palo Alto, CA, USA). The cartridge was rinsed with 50 μL 2% acetic acid and the flow through was dried and neutralized with dilute ammonia.

Glycans were reduced with 5 μL of a 1% solution of $NaBH_4$ at room temperature for 4 h. To remove excess salt and remaining peptides, an additional purification step with a 10 mg HyperSep Hypercarb SPE cartridge (Thermo Scientific, Waltham, MA, USA) was performed. Deviating from the original work [37], glycans were eluted with 1% ammonia in 60% ACN.

2.3 LC-ESI-MS of glycopeptides (RPC-MS)

Analysis of peptides and glycopeptides was performed on a capillary LC-ESI-MS system consisting of an Aquasil C-18 precolumn (30 mm × 0.32 mm, 5 μm, Thermo Scientific), a BioBasic C18 analytical column (150 mm × 0.18 mm, 5 μm, Thermo Scientific), a Waters CapLC, a Rheodyne 10-port valve and a Waters Q-TOF Ultima with standard ESI-source. Deviating from recent work [33, 36], solvent A consisted of 65 mM ammonium formate of pH 3.0 and solvent B was 80% ACN in solvent A as also used for PGC-LC [35]. The precolumn was equilibrated and loaded in the absence of ACN. Thereafter, a gradient from 6.3 to 62.5% solvent B was developed over 45 min. Positive ions in the range from m/z 150 to 1800 were measured. Capillary voltage was 3.2 kV and cone voltage 35 V, source temperature was 100°C and desolvation temperature 120°C.

Data were evaluated using MassLynx 4.0 software and herein, notably, the MaxEnt3 deconvolution/deisotoping feature [38].

2.4 LC-ESI-MS of glycans (PGCC-MS)

Analysis of borohydride-reduced oligosaccharides by positive-ion LC-ESI-MS was performed with a 100 mm × 0.32 mm PGC column (Thermo) and a flow rate of 5 μL/min but otherwise essentially as described recently [35].

2.5 Reference oligosaccharides

Human polyclonal IgG was used as a source of core-α1,6-fucosylated, variously β1,4- galactosylated glycans without ($GnGnF^6$, A^4GnF^6, GnA^4F^6, $A^4A^4F^6$ and GnGn, A^4Gn, GnA^4, A^4A^4 = structures **1**–**4** and **9**–**12** in Fig. 1) and with bisecting GlcNAc ($GnGnF^6$bi, A^4GnF^6bi = **5** and **6**) and $A^4A^4F^6$bi (**8**) plus the nonfucosylated versions thereof) [8, 11]. The isomers A^4GnF^6 and GnA^4F^6 (**2**, **3**) were isolated by micro-preparative HPLC on the capillary column. The α2,6-sialylated reference glycans $A^4Na^{6\text{-}4}F^6$ and $Na^{6\text{-}4}Na^{6\text{-}4}F^6$ (structures **21** and **20**) were obtained from porcine fibrin [35]. A 30 min hydrolysis at 100°C with 0.1 M formic acid yielded the $Na^{6\text{-}4}A^4F^6$ isomer (**19**). From these monosialylated glycans, two further references ($Na^{6\text{-}4}GnF^6$ and $GnNa^{6\text{-}4}F^6$ = structures **17** and **18**) were made by galactosidase digestion [35]. The α2,3-sialylated reference glycan $Na^{3\text{-}4}Na^{3\text{-}4}F^6$ (**49**) was obtained from recombinant erythropoietin (Janssen-Silag Pharma, Vienna, Austria) by semi-preparative HPLC on a 150 mm × 3 mm PGC column with UV detection at 210 nm and a flow rate of 0.6 mL/min using the same solvents as used for PGCC-MS. The α2,3-monosialylated glycans were obtained by incubation with 0.1% formic acid for only 5 min. The two isomers were isolated by micro-preparative HPLC and separately digested with β-galactosidase to obtain $Na^{3\text{-}4}GnF^6$ and $GnNa^{3\text{-}4}F^6$ (structures **44** and **45**). Unfucosylated α2,6- and α2,3-sialylated glycans were similarly prepared from bovine fibrin with the help of α2,3-sialyltransferase (Calbiochem/Merck Biosciences, Nottingham, UK) as described in the previous paper [35]. The strategies for assignment of isobaric compounds are described in Section 3.4. Man5, Man5Gn, MGn, and GnM were available from the previous study [35]. Retention times for $GnMF^6$ and $MGnF^6$ (structures **32** and **33**) were obtained by *N*-acetylglucosaminidase digest of IgG glycans.

2.6 NP-HPLC of labeled glycans

Oligosaccharides were prepared as described above and labeled with AB [39] with the deviation that cyano-SPE cartridges (100 mg, AccuBond II CN, Agilent) were used for hydrophilic purification. NP-HPLC was performed as described recently [7, 40].

3 Results

3.1 LC-ESI analysis of IgG glycopeptides

For the "RPC-MS" approach, a standard proteomics setting involving tryptic glycopeptides, a reversed-phase precolumn, and an analytical column, coupled to an ESI source and a Q-TOF MS instrument were used. It seems that not each and every combination of pre- and main-column is suitable for the retention of the rather hydrophilic IgG glycopeptide(s), which may be one possible reason for the low dissemination of the RPC-MS strategy. Deviating from previous works, we have herein used a 65 mM ammonium formate buffer (pH 3.0) as LC solvent. While the primary intention was to avoid changing of solvents when switching from RP to PGC and back, it soon turned out that this buffered solvent had two advantages over the usually used 0.1% formic acid. First, the peaks became narrower and hence the S/N improved and, second, glycopeptides with oligomannosidic structures exhibited much less fragmentation. This was experienced with maize-produced antibodies containing mainly oligomannosidic glycans, where a considerable amount of fragmented glycans was found with formic acid [41], but not with ammonium formate (results not shown). Care must, however, be taken as the nitrogen content facilitates microbial growth in the solvent containers. Whatever the aqueous component, success depends on the use of a precolumn able to operate with 100% aqueous solvent and to retain rather polar peptides.

The primary result is an LC-ESI-MS chromatogram file of several 100 kB from which the glycopeptide profile has to be skillfully extracted. The glycopeptides of IgG1 and IgG2 are rather hydrophilic and elute amongst the first peptides in the LC profile of the tryptic IgG (glyco-)peptides (Fig. 2). The oligosaccharide moiety exerts a small contribution to the low binding of the glycopeptides to the RP, as deglycosylated peptides elute about 2.5 min later (data not shown).

The tryptic cleavage site N-terminal to glycosylated Asn is only slowly digested (possibly due to steric effects of the oligosaccharide, as judged from analysis of partially unglycosylated IgG (data not shown)), which leads to the habitual appearance of a fully and an incompletely digested glycopeptide for both IgG1 (gp1a = EEQYNSTYR, gp1b = TKPREEQYNSTYR) and IgG2 (gp2a = EEQFNSTFR, gp2b = TKPREEQFNSTFR) in varying amounts (Fig. 2). While this results in a doubling of the peak set, which seemingly complicates interpretation, it yet strengthens plausibility of peak assignment. We routinely combine all spectra containing glycopeptide peaks of a given IgG subclass. This "summed spectrum" contains glycopeptide peaks of charge states from +2 to +4 and several other peaks of nonglycosylated peptides. In order to obtain a simplified and quantifiable result, the summed spectrum is deconvoluted and deisotoped (Fig. 3). This simplified spectrum is then searched for peaks potentially representing glycopeptides with the help of a spreadsheet calculating the masses of all possible glycopeptides by adding peptide (short and long form) and glycan masses (Fig. 3 and Table 1) [10, 11, 13].

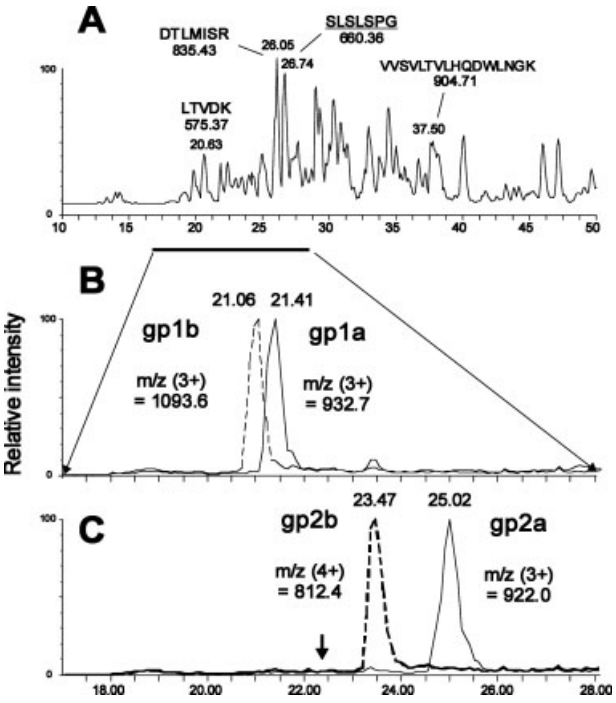

Figure 2. RPC-MS of peptides and glycopeptides from human IgG. (A) Part of the base peak intensity (BPI) chromatogram of hIgG. Four easily identifiable IgG1 peptides are labeled, among them the Lys-truncated C-terminal peptide SLSLSPG. (B) The SIM traces for the IgG1 glycopeptides gp1a and gp1b being the fully cleaved peptide (EEQYNSTYR) and the incompletely digested version with an N-terminal TKPR. (C) The SIM trace for IgG2, where gp2a has the sequence EEQFNSTFR and gp2b is the longer versions. The arrow in (C) gives the elution position of the glycopeptide from IgG4 (and the A2NU35 variant of IgG3).

The polyclonal antibody sample contained a third glycopeptide peak eluting just between those of IgG1 and IgG2 (Fig. 2). This glycopeptide arises from IgG4 (gp4a = EEQFNSTYR) and also from allelic variants of IgG3 [42], which, however, was recently found to occur only as the variant with an EEQYNSTYR sequence (UniProt Q8N4Y9) isobaric with the IgG2 glycopeptide [43].

The spectra obtained for human polyclonal IgG1 and IgG2 both show a variety of structures arising from absence or presence of fucose, bisecting GlcNAc, terminal galactose, and sialic acid (Fig. 3). Notably, in all subclasses only traces of structures bearing two sialic acid residues or lacking fucose were detected. These profiles therefore deviate from the results obtained from the analyses of free oligosaccharides by either HPLC, MALDI-TOF MS, or PGCC-MS (see Section 3.4). These differences, however, are not a methodological artifact, but rather the necessary result of a fraction of IgG being glycosylated in the variable regions of the light and heavy chains [11, 44]. The *N*-glycans located to the variable chains differ structurally from those attached to the constant region of hIgG and mAbs [11, 19, 34]. A site-specific analysis of *N*-glycans attached to the variable region of mAbs can be accomplished by RPC-MS (see Section 3.2).

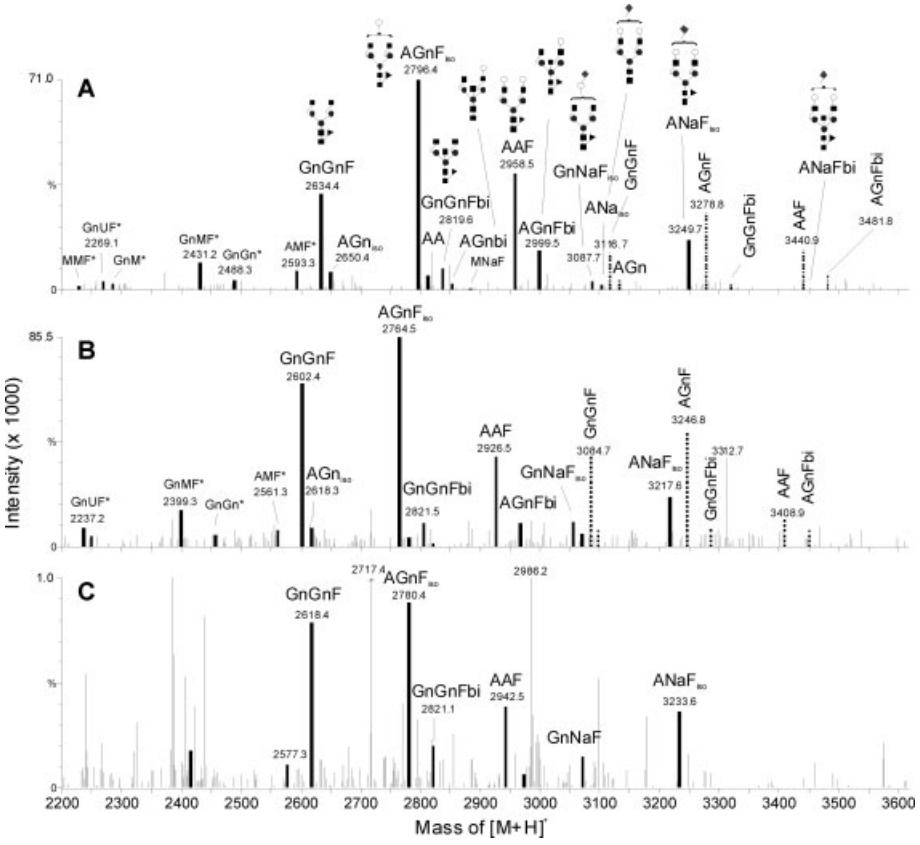

Figure 3. Glycopeptide spectra from human IgG. The relevant parts of the deconvoluted spectra of (A) IgG1, (B) IgG2 (plus IgG3), and (C) IgG4 are shown. Peaks assignable to the fully cleaved peptides are shown in bold lines. The partially digested glycopeptides are shown as dashed lines. Peak names with an asterisk denote species which are essentially in-source fragments as judged from charge state comparison. Peak abbreviations are given without linkage superscripts to save space. The suffix "iso" to glycan abbreviations denotes the probable presence of isomers, space permitting. Unassigned peaks are background, mainly from coeluting peptides of IgG, trypsin, or other serum proteins.

Unfortunately, most of the small m/z peaks indicating, *e.g.*, the presence of $MGnF^6$ (structure **33**) or its isomer must be regarded as artifacts (Fig. 3). Although $MGnF^6$ has been reported for myeloma IgG2 and IgG3 [45], the respective signals are not found in PGCC-MS of polyclonal hIgG. The (at least partial) artifactual nature of these peaks is revealed by close inspection of the charge states in which they are detected. In the hIgG spectrum, a strong signal of $m/z = 1216.0$ for the respective doubly charged glycopeptide was present, yet no triply charged species was detected. In contrast, PGCC-MS revealed the presence of small amounts of $MGnF^6$ in most mAbs, and indeed, these samples exhibited the respective triply charged glycopeptide. The occurrence of $[M + 3H]^{3+}$ ions and their ratio to $[M + 2H]^{2+}$ ions thus helps to distinguish real components of the sample from in-source fragments.

An exception was posed by the oligomannosidic glycopeptides, namely with Man5, which predominantly occurred doubly charged – although it undoubtedly existed in the sample.

3.2 mAb and animal IgGs

A panel of seven CHO-cell derived mAbs was investigated. The four commercial antibodies (Rituximab, Omalizumab, Trastuzumab, Infliximab) exhibited a very simple glycan profile essentially consisting of neutral, fucosylated glycans, with a rather low degree of galactosylation (Table 1), similar to what has recently been reported for Rituximab [14, 23]. In addition, small amounts of Man5 were identified in some mAbs (Table 1). Surprisingly, galactosylation was much higher in the CHO-cell derived anti-HIV mAb 4E10 (Table 1 and Fig. 4), which was, most noteworthy, the only CHO-produced mAb containing considerable amounts of sialylated *N*-glycans. The same mAb, produced in human lymphoma cells, however, yielded even higher degrees of galactosylation and sialylation (Table 1).

Daclizumab (Zenapax) is produced in mouse NS0 cells and contains only neutral glycans, similar to the previously mentioned CHO-produced mAbs. Here, however, the Fc-linked *N*-glycans also included species with one or two α1,3-Gal residues, in agreement with earlier studies of another mAb made in NS0 cells [30]. Hybrid type glycans were found and identified as MA^4F^6 and $MA^{a3\text{-}4}F^6$ (structures **35** and **40**) as will be detailed later.

In contrast to all the other mAbs analyzed in this study, cetuximab contains a glycosylation site located to the variable domain of the heavy chain. Whereas the CH_2-linked glycans were found to exhibit roughly the same profile as those found on other mAbs, consisting mainly of $GnGnF^6$ (structure **1**, Table 1 and Fig. 5), the V_H-linked glycans, however, were found to be larger and to contain α-Gal or sialic acid in

Table 1. Results of analyses by RPC-MS of various mAbs

Mass[a)] [M+H]+	Composition[b)]	Structure interpreted	Number, see Fig. 1	Omalizumab (Xolair)	Trastuzumab (Herceptin)	Cetuximab (Erbitux)	Infliximab (Remicade)	Daclizumab (Zenapax)	Rituximab (Rituxan)	4E10 CHO	4E10 hum	IGHG 1 Fc region	IGHG 2 Fc region[d)]
2405.94	H_5N_2	Man5	36	**8.5**	**13.8**	**27.6**	**7.3**	3.8	2.2	–	1.4	–	–
2430.97	$H_3N_3F_1$	$GnMF^6$/$MGnF^6$	32, 33	2.8	3.5	4.1	1.6	2.3	1.9	–	–	–	–
2488.00	H_3N_4	GnGn	9	**5.5**	**19.4**	3.5	3.0	–	2.6	2.4	1.5	5.2	5.7
2567.99	H_6N_2	Man6	–	–	4.3	**5.3**	–	–	–	–	–	–	–
2609.01	H_5N_3	Man5Gn	37	3.3	–	1.4	3.4	1.0	–	–	1.3	0.2	–
2634.05	$H_3N_4F_1$	$GnGnF^6$	1	**100.0**	**100.0**	**100.0**	**60.6**	**100.0**	**100.0**	**44.2**	1.5	**46.8**	77.9
2650.04	H_4N_4	A^4Gn/GnA^4	10, 11	3.6	4.9	2.4	3.9	1.4	1.2	6.8	1.1	10.7	8.9
2730.04	H_7N_2	Man7	–	–	2.9	2.0	–	–	–	–	–	–	–
2755.07	$H_5N_3F_1$	[c)]	38, 40	–	1.8	**6.1**	3.5	**7.4**	–	–	1.4	–	–
2796.10	$H_4N_4F_1$	A^4GnF^6/GnA^4F^6	2, 3	**58.4**	**99.9**	**68.7**	**100.0**	**97.0**	**78.1**	**100.0**	**16.2**	**100.0**	**100.0**
2812.09	H_5N_4	A^4A^4	12	1.2	2.5	1.0	2.7	–	–	5.9	1.4	7.9	4.3
2849.33	$H_3N_5F_1$	$GnGnF^6$bi	5	–	–	–	–	–	–	–	–	**8.0**	**11.3**
2853.12	H_4N_5	A^4Gnbi	14 (15)	–	–	–	–	–	–	–	–	2.5	1.6
2958.15	$H_5N_4F_1$	$A^4A^4F^6$	4	**6.9**	**22.1**	**11.2**	**33.8**	**28.4**	**12.1**	**83.7**	**41.0**	**53.8**	**42.5**
2999.18	$H_4N_5F_1$	A^4GnF^6bi	6 (7)	–	–	–	–	2.4	–	–	–	**18.0**	**11.0**
3087.19	$H_4N_4F_1S_1$	$GnNaF^6$	18 (17), 44, 45	–	–	–	–	–	–	**8.0**	2.5	3.8	**11.8**
3120.20	$H_6N_4F_1$	$A^4A^4F^6$ + Hex[e)]	41, 42	–	6.2	–	–	**8.4**	–	–	–	–	–
3161.23	H_5N5F_1	$A^4A^4F^6$bi	8	–	–	–	–	–	–	–	–	4.2	4.8
3249.25	$H_5N_4F_1S_1$	$A^4Na^{6\text{-}4}F^6$/$A^4Na^{3\text{-}4}F^6$	19, 20, 46, 47	–	2.8	–	–	–	–	**37.7**	**100.0**	**24.0**	**23.6**
3282.25	$H_7N_4F_1$	$A^4A^4F^6$ + 2 Hex[e)]	43	–	1.7	–	–	3.6	–	–	–	–	–
3394.28	$H_5N_4S_2$	$Na^{6\text{-}4}Na^{6\text{-}4}$/$Na^{3\text{-}4}Na^{3\text{-}4}$	31, 48	–	–	–	–	–	–	3.5	–	–	–
3452.33	$H_5N_5F_1S_1$	$A^4Na^{6\text{-}4}F^6$bi	25 (24)	–	–	–	–	–	–	–	–	0.9	0.1
3540.34	$H_5N_4F_1S_2$	$Na^{6\text{-}4}Na^{6\text{-}4}F^6$/$Na^{3\text{-}4}Na^{3\text{-}4}F^6$	21, 49	–	0.9	–	–	–	–	**35.7**	**10.9**	–	–

The third column gives the most likely interpretation of the compositions in terms of *N*-glycans occurring on antibodies (Fig. 1). The numbers in the following rows give the percentages of each glycan species normalized to the most abundant glycan structure as deduced from peak height in the deconvoluted spectrum. Signals derived from in-source fragments have not been considered in this list. Glycans amounting to more than 5% relative to the major signal are printed bold. The SDs in abundance of these major compounds were smaller than 1% (relative to the major compound) as determined for Rituximab and 4E10 CHO.

a) Ion masses for gp1a = fully cleaved glycopeptide from human IgG.

b) Monosaccharides added to peptide, H = hexose, N = *N*-acetylhexosamine, F = fucose, S = sialic acid.

c) Identified as $Man5GnF^6$ in Herceptin and as $MA^{a3\text{-}4}F^6$ in Zenapax (see Fig. 9).

d) Contains also glycopeptides from major IgG3 variant Q8N4Y9.

e) Presumably α1,3-linked galactose, amounts inferred from PGCC data.

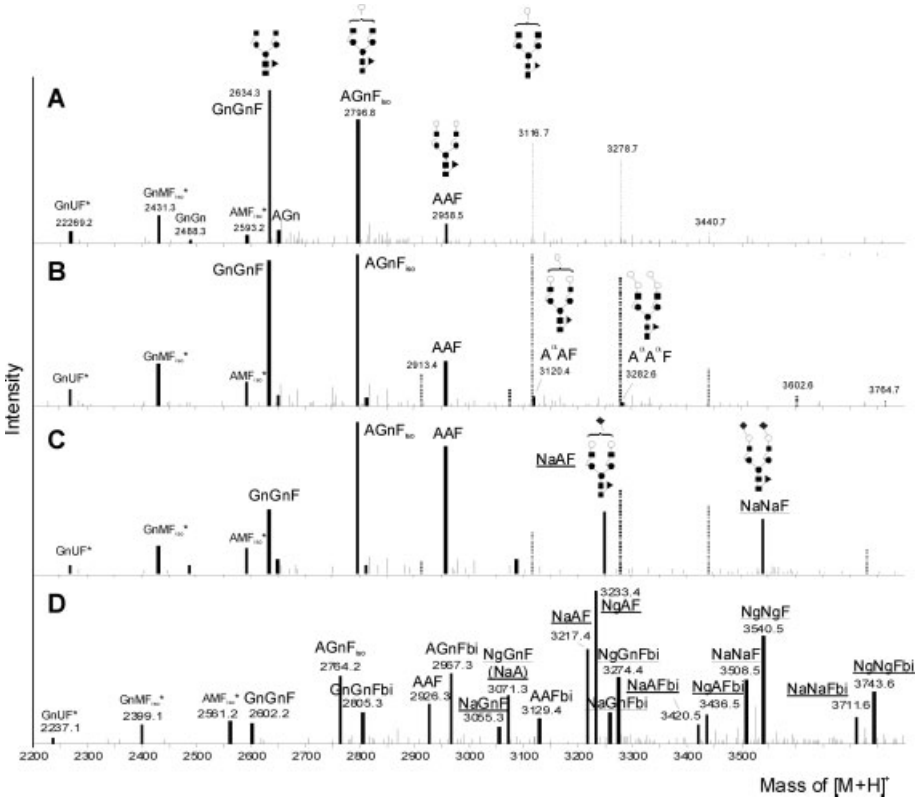

Figure 4. Glycopeptides from three mAbs and from guinea pig IgG: the deconvoluted spectra of (A) the glycopeptides from Rituximab; (B) Daclizumab (Zenapax); (C) CHO-cell produced 4E10; and (D) guinea pig IgG are shown. See also legend to Fig. 3. Note the absence of sialylated glycans in A, the occurrence of "hypergalactosylated" species in B, and the dominance of sialylated glycans in D (Na and Ng standing for *N*-acetyl- and *N*-glycolylneuraminic acid, respectively).

the form of *N*-glycolylneuraminic acid; which is in perfect agreement with a recently published, large and thorough study on Cetuximab glycosylation [19]. The variable region's glycopeptide spectrum indicated a very similar glycan profile as was obtained by Qian *et al.* [19] using NP-HPLC of free glycans after separation of the mAb's Fab and Fc fragments.

Actually, the site-specific glycopeptide data obtained from one SDS-PAGE band presented even more low abundance species, *e.g.*, a triantennary, trisialylated glycan (Fig. 5), and yet only required a single LC-MS run.

Finally, the remarkable case of guinea pig IgG shall be presented. In striking contrast to other species' IgG, these small animals furnish their antibodies with mainly sialylated, often disialylated *N*-glycans and a good portion of bisecting GlcNAc (Fig. 4). The glycan structures of guinea pig IgG have been reported earlier [9], but the separate analysis of neutral and sialylated glycans masked the strong prevalence of sialylated glycans.

3.3 RPC-MS as a tool for the quantification of glycan structures

The reproducibility of the RPC-MS analysis was evaluated by the parallel processing of three individual SDS-PAGE gel bands from 1 μg of CHO-4E10 each. With the intensities of the major peak (*i.e.*, monogalactosylated, fucosylated glycan) set to 100%, the SD of the other peaks was 0.5% or less for all components except for the essentially artifactual peaks of smaller m/z. Although these peaks accounted for remarkable 15% of the total signal, we assume their distorting effect on the relative glycoform abundances to be mild, as these fragments most likely stem from all the larger glycopeptides.

Eventually, RPC-MS data were compared with those from NP-HPLC of AB-labeled glycans, a "standard method, in which the fluorescence correlates with the amounts of individual components" according to Wada *et al.* [13]. The knowledge of the structures occurring on 4E10 from RPC-MS and a detailed work on the elution order of IgG glycans in NP-HPLC [7] together allowed the assignment of the respective HPLC peaks with reasonable confidence. The simplicity of the mAbs glycan profile additionally minimized problems with overlapping peaks (Fig. 6A). The comparison revealed a high, but not perfect quantitative match (Fig. 6B). However imperfect, the match seems to be as good or better than that obtained in a comparison of NP-HPLC with MALDI, where no quantitation of sialylated glycans was presented [13]. Whatever the reason for the small differences may be, the search for the true value goes beyond the scope of this work. The NP-HPLC chromatogram shows a number of unidentified peaks, some of which may be non-fucosylated glycans, *e.g.*, GnGn and NaNa that cannot be found in the RPC-MS spectra.

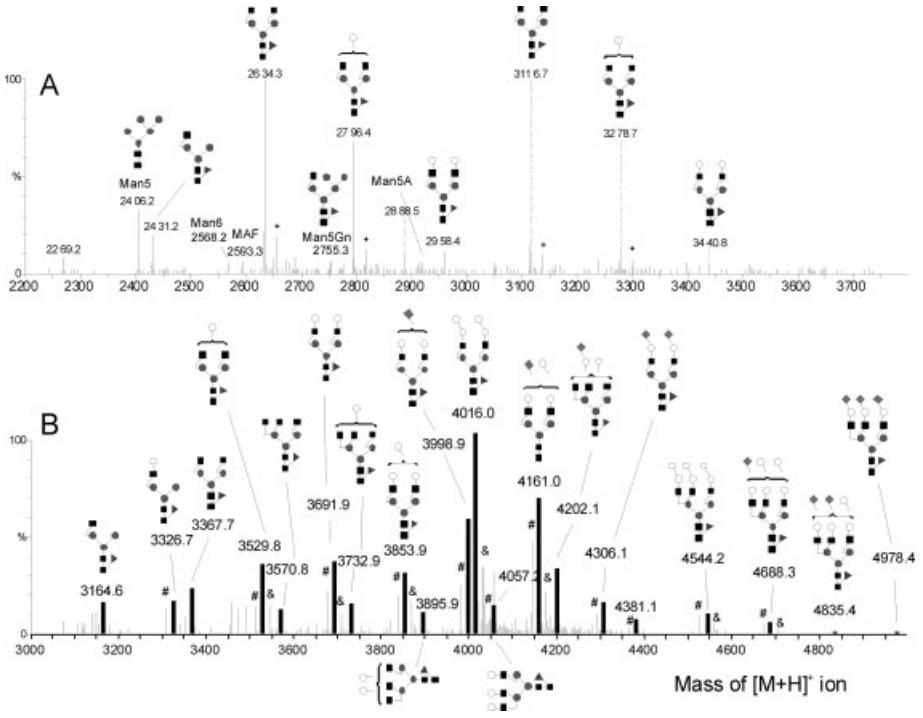

Figure 5. Site-specific glycosylation analysis of Cetuximab (Erbitux). (A) Glycopeptides from the constant region. (B) Glycopeptides from the variable region of the heavy chain displaying a drastically different glycan profile. Note the occurrence of triantennary glycans, the α1,3-galactosylated and the α2,3-sialylation. All sialic acids were *N*-glycolylneuraminic acids.

3.4 Oligosaccharide LC-ESI-MS

Our recent work on fibrin *N*-glycans revealed the ability of a buffered, acidic solvent to procure separation of isomeric glycans on a PGC column and concomitant sensitive detection of both neutral and sialylated glycans [35]. For this work, we established a retention time library for the relevant structures to be expected on antibodies produced in mammalian cells (*i.e.*, permutations of diantennary glycans with or without fucose, galactose on either arm, bisecting GlcNAc, and sialic acid in α2,6-linkage as typical for humans or α2,3-linkage as obtained, *e.g.*, in CHO cells (Figs. 1 and 7).

PGCC allowed separation of all isobaric structures occurring on the antibodies used in this study (Fig. 7). The rationale for assignment of the peaks was based on the literature data of human IgG glycans [7–13], on previous experiments as in the case of monosialylated nonfucosylated glycans [35], and on glycan mass. The assignment of isomeric glycans in the reference mixture shown in Fig. 7 was attained as follows. The two IgG glycans with one galactose residue were isolated and separately digested with β-*N*-acetylglucosaminidase and, after heat treatment, β-galactosidase and α-fucosidase as described previously and the products were compared with standards of known retention time [35].The major isomer yielded GnM and not MGn and thus was assigned to be A^4GnF^6 (structure **2**). A^4GnF^6 has been described as the prevailing isomer in human IgG [9, 39, 45]. In accordance with Raju *et al.* [9], bovine IgG exhibited the opposite distribution (data not shown). In the case of the bisected glycans (structures **6** and **14**), the assignment of galactose to the 6-arm is based on the reported almost exclusive occurrence of these isomers in human IgG [8, 11, 45]. In accordance with the strong arm preference of α2,6-sialyltransferase, only one α2,6-monosialylated isomer is present in porcine fibrin [35]. The other isomer $Na^{6\text{-}4}A^4F^6$ (structure **19**) was generated by desialylation of $Na^{6\text{-}4}Na^{6\text{-}4}F^6$ (**21**). Similarly, the α2,3-monosialylated glycans $Na^{3\text{-}4}A^4F^6$ and $A^4Na^{3\text{-}4}F^6$ (structures **46** and **47**) were obtained from disialylated erythropoeitin glycan. Their elution order was determined by fractionation followed by galactosidase treatment and desialylation. The glycan with α2,3-Neu5Ac on the 6-arm eluted earlier just as found for α2,6-sialylated structures with fucose and, previously, for nonfucosylated glycans [35]. The monogalactosylated peaks $Na^{3\text{-}4}GnF^6$ and $GnNa^{3\text{-}4}F^6$ (structures **44** and **45**) exhibited an increased peak spacing, which agrees with the effect of the galactose residues on elution time as seen with A^4GnF^6/GnA^4F^6 (structures **2** and **3**). Apparently, the elution order of monosialylated or monogalactosylated isomers remains the same irrespective of core α1,6-fucosylated. The elution order of the isomers A^4Gn and GnA^4 (structures **10** and **11**) and the monosialylated unfucosylated glycans have previously been determined by exo-glycosidase digestions [35].

Daclizumab is produced in a mouse cell line. Other mouse products have been found to contain α1,3-galactose residues [30]. Indeed, peaks interpretable as diantennary glycans with one or two α1,3-galactose residues appeared in the PGC chromatogram (Fig. 8). No definitive isomeric assignment was performed for the α1,3-galactosylated

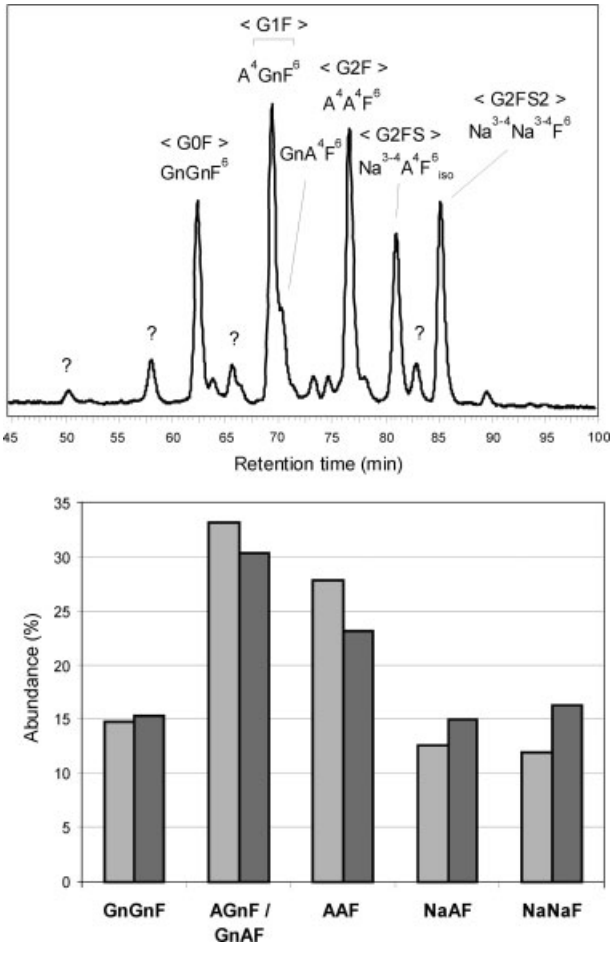

Figure 6. Comparison of PGC-MS with NP-HPLC: (A) chromatogram of AB-labeled glycans from CHO-derived 4E10. The peaks are labeled both with proglycan names and according to the "Oxford system" [7], where G stands for galactose, F for fucose, and S for sialic acid. A number of peaks (some marked "?") remained unidentified. Some of these could be defucosylation artifacts of major compounds. The bar encompasses the two isomers to which the term "G1F" applies. (B) Comparison of the quantities of the major compounds as obtained with the two methods. Dark bars represent the results from a single NP-HPLC run. Light bars are the results of three analyses (including SDS-PAGE) by RPC-MS with SDs of 0.6% or less (notably for the major peaks of a pure protein). Structure abbreviations are given without linkage superscripts.

glycans occurring in Daclizumab (Zenapax), but considering peak distances and heights in comparison with the A^4GnF^6/GnA^4F^6 (structures **2** and **3**) it is highly likely, that glycans with α1,3-galactose on the 6-arm eluted first. In the case of a glycan with the composition $Hex_5HexNAc_3dHex_1$, however, we dare to identify it as a hybrid-type glycan with an α1,3-galactose (Fig. 9). This assignment is based on the isomeric structure of the presumable precursors $MGnF^6$ and MA^4F^6, for which reference compounds could be prepared and the likely elution times of the isobaric glycans $Man5GnF^6$ (structure **38**) and $Man4A^4F^6$ (= **38** − Man + Gal). Interestingly, traces of α-galactosylation could be found in CHO-derived antibodies, especially Trastuzumab (Herceptin), by PGCC while they have been masked by neighboring signals in the RPC-ESI data of glycopeptides.

Once the elution order is established, the structural identities of all *N*-glycan structures on polyclonal or mAb can be deduced from the data of a single LC-MS experiment (Fig. 10), which also yields a quantitatively meaningful profile for both neutral and sialylated glycans (Data not shown; A comprehensive study on the relative signal intensity of neutral, mono- to tetrasialylated glycans is in preparation). Notable exceptions are larger glycans from the Fab-region of, *e.g.*, Cetuximab, for which no complete glycan library has yet been constructed.

The analysis of, *e.g.*, antibody 4E10 produced in either CHO or human lymphoma cells demonstrates the impact of the sialic acid linkage type on the elution position (Fig. 8).

ESI of glycans can induce some in-source fragmentation, the extent being inversely related to the size of the respective glycan [35]. The separation of glycans, however, allows the discrimination of fragments and real glycans. Additionally, in analogy to the situation with glycopeptides, fragments do only occur singly charged. A perspicuous example is the ample occurrence of singly charged ions having the mass of $MGnF^6$ (structure **33**) coinciding with the major peaks $GnGnF^6$, $AGnF^6$ (**1**, **2**), *etc.* in all samples. However, only in some mAbs, doubly charged ions at the elution position of $MGnF^6$ (**33**) were observed (data not shown but considered in Table 1). The exclusive occurrence of this isomer indicates lack of action of GlcNAc-transferase II rather than degradation by a later acting hexosaminidase as is found in plants or insect cells [46, 47]. While small low-mass peaks in glycopeptide (as in this work) or glycan spectra (as in [13]) are always suspect of being in-source fragments, the occurrence of separate peaks for the rare species of nonfucosylated bisected glycans underpins their actual existence in human IgG. Higher amounts of these structures are found in glycoengineered antibodies held to possess higher physiological efficacy [48].

Finally, it remains to be stated that the standard analysis started with 3 μg of antibody containing a calculated 20 pmol of *N*-glycans, of which one-third was injected. Despite the use of a column with a rather large diameter (0.32 mm), glycan species amounting to hardly more than 1% of the mixture (*e.g.*, A^4A^4bi = structure **16**) could be identified.

4 Discussion

In addition to the recently evaluated *N*-glycan profiling methods [13], we here describe two alternative, rather novel strategies.

The first, technically the common proteomic approach based on RP-LC-ESI-MS of tryptic (glyco-)peptides, has already been successfully used in our laboratory in order to

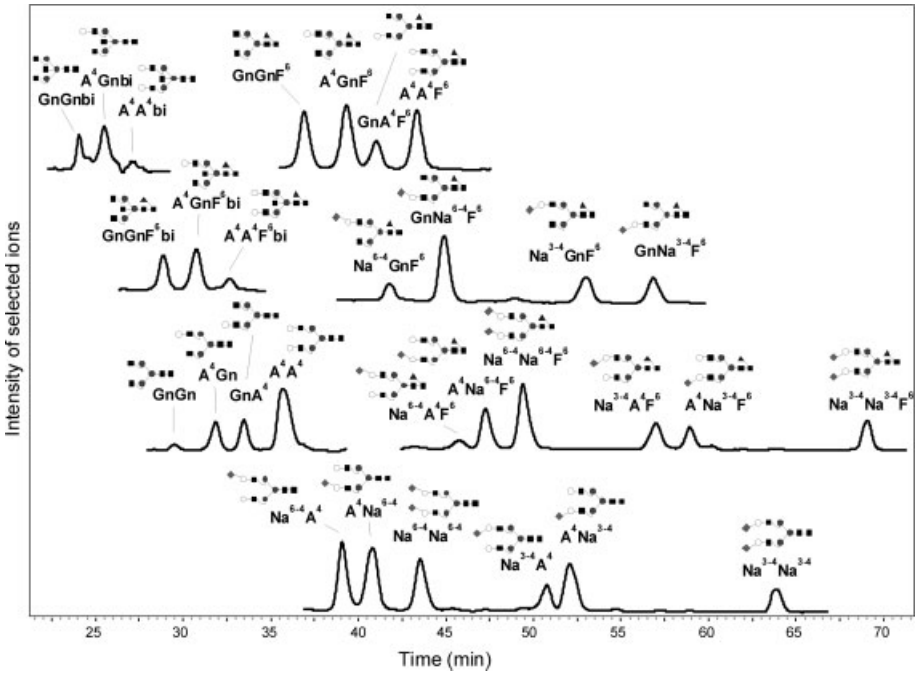

Figure 7. PGCC-ESI-MS of reference oligosaccharides. The elution order of reduced reference oligosaccharides was obtained from a single chromatographic experiment. The results are shown in groups comprising the SIM traces of related glycans with two or three different SIM masses, *i.e.*, the exact masses of the doubly charged ions.

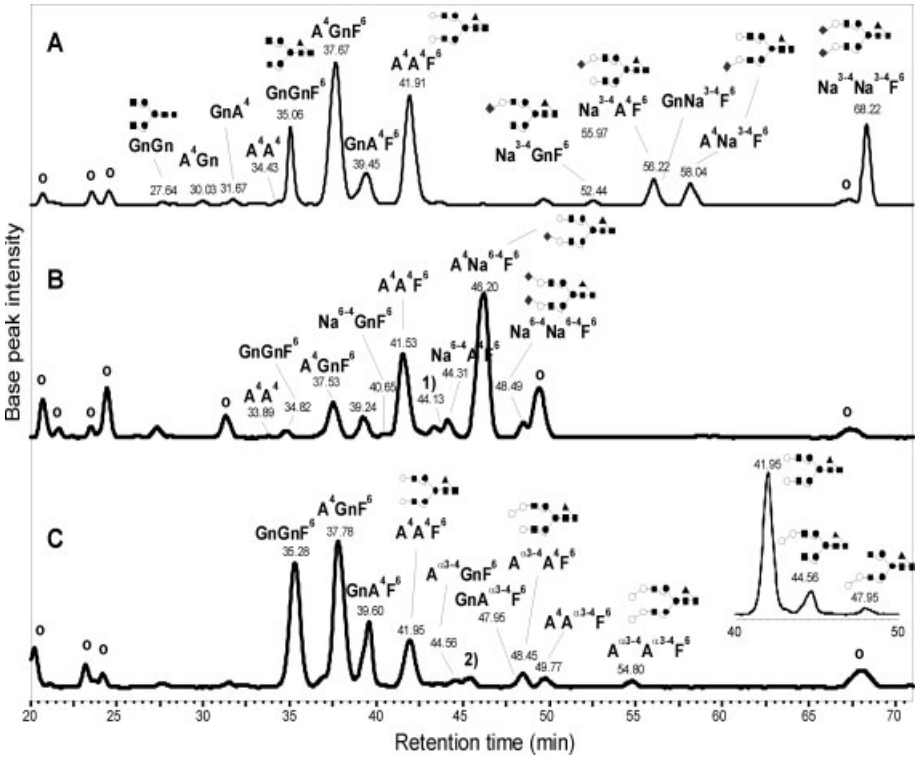

Figure 8. PGCC-MS of selected mAbs: the BPI chromatograms of the PGCC separations of *N*-glycans from (A) CHO-produced 4E10; (B) human lymphoma cell produced 4E10; and (C) daclizumab (Zenapax), which is made in mouse cells. The inset in (C) shows an SIM chromatogram for the doubly charged on of $A^4A^4F^6$ (structure **4**) eluting at 41.96 min, which in the case of Zenapax also picks up two isomers with α1,3-galactose. The tentative isomeric assignment in the inset is based on the similar isomer ratio of the substrates for this α-galactosylation. Small circles denote nonglycan peaks. The 1 in (B) denotes the peak of $GnNa^{6\text{-}4}F^6$ (structure **18**). Peak 2 in (C) exhibited the composition H_5N_3F, which from its elution time is interpreted as $MA^{a3\text{-}4}F^6$ (**40**, see Fig. 9) rather than $Man5GnF^6$ (structure **38**), which would elute at about 35 min rather than 46 min (see Fig. 9).

analyze various glycoproteins such as antibodies [5, 33] and others [36]. The major advantage of this strategy lies in the concomitant acquisition of additional information regarding underglycosylation, C-terminal (Lys-)truncations, N-terminal integrity and other potential protein chain modifications. The RPC-MS approach thus provides a comprehensive analysis of the molecule's primary structure, which proved particularly valuable for antibodies in the research pipeline; all the more when produced in novel expression systems [5, 33]. The method

was applied to bands from SDS-PAGE but protein A/G purified IgG may be an equally suitable starting material.

