Summary

The aim of the present work was the establishment of suitable methods for effective isolation and detailed structural characterisation of arabinogalactan proteins (AGP), concerning the respective carbohydrate and protein part, by advanced methods. Special focus was laid on the structural features of the connection between sugar and peptide parts of the polymers. This was performed exemplarily for an physiological active AGP from Jatropha curcas L. seeds.

Within optimisation of the isolation protocol, raw polysaccharides were extracted from J. curcas endosperm, purified by heat-depletion from co-extracted protein, fractionated by ion-exchange chromatography and finally, pure AGP was precipitated in good yields with β-D-glucosyl Yariv reagent, synthesised by an effective 2-step procedure. From this protocol highly purified AGP, named Y2, was obtained (yield 0.01 %, related to the whole seeds). HP-SEC analysis indicated the presence of one single polymer with a mean MW of 108 kDa.

Y2 was composed of Gal, Ara, GluA, Rha, Man and GlcNAc (48.2, 39.3, 5.3, 4.1, 2.2, 0.9 mol %); Fuc and Xyl were found in traces. Different methods of quantification were compared and optimised/validated concerning their potential use for AGP characterisation. As determined by methylation analysis of native and carboxyl-reduced (deuterated) Y2, Gal was present as β-(1,3)-, β-(1,6)- and β-(1,3,6)-linked D-galactopyranosyl (D-Galp) (10.7, 2.3 and 35.2 mol %), Ara as α-(1,5)-linked and terminal L-arabinofuranosyl (L-Araf) residues (10.0 and 29.3 mol %). Rha occurred only as terminal α-L-Rhap, whereas GluA was mainly present as β-(1,4)-linked D-GlupA and in traces as terminal β-D-GlupA (4.5 and 0.8 mol %). Man was identified in 1,2-linkage. The α- and β-configuration was determined by NMR spectroscopy, the D- and L-enantiomers were distinguished after derivatisation and separation by CZE.

Enzymatic cleavage by exo-α-L-arabinofuranosidase released all Araf residues with an intense increase of terminal Galp and β-(1,6)-Galp, and a slightly growth of β-(1,3)-Gal, while the amount of β-(1,3,6)-Galp was diminished accordingly. This indicated that all α-(1,5)-Araf residues are decorated by terminal Araf. The existence of many short side chains can be deduced, connected via β-(1→6) to a β-(1→3)-Galp backbone, terminated by Araf residues. These side chains, as well as the backbone, are again decorated by terminal Araf.

Enzymatic digest by endo-arabinase affirmed this result: α-(1,5)-Araf was removed nearly completely, while terminal Araf was dimished only by 42 %. The endo-arabinanase cleavage resulted in the same amount of 1,6-Galp, compared to the exo-digest. This indicates, that nearly all α-(1→5)-Araf branches are located at the (1→6)-Galp prolonged side chains.

The β-(1→3)-galactose backbone was degraded by an exo-β-(1→3)-galactanase, semi-purified from Driselase®, releasing the (1→6)-galactose side chains. These oligosaccharides were analysed by

nanoESI-Q-TOF-MS/MS. MS-sequencing of neutral and acidic oligosaccharides revealed typical fragmentation patterns of dimeric to nonameric oligosaccharides, and gave clear structural features for the spacer systems, linking the carbohydrate part to the peptide moiety.

Combining all gained data, a structure proposal was assumed: the β-(1→3)-D-Galp backbone is branched via position C-6. The backbone is decorated with few α-L-Araf monomers. The β-(1→6)-D-Galp side chains consist of not more than six Galp residues. Most of the arabinose is located within the side chains, terminating the β-(1→6)-D-galactose chains, or being linked in position C-3 to Galp of the side chains. Some of the α-L-Araf residues are prolonged by further α-L-Araf, connected by α-(1→5)-Araf linkage. The α-L-Rhap is terminally linked to position C-4 of β-D-GlcpA, which are connected to the β-(1→6)-D-Galp side chains, as well. Some very few β-D-GlcpA are terminal residues. These β-D-GlcpA carrying Gal chains, are not longer than two units. A typical N-glycan linker of the complex type was identified, with $GlcNAc_4Man_3Fuc_1Xyl_1$.

The protein moiety of Y2 accounted for 4.8 % (w/w), with Ala (17.7 mol %), Ser (12.9 mol %), Gly (11.5 mol %), Hyp (9.2 mol %) and Glu (9.0 mol %) as the most abundant amino acids. Lys, Thr, Leu, Asp, Val, Pro and Ile occurred in lower amounts.

After deglycosylation by TFMS and alkaline β-elimination, the presence of O-glycans were identified mostly in linkage to Hyp and in small concentrations to Thr respectively Ser. PNGase F was not able to cleave, in contrast to N-glycosidase A, which confirms the presence of additional N-glycans from the complex type, carrying at the innermost GlcNAc a terminal α-Fuc at position C-3.

After tryptic digest of Y2, the peptides were analysed by Q-TOF Premier nanoAcquity LC. Strong homologies to three proteins were found, namely to fasciclin, xylogen and LysM. Fasciclin and xylogen were described as arabinogalactan proteins before, but LysM is described for the first time as part of an AGP.

The data described in the present work are the base for a quick and efficient way for isolation and analysis of AGPs. Carbohydrate and peptide characterisation can be much faster by using the described protocols for MS-fragmentation and methods for deglycosylation and enzymatic degradation.