The complexity of the raw data may have detained potential users to follow this approach. Use of the expensive Lys-C [30, 34] or prefractionation of glycopeptides by hydrophilic interaction chromatography may lead to a substantial simplification of the LC-MS data [24, 49]. Although this has proven useful for other glycoproteins [24, 49], prefractionation is not necessary in the case of IgG. Above and beyond this extra step will almost certainly reduce sensitivity and possibly bias the glycan profile.

A second benefit of the RPC-MS method is, despite a certain degree of fragmentation of the glycopeptides, the true profiling of all components, including sialylated glycans. Even though glycopeptides are prone to in-source fragmentation (predominantly occurring at GlcNAc–Man linkages), the fragmentation-derived artifacts can be identified in and eliminated from the spectra. After elimination of fragmentation artifacts, RPC-MS results compare to those from NP-HPLC of fluorescently labeled glycans, and thus qualify RPC-MS as a reproducible quantitative method. However, it still remains to be elucidated whether the ratio of neutral to sialylated glycans is portrayed accurately, or if the use of correction factors is mandatory, - a reservation which pertains to other methods including NP-HPLC as well.

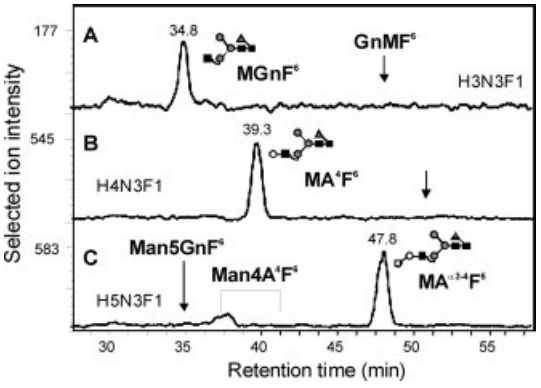

Figure 9. Detail from the oligosaccharide analysis of Daclizumab (Zenapax): Structural assignment of hybrid-type *N*-glycans. (A) The SIM trace of $MGnF^6$/$GnMF^6$ (structures **32**, **33**), where one peak appeared at the position of $MGnF^6$ (and not at that of $GnMF^6$, which both were enzymatically generated from $GnGnF^6$). (B) The SIM trace for the mass of MA^4F^6/A^4MF^6 (and theoretically $Man4GnF^6$ = structure **38** − Man + Gal) showing, again only one of the isomers obtained by *N*-acetylglucosaminidase digestion of an A^4GnF^6/GnA^4F^6 mixture (structures **2** and **3**). The arrows in A and B indicate the elution positions of the respective other isomers. Panel C is the SIM trace for $MA^{a3\text{-}4}F^6$ (structure **40**), which is, however, isobaric with the glycans $Man5GnF^6$ (structure **38**) or $Man4A^4F^6$ (= 38 − Man + Gal). The elution position of $Man5GnF^6$ as found in Herceptin is indicated as well as the elution region for $Man4A^4F^6$ as predicted from the usual retention time increment of a 3-arm galactose (±2 min uncertainty). The isomeric assignment of $MA^{a3\text{-}4}F^6$ is corroborated by the exclusive availability of MA^4F^6 as the substrate for α1,3-galactosyltransferase.

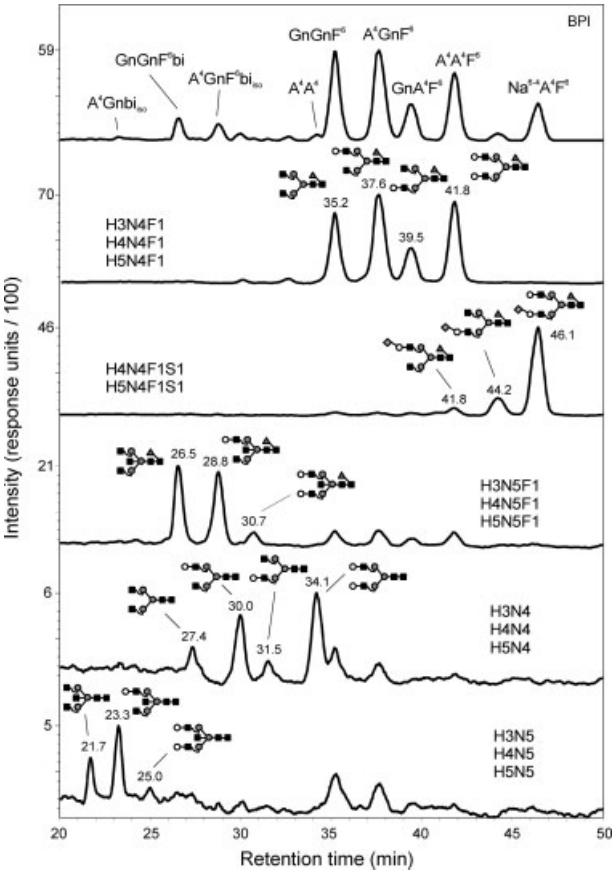

Figure 10. PGCC-MS analysis of IgG. The upper trace shows the BPI chromatogram of a sample of reduced *N*-glycans from human polyclonal IgG subjected to PGC separation with MS detection. The following lines are SIM chromatograms showing the intensity of the doubly charged ions for distinct groups of *N*-glycans. The lettering indicates the numbers of hexoses, *N*-acetylhexoses, fucoses, and sialic acids (H, N, F, and S, respectively) comprised by the SIM masses. The cartoons depict the real structure of the glycans in each peak. The relative intensity can be estimated from the *y*-axis labels.

Notably, while IgG-Fc glycopeptides with neutral and sialylated glycans eluted simultaneously from the BioBasic 18 column used in this study, a substantial separation was observed on an Acclaim PepMap100 C18 column (Dionex/LC-Packings, Sunnyvalley, CA, USA) in a very recent study [43].

As an elegant alternative to "traditional" multidimensional methods [10, 15, 23, 50] we here describe the application of PGC-LC-ESI-MS (PGCC-MS) for the full structural assignment of glycans occurring on IgG molecules. The PGC approach is a nondiscriminative method which likewise detects neutral and sialylated glycans, and even sulfated species (Pabst, unpublished results). Nevertheless, certain disadvantages are posed by the difficulties of identifying unusual structures among the hitherto unavoidable nonglycan peaks, and the larger Fab-domain glycans for which the elution positions of, *e.g.*, triantennary glycans or structures with Lewis fucoses have not yet

been acquired. Here, the combination of PGCC-MS with MS/MS experiments will prove valuable. Otherwise, MS/MS can be applied to corroborate the initial structural assignments made with the help of retention times.

The present study revealed vast differences in the glycosylation of recombinantly produced antibodies. Most remarkably, these differences have not only been detected between mAbs expressed in different cell lines (*i.e.*, CHO, SP2/0, NS0, and Burkitt lymphoma cells), but also within the large group of CHO cell-derived mAbs.

Furthermore, in contrast to all the investigated commercial mAbs, which are essentially devoid of sialic acid, the preclinical anti-HIV mAb 4E10 exhibited a considerable amount of sialylated glycans and contained generally more galactose. While an influence of fermentation conditions on sialylation has been reported [20], it remains to be shown why such drastic differences arise under conditions of mass production. The discovery of hybrid-type *N*-glycans in several mAbs (employing both methods) constitutes a nice detail that demonstrates the potency of these MS methods.

This work was to a large extent supported by the EC FP6 project "Pharma-Planta". We gratefully acknowledge the receipt of materials from Dr. Gottfried Himmler, Dr. Anton Stift, and Dr. Harald Höger and the skilful technical assistance of Ing. Karin Polacsek and Ing. Thomas Dalik.

The authors have declared no conflict of interest.

5 References

[1] Satoh, M., Iida, S., Shitara, K., Nonfucosylated therapeutic antibodies as next-generation therapeutic antibodies. *Expert Opin. Biol. Ther.* 2006, *6*, 1161–1173.

[2] Cox, K. M., Sterling, J. D., Regan, J. T., Gasdaska, J. R. *et al.*, Glycan optimization of a human monoclonal antibody in the aquatic plant Lemna minor. *Nat. Biotechnol.* 2006, *24*, 1591–1597.

[3] Li, H., Sethuraman, N., Stadheim, T. A., Zha, D. *et al.*, Optimization of humanized IgGs in glycoengineered Pichia pastoris. *Nat. Biotechnol.* 2006, *24*, 210–215.

[4] Ma, J. K., Drake, P. M., Chargelegue, D., Obregon, P. *et al.*, Antibody processing and engineering in plants, and new strategies for vaccine production. *Vaccine* 2005, *23*, 1814–1818.

[5] Schahs, M., Strasser, R., Stadlmann, J., Kunert, R. *et al.*, Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern. *Plant Biotechnol. J.* 2007, *5*, 657–663.

[6] Schuster, M., Jost, W., Mudde, G. C., Wiederkum, S. *et al.*, In vivo glyco-engineered antibody with improved lytic potential produced by an innovative nonmammalian expression system. *Biotechnol. J.* 2007, *2*, 700–708.

[7] Guile, G. R., Rudd, P. M., Wing, D. R., Prime, S. B. *et al.*, A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. *Anal. Biochem.* 1996, *240*, 210–226.

[8] Takahashi, N., Ishii, I., Ishihara, H., Mori, M. *et al.*, Comparative structural study of the N-linked oligosaccharides of human normal and pathological immunoglobulin G. *Biochemistry* 1987, *26*, 1137–1144.

[9] Raju, T. S., Briggs, J. B., Borge, S. M., Jones, A. J., Species-specific variation in glycosylation of IgG: Evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. *Glycobiology* 2000, *10*, 477–486.

[10] Saba, J. A., Kunkel, J. P., Jan, D. C., Ens, W. E. *et al.*, A study of immunoglobulin G glycosylation in monoclonal and polyclonal species by electrospray and matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Biochem.* 2002, *305*, 16–31.

[11] Holland, M., Yagi, H., Takahashi, N., Kato, K. *et al.*, Differential glycosylation of polyclonal IgG, IgG-Fc and IgG-Fab isolated from the sera of patients with ANCA-associated systemic vasculitis. *Biochim. Biophys. Acta* 2006, *1760*, 669–677.

[12] Krapp, S., Mimura, Y., Jefferis, R., Huber, R. *et al.*, Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J. Mol. Biol.* 2003, *325*, 979–989.

[13] Wada, Y., Azadi, P., Costello, C. E., Dell, A. *et al.*, Comparison of the methods for profiling glycoprotein glycans–HUPO human disease glycomics/proteome initiative multi-institutional study. *Glycobiology* 2007, *17*, 411–422.

[14] Kanda, Y., Yamada, T., Mori, K., Okazaki, A. *et al.*, Comparison of biological activity among nonfucosylated therapeutic IgG1 antibodies with three different N-linked Fc oligosaccharides: The high-mannose, hybrid, and complex types. *Glycobiology* 2007, *17*, 104–118.

[15] Geyer, H., Geyer, R., Strategies for analysis of glycoprotein glycosylation. *Biochim. Biophys. Acta* 2006, *1764*, 1853–1869.

[16] Parekh, R. B., Dwek, R. A., Sutton, B. J., Fernandes, D. L. *et al.*, Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* 1985, *316*, 452–457.

[17] Omtvedt, L. A., Royle, L., Husby, G., Sletten, K. *et al.*, Glycan analysis of monoclonal antibodies secreted in deposition disorders indicates that subsets of plasma cells differentially process IgG glycans. *Arthritis Rheum.* 2006, *54*, 3433–3440.

[18] Siemiatkoski, J., Lyubarskaya, Y., Houde, D., Tep, S. *et al.*, A comparison of three techniques for quantitative carbohydrate analysis used in characterization of therapeutic antibodies. *Carbohydr. Res.* 2006, *341*, 410–419.

[19] Qian, J., Liu, T., Yang, L., Daus, A. *et al.*, Structural characterization of N-linked oligosaccharides on monoclonal antibody cetuximab by the combination of orthogonal matrix-assisted laser desorption/ionization hybrid quadrupole-quadrupole time-of-flight tandem mass spectrometry and sequential enzymatic digestion. *Anal. Biochem.* 2007, *364*, 8–18.

[20] Kunkel, J. P., Jan, D. C., Butler, M., Jamieson, J. C., Comparisons of the glycosylation of a monoclonal antibody produced under nominally identical cell culture conditions in two different bioreactors. *Biotechnol. Prog.* 2000, *16*, 462–470.

[21] Weitzhandler, M., Hardy, M., Co, M. S., Avdalovic, N., Analysis of carbohydrates on IgG preparations. *J. Pharm. Sci.* 1994, *83*, 1670–1675.

[22] Muthing, J., Kemminer, S. E., Conradt, H. S., Sagi, D. *et al.*, Effects of buffering conditions and culture pH on production rates and glycosylation of clinical phase I anti-melanoma

mouse IgG3 monoclonal antibody R24. *Biotechnol. Bioeng.* 2003, *83*, 321–334.

[23] Mechref, Y., Novotny, M. V., Miniaturized separation techniques in glycomic investigations. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2006, *841*, 65–78.

[24] Wada, Y., Tajiri, M., Yoshida, S., Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal. Chem.* 2004, *76*, 6560–6565.

[25] Wuhrer, M., Catalina, M. I., Deelder, A. M., Hokke, C. H., Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2007, *849*, 115–128.

[26] Haslam, S. M., North, S. J., Dell, A., Mass spectrometric analysis of N- and O-glycosylation of tissues and cells. *Curr. Opin. Struct. Biol.* 2006, *16*, 584–591.

[27] Kuster, B., Wheeler, S. F., Hunter, A. P., Dwek, R. A. *et al.*, Sequencing of N-linked oligosaccharides directly from protein gels: In-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high-performance liquid chromatography. *Anal. Biochem.* 1997, *250*, 82–101.

[28] Papac, D. I., Wong, A., Jones, A. J., Analysis of acidic oligosaccharides and glycopeptides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Chem.* 1996, *68*, 3215–3223.

[29] Hirayama, K., Yuji, R., Yamada, N., Kato, K. *et al.*, Complete and rapid peptide and glycopeptide mapping of mouse monoclonal antibody by LC/MS/MS using ion trap mass spectrometry. *Anal. Chem.* 1998, *70*, 2718–2725.

[30] Sheeley, D. M., Merrill, B. M., Taylor, L. C., Characterization of monoclonal antibody glycosylation: Comparison of expression systems and identification of terminal alpha-linked galactose. *Anal. Biochem.* 1997, *247*, 102–110.

[31] Beck, A., Bussat, M. C., Zorn, N., Robillard, V. *et al.*, Characterization by liquid chromatography combined with mass spectrometry of monoclonal anti-IGF-1 receptor antibodies produced in CHO and NS0 cells. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2005, *819*, 203–218.

[32] Rouwendal, G. J., Wuhrer, M., Florack, D. E., Koeleman, C. A. *et al.*, Efficient introduction of a bisecting GlcNAc residue in tobacco N-glycans by expression of the gene encoding human N-acetylglucosaminyltransferase III. *Glycobiology* 2007, *17*, 334–344.

[33] Van Droogenbroeck, B., Cao, J., Stadlmann, J., Altmann, F. *et al.*, Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic Arabidopsis seeds. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 1430–1435.

[34] Huang, L., Biolsi, S., Bales, K. R., Kuchibhotla, U., Impact of variable domain glycosylation on antibody clearance: An LC/MS characterization. *Anal. Biochem.* 2006, *349*, 197–207.

[35] Pabst, M., Bondili, J. S., Stadlmann, J., Mach, L. *et al.*, Mass + Retention Time = Structure: A Strategy for the Analysis of N-Glycans by Carbon LC-ESI-MS and Its Application to Fibrin N-Glycans. *Anal. Chem.* 2007, *79*, 5051–5057.

[36] Kolarich, D., Weber, A., Turecek, P. L., Schwarz, H. P. *et al.*, Comprehensive glyco-proteomic analysis of human alpha1-antitrypsin and its charge isoforms. *Proteomics* 2006, *6*, 3369–3380.

[37] Packer, N. H., Lawson, M. A., Jardine, D. R., Redmond, J. W., A general approach to desalting oligosaccharides released from glycoproteins. *Glycoconj. J.* 1998, *15*, 737–747.

[38] Zhang, X., Asara, J. M., Adamec, J., Ouzzani, M. *et al.*, Data preprocessing in liquid chromatography-mass spectrometry-based proteomics. *Bioinformatics* 2005, *21*, 4054–4059.

[39] Bigge, J. C., Patel, T. P., Bruce, J. A., Goulding, P. N. *et al.*, Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. *Anal. Biochem.* 1995, *230*, 229–238.

[40] Kolarich, D., Turecek, P. L., Weber, A., Mitterer, A. *et al.*, Biochemical, molecular characterization, and glycoproteomic analyses of alpha(1)-proteinase inhibitor products used for replacement therapy. *Transfusion* 2006, *46*, 1959–1977.

[41] Rademacher, T., Sack, M., Arcalis, E., Stadlmann, J. *et al.*, Recombinant antibody 2G12 produced in maize endosperm efficiently neutralizes HIV-1 and contains predominantly single-GlcNAc N-glycans. *Plant Biotechnol. J.* 2008, *6*, 190–201.

[42] Lefranc, M. P., Giudicelli, V., Kaas, Q., Duprat, E. *et al.*, IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res.* 2005, *33*, D593–D597.

[43] Wuhrer, M., Stam, J. C., van de Geijn, F. E., Koeleman, C. A. *et al.*, Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum. *Proteomics* 2007, *7*, 4070–4081.

[44] Abel, C. A., Spiegelberg, H. L., Grey, H. M., The carbohydrate contents of fragments and polypeptide chains of human gamma-G-myeloma proteins of different heavy-chain subclasses. *Biochemistry* 1968, *7*, 1271–1278.

[45] Jefferis, R., Lund, J., Mizutani, H., Nakagawa, H. *et al.*, A comparative study of the N-linked oligosaccharide structures of human IgG subclass proteins. *Biochem. J.* 1990, *268*, 529–537.

[46] Leonard, R., Rendic, D., Rabouille, C., Wilson, I. B. *et al.*, The Drosophila fused lobes gene encodes an N-acetylglucosaminidase involved in N-glycan processing. *J. Biol. Chem.* 2006, *281*, 4867–4875.

[47] Strasser, R., Bondili, J. S., Schoberer, J., Svoboda, B. *et al.*, Enzymatic properties and subcellular localization of Arabidopsis beta-N-acetylhexosaminidases. *Plant Physiol.* 2007, *145*, 5–16.

[48] Ferrara, C., Stuart, F., Sondermann, P., Brunker, P. *et al.*, The carbohydrate at FcgammaRIIIa Asn-162. An element required for high affinity binding to nonfucosylated IgG glycoforms. *J. Biol. Chem.* 2006, *281*, 5032–5036.

[49] Takegawa, Y., Deguchi, K., Ito, H., Keira, T. *et al.*, Simple separation of isomeric sialylated N-glycopeptides by a zwitterionic type of hydrophilic interaction chromatography. *J. Sep. Sci.* 2006, *29*, 2533–2540.

[50] Takahashi, N., Three-dimensional mapping of N-linked oligosaccharides using anion-exchange, hydrophobic and hydrophilic interaction modes of high-performance liquid chromatography. *J. Chromatogr. A* 1996, *720*, 217–225.

[51] Takegawa, Y., Deguchi, K., Nakagawa, H., Nishimura, S., Structural analysis of an N-glycan with "beta1-4 bisecting branch" from human serum IgG by negative-ion MSn spectral matching and exoglycosidase digestion. *Anal. Chem.* 2005, *77*, 6062–6068.

Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern.

Schahs, M; Strasser, R; Stadlmann, J; Kunert, R; Rademacher, T and Steinkellner, H (2007). Plant Biotechnol J 5(5): 657-63.

Plant Biotechnology Journal (2007) **5**, pp. 657–663 doi: 10.1111/j.1467-7652.2007.00273.x

Production of a monoclonal antibody in plants with a humanized *N*-glycosylation pattern

Matthias Schähs[1], Richard Strasser[1], Johannes Stadlmann[2], Renate Kunert[3], Thomas Rademacher[4] and Herta Steinkellner[1,*]

[1]*Institute of Applied Genetics and Cell Biology, BOKU-Wien, Austria*
[2]*Department of Chemistry, BOKU-Wien, Austria*
[3]*Institute of Applied Microbiology, BOKU-Wien, Austria*
[4]*Rheinisch-Westfälische Technische Hochschule (RWTH), Aachen, Germany*

Received 2 March 2007;
revised 5 May 2007;
accepted 10 May 2007.
**Correspondence* (fax 43 1 36006 6392;
e-mail herta.steinkellner@boku.ac.at)

Keywords: antibodies, β1,2-xylose and core α1,3-fucose, plant glycan engineering, recombinant proteins.

Summary

In recent years, plants have become an attractive alternative for the production of recombinant proteins. However, their inability to perform authentic mammalian *N*-glycosylation may cause limitations for the production of therapeutics. A major concern is the presence of β1,2-xylose and core α1,3-fucose residues on complex *N*-linked glycans, as these *N*-glycan epitopes are immunogenic in mammals. In our attempts towards the humanization of plant *N*-glycans, we have generated an *Arabidopsis thaliana* knockout line that synthesizes complex *N*-glycans lacking immunogenic xylose and fucose epitopes. Here, we report the expression of a monoclonal antibody in these glycan-engineered plants that carry a homogeneous mammalian-like complex *N*-glycan pattern without β1,2-xylose and core α1,3-fucose. Plant and Chinese hamster ovary (CHO)-derived immunoglobulins (IgGs) exhibited no differences in electrophoretic mobility and enzyme-linked immunosorbent specificity assays. Our results demonstrate the feasibility of a knockout strategy for *N*-glycan engineering of plants towards mammalian-like structures, thus providing a significant improvement in the use of plants as an expression platform.

Introduction

Although the number of interesting recombinant glycoproteins for human therapy is steadily increasing, the current production capacities act as a bottleneck. Therefore, alternative production systems are currently the subject of intense research. Over recent years, plants have been considered as an attractive vehicle for the expression of recombinant pharmaceutical proteins as they are inexpensive and versatile systems, amenable to rapid and economical scale-up. Production levels of up to 60% of total soluble protein (TSP) have been reported recently, which makes the system of economic interest (De Jaeger *et al*., 2002; Marillonnet *et al*., 2004, 2005). A major advantage of plants is that they are higher eukaryotes that possess an endomembrane system and a secretory pathway similar to that of mammalian cells. Therefore, complex proteins are correctly folded and assembled, which has been demonstrated very impressively for a secretory immunoglobulin A (IgA) produced in tobacco (Ma *et al*., 1995). In addition, the increasing awareness of regulatory agencies of correct protein glycosylation requires a controlled glycosylation machinery. In this respect, plants are interesting, as they synthesize a core complex *N*-glycan structure similar to that in mammalian cells (Figure 1). However, by contrast with the common core structure, the terminal residues of plant and mammalian complex *N*-glycans differ; plant *N*-glycans are generally much smaller, lack β1,4-galactose and sialic acid, and, instead, carry plant-specific β1,2-xylose and core α1,3-fucose residues. However, the immunogenicity of β1,2-xylose and core α1,3-fucose residues in mammals is well documented (Kurosaka *et al*., 1991; Faye *et al*., 1993; Bardor *et al*., 2003; Bencurova *et al*., 2004; Jin *et al*., 2006). Furthermore, altered *N*-glycosylation may affect the function of a protein (for a recent review, see Jefferis, 2005). Thus, the unique *N*-glycans added by plants could impact on both the immunogenicity and functional activity, and, consequently,

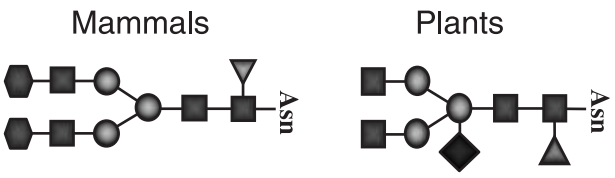

Figure 1 Schematic representation of complex *N*-glycans from mammals and plants (simplified). ■, *N*-acetylglucosamine; ●, mannose; ◆, β1,2-xylose; ▲, α1,3-fucose; ▼, α1,6-fucose; ⬢, α1,4-galactose.

may represent a limitation for plants to be used as a production platform.

Various strategies have been applied to avoid plant-specific *N*-glycosylation. In this respect, targeting of proteins to specific subcellular compartments with defined *N*-glycan structures has been reported (Schouten *et al*., 1996). For example, retention of recombinant proteins in the endoplasmic reticulum has resulted in their accumulation, carrying mainly oligo-mannosidic *N*-glycans, typical of endoplasmic reticulum-resident proteins. However, these structures may lead to a dramatic decrease in the *in vivo* half-life of the target protein, as recently reported for a plant-produced antibody (Ko *et al*., 2003). Another promising strategy is based on the specific manipulation of the *N*-glycosylation pathway in host plants. The over-expression of human β1,4-galactosyltransferase, which competes for the same acceptor substrate as β1,2-xylosyltransferase (XT) and core α1,3-fucosyltransferase (FT), resulted in a significant decrease in β1,2-xylose and core α1,3-fucose (Palacpac *et al*., 1999; Bakker *et al*., 2001, 2006). However, the complete elimination of these glycan epitopes was not achieved. Nevertheless, it is possible to partially elongate plant *N*-glycans with β1,4-galactose, a terminal residue present on many mammalian *N*-linked glycans, but absent in plants. The importance of this glycan residue is twofold: first, it may affect the stability and/or *in vivo* activity of a protein and, second, it is the prerequisite structure for the attachment of sialic acid.

Recently, a knockout (k.o.) line from the model plant *Arabidopsis thaliana* was generated in our laboratory, with deficiency of active XT and core FT, the enzymes responsible for the transfer of β1,2-xylose and core α1,3-fucose (Strasser *et al*., 2004). This line (XT/FT k.o. line) lacks immunogenic β1,2-xylose and core α1,3-fucose residues and, instead, synthesizes predominantly humanized structures with terminal *N*-acetylglucosamine (GlcNAc) residues (GnGn). Interestingly, XT/FT k.o. plants are viable and reveal no obvious morphological phenotype under standard growth conditions. In addition, an RNA interference (RNAi) strategy has been applied to eliminate xylose and fucose residues in the aquatic plant *Lemna minor* (Cox *et al*., 2006). An IgG produced in these 'glycan-optimized' plants exhibited a homogeneous complex *N*-glycan (GnGn) structure which showed an up to 50-fold increase in antibody-dependent cellular cytotoxicity (ADCC). Core α1,3-fucose seems to have a similar effect to mammalian α1,6-fucose, as the removal of this residue in Chinese hamster ovary (CHO)-derived antibodies dramatically increases Fcγ receptor binding and ADCC activity (Shields *et al*., 2001).

Here, we describe the feasibility of *A. thaliana* XT/FT k.o. lines for the production of recombinant proteins, with a monoclonal antibody as a model glycoprotein. K.o. lines were transformed with genes from heavy and light chains, and selected according to their IgG expression. As expected, antibodies purified from the mutant line carry complex *N*-glycan structures, but completely lack xylose and fucose, as determined by immunoblotting and mass spectrometry (MS). The IgG obtained assembled correctly and exhibited an unaltered antigen-binding affinity compared with CHO-derived IgG.

Results and discussion

Expression and characterization of 2G12 in *A. thaliana* wild-type (wt) and XT/FT k.o. plants

We have recently generated an *A. thaliana* triple k.o. line (XT/FT k.o.) that synthesizes complex *N*-glycans completely devoid of immunogenic β1,2-xylose and core α1,3-fucose (Strasser *et al*., 2004). A monoclonal antibody of the class IgG1 against human immunodeficiency virus (HIV) (2G12; Trkola *et al*., 1995) was chosen to investigate the feasibility of these k.o. lines for the production of recombinant proteins. The two complementary DNAs (cDNAs) were cloned in tandem into a binary vector (Figure 2a) that also contained DsRed, a fluorescent protein, as selection marker. *A. thaliana* wt and XT/FT k.o. mutants were transformed, and the resulting seeds were monitored according to their DsRed expression using fluorescence microscopy. Seeds that exhibited a strong DsRed signal were chosen for further propagation (Figure 2b). To test whether 2G12 antibodies were actually expressed in these plants, soluble proteins were extracted from leaves and subjected to Western blot analysis using antibodies against heavy and light chains (C_H, C_L). The expression of both chains was clearly visible in wt and XT/FT k.o. plants (Figure 3a). Subsequently, 2G12 antibodies were purified from leaves via protein A affinity chromatography, and subsequent sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis under reducing conditions revealed the presence of two bands that correlated with the expected molecular masses of C_H and C_L (55 kDa and 25 kDa, respectively; Figure 3b). No

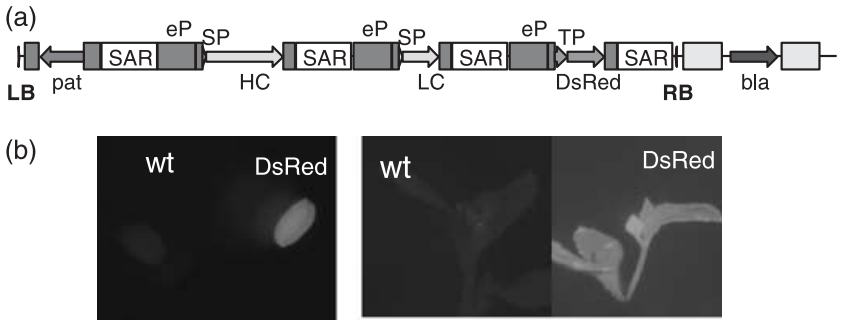

Figure 2 (a) Plant expression vector used for the expression of immunoglobulin G (IgG) antibodies: bla, ampicillin resistance gene; DsRed, red fluorescent protein; eP, enhanced cauliflower mosaic virus (CaMV)-35S promoter; HC, LC, heavy and light chain; LB, RB, left and right border; pat, basta resistance gene with promoter and terminator of the nopaline synthase gene; SAR, scaffold attachment region; SP, signal peptide; TP, transit peptide. (b) Seeds and seedlings examined under fluorescence microscopy: DsRed, seed transformed with the binary vector described above; wt, non-transformed seed.

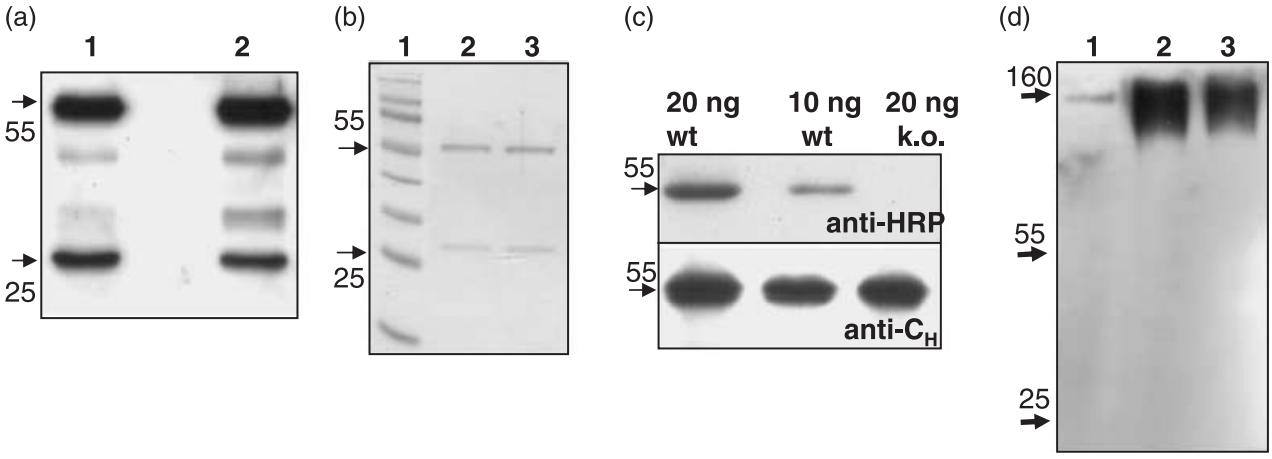

Figure 3 (a) Western blot analysis of proteins extracted from 2G12-transformed wild-type (wt) (lane 1) and β1,2-xylosyltransferase and α1,3-fucosyltransferase knockout (XT/FT k.o.) (lane 2) plants probed with rabbit anti-human immunoglobulin G (IgG) (C_H and C_L-specific). (b) Sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (Coomassie Blue stained): lane 1, protein marker; 2G12 purified from wt *Arabidopsis thaliana* (lane 2) and k.o. line (lane 3). (c) Western blot analysis of purified 2G12 using anti-horseradish peroxidase (anti-HRP) antibodies (recognizing xylose and fucose) and goat anti-human IgG-(γ-chain-specific, anti-C_H): wt, 2G12 derived from *A. thaliana* wt; k.o., 2G12 derived from XT/FT k.o. line. (d) Western blot analysis of purified 2G12 after native PAGE using rabbit anti-human IgG (C_H and C_L-specific): lane 1, Chinese hamster ovary (CHO)-derived 2G12; 2G12 derived from wt *A. thaliana* (lane 2) and k.o. line (lane 3). Numbers next to the gels indicate the protein masses in kilodaltons.

further bands, degradation products or impaired heavy or light chains in PAGE under native conditions were visible, indicating the purity and high rate of antibody assembly. The weak extra bands between C_H and C_L of 2G12 that were visible when TSP extracts were used for immunoblotting (Figure 3a) may represent degradation products from C_H within the plant cell; however, they disappeared completely after 2G12 purification. In addition, purified CHO- and plant-derived IgGs were subjected to immunoblotting using antibodies that recognized C_H- and plant-specific glycan epitopes [anti-horseradish peroxidase (anti-HRP) antibody], respectively (Figure 3c). As expected, a single band with a molecular mass of about 55 kDa was detected in all samples using C_H antibodies. By contrast, no signal was obtained in XT/FT k.o.-derived IgG using anti-HRP anti-serum, indicating the absence of immunogenic glycan epitopes.

The expression level of 2G12 in *A. thaliana* wt and k.o. lines varied between 0.2 and 0.05% of TSP in young plants (up to the six-leaf stage). As plants aged, the expression level dropped below 0.05% TSP. 2G12-expressing plants, both wt and k.o. lines, were normal and viable and did not show any obvious morphological phenotype under standard growth conditions.

In previous studies, it has been shown that 2G12 recognizes a specific epitope at the HIV envelope protein gp120 (Trkola *et al*., 1996). In order to test the antigen-binding specificities of plant-produced 2G12, a specificity enzyme-linked immunosorbent assay (ELISA) coated with HIV gp160, which encompasses gp120, was carried out. Compared with CHO-derived 2G12, the binding capacity of the plant-derived counterpart was 100% or slightly above, demonstrating unaltered antigen-binding specificity.

Determination of *N*-glycan composition of 2G12 from *A. thaliana* wt and XT/FT k.o. lines

To determine the exact *N*-glycan composition of 2G12, the corresponding heavy chains were subjected to tryptic digestion and MS. Owing to incomplete digestion of C_H, two glycopeptides, which contain the carbohydrate at asparagine-297 but differ in four amino acids (482.3 Da), are present. The absence of xylose and fucose residues on *N*-glycans can be monitored by the decrease in mass of the respective peaks (132 mass units for xylose and 146 mass units for fucose).

The mass spectrum of wt-derived C_H exhibited three major peaks for both glycopeptides (Figure 4a). The major peak was derived from the complex *N*-glycan GnGnXF with two terminal GlcNAc residues, followed by two minor peaks with one GlcNAc residue (GnMXF/MGnXF) and without any terminal GlcNAc residues (MMXF). All of these detected glycoforms contained immunogenic β1,2-xylose and core α1,3-fucose residues. By contrast, XT/FT k.o.-derived 2G12 lacked the three major peaks carrying β1,2-xylose and core α1,3-fucose (Figure 4b). Instead, GnGn was the most abundant *N*-glycan structure. Apart from the complex *N*-glycans, two minor glycopeptides with oligomannosidic structures (Man7, Man8) were identified in both wt- and XT/FT k.o.-derived IgG, which indicates that not all glycans were completely processed to complex-type forms. *N*-Glycan analysis of CHO-derived 2G12 revealed a mixture of structures containing either no (GnGnF), one (AGnF, GnAF) or two (AAF) terminal galactose residues and, in addition, minor amounts of glycopeptides with sialic acid (NaAF, NaNaF). Interestingly, all detected glycoforms contained a fucose residue linked α1,6 to the innermost GlcNAc. Importantly, it has been shown that the absence of this fucose residue significantly increases Fc receptor-mediated ADCC (Shields *et al*., 2001).

Our results clearly demonstrate that engineering of complex plant *N*-glycans towards a mammalian type is possible without any deleterious effects in growth behaviour. We have shown that glyco-engineered plants are able to produce monoclonal antibodies with a defined humanized *N*-glycan profile, not only lacking plant-specific glycan epitopes, but also α1,6-fucose, a glycan residue that seems to negatively influence the ADCC potency of antibodies (Shields *et al*., 2001). Indeed, recently, it has been demonstrated very impressively that such 'glycan-optimized' plant-derived antibodies carrying a homogeneous GnGn structure exhibit a significant ADCC enhancement compared with the CHO-produced counterpart (Cox *et al*., 2006). These results and recent improvements in plant expression systems using *A. thaliana* seeds and *Nicotiana* species, where IgG and antibody fragment levels of up to 0.5 g/kg leaf material were obtained (Giritch *et al*., 2006; Van Droogenbroeck *et al*., 2007), demonstrate the great potential of plants as an alternative expression system for therapeutics, not only in terms of quantity, but also in quality improvement.

Experimental procedures

Expression of 2G12 antibody in *A. thaliana* plants

The cDNAs of 2G12 heavy and light chains were isolated, together with the sequence for the signal peptide, from their eukaryotic expression vectors (Trkola *et al*., 1995). Each chain was subcloned into an expression cassette consisting of the duplicated cauliflower mosaic virus (CaMV)-35S promoter, the tobacco etch virus (TEV) 5′-untranslated region and the CaMV-35S transcriptional terminator. For co-expression of heavy and light chains, the two cassettes were inserted head-to-tail into the binary vector pTRAp, a derivative of pPAM (GenBank: AY027531). The tandem expression units were separated by scaffold attachment regions (SARs) of the tobacco RB7 gene (GenBank: U67919). In addition, an expression cassette for DsRed (Jach *et al*., 2001), fused to a plastid transit peptide, was cloned behind the antibody chain cassettes to create the final construct pTRAp-2G12-Ds (Figure 2a). The plasmid was transformed into *Agrobacterium* strain GV3101::pMP90RK.

Transgenic *A. thaliana* plants were generated by floral dipping of Col-0 wt and XT/FT k.o. lines (Strasser *et al*., 2004). A Leica MZ FLIII stereomicroscope (Leica, Wetzlar, Germany), equipped with a Leica DC 500 digital camera and a DsRed filter set (No. 10447079), was used to detect DsRed fluorescence in seeds and young seedlings.

Purification of recombinant 2G12 antibody from *A. thaliana leaves*

Leaves (200 mg) were frozen in liquid nitrogen and ground in a mixer mill (Retsch MM2000, Haan, Germany). The powder was dissolved in 400 μL of 1 × phosphate-buffered saline (PBS) (pH 6) containing 1% protease (v/v) inhibitor cocktail (Sigma, P9599, St. Louis, MO, USA), and incubated at 4 °C for 15 min. After centrifugation for 30 min at 16 000 ***g***, the supernatant (SN1) was incubated at 4 °C for 90 min with 15 μL of rProteinA Sepharose Fast Flow (GE Healthcare, Chalfont, UK) using an orbital shaker. The incubated slurry was transferred into a Micro Bio-Spin chromatography column (Biorad, Hercules, CA, USA) and washed three times with 250 μL of 1 × PBS. Elution was performed with 15 μL of 0.1 M glycine–HCl buffer (pH 3.0).

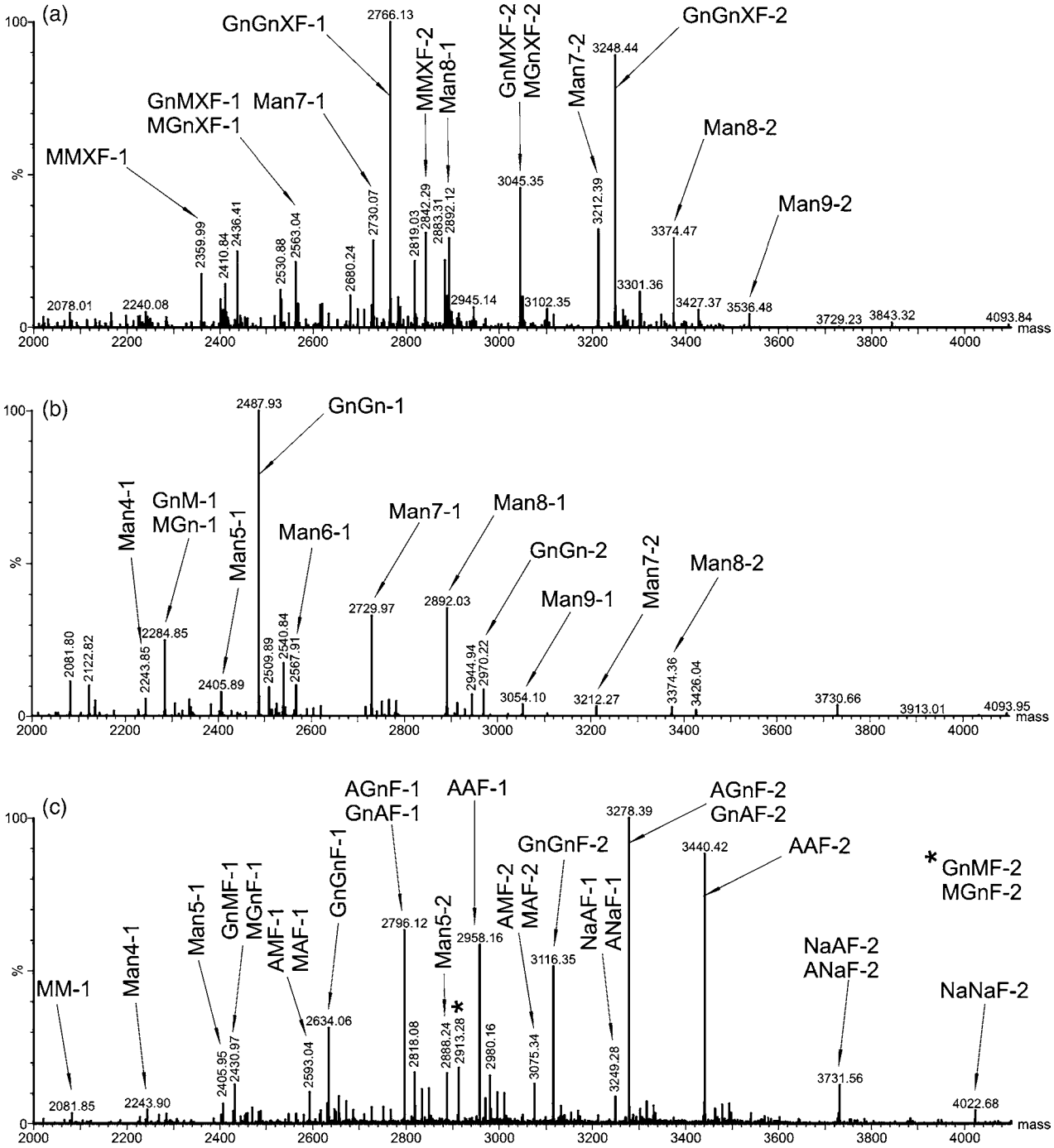

Figure 4 Liquid chromatography-mass spectrometry (LC-MS) of glycosylated peptides (-1, EEQY*N*STYR; -2, TKPREEQY*N*STYR) containing asparagine-297 from the 2G12 Fc region: (a) glycopeptides derived from wild-type (wt)-produced 2G12; (b) glycopeptides derived from β1,2-xylosyltransferase and α1,3-fucosyltransferase knockout (XT/FT k.o.) plant-produced 2G12; (c) glycopeptides derived from Chinese hamster ovary (CHO)-produced 2G12. See http://www.proglycan.com for an explanation of *N*-glycan abbreviations.

Immunoblot analysis

SN1 and purified IgG (100 ng) samples were subjected to SDS-PAGE (12.5% polyacrylamide) under reducing and non-reducing conditions. The separated proteins were blotted to Hybond-ECL membranes (GE Healthcare) and detected with either a goat anti-human IgG ($C_H + C_L$-specific) antibody conjugated to peroxidase (Promega, Madison, WI, USA) or a goat anti-human IgG (γ-chain-specific) antibody conjugated to peroxidase (Sigma, A8775), both diluted 1 : 5000 in 1 × PBS (pH 7.4) containing 1% (w/v) bovine serum albumin (BSA). For the detection of N-linked glycans with β1,2-xylose and core α1,3-fucose, rabbit anti-HRP antibody was used as described

previously (Strasser *et al*., 2004). The detection of bound antibodies was performed using SuperSignal West Pico chemiluminescent substrate (Pierce, Rockford, IL, USA).

2G12 quantification and antigen-binding assay

To quantify and determine the antigen-binding specificity of plant-derived 2G12, a Standard Operating Procedure (SOP) developed by Polymun Scientific was carried out. The IgG concentration and gp160 binding were determined by coating ELISA plates with a polyclonal antibody 'goat anti-human IgG' (γ-chain, Sigma; for quantification) and with 100 ng purified gp160 [provided by Polymun Scientific (Vienna, Austria); for antigen-binding assays]. After washing, 2G12 samples were diluted to a concentration of about 200 ng/mL. Subsequently, a 1 : 2 serial dilution was carried out (eight dilutions) and 50 μL was transferred per well. Recombinant 2G12 purified from CHO culture supernatants (provided by Polymun Scientific) was used as standard, and applied identically to plant-derived 2G12. Plates were incubated for 1 h at room temperature (gentle shaking) and, after washing, a polyclonal antibody 'goat anti-human IgG' (κ-chain) conjugated to alkaline phosphatase (1 : 1000, Sigma) was incubated. After intensive washing, alkaline phosphatase was detected by *p*-nitro-phenylphosphate substrate solution using an ELISA reader at 405 nm, with 620 nm as reference wavelength. All analyses were carried out in duplicate, and a four-parameter polynomial was applied for curve fitting and data analysis.

N-Glycan analysis

Purified 2G12 IgG1 (0.5 μg) was separated by SDS-PAGE (12.5% polyacrylamide) analysis under reducing conditions, and polypeptides were detected by Coomassie Blue staining. The heavy chain was excised from the gel, destained, carbamidomethylated and in-gel trypsin digested as described previously (Kolarich and Altmann, 2000). Tryptic peptides were dried in a Speed Vac concentrator and reconstituted with water containing 0.1% (v/v) formic acid. Mass spectrometric analysis was performed on a quadrupole time-of-flight (Q-TOF) Ultima Global (Waters Micromass, Milford, MA, USA) equipped with a standard electro-spray unit, a Cap-LC system (Waters Micromass) and a 10-port solvent switch module (Rheodyne, Rohnert Park, MA, USA). Samples were first captured by an Aquasil C18 precolumn (30 × 0.32 mm, Thermo Electron, Waltham, MA, USA) using water as the solvent. The analytical column was held at 5% acetonitrile before solvent switching, and then a linear gradient from 5 to 50% acetonitrile was applied at a flow rate of 2 μL/min. All eluents contained 0.1% formic acid. The mass spectrometer had been previously tuned with [Glu1]-fibrinopeptide B to give the highest possible sensitivity and a resolution of *c*. 10 000 (full width at half-maximum, FWHM). Mass tuning of the TOF analyser was performed in the tandem MS mode using, again, [Glu1]-fibrinopeptide B. Samples were analysed in the MS mode. Because no switching between MS and tandem MS mode was performed, no loss of signal, especially for the analysis of the glycopeptides, occurred. Data analysis was performed with MassLynx 4.0 SP4 Software (Waters Micromass).

The mass spectrometric data of tryptic peptides were analysed against the *in silico*-generated tryptic digestion of the 2G12 C_H amino acid sequence, employing the program 'PeptideMass' (http://www.expasy.org/tools/peptide-mass.html). Based on the tryptic peptide data set, the tryptic glycopeptide data sets ('glycopeptide 1' and 'glycopeptide 2', representing the 'perfectly' cleaved tryptic glycopeptide 'EEQYNSTYR', and the tryptic glycopeptide bearing one missed cleavage site 'TKPREEQYNSTYR', respectively) were generated by the addition of the respective glycan masses to the tryptic peptide masses of glycopeptide 1 and glycopeptide 2.

Acknowledgements

We thank Chunsheng Jin and Robert Weik (BOKU-Wien, Austria) for assistance in antibody purification, and Ali Assadian and Gabriele Stiegler (BOKU-Wien, Austria) for IgG quantification and gp160-binding assays. This work was supported by a Wiener Wirtschafts-, Forschung-, und Technologiefonds-funded project (www.wwtf.at) and the European Community (www.pharma-planta.com).

References

Bakker, H., Bardor, M., Molthoff, J.W., Gomord, V., Elbers, I., Stevens, L.H., Jordi, W., Lommen, A., Faye, L., Lerouge, P. and Bosch, D. (2001) Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl. Acad. Sci. USA*, **98**, 2899–2904.

Bakker, H., Rouwendal, G.J., Karnoup, A.S., Florack, D.E., Stoopen, G.M., van Helsper, J.P., van Ree, R., Die, I. and Bosch, D. (2006) An antibody produced in tobacco expressing a hybrid beta-1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc. Natl. Acad. Sci. USA*, **103**, 7577–7582.

Bardor, M., Faveeuw, C., Fitchette, A.C., Gilbert, D., Galas, L., Trottein, F., Faye, L. and Lerouge, P. (2003) Immunoreactivity in mammals of two typical plant glyco-epitopes, core alpha (1,3)-fucose and core xylose. *Glycobiology*, **13**, 427–434.

Bencurova, M., Hemmer, W., Focke-Tejkl, M., Wilson, I.B. and Altmann, F. (2004) Specificity of IgG and IgE antibodies against plant and insect glycoprotein glycans determined with artificial glycoforms of human transferrin. *Glycobiology*, **14**, 457–466.

Cox, K.M., Sterling, J.D., Regan, J.T., Gasdaska, J.R., Frantz, K.K., Peele, C.G., Black, A., Passmore, D., Moldovan-Loomis, C., Srinivasan, M., Cuison, S., Cardarelli, P.M. and Dickey, L.F. (2006) Glycan optimization of a human monoclonal antibody in the aquatic plant Lemna minor. *Nat. Biotechnol*. **24**, 1591–1597.

De Jaeger, G., Scheffer, S., Jacobs, A., Zambre, M., Zobell, O., Goossens, A., Depicker, A. and Angenon. G. (2002) Boosting heterologous protein production in transgenic dicotyledonous seeds using Phaseolus vulgaris regulatory sequences. *Nat. Biotechnol*. **20**, 1265–1268.

Faye, L., Gomord, V., Fitchette-Laine, A.C., and Chrispeels, M.J. (1993) Affinity purification of antibodies specific for Asn-linked glycans containing alpha 1→3 fucose or beta 1→2 xylose. *Anal. Biochem*. **209**, 104–108.

Giritch, A., Marillonnet, S., Engler, C., van Eldik, G., Botterman, J., Klimyuk, V. and Gleba, Y. (2006) Rapid high-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors. *Proc. Natl. Acad. Sci. USA*, **103**, 14 701–14 706.

Jach, G., Binot, E., Frings, S., Luxa, K. and Schell, J. (2001) Use of red fluorescent protein from Discosoma sp. (dsRED) as a reporter for plant gene expression. *Plant J*. **28**, 483–491.

Jefferis, R. (2005) Glycosylation of recombinant antibody therapeutics. *Biotechnol. Prog*. **21**, 11–16.

Jin, C., Bencurova, M., Borth, N., Ferko, B., Jensen-Jarolim, E., Altmann, F. and Hantusch, B. (2006) Immunoglobulin G specifically binding plant N-glycans with high affinity could be generated in rabbits but not in mice. *Glycobiology*, **16**, 349–357.

Ko, K., Tekoah, Y., Rudd, P.M., Harvey, D.J., Dwek, R.A., Spitsin, S., Hanlon, C.A., Rupprecht, C., Dietzschold, B., Golovkin, M. and Koprowski, H. (2003) Function and glycosylation of plant-derived antiviral monoclonal antibody. *Proc. Natl. Acad. Sci. USA*, **100**, 8013–8018.

Kolarich, D. and Altmann, F. (2000) N-Glycan analysis by matrix-assisted laser desorption/ionization mass spectrometry of electrophoretically separated nonmammalian proteins: application to peanut allergen Ara h 1 and olive pollen allergen Ole e 1. *Anal. Biochem*. **285**, 64–75.

Kurosaka, A., Yano, A., Itoh, N., Kuroda, Y., Nakagawa, T. and Kawasaki, T. (1991) The structure of a neural specific carbohydrate epitope of horseradish peroxidase recognized by anti-horseradish peroxidase antiserum. *J. Biol. Chem*. **266**, 4168–4172.

Leonard, R., Rendic, D., Wilson, I.B.H. and Altmann, F. (2005) Adaptation of the 'in gel release method' to glycom studies from a very little amount of material. *Glycoconjugate J*. **22**, 238.

Ma, J.K., Hiatt, A., Hein, M., Vine, N.D., Wang, F., van Stabila, P., Dolleveerd. C., Mostov, K. and Lehner, T. (1995) Generation and assembly of secretory antibodies in plants. *Science*, **268**, 716–719.

Marillonnet, S., Giritch, A., Gils, M., Kandzia, R., Klimyuk, V. and Gleba, Y. (2004) In planta engineering of viral RNA replicons: efficient assembly by recombination of DNA modules delivered by Agrobacterium. *Proc. Natl. Acad. Sci. USA*, **101**, 6852–6857.

Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V. and Gleba, Y. (2005) Systemic Agrobacterium tumefaciens-mediated transfection of viral replicons for efficient transient expression in plants. *Nat. Biotechnol*. **23**, 718–723.

Palacpac, N.Q., Yoshida, S., Sakai, H., Kimura, Y., Fujiyama, K., Yoshida, T. and Seki, T. (1999) Stable expression of human beta1,4-galactosyltransferase in plant cells modifies N-linked glycosylation patterns. *Proc. Natl. Acad. Sci. USA*, **96**, 4692–4697.

Schouten, A., Roosien, J., van Engelen, F.A., de Jong, G.A., Borst-Vrenssen, A.W., Zilverentant, J.F., Bosch, D., Stiekema, W.J., Gommers, F.J., Schots, A. and Bakker, J. (1996) The C-terminal KDEL sequence increases the expression level of a single-chain antibody designed to be targeted to both the cytosol and the secretory pathway in transgenic tobacco. *Plant Mol. Biol*. **30**, 781–793.

Shields, R.L., Lai, J., Keck, R., O'Connell, L.Y., Hong, K., Meng, Y.G., Weikert, S.H. and Presta, L.G. (2001) Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *J. Biol. Chem*. **277**, 26 733–26 740.

Strasser, R., Altmann, F., Mach, L., Glossl, J. and Steinkellner, H. (2004) Generation of Arabidopsis thaliana plants with complex N-glycans lacking beta 1,2-linked xylose and core alpha 1,3-linked fucose. *FEBS Lett*. **561**, 132–136.

Trkola, A., Pomales, A.B., Yuan, H., Korber, B., Maddon, P.J., Allaway, G.P., Katinger, H., Barbas, C.F., Burton, D.R., Ho, D.D. and Moore, J.P. (1995) Cross-clade neutralization of primary isolates of human immunodeficiency virus type 1 by human monoclonal antibodies and tetrameric CD4-IgG. *J. Virol*. **69**, 6609–6617.

Trkola, A., Purtscher, M., Muster, T., Ballaun, C., Buchacher, A., Sullivan, N., Srinivasan, K., Sodroski, J., Moore, J.P. and Katinger, H. (1996) Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. *J. Virol*. **70**, 1100–1108.

Van Droogenbroeck, B., Cao, J., Stadlmann, J., Altmann, F., Colanesi, S., Hillmer, S., Robinson, D.G., Van Lerberge, E., Terryn, N., Van Montagu, M., Liang, M., Depicker, A. and De Jaeger, G. (2007) Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic Arabidopsis seeds. *Proc. Natl Acad. Sci. USA*, **104**, 1430–1435.

Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure.

Strasser, R; Stadlmann, J; Schahs, M; Stiegler, G; Quendler, H; Mach, L; Glossl, J; Weterings, K; Pabst, M and Steinkellner, H (2008). Plant Biotechnol J 6(4): 392-402.

Plant Biotechnology Journal (2008) **6**, pp. 392–402 doi: 10.1111/j.1467-7652.2008.00330.x

Generation of glyco-engineered *Nicotiana benthamiana* for the production of monoclonal antibodies with a homogeneous human-like *N*-glycan structure

Richard Strasser[1], Johannes Stadlmann[2], Matthias Schähs[1], Gabriela Stiegler[3], Heribert Quendler[3], Lukas Mach[1], Josef Glössl[3], Koen Weterings[4], Martin Pabst[2] and Herta Steinkellner[1,*]

[1]*Institute of Applied Genetics and Cell Biology, University of Natural Resources and Applied Life Sciences, Muthgasse 18, 1190 Vienna, Austria*
[2]*Department of Chemistry, University of Natural Resources and Applied Life Sciences, Muthgasse 18, 1190 Vienna, Austria*
[3]*Polymun Scientific GmbH, Muthgasse 18, 1190 Vienna, Austria*
[4]*Bayer BioScience N.V. Technologiepark 38, B-9052 Gent, Belgium*

Received 4 December 2007;
revised 18 January 2008;
accepted 21 January 2008.
**Correspondence* (fax 43-1-36006-6392;
e-mail herta.steinkellner@boku.ac.at)

Keywords: 2G12 anti-HIV antibody, *Nicotiana benthamiana*, plant *N*-glycosylation, recombinant proteins, RNAi glyco-engineering.

Summary

A common argument against using plants as a production system for therapeutic proteins is their inability to perform authentic human *N*-glycosylation (i.e. the presence of β1,2-xylosylation and core α1,3-fucosylation). In this study, RNA interference (RNAi) technology was used to obtain a targeted down-regulation of the endogenous β1,2-*xylosyltransferase (XylT)* and α1,3-*fucosyltransferase (FucT)* genes in *Nicotiana benthamiana*, a tobacco-related plant species widely used for recombinant protein expression. Three glyco-engineered lines with significantly reduced xylosylated and/or core α1,3-fucosylated glycan structures were generated. The human anti HIV monoclonal antibody 2G12 was transiently expressed in these glycosylation mutants as well as in wild-type plants. Four glycoforms of 2G12 differing in the presence/absence of xylose and core α1,3-fucose residues in their *N*-glycans were produced. Notably, 2G12 produced in XylT/FucT-RNAi plants was found to contain an almost homogeneous *N*-glycan species without detectable xylose and α1,3-fucose residues. Plant-derived glycoforms were indistinguishable from Chinese hamster ovary (CHO)-derived 2G12 with respect to electrophoretic properties, and exhibited functional properties (i.e. antigen binding and HIV neutralization activity) at least equivalent to those of the CHO counterpart. The generated RNAi lines were stable, viable and did not show any obvious phenotype, thus providing a robust tool for the production of therapeutically relevant glycoproteins in plants with a humanized *N*-glycan structure.

Introduction

Several studies have demonstrated that plants can produce efficiently complex mammalian proteins, such as monoclonal antibodies (mAbs; for recent reviews, see Stoger *et al*., 2002, 2004). The conserved secretory pathway between plants and mammals provides the post-translational modifications and correct folding needed to assemble functional antibodies. One advantage of using plants as an expression system is their ability to perform *N*-glycosylation similar to mammalian cells. Although plants are able to synthesize an *N*-glycan core structure identical to that of mammalian cells (GnGn), the terminal residues differ, mainly because plant complex *N*-glycans lack β1,4-galactose (and sialic acid) and core α1,6-fucose; however, they carry instead β1,2-xylose and core α1,3-fucose, which are absent in mammals. Typically, an *N*-glycan profile from a plant-derived mAb consists mainly of complex *N*-glycans carrying β1,2-xylose and core α1,3-fucose residues (for example, Cabanes-Macheteau *et al*., 1999; Bakker *et al*., 2001; Bardor *et al*., 2003; Schähs *et al*., 2007). Although it seems that the natural lack of core α1,6-fucose may provide an advantage for a possible therapeutic use (for example, Shields *et al*., 2002), unwanted *in vivo* side-effects of the non-mammalian sugars β1,2-xylose and core

α1,3-fucose cannot be excluded, in particular when mAbs are intended to be used for systemic application. The immunogenicity of these *N*-glycan epitopes is well documented and their role in allergy has not yet been clarified (for a recent review, see Altmann, 2007). RNA interference (RNAi) has proven to be a valuable tool for the elimination of core α1,3-fucosyltransferase (FucT) and β1,2-xylosyltransferase (XylT) expression (Cox *et al*., 2006), the two enzymes responsible for the attachment of β1,2-xylose and core α1,3-fucose. The surprisingly high efficiency of this strategy led to the production of a mAb with a single *N*-glycan species (GnGn) in the aquatic plant species *Lemna minor*. More importantly, the effector activities of antibodies with such a glycan-optimized profile were significantly enhanced compared with their Chinese hamster ovary (CHO)-derived homologues (Cox *et al*., 2006). Thus, it seems that the lack of core α1,3-linked fucose on a plant-made antibody has the same positive effect on its antibody-dependent cellular cytotoxicity activity as the removal of α1,6-fucose on CHO-derived antibodies (Shields *et al*., 2002).

Because of several unique features of human immunodeficiency virus (HIV), immunogens able to elicit neutralizing antibodies to a broad range of primary HIV-1 isolates have not yet been found. Nevertheless, a cocktail of three mAbs broadly neutralizing different HIV-1 strains exhibits *in vivo* inhibitory viral activity in patients (Trkola *et al*., 2005). The main antiviral effect of this cocktail was attributable to 2G12, making this antibody particularly interesting. Currently, recombinant 2G12 used for *in vitro* and *in vivo* studies is produced in CHO cells (for example, Armbruster *et al*., 2002). Recently, we reported the expression of 2G12 in the model plant *Arabidopsis thaliana* and a glycosylation mutant thereof (XT/FT k.o.), synthesizing complex *N*-glycans without xylose and core α1,3-fucose (Strasser *et al*., 2004a). Although *N*-glycans from wild-type *A. thaliana* 2G12 exhibit mainly xylosylated and core α1,3-fucosylated *N*-glycans (GnGnXF), 2G12 derived from the XT/FT k.o. line carries a homogeneous glycan species (GnGn) devoid of plant-specific glycan epitopes (Schähs *et al*., 2007). *Arabidopsis thaliana*- and CHO-derived 2G12 exhibit no differences in electrophoretic mobility and antigen-binding capacity. However, no further characterization of *A. thaliana*-derived 2G12 was reported, mainly because of the relatively low expression levels and low biomass. Functional analyses are particularly important as, for unknown reasons, another anti-HIV antibody (2F5) produced in tobacco cells revealed a significant decrease in neutralization activity compared with the CHO-derived counterpart (Sack *et al*., 2007). Although recent progress has been made in enhancing the expression levels of recombinant proteins in *A. thaliana* seeds (De Jaeger *et al*., 1999; Van Droogenbroeck *et al*., 2007), this is currently not the species of choice if a large amount of recombinant proteins is needed (for example, detailed functional *in vitro* and *in vivo* studies). Tobacco and related species have been proven to meet many of the requirements needed for a rapid, efficient and cost-effective production platform. In particular, *Nicotiana benthamiana* is widely used for the efficient expression of recombinant proteins. Recent advances in expression technologies have enabled the production of recombinant proteins at extremely high levels in this plant species (Marillonnet *et al*., 2004). These results, together with a relatively fast growth rate and reasonable growth conditions, make this plant species particularly economically interesting.

In this study we modulated the *N*-glycan composition in *Nicotiana benthamiana* using RNAi constructs to obtain a targeted downredulation of the expression of endogenous *FucT* and *XylT* genes. These glycosylation mutants were used to produce various 2G12 glycoforms differing in the presence/absence of plant-specific glycan epitopes. All plant-derived 2G12 glycoforms were investigated for their structural and functional integrity using sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE), liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS), antigen binding and virus neutralization activity.

Results

Generation of *N. benthamiana* glycosylation mutants

To obtain a targeted down-regulation of endogenous *XylT* and *FucT* genes, two binary constructs with appropriate RNAi sequences were used to transform *N. benthamiana* wild-type plants (Figure 1a). Independent transgenic XylT- and FucT-RNAi lines were screened for the presence/absence of the respective products: *N*-glycans with xylose or core α1,3-fucose residues, respectively. Protein extracts from RNAi lines were subjected to Western blot analyses using xylose- and core α1,3-fucose-specific antibodies, respectively (Figure 1b). Various staining intensities were obtained, indicating the presence of different amounts of xylose and fucose residues, respectively. One XylT-RNAi line (line 1, named ΔXT) and one FucT-RNAi line (line 14, named ΔFT), which exhibited no or very weak staining with the corresponding antibody, were brought to a homozygous stage and used for crossing. The progeny thereof were subjected to Western blotting using anti-horseradish peroxidase (anti-HRP) antibodies, which recognize β1,2-xylose- and core α1,3-fucose-containing

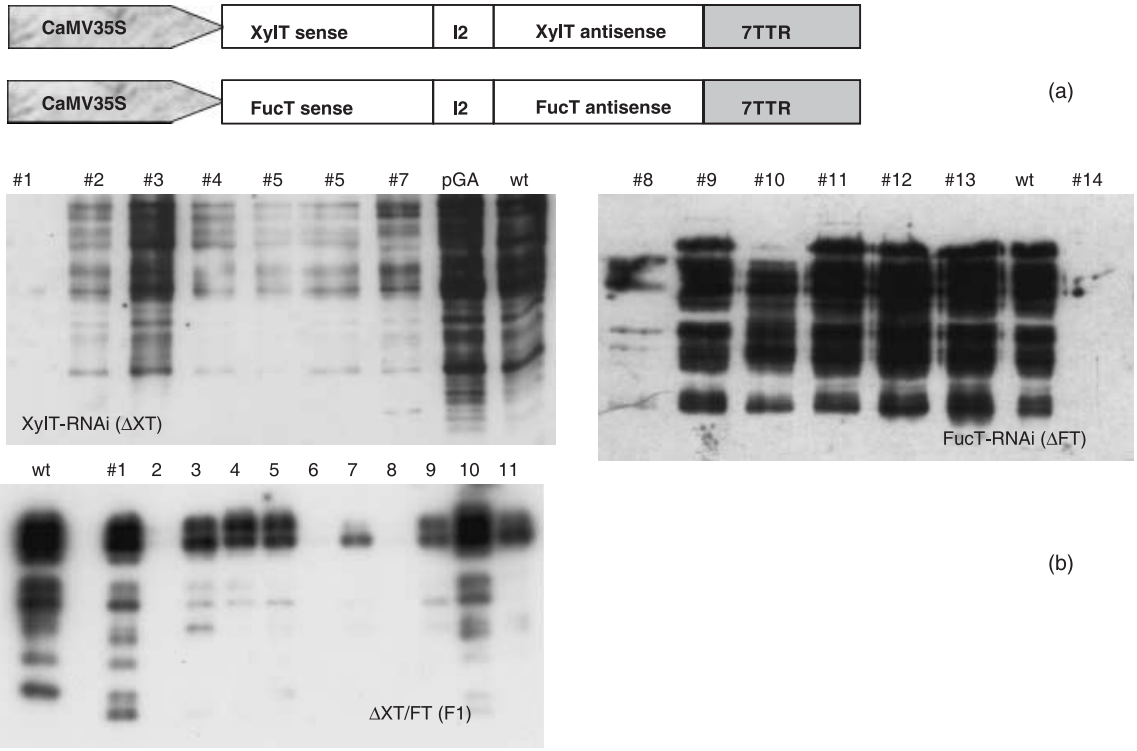

Figure 1 (a) XylT- and FucT-RNAi constructs [RNA interference constructs targeting the expression of endogenous β1,2-*xylosyltransferase (XylT)* and α1,3-*fucosyltransferase (FucT)* genes] used for *Nicotiana benthamiana* transformation. CaMV35S, cauliflower mosaic virus 35S promoter; I2, intron; 7TTR, transcriptional termination sequence. A detailed description of the cloning procedure and length of *XylT* and *FucT* sense–intron–antisense constructs is given in 'Experimental procedures'. (b) Western blot analysis of total protein extracts derived from XylT- and FucT-RNAi transformed *N. benthamiana* leaves using anti-xylose (in ΔXT lines) and anti-fucose (in ΔFT lines) specific antibodies. ΔXT/FT, total proteins extracted from progeny of crossed ΔXT and ΔFT. The presence/absence of xylose and fucose was detected with anti-horseradish peroxidase (anti-HRP) antiserum. Numbers indicate the different transformed lines, pGA, *N. benthamiana* transformed with empty binary vector; wt, non-transformed plants.

structures (Wilson and Altmann, 1998). Several plantlets of this F_1 generation did not exhibit any staining signal (Figure 1b), indicating the lack of both *N*-glycan epitopes. One line (line 6) was grown to maturity and used for further propagation. Although selfed progeny (F_2) and F_3 seedlings still exhibited a heterogeneous staining pattern in immunoblotting using anti-HRP antibodies, F_4 progeny showed a uniform negative staining (data not shown), indicating the stable down-regulation of both *XylT* and *FucT*. This line was named ΔXT/FT. In addition, xylose and fucose reduction was stable for at least three generations in ΔXT and ΔFT single RNAi lines, respectively. All RNAi lines were viable and did not exhibit any obvious phenotype under standard growth conditions.

N-Glycan analysis of total endogenous proteins from *N. benthamiana* wild-type and RNAi lines

To reveal the exact *N*-glycan composition of the total endogenous proteins, wild-type and all three glycosylation mutants of *N. benthamiana* were subjected to matrix-assisted laser desorption ionization-time of flight/time of flight mass spectrometry (MALDI-TOF/TOF MS) analysis (Figure 2). As expected, the mass spectrum of wild-type plants revealed that the vast majority of the *N*-glycan species carried xylose and α1,3-fucose residues (over 80%). In contrast, the *N*-glycan composition in ΔXT and ΔFT exhibited a significant reduction of xylose and fucose, respectively. Surprisingly, the specific down-regulation was more pronounced in the ΔXT, where only traces of one xylose-carrying glycoform were detected (GnGnXF: approximately 3%). However, significant amounts of α1,3-fucose carrying glycoforms were present in the ΔFT (about 20% of MMXF and GnGnXF). The *N*-glycan composition of the ΔXT/FT line exhibited a profile with complex-type *N*-glycans virtually lacking xylose and a significant reduction of α1,3-fucose (Figure 2).

Expression and purification of immunoglobulin G (IgG) glycoforms

ΔXT, ΔFT, ΔXT/FT and wild-type *N. benthamiana* plants were used to transiently express 2G12, a human IgG1 mAb. A binary vector that carries the two cDNAs of 2G12 light and

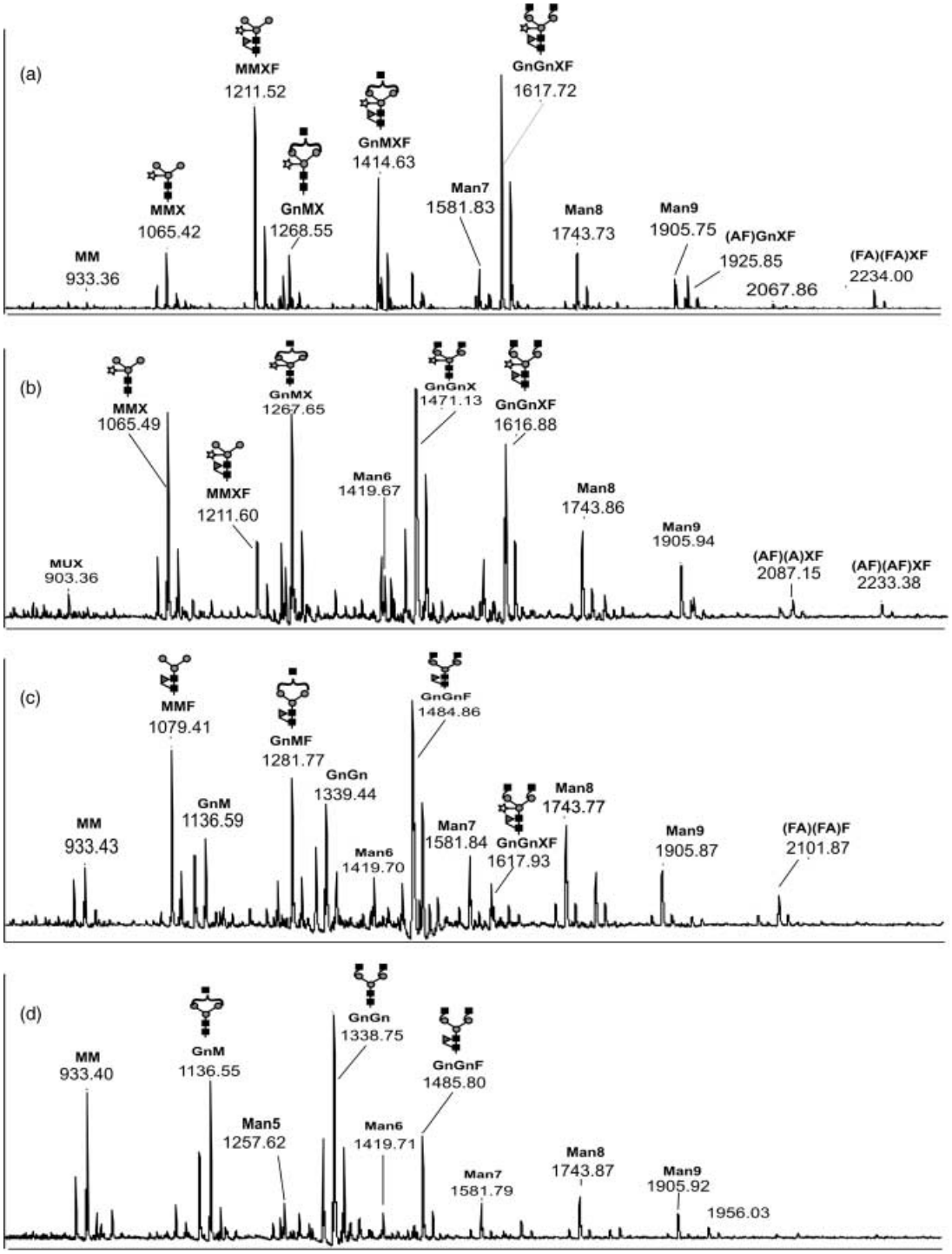

Figure 2 Matrix-assisted laser desorption ionization-time of flight/time of flight mass spectrometry (MALDI-TOF/TOF MS) of oligosaccharides from endogenous proteins extracted from *Nicotiana benthamiana* leaves: (a) wild-type; (b) ΔFT; (c) ΔXT; (d) ΔXT/FT. The labelled peaks represent $(M + Na)^+$ ions. Other peaks are potassium adducts of the same glycans. (See http://www.proglycan.com for an explanation of *N*-glycan abbreviations.)

heavy chain cloned in tandem (Schähs *et al.*, 2007) was used to agroinfiltrate *N. benthamiana* plants. Three days after infiltration, the leaves were initially monitored for 2G12 expression by immunoblotting using anti-human IgG antibodies. Two bands at positions 55 and 25 kDa, the expected size of the heavy and light chains, respectively, were detected (data not shown), demonstrating the functional expression of the two cDNAs. On average, an expression level

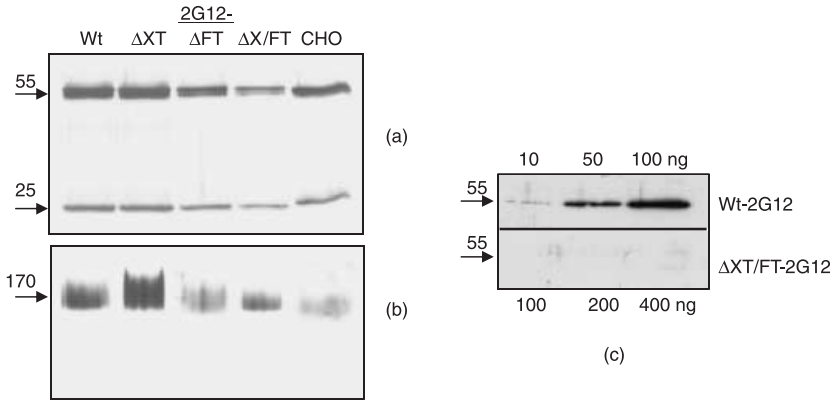

Figure 3 Silver-stained sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) of purified 2G12 derived from wild-type *Nicotiana benthamiana* (Wt), different RNA interference (RNAi) lines and Chinese hamster ovary (CHO) cells under reducing (a, 50 ng) and non-reducing (b, 200 ng) conditions. (c) Western blot analysis of the heavy chain from purified wild-type 2G12 (Wt-2G12) and ΔXT/FT-2G12 using anti-horseradish peroxidase antiserum. Please note that different amounts of 2G12 were loaded.

of about 0.5% of total soluble proteins was detected, which corresponds to 110 μg of 2G12 per gram of leaf material. Subsequently, agroinfiltrated leaf material was used to purify 2G12 by protein A affinity chromatography, yielding 50 μg of purified 2G12 per gram of biomass. In this respect, it should be mentioned that neither the infiltration process nor any purification step was systematically optimized. SDS-PAGE analysis of purified 2G12 revealed the presence of two bands that correlated with the expected molecular masses of the two chains (Figure 3a). An additional weak band was detected slightly below the band for the heavy chain, representing (LC-ESI-MS data of the glycopeptides) an unglycosylated form of the heavy chain. No further bands, degradation products or free heavy or light chains were visible in reducing and non-reducing gels, indicating the purity and structural integrity of all plant-produced 2G12 glycoforms (Figure 3a,b).

N-Glycan analyses of 2G12 glycoforms

To determine the exact *N*-glycan composition of CHO- and plant-derived 2G12 glycoforms, the corresponding heavy chains, excluding the unglycosylated fraction, were subjected to LC-ESI-MS. Because of incomplete trypsin digestion of the heavy chain, two glycopeptides, which contain the carbohydrate attached to the side-chain asparagine-297, but differ in four amino acids (482.3 Da), were present. The absence of xylose and fucose residues on *N*-glycans can be monitored by a decrease in the mass of the respective peaks (132 mass units for xylose and 146 mass units for fucose).

The mass spectrum of wild-type 2G12 exhibited two major peaks for both glycopeptides (Figure 4b, Table 1), which were assigned to the structure GnGnXF. This *N*-glycan species encompassed over 68% of all assigned structures. Only two other structures (GnMXF and GnUXF) slightly exceeded a share of 5% of the total *N*-glycan structures. Noteworthy, that over 90% of all structures carried xylose and core α1,3-fucose residues. The mass spectra of ΔXT- and ΔFT-2G12 exhibited two major peaks identified as GnGnF and GnGnX, respectively (Figure 4c,d, Table 1). However, in both lines, GnGnXF was also found: 20% in ΔFT-2G12 and 10% in ΔXT-2G12. This result demonstrates a significant decrease, but not complete elimination, of xylose and fucose residues in the respective RNAi lines. For so far unknown reasons, interestingly, an increased level of GnGn structures was detected in ΔXT-2G12. By far the most abundant *N*-glycan structure present in ΔXT/FT-2G12 was GnGn (80%; Figure 4e, Table 1). Surprisingly, and in contrast with the *N*-glycan structure of total endogenous proteins, no xylose- and fucose-containing *N*-glycan structures were detected. As traces of xylose- and fucose-containing glycoforms could not be excluded, there was a need to estimate at what level plant-specific glycoforms could be detected using the method described here. A serial dilution of wild-type 2G12 in CHO-2G12 was analysed, and revealed no detection of plant-specific *N*-glycans when the wild-type 2G12 content was below 5% (data not shown). To further confirm this result, an immunoblot

Table 1 Liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) analysis of the glycosylated Fc peptides EEQYN297STYR and TKPREEQYN297STYR from *Nicotiana benthamiana* wild-type (wt) 2G12, ΔXT-2G12, ΔFT-2G12 and ΔXT/FT-2G12

	Relative abundance (%)*			
N-glycan	wt	ΔXT	ΔFT	ΔXT/FT
GnGn	1.9	20.8	5.8	80.6
GnGnF	2.1	39.5	1.2	< 1
GnGnX	5.5	1.3	41.8	< 1
GnGnXF	67.6	10.4	19.4	< 1

*Numbers are the sum of the relative abundances of a specific glycoform found on each of the two immunoglobulin G (IgG) glycopeptides. The data were deduced from the LC-ESI-MS analyses illustrated in Figure 4.

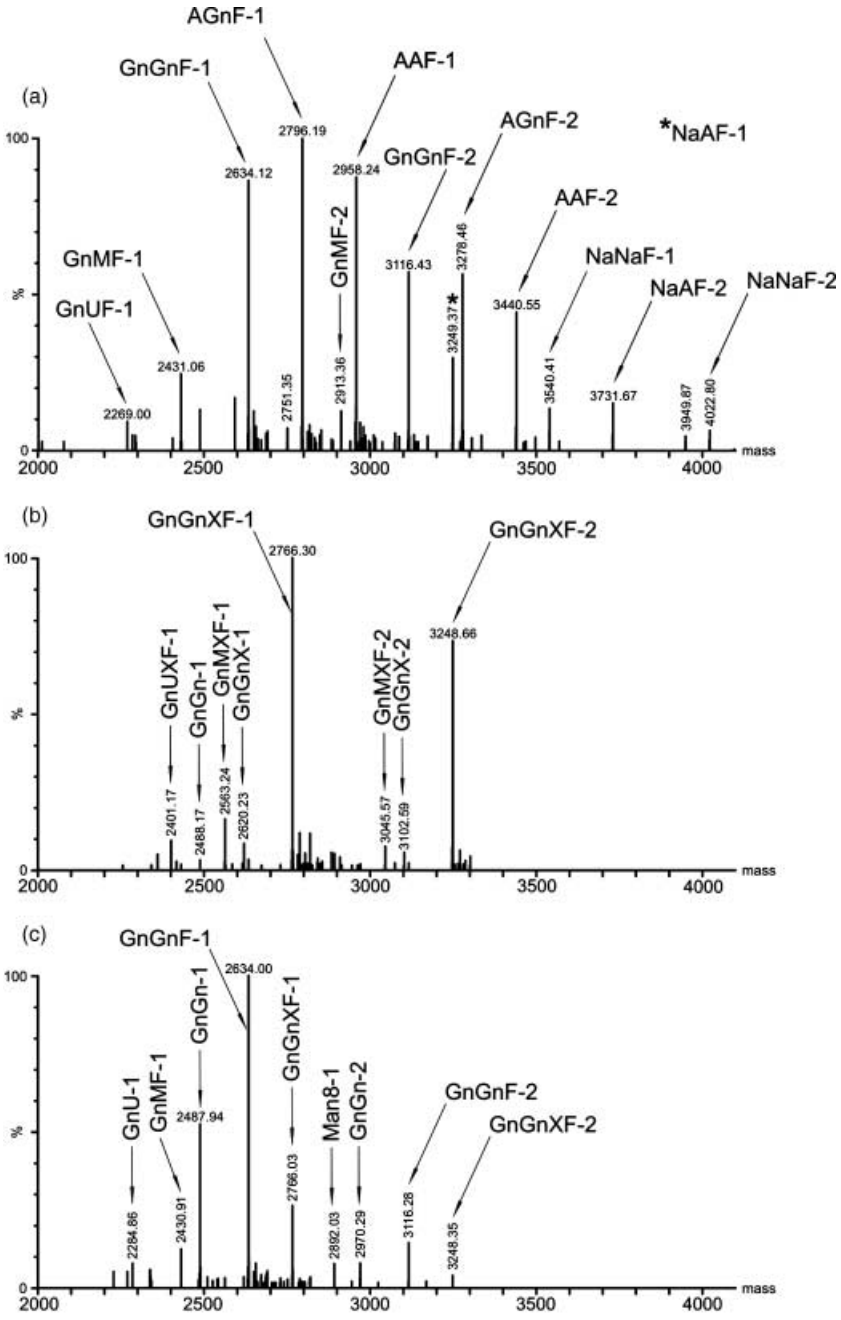

Figure 4 Liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) of glycosylated peptides (1, EEQYNSTYR; 2, TKPREEQYNSTYR) containing asparagine-297 (Asn-297) from the 2G12 Fc region. Glycopeptides derived from Chinese hamster ovary (CHO)-2G12 (a), *Nicotiana benthamiana* wild-type 2G12 (b), ΔXT-2G12 (c), ΔFT-2G12 (d), ΔXT/FT-2G12 (e). (See http://www.proglycan.com for an explanation of *N*-glycan abbreviations.) For CHO-2G12, ΔXT-2G12 and ΔFT-2G12, 500 ng of heavy chain was used; for ΔXT/FT-2G12, 1 μg.

using anti-HRP antiserum, which efficiently recognizes plant-specific *N*-glycan structures, was carried out (Figure 3c). Although anti-HRP recognizes 10 ng of wild-type 2G12, no reaction was obtained with ΔXT/FT-2G12, even when 400 ng was loaded (Figure 4c). This result indicates a detection limit of plant-specific *N*-glycans below 2.5%, and demonstrates an efficient elimination of both xylose and core α1,3-fucose in ΔXT/FT-2G12. *N*-Glycan analysis of CHO-2G12 revealed a mixture of structures containing either no (GnGnF), one (AGnF, GnAF) or two (AAF) terminal galactose residues and minor amounts of glycopeptides with terminal sialic acid (NaAF, NaNaF). Interestingly, all detected glycoforms contained a fucose residue α1,6-linked to the innermost *N*-acetylglucosamine (GlcNAc).

Functional analyses of 2G12 glycoforms

In order to test the functional integrity of the 2G12 glycoforms, specific antigen-binding and HIV neutralization assays were carried out. In previous studies, it has been shown that 2G12 recognizes an epitope within the glycoprotein 120 (gp120) subunit of the HIV envelope glycoprotein 160 (gp160) (Trkola

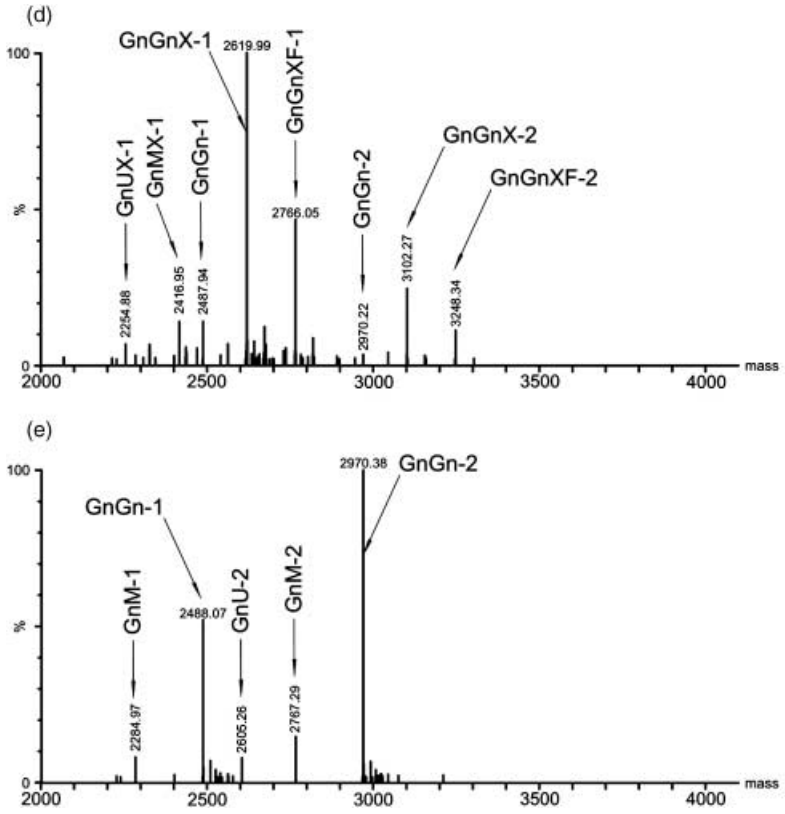

Figure 4 *Continued*

Table 2 Analysis of functional integrity of 2G12 glycoforms. Antigen-binding capacity was detected by a specificity enzyme-linked immunosorbent assay (ELISA) coated with purified human immunodeficiency virus (HIV) glycoprotein 160 (gp160). *In vitro* HIV-1 neutralization activity of 2G12 glycoforms was detected by a syncytium inhibition assay using AA-2 cells

2G12 glycoform	Antigen binding (%) (± SD)	IC_{50} (μg/mL)	IC_{50} (μg/mL)	IC_{50} (μg/mL)	Average IC_{50} (SD) (μg/mL)
Wild-type	117 (± 17)	5.3	1.3	1.6	2.7 (2.2)
ΔXT	115 (± 16)	4.4	1.9	1.9	2.7 (1.4)
ΔFT	138 (± 20)	4.4	0.6	6.3	3.8 (2.9)
ΔXT/FT	130 (± 19)	3.1	2.2	2.2	2.5 (0.5)
CHO	100 (± 15)	18	4.4	12.6	11.7 (6.8)

CHO, Chinese hamster ovary; IC_{50}, 50% inhibitory concentration; SD, standard deviation.

et al., 1996). In order to test the antigen binding of plant-produced 2G12 glycoforms, a specificity enzyme-linked immunosorbent assay (ELISA), using purified gp160, was carried out. Compared with CHO-2G12, the binding capacity of the plant-derived antibodies was 100% or slightly above (Table 2), demonstrating that the antigen-binding capacity was at least as efficient as that of the CHO-derived protein.

Finally, the ability of plant-derived 2G12 glycoforms to neutralize HIV was examined using a syncytium inhibition assay (Table 2). All plant-derived 2G12 glycoforms exhibited relatively uniform 50% inhibitory concentration (IC_{50}) values. No significant differences were observed between the four glycoforms during three independent virus neutralization assays using three different 2G12 batches. Notably, in all tests, CHO-2G12 displayed a three- to fourfold greater IC_{50} value.

Our results demonstrate that all plant-derived 2G12 glycoforms are fully functional antibodies and exhibit activities at least comparable with their CHO-produced counterpart.

Discussion

The feasibility of glyco-engineering in *N. benthamiana*, a plant species commonly used for the high-level expression of recombinant proteins, has been demonstrated. For this, three glycosylation mutants were generated using RNAi-targeted

down-regulation of the endogenous *XylT* and *FucT* genes. Silencing was stable and plants were viable and did not show an obvious phenotype under laboratory conditions. More importantly, the human mAb 2G12 against HIV was efficiently expressed, secreted and assembled in all three glycosylation mutants, as well as in wild-type *N. benthamiana*. Although the *N*-glycans of wild-type 2G12 were mainly decorated with xylose and core α1,3-fucose residues, ΔXT- and ΔFT-2G12 exhibited a significant decrease in the respective glycan residues. Interestingly, ΔXT/FT-2G12 carried a major complex *N*-glycan species with undetectable plant-specific carbohydrate modifications. The absence of xylosylated and core α1,3-fucosylated *N*-glycan structures was confirmed by MS analysis and immunoblotting of purified ΔXT/FT-2G12. To roughly estimate the detection limits of the methods used, a serial dilution wt-2G12 in CHO-2G12 was analysed with MS, and revealed a detection limit below 5% (data not shown). Furthermore, plant-specific carbohydrates were recognized by anti-HRP antibodies when only 10 ng of wild-type 2G12 was used. No reaction was obtained using 400 ng of ΔXT/FT-2G12, indicating that, if any glycan epitopes were present, the level was below the detection limit of 2.5%. The fact that neither xylose nor α1,3-fucose residues were detected in ΔXT/FT-2G12 is surprising, as *N*-glycans of total endogenous proteins, as well as 2G12 expressed in single RNAi lines, did not exhibit a complete elimination of the *N*-glycan epitopes. This was particularly true for α1,3-fucose, for which amounts of about 20% were still present on total endogenous proteins in corresponding RNAi lines as well as in ΔFT-2G12. These results suggest that an incomplete down-regulation of XylT and FucT activity is sufficient to allow the generation of IgGs carrying a glycan-optimized structure. One explanation of the relatively low down-regulation of FucT compared with XylT could be the presence of several copies of *FucT* genes, as described for other plant species (Strasser *et al*., 2004a; Bondili *et al*., 2006). Another unexpected observation was the presence of a significant portion of non-glycosylated heavy chain. However, this does not seem to influence antigen binding and HIV neutralization. This is not surprising as the importance of *N*-glycosylation status mainly lies in the *in vivo* activities of IgGs (for example, antibody-dependent cellular cytotoxicity and complement-dependent cytotoxicity). Possible *in vivo* effects of this incomplete glycosylation remain to be investigated.

The presence of a major *N*-glycan species in ΔXT/FT-2G12 (GnGn) indicates that the strategy applied in this study has, beside the elimination of potentially immunogenic *N*-glycan residues, the additional benefit of producing a monoclonal antibody with a widely homogeneous *N*-glycan profile. This may serve as a tremendous advantage over CHO-produced 2G12, where a variety of glycoforms are present. Notably, nearly all glycan structures of CHO-2G12 carry α1,6-fucose, a glycan residue which negatively interferes with the effector activities of IgGs (Shields *et al*., 2002; Shinkawa *et al*., 2003; Ferrara *et al*., 2006). In this respect, the absence of plant-specific *N*-glycans on ΔXT/FT-2G12 may not only eliminate potential negative side-effects, when systemically applied in humans, but may also positively influence effector functions, as recently described for other plant-derived glycan-optimized IgGs (Cox *et al*., 2006; Schuster *et al*., 2007).

All plant-derived 2G12 glycoforms were indistinguishable from the CHO-produced counterpart with regard to electrophoretic mobility and IgG assembly, and showed at least the same antigen-binding capacity. Moreover, for all plant-produced 2G12 glycoforms, HIV neutralization activity, as determined by syncytium inhibition assays, was at least the same as that of the CHO-derived protein. This is particularly interesting as, recently, another HIV antibody produced in tobacco cells exhibited a significantly reduced HIV neutralization activity compared with the CHO-derived counterpart (Sack *et al*., 2007).

Therefore, in this study, it has been demonstrated that glyco-engineered *N. benthamiana* may serve as a robust and efficient expression platform for the production of human mAbs without detectable plant-specific *N*-glycan residues. ΔXT/FT and the single RNAi lines, together with the recently developed viral-based expression vectors (Marillonnet *et al*., 2004; Giritch *et al*., 2006), which allow the generation of high titres of IgG (0.5 g of IgG per kilogram of fresh leaf biomass; Giritch *et al*., 2006), will pave the way to the transfer of knowledge gained in the laboratory to a clinical trial scale, where grams of purified IgGs are needed.

Experimental procedures

RNAi constructs

For the XylT-RNAi construct, a β1,2-*xylosyltransferase* fragment was amplified from *N. benthamiana* leaf cDNA using the forward primer XT24 (5′-TATATGTCGACTCTAGATTAGCAATGAAGAGCAAGTA-3′) and reverse primer Tom-XT23 (5′-AGCAGCCAAGACTCCTCAAAAT-3′), which were designed on the basis of sequence homology to already known (for example, Strasser *et al*., 2004a; Bondili *et al*., 2006) or annotated XylT sequences in the database. The polymerase chain reaction (PCR) product (304 bp) was *Sal*I/*Bam*HI digested and ligated into the *Sal*I/*Bam*HI sites of cloning vector puc18XTI2 (Strasser *et al*., 2007) to create puc18Xsi. The antisense fragment was obtained by PCR using the primers XT25 (5′-TATATGAATTCTAGATTAGCAATGAAGAGCAAGTA-3′) and XT26 (5′-ATTGCGGTACCGCATAAGACCCCTCCA-3′), and cloned into the *Eco*RI/*Kpn*I site of puc18Xsi to create puc18Xsias.

Subsequently, the 'sense–intron–antisense' cassette was excised by *Xba*I digestion and cloned into *Xba*I linearized plant expression vector pGA643. FucT cDNA (414 bp) was amplified by PCR using the primers NbFT1 (5′-TTATGGTACCGGATCCTTGGCAGCGGCTTTCATTT-3′) and NbFT2 (5′-AATTGGTACCGGATCCATCAGATGGGCCCT-CAAACT-3′), also designed on the basis of sequence homology with already known *FucT* genes (for example, Strasser *et al*., 2004a; Bondili *et al*., 2006), and cloned into the *Bam*HI site of puc18XTI2 to create puc18Fsi. The antisense fragment was amplified by PCR using primers NbFT2 and NbFT4 (5′-TTATGGTACCTCTAGATTGGCAGCG-GCTTTCATTT-3′), and cloned into the *Kpn*I digested vector puc18Fsi to generate puc18Fsias. Subsequently, the 'sense–intron–antisense' cassette was excised by *Xba*I digestion and cloned into *Xba*I linearized plant expression vector pGA643.

Screening of transgenic plants

Leaf disc transformation was used to transform *N. benthamiana* plants with appropriate RNAi constructs (Strasser *et al*., 2004b). Kanamycin-resistant plants were screened by PCR to confirm the genomic insertion of the RNAi constructs. Total soluble proteins of transgenic plants were extracted in phosphate-buffered saline (PBS) (w/v) and subjected to Western blot analysis as described previously (Strasser *et al*., 2004a) using rabbit anti-xylose and anti-fucose antiserum (1 : 500). Crossed F_1 and further selfed progeny were screened using rabbit anti-HRP antiserum (1 : 15 000; Sigma-Aldrich, St Louis, MO, USA). As detection antibody, HRP-conjugated goat anti-rabbit IgG (1 : 100 000; Sigma-Aldrich) was used. Peroxidase detection was carried out with Super Signal West Pico Chemiluminescent Substrate Kits (Pierce, Rockford, IL, USA).

Preparation of *N*-linked glycans and MALDI-TOF/TOF MS analysis

The preparation and purification of glycans derived from total endogenous leaf proteins were performed as described by Strasser *et al*. (2004a). Mass spectrometric analysis of endogenous glycans was carried out with a Bruker Ultraflex MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) in reflector mode at positive ionization. 2,5-Dihydroxybenzoic acid (DHB) [2% (w/v) DHB in 50% (v/v) acetonitrile in H_2O] was used as matrix.

Agroinfiltration and 2G12 purification

Agrobacterium strain GV3101::pMP90RK containing 2G12 binary vector, described recently (Schähs *et al*., 2007), was grown overnight [optical density (OD), 1.0–1.5] and diluted in infiltration buffer [50 mg D-glucose 10 mL, 50 mM 2-(*N*-morpholino)ethanesulphonic acid (MES), 2 mM $Na_3PO_4 \cdot 12H_2O$, 0.1 mM acetosyringone] to an OD of 0.2. Using a 1-mL syringe without a needle, *Agrobacterium* suspensions were injected through a small puncture into fully expanded leaves of 6–8-week-old *N. benthamiana*. For purification, 30 g of infiltrated leaf material was homogenized in liquid nitrogen, resuspended in 60 mL of pre-cooled extraction buffer (1 × PBS, pH 6.0) and incubated for 15 min at 4 °C. Insoluble material was removed by centrifugation at 21 000 ***g*** for 30 min at 4 °C. The supernatant was filtered and sedimented as above. Subsequently, an additional isoelectric precipitation by reducing the pH to 4.5 (1 M HCl) was carried out. The supernatant was centrifuged for 15 min at 4 °C. The supernatant obtained was brought to pH 6.0–7.0 with 1 M KOH and applied to a disposable column (Bio-Rad, Hercules, CA, USA) packed with 1 mL rProtein A Sepharose Fast Flow (GE Healthcare, Freiburg, Germany). IgG was eluted with 100 mM glycine, pH 2.5. The elution fractions were immediately adjusted to pH 7.0 by the addition of 1 M tris(hydroxymethyl)aminomethane (Tris).

2G12 quantification and antigen-binding assay

To quantify and determine the antigen-binding specificity of plant-derived 2G12, a sandwich ELISA was carried out. The IgG concentration and gp160 binding were determined by coating ELISA plates with a polyclonal antibody 'goat anti-human IgG' (γ-chain, Sigma; for quantification) or with 100 ng of purified gp160 (provided by Polymun Scientific, GmbH, Vienna, Austria). 2G12 samples were diluted to a concentration of about 200 ng/mL. Subsequently, a 1 : 2 serial dilution was carried out (eight dilutions) and 50 μL was transferred per well. Recombinant 2G12 purified from CHO culture supernatants (provided by Polymun Scientific) was used as standard, and applied identically to plant-derived 2G12. Plates were incubated for 1 h at room temperature (gentle shaking) and, after washing, a polyclonal antibody 'goat anti-human IgG' (κ-chain) conjugated to alkaline phosphatase (1 : 1000, Sigma) was added. After intensive washing, bound alkaline phosphatase was detected by *p*-nitrophenylphosphate substrate solution using an ELISA reader at 405 nm, with 620 nm as reference wavelength. All analyses were carried out in duplicate, and a four-parameter polynomial was applied for curve fitting and data analysis.

N-Glycan analysis of 2G12 by LC-ESI-MS

In order to allow a highly reliable identification of glycan structures, together with rapid and sensitive quantification, *N*-glycan analysis of 2G12 glycoforms was carried out by LC-ESI-MS of tryptic glycopeptides (Schuster *et al*., 2007; Van Droogenbroeck *et al*., 2007). Briefly, the heavy chain (500 ng for ΔXT- and ΔFT-2G12, and 1 μg for ΔXT/FT-2G12) of SDS-PAGE-separated, purified 2G12 was excised from the gel, *S*-alkylated and digested with trypsin. The tryptic peptides and glycopeptides were trapped on an Aquasil C18 pre-column (30 × 0.32 mm, Thermo Electron; Thermo Scientific, Waltman, MA, USA) using water as the solvent, and then separated on a Biobasic C18 column (100 × 0.18 mm, Thermo Electron) with a gradient of 5%–50% acetonitrile containing 0.1 M formic acid. Positive ions of m/z = 200–2000 were monitored with a Q-TOF Ultima Global mass spectrometer (Waters, Milford, MA, USA). Identification and glycoform quantification were performed on the summed and deconvoluted spectrum of the glycopeptide elution region.

Western blotting of wild-type 2G12 and ΔXT/FT-2G12

Purified wild-type 2G12 and ΔXT/FT-2G12 were separated by 10% SDS-PAGE under reducing conditions, transferred to nitrocellulose membrane and incubated with rabbit anti-HRP (1 : 15 000; Sigma-Aldrich). Detection was carried out with HRP-conjugated goat anti-rabbit IgG (1 : 100 000; Sigma-Aldrich) and Super Signal West Pico Chemiluminescent Substrate Kits (Pierce).

HIV neutralization assay

To test the ability of 2G12 to neutralize HIV-1, a syncytium inhibition assay was carried out (Trkola *et al*., 2005). Syncytium inhibition was assessed using AA-2 cells as indicator cell line with syncytium formation as read-out. Briefly, 10 twofold serial dilutions of antibodies (starting concentration, 50 μg/mL) in polybrene-containing cell culture medium (5 μg/mL; Sigma) were pre-incubated with 10^2–10^3 50% tissue culture infective dose ($TCID_{50}$) per millilitre of HIV-1 for 1 h at 37 °C before the addition of AA-2 cells. The cells were incubated for 5 days before the assessment of syncytium formation. Experiments were performed with eight replicates per dilution step. The presence of at least one syncytium per well was considered as indication for HIV-1 infection. IC_{50} was calculated according to the method of Reed and Muench (1938). All assays included a virus titration of the inoculum to confirm its infectious titre.

Acknowledgements

We would like to thank Jakub Jez and Pia Gattinger for excellent technical assistance, and Mohammad-Ali Assadian-Tschalschotori (Polymun Scientific GmbH, Vienna, Austria) for 2G12 quantification and HIV gp160 binding assays.

This work was funded by the Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF, www.wwtf.at, Project number: LS154), the European Union (Pharma Planta: www.pharma-planta.com) and Bayer BioScience N.V. Gent, Belgium.

References

Altmann, F. (2007) The role of protein glycosylation in allergy. *Int. Arch. Allergy Immunol*. **142**, 99–115.

Armbruster, C., Stiegler, G.M., Vcelar, B.A., Jager, W., Michael, N.L., Vetter, N. and Katinger, H.W. (2002) A phase I trial with two human monoclonal antibodies (hMAb 2F5, 2G12) against HIV-1. *AIDS*, **16**, 227–233.

Bakker, H., Bardor, M., Molthoff, J.W., Gomord, V., Elbers, I., Stevens, L.H., Jordi, W., Lommen, A., Faye, L., Lerouge, P. and Bosch, D. (2001) Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl. Acad. Sci. USA* **98**, 2899–2904.

Bardor, M., Loutelier-Bourhis, C., Paccalet, T., Cosette, P., Fitchette, A.C., Vézina, L.P., Trépanier, S., Dargis, M., Lemieux, R., Lange, C., Faye, L. and Lerouge, P. (2003) Monoclonal C5-1 antibody produced in transgenic alfalfa plants exhibits an *N*-glycosylation that is homogeneous and suitable for glyco-engineering into human-compatible structures. *Plant Biotechnol. J*. **1**, 451–462.

Bondili, J.S., Castilho, A., Mach, L., Glossl, J., Steinkellner, H., Altmann, F. and Strasser, R. (2006) Molecular cloning and heterologous expression of beta1,2-xylosyltransferase and core alpha1,3-fucosyltransferase from maize. *Phytochemistry*, **67**, 2215–2224.

Cabanes-Macheteau, M., Fitchette-Laine, A.C., Loutelier-Bourhis, C., Lange, C., Vine, N.D., Ma, J.K., Lerouge, P. and Faye, L. (1999) N-Glycosylation of a mouse IgG expressed in transgenic tobacco plants. *Glycobiology*, **9**, 365–372.

Cox, K.M., Sterling, J.D., Regan, J.T., Gasdaska, J.R., Frantz, K.K., Peele, C.G., Black, A., Passmore, D., Moldovan-Loomis, C., Srinivasan, M., Cuison, S., Cardarelli, P.M. and Dickey, L.F. (2006) Glycan optimization of a human monoclonal antibody in the aquatic plant Lemna minor. *Nat. Biotechnol*. **24**, 1591–1597.

De Jaeger, B., Buys, E., Eeckhout, D., De Wilde, C., Jacobs, A., Kapila, J., Angenon, G., Van Montagu, M., Gerats, T. and Depicker, A. (1999) High level accumulation of single-chain variable fragments in the cytosol of transgenic Petunia hybrida. *Eur. J. Biochem*. **259**, 426–434.

Ferrara, C., Stuart, F., Sondermann, P., Brunker, P. and Umana, P. (2006) The carbohydrate at FcgammaRIIIa Asn-162. An element required for high affinity binding to non-fucosylated IgG glycoforms. *J. Biol. Chem*. **281**, 5032–5036.

Giritch, A., Marillonnet, S., Engler, C., van Eldik, G., Botterman, J., Klimyuk, V. and Gleba, Y. (2006) Rapid high-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors. *Proc. Natl. Acad. Sci. USA*, **103**, 14 701–14 706.

Marillonnet, S., Giritch, A., Gils, M., Kandzia, R., Klimyuk, V., Gleba, Y., Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V. and Gleba, Y. (2004) In planta engineering of viral RNA replicons: efficient assembly by recombination of DNA modules delivered by Agrobacterium. *Proc. Natl. Acad. Sci. USA*, **101**, 6852–6857.

Reed, L.J. and Muench, H.A. (1938) Simple method of estimating fifty per cent endpoints. *Am. J. Hyg.* **27**, 493–497.

Sack, M., Paetz, A., Kunert, R., Bomble, M., Hesse, F., Stiegler, G., Fischer, R., Katinger, H., Stoeger, E. and Rademacher, T. (2007) Functional analysis of the broadly neutralizing human anti-HIV-1 antibody 2F5 produced in transgenic BY-2 suspension cultures. *FASEB J*. **21**, 1655–1664.

Schähs, M., Strasser, R., Stadlmann, J., Kunert, R., Rademacher, T. and Steinkellner, H. (2007) Production of a monoclonal antibody in plants with a humanized *N*-glycosylation pattern. *Plant Biotechnol. J*. **5**, 657–663.

Schuster, M., Jost, W., Mudde, G.C., Wiederkum, S., Schwager, C., Janzek, E., Altmann, F., Stadlmann, J., Stemmer, C. and Gorr, G. (2007) In vivo glyco-engineered antibody with improved lytic potential produced by an innovative non-mammalian expression system. *Biotechnol. J*. **2**, 700–708.

Shields, R.L., Lai, J., Keck, R., O'Connell, L.Y., Hong, K., Meng, Y.G., Weikert, S.H. and Presta, L.G. (2002) Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *J. Biol. Chem*. **277**, 26 733–26 740.

Stoger, E., Sack, M., Fischer, R. and Christou, P. (2002) Plantibodies: applications, advantages and bottlenecks. *Curr. Opin. Biotechnol*. **13**, 161–166.

Stoger, E., Schillberg, S., Twyman, R.M., Fischer, R. and Christou, P. (2004) Antibody production in transgenic plants. *Methods Mol. Biol*. **248**, 301–318.

Strasser, R., Altmann, F., Mach, L., Glossl, J. and Steinkellner, H. (2004a) Generation of Arabidopsis thaliana plants with complex N-glycans lacking beta 1,2-linked xylose and core alpha 1,3-linked fucose. *FEBS Lett*. **561**, 132–136.

Strasser, R., Altmann, F., Glossl, J. and Steinkellner, H. (2004b) Unaltered complex N-glycan profiles in Nicotiana benthamiana despite drastic reduction of beta 1,2-N-acetylglucosaminyltransferase I activity. *Glycoconj. J*. **21**, 275–282.

Strasser, R., Bondili, J.S., Vavra, U., Schoberer, J., Svoboda, B., Glossl, J., Leonard, R., Stadlmann, J., Altmann, F., Steinkellner, H. and Mach, L. (2007) A unique {beta}1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis A structures in Arabidopsis thaliana. *Plant Cell*, **19**, 2278–2292.

Trkola, A., Purtscher, M., Muster, T., Ballaun, C., Buchacher, A., Sullivan, N., Srinivasan, K., Sodroski, J., Moore, J.P. and Katinger, H. (1996) Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. *J. Virol*. **70**, 1100–1108.

Trkola, A., Kuster, H., Rusert, P., Joos, B., Fischer, M., Leemann, C., Manrique, A., Huber, M., Rehr, M., Oxenius, A. *et al*. (2005) Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nat. Med*. **11**, 615–622.

Unger, J.B., Cruz, T., Shakib, S., Mock, J., Shields, A., Baezconde-Garbanati, L., Palmer, P., Cruz, J.D., Edsall, E., Gritz, E.R., Glynn, T. and Johnson, C.A. (2003) The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J. Biol. Chem*. **278**, 3466–3473.

Van Droogenbroeck, B., Cao, J., Stadlmann, J., Altmann, F., Colanesi, S., Hillmer, S., Van Robinson, D.G.L.E., Van Terryn, N.M.M., Liang, M., Depicker, A. and De Jaeger, G. (2007) Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic Arabidopsis seeds. *Proc. Natl. Acad. Sci. USA* **104**, 1430–1435.

Wilson, I.B. and Altmann, F. (1998) Structural analysis of N-glycans from allergenic grass, ragweed and tree pollens: core alpha1,3-linked fucose and xylose present in all pollens examined. *Glycoconj. J*. **15**, 1055–1070.

Cost-effective production of a vaginal protein microbicide to prevent HIV transmission.
Ramessar, K; Rademacher, T; Sack, M; Stadlmann, J; Platis, D; Stiegler, G; Labrou, N; Altmann, F; Ma, J; Stoger, E; Capell, T and Christou, P (2008).
Proc Natl Acad Sci U S A 105(10): 3727-32.

Cost-effective production of a vaginal protein microbicide to prevent HIV transmission

Koreen Ramessar*, Thomas Rademacher†, Markus Sack†, Johannes Stadlmann‡, Dimitris Platis§, Gabriela Stiegler¶, Nikos Labrou§, Fritz Altmann‡, Julian Ma‖, Eva Stöger†, Teresa Capell*, and Paul Christou*,**††

*Departament de Producció Vegetal i Ciència Forestal, Universitat de Lleida, Avenida Alcalde Rovira Roure, 191, Lleida, 25198, Spain; †Institute for Molecular Biotechnology, Biology VII, RWTH Aachen, Worringerweg 1, 52074 Aachen, Germany; ‡Department of Chemistry, Glycobiology Division, University of Natural Resources and Applied Life Sciences, Muthgasse 18, 1190 Vienna, Austria; §Department of Agricultural Biotechnology, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece; ¶Polymun Scientific, Nussdorfer Laende 11, 1190 Vienna, Austria; ‖Saint George's University of London, Cranmer Terrace, London SW17 0RE, United Kingdom; and **Institucio Catalana de Recerca i Estudis Avancats, 08010 Barcelona, Spain

Communicated by M. S. Swaminathan, Taramani Institutional Area, Madras, India, September 22, 2007 (received for review July 24, 2007)

A series of small-molecule microbicides has been developed for vaginal delivery to prevent heterosexual HIV transmission, but results from human clinical trials have been disappointing. Protein-based microbicides, such as HIV-specific monoclonal antibodies, have been considered as an alternative approach. Despite their promising safety profile and efficacy, the major drawback of such molecules is the economy of large-scale production in mammalian cells, the current system of choice. Here, we show that an alternative biomanufacturing platform is now available for one of the most promising anti-HIV antibodies (2G12). Our data show that the HIV-neutralization capability of the antibody is equal to or superior to that of the same antibody produced in CHO cells. We conclude that this protein production system may provide a means to achieve microbicide ingredient manufacture at costs that would allow product introduction and manufacture in the developing world.

The number of people infected with HIV rose to just more than 40 million in 2006, an increase of over 2 million since 2004 (1). More than 60% of infected people live in sub-Saharan Africa, where at least 2 million deaths from HIV/AIDS occurred in 2006. Novel therapeutic strategies are urgently needed for deployment alongside conventional antiretroviral drugs, vaccines, and microbicides to prevent the spread of the disease. Broadly reactive human monoclonal antibodies (mAbs) against HIV could be used as both prophylactic and therapeutic modalities.

HIV-1 entry into susceptible cells is mediated by the envelope protein (Env), which comprises a trimer of gp120/gp41 heterodimers, with gp120 acting as the external surface subunit responsible for engaging cellular receptors, and gp41 as the transmembrane subunit that mediates membrane fusion (2). Infection occurs when gp120 interacts with cellular CD4 and then a coreceptor, usually CCR5 or CXCR4. Env is therefore an ideal target for neutralizing antibodies, and four mAbs with broad HIV-neutralizing activity have been characterized: the anti-gp120 antibodies b12 (3) and 2G12 (4) and the anti-gp41 antibodies 2F5 (5) and 4E10 (6).

The Env protein has evolved defenses to prevent neutralization, which few antibodies can overcome (7). The 2G12 antibody is an exception (8) and achieves neutralization by recognizing a unique gp120 epitope that although not directly associated with the receptor-binding sites (9) still prevents the virus interacting with its receptors (10). As well as neutralizing HIV-1 *in vitro*, 2G12 can protect macaques from simian-HIV challenge in passive immunization experiments, usually in combination with other antibodies (11). The crystal structure of the gp120 core shows that the epitope is highly conserved despite being located on the relatively variable surface of the outer domain, opposite to the CD4-binding site (12). The epitope contains high-mannose glycans to which the antibody probably binds (12).

The 2G12, 2F5 and 4E10 antibodies have been produced in Chinese hamster ovary (CHO) cells for prophylactic (13) and therapeutic (14) use, but the limitations of this platform (including high cost, low capacity, low scalability, and safety issues) (15) mean that large-scale production for use in developing countries is unfeasible (16). Among the alternative systems available for biopharmaceutical manufacture, only plants provide the scalability and economy required to meet the anticipated demand for such products in the HIV-endemic regions of Africa and Asia at a price the market can bear (17–20). Cereal seeds are likely to be the most suitable platform for deployment in such areas, because the infrastructure for large-scale cultivation and harvesting is already in place, and the dry seeds favor product stability (21). Maize is particularly attractive, because it has been developed as a commercial platform for recombinant protein production by a number of companies, and its success as a production system for pharmaceutical proteins is widely documented (22–24).

Because of the lack of direct comparative studies, it is unclear whether plant-derived antibodies are functionally comparable to antibodies produced in mammalian cells, particularly because of the differences in glycan structures between plants and mammals. Here, we show that a monoclonal antibody produced in maize seeds is not only comparable in terms of antigen-binding activity with its CHO-derived counterpart, but it is at least as active, if not superior, in HIV-neutralization assays. We also show that it is possible to isolate the antibody from maize seeds efficiently and to achieve 90% purity, using a simple two-step extraction process, which could be scaled up much more economically than the Protein A chromatography systems currently favored for industrial-scale antibody production in mammalian cells. Furthermore, by selecting a maize line with high-level antibody expression and subjecting it to a dedifferentiation–regeneration cycle, we were able to increase the accumulation of 2G12 to well over 100 μg per gram of dry seed weight, much higher than previously reported. We conclude that both the product and process can be economically advantageous when maize is used to express recombinant pharmaceutical proteins for humans.

Results

Transgenic Maize Plants. A large number of transgenic events were generated in the elite maize cultivar M37W by particle bombardment with vectors pTRAGtiGH and pTRAGtiGL (containing secretable forms of the 2G12 heavy and light chain genes,

Author contributions: J.M., E.S., and P.C. designed research; K.R., T.R., M.S., J.S., D.P., G.S., and T.C. performed research; K.R., M.S., J.S., D.P., N.L., F.A., J.M., E.S., T.C., and P.C. analyzed data; and K.R. and P.C. wrote the paper.

The author declares no conflict of interest.

††To whom correspondence should be addressed. E-mail: christou@pvcf.udl.cat.

This article contains supporting information online at www.pnas.org/cgi/content/full/0708841104/DC1.

respectively, under the control of an endosperm-specific promoter) and pTRAuxBar (containing the *bar* selectable marker gene, expressed under a constitutive promoter to facilitate phosphinothricin selection). The plants were either selfed or crossed with wild-type M37W, and Southern blot analysis was used to confirm transgene integration and stability. Most of the transgenic plants contained all three input transgenes linked within a single genetic locus, as anticipated (25, 26). Single genetic loci are beneficial, because they favor high-level and stable transgene expression over future generations (27–31).

2G12 Expression in Seeds and Selection of a Highly Expressing Plant Line. Transgenic seeds (T_1 generation) were separated into embryo and endosperm and screened by dot-blot analysis [supporting information (SI) Fig. 3], and the amount of 2G12 in the endosperm was estimated by ELISA. Western blot analyses were carried out under reducing conditions to confirm the presence of the 50-kDa heavy chain and 25-kDa light chain. Embryos from such seeds with high 2G12 expression levels were germinated, and the resulting plants self-pollinated. The new seeds (T_2 generation) were then analyzed by dot-blot analysis (SI Fig. 4) and ELISA, identifying event 3C as the highest expressing line. The endosperm tissue from 30 additional 3C seeds was tested by dot-blot analysis and all seeds were shown to be high expressers. Biacore surface plasmon resonance (SPR) spectroscopy showed considerable variation in the concentration of 2G12 among seeds tested from event 3C (SI Fig. 5). Immature zygotic embryos from event 3C were dedifferentiated *in vitro*, and numerous plants were regenerated from the resulting callus tissue. These plants accumulated even higher levels of 2G12 (see below).

Purification and Quantitation of 2G12. ***Affinity purification with Protein A.*** The 2G12 antibody was purified from maize by protein A affinity chromatography, which specifically binds the Fc portion of the antibody heavy chain. The flow-through and elution fractions were analyzed on-line by Biacore SPR spectroscopy. Light chain was detected in the flow-through fraction, which indicated that unassembled light chain was present *in planta* (i.e., the light chain was produced in excess). No heavy chain was detected in the flow-through, indicating that all of the heavy chain produced *in planta* was folded correctly (at least in the vicinity of the Fc region) and retained on the affinity matrix. Of the possible intact species that might be found in the elution fraction (H2L2, H2L, and H2), only H2L2 was detected, indicating highly efficient antibody assembly *in planta*. The antibody was eluted into a low-pH buffer to prevent aggregation and precipitation.

The fate of the antibody chains was also monitored by SDS/PAGE (Fig. 1*A*) and Western blot (Fig. 1*B*), confirming the presence of the heavy and light antibody chains (revealed as 50- and 25-kDa bands, respectively) before dialysis and concentration to 3.42 mg/ml (as determined by Biacore SPR spectroscopy using a protein L surface). Some degradation products (≈40 kDa) were visible in the pooled protein samples (indicated by red arrows in Fig. 1). The calculated extractable yield of 2G12 was 75 μg per gram of dry seed weight, 75% of which was recovered. Levels of 2G12 in seeds of plants regenerated from dedifferentiated immature zygotic embryos exceeded 100 μg per gram of dry seed weight. Furthermore, the dedifferentiation-regeneration cycle eliminated the wide variability in 2G12 expression we had observed previously in seeds of primary transformants and their progeny.

Non-protein A purification. The 2G12 antibody was also purified by using S-Sepharose FF cation exchange chromatography, followed by immobilized metal affinity chromatography (IMAC) on Zn^{2+}- iminodiacetic acid (IDA)-agarose. The antibody was adsorbed at pH 6.0 and subsequently eluted in a stepwise pH gradient (pH 7.0–8.0). Bulk maize endosperm protein (acidic pI) either did not bind to the cation exchanger or eluted at pH <7.2 (Fig. 2). Elution fraction corresponding to pH values between 7.4 and 7.6 in the first step separation favor the final obtainment of a highly pure antibody preparation. This procedure resulted in 10- to 15-fold purification of 2G12 with >70% recovery. Pooled eluted fractions (pH 7.4, 7.5, and 7.6) were loaded directly onto Zn^{2+}-IDA-agarose without any pretreatment and 2G12 was eluted in a stepwise pH gradient (pH 6.5 and 6.0) with step-recovery of ≈80%. Analysis of the eluted fractions by SDS/PAGE (Fig. 2) showed that the antibody was ≈90% pure. The overall recovery was 50–60%.

Antigen Binding Assays. Biacore SPR spectroscopy was used to compare the antigen-binding properties of the maize-derived antibody with its CHO-derived counterpart. Three surfaces were used: protein A (detects any species containing a correctly folded heavy chain), protein L (detects any species containing a correctly folded light chain), and gp120 (detects the assembled antibody only, not the free chains). The three surfaces also detect related degradation products with intact binding domains, although the signal would be weaker because of the lower molecular mass. By measuring ratios between the signals, i.e., the protein L/protein A ratio (SI Fig. 6*A*), the gp120/protein A ratio (SI Fig. 6*B*), and the gp120/protein L ratio (SI Fig. 6*C*), it is possible to characterize the binding properties of each antibody accurately. The results (Table 1) show that the gp120/protein L ratios of the maize and CHO 2G12 antibodies are the same, whereas the protein L/protein A and gp120/protein A ratios differ by the same relative amount. This deviation is caused by an increased protein A signal, meaning that the maize-derived 2G12 preparation contains an excess of heavy chain. Because we have already established that an excess of light chain is produced *in planta* and that partially assembled intact antibody forms are not eluted, the most likely explanation for the excess heavy chain in the eluted antibody preparation is the degradation products visible in Fig. 1. These heavy-chain degradation products do not bind to the antigen, and the gp120/protein L ratio confirms that the intact assembled antibody has full antigen-binding activity.

Glycan Analysis. Tryptic glycopeptides from maize 2G12 were analyzed by ESI-Q-TOF mass spectrometry, yielding valuable additional data on the protein sequence. The heavy chain from the protein A-purified antibody contained the expected, correctly processed N-terminal peptide EVQLVESGGGLVK, which was absent from the degradation product. The full-size heavy chain partially lacked the C-terminal lysine residue, as was recently found in the IGN314 antibody produced in moss (32), whereas this residue was absent from most of the degradation product. The continuous peptide map of the degradation product ran from L123 to K472, with the proteolytic cleavage site presumably located between G119 and L123.

Tryptic in-gel digestion of the heavy chains yielded two, a fully and an incompletely cleaved tryptic glycopeptides (i.e., EEQYN297STYR and TKPREEQYN297STYR, respectively), which both represent the same glycan acceptor site. There were striking differences in the glycan profiles of the two heavy chains (Table 2). More than half (54.5%) of the full-sized chains contained a single GlcNAc residue at the acceptor site, whereas only 17% of the degraded product carried this residue. Approximately 6% of the full-size product carried high-mannose glycans, whereas 28% carried complex type glycans, including fucose and xylose residues. In contrast, no high-mannose glycans were detected on the degraded heavy chain, and 67% carried complex glycans. A similar proportion of both types of heavy chain were nonglycosylated (11.5% for the full-size chain, 16% for the degradation product).

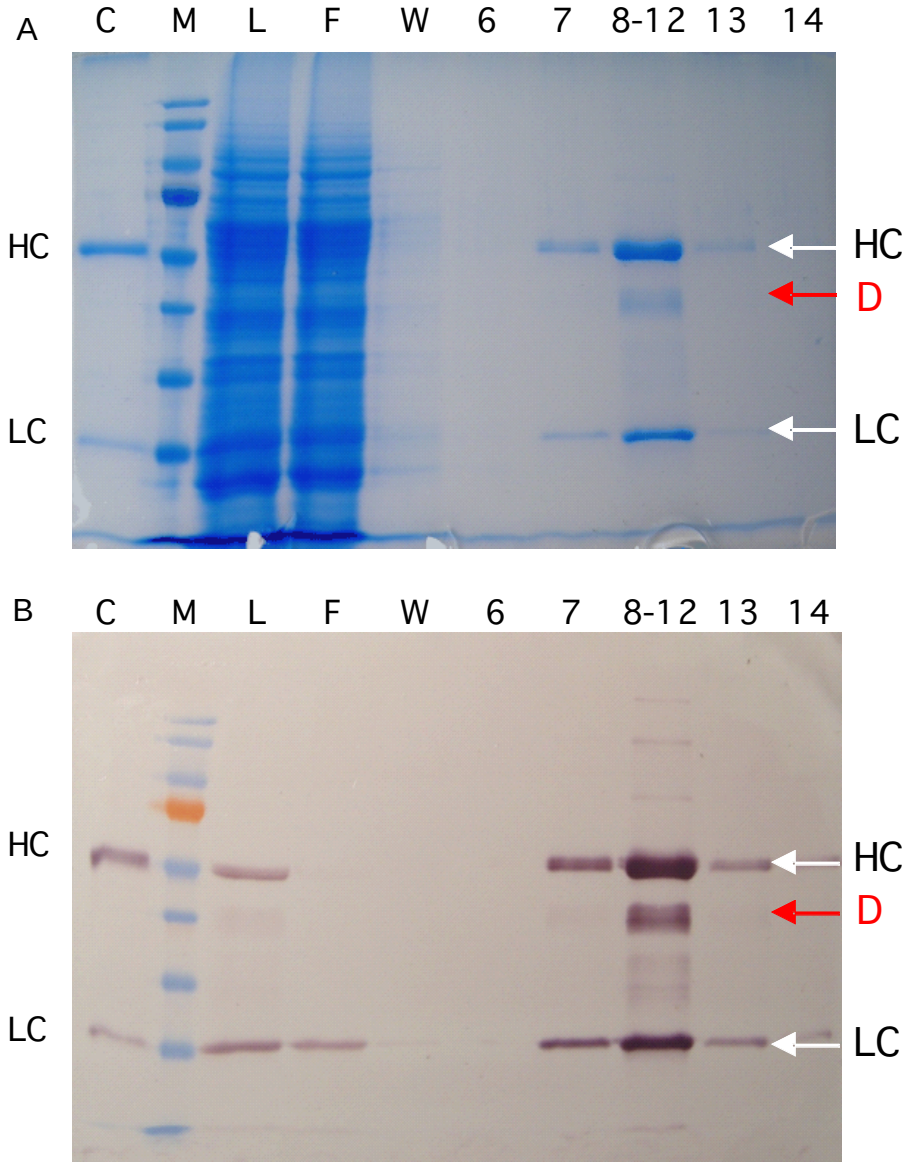

Fig. 1. Analysis of the 2G12 purification process. (*A*) Coomassie staining after SDS/PAGE (reducing conditions) of fractions taken during 2G12 purification by protein A chromatography. (*B*) Western blot (reducing conditions) of the 2G12 heavy and light chains from fractions taken during the same purification process. The presence of the heavy chain (HC), light chain (LC), and degradation products (D) are shown with arrows. An excess of light chain can be seen in the flow-through lane. Lanes are marked as follows: C, control; M, markers; L, load; F, flow-through; and W, wash. Numbers represent elution fractions 6, 7, 8–12 pooled, 13, and 14.

Virus Neutralization Assays. Virus-neutralization assays were performed with matched preparations of 2G12 derived from maize and CHO cells, i.e., preparations matched in terms of the amount of antibody capable of binding to the antigen, as established by Biacore analysis. HIV-1 neutralization was determined by using a syncytium inhibition assay. The 50% inhibiting concentration (IC_{50}) of maize-derived 2G12 was 2.88 μg/ml, whereas that of the control (CHO) antibody was 8.11 μg/ml, demonstrating that the maize 2G12 preparation was nearly three times as potent as its CHO counterpart. This apparent increase in neutralization efficiency probably reflects the presence of dimers and other aggregates in the maize preparation, which are known to have higher neutralization ability than monomeric antibody forms (33).

Discussion

Protein microbicides against HIV are promising alternatives to the current generation of small-molecule drugs and have performed well in preclinical and clinical studies (34). However, producing and distributing such drugs in key HIV-endemic areas is difficult and expensive because of the lack of fermenter capacity and the absence of a cold chain, meaning that those people most likely to benefit from the microbicides are effectively excluded on economic grounds. Many therapeutic heterologous proteins have been expressed in plants, including a large number of vaccine candidates and various recombinant antibody formats for veterinary and human use (35). The advantages of plants over fermenter-based systems are described in refs. 18, 19, and 36, but, in the context of our study, the most important benefits of using maize seeds include the enhanced stability of recombinant proteins accumulating in the endosperm, which means that a cold chain is not necessary for product distribution (17, 37, 38), and the economy of large-scale production both in terms of the upstream phase and downstream processing, which will make the product affordable without detracting from its safety or performance (17, 21).

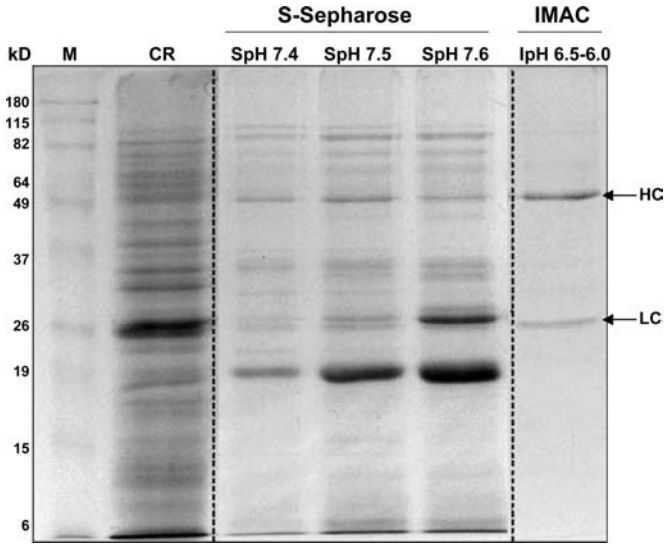

Fig. 2. SDS/PAGE of selected fractions from non-protein A purification procedure. Lanes are marked as follows: M, markers; CR, transgenic maize crude extract; SpH 7.4-SpH 7.6, pool of eluted fractions at pH 7.4–7.6 from S-Sepharose FF cation exchanger; and IpH 6.5–6.0, pool of eluted fractions at pH 6.5–6.0 from IMAC.

One often-cited limitation of plant-based production systems is the low yield of recombinant protein (21). We have previously demonstrated that the methylation clock in plants can be reset by dedifferentiation, resulting in significant increases in levels of transgene expression (39). By subjecting immature zygotic embryos from one 2G12 highly expressing line to dedifferentiation and subsequent regeneration, we were able to not only boost accumulation levels by 30–40% compared with plants restricted to the sexual reproduction cycle but also to eliminate most of the seed-to-seed variation.

Another criticism of plant-based platforms for pharmaceutical protein production is that, although plants process recombinant proteins in a very similar way to mammalian cells, differences in glycosylation result in plant-specific glycan structures that could modify the biological properties of recombinant human glycoproteins produced in plants. High mannose glycans are the same in plants and mammalian cells, but the processing of complex glycans is distinct, resulting in the inability of plants to incorporate sialic acid on N-linked glycans and the introduction of fucose and xylose linkages that are not found in mammalian cells (40). For injectable products, the presence of plant glycans might result in immunogenic reactions, but because our product is intended for topical application, the potential immunogenicity should not be an issue (21).

To investigate these differences functionally, we set out to produce a well characterized therapeutic antibody in maize and test its efficacy against the same antibody produced in CHO cells. The 2G12 heavy and light chain genes were introduced into maize, and one homozygous line (3C) was identified after screening over two generations that expressed 2G12 at levels exceeding 75 μg per gram of dry seed weight, which is considerably higher than the typical levels for pharmaceutical proteins expressed from the plant nuclear genome.

Table 2. Q-TOF analysis of tryptic glycopeptides from the purified, assembled 2G12 antibody isolated from maize seed endosperm and the 40-kDa degradation product thereof

	Intensity	
Structure	Full-size 2G12	Degraded 2G12
Nonglycosylated	11.5	16.1
Oligomannose-type N-glycans	5.9	nd
Complex-type N-glycans	27.9	67.0
Single GlcNAc	54.6	16.9

The group of oligomannose-type glycans consisted of the N-glycans $Man_5GlcNAc_2$ to $Man_9GlcNAc_2$ in similar orders of magnitude. Complex-type N-glycans comprised the structures $MUXF^3$, $MMXF^3$, $GnMXF^3$/$MGnXF^3$, and $GnGnXF^3$ in roughly similar amounts in both samples. nd, none detected.

The first step was to obtain a preparation of the antibody equivalent in terms of purity to the product from CHO cells. The downstream processing of antibodies produced in plants plays an important role in the technical and economical feasibility of large-scale applications. For this reason, in addition to the use of a standard protein A affinity chromatography purification method, we evaluated a two-step non-protein A method as an alternative. Protein A binds with great affinity and specificity to the Fc portion of full-length antibodies and is widely used in commercial antibody production. However, the resin is very expensive (particularly when used in industrial scale chromatography columns) and can undergo degradation (41) and leach into the final product, causing toxicity (42). However, conventional chromatographic methods are suitable for large-scale processes because they are less expensive, easier to use, and largely resistant to chemical and biological degradation (43, 44). Therefore, we developed a two-step strategy that exploited two intrinsic properties of IgG1 monoclonal antibodies: their pI and their ability to bind metal ions because of the presence of a surface-accessible histidine-rich region near the C terminus. The procedure consisted of an initial S-Sepharose FF cation exchange chromatography step, followed by an IMAC step, using Zn^{2+}-IDA-agarose, yielding 2G12 at 90% purity and achieving ≈50–60% recovery. As well as using inexpensive media, the antibody could be recovered under mild nondenaturing conditions, a significant advantage for therapeutic proteins.

Once purified, the maize 2G12 was characterized by SDS/PAGE and Western blot analysis. The heavy and light chains comigrated with their counterparts from the CHO antibody (50 kDa for the heavy chain, 25 kDa for the light chain) confirming their structural integrity and approximately equivalent molecular masses. Western blot analysis suggested that the light chain was produced in excess (visible in the flow-through lane in Fig. 1*B*) as has been reported for other full-length antibodies (45, 46).

A comparison of the functional properties of maize 2G12 and its CHO counterpart was carried out by performing antigen

Table 1. Biacore characterization of the antigen-binding properties of the maize and CHO antibodies

Binding ratios	^{CHO}2G12	^{zm}2G12	CHO, %
protein-L/protein-A	0.7114 ± 0.0051	0.6147 ± 0.0038	86.4 ± 0.5
gp120/protein-A	0.2849 ± 0.0018	0.2473 ± 0.0017	86.8 ± 0.6
gp120/protein-L	0.4002 ± 0.0023	0.4022 ± 0.0022	100.5 ± 0.6

Protein-L/protein-A measures the presence of degradation products and unassembled or partially assembled antibody chains. gp120 surfaces/protein-A and gp20/protein-L measure the specific antigen binding activity.

binding assays, using SPR spectroscopy and virus-neutralization assays based on syncytium inhibition. We found no significant difference between the two intact antibodies in terms of their binding activity as determined through Biacore analysis, although binding to the protein A surface was higher in the case of the maize antibody because of the presence of heavy chain degradation products that were unable to bind the antigen. We anticipated that the antibodies would also be equivalent in terms of their neutralization activity, but we found that the maize antibody was three times more efficient than the commercially available CHO 2G12. The most likely explanation for this is the presence of aggregates in the maize-derived 2G12, because aggregates are known to have greater virus neutralizing activity. The absence of aggregates and degradation products in the CHO preparation reflects the addition processing steps that have been carried out on this antibody, specifically the gel filtration step, which removes products with identical physicochemical and functional properties but distinct molecular masses.

We conclude that 2G12, an HIV-neutralizing monoclonal antibody, has been expressed successfully in maize endosperm at a mean concentration of $\approx$75 μg per gram of dry seed weight. The antibody can be recovered efficiently from maize kernels at 90% purity, using two distinct processing strategies, one based on protein A affinity chromatography and the other based on a more economical two-step process, using ion exchange chromatography and IMAC, which should be applicable on a large scale. The purified maize-derived 2G12 was physically identical to its CHO-derived counterpart with the exception of its glycan structure, and in functional terms it showed identical antigen-binding activity but nearly three times the efficacy in HIV-neutralization assays, probably because of the presence of aggregates. Further antibodies produced in maize are currently undergoing testing to see whether the encouraging results achieved with 2G12 can be repeated for other molecules. If so, this would represent a remarkable opportunity to produce therapeutic antibodies inexpensively and on a massive scale, making it much more likely that antibody-based therapeutics could be made more widely available at a much lower cost than currently possible.

Methods

Transformation Vectors. All transformation constructs were based on the binary vector pTRA, a derivative of pPAM (GenBank accession no. AY027531) containing two tobacco RB7 scaffold attachment regions flanking the expression cassette (47). The coding regions of the 2G12 heavy and light chains (obtained from Polymun) contained an N-terminal signal sequence targeting the polypeptide to the secretory pathway. The expression cassette comprised the endosperm-specific rice glutelin-1 promoter, the Tobacco Etch Virus 5′ leader, the coding region, and the CaMV 35S terminator, resulting in final constructs pTRAgtiGH and pTRAgtiGL. The third construct (pTRAuxbar) contained the *bar* gene between the constitutive maize ubiquitin-1 promoter and 35S terminator. All expression cassettes contained the maize ubiquitin-1 first intron.

Transformation, Selection, and Regeneration of Transgenic Plants. Immature zygotic embryos of the South African elite white maize genotype M37W were transformed by particle bombardment at 10–14 days after pollination as described in ref. 26. Bombarded embryos were transferred in the dark every 2 weeks to fresh N6-based medium containing 3 mg/liter phosphinothricin (PPT). Four to 6 weeks after bombardment, pieces of PPT-resistant embryogenic type I callus were transferred to regeneration medium containing 3 mg/liter PPT for 2–4 weeks with a 16-h/8-h (day/night) photoperiod. Developing plantlets with well formed shoots and roots were hardened off in soil. All experiments were performed at 25°C. Embryos from a lead event (3C) were subjected to a dedifferentiation regeneration cycle by reintroduction into callus culture following the procedure described above for transformation and regeneration, in the absence of selection. Numerous regenerants were recovered, and all were shown by DNA analysis to be clones of the original transformant 3C.

Screening Seeds for 2G12 Expression. We analyzed five T_1 seeds per transgenic event by removing and germinating the embryos and extracting total soluble protein from the endosperm, using three volumes of PBS. After centrifugation, 4 μl of supernatant was spotted individually onto a nitrocellulose membrane and blocked for 1 h. The membrane was then incubated at room temperature overnight with an alkaline phosphatase-labeled antibody [goat anti-human IgG Fc chain (Sigma A9544) or goat anti-human kappa chain (Sigma A3813)]. Blots were washed three times with PBS buffer supplemented with 2% Tween-20 and the signal detected with Sigma FAST BCIP/NBT dissolved in distilled water.

SDS/PAGE, Western blot, and ELISAs. Proteins were separated by SDS/PAGE in a 12% denaturing gel, and Western blots were carried out by using standard methods and the same antibodies as used in the dot-blot anaylses described above. For quantitation of 2G12 expression, 10 T_2 seeds each from events 3C (high expression) and 1G (low expression) were analyzed by ELISA. The capture antibody [goat anti-human kappa chain (Sigma K3502) or goat anti-human IgG Fc chain (Sigma K2136) at 1:200 dilution] was coated directly onto 96-well Nunc-Immuno Maxisorp surface plates (Nalge Nunc) and incubated overnight at 4°C. Plates were blocked with 5% nonfat milk for 2 h and then washed with PBS containing 0.2% Tween-20. Samples diluted in this buffer were added to the wells and incubated for 2–3 h at room temperature. After washing, the alkaline phosphatase-labeled antibodies used for dot-blot analysis were added to the plates at 1:1,000 dilution and incubated at room temperature for 2 h before the signal was detected with *p*-nitrophenyl phosphate. After 4–5 min, the signal was quantified by measuring the absorbance at 415 nm.

Protein A Affinity Purification. Maize seeds were crushed to a fine powder and extracted overnight at 4°C in five volumes of buffer [1$\times$ PBS, 5 mM EDTA, and 1 mM 2-mercaptoethanol (pH 7.4)]. Insoluble material and fat deposits were removed by centrifuging twice at 8,000 $\times$ *g* for 30 min at 4°C. The supernatant was filtered and the antibody concentration determined by Biacore analysis. The sample was loaded onto a protein A affinity column (ceramic HyperDF) at a flow rate of 2 ml/min. The column was washed with PBST and PBS, and the antibody was eluted with 100 mM glycine (pH 3.6) containing 100 mM fructose, and buffered with one-tenth volume of 1 M acetate/sodium acetate (pH 4.75). Antibody-containing fractions were identified by the droplet Bradford method, and 2G12 concentrations were determined by Biacore. Fractions containing $>$50 μg/ml were pooled and dialyzed for 2 days against 10 mM acetate/sodium acetate (pH 4.75) containing 1 mM EDTA and 0.1 mM 2-mercaptoethanol (omitted in final dialysis step). The dialyzed antibody was concentrated by ultrafiltration, using spin-columns with a 30-kDa molecular mass cut-off. Biacore was used to monitor all steps. The final antibody concentration was determined by Biacore, using CHO-2G12 as standard.

Non-Protein A Purification. Maize seeds (5 g) were extracted by blending for 10 min in 20 ml of 20 mM sodium phosphate buffer (pH 6.0) followed by rotary mixing for 1 h at room temperature. After centrifuging at 10,000 $\times$ *g* for 30 min, the supernatant was collected and passed through a 0.45-μm filter. Cation exchange chromatography was carried out on 1 ml of filtered extract, using S-Sepharose FF (0.5 ml of moist adsorbent) at 4°C after equilibration with 20 mM sodium phosphate buffer (pH 6.0). The column was then washed with a series of buffers of increasing pH (20 mM sodium phosphate at pH 6.0, then pH 6.5, then pH 6.75), and the antibody was eluted by using stepwise increases in pH from pH 7.0 to 8.0. All chromatographic steps were performed at a flow rate of 0.5 ml/min. Collected fractions (1.5 ml) were analyzed by ELISA, and total protein was determined by the Bradford assay. Pooled eluted fractions at pH 7.4–7.6 (4.5 ml) were loaded onto IDA-agarose resin (0.5 ml of moist adsorbent) at 4°C charged with Zn^{2+}, using a solution of zinc chloride (2 ml, 0.1 M) and equilibrated with sodium phosphate buffer (50 mM, pH 7.5) containing 1 M NaCl. The column was washed with 10 ml of equilibration buffer, and elution was achieved by using a stepwise pH gradient (pH 6.5–5.0 in 50 mM sodium phosphate buffer containing 1 M NaCl). Collected fractions were analyzed by ELISA, and total protein was determined by the Bradford assay.

Surface Plasmon Resonance (SPR) spectroscopy. SPR assays were performed at 25°C with a flow rate of 30 μl/min, using a BIACORE 2000 (GE Healthcare), a CM5-rg sensor chip and HBS-EP as the running buffer. Ligands were immobilized by using the standard EDC/NHS protocol. The first flow cell was activated/deactivated and used for blank subtraction. Immobilization buffers and levels were 4,200 resonance units (RU) for protein A [200 μg/ml and 10 mM sodium acetate (pH 4.5)], 2,300 RU for protein L [100 μg/ml and 10 mM sodium acetate (pH 4.5)] and 12,000 RU for gp120 [(10 mM sodium acetate (pH 4.75)]. Protein A and protein L surfaces were regenerated with a 30-s pulse of 30 mM HCl,

whereas the gp120 surface was regenerated with 0.5 M citrate, pH 3.0. CHO-derived 2G12 was used as standard. Triplicate dilution series were injected simultaneously over the four flow cells. The blank-subtracted endpoint signals were used to evaluate the quality of the preparations. All surfaces exhibited high binding capacities and strong mass-transport limitation, resulting in linear dose-response relationships and constant binding rates up to an antibody concentration of 1 μg/ml. Data were evaluated by using BIAevaluation software, Version 4.0 and Microcal Origin software, Version 5.0 software. The binding signal ratios were derived from the endpoint responses for extracts or determined by pair wise plotting of the binding signals obtained at different dilutions and linear regression analysis for purified preparations.

N-Glycan Analysis. Antibody bands from Coomassie-stained SDS/PAGE were destained, carbamidomethylated, digested with trypsin, and extracted from gel pieces as described in ref. 44. The subsequent fractionation of the peptides by capillary reversed phase chromatography with detection by a Q-TOF Ultima Global (Waters Micromass) mass spectrometer was performed as described in refs. 48 and 49.

The MS-data from the tryptic peptides were analyzed as described in ref. 49 and compared with datasets generated by *in silico* tryptic digestion of the 2G12 coding, using the PeptideMass program (available at www.expasy.org/tools/peptide-mass.html). Based on the tryptic peptide datasets, tryptic glycopeptide datasets were generated by the addition of glycan mass increments to the masses of the two identified glycopeptides.

Neutralization Assays. HIV-1 neutralization was assessed by using a syncytium inhibition assay. Ten twofold serial dilutions (start concentration: 100 μg/ml) of ^{Zm}2G12, ^{CHO}2G12, and a nonneutralizing control were preincubated with HIV-1 RF at 10 half-maximum tissue culture infectious dose (2, 42) per milliliter for 1 h at 37°C. CD4-positive human AA-2 cells were added at a density of 4×10^5 cells per milliliter and incubated for a further 5 days. Experiments were performed with eight replicates per antibody dilution step. The presence of one or more syncytium per well after 5 days was scored as positive for infection. The 50% inhibiting concentrations (IC_{50}) were calculated by the method of Reed and Muench (50), using the concentrations present during the antibody-virus preincubation step.

ACKNOWLEDGMENTS. This work was supported by European Union Framework 6 Program—The Pharma-Planta Integrated Project Grant LSH-2002–1.2.5–2; Acciones Complementarias (MEC) Grant BIO2005–24826-E; Generalitat de Catalunya Grant 2005SGR118; the Ramon y Cajol Program; and Center CONSOLIDER, Ministerio de Educación y Ciencia, Spain.

1. WHO/UNAIDS (2006) *AIDS Epidemic Update: Special Report on HIV/AIDS, December 2006* (UN/WHO, Geneva).
2. Poignard P, Saphire EO, Parren PW, Burton DR (2001) *Annu Rev Immunol* 19:253–274.
3. Burton DR, Pyati J, Koduri R, Sharp SJ, Thornton GB, Parren PW, Sawyer LS, Hendry RM, Dunlop N, Nara PL, *et al.* (1994) *Science* 266:1024–1027.
4. Trkola A, Purtscher M, Muster T, Ballaun C, Buchacher A, Sullivan N, Srinivasan K, Sodroski J, Moore JP, Katinger H (1996) *J Virol* 70:1100–1108.
5. Conley AJ, Kessler JA, II, Boots LJ, Tung JS, Arnold BA, Keller PM, Shaw AR, Emini EA (1994) *Proc Natl Acad Sci USA* 91:3348–3352.
6. Zwick MB, Labrijn AF, Wang M, Spenlehauer C, Saphire EO, Binley JM, Moore JP, Stiegler G, Katinger H, Burton DR, *et al.* (2001) *J Virol* 75:10892–10905.
7. Kwong PD, Wyatt R, Majeed S, Robinson J, Sweet RW, Sodroski J, Hendrickson WA (2000) *Structure* 8:1329–1339.
8. Trkola A, Pomales AB, Yuan H, Korber B, Maddon PJ, Allaway GP, Katinger H, Barbas CF, Burton DR, Ho DD, *et al.* (1995) *J Virol* 69:6609–6617.
9. Ugolini S, Mondor I, Parren PW, Burton DR, Tilley SA, Klasse PJ, Sattentau QJ (1997) *J Exp Med* 186:287–1298.
10. Fouts TR, Binley JM, Trkola A, Robinson JE, Moore JP (1997) *J Virol* 71:2779–2785.
11. Mascola JR, Stiegler G, VanCott TC, Katinger H, Carpenter CB, Hanson CE, Beary H, Hayes D, Frankel SS, Birx DL, *et al.* (2000) *Nature Med* 6:207–210.
12. Wyatt R, Kwong PD, Desjardins E, Sweet RW, Robinson J, Hendrickson WA, Sodroski JG (1998) *Nature* 393:705–711.
13. Hofmann-Lehmann R, Vlasak J, Rasmussen RA, Jiang S, Li PL, Baba TW, Montefiori DC, Bernacky BJ, Rizvi TA, Schmidt R, *et al.* (2001) *J Virol* 75:7470–7480.
14. Stiegler G, Armbruster C, Vcelar B, Stoiber H, Kunert R, Michael NL, Jagodzinski LL, Ammann C, Jäger W, Jacobson J, *et al.* (2002) *AIDS* 16:2019–2025.
15. Chu L, Robinson D K (2001) *Curr Opin Biotechnol* 12:180–187.
16. WHO/UNAIDS (2006) *Progress in Scaling up Access to HIV Treatment in Low and Middle-income Countries (June Fact Sheet)* (WHO/UNAIDS, Geneva).
17. Stoger E, Ma JKC, Fischer R, Christou P (2005) *Curr Opin Biotechnol* 16:167–173.
18. Ma JKC, Drake PMW, Christou P (2003) *Nat Rev Genet* 4:794–805.
19. Fischer R, Stoger E, Schillberg S, Christou P, Twyman RM (2004) *Curr Opin Plant Biol* 7:1–7.
20. Stoger E, Sack M, Fischer R, Christou P (2002) *Curr Opin Biotechnol* 13:161–166.
21. Twyman RM, Schillberg S, Fischer R (2005) *Expert Opin Emerg Drugs* 10:185–218.
22. Hood EE, Witcher DR, Maddock S, Meyer T, Baszczynski C, Bailey M, Flynn P, Register J, Marshall L, Bond D, *et al.* (1997) *Mol Breeding* 3:291–306.
23. Witcher DR, Hood EE, Peterson D, Bailey M, Bond D, Kusnadi A, Evangelista R, Nikolov Z, Wooge C, Mehigh R, *et al.* (1998) *Mol Breeding* 4:301–312.
24. Hood EE, Woodard SL, Horn ME (2002) *Curr Opin Biotechnol* 13:630–635.
25. Nicholson L, Gonzalez-Melendi P, van Dolleweerd C, Tuck H, Perrin Y, Ma JK, Fischer R, Christou P, Stoger E (2005) *Plant Biotechnol J* 3:115–127.
26. Drakakaki G, Marcel S, Glahn RP, Lund EK, Pariagh S, Fischer R, Christou P, Stoger E (2005) *Plant Mol Biol* 59:869–880.
27. Kohli A, Leech M, Vain P, Laurie DA, Christou P (1998) *Proc Natl Acad Sci USA* 95:7203–7208.
28. Kohli A, Gahakwa D, Vain P, Laurie DA, Christou P (1999) *Planta* 208:88–97.
29. Abranches R, Santos AP, Wegel E, Williams S, Castilho A, Christou P, Shaw PJ, Stöger E (2000) *Plant J* 24:713–723.
30. Kohli A, Twyman RM, Abranches R, Wegel E, Stoger E, Christou P (2003) *Plant Mol Biol* 52:247–258.
31. Kohli A, Gonzales-Melendi P, Abranches R, Capell T, Stoger E, Christou P (2006) *Plant Signal Behav* 1:185–195.
32. Schuster M, Jost W, Mudde GC, Wiederkum S, Schwager C, Janzek E, Altmann F, Stadlmann J, Stemmer C, Gorr G (2007) *Biotechnol J* 2:700–708.
33. Wolbank S, Kunert R, Stiegler G, Katinger H (2003) *J Virol* 77:4095–4103.
34. De Clercq E (2007) *Nat Rev Drug Discov* 6:1001–1018.
35. Floss DM, Falkenburg D, Conrad U (2007) *Transgen Res* 16:315–332.
36. Stoger E, Sack M, Perrin Y, Vaquero C, Torres E, Twyman R, Christou P, Fischer R (2002) *Mol Breeding* 9:149–158.
37. Stoger E, Vaquero C, Torres E, Sack M, Nicholson L, Drossard J, Williams S, Keen D, Perrin Y, Christou P, *et al.* (2000) *Plant Mol Biol* 42:583–590.
38. Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R (2003) *Trends Biotechnol* 21:570–578.
39. Kohli A, Prynne MQ, Miro B, Pereira A, Twyman RM, Capell T, Christou P (2004) *Mol Breeding* 13:177–191.
40. Bakker H, Bardor M, Molthoff JW, Gomord V, Elbers I, Stevens LH, Jordi W, Lommen A, Faye L, Lerouge P, *et al.* (2001) *Proc Natl Acad Sci USA* 98:2899–2904.
41. Platis D, Sotriffer CA, Labrou NE (2006) *J Chromatogr A* 1128:138–151.
42. Terman DS, Bertram JH (1985) *Eur J Cancer Clin Oncol* 21:1115–1122.
43. Labrou NE (2003) *J Chromatogr* B 790:67–78.
44. Shukla AA, Hubbard B, Tressel T, Guhan S, Low D (2007) *J Chromatogr B* 841:28–39.
45. Giritch A, Marillonnet S, Engler C, van Eldik G, Botterman J, Klimyuk V, Gleba Y (2006) *Proc Natl Acad Sci USA* 103:14701–14706.
46. Law RD, Russell DA, Thompson LC, Schroeder SC, Middle CM, Tremaine MT, Jury TP, Delannay X, Slater SC (2006) *Biochim Biophys Acta* 1760:1434–1444.
47. Sack M, Paetz A, Kunert R, Bomble M, Hesse F, Stiegler G, Fischer R, Katinger H, Stoeger E, Rademacher T (2007) *FASEB J* 21:1655–1664.
48. Kolarich D, Altmann F (2000) *Anal Biochem* 285:64–75.
49. Van Droogenbroeck B, Cao J, Stadlmann J, Altmann F, Colanesi S, Hillmer S, Robinson DG, Van Lerberge E, Terryn N, Van Montagu M, *et al.* (2006) *Proc Natl Acad Sci USA* 104:1430–1435.
50. Reed LJ, Muench H (1938) *Am J Hyg* 27:493–497.

Recombinant antibody 2G12 produced in maize endosperm efficiently neutralizes HIV-1 and contains predominantly single-GlcNAc N-glycans.

Rademacher, T; Sack, M; Arcalis, E; Stadlmann, J; Balzer, S; Altmann, F; Quendler, H; Stiegler, G; Kunert, R; Fischer, R and Stoger, E (2008).
Plant Biotechnol J 6(2): 189-201.

Plant Biotechnology Journal (2008) **6**, pp. 189–201 doi: 10.1111/j.1467-7652.2007.00306.x

Recombinant antibody 2G12 produced in maize endosperm efficiently neutralizes HIV-1 and contains predominantly single-GlcNAc *N*-glycans

Thomas Rademacher[1], Markus Sack[1], Elsa Arcalis[1], Johannes Stadlmann[2], Simone Balzer[1], Friedrich Altmann[2], Heribert Quendler[3], Gabriela Stiegler[3], Renate Kunert[3], Rainer Fischer[1] and Eva Stoger[1,*]

[1]*Institute for Molecular Biotechnology, Biology VII, RWTH Aachen, Worringerweg 1, 52074 Aachen, Germany*
[2]*Department for Chemistry, Glycobiology Division, University of Natural Resources and Applied Life Sciences, Muthgasse 18, 1190 Vienna, Austria*
[3]*Polymun Scientific, Nussdorfer Laende 11, 1190 Vienna, Austria*

Received 30 June 2007;
revised 6 September 2007;
accepted 27 September 2007.
**Correspondence* (fax +492418020145;
e-mail stoger@molbiotech.rwth-aachen.de)

Summary

Antibody 2G12 is one of a small number of human immunoglobulin G (IgG) monoclonal antibodies exhibiting potent and broad human immunodeficiency virus-1 (HIV-1)-neutralizing activity *in vitro*, and the ability to prevent HIV-1 infection in animal models. It could be used to treat or prevent HIV-1 infection in humans, although to be effective it would need to be produced on a very large scale. We have therefore expressed this antibody in maize, which could facilitate inexpensive, large-scale production. The antibody was expressed in the endosperm, together with the fluorescent marker protein *Discosoma* red fluorescent protein (DsRed), which helps to identify antibody-expressing lines and trace transgenic offspring when bred into elite maize germplasm. To achieve accumulation in storage organelles derived from the endomembrane system, a KDEL signal was added to both antibody chains. Immunofluorescence and electron microscopy confirmed the accumulation of the antibody in zein bodies that bud from the endoplasmic reticulum. In agreement with this localization, *N*-glycans attached to the heavy chain were mostly devoid of Golgi-specific modifications, such as fucose and xylose. Surprisingly, most of the glycans were trimmed extensively, indicating that a significant endoglycanase activity was present in maize endosperm. The specific antigen-binding function of the purified antibody was verified by surface plasmon resonance analysis, and *in vitro* cell assays demonstrated that the HIV-neutralizing properties of the maize-produced antibody were equivalent to or better than those of its Chinese hamster ovary cell-derived counterpart.

Keywords: biopharmaceutical, human immunodeficiency virus antibody, N-glycosylation, protein body, red fluorescent protein, transgenic maize seed.

Introduction

Plants are being developed as inexpensive, large-scale production systems to meet the future demand for complex biopharmaceuticals (Daniell *et al*., 2001; Ma *et al*., 2005; Schillberg *et al*., 2005). Numerous studies have shown that transgenic plants can express correctly folded and functional antibodies, including full-size immunoglobulin G (IgG) and secretory immunoglobulin A (sIgA) molecules (Ludwig *et al*., 2004). Plant systems could therefore be ideal for the production of molecules such as 2G12, a human monoclonal antibody that recognizes oligomannose-type *N*-glycans on the human immunodeficiency virus-1 (HIV-1) gp120 envelope protein. This antibody has a potent and broad HIV-1-neutralizing activity *in vitro*, and passive transfer studies in primates show that it is able to control infection (Mascola *et al*., 1999, 2000; Mascola, 2002) and prevent transmission, both parenterally and through mucosal tissues (Veazey *et al*., 2003). Oral challenge studies in neonatal macaque monkeys support the use of 2G12 to prevent the transmission of HIV-1 to human infants (Hofmann-Lehmann *et al*., 2001; Ferrantelli *et al*., 2004), and antibody cocktails including

2G12 reduce the rate of viral rebound after the cessation of antiretroviral treatment in some human patients (Trkola *et al*., 2005). The use of such antibody cocktails *in vivo* for passive immunotherapy and prophylaxis requires multiple high doses (Trkola *et al*., 2005). The global demand for such antibodies would therefore quickly overwhelm current capacity in traditional, fermenter-based production systems, but could be met using plants.

One challenge that needs to be addressed before therapeutic antibodies can be produced commercially in plants is glycan chain structure. *N*-glycans are correctly attached to the asparagine-297 (Asn297) residue of full-size plant-derived antibodies, but their structures differ from the glycoforms found in humans, raising concerns that this might limit their therapeutic potential. Whether plant glycans are immunogenic or allergenic is still a matter of debate, but they can be immunoreactive and, in particular, β1,2-xylose and α1,3-fucose residues have been identified within epitopes occasionally recognized by the IgEs of patients with allergies to plant food or pollen (van Ree *et al*., 2000; Fotisch and Vieths, 2001). One strategy for producing antibodies lacking plant-specific complex-type *N*-glycans is to target the antibody for retention in the endoplasmic reticulum (ER), as ER-associated *N*-glycosylation generates oligomannose-type (OMT) *N*-glycans, which are identical in mammals and plants (Gomord *et al*., 2005). An additional advantage of this strategy is that antibodies accumulate to higher levels in the ER than in other compartments (Schouten *et al*., 1996). ER retention can be achieved by fusing a KDEL sequence to the C-terminus of a normally secreted protein. The efficiency of antibody retrieval to the ER and the resulting *N*-glycan profiles have been examined in several studies, although these have focused mainly on leaf-based expression in dicot species. More recently, the fate of a KDEL-tagged antibody in tobacco seeds has been described (Petruccelli *et al*., 2006). By contrast with leaves, where the antibody was found exclusively in the ER, the seed-derived antibody was partially secreted and sorted to protein storage vacuoles (PSVs).

Cereals are advantageous for the production of pharmaceutical proteins, but little is known about the characteristics of antibodies retained in the ER of cereal seed endosperm, a highly specialized storage tissue with a triploid genome. The majority of proteins in the cereal endosperm accumulate in defined protein storage compartments: PSVs and ER-derived protein bodies (PBs). Although most of the proteins found in PSVs pass through the Golgi, where plant-specific *N*-glycan modification occurs, the proteins in PBs (mostly prolamins) originate directly from the ER (Lending and Larkins, 1989). The specialized protein storage function makes the endosperm of cereal seeds particularly useful for the expression of recombinant proteins. The possibility of long-term storage in dry seed offers more flexibility in manufacturing, and the low content of secondary metabolites and contaminating proteins simplifies purification. Although PSVs and PBs are both abundant in other cereal species, such as rice, the large endosperm of maize seeds is dominated by prolamin-type storage proteins deposited in ER-derived PBs (Lending and Larkins, 1989). Important components of the study described in this article were therefore to determine the subcellular localization of 2G12 expressed in maize endosperm, to describe its *N*-glycan structures and to compare the performance of the plant-derived antibody with its counterpart produced using the standard Chinese hamster ovary (CHO) cell system.

Maize is favoured amongst the four major cereals used for recombinant pharmaceutical production because of its high biomass yields. However, as maize is a wind-pollinated species, there is a risk of outcrossing to food crops, as well as the inadvertent post-harvest mixing of seeds when pharmaceutical maize is grown on an agricultural scale. In order to alleviate these concerns, a visual marker for identity preservation is desirable. Green fluorescent protein (GFP) is used as a visual marker for the production and detection of transgenic plants (Harper *et al*., 1999; Stewart, 2005). In this paper, it is shown that the *Discosoma* red fluorescent protein (DsRed) is suitable as an identity marker for pharmaceutical maize. The ability to detect DsRed with quite simple equipment in whole plants, pollen and seeds makes it a valuable tool for the detection of unintentional outcrosses and admixtures. In addition, DsRed can be used to facilitate and accelerate breeding into elite lines.

In this study, it has been demonstrated that maize is a suitable production host for antibody 2G12. The antigen-binding properties of the plant-derived antibody and its potency for neutralizing HIV were investigated in direct comparison with results for its CHO cell-derived counterpart, and the full functionality of the corn-derived antibody was confirmed.

Results

Generation of identity-preserved transgenic maize plants producing 2G12

Both antibody chains were tagged at the C-terminus with the SEKDEL motif and expressed under the control of the endosperm-specific rice glutelin-1 (*gt-1*) promoter. A co-transformation strategy was chosen so that each antibody gene was coupled to a constitutively expressed marker gene. The heavy chain (HC) gene was arranged in tandem with the

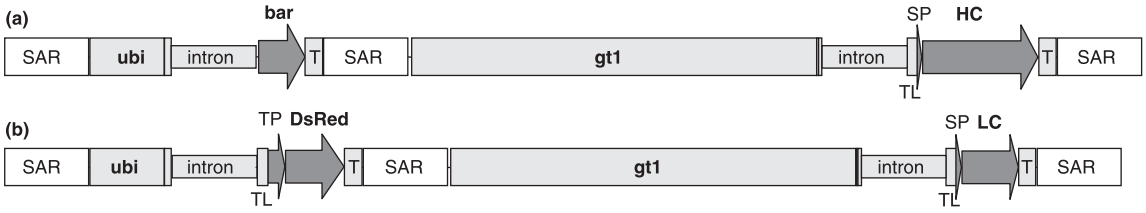

Figure 1 Plant expression cassettes in pTRAb-gGHER (a) and pTRAds-gGLER (b). bar, bialaphos resistance; DsRed, red fluorescent protein from *Discosoma* sp.; gt1, rice glutelin-1 promoter; HC and LC, coding sequence of heavy and light chain of the 2G12 antibody with endoplasmic reticulum (ER) retention signal; intron, intron1 of the maize ubiquitin-1 promoter; SAR, scaffold attachment region; SP, signal peptide; T, terminator of the cauliflower mosaic virus (CaMV) 35S gene; TL, 5′-untransformed region (5′-UTR) of the tobacco etch virus; TP, transit peptide; ubi, maize ubiquitin-1 promoter.

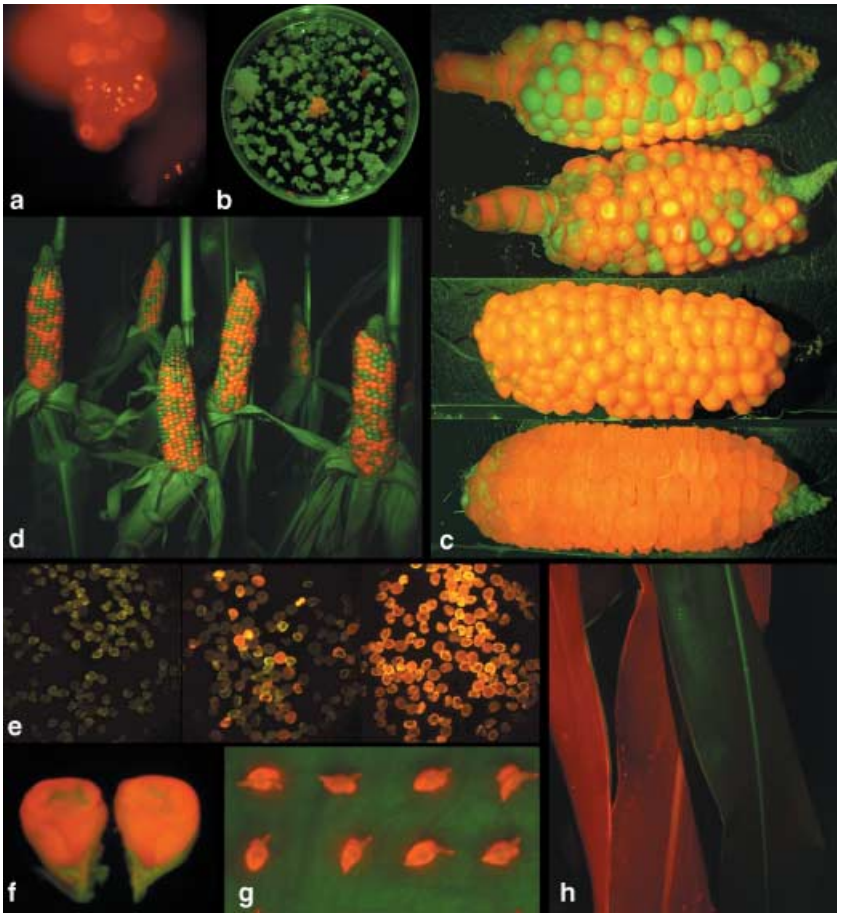

Figure 2 Monitoring of *Discosoma* red fluorescent protein (DsRed) in transgenic maize. (a) Embryogenic callus 48 h after bombardment. (b) Selection of herbicide-resistant callus with high DsRed fluorescence. (c) Identification of transgenic seeds (top to bottom: F_1, T_1, homozygous T_3, homozygous sweetcorn). (d) Wild-type pollinated with hemizygous transgenic. (e) Pollen segregation (left to right: non-transgenic, hemizygous transgenic and homozygous transgenic). (f) Distribution of DsRed in seeds. (g) Isolated embryos from transgenic seeds. (h) DsRed in leaves of mature plants (left, transgenic; right, wild-type).

selectable *bar* gene, and the light chain (LC) gene in tandem with the gene for the screenable marker DsRed (Figure 1). DsRed fluorescence was monitored throughout the transformation and regeneration process. After bombardment of maize callus (*Zea mays* cv. Hill), numerous individual fluorescent cells could be observed, confirming the integration of the DsRed-LC construct (Figure 2a). As expected, only a fraction of the fluorescent cells developed further under selection, indicating co-integration of the second construct containing the selectable marker and HC. Only callus displaying both strong, macroscopically visible DsRed fluorescence and fast growth under selection was propagated further for shoot regeneration (Figure 2b).

After this rigorous selection process, 10 siblings from two independent events showing strong DsRed expression in leaves were transferred to the glasshouse to produce seeds (Figure 2h). In addition to repeated selfing and selection, the initial transgenic Hill line was backcrossed into different elite starch germplasm, as well as into a sugary-type sweetcorn background, as cv. Hill has little agronomic relevance and poor

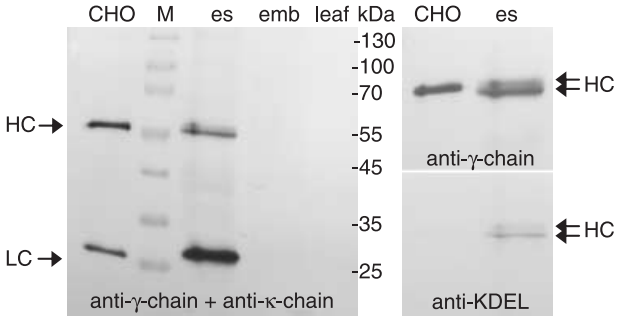

Figure 3 Immunoblots of ^{CHO}2G12 and extracts of different tissues from a ^{Zm}2G12$_{SEKDEL}$ line (es, endosperm; emb, embryo). The blots were probed with anti-γ-chain, anti-κ-chain and anti-KDEL antibodies as indicated. Heavy (HC) and light (LC) chains are indicated by arrows.

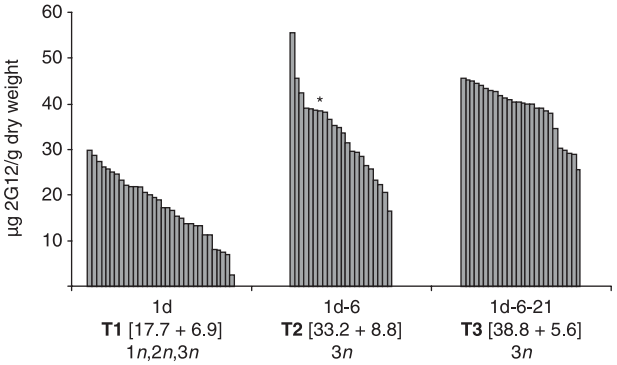

Figure 4 Antibody expression over three Hill generations. T_1 generation contains homozygous and hemizygous seeds (1*n*–3*n* in the triploid endosperm). The subsequent generations contain only homozygous seeds (3*n*). The T_2 plant used for the generation of T_3 seeds is indicated by an asterisk.

yields. Half of the pollen from transgenic T_0 lines expressed DsRed, indicating a single-locus integration of the *DsRed-2G12LC$_{SEKDEL}$* construct (Figure 2e, middle). Screening complete mature ears for DsRed confirmed Mendelian segregation, i.e. 3 : 1 in T_1 seeds resulting from self-pollination and 1 : 1 in the F_1 hybrids for both transgenic events (Figure 2c). Co-segregation of DsRed with both antibody chains was confirmed by immunoblot and surface plasmon resonance analysis of endosperm extracts from fluorescent and non-fluorescent seeds (data not shown). Neither HC nor LC was detectable in non-fluorescent seeds in either of the events, confirming the co-integration of *DsRed-2G12LC$_{SEKDEL}$* and *bar-2G12HC$_{SEKDEL}$* at a single locus. No antibody chain was detected in the embryos of fluorescent seeds, confirming the endosperm-specific activity of the rice *gt-1* promoter in maize (Figures 2f and 3). Immunoblot analysis also showed that both antibody chains were intact (Figure 3). There was a slight difference in mobility between LC and its CHO-derived counterpart, as expected because of the additional KDEL sequence. HC was represented by a double band, revealing two variants differing slightly in molecular mass. This was not caused by C-terminal degradation, as both variants were also detected by an anti-KDEL antibody (Figure 3).

Quantification of ^{Zm}2G12$_{SEKDEL}$ in maize seeds

The accumulation of 2G12 in the endosperm of individual seeds was quantified by surface plasmon resonance spectroscopy. The antibody content in T_1 seeds reached a maximum of 30 µg/g dry weight. Because seeds from the first event produced more antibody than those from the second event, further breeding and detailed antibody characterization was carried out with progeny from event 1. In subsequent generations, individual seeds with the highest antibody levels were selected for further propagation (Figure 2g). The zygosity of the selected plants was determined by DsRed segregation in the pollen (Figure 2e). Because endosperm is triploid, the transgene dose varied from 1*n* to 3*n*. In the inbred lines, the antibody levels increased from 17.7 µg/g in T_1 to 33.2 µg/g in T_2 (most probably because of the higher gene dose in the homozygous T_2 ear), and then to an average of 38.8 µg/g in T_3 (maximum, 50–60 µg/g), with a clear decrease in seed-to-seed variability (Figure 4).

^{Zm}2G12$_{SEKDEL}$ accumulates in protein storage organelles

The subcellular localization of recombinant 2G12 in endosperm tissue was analysed by fluorescence and electron microscopy. The experiments were performed using sections of immature seed, which contain cells at different developmental stages. Our observations were concentrated on young cells in the subaleurone and mid-endosperm (Figure 5a). These cells contain starch grains, a large number of small, spherical zein bodies and several PSVs (Figure 5b). As expected for a protein with a signal peptide and a C-terminal KDEL sequence, intracellular retention was observed, primarily within spherical PBs (Figure 5c). In order to define these structures precisely, ultra-thin sections were analysed by electron microscopy. Gold labelling clearly appeared in the ER-derived zein bodies (Figure 5d), whereas no significant labelling was observed in PSVs (Figure 5e) or any other cell compartment, including the intercellular space. Sections of non-transgenic seeds were not labelled significantly (data not shown).

Purification and characterization of ^{Zm}2G12$_{SEKDEL}$

The antibody was purified from several 50–100-g seed batches using protein A affinity, with continuous monitoring by surface plasmon resonance spectroscopy. Neither the

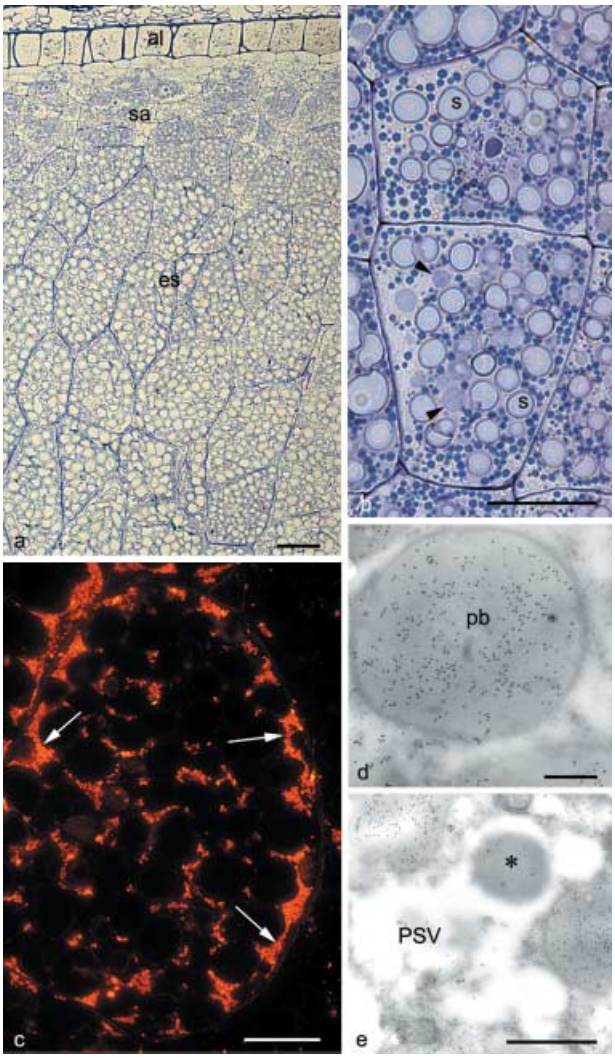

Figure 5 Localization of ^{Zm}2G12$_{SEKDEL}$. (a, b) Spurr sections. Light microscopy, toluidine blue. Note the aleurone (al), the young cells in the subaleurone (sa) and the larger cells in the endosperm (es), packed with starch. In the enlargement in (b), the small, spherical zein bodies, stained in blue, can be identified easily. Note also the presence of protein storage vacuoles (PSVs), stained in lighter blue (arrowhead) and similar in size to the starch grains (s). (c) LRWhite section. Fluorescence microscopy, localization of κ-chain. See the strong labelling on the zein bodies (arrows). No significant labelling can be observed in other cell compartments. (d, e) LRWhite section. Electron microscopy, localization of κ-chain. Abundant gold probes restricted to the α-zein portion of the protein bodies (pb). No significant labelling can be found within the PSV. Globulin inclusion (*). Scale bars: (a) 50 μm; (b, c) 20 μm; (d, e) 0.25 μm.

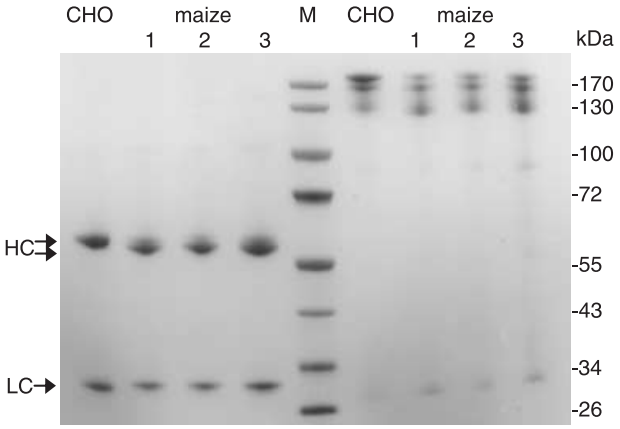

Figure 6 Sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) of purified 2G12 from Chinese hamster ovary (CHO) cells and maize seeds. Reduced (lanes 1–4) and non-reduced (lanes 6–9) samples (0.5 μg) were separated on 4%–12% precast gels and stained with Coomassie. Lanes 1 and 6, ^{CHO}2G12; lanes 2–4 and 7–9, three independent preparations of ^{Zm}2G12$_{SEKDEL}$.

flow-through nor wash fractions contained measurable amounts of 2G12, although free LC was detected in the flow-through, showing that LC was produced in excess (data not shown). Recovery was typically 60%–80% in this bench-scale process. Sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 6) showed that the preparations were very pure, containing negligible HC degradation products. As in the immunoblots mentioned above, HC was represented by two bands, that with the lowest molecular mass being more abundant. The two bands were excised together or separately from the gel, digested with trypsin and subjected to peptide mass fingerprinting. This analysis confirmed that both bands were 2G12 HC with correctly processed N-terminal ends, as indicated by the presence of peptide 'EVQLVESGGGLVK', resulting from the cleavage of the signal peptide. As the C-terminal end had also been shown to be intact via immunological detection of the KDEL sequence, differences in *N*-glycan structure were considered to be the most probable reason for the variation in molecular mass.

N-glycan structures attached to ^{Zm}2G12$_{SEKDEL}$ are extensively trimmed

The *N*-glycan structures of the two ^{Zm}2G12$_{SEKDEL}$ HC variants were analysed by mass spectrometry. The higher molecular mass variant was found to contain mainly OMT glycans, as expected for a protein with an ER retrieval signal (Figure 7a). In addition, small amounts of vacuolar-type *N*-glycans (MMXF and MUXF; Table 1) were identified, indicating that a small proportion of the antibody had escaped from the ER and passed through the Golgi. Approximately 20% of the more abundant lower molecular mass HC variant had no *N*-glycans at all, whereas the remainder contained unusually short glycans consisting of a single *N*-acetylglucosamine (GlcNAc) residue (Figure 7b). To determine the relative amounts of the different glycan structures in the total antibody preparation, the HC

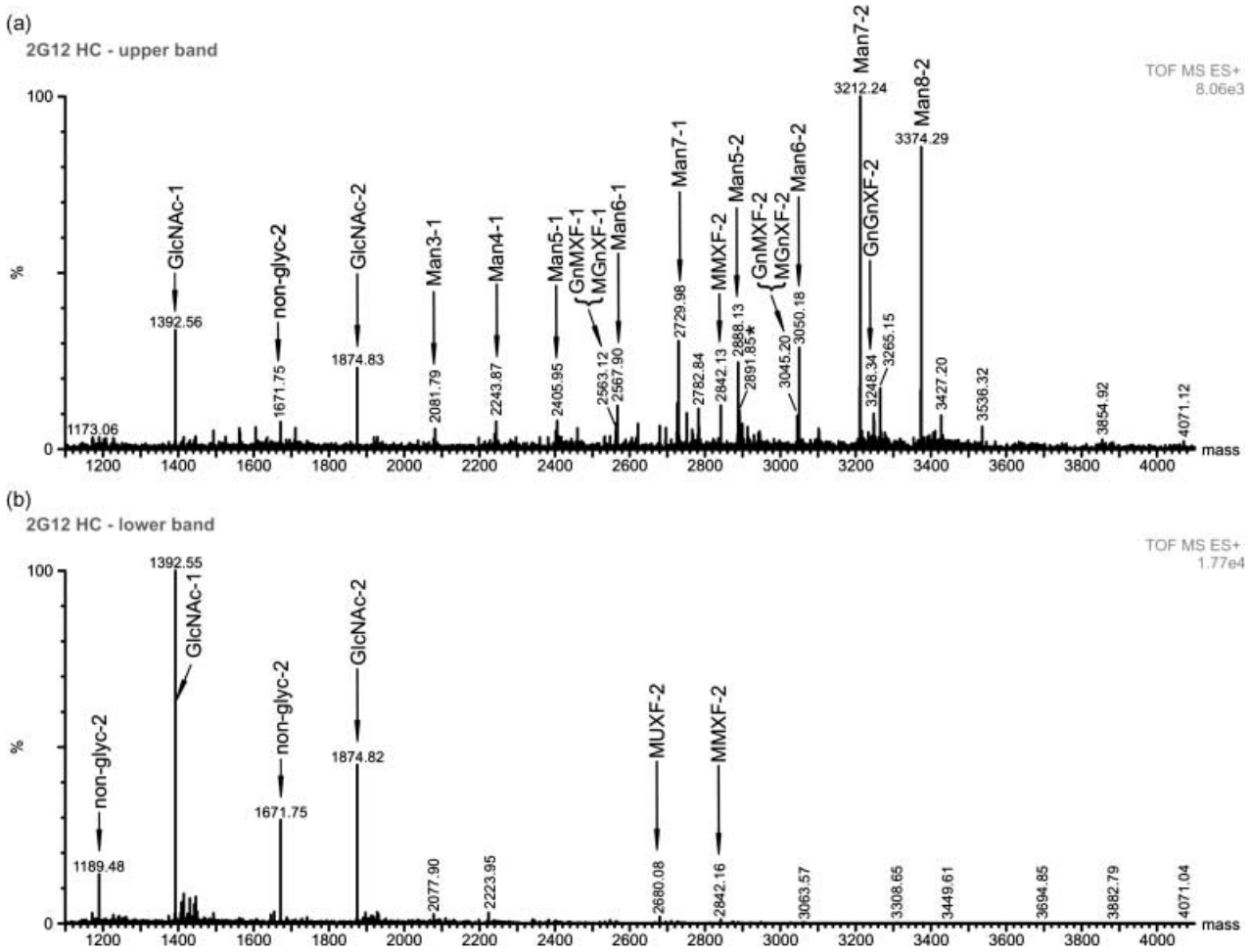

Figure 7 Deconvoluted liquid chromatography-mass spectra of glycosylated peptides (-1, EEQY*N*STYR; -2, TKPREEQY*N*STYR) containing asparagine-297 (Asn-297) from the 2G12 Fc region. (a) Glycopeptides derived from the 2G12 heavy chain (HC) upper band. (b) Glycopeptides derived from the 2G12 HC lower band. See http://www.proglycan.com for an explanation of *N*-glycan structure abbreviations.

bands were combined prior to analysis. The results, shown in Table 1, are representative of more than five experiments that were carried out with independent samples obtained from different crosses. Only small quantitative variations in the relative amounts of glycoforms were observed.

$^{Zm}2G12_{SEKDEL}$ has full *in vitro* antigen-binding activity

The antigen-binding activity of purified maize-derived $^{Zm}2G12_{SEKDEL}$ was quantified precisely and compared with the reference $^{CHO}2G12$ using a BIACORE 2000 instrument (Biacore, GE Healthcare, Uppsala, Sweden) with protein A and antigen (gp120) surfaces. Both surfaces were strongly mass transport limited, and therefore showed a linear dose–response and constant binding rates up to an antibody concentration of 1 μg/mL. Dilution series were measured in triplicate and showed low intra-assay variation. The antigen-binding activity was derived by plotting the gp120 signals directly against the protein A signals and performing a linear regression analysis (Figure 8). Using this approach, it is only important to ensure that the binding signals are in the linear range, as both the protein A binding signal (a measure of the total antibody concentration) and the gp120 binding signal (a measure of the active antibody concentration) are determined. The ratio of the binding signals, i.e. gp120/protein A, is equal to the slope derived by linear regression, and represents the antigen-binding activity of the antibody preparation. The results for $^{CHO}2G12$ and for two independent $^{Zm}2G12_{SEKDEL}$ preparations obtained from different crosses are illustrated in Figure 8. Both preparations show antigen-binding activities that are equivalent to the binding activity of $^{CHO}2G12$, with slopes of 0.1763 ± 0.0006 and 0.1737 ± 0.0004 compared with 0.1791 ± 0.0006. The maize antibodies therefore showed relative antigen-binding activities of $98.5\% \pm 0.34\%$ and

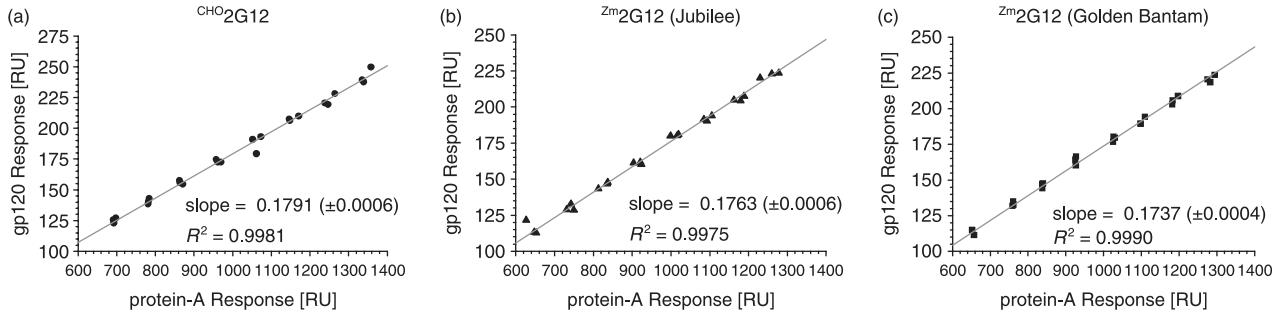

Figure 8 Signals obtained for 2G12 binding to protein A and gp120 surface. The gp120 response was plotted against the protein A response, and the slope representing the antigen-binding activity was determined by linear regression. The standard errors of the slopes are shown in parentheses. The reference ^{CHO}2G12 (a) and two independent purifications of ^{Zm}2G12$_{SEKDEL}$ from backcrosses into Jubilee (b) and Golden Bantam (c) are shown.

Table 1 Mass spectrometric analysis of the glycopeptides EEQYN[297]STYR and TKPREEQYN[297]STYR from tryptic digests of 2G12 purified from different maize hybrids

	Relative abundance (%)	
Glycoform	Starchcorn	Sweetcorn
Non-glycosylated	13.3	10.6
Single GlcNAc (Gn)	51.7	51.4
Man3 (M3Gn2)	0.9	1.2
Man4 (M4Gn2)	1.5	2.2
Man5 (M5Gn2)	2.7	3.5
Man6 (M6Gn2)	1.6	1.8
Man7 (M7Gn2)	9.3	13.7
Man8 (M8Gn2)	5.2	8.6
Man9 (M9Gn2)	0.4	1.3
Oligomannose-type (Σ)	21.6	32.3
MUXF (M2Gn2X1F1)	2.8	1.2
MMXF (M3Gn2X1F1)	6.0	3.1
GnMXF (M3Gn3X1F1)	2.2	0.7
GnGnXF (M3Gn4X1F1)	2.4	0.7
Complex type (Σ)	13.4	5.7

F, fucose; GlcNAc, *N*-acetylglucosamine (Gn); Man, mannose (M); X, xylose.

97% ± 0.24%, clearly demonstrating that ^{Zm}2G12$_{SEKDEL}$ is correctly folded and assembled in the maize endosperm, and that *in vitro* antigen binding is not affected by the presence of either the KDEL tag or the different *N*-glycans.

^{Zm}2G12$_{SEKDEL}$ neutralizes HIV-1

The total and active antibody concentrations in the purified samples were measured by enzyme-linked immunosorbent assay (ELISA), confirming that the mean relative activities of the plant-derived and CHO-derived antibodies were identical within experimental fluctuations. HIV-1 neutralization of three independent maize-derived 2G12 preparations was assessed in a syncytium inhibition assay in direct comparison with ^{CHO}2G12. The 50% inhibitory concentrations (IC_{50}) of ^{Zm}2G12$_{SEKDEL}$ were significantly lower than those of ^{CHO}2G12 in all three experiments (Table 2).

Table 2 Virus neutralization of maize-derived 2G12

Assay	IC_{50} ^{Zm}2G12$_{SEKDEL}$ (μg/mL)	IC_{50} ^{CHO}2G12 (μg/mL)	Ratio*
1	3.9	17.7	4.5
2	8.3	22.9	2.8
3	1.7	8.1	4.8

IC50, 50% inhibitory concentration.
*IC_{50} ^{CHO}2G12 divided by IC_{50} ^{Zm}2G12$_{SEKDEL}$.

Discussion

Full-size antibodies are currently produced in fermenters, using one of a small selection of approved mammalian cell lines (Wurm, 2004). This production platform is expensive, inflexible and lacks capacity, especially for the production of antibodies required in large amounts. With a view to producing such antibodies in plants, which can be grown inexpensively on a large scale, a well-characterized antibody against HIV was expressed in maize endosperm, and structural and functional analysis of the antibody in comparison with its counterpart produced in the established CHO system was carried out. Importantly, the antibody was targeted for retention in the ER so that plant-specific *N*-glycosylation could be avoided. The antibody was co-expressed with a fluorescent marker protein to facilitate the breeding programme and to preserve the identity of the pharmaceutical maize lines.

To restrict the recombinant protein to the endosperm, the rice *gt*-1 promoter was used, whose specificity has been

confirmed in several studies (Russell and Fromm, 1997; Ludwig *et al*., 2004; Law *et al*., 2006). The identification of plant lines strongly expressing a recombinant protein is more challenging when using a seed-specific promoter, as expression levels can be analysed only after seed set. The challenge is even greater when two different protein subunits, such as HC and LC of an antibody, need to be co-expressed at high levels. Because the collection and testing of seeds from numerous transgenic maize lines is a resource-heavy task, an approach was chosen that allowed the earliest possible identification of events in which both antibody genes had integrated. The HC gene was linked to a constitutively expressed *bar* gene and the LC gene to a constitutively expressed *DsRed* gene. Thus, growth under selection indicates the presence of the HC gene, and concomitant red fluorescence indicates the presence of the LC gene.

As transgenes integrate randomly into the plant genome, linkage to a fluorescent marker also allows the positive selection of events in which the transgene has integrated at a site allowing strong and stable expression (Stewart, 2001). DsRed was chosen because it can be excited with both green and blue light, making identification more reliable by reducing autofluorescence artefacts. In addition, green light is more user-friendly, penetrates plant tissue better than blue or UV light and causes less damage to the plant. During transformation and selection, DsRed is a useful marker to identify transgenic lines. Later in a breeding programme, however, DsRed also enables the identification of antibody-producing plants and helps to determine zygosity at the pollen stage. This provides a powerful tool to save time and resources in breeding tasks, such as crossing transgenic events into elite germplasm, which is necessary to improve agronomic performance and to establish parent lines for the production of hybrid maize. Breeding programmes are beneficial because the recombinant protein content of seeds can be increased over several plant generations by the continuous selection of individuals showing high transgene expression (Hood *et al*., 1997, 2003). A similar trend was observed over three generations in our maize line.

DsRed also provides a safety mechanism, as it can be used for the unambiguous and sensitive macroscopic detection of transgenic pollen and seed. Maize is beneficial for the production of pharmaceutical proteins, but is also widely cultivated for food and feed, and so additional safety measures must be put in place to avoid outcrossing and admixture. In this context, easily traceable visual markers greatly facilitate the monitoring process (a purpose for which DsRed is eminently suitable). DsRed may also serve as a built-in quality marker at different stages of production and downstream processing to rapidly detect the presence of non-transgenic material. Simple and portable equipment can be used to inspect production sites.

One of the most compelling advantages of seed-based production systems is that the seeds provide a protective environment for the recombinant protein, and allow prolonged storage at ambient temperatures without degradation. To take full advantage of the unique morphological storage characteristics of cereal endosperm cells, accumulation of the recombinant antibody in the main protein storage organelles is desirable. These contain several classes of zeins, which, like other prolamins, remain within the ER and aggregate into PBs by direct enlargement of the ER network (Lending and Larkins, 1989). Consequently, maize endosperm cells contain mainly ER-derived (pre-Golgi) prolamin bodies, and only a few globulin-like storage proteins that accumulate in post-Golgi PSVs (Woo *et al*., 2001). To achieve 2G12 deposition in the ER-derived PBs, a C-terminal ER retrieval signal was added to both antibody chains. As expected, $^{Zm}2G12_{SEKDEL}$ accumulated in the prolamin bodies, but not in the PSV-like structures. This supports our earlier observations in rice, where a KDEL-tagged single-chain antibody fragment accumulated predominantly in the ER-derived prolamin bodies, and only to a small extent in the PSVs, known in rice as glutelin bodies (Torres *et al*., 2001). Similarly, human serum albumin with an added KDEL signal was detected in prolamin aggregates in wheat (Arcalis *et al*., 2004). Although $^{Zm}2G12_{SEKDEL}$ co-localized with the zein aggregates, aqueous extraction with simple, non-reducing buffers was efficient, in agreement with earlier studies in rice (Nicholson *et al*., 2005) and wheat (Stoger *et al*., 2000).

In order to assess the suitability of maize endosperm as a production system for antibodies, it is important to compare the structural and functional properties of the maize-derived antibody with its counterpart produced in CHO cells. Structural differences between native and recombinant human proteins tend to reflect differing glycosylation patterns in plants and mammalian cells. However, plant-specific modifications take place in the Golgi, whereas the addition of a KDEL signal prevents passage through this compartment, leading to the formation of OMT glycans only. When HC was extracted from maize seeds, it separated into two variants with different molecular masses, but intact C- and N-termini. This suggested that the two variants differed in terms of their glycan structures. The variant with the higher molecular mass contained mainly OMT glycans, predominantly mannose-7 (Man7) and Man8. These glycans are typically found on ER-resident proteins, and their presence is consistent with the subcellular localization of $^{Zm}2G12_{SEKDEL}$ in the ER and in

ER-derived structures. In repeat measurements of independent antibody preparations, small amounts of complex glycan structures, such as MMXF, MUXF and GnMXF (see Table 1), were also identified, which are typical of vacuolar proteins, but their total amount rarely exceeded 10%. This agrees with previous reports showing that the *N*-glycans of KDEL-tagged antibodies produced in tobacco mainly contained OMT glycans, but also a small amount of complex glycans (Ko *et al*., 2003; Ramirez *et al*., 2003; Triguero *et al*., 2005). Other studies suggested that all the glycans present were of the OMT class, indicating more efficient ER retrieval (Pagny *et al*., 2000; Sriraman *et al*., 2004).

The accessibility of the KDEL sequence may explain differences in the retrieval efficiency for different KDEL-tagged antibodies (Sriraman *et al*., 2004). Alternatively, these differences could reflect detection sensitivity, or variations in the type or physiological state of the tissue used for antibody production. In line with this, Petruccelli *et al*. (2006) observed a difference in the efficiency of ER retrieval between leaves and seeds, and speculated that seed-specific factors may be responsible for this behaviour. In our study, fewer complex glycan structures were detected in sweetcorn hybrids, possibly indicating germplasm influences (Table 1).

Approximately 10% of HC was non-glycosylated, which is far lower than that reported recently for a KDEL-tagged single-chain Fv-Fc antibody produced in *Arabidopsis* seeds (Van Droogenbroeck *et al*., 2007). However, it falls within the normal range of glycosylation site occupancy observed with recombinant glycoproteins derived from mammalian cells, depending on the culture environment, in particular the availability of dolichol (Rosenwald *et al*., 1990; Crick and Waechter, 1994) and the ambient glucose concentration (Hayter *et al*., 1992, 1993; Tachibana *et al*., 1994). Most surprisingly, more than half of the *N*-glycans comprised only a single GlcNAc residue. A smaller amount (11.7%) of the same glycoform has been reported previously for an antibody produced in tobacco expressing a hybrid galactosyltransferase (Bakker *et al*., 2006). Only a very few endogenous plant proteins are known to carry monosaccharide *N*-glycans, including ribosome-inactivating proteins from pokeweed seeds (Islam *et al*., 1991; Zeng *et al*., 2003). The existence of single GlcNAc residues suggests that OMT glycans are processed by an endoglycanase (ENGase), consistent with our observation that OMT glycans remained on only 20%–30% of $^{Zm}2G12_{SEKDEL}$ HC. ENGase activity in cereal seeds has been reported (Chang *et al*., 2000; Vuylsteker *et al*., 2000), and free OMT glycan structures, which are released by ENGase activity, have indeed been identified in seeds of various plant species (Kimura and Kitahara, 2000; Kimura *et al*., 2002). A physiological role for these free OMT *N*-glycans in plant development, fruit and seed maturation has been proposed (Priem and Gross, 1992; Kimura and Kitahara, 2000).

As shown by immunolocalization, the bulk of $^{Zm}2G12_{SEKDEL}$ is located in PBs. There are two possible explanations for the accumulation of antibodies with trimmed *N*-glycans in PBs. Either the ENGase removes the OMT glycans in the ER before translocation into PBs, and/or the ENGase is active within PBs. Li and Larkins (1996) have shown that two forms of protein disulphide isomerase, a native ER-resident protein with a KDEL tag, exist in maize endosperm. One form has OMT glycans and resides in the endomembrane system, whereas the other lacks OMT glycans and resides in PBs.

It would be interesting to investigate the implications of such glycan modification on various *in vivo* antibody functions. It is probable that antibody effector functions and antibody-dependent cell-mediated cytotoxicity would be affected (Umana *et al*., 1999; Schuster *et al*., 2007), and this would need to be considered before systemic administration. However, this is not envisaged in our project, as $^{Zm}2G12_{SEKDEL}$ is intended for prophylactic mucosal application.

The quality of the antibody preparations was validated using a novel binding signal ratio assay based on surface plasmon resonance. The relative standard errors were very low (0.24%–0.33%) and demonstrated the excellent repeatability of the assay. The surface plasmon resonance assay is both precise and simple to perform and evaluate. The assay demonstrated clearly that the *in vitro* antigen-binding activity of the maize-derived antibody was nearly identical to that of its CHO counterpart, showing that neither the *N*-glycan truncations nor the presence of the C-terminal SEKDEL tag had any negative impact on folding, assembly or antibody function. In a direct comparison with the CHO-derived antibody, $^{Zm}2G12_{SEKDEL}$ demonstrated an excellent neutralization capacity, and the IC_{50} value was an average of four times lower. This might reflect the presence of dimeric forms in the plant-derived antibody samples, as these have a higher neutralization capacity, similar to polymeric forms of 2G12 (Wolbank *et al*., 2003).

In this investigation, maize was evaluated as a potential large-scale production system for therapeutic antibodies against HIV. HIV-neutralizing antibodies are promising candidates for the development of multicomponent topical microbicides, and are thus attracting increasing attention as a prevention strategy for HIV. However, such antibodies would need to be administered frequently in high doses, placing an unusual demand on current production technologies (Trkola *et al*., 2005). Our data show that maize can be used to produce functional 2G12, and, despite differences in glycan structure,

the performance of the antibody matches, if not exceeds, that of an identical antibody expressed in mammalian cells. The maize endosperm system provides the capacity for large-scale production, and directing the recombinant protein to accumulate in ER-derived PBs minimizes the addition of plant-specific glycans, therefore reducing any immunoreactive potential from the recombinant proteins. The co-expression of a visible marker was also described, which allows the rapid identification of promising transgenic events and provides traceability and identity preservation for pharmaceutical maize crops. Taken together, the advantages of maize in terms of high yield and storage capability, the use of ER retrieval to control glycan structure and the use of dual visual selection markers to facilitate the identification and breeding of superior transgenic lines expressing two antibody chains provide an excellent system for the large-scale production of therapeutic antibodies, for which demand will almost certainly outstrip the capacity of current production platforms.

Experimental procedures

Plant expression vectors and maize transformation

All cloning steps were carried out with the binary vector pTRA, a derivative of pPAM (GenBank: AY027531) (Sack *et al*., 2007). The *bar-2G12HC*$_{SEKDEL}$ vector (pTRAb-gGHER) and the *DsRed-2G12LC*$_{SEKDEL}$ vector (pTRAds-gGLER), shown in Figure 1, were constructed as follows. The genes for the two 2G12 antibody chains were obtained from Polymun (Vienna, Austria). The codons for the SEKDEL tag were fused to both antibody chain genes by polymerase chain reaction (PCR), and the coding regions, including the signal peptides, were subcloned between the tobacco etch virus (TEV) 5′-untransformed region (5′-UTR) and the cauliflower mosaic virus (CaMV) 35S terminator in a *gt-1* promoter expression cassette. The *2G12HC* cassette was then inserted downstream of the ubiquitin-1 promoter *bar* cassette. The *2G12LC* cassette was cloned downstream of the ubiquitin-1 promoter *DsRed* cassette containing a plastid transit peptide sequence. All expression cassettes contained the maize ubiquitin-1 first intron. The expression cassettes were oriented head to tail and flanked by tobacco RB7 scaffold attachment regions in order to avoid interference of surrounding sequences. The plasmids for HC and LC were coated on to gold particles in an equimolar ratio, and transgenic plants were regenerated from bombarded embryogenic Hill (A188 × B73) callus, as described previously (Frame *et al*., 2000).

Plant cultivation

Maize plants were grown in a glasshouse or a phytotron at 28/25 °C day/night temperature with a 14-h photoperiod and 50%–70% relative humidity. Primary Hill transformants were selfed or pollinated with African elite cultivars SSG62B, NSP5120A1-2 or K64r. Jubilee and Golden Bantam cultivars were used for breeding into sweetcorn. T_1 and F_1 plants were either self-pollinated or further backcrossed to the elite lines.

Fluorescence detection

For microscopic analysis, standard epifluorescence and dissecting microscopes were used with appropriate filter sets (Leica, Wetzlar, Germany; Olympus, Tokyo, Japan). To visualize DsRed fluorescence macroscopically, excitation was conducted with a cold light source with 1 m fiber optics and a green excitation filter (Schott, Mainz, Germany). Red fluorescence was observed and photographed through a simple red filter (LEE filters, Andover, UK).

Antibody extraction and purification

For single seed analysis, seeds were soaked overnight in tap water and separated into endosperm and embryo. The endosperm portions were dried and individually ground with a mortar and pestle. Proteins were extracted with three volumes (v/w) of buffer [phosphate-buffered saline (PBS) pH 7.4, 5 mM ethylenediaminetetraacetic acid (EDTA), 5 mM β-mercaptoethanol] for at least 2 h at 22 °C on a shaker, followed by incubation overnight at 4 °C. After additional mixing, the suspension was centrifuged and the supernatant was used for analysis. The embryos were rescued on cotton pads and those from seeds with high antibody levels were retained.

For antibody purification, pooled dry seeds were ground twice for 30 s in a Warring blender and extracted in five volumes (v/w) of buffer (PBS pH 7.4, 5 mM EDTA, 1 mM β-mercaptoethanol). After stirring overnight at 4 °C, the antibody was purified from the clarified supernatant by protein A chromatography as described previously (Sack *et al*., 2007), with the exception that 100 mM glycine pH 3.6 with 100 mM fructose was used as the elution buffer. The dialysed antibody was concentrated by ultrafiltration (molecular weight cut-off, 30 kDa) and the clear supernatant was stored at 4 °C.

SDS-PAGE, immunoblot analysis and ELISAs were performed as described previously (Sack *et al*., 2007).

Surface plasmon resonance analysis

Antibody quantification and antigen-binding assays were performed using a BIACORE 2000 instrument (Biacore, GE Healthcare). Recombinant protein A (Sigma-Aldrich, St. Louis, MO, USA; 100 μg/mL in 10 mM sodium acetate pH 4.5) and gp120 [Centre for AIDS Reagents (CFAR/NIBSC, Potters Bar, Hertfordshire, UK), USA; EVA 607, 25 μg/mL in 10 mM sodium acetate pH 4.75] were coupled to a CM5-rg sensorchip following the standard 1-ethyl-3-diaminopropyl-carbodiimide/*N*-hydroxylsuccinimide (EDC/NHS) protocol. An activated/deactivated surface was used as the reference for blank subtraction. About 4 kRU (RU = resonance units) of protein A and 12 kRU of gp120 were coupled, resulting in surfaces with a high binding capacity and high mass transport limitation. Regeneration was achieved with a 30-s pulse of 0.5 M citrate pH 3.0 for gp120 and 30 mM HCl for protein A. The surfaces were stable for several hundred cycles. All measurements were performed at 25 °C at a flow rate of 30 μL/min using HBS-EP as the running buffer (10 mM HEPES, pH 7.4, 150 M NaCl, 3 mM EDTA, 0.005% polysorbate 20). Samples were diluted such that binding signals were in the linear range. Antibody concentrations were calculated using CHO-derived 2G12 as standard.

The antigen-binding activity was determined by linear regression of the gp120 responses plotted against the protein A responses. The relative antigen-binding activity was derived by dividing the slope

obtained for the samples by the slope obtained for the ^{CHO}2G12 reference. Data evaluation was performed using BIAevaluation version 4.0 (GE Healthcare, Uppsala, Sweden) and Microcal Origin version 5.0 (OriginLab, Northampton, MA, USA).

Liquid chromatography-mass spectrometry (LC-MS) analysis

Coomassie-stained bands representing HC were excised, destained, carbamidomethylated, digested with trypsin and extracted from gel pieces, as described previously (Kolarich and Altmann, 2000; Kolarich *et al*., 2006). The subsequent fractionation of the peptides by capillary reversed-phase chromatography, with detection by a quadrupole time-of-flight (Q-TOF) Ultima Global (Waters Micromass, Manchester, UK) mass spectrometer, was performed as described previously (Kolarich and Altmann, 2000; Van Droogenbroeck *et al*., 2007).

The MS data from the tryptic peptides were compared with data sets generated by *in silico* tryptic digestion of the 2G12 coding using the PeptideMass program (http://www.expasy.org/tools/peptide-mass.html). Based on the tryptic peptide data sets, tryptic glycopeptide data sets were generated by the addition of glycan mass increments to the masses of the two identified glycopeptides.

HIV neutralization assay

HIV-1 neutralization of 2G12 was assessed using a syncytium inhibition assay. Ten twofold serial dilutions (starting concentration, 100 μg/mL) of ^{Zm}2G12$_{SEKDEL}$, ^{CHO}2G12 and a non-neutralizing control were pre-incubated with HIV-1 strain RF at 10^2–10^3 50% tissue culture infective dose ($TCID_{50}$)/mL for 1 h at 37 °C. CD4-positive human AA-2 cells were added at a density of 4×10^5 cells/mL and further incubated for 5 days. Experiments were performed with eight replicates per antibody dilution step. The presence of one or more syncytia per well after 5 days was scored as positive infection. The IC_{50} values were calculated by the method of Reed and Muench (1938) using the concentrations present during the antibody virus pre-incubation step.

Immunofluorescence and electron microscopy

Developing maize grains (25 days after fertilization) were bisected longitudinally and the embryo was removed. One half of the grain was processed for recombinant protein analysis and the other half was fixed and processed for microscopy as described previously (Arcalis *et al*., 2004). Semi-thin sections were mounted on glass slides and stained with toluidine blue for the identification of the cell layers in the endosperm. Sections mounted on glass slides for fluorescence microscopy and on gold grids for electron microscopy were treated as described previously (Arcalis *et al*., 2004), and incubated with polyclonal antiserum against human κ-chain. Sections were then treated with secondary antibody labelled with Alexa Fluor 594 for fluorescence microscopy, or with 10-nm gold particles for electron microscopy. Following immunolocalization, counterstaining of ultra-thin sections was performed in 2% (w/v) aqueous uranyl acetate. The sections were observed using a Philips EM-400 transmission electron microscope (Philips, Eindhoven, The Netherlands).

Acknowledgements

This work was supported by the EU project Pharma-Planta. We thank Dr R. M. Twyman for help with the preparation of the manuscript, and Drs T. R. Rocheford and C. Paul for providing Hill ears as donor material.

References

Arcalis, E., Marcel, S., Altmann, F., Kolarich, D., Drakakaki, G., Fischer, R., Christou, P. and Stoger, E. (2004) Unexpected deposition patterns of recombinant proteins in post-endoplasmic comprised reticulum compartments of wheat endosperm. *Plant Physiol*. **136**, 3457–3466.

Bakker, H., Rouwendal, G.J., Karnoup, A.S., Florack, D.E., Stoopen, G.M., Helsper, J.P., van Ree, R., van Die, I. and Bosch, D. (2006) An antibody produced in tobacco expressing a hybrid beta-1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc. Natl. Acad. Sci. USA*, **103**, 7577–7582.

Chang, T., Kuo, M.C., Khoo, K.H., Inoue, S. and Inoue, Y. (2000) Developmentally regulated expression of a peptide: N-glycanase during germination of rice seeds (*Oryza sativa*) and its purification and characterization. *J. Biol. Chem*. **275**, 129–134.

Crick, D.C. and Waechter, C.J. (1994) Long-chain cis-isoprenyltransferase activity is induced early in the developmental program for protein N-glycosylation in embryonic rat brain cells. *J. Neurochem*. **62**, 247–256.

Daniell, H., Streatfield, S.J. and Wycoff, K. (2001) Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends Plant Sci*. **6**, 219–226.

Ferrantelli, F., Rasmussen, R.A., Buckley, K.A., Li, P.L., Wang, T., Montefiori, D.C., Katinger, H., Stiegler, G., Anderson, D.C., McClure, H.M. and Ruprecht, R.M. (2004) Complete protection of neonatal rhesus macaques against oral exposure to pathogenic simian-human immunodeficiency virus by human anti-HIV monoclonal antibodies. *J. Infect. Dis*. **21**, 820–825.

Fotisch, K. and Vieths, S. (2001) N- and O-linked oligosaccharides of allergenic glycoproteins. *Glycoconj. J*. **18**, 373–390.

Frame, B.R., Zhang, H., Cocciolone, S.M., Sidorenko, L.V., Dietrich, C.R., Pegg, S.E., Zhen, S., Schnable, P.S. and Wang, K. (2000) Production of transgenic maize from bombarded type II callus: effect of gold particle size and callus morphology on transformation efficiency. *In Vitro Cellular Dev. Biol. Plant*, **36**, 21.

Gomord, V., Chamberlain, P., Jefferis, R. and Faye, L. (2005) Biopharmaceutical production in plants: problems, solutions and opportunities. *Trends Biotechnol*. **23**, 559–565.

Harper, B.K., Mabon, S.A., Leffel, S.M., Halfhill, M.D., Richards, H.A., Moyer, K.A. and Stewart, C.N. Jr (1999) Green fluorescent protein as a marker for expression of a second gene in transgenic plants. *Nat. Biotechnol*. **17**, 1125–1129.

Hayter, P.M., Curling, E.M.A., Baines, A.J., Jenkins, N., Salmon, I., Strange, P.G., Tong, J.M. and Bull., A.T. (1992) Glucose-limited chemostat culture of Chinese hamster ovary cells producing recombinant human interferon-γ. *Biotechnol. Bioeng*. **39**, 327–335.

Hayter, P.M., Curling, E.M.A., Gould, M.L., Baines, A.J., Jenkins, N., Salmon, I., Strange, P.G. and Bull., A.T. (1993) The effect of the dilution rate on CHO cell physiology and recombinant interferon-γ

production in glucose-limited chemostat culture. *Biotechnol. Bioeng*. **42**, 1077–1085.

Hofmann-Lehmann, R., Vlasak, J., Rasmussen, R.A., Smith, B.A., Baba, T.W., Liska, V., Ferrantelli, F., Montefiori, D.C., McClure, H.M., Anderson, D.C., Bernacky, B.J., Rizvi, T.A., Schmidt, R., Hill, L.R., Keeling, M.E., Katinger, H., Stiegler, G., Cavacini, L.A., Posner, M.R., Chou, T.C., Andersen, J. and Ruprecht, R.M. (2001) Postnatal passive immunization of neonatal macaques with a triple combination of human monoclonal antibodies against oral simian-human immunodeficiency virus challenge. *J. Virol*. **75**, 7470–7480.

Hood, E.E., Witcher, D.R., Maddock, S., Meyer, T., Baszczynski, C., Bailey, M., Flynn, P., Register, J., Marshall, L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehigh, R.J., Hernan, R., Kappel, W.K., Ritland, D., Li, C.P. and Howard, J.A. (1997) Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Mol. Breed*. **3**, 291.

Hood, E.E., Bailey, M.R., Beifuss, K., Magallanes-Lundback, M., Horn, M.E., Callaway, E., Drees, C., Delaney, D.E., Clough, R. and Howard, J.A. (2003) Criteria for high-level expression of a fungal laccase gene in transgenic maize. *Plant Biotechnol. J*. **1**, 129–140.

Islam, M.R., Kung, S.S., Kimura, Y. and Funatsu, G. (1991) N-acetyl-D-glucosamine-asparagine structure in ribosome-inactivating proteins from the seeds of *Luffa cylindrica* and *Phytolacca americana*. *Agric. Biol. Chem*. **55**, 1375–1381.

Kimura, Y. and Kitahara, E. (2000) Structural analysis of free N-glycans occurring in soybean seedlings and dry seeds. *Biosci. Biotechnol. Biochem*. **64**, 1847–1855.

Kimura, Y., Matsuo, S., Tsurusaki, S., Kimura, M., Hara-Nishimura, I. and Nishimura, M. (2002) Subcellular localization of endo-beta-N-acetylglucosaminidase and high-mannose type free N-glycans in plant cell. *Biochim. Biophys. Acta*, **1570**, 38–46.

Ko, K., Tekoah, Y., Rudd, P.M., Harvey, D.J., Dwek, R.A., Spitsin, S., Hanlon, C.A., Rupprecht, C., Dietzschold, B., Golovkin, M. and Koprowski, H. (2003) Function and glycosylation of plant-derived antiviral monoclonal antibody. *Proc. Natl. Acad. Sci. USA*, **100**, 8013–8018.

Kolarich, D. and Altmann, F. (2000) N-Glycan analysis by matrix-assisted laser desorption/ionization mass spectrometry of electrophoretically separated nonmammalian proteins: application to peanut allergen Ara h 1 and olive pollen allergen Ole e 1. *Anal. Biochem*. **285**, 64–75.

Kolarich, D., Altmann, F. and Sunderasan, E. (2006) Structural analysis of the glycoprotein allergen Hev b 4 from natural rubber latex by mass spectrometry. *Biochim. Biophys. Acta*, **1760**, 715–720.

Law, R.D., Russell, D.A., Thompson, L.C., Schroeder, S.C., Middle, C.M., Tremaine, M.T., Jury, T.P., Delannay, X. and Slater, S.C. (2006) Biochemical limitations to high-level expression of humanized monoclonal antibodies in transgenic maize seed endosperm. *Biochim. Biophys. Acta*, **1760**, 1434–1444.

Lending, C.R. and Larkins, B.A. (1989) Changes in the zein composition of protein bodies during maize endosperm development. *Plant Cell*, **1**, 1011–1023.

Li, C.P. and Larkins, B.A. (1996) Expression of protein disulfide isomerase is elevated in the endosperm of the maize floury-2 mutant. *Plant Mol. Biol*. **30**, 873–882.

Ludwig, D.L., Witte, L., Hicklin, D.J., Prewett, M., Bassi, R., Burtrum, D., Pereira, D.S., Jimenez, X., Fox, F., Saxena, B., Zhou, Q., Ma, Y., Kang, X., Patel, D., Barry, M., Kussie, P., Zhu, Z., Russell, D.A., Petersen, W.L., Jury, T.P., Gaitan-Gaitan, F., Moran, D.L., Delannay, X., Storrs, B.S., Tou, J., Zupec, M.E., Gustafson, K.S., McIntyre, J., Tarnowski, S.J. and Bohlen, P. (2004) Conservation of receptor antagonist anti-tumor activity by epidermal growth factor receptor antibody expressed in transgenic corn seed. *Hum. Antibodies*, **13**, 81–90.

Ma, J.K., Drake, P.M., Chargelegue, D., Obregon, P. and Prada, A. (2005) Antibody processing and engineering in plants, and new strategies for vaccine production. *Vaccine*, **23**, 1814–1818.

Mascola, J.R. (2002) Passive transfer studies to elucidate the role of antibody-mediated protection against HIV-1. *Vaccine*, **20**, 1922–1925.

Mascola, J.R., Lewis, M.G., Stiegler, G., Harris, D., VanCott, T.C., Hayes, D., Louder, M.K., Brown, C.R., Sapan, C.V., Frankel, S.S., Lu, Y., Robb, M.L., Katinger, H. and Birx, D.L. (1999) Protection of Macaques against pathogenic simian/human immunodeficiency virus 89.6PD by passive transfer of neutralizing antibodies. *J. Virol*. **73**, 4009–4018.

Mascola, J.R., Stiegler, G., VanCott, T.C., Katinger, H., Carpenter, C.B., Hanson, C.E., Beary, H., Hayes, D., Frankel, S.S., Birx, D.L. and Lewis, M.G. (2000) Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nat. Med*. **6**, 207–210.

Nicholson, L., Gonzalez-Melendi, P., van Dolleweerd, C., Tuck, H., Perrin, Y., Ma, J.K.-C., Fischer, R., Christou, P. and Stoger, E. (2005) A recombinant multimeric immunoglobulin expressed in rice shows assembly-dependent subcellular localization in endosperm cells. *Plant Biotechnol. J*. **3**, 115–127.

Pagny, S., Cabanes-Macheteau, M., Gillikin, J.W., Leborgne-Castel, N., Lerouge, P., Boston, R.S., Faye, L. and Gomord, V. (2000) Protein recycling from the Golgi apparatus to the endoplasmic reticulum in plants and its minor contribution to calreticulin retention. *Plant Cell*, **12**, 739–756.

Petruccelli, S., Otegui, M.S., Lareu, F., Tran Dinh, O., Fitchette, A.-C., Circosta, A., Rumbo, M., Bardor, M., Carcamo, R., Gomord, V. and Beachy, R.N. (2006) A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. *Plant Biotechnol. J*. **4**, 511–527.

Priem, B. and Gross, K.C. (1992) Mannosyl- and xylosyl-containing glycans promote tomato (*Lycopersicon esculentum* Mill.) fruit ripening. *Plant Physiol*. **98**, 399–401.

Ramirez, N., Rodriguez, M., Ayala, M., Cremata, J., Perez, M., Martinez, A., Linares, M., Hevia, Y., Paez, R., Valdes, R., Gavilondo, J.V. and Selman-Housein, G. (2003) Expression and characterization of an anti-(hepatitis B surface antigen) glycosylated mouse antibody in transgenic tobacco (*Nicotiana tabacum*) plants and its use in the immunopurification of its target antigen. *Biotechnol. Appl. Biochem*. **38**, 223–230.

van Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J.P., Loutelier-Bourhis, C., Rayon, C., Villalba, M., Koppelman, S., Aalberse, R., Rodriguez, R., Faye, L. and Lerouge, P. (2000) Beta (1,2)-xylose and alpha (1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J. Biol. Chem*. **275**, 11 451–11 458.

Reed, L.J. and Muench, H.A. (1938) Simple method of estimating fifty per cent endpoints. *Am. J. Hyg*. **27**, 493–497.

Rosenwald, A.G., Stoll, J. and Krag, S.S. (1990) Regulation of glycosylation. Three enzymes compete for a common pool of dolichyl phosphate in vivo. *J. Biol. Chem*. **265**, 14 544–14 553.

Russell, D.A. and Fromm, M.E. (1997) Tissue-specific expression in transgenic maize of four endosperm promoters from maize and rice. *Transgenic Res*. **6**, 157–168.

Sack, M., Paetz, A., Kunert, R., Bomble, M., Hesse, F., Stiegler, G., Fischer, R., Katinger, H., Stoeger, E. and Rademacher, T. (2007) Functional analysis of the broadly neutralizing human anti-HIV-1 antibody 2F5 produced in transgenic BY-2 suspension cultures. *Faseb J*. **21**, 1655–1664.

Schillberg, S., Twyman, R.M. and Fischer, R. (2005) Opportunities for recombinant antigen and antibody expression in transgenic plants – technology assessment. *Vaccine*, **23**, 1764–1769.

Schouten, A., Roosien, J., van Engelen, F.A., de Jong, G.A., Borst-Vrenssen, A.W., Zilverentant, J.F., Bosch, D., Stiekema, W.J., Gommers, F.J., Schots, A. and Bakker, J. (1996) The C-terminal KDEL sequence increases the expression level of a single-chain antibody designed to be targeted to both the cytosol and the secretory pathway in transgenic tobacco. *Plant Mol. Biol*. **30**, 781–793.

Schuster, M., Jost, W., Mudde, G.C., Wiederkum, S., Schwager, C., Janzek, E., Altmann, F., Stadlmann, J., Stemmer, C. and Gorr, G. (2007) In vivo glyco-engineered antibody with improved lytic potential produced by an innovative non-mammalian expression system. *Biotechnol. J*. **2**, 700–708.

Sriraman, R., Bardor, M., Sack, M., Vaquero, C., Faye, L., Fischer, R., Finnern, R. and Lerouge, P. (2004) Recombinant anti-hCG antibodies retained in the endoplasmic reticulum of transformed plants lack core-xylose and core (1,3)-fucose residues. *Plant Biotechnol. J*. **2**, 279–287.

Stewart, C.N. Jr (2001) The utility of green fluorescent protein in transgenic plants. *Plant Cell Rep*. **20**, 376–382.

Stewart, C.N. Jr (2005) Monitoring the presence and expression of transgenes in living plants. *Trends Plant Sci*. **10**, 390–396.

Stoger, E., Vaquero, C., Torres, E., Sack, M., Nicholson, L., Drossard, J., Williams, S., Keen, D., Perrin, Y., Christou, P. and Fischer, R. (2000) Cereal crops as viable production and storage systems for pharmaceutical scFv antibodies. *Plant Mol. Biol*. **42**, 583–590.

Tachibana, H., Taniguchi, K., Ushio, Y., Teruya, K., Osada, K. and Murakami, H. (1994) Changes of monosaccharide availability of human hybridoma lead to alteration of biological properties of human monoclonal antibody. *Cytotechnology*, **16**, 151–157.

Torres, E., Gonzalez-Melendi, P., Stoger, E., Shaw, P., Twyman, R.M., Nicholson, L., Vaquero, C., Fischer, R., Christou, P. and Perrin, Y. (2001) Native and artificial reticuloplasmins co-accumulate in distinct domains of the endoplasmic reticulum and in post-endoplasmic reticulum compartments. *Plant Physiol*. **127**, 1212–1223.

Triguero, A., Cabrera, G., Cremata, J.A., Yuen, C.-T., Wheeler, J. and Ramirez, N.I. (2005) Plant-derived mouse IgG monoclonal antibody fused to KDEL endoplasmic reticulum-retention signal is *N*-glycosylated homogeneously throughout the plant with mostly high-mannose-type *N*-glycans. *Plant Biotechnol. J*. **3**, 449–457.

Trkola, A., Kuster, H., Rusert, P., Joos, B., Fischer, M., Leemann, C., Manrique, A., Huber, M., Rehr, M., Oxenius, A., Weber, R., Stiegler, G., Vcelar, B., Katinger, H., Aceto, L. and Gunthard, H.F. (2005) Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nat. Med*. **11**, 615–622.

Umana, P., Jean-Mairet, J., Moudry, R., Amstutz, H. and Bailey, J.E. (1999) Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nat. Biotechnol*. **17**, 176–180.

Van Droogenbroeck, B., Cao, J., Stadlmann, J., Altmann, F., Colanesi, S., Hillmer, S., Robinson, D.G., Van Lerberge, E., Terryn, N., Van Montagu, M., Liang, M., Depicker, A. and De Jaeger, G. (2007) Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic Arabidopsis seeds. *Proc. Natl. Acad. Sci. USA*, **104**, 1430–1435.

Veazey, R.S., Shattock, R.J., Pope, M., Kirijan, J.C., Jones, J., Hu, Q., Ketas, T., Marx, P.A., Klasse, P.J., Burton, D.R. and Moore, J.P. (2003) Prevention of virus transmission to macaque monkeys by a vaginally applied monoclonal antibody to HIV-1 gp120. *Nat. Med*. **9**, 343–346.

Vuylsteker, C., Cuvellier, G., Berger, S., Faugeron, C. and Karamanos, Y. (2000) Evidence of two enzymes performing the de-N-glycosylation of proteins in barley: expression during germination, localization within the grain and set-up during grain formation. *J. Exp. Bot*. **51**, 839–845.

Wolbank, S., Kunert, R., Stiegler, G. and Katinger, H. (2003) Characterization of human class-switched polymeric (immunoglobulin M [IgM] and IgA) anti-human immunodeficiency virus type 1 antibodies 2F5 and 2G12. *J. Virol*. **77**, 4095–4103.

Woo, Y.M., Hu, D.W., Larkins, B.A. and Jung, R. (2001) Genomics analysis of genes expressed in maize endosperm identifies novel seed proteins and clarifies patterns of zein gene expression. *Plant Cell*, **13**, 2297–2317.

Wurm, F.M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol*. **22**, 1393–1398.

Zeng, Z.H., He, X.L., Li, H.M., Hu, Z. and Wang, D.C. (2003) Crystal structure of pokeweed antiviral protein with well-defined sugars from seeds at 1.8 Å resolution. *J. Struct. Biol*. **141**, 171–178.

Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic Arabidopsis seeds.

Van Droogenbroeck, B; Cao, J; Stadlmann, J; Altmann, F; Colanesi, S; Hillmer, S; Robinson, DG; Van Lerberge, E; Terryn, N; Van Montagu, M; Liang, M; Depicker, A and De Jaeger, G (2007). Proc Natl Acad Sci U S A 104(4): 1430-5.

Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic *Arabidopsis* seeds

Bart Van Droogenbroeck*, Jingyuan Cao†, Johannes Stadlmann‡, Friedrich Altmann‡, Sarah Colanesi§, Stefan Hillmer§, David G. Robinson§, Els Van Lerberge*, Nancy Terryn¶, Marc Van Montagu¶‖, Mifang Liang†‖, Ann Depicker*‖, and Geert De Jaeger*

*Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology, Ghent University, Technologiepark 927, B-9052 Ghent, Belgium; †State Key Laboratory for Infectious Disease Control and Prevention, National Institute for Viral Disease Control and Prevention, 100052 Beijing, China; ‡Department of Chemistry, University of Natural Resources and Applied Life Sciences, A-1190 Vienna, Austria; §Heidelberg Institute for Plant Sciences, University of Heidelberg, D-69120 Heidelberg, Germany; and ¶Institute for Plant Biotechnology for Developing Countries, Ghent University, B-9000 Ghent, Belgium

Contributed by Marc Van Montagu, November 27, 2006 (sent for review September 21, 2006)

Production of high-value recombinant proteins in transgenic seeds is an attractive and economically feasible alternative to conventional systems based on mammalian cells and bacteria. In contrast to leaves, seeds allow high-level accumulation of recombinant proteins in a relatively small volume and a stable environment. We demonstrate that single-chain variable fragment (scFv)-Fc antibodies, with N-terminal signal sequence and C-terminal KDEL tag, can accumulate to very high levels as bivalent IgG-like antibodies in *Arabidopsis thaliana* seeds and illustrate that a plant-produced anti-hepatitis A virus scFv-Fc has similar antigen-binding and *in vitro* neutralizing activities as the corresponding full-length IgG. As expected, most scFv-Fc produced in seeds contained only oligomannose-type *N*-glycans, but, unexpectedly, 35–40% was never glycosylated. A portion of the scFv-Fc was found in endoplasmic reticulum (ER)-derived compartments delimited by ribosome-associated membranes. Additionally, consistent with the glycosylation data, large amounts of the recombinant protein were deposited in the periplasmic space, implying a direct transport from the ER to the periplasmic space between the plasma membrane and the cell wall. Aberrant localization of the ER chaperones calreticulin and binding protein (BiP) and the endogenous seed storage protein cruciferin in the periplasmic space suggests that overproduction of recombinant scFv-Fc disturbs normal ER retention and protein-sorting mechanisms in the secretory pathway.

glycosylation | molecular farming | recombinant antibody | subcellular localization

Transgenic plants for the production of high-value recombinant proteins are a promising alternative to conventional recombinant protein production systems, such as bacteria, yeast, animal, and insect cell cultures (1). One of the most important factors driving research in this field is yield improvement, because of its significant impact on economic feasibility (2). Strategies to increase recombinant protein yield in plants include development of better expression cassettes, improvement of protein stability and accumulation by using specific subcellular targeting signals, and development of downstream processing technologies (3). In this perspective, seed-based platforms are particularly attractive because they allow recombinant proteins to stably accumulate at a relatively high concentration in a compact biomass, which is beneficial for extraction and downstream processing (4). By using a seed-specific expression cassette based on the regulatory signals of seed storage proteins of common bean (*Phaseolus vulgaris*), and by targeting the recombinant protein to the endoplasmic reticulum (ER), we obtained the highest yields of recombinant proteins in plants described so far: a single-chain variable fragment (scFv) accumulated to levels in excess of 36% of total soluble protein (TSP) in homozygous *Arabidopsis* seeds, while retaining its antigen-binding activity and affinity (5).

For some applications, fusion of the scFv with the Fc domain is useful to restore antigen-binding avidity and antibody effector functions and to reach a prolonged serum half-life (6–8). In addition, scFv-Fc antibodies benefit from the Fc domain as a convenient affinity handle for purification and immunochemistry, eliminating the need for a proteolytically sensitive epitope and/or affinity tag (8). Moreover, a single cloning step between phagemid and scFv-Fc format allows high-throughput screening and accelerated validation of phage lead scFv candidates in a format mimicking the bivalent properties of IgG molecules (9, 10). Given these characteristics, together with the therapeutic potential of this recombinant antibody format (11, 12), the power of the seed-specific expression cassette was evaluated for the more complex scFv-Fc antibody format in *Arabidopsis*. The fusions are similar to those made in *Pichia pastoris*, mammalian, and insect cells (8, 12, 13) but instead expressed in plants. Because of the importance of subcellular localization (14) and N-glycosylation (15) on the stability, correct folding, and biological activity of recombinant antibodies, the N-terminal sequence, the *N*-glycan structures, and the intracellular localization of scFv-Fc antibodies expressed in *Arabidopsis* seeds were determined. Finally, we demonstrate the functionality of a seed-produced anti-hepatitis A scFv-Fc.

Results

scFv-Fc Production in *Arabidopsis* Seeds. Four different scFv-Fc fusion proteins (see *Materials and Methods*) were cloned in the binary vector pPhas (Fig. 1*A*). Transgenic *Arabidopsis* plants were obtained by *Agrobacterium*-mediated floral dip transformation. For three of the four scFv-Fc constructs, i.e., MBP10, EHF34, and HA78, high-level accumulation of the recombinant protein was detected by SDS/PAGE in the protein extracts of T2-segregating seed stocks (Fig. 1*B* and Table 1). Levels of

Author contributions: B.V.D. and J.C. contributed equally to this work. B.V.D., N.T., M.V.M., M.L., A.D., and G.D.J. designed research; B.V.D., J.C., and E.V.L. performed research; J.S., F.A., S.C., S.H., and D.G.R. contributed new reagents/analytic tools; B.V.D., M.V.M., A.D., and G.D.J. analyzed data; and B.V.D., A.D., and G.D.J. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: scFv, single-chain variable fragment; ER, endoplasmic reticulum; TSP, total soluble protein; PS, periplasmic space; PSV, protein storage vacuole; BiP, binding protein; ProtA, protein A; HAV, hepatitis A virus; HTNV, Hantaan virus; $TCID_{50}$, tissue culture 50% infective dose.

‖To whom correspondence may be addressed. E-mail: ann.depicker@psb.ugent.be, mifangl@vip.sina.com, or marc.vanmontagu@agent.ac.be.

This article contains supporting information online at www.pnas.org/cgi/content/full/0609997104/DC1.

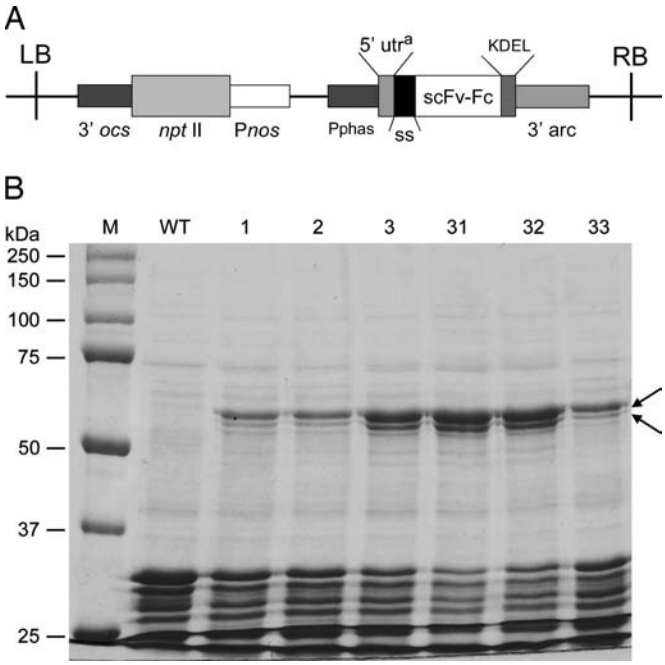

Fig. 1. Expression of scFv-Fc in transgenic *Arabidopsis* plants. (*A*) Schematic diagram of the T-DNA region in the binary vector pPhas. P*nos*, nopaline synthase gene promoter; *npt* II, coding sequence of the neomycin phosphotransferase II gene; 3′ *ocs*, 3′ end of the octopine synthase gene; Pphas, β-phaseolin gene promoter (−1 to −1,470; GenBank accession no. J01263); 5′ utr[a], 5′ UTR of *arc5-I* gene (13 bp; part of GenBank accession no. Z5020) (42); 3′ arc, 3′-flanking regulatory sequences of the *arc5-I* genomic clone (3,900 bp; part of GenBank accession no. Z50202) (42); ss, signal peptide of the *Arabidopsis* 2S2 seed storage protein gene (43); scFv-Fc, scFv-Fc coding sequence; KDEL, ER retention signal (16); RB and LB, T-DNA right and left border, respectively. (*B*) Coomassie-stained, reducing (+DTT) SDS/PAGE gel on which 40 μg was loaded of total seed protein extracts from wild-type (WT), and transgenic MBP10-producing *Arabidopsis* lines 1, 2, 3, 31, 32, and 33. The arrows indicate the doublet of scFv-Fc monomer bands. M, Precision Plus Protein Dual Color Standards (Bio-Rad).

HA16 were much lower and could only be detected by protein gel blots, in which multiple low-molecular-mass bands also were observed, indicating that HA16 probably was proteolytically degraded (data not shown). Therefore, HA16 lines were not included for further analysis. Under reducing conditions, scFv-Fc monomers migrated as doublets of major bands at the expected molecular mass of 55 kDa (Fig. 1*B*). Homodimers of

Table 1. Accumulation of scFv-Fc as a percentage of TSP content (±SD) in transgenic *Arabidopsis* seed stocks

Line	Seg T2*	Seg T3*	Ho T3†
MBP10–31	12.4 ± 1.9	7.4 ± 1.7	9.2 ± 0.3
		3.8 ± 0.6	8.8 ± 1.4
MBP10–39	11.6 ± 1.2	5.1 ± 3.4	9.4 ± 3.6
		5.4 ± 1.1	7.7 ± 0.7
HA78–6	8.4 ± 0.7	6.3 ± 1.8	12.5 ± 1.8
		4.8 ± 0.4	5.5 ± 0.7
HA78–8	9.0 ± 1.4	5.5 ± 0.7	13.1 ± 1.8
		4.8 ± 1.6	8.6 ± 0.6
EHF34–7	7.1 ± 0.6	7.0 ± 2.8	11.1 ± 0.6
		5.6 ± 0.9	9.6 ± 2.0
EHF34–10	6.9 ± 0.9	5.7 ± 1.0	13.9 ± 1.5
		5.0 ± 0.0	11.3 ± 3.5

*Segregating seed stocks.
†Homozygous seed stocks.

Table 2. LC-ESI-MS analysis of the glycosylated Fc peptides EEQYN[297]STYR and TKPREEQYN[297]STYR from scFv-Fc-producing *Arabidopsis* lines

	Relative abundance,* %					
	MBP10		HA78		EHF34	
Glycoform	Upper	Lower	Upper	Lower	Upper	Lower
None	2.1	76.5	8.2	93.4	1.9	97.2
$Man_5GlcNAc_2$	9.0	2.6	6.4	0.0	6.4	1.4
$Man_6GlcNAc_2$	14.4	2.9	7.6	0.2	9.5	0.0
$Man_7GlcNAc_2$	27.6	5.5	21.9	2.0	21.0	0.0
$Man_8GlcNAc_2$	28.5	9.1	39.9	2.8	44.9	0.0
$Man_9GlcNAc_2$	0.7	0.3	3.2	0.0	0.0	0.0

*Numbers are the sum of the relative abundances of a specific glycoform found on each of the two glycopeptides.

≈2-fold the apparent molecular weight were detected under nonreducing conditions [supporting information (SI) Fig. 4], confirming that the scFv-Fc proteins assembled in the seeds as IgG-like dimers through disulfide bonds formed by cysteine residues present in the hinge region of the Fc. T2-segregating seeds stocks with the highest density of scFv-Fc bands on Coomassie-stained gels were selected out of 20 independent transformants and tested for single-locus T-DNA insertion. For each of the three constructs, the scFv-Fc concentration was estimated in the two best lines with a single T-DNA locus, and their segregating and homozygous offspring were determined by visual comparison with a standard on Coomassie-stained gels (SI Fig. 5). Accumulation levels of the three different constructs were in the same range, varying from 7.0% to 12.5% of TSP for EHF34 and MBP10 in T2-segregating seed stocks, respectively. In some lines, the expected dosage effects for T3 homozygous versus segregating seed stocks resulted in slightly higher expression levels in the homozygous seed stocks (Table 1). The scFv-Fc accumulation levels observed in these homozygous lines was retained over several subsequent generations (data not shown).

MS Analysis of Tryptic Peptides and Glycopeptides. scFv-Fc proteins have a conserved N-linked glycosylation site in the C_H2 domain of the Fc. To verify whether the doublet of the scFv-Fc bands observed on SDS/PAGE gels under nonreducing conditions (Fig. 1*B*) was attributable to differential glycosylation, total seed protein extracts of a representative transgenic line for each scFv-Fc were treated with PNGase F, which releases all *N*-glycans devoid of a core α(1,3)-fucosyl residue that are essentially oligomannosidic glycans in plants. Upon treatment with PNGase F, the doublet of recombinant bands had collapsed to a single band (SI Fig. 6), running at the same position as the lower band of the original doublet, indicating that the lower band represented a nonglycosylated scFv-Fc. The doublet bands were isolated separately from SDS/PAGE gels and subjected to a detailed glycoproteomic analysis by liquid chromatography/electrospray ionization MS (LC-ESI-MS). The lower scFv-Fc bands predominantly contained nonglycosylated peptides (Table 2 and SI Fig. 7*A*), confirming the results found with PNGase F. The respective peptides isolated from the upper bands were N-glycosylated to a large extent. The *N*-glycans were exclusively of the high-mannose type with five to nine mannose residues. The most abundant structures were $Man_8GlcNac_2$ and $Man_7GlcNac_2$ (Table 2 and SI Fig. 7*B*). As expected, no complex xylose-containing or α(1,3)-fucose-containing *N*-glycans were detected in the mass spectra of the KDEL-tagged scFv-Fc proteins. When doublet bands were not excised separately from gel and analyzed as one sample, the unglycosylated isoform represented between 35% and 40% in each sample (data not

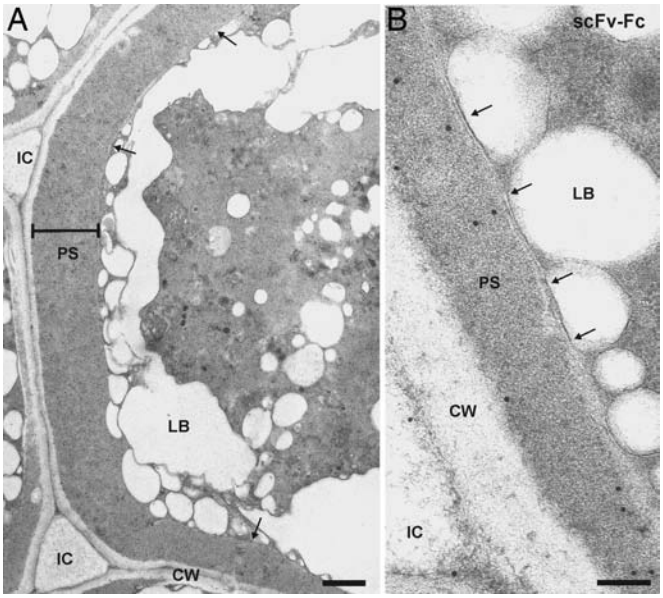

Fig. 2. Localization of scFv-Fc in imbibed transgenic *Arabidopsis* seeds. (*A*) Overview of a storage parenchyma cell in the cotyledon. A striking feature is the presence of an extensive electron opaque layer (PS) between the cell wall (CW) and the plasma membrane (arrows). (*B*) High magnification showing the immunogold (15-nm) detection of scFv-Fc molecules in the PS. LB, lipid body; IC, intercellular space. (Scale bars: *A*, 750 nm; *B*, 150 nm.)

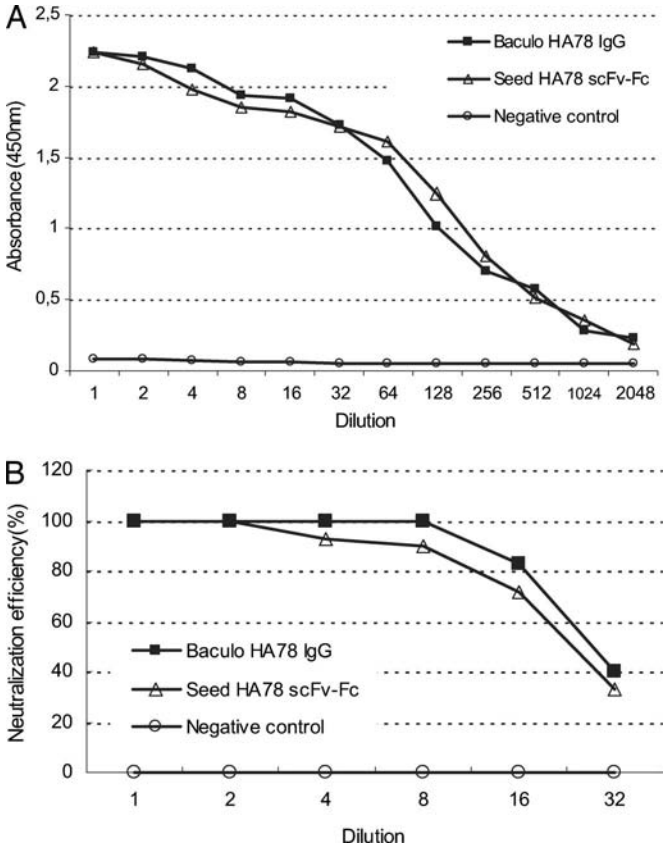

Fig. 3. Functional analysis of seed-produced HA78. (*A*) ELISA analysis of antigen-binding activity of HA78 produced and purified from transgenic *Arabidopsis* seed. ELISA was conducted with 96-well plates coated with HAV antigen purified from the Chinese strain JN. Seed HA78 scFv-Fc, 100 μl (2.5 μg/ml) of purified HA78 from transgenic *Arabidopsis* seed; Baculo HA78 IgG, 100 μl (2.5 μg/ml) of purified anti-hepatitis HA78 IgG from the baculovirus-infected insect cell supernatant; negative control, 100 μl of PBS. (*B*) *In vitro* neutralization activity of *Arabidopsis* seed-produced HA78 against HAV strain JN. A dose of 50× $TCID_{50}$ of HAV strain JN was incubated with a 2-fold serial dilution of ProtA-purified *Arabidopsis* HA78 antibody (seed HA78 scFv-Fc), purified full-length anti-HAV HA78 IgG (baculo HA78 IgG), and anti-HTNV AH100 IgG (negative control), starting from 80 μg/ml for all three antibodies. After inoculation of FRhK-4 cells with the virus-antibody mix and 21 days of growth, cells were harvested, and the presence of HAV was monitored by ELISA.

shown), roughly corresponding to the ratio of upper and lower band intensities on Coomassie-stained SDS/PAGE gels (Fig. 1*B* and SI Fig. 5). Furthermore, LC-ESI-MS analysis showed that the unglycosylated scFv-Fc fractions never had been glycosylated because they had retained the Asn residue at the respective position. Deglycosylation would have led to the formation of an Asp residue and, consequently, to a detectable mass increase of 1 Da. Finally, the N-terminal sequence of the scFv-Fc recombinant protein present in the upper and lower bands of a representative transgenic line was verified for each of the three different constructs. In all cases, the correctly processed N-terminal peptide was found, and its identity was established by tandem MS analysis (SI Fig. 8), confirming correct ER translocation and cleavage of the 2S2 signal peptide.

EM Localization Study. Because of the high scFv-Fc accumulation levels and significant impact of the subcellular localization on structure, stability, and yield of recombinant proteins produced in plants, we performed an immunogold EM localization study on ultrathin sections of imbibed mature MBP10-producing transgenic *Arabidopsis* seeds, prepared by high-pressure freezing/freeze substitution. Unexpectedly, electron micrographs revealed the presence of scFv-Fc in huge deposits in the periplasmic space (PS), between the plasma membrane and the cell wall (Fig. 2), but not in the protein storage vacuole (PSV) of transgenic seeds. To determine whether the periplasmic deposits consisted uniquely of recombinant scFv-Fc, the distribution of two ER-resident chaperones, calreticulin and binding protein (BiP), and the endogenous seed storage protein cruciferin (11S globulin) was characterized by immunogold labeling with specific antisera. Surprisingly, these three endogenous proteins also were found in the PS (SI Fig. 9). Furthermore, EM of wild-type seeds did not reveal the presence of protein deposits outside of the plasma membrane, and cruciferin was located exclusively in the PSVs (SI Fig. 10).

Aside from their aberrant localization in the PS, recombinant scFv-Fc, calreticulin, and BiP also were detected in internal compartments delimited by ribosome-associated membranes, suggesting that they are dilated ER (SI Fig. 11).

Functionality of Seed-Produced HA78. HA78 was purified with protein A (ProtA) from pooled seed protein extracts obtained from T2-segregating *Arabidopsis* seed stocks and yielded 1 mg of scFv-Fc from <200 mg of dry seeds. Similar yields were obtained with pooled seed protein extracts from T2-segregating seed stocks producing MBP10 or EHF34. No degradation products were found in the purified scFv-Fc samples via protein gel blots with an anti-human IgG (Fc-specific) antibody (data not shown). ProtA-purified HA78, produced in *Arabidopsis* seeds, and ProtG-purified anti-HA78 full-length IgG, produced in insect cells with the baculovirus expression system, were applied to ELISA plates coated with hepatitis A virus (HAV) antigen purified from Chinese strain JN. The absorbance values of *Arabidopsis* HA78 were similar to those of the full-length anti-HA78 IgG at all dilutions (Fig. 3*A*). These results confirm the binding activity of the *Arabidopsis*-produced scFv-Fc to HAV antigen. The neutralizing activity of seed-purified HA78 and

anti-HA78 IgG, purified from insect cells against HAV strain JN, was compared *in vitro* with a cell-culture assay. Neutralization efficiency of *Arabidopsis* HA78 was comparable to that of the full-length anti-HA78 IgG, showing a similar dose-dependent response, whereas anti-Hantaan virus (HTNV) IgG used as negative control did not neutralize HAV (Fig. 3*B*). These data demonstrate that the HAV-neutralizing activity of the *Arabidopsis*-produced HA78 is as high as that of the full-length IgG HA78 produced in insect cells.

Discussion

As an alternative antibody format, scFv antibody fragments were fused to the Fc fragment of a human IgG. The production of scFv-Fc monomers, as single-gene products, results in correctly assembled IgG-like bivalent homodimers *in planta*. The use of regulatory sequences of common bean resulted in outstanding accumulation levels for three of the four scFv-Fc fusion proteins in T3-homozygous *Arabidopsis* seed stocks, ranging from 9.5% to 14.0% of TSP, equivalent to 19–28 μg of scFv-Fc per mg of seed. For all three constructs, the variation in transgene expression was low, and high-accumulating lines could be identified in a small group of 20 screened independent transformants. These results confirm the scFv levels of 20% of TSP obtained previously in *Arabidopsis* seeds with the same β-phaseolin promoter construct (5).

Understanding how proteins are targeted intracellularly is important when endeavoring to exploit the plant secretory pathway for heterologous protein production. The EM localization study revealed that the KDEL-tagged MBP10 accumulated both in ER-derived spherical bodies and also, unexpectedly, in the PS between the plasma membrane and the cell wall, suggesting a direct transport of the MBP10 fusion proteins from the ER to the PS that bypasses the Golgi. These observations are consistent with the glycosylation analysis: glycosylated scFv-Fc proteins contained only high-mannose *N*-glycans and no complex glycans. The presence of $Man_5GlcNac_2$ to $Man_7GlcNac_2$ suggests that the *N*-glycans have been trimmed by α-mannosidase I, an enzyme located in the *cis*-Golgi, and probably are sent back to the ER via retrograde transport mediated by the KDEL receptor (16). A similar *N*-glycan profile was found for a full-length IgG produced in tobacco (*Nicotiana tabacum*) leaves in which both heavy and light chains were fused via a long linker motif to the KDEL sequence to increase accessibility of the KDEL tag (17). Although undetected here, complex glycans were found on 10–20% of the full-length KDEL-tagged IgGs, produced in tobacco leaves (18, 19). Recently, a full-length antibody produced in leaves and seeds of tobacco has been characterized in detail (20). A SEKDEL tag was attached to both heavy and light chains. As expected, the leaf-produced IgG was efficiently retained in the ER and carried only high-mannose *N*-glycans, whereas, in contrast to our findings, the seed-produced SEKDEL-tagged antibody was found in both the PSV and the intercellular space, bearing complex glycans, suggesting a Golgi-dependent transport pathway in tobacco seeds. The same SEKDEL signal used in tobacco leaves did not prevent the transport of protein to late Golgi compartments in tobacco seeds (20). Similar tissue-dependent differences were observed when subcellular accumulation and glycosylation of a model glycoprotein were studied in rice (*Oryza sativa*). Although the recombinant protein was efficiently secreted from leaf cells, it was present in ER-derived prolamin bodies and PSVs within the endosperm. Consistent with the immunolocalization data, the endosperm-produced protein possessed oligomannose and vacuole-type *N*-glycans, whereas the leaf-produced protein contained predominantly secretion-type *N*-glycans (21). Our data indicate that the shorter KDEL tag is sufficient for efficient ER retention/cis-Golgi retrieval of scFv-Fc proteins produced in *Arabidopsis* seeds but that their high-level accumulation seems to overwhelm the ER storage capacity, inducing an export pathway of the recombinant protein to the PS. The presence of calreticulin and BiP in the PS, as well as cruciferin that normally accumulates in the PSV (22), suggests that the overproduction of the recombinant scFv-Fc also disturbs the retention and sorting mechanisms of endogenous proteins. Given their strong biological and physical association (23), the colocalization of calreticulin and BiP in the PS is not surprising. Under normal circumstances, these ER-resident proteins are not found outside the plasma membrane (24). The localization of cruciferin in the PS indicates that the receptor-mediated transport of *Arabidopsis* seed-storage proteins to PSV (22) also is affected by the high scFv-Fc production level. Altogether, these findings suggest that, aside from accessibility and density of the KDEL sequence on the recombinant protein and tissue-specific differences in ER retention/retrieval efficiency, protein- and species-specific factors also influence the fate of a recombinant protein in the secretory pathway.

An explanation for the observation of a nonglycosylated scFv-Fc fraction might be that the plant ER translocation and glycosylation machinery is not able to hold all of the heterologous scFv-Fc proteins in a flexible condition. Nascent eukaryotic proteins are known to have to be in such a flexible state to be susceptible to oligosaccharyl transferase complex-induced conformational changes and subsequent attachment of the $Glc_3Man_9GlcNac_2$ precursor (25). Saturation of the glycosylation machinery by the high rate of recombinant protein synthesis seems less likely because this hypothesis is in contrast with our observation that a fraction of unglycosylated scFv-Fc also is found in transgenic lines producing proteins at levels below 0.5% of TSP (B.V.D., A.D., and G.D.J., unpublished data), which we consider too low to be "saturating" for the glycosylation machinery. Alternatively, the underglycosylation could be related specifically to the scFv-Fc format because a fraction of nonglycosylated scFv-Fc occurred when scFv-Fc antibodies were produced in *Pichia* (8). Possibly, folding and coupling of two scFv-Fc monomers result in an Fc pocket smaller than that of the full-length IgGs. The presence of the large triantennary high-mannose-type *N*-glycans on the Fc domain of one glycosylated scFv-Fc monomer, attached both in *Pichia* (26) and plants, could prevent the pairing with another glycosylated scFv-Fc but only allow coupling to nonglycosylated scFv-Fc monomers. Such a process could explain that about the same ratio of glycosylated versus nonglycosylated scFv-Fc fusion protein is observed in all transgenic lines, independent of the expression levels. Future investigations will be necessary to resolve this phenomenon.

Finally, the binding assay of seed-produced HA78 and corresponding full-length anti-HA78 IgG, produced in insect cells, illustrated similar specificities in recognizing the HAV-purified antigen. These results confirm that neither the conversion of a scFv or full-length IgG into the scFv-Fc format (12, 27) nor the altered glycosylation on the C_H2 domain of the plant-produced antibody affects the antigen-specific binding activity (18, 28). In addition, the seed-derived HA78 had an *in vitro* neutralizing activity against the Chinese HAV strain JN comparable to that of the full-length IgG produced in insect cells. Therefore, recombinant plant-produced HA78 could provide an alternative source of neutralizing antibody to be used in pre- and postexposure HAV prophylaxis (29). The importance of steric hindrance and valency in antibody-dependent virus neutralization has been demonstrated conclusively (30). Our results are consistent with those of other successful *in vitro* and *in vivo* virus neutralization based on the use of bivalent scFv-Fc fusion proteins (12). Altogether, we illustrate the value of a seed-specific expression platform for the safe and inexpensive production of antibodies in the scFv-Fc format, which are easier to produce and have the potential to replace IgGs in specific detection, purification, diagnostic, and even therapeutic applications.

Materials and Methods

scFv-Fc Construction. Four scFv-Fc fusion proteins were used: MBP10 is based on an anti-MBP scFv, selected from a human synthetic scFv phage display library (9); HA16 and HA78 are based on two different human-derived, neutralizing Fab antibodies to HAV (31); and EHF34 is based on a human-derived Fab antibody to HTNV nucleocapsid protein (32). For the latter three, the V_H and V_L genes were amplified by PCR with forward and reverse primers containing NcoI and XhoI and ApaLI and NotI sites in the 5′ and 3′ ends, respectively, and aligned as scFv sequences in the pHEN2 phagemid vector. scFv-encoding sequences were cut from pHEN2-scFv by SfiI and NotI and fused to a human IgG1 Fc domain (hinge, C_H2 and C_H3) in pPICZαFc (9). scFv-Fc expression cassettes were cloned into pPphas (Fig. 1*A*).

Plant Transformation. *Arabidopsis thaliana* (L.) Heyhn., ecotype Columbia 0, was transformed by floral dip (33) with *Agrobacterium tumefaciens* strain C58C1Rif^R (34). Transgenic T1 plants were selected on kanamycin-containing medium and maintained in soil. Single-locus transgenic lines and homozygous T3 seed stocks were identified as described in ref. 5.

scFv-Fc Detection and Quantification. *Arabidopsis* seeds were extracted, and TSP concentrations were determined as described in ref. 5. Essentially, soluble proteins were separated by SDS/PAGE, under reducing conditions (+DTT), and visualized by Coomassie-blue staining. Accumulation levels of scFv-Fc recombinant proteins in the best lines were estimated (as percentage of TSP) by comparing the intensity of the scFv-Fc band in the seed extract with that of different standards that contained different amounts of MBP10 proteins, produced in *Pichia* and purified with ProtA affinity chromatography (see below). To verify the *in vivo* dimerization of scFv-Fc monomers, soluble proteins were separated in NuPAGE Novex Bis-Tris 4–12% gradient gels (Invitrogen, Carlsbad, CA) under nonreducing conditions.

MS Analysis of Tryptic Peptides and Glycopeptides. PNGase F digests were performed on 10 μg of total seed protein extract with 500 units of enzyme (New England Biolabs, Ipswich, MA), according to the manufacturer's instructions. For MS analysis of *N*-glycans, Coomassie-stained bands were excised, destained, carbamidomethylated, digested with trypsin, extracted from gel pieces (35), and analyzed (36). First, the samples were analyzed in plain MS mode to facilitate detection of the sometimes less intense glycopeptide signals. Sequences of the N-terminal peptides were confirmed by tandem MS experiments. Data were analyzed with MassLynx 4.0 SP4 Software (Waters Micromass, Milford, MA). Possible glycopeptide masses were obtained from *in silico* tryptic digests by the addition of relevant glycan masses.

EM Localization Study. *Arabidopsis* seeds were imbibed with water for 20 h at 4°C. Seed coats were removed after seed incubation in hexadecene. Five to eight submerged seeds were mounted onto planchettes and frozen in a high-pressure freezer (HPF010; Bal-Tec, Balzers, Liechtenstein). Freeze substitution was carried out as described in ref. 37, except for an extra day of incubation at −85°C and two extra washing steps in 100% ethanol for 60 min before embedding in Lowicryl HM20 (Electron Microscopy Sciences, Hatfield, PA). To increase sectioning quality, the blocks were hardened with UV light for 3 days at room temperature. Ultrathin sections were cut on an Ultracut S (Leica, Bannockburn, IL) and incubated with antibodies against cruciferin (ProtA-purified serum at 1:50 dilution), affinity-purified BiP (at 1:25–1:50 dilution), and calreticulin (at 1:50–1:100 dilution) antibodies, followed by incubation with 10-nm gold-coupled secondary antibodies (GAR10; British Biocell International, Cardiff, U.K.) at a dilution of 1:50 in PBS supplemented with 1% BSA or directly with a ProtA-gold conjugate (EM.PAG10; British Biocell International). Sections were poststained with aqueous uranyl acetate/lead citrate and examined by transmission electron microscopy (CM10; Philips, Amsterdam, The Netherlands). For conventional fixation, imbibed, testa-free seeds were immersed in a phosphate-buffered 2% glutaraldehyde solution at room temperature for 2 h, washed in buffer, postfixed in 2% osmium tetroxide for 4 h at room temperature, washed again in buffer, and dehydrated in an acetone series. Specimens were embedded in Spurr's resin.

Purification of scFv-Fc. *Arabidopsis* seed protein extracts were filtered twice over a 0.45-μm filter (Alltech, Nicholasville, KY) and diluted to 1 mg/ml TSP with 10% glycerol (VWR, West Chester, PA). In *Pichia*, production of scFv-Fc was as described in ref. 9. Basically, it was purified over a ProtA Sepharose 4 Fast Flow column (GE Healthcare, Fairfield, CT) according to the manufacturer's instructions, eluted with 0.1 M glycine·HCl (pH 3.0), and dialyzed overnight with 1× PBS.

HA78 Functional Assays. HAV antigen was produced by infecting FRhK-4 cells with HAV Chinese JN strain, culturing the infected cells, and purifying the HAV particles by sucrose-gradient sedimentation (38). ELISA plates were coated overnight at 4°C with purified HAV antigens, blocked with 2% BSA in PBS for 1.5 h, and incubated with a 2-fold dilution series of *Arabidopsis* HA78, starting at 2.5 μg/ml. Full-length anti-HA78 IgG, produced in insect cells (32), was included as positive control. After washing, bound scFv-Fc or IgG was detected by goat anti-human IgG (Fc-specific), conjugated to HRP (1:1,000 dilution; Sigma–Aldrich, St. Louis, MO). The ELISA was developed with 3,3′,5,5′-tetramethylbenzidine as a substrate, and absorbance was measured at 405 nm.

HAV neutralization was assayed as described in ref. 39 with modifications. Chloroform-treated HAV JN strain was diluted in 10-fold steps from 10^{-1} to 10^{-5} to determine the tissue culture 50% infective dose ($TCID_{50}$) in FRhK-4 cells (40). For the titer of infectious virus, the method of Reed and Muench (41) and 50× $TCID_{50}$ of HAV were used. A 2-fold serial dilution of *Arabidopsis* HA78 antibody was incubated with HAV JN strain at 37°C for 2 h. The full-length anti-HA78 IgG, produced as described above, was included as positive control, and anti-HNTV IgG antibody AH100 (32) was included as negative control. Mixtures of antibody and virus were added to monolayers of FRhK-4 cells growing in 24-well plates. Mock-infected cells were included as negative control. Plates were incubated at 37°C for 1.5 h, the medium was removed, and 1.0 ml of fresh culture medium was added per well. Monolayers were incubated for 21 days at 33°C, with change of culture medium every 5–6 days. Cells were harvested after 21 days with 0.2 ml of PBS plus 0.05% Tween 20 and disrupted by three cycles of freeze-thawing. HAV titers were monitored by ELISA with human anti-HAV PcAb as coating antibody (1:1,000 dilution) and HRP-labeled human anti-HAV PcAb as detecting antibody (1:1,000 dilution). Neutralization efficiency was calculated according to the method of Reed and Muench (41).

We thank Geert Persiau, Dominique Eeckhout, Eveline Van De Slijke, and Wilson Ardiles-Diaz for technical help or expertise and Martine De Cock for help in preparing the manuscript. This work was supported by the Pharma–Planta Consortium, which is funded by the European Union through the Framework 6 research program, and a Chinese–Flemish bilateral scientific and technological cooperation, funded by the European Union through the Framework 6 Research Program Grant 011S0203.

1. Ma JK-C, Barros E, Bock R, Christou P, Dale PJ, Dix PJ, Fischer R, Irwin J, Mahoney R, Pezzotti M, *et al.* (2005) *EMBO Rep* 6:593–599.
2. Abranches R, Marcel S, Arcalis E, Altmann F, Fevereiro P, Stoger E (2005) *J Biotechnol* 120:121–134.
3. Menkhaus TJ, Bai Y, Zhang C, Nikolov ZL, Glatz CE (2004) *Biotechnol Prog* 20:1001–1014.
4. Stoger E, Ma JK-C, Fischer R, Christou P (2005) *Curr Opin Biotechnol* 16:167–173.
5. De Jaeger G, Scheffer S, Jacobs A, Zambre M, Zobell O, Goossens A, Depicker A, Angenon G (2002) *Nat Biotechnol* 20:1265–1268.
6. Shu L, Qi C-F, Schlom J, Kashmiri SVS (1993) *Proc Natl Acad Sci USA* 90:7995–7999.
7. Li S-L, Liang S-J, Guo N, Wu AM, Fujita-Yamaguchi Y (2000) *Cancer Immunol Immunother* 49:243–252.
8. Powers DB, Amersdorfer P, Poul M-A, Nielsen UB, Shalaby MR, Adams GP, Weiner LM, Marks JD (2001) *J Immunol Methods* 251:123–135.
9. Eeckhout D, De Clercq A, Van De Slijke E, Van Leene J, Stals H, Casteels P, Persiau G, Vercammen D, Van Breusegem F, Zabeau M, *et al.* (2004) *J Immunol Methods* 294:181–187.
10. Valadon P, Garnett JD, Testa JE, Bauerle M, Oh P, Schnitzer JE (2006) *Proc Natl Acad Sci USA* 103:407–412.
11. Wang B, Chen Y-B, Ayalon O, Bender J, Garen A (1999) *Proc Natl Acad Sci USA* 96:1627–1632.
12. Gould LH, Sui J, Foellmer H, Oliphant T, Wang T, Ledizet M, Murakami A, Noonan K, Lambeth C, Kar K, *et al.* (2005) *J Virol* 79:14606–14613.
13. Brocks B, Rode H-J, Klein M, Gerlach E, Dübel S, Little M, Pfizenmaier K, Moosmayer D (1997) *Immunotechnology* 3:173–184.
14. Vitale A, Pedrazzini E (2005) *Mol Interventions* 5:216–225.
15. Gomord V, Sourrouille C, Fitchette A-C, Bardor M, Pagny S, Lerouge P, Faye L (2004) *Plant Biotechnol J* 2:83–100.
16. Lerouge P, Cabanes-Macheteau M, Rayon C, Fischette-Lainé A-C, Gomord V, Faye L (1998) *Plant Mol Biol* 38:31–48.
17. Triguero A, Cabrera G, Cremata JA, Yuen, C-T, Wheeler J, Ramírez NI (2005) *Plant Biotechnol J* 3:449–457.
18. Ko K, Tekoah Y, Rudd PM, Harvey DJ, Dwek RA, Spitsin S, Hanlon CA, Rupprecht C, Dietzschold B, Golovkin M, *et al.* (2003) *Proc Natl Acad Sci USA* 100:8013–8018.
19. Sriraman R, Bardor M, Sack M, Vaquero C, Faye L, Finnern R, Lerouge P (2004) *Plant Biotechnol J* 2:279–287.
20. Petruccelli S, Otegui MS, Lareu F, Tran Dinh O, Fitchette A-C, Circosta A, Rumbo M, Bardor M, Carcamo R, Gomord V, Beachy RN (2006) *Plant Biotechnol J* 4:511–527.
21. Drakakaki G, Marcel S, Arcalis E, Altmann F, Gonzalez-Melendi P, Fischer R, Christou P, Stoger E (2006) *Plant Physiol* 141:578–586.
22. Shimada T, Fuji K, Tamura K, Kondo M, Nishimura M, Hara-Nishimura I (2003) *Proc Natl Acad Sci USA* 100:16095–16100.
23. Crofts AJ, Leborgne-Castel N, Pesca M, Vitale A, Denecke J (1998) *Plant Cell* 10:813–823.
24. Pimpl P, Taylor JP, Snowden C, Hillmer S, Robinson DG, Denecke J (2006) *Plant Cell* 18:198–211.
25. Bulleid NJ, Bassel-Dubym RS, Freedman RB, Sambrook JF, Gething MJ (1992) *Biochem J* 286:275–280.
26. Grinna LS, Tschopp JF (1989) *Yeast* 5:107–115.
27. Bose B, Khanna N, Acharya SK, Sinha S (2006) *J Gastroenterol Hepatol* 21:1439–1447.
28. Ko K, Steplewski Z, Glogowska M, Koprowski H (2005) *Proc Natl Acad Sci USA* 102:7026–7030.
29. Fiore AE, Wasley A, Bell BP (2006) *Morbid Mortal Wkly Rep* 55(RR-7):1–23.
30. Burton DR (2002) *Nat Rev Immunol* 2:706–713.
31. Cao JY, Liang MF, Meng QL, Wang XF, Xu YG, Guo KQ, Zhan MY, Bi SL, Li DX (2004) *Chin J Exp Clin Virol* 18:20–23 (in Chinese).
32. Liang M, Mahler M, Koch J, Ji Y, Li D, Schmaljohn C, Bautz EKF (2003) *J Med Virol* 69:99–107.
33. Clough SJ, Bent AF (1998) *Plant J* 16:735–743.
34. Koncz C, Schell J (1986) *Mol Gen Genet* 204:383–396.
35. Kolarich D, Altmann F (2000) *Anal Biochem* 285:64–75.
36. Kolarich D, Altmann F, Sunderasan E (2006) *Biochim Biophys Acta* 1760:715–720.
37. Oliviusson P, Heinzerling O, Hillmer S, Hinz G, Tse YC, Jiang L, Robinson DG (2006) *Plant Cell* 18:1239–1252.
38. Hughes JV, Stanton LW, Tomassini JE, Long WJ, Scolnick EM (1984) *J Virol* 52:465–473.
39. MacGregor A, Kornitschuk M, Hurrell JGR, Lehmann NI, Coulepis AG, Locarnini SA, Gust ID (1983) *J Clin Microbiol* 18:1237–1243.
40. Lemon SM, Binn LN (1985) *J Gen Virol* 66:2501–2505.
41. Reed LJ, Muench H (1938) *Am J Hyg* 27:493–497.
42. Goossens A, Dillen W, De Clercq J, Van Montagu M, Angenon G (1999) *Plant Physiol* 120:1095–1104.
43. Krebbers E, Herdies L, De Clercq A, Seurinck J, Leemans J, Van Damme J, Segura M, Gheysen G, Van Montagu M, Vandekerckhove J (1988) *Plant Physiol* 87:859–866.

REFERENCES

1. Jenkins, N., L. Murphy, and R. Tyther, *Post-translational modifications of recombinant proteins: significance for biopharmaceuticals.* Mol Biotechnol, 2008. **39**(2): p. 113-8.
2. Walsh, G. and R. Jefferis, *Post-translational modifications in the context of therapeutic proteins.* Nat Biotechnol, 2006. **24**(10): p. 1241-52.
3. Dwek MV, S.U., *Functional and Molecular Glycobiology*. 2002: Garland Science. 256.
4. Taylor ME, D.K., *Introduction to Glycobiology*. 2003: Oxford University Press, USA. 160.
5. Varki A, C.R., Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME, ed. *Essentials of Glycobiology*. 2nd ed. 2008, Cold Spring Harbor Laboratory Press: Plainview (NY). 784.
6. Shi, X. and D.L. Jarvis, *Protein N-glycosylation in the baculovirus-insect cell system.* Curr Drug Targets, 2007. **8**(10): p. 1116-25.
7. Paschinger, K., et al., *The N-glycosylation pattern of Caenorhabditis elegans.* Carbohydr Res, 2008. **343**(12): p. 2041-9.
8. Deshpande, N., et al., *Protein glycosylation pathways in filamentous fungi.* Glycobiology, 2008. **18**(8): p. 626-37.
9. Werner, R.G., K. Kopp, and M. Schlueter, *Glycosylation of therapeutic proteins in different production systems.* Acta Paediatr Suppl, 2007. **96**(455): p. 17-22.
10. Ahn, W.S., et al., *Effect of culture temperature on erythropoietin production and glycosylation in a perfusion culture of recombinant CHO cells.* Biotechnol Bioeng, 2008. **101**(6): p. 1234-44.
11. Stadlmann, J., et al., *Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides.* Proteomics, 2008. **8**(14): p. 2858-71.
12. Qian, J., et al., *Structural characterization of N-linked oligosaccharides on monoclonal antibody cetuximab by the combination of orthogonal matrix-assisted laser desorption/ionization hybrid quadrupole-quadrupole time-of-flight tandem mass spectrometry and sequential enzymatic digestion.* Anal Biochem, 2007. **364**(1): p. 8-18.
13. Kunkel, J.P., et al., *Comparisons of the glycosylation of a monoclonal antibody produced under nominally identical cell culture conditions in two different bioreactors.* Biotechnol Prog, 2000. **16**(3): p. 462-70.
14. Kunkel, J.P., et al., *Dissolved oxygen concentration in serum-free continuous culture affects N-linked glycosylation of a monoclonal antibody.* J Biotechnol, 1998. **62**(1): p. 55-71.
15. Muthing, J., et al., *Effects of buffering conditions and culture pH on production rates and glycosylation of clinical phase I anti-melanoma mouse IgG3 monoclonal antibody R24.* Biotechnol Bioeng, 2003. **83**(3): p. 321-34.
16. Trombetta, E.S., *The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis.* Glycobiology, 2003. **13**(9): p. 77R-91R.
17. Trombetta, E.S. and A.J. Parodi, *Quality control and protein folding in the secretory pathway.* Annu Rev Cell Dev Biol, 2003. **19**: p. 649-76.

18. Jefferis, R., *Human IgG Glycosylation in Inflammation and Inflammatory Disease,* in *Comprehensive glycoscience: from chemistry to systems biology*, J.P. Kamerling, Editor. 2007, Elsevier: Amsterdam. p. 373-392.
19. Jefferis, R. and J. Lund, *Interaction sites on human IgG-Fc for FcgammaR: current models.* Immunol Lett, 2002. **82**(1-2): p. 57-65.
20. Jefferis, R. and J. Lund, *Glycosylation of antibody molecules: structural and functional significance.* Chem Immunol, 1997. **65**: p. 111-28.
21. Jefferis, R. and D.S. Kumararatne, *Selective IgG subclass deficiency: quantification and clinical relevance.* Clin Exp Immunol, 1990. **81**(3): p. 357-67.
22. Lund, J., et al., *Oligosaccharide-protein interactions in IgG antibody molecules: structural and functional consequences.* Biochem Soc Trans, 1995. **23**(1): p. 102S.
23. Lund, J., et al., *Oligosaccharide-protein interactions in IgG can modulate recognition by Fc gamma receptors.* Faseb J, 1995. **9**(1): p. 115-9.
24. Jefferis, R., J. Lund, and M. Goodall, *Recognition sites on human IgG for Fc gamma receptors: the role of glycosylation.* Immunol Lett, 1995. **44**(2-3): p. 111-7.
25. Jefferis, R., J. Lund, and J.D. Pound, *IgG-Fc-mediated effector functions: molecular definition of interaction sites for effector ligands and the role of glycosylation.* Immunol Rev, 1998. **163**: p. 59-76.
26. Jefferis, R., et al., *Effector mechanisms activated by human IgG subclass antibodies: clinical and molecular aspects. Review article.* Ann Biol Clin (Paris), 1994. **52**(1): p. 57-65.
27. Lund, J., et al., *Multiple interactions of IgG with its core oligosaccharide can modulate recognition by complement and human Fc gamma receptor I and influence the synthesis of its oligosaccharide chains.* J Immunol, 1996. **157**(11): p. 4963-9.
28. Mimura, Y., et al., *The molecular specificity of IgG-Fc interactions with Fc gamma receptors.* Adv Exp Med Biol, 2001. **495**: p. 49-53.
29. Mimura, Y., et al., *Role of oligosaccharide residues of IgG1-Fc in Fc gamma RIIb binding.* J Biol Chem, 2001. **276**(49): p. 45539-47.
30. Takegawa, Y., et al., *Simple separation of isomeric sialylated N-glycopeptides by a zwitterionic type of hydrophilic interaction chromatography.* J Sep Sci, 2006. **29**(16): p. 2533-40.
31. Holland, M., et al., *Differential glycosylation of polyclonal IgG, IgG-Fc and IgG-Fab isolated from the sera of patients with ANCA-associated systemic vasculitis.* Biochim Biophys Acta, 2006. **1760**(4): p. 669-77.
32. Holliger, P. and P.J. Hudson, *Engineered antibody fragments and the rise of single domains.* Nat Biotechnol, 2005. **23**(9): p. 1126-36.
33. Schahs, M., et al., *Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern.* Plant Biotechnol J, 2007. **5**(5): p. 657-63.
34. Yoo, E.M., et al., *Myeloma expression systems.* J Immunol Methods, 2002. **261**(1-2): p. 1-20.
35. van Berkel, P.H., et al., *N-linked glycosylation is an important parameter for optimal selection of cell lines producing biopharmaceutical human IgG.* Biotechnol Prog, 2009. **25**(1): p. 244-51.
36. Sethuraman, N. and T.A. Stadheim, *Challenges in therapeutic glycoprotein production.* Curr Opin Biotechnol, 2006. **17**(4): p. 341-6.

37. Verma, R., E. Boleti, and A.J. George, *Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems.* J Immunol Methods, 1998. **216**(1-2): p. 165-81.
38. Hamilton, S.R., et al., *Humanization of yeast to produce complex terminally sialylated glycoproteins.* Science, 2006. **313**(5792): p. 1441-3.
39. Li, H., et al., *Optimization of humanized IgGs in glycoengineered Pichia pastoris.* Nat Biotechnol, 2006. **24**(2): p. 210-5.
40. Cox, M.M., *Progress on baculovirus-derived influenza vaccines.* Curr Opin Mol Ther, 2008. **10**(1): p. 56-61.
41. Harrison, R.L. and D.L. Jarvis, *Protein N-glycosylation in the baculovirus-insect cell expression system and engineering of insect cells to produce "mammalianized" recombinant glycoproteins.* Adv Virus Res, 2006. **68**: p. 159-91.
42. Fischer, R., et al., *Plant-based production of biopharmaceuticals.* Curr Opin Plant Biol, 2004. **7**(2): p. 152-8.
43. Ma, J.K., et al., *Molecular farming for new drugs and vaccines. Current perspectives on the production of pharmaceuticals in transgenic plants.* EMBO Rep, 2005. **6**(7): p. 593-9.
44. Stoger, E., et al., *Plantibodies: applications, advantages and bottlenecks.* Curr Opin Biotechnol, 2002. **13**(2): p. 161-6.
45. Twyman, R.M., et al., *Molecular farming in plants: host systems and expression technology.* Trends Biotechnol, 2003. **21**(12): p. 570-8.
46. Liu, H., et al., *Heterogeneity of monoclonal antibodies.* J Pharm Sci, 2008. **97**(7): p. 2426-47.
47. Liu, Y.D., et al., *Human IgG2 antibody disulfide rearrangement in vivo.* J Biol Chem, 2008. **283**(43): p. 29266-72.
48. Antes, B., et al., *Analysis of lysine clipping of a humanized Lewis-Y specific IgG antibody and its relation to Fc-mediated effector function.* J Chromatogr B Analyt Technol Biomed Life Sci, 2007. **852**(1-2): p. 250-6.
49. Liu, H., G. Gaza-Bulseco, and C. Chumsae, *Glutamine deamidation of a recombinant monoclonal antibody.* Rapid Commun Mass Spectrom, 2008. **22**(24): p. 4081-8.
50. Chen, X., Y.D. Liu, and G.C. Flynn, *The effect of Fc glycan forms on human IgG2 antibody clearance in humans.* Glycobiology, 2009. **19**(3): p. 240-9.
51. Malhotra, R., et al., *Glycosylation changes of IgG associated with rheumatoid arthritis can activate complement via the mannose-binding protein.* Nat Med, 1995. **1**(3): p. 237-43.
52. Chung, C.H., et al., *Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose.* N Engl J Med, 2008. **358**(11): p. 1109-17.
53. Shields, R.L., et al., *High resolution mapping of the binding site on human IgG1 for Fc gamma RI, Fc gamma RII, Fc gamma RIII, and FcRn and design of IgG1 variants with improved binding to the Fc gamma R.* J Biol Chem, 2001. **276**(9): p. 6591-604.
54. Shields, R.L., et al., *Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity.* J Biol Chem, 2002. **277**(30): p. 26733-40.
55. Scallon, B.J., et al., *Higher levels of sialylated Fc glycans in immunoglobulin G molecules can adversely impact functionality.* Mol Immunol, 2007. **44**(7): p. 1524-34.

56. Zhou, Q., et al., *Development of a simple and rapid method for producing non-fucosylated oligomannose containing antibodies with increased effector function.* Biotechnol Bioeng, 2008. **99**(3): p. 652-65.
57. Kanda, Y., et al., *Comparison of cell lines for stable production of fucose-negative antibodies with enhanced ADCC.* Biotechnol Bioeng, 2006. **94**(4): p. 680-8.
58. Shinkawa, T., et al., *The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity.* J Biol Chem, 2003. **278**(5): p. 3466-73.
59. Yamane-Ohnuki, N., et al., *Establishment of FUT8 knockout Chinese hamster ovary cells: an ideal host cell line for producing completely defucosylated antibodies with enhanced antibody-dependent cellular cytotoxicity.* Biotechnol Bioeng, 2004. **87**(5): p. 614-22.
60. Kanda, Y., et al., *Establishment of a GDP-mannose 4,6-dehydratase (GMD) knockout host cell line: a new strategy for generating completely non-fucosylated recombinant therapeutics.* J Biotechnol, 2007. **130**(3): p. 300-10.
61. Imai-Nishiya, H., et al., *Double knockdown of alpha1,6-fucosyltransferase (FUT8) and GDP-mannose 4,6-dehydratase (GMD) in antibody-producing cells: a new strategy for generating fully non-fucosylated therapeutic antibodies with enhanced ADCC.* BMC Biotechnol, 2007. **7**: p. 84.
62. Davies, J., et al., *Expression of GnTIII in a recombinant anti-CD20 CHO production cell line: Expression of antibodies with altered glycoforms leads to an increase in ADCC through higher affinity for FC gamma RIII.* Biotechnol Bioeng, 2001. **74**(4): p. 288-94.
63. Umana, P., et al., *Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity.* Nat Biotechnol, 1999. **17**(2): p. 176-80.
64. Decker, E.L. and R. Reski, *Current achievements in the production of complex biopharmaceuticals with moss bioreactors.* Bioprocess Biosyst Eng, 2008. **31**(1): p. 3-9.
65. Strasser, R., et al., *Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure.* Plant Biotechnol J, 2008. **6**(4): p. 392-402.
66. Khoudi, H., et al., *Production of a diagnostic monoclonal antibody in perennial alfalfa plants.* Biotechnol Bioeng, 1999. **64**(2): p. 135-43.
67. Cox, K.M., et al., *Glycan optimization of a human monoclonal antibody in the aquatic plant Lemna minor.* Nat Biotechnol, 2006. **24**(12): p. 1591-7.
68. Rademacher, T., et al., *Recombinant antibody 2G12 produced in maize endosperm efficiently neutralizes HIV-1 and contains predominantly single-GlcNAc N-glycans.* Plant Biotechnol J, 2008. **6**(2): p. 189-201.
69. Ramessar, K., et al., *Cost-effective production of a vaginal protein microbicide to prevent HIV transmission.* Proc Natl Acad Sci U S A, 2008. **105**(10): p. 3727-32.
70. Wildt, S. and T.U. Gerngross, *The humanization of N-glycosylation pathways in yeast.* Nat Rev Microbiol, 2005. **3**(2): p. 119-28.

71. Schuster, M., et al., *In vivo glyco-engineered antibody with improved lytic potential produced by an innovative non-mammalian expression system.* Biotechnol J, 2007. **2**(6): p. 700-8.
72. Koprivova, A., et al., *Targeted knockouts of Physcomitrella lacking plant-specific immunogenic N-glycans.* Plant Biotechnol J, 2004. **2**(6): p. 517-23.
73. Huether, C.M., et al., *Glyco-engineering of moss lacking plant-specific sugar residues.* Plant Biol (Stuttg), 2005. **7**(3): p. 292-9.
74. Warnock, D., et al., *In vitro galactosylation of human IgG at 1 kg scale using recombinant galactosyltransferase.* Biotechnol Bioeng, 2005. **92**(7): p. 831-42.
75. Hodoniczky, J., Y.Z. Zheng, and D.C. James, *Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling in vitro.* Biotechnol Prog, 2005. **21**(6): p. 1644-52.
76. Raju, T.S., et al., *Glycoengineering of therapeutic glycoproteins: in vitro galactosylation and sialylation of glycoproteins with terminal N-acetylglucosamine and galactose residues.* Biochemistry, 2001. **40**(30): p. 8868-76.
77. Geyer, H. and R. Geyer, *Strategies for analysis of glycoprotein glycosylation.* Biochim Biophys Acta, 2006. **1764**(12): p. 1853-69.
78. Huhn, C., et al., *IgG glycosylation analysis.* Proteomics, 2009. **9**(4): p. 882-913.
79. Gaza-Bulseco, G., et al., *Effect of the conserved oligosaccharides of recombinant monoclonal antibodies on the separation by protein A and protein G chromatography.* J Chromatogr A, 2009. **1216**(12): p. 2382-7.
80. Wada, Y., et al., *Comparison of the methods for profiling glycoprotein glycans--HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study.* Glycobiology, 2007. **17**(4): p. 411-22.
81. Anumula, K.R., *Advances in fluorescence derivatization methods for high-performance liquid chromatographic analysis of glycoprotein carbohydrates.* Anal Biochem, 2006. **350**(1): p. 1-23.
82. Guile, G.R., et al., *A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles.* Anal Biochem, 1996. **240**(2): p. 210-26.
83. Omtvedt, L.A., et al., *Glycan analysis of monoclonal antibodies secreted in deposition disorders indicates that subsets of plasma cells differentially process IgG glycans.* Arthritis Rheum, 2006. **54**(11): p. 3433-40.
84. Takahashi, N., et al., *Comparative structural study of the N-linked oligosaccharides of human normal and pathological immunoglobulin G.* Biochemistry, 1987. **26**(4): p. 1137-44.
85. Siemiatkoski, J., et al., *A comparison of three techniques for quantitative carbohydrate analysis used in characterization of therapeutic antibodies.* Carbohydr Res, 2006. **341**(3): p. 410-9.
86. Watson, M., et al., *Sugar printing rheumatic diseases: a potential method for disease differentiation using immunoglobulin G oligosaccharides.* Arthritis Rheum, 1999. **42**(8): p. 1682-90.
87. Haslam, S.M., S.J. North, and A. Dell, *Mass spectrometric analysis of N- and O-glycosylation of tissues and cells.* Curr Opin Struct Biol, 2006. **16**(5): p. 584-91.
88. Kuster, B., et al., *Structural determination of N-linked carbohydrates by matrix-assisted laser desorption/ionization-mass spectrometry following enzymatic release*

within sodium dodecyl sulphate-polyacrylamide electrophoresis gels: application to species-specific glycosylation of alpha1-acid glycoprotein. Electrophoresis, 1998. **19**(11): p. 1950-9.
89. Harvey, D.J., et al., *Differentiation between isomeric triantennary N-linked glycans by negative ion tandem mass spectrometry and confirmation of glycans containing galactose attached to the bisecting (beta1-4-GlcNAc) residue in N-glycans from IgG.* Rapid Commun Mass Spectrom, 2008. **22**(7): p. 1047-52.
90. Creaser, C.S., J.C. Reynolds, and D.J. Harvey, *Structural analysis of oligosaccharides by atmospheric pressure matrix-assisted laser desorption/ionisation quadrupole ion trap mass spectrometry.* Rapid Commun Mass Spectrom, 2002. **16**(3): p. 176-84.
91. Weiskopf, A.S., P. Vouros, and D.J. Harvey, *Characterization of oligosaccharide composition and structure by quadrupole ion trap mass spectrometry.* Rapid Commun Mass Spectrom, 1997. **11**(14): p. 1493-504.
92. Wagner-Rousset, E., et al., *The way forward, enhanced characterization of therapeutic antibody glycosylation: comparison of three level mass spectrometry-based strategies.* J Chromatogr B Analyt Technol Biomed Life Sci, 2008. **872**(1-2): p. 23-37.
93. Deguchi, K., et al., *Characteristic pH and electrolyte concentration dependences of 2-aminopyridine-derivatized oligosaccharides (N-glycans) in sonic-spray ionization mass spectrometry.* Rapid Commun Mass Spectrom, 2005. **19**(16): p. 2325-30.
94. Takegawa, Y., et al., *Simultaneous analysis of 2-aminopyridine-derivatized neutral and sialylated oligosaccharides from human serum in the negative-ion mode by sonic spray ionization ion trap mass spectrometry.* Anal Chem, 2005. **77**(7): p. 2097-106.
95. Chen, X. and G.C. Flynn, *Analysis of N-glycans from recombinant immunoglobulin G by on-line reversed-phase high-performance liquid chromatography/mass spectrometry.* Anal Biochem, 2007. **370**(2): p. 147-61.
96. Ruhaak, L.R., et al., *Hydrophilic interaction chromatography-based high-throughput sample preparation method for N-glycan analysis from total human plasma glycoproteins.* Anal Chem, 2008. **80**(15): p. 6119-26.
97. Pabst, M. and F. Altmann, *Influence of electrosorption, solvent, temperature, and ion polarity on the performance of LC-ESI-MS using graphitic carbon for acidic oligosaccharides.* Anal Chem, 2008. **80**(19): p. 7534-42.
98. Pabst, M., et al., *Mass + retention time = structure: a strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans.* Anal Chem, 2007. **79**(13): p. 5051-7.
99. Itoh, S., et al., *N-linked oligosaccharide analysis of rat brain Thy-1 by liquid chromatography with graphitized carbon column/ion trap-Fourier transform ion cyclotron resonance mass spectrometry in positive and negative ion modes.* J Chromatogr A, 2006. **1103**(2): p. 296-306.
100. Karlsson, N.G., et al., *Negative ion graphitised carbon nano-liquid chromatography/mass spectrometry increases sensitivity for glycoprotein oligosaccharide analysis.* Rapid Commun Mass Spectrom, 2004. **18**(19): p. 2282-92.
101. Bardor, M., et al., *N-glycosylation of plant recombinant pharmaceuticals.* Methods Mol Biol, 2009. **483**: p. 239-64.
102. De Muynck, B., et al., *Different subcellular localization and glycosylation for a functional antibody expressed in Nicotiana tabacum plants and suspension cells.* Transgenic Res, 2009. **18**(3): p. 467-82.

103. Van Droogenbroeck, B., et al., *Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic Arabidopsis seeds.* Proc Natl Acad Sci U S A, 2007. **104**(4): p. 1430-5.
104. Dick, L.W., Jr., et al., *C-terminal lysine variants in fully human monoclonal antibodies: investigation of test methods and possible causes.* Biotechnol Bioeng, 2008. **100**(6): p. 1132-43.

APPENDIX

Curriculum Vitae - Johannes STADLMANN

Date of birth: June 26th, 1978	Citizenship: Austria

Home Address: Wiener Straße 41 A-2514 Traiskrichen AUSTRIA / EU	Term Address: Urban-Loritz-Platz 4/8 A-1070 Vienna AUSTRIA / EU
E-mail: johannes.stadlmann@gmail.com mobile: +43-(0)699-19245931	

EDUCATION:

Nov. 2004 – Jul. 2009:	Univ. of Natural Resources and Applied Life Sciences, Austria. PhD Thesis: "Applications of a Mass-spectrometric Method for the N-glycosylation Analysis of Immunoglobulin G molecules"
28-30 Aug. 2008	Novartis Biotech Leadership Camp 2008, Basel, Switzerland.
12-18 Aug. 2007	1st EU Summer School in Proteomic Basics (HUPO/ EuPA), Italy.
18-21 Jun. 2007	Tutorial for tutors: Bioinformatics for MS analysis (EuPA/ SIB), Geneva, Switzerland.
Oct. 1996 – Oct. 2004:	Univ. of Natural Resources and Applied Life Sciences, Austria. Diploma Thesis: "The Molecular and Biochemical Characterisation of the *A. thaliana* N-acetlyglucosaminyltransferase I mutant cgl"
Jul. 2001- Aug. 2001	Oeresund Summer University, Copenhagen, DK and Lund, Sweden.
Feb. 2001- May 2001	University of Nottingham, Nottingham, U.K. Research-project at Dept. of Nutritional Biochemistry: 'The Biochemical Analysis of the Tomato-Ripening Mutant T38'
Oct. 2000 – Jan. 2001:	University of Nottingham, Nottingham, U.K. ERASMUS-Exchange Program.
Sept. 1986 – Jun. 1996:	Gymnasium d. Salesianer Don Boscos, Unterwaltersdorf, Austria.

LANGUAGES SKILLS:

German: mother tongue	French: fluently spoken, well written
English: fluently spoken and written	Spanish: basic knowledge

PROFICIENT COMPUTER SKILLS:

MS Office Professional; OpenOffice; Windows 2000, NT and XP; Linux (SuSE, RedHat); Mac OS X; HTML-Programming; Visual Basic; Java; various scientific software applications (MassLynx, Dynamo, mMass, Chromeleon, X!tandem etc.)

PROFESSIONAL EXPERIENCE:

Summer 2002	Novartis, Kundl, Tirol. - Environmental and working-environmental analytics
March to September 2003	Zentrum f. Angewandte Genetik, Wien. - Plant tissue-culture
Summer 2000	Q-Lab Austria, Wien. - Qualitative analytics of food-stuff
Summer 1999	Zuckerforschung Tulln, Tulln, NÖ. - Qualitative analysis of epoxid-residues in cationic starches and the evaluation of the respective analytical methdos employed

TEACHING ACTIVITES:

Since Jan. 2009	Lecturer “Biochemie Methoden” FH Campus Wien, Vienna, Austria.
Oct. 2008	“Protein Sequence Data Bases & MS-Data Search Engines”. AuPA Seminar, Technical University Vienna, Austria.
Oct. 2006 – Jun. 2009	Lector “Practical Biochemistry II - Proteomics” Univ. of Natural Resources and Applied Life Sciences, Vienna, Austria.
Oct. 2005 – Jun. 2009	Tutor “Practical Protemics” Univ. of Natural Resources and Applied Life Sciences, Vienna, Austria.
Oct. 2004 – Jun. 2009	Tutor “Practical Biochemistry I” Univ. of Natural Resources and Applied Life Sciences, Vienna, Austria.

Comprehensive List of Publications

1: Strasser R, Castilho A, Stadlmann J, Kunert R, Quendler H, Gattinger P, Rademacher T, Altmann F, Mach L, Steinkellner H. *Improved virus neutralization by plant-produced anti-HIV antibodies with a homogeneous beta1,4-galactosylated N-glycan profile*. JBC 2009; in press.

2: Drake PM, Barbi T, Sexton A, McGowan E, Stadlmann J, Navarre C, Paul MJ, Ma JK. *Development of rhizosecretion as a production system for recombinant proteins from hydroponic cultivated tobacco*. FASEB J. 2009 May 26. [Epub ahead of print]

3: Leitsch D, Kolarich D, Binder M, Stadlmann J, Altmann F, Duchêne M. *Trichomonas vaginalis: metronidazole and other nitroimidazole drugs are reduced by the flavin enzyme thioredoxin reductase and disrupt the cellular redox system. Implications for nitroimidazole toxicity and resistance*. Mol Microbiol. 2009 Apr;72(2):518-36.

4: Bortesi L, Rossato M, Schuster F, Raven N, Stadlmann J, Avesani L, Falorni A, Bazzoni F, Bock R, Schillberg S, Pezzotti M. *Viral and murine interleukin-10 are correctly processed and retain their biological activity when produced in tobacco*. BMC Biotechnol. 2009 Mar 19;9:22.

5: Dragosits M, Stadlmann J, Albiol J, Baumann K, Maurer M, Gasser B, Sauer M, Altmann F, Ferrer P, Mattanovich D. *The Effect of Temperature on the Proteome of Recombinant Pichia pastoris*. J Proteome Res. 2009 Feb 13. [Epub ahead of print]

6: Bardor M, Cabrera G, Stadlmann J, Lerouge P, Cremata JA, Gomord V, Fitchette AC. *N-glycosylation of plant recombinant pharmaceuticals*. Methods Mol Biol. 2009;483:239-64.

7: De Muynck B, Navarre C, Nizet Y, Stadlmann J, Boutry M. *Different subcellular localization and glycosylation for a functional antibody expressed in Nicotiana tabacum plants and suspension cells*. Transgenic Res. 2009 Jun;18(3):467-82. Epub 2009 Jan 14.

8: Zamocky M, Furtmüller PG, Bellei M, Battistuzzi G, Stadlmann J, Vlasits J, Obinger C. *Intracellular catalase/peroxidase from the phytopathogenic rice blast fungus Magnaporthe grisea: expression analysis and biochemical characterization of the recombinant protein*. Biochem J. 2009 Mar 1;418(2):443-51.

9: Schoberer J, Vavra U, Stadlmann J, Hawes C, Mach L, Steinkellner H, Strasser R. *Arginine/lysine residues in the cytoplasmic tail promote ER export of plant glycosylation enzymes*. Traffic. 2009 Jan;10(1):101-15. Epub 2008 Nov 1.

10: Stadlmann J, Pabst M, Kolarich D, Kunert R, Altmann F. *Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides*. Proteomics. 2008 Jul;8(14):2858-71.

11: Schähs P, Weidinger P, Probst OC, Svoboda B, Stadlmann J, Beug H, Waerner T, Mach L. *Cellular repressor of E1A-stimulated genes is a bona fide lysosomal protein which undergoes proteolytic maturation during its biosynthesis*. Exp Cell Res. 2008 Oct 1;314(16):3036-47. Epub 2008 Jun 28.

12: Strasser R, Stadlmann J, Schähs M, Stiegler G, Quendler H, Mach L, Glössl J, Weterings K, Pabst M, Steinkellner H. *Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure*. Plant Biotechnol J. 2008 May;6(4):392-402. Epub 2008 Mar 13.

13: Ramessar K, Rademacher T, Sack M, Stadlmann J, Platis D, Stiegler G, Labrou N, Altmann F, Ma J, Stöger E, Capell T, Christou P. *Cost-effective production of a vaginal protein microbicide to prevent HIV transmission*. Proc Natl Acad Sci U S A. 2008 Mar 11;105(10):3727-32. Epub 2008 Mar 3.

14: Floss DM, Sack M, Stadlmann J, Rademacher T, Scheller J, Stöger E, Fischer R, Conrad U. *Biochemical and functional characterization of anti-HIV antibody-ELP fusion proteins from transgenic plants*. Plant Biotechnol J. 2008 May;6(4):379-91. Epub 2008 Feb 28. Erratum in: Plant Biotechnol J. 2008 May;6(4):424.

15: Kolarich D, Weber A, Pabst M, Stadlmann J, Teschner W, Ehrlich H, Schwarz HP, Altmann F. *Glycoproteomic characterization of butyrylcholinesterase from human plasma*. Proteomics. 2008 Jan;8(2):254-63.

16: Rademacher T, Sack M, Arcalis E, Stadlmann J, Balzer S, Altmann F, Quendler H, Stiegler G, Kunert R, Fischer R, Stoger E. *Recombinant antibody 2G12 produced in maize endosperm efficiently neutralizes HIV-1 and contains predominantly single-GlcNAc N-glycans*. Plant Biotechnol J. 2008 Feb;6(2):189-201. Epub 2007 Nov 3.

17: Jin C, Hantusch B, Hemmer W, Stadlmann J, Altmann F. *Affinity of IgE and IgG against cross-reactive carbohydrate determinants on plant and insect glycoproteins*. J Allergy Clin Immunol. 2008 Jan;121(1):185-190.e2. Epub 2007 Sep 19.

18: Schähs M, Strasser R, Stadlmann J, Kunert R, Rademacher T, Steinkellner H. *Production of a monoclonal antibody in plants with a humanized N-glycosylation pattern*. Plant Biotechnol J. 2007 Sep;5(5):657-63.

19: Strasser R, Bondili JS, Vavra U, Schoberer J, Svoboda B, Glössl J, Léonard R, Stadlmann J, Altmann F, Steinkellner H, Mach L. *A unique beta1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in Arabidopsis thaliana*. Plant Cell. 2007 Jul;19(7):2278-92. Epub 2007 Jul 13.

20: Asuthkar S, Velineni S, Stadlmann J, Altmann F, Sritharan M. *Expression and characterization of an iron-regulated hemin-binding protein, HbpA, from Leptospira interrogans serovar Lai*. Infect Immun. 2007 Sep;75(9):4582-91. Epub 2007 Jun 18.

21: Pabst M, Bondili JS, Stadlmann J, Mach L, Altmann F. *Mass + retention time = structure: a strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans*. Anal Chem. 2007 Jul 1;79(13):5051-7. Epub 2007 Jun 1.

22: Schuster M, Jost W, Mudde GC, Wiederkum S, Schwager C, Janzek E, Altmann F, Stadlmann J, Stemmer C, Gorr G. *In vivo glyco-engineered antibody with improved lytic potential produced by an innovative non-mammalian expression system*. Biotechnol J. 2007 Jun;2(6):700-8.

23: Van Droogenbroeck B, Cao J, Stadlmann J, Altmann F, Colanesi S, Hillmer S, Robinson DG, Van Lerberge E, Terryn N, Van Montagu M, Liang M, Depicker A, De Jaeger G. *Aberrant localization and underglycosylation of highly accumulating single-chain Fv-Fc antibodies in transgenic Arabidopsis seeds*. Proc Natl Acad Sci U S A. 2007 Jan 23;104(4):1430-5. Epub 2007 Jan 16.

24: Strasser R, Stadlmann J, Svoboda B, Altmann F, Glössl J, Mach L. *Molecular basis of N-acetylglucosaminyltransferase I deficiency in Arabidopsis thaliana plants lacking complex N-glycans*. Biochem J. 2005 Apr 15;387(Pt 2):385-91.