Pure & Appl. Chem., Vol. 65, No. 6, pp. 1223–1232, 1993.
Printed in Great Britain.

Glycopeptides of biological interest: A challenge for chemical synthesis

Horst Kunz

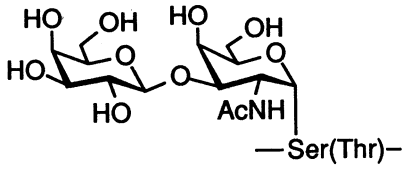
Institut für Organische Chemie, Universität Mainz, Becher-Weg 18, D-6500 Mainz, Germany

Abstract. The synthesis of glycopeptides which constitute partial structures of glycoproteins, e. g. tumor-associated antigens or virus envelope glycoproteins, is presented. The problems of the selective deblocking and chain extension of such glycopeptides which contain acetalic glycosidic bonds were solved by application of the fluorenylmethoxycarbonyl (Fmoc) group cleavable with the weak base morpholine and by the introduction of the allylic protecting groups which can be removed under almost neutral conditions via palladium(0)-catalyzed allyl transfer to weakly basic nucleophiles. The allylic protection was extended to a new allylic anchoring principle in the solid phase synthesis. It allows to detach the synthesized peptides and glycopeptides from the polymer under neutral conditions. The synthesis of glycopeptides with branched and fucose containing saccharide side chains demands an even more chemoselective strategy which includes the exchange of protecting groups within the saccharide portion. In these syntheses, the azido group was used as the protection and precursor of the anomeric amino function required for the construction of the N-glycosyl asparagine linkage. An important indirect protecting effect is exhibited by the O-acetyl protecting groups within the saccharide part which strongly stabilize the glycosidic bonds. Using these methods, glycopeptides which contain tumor-associated T_N and T antigen or Lewis antigen-type side chains and represent partial sequences of virus envelope proteins were synthesized and linked to carrier proteins to give synthetic glycoprotein antigens.

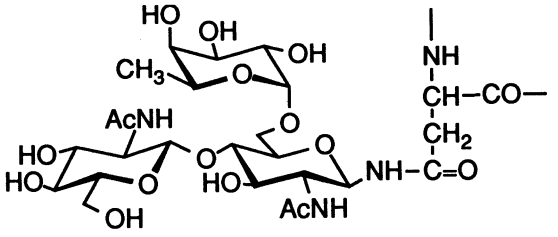
BIOLOGICAL SELECTIVITY WRITTEN IN CARBOHYDRATES

During the past decades it has been recognized that most of the natural proteins of mammalians carry covalently linked saccharide side chains. The carboyhydrate portions of the glycoproteins as well as those of glycolipids obviously play key roles in biological recognition processes such as the selective transport of serum components through membranes, the adaption of enzymes or toxines to membranes, the docking of bacteria and viruses to cells and the cell-cell communication (ref. 1). Carbohydrates of membrane glycoproteins constitute recognition signals in the regulation of the cell growth (ref. 2). It is logical, that normal cells and tumor cells have been found to be quite different in the glycoprotein profiles of their cell membranes (ref. 3). Altered glycoproteins of the tumor membranes have been reported to be tumor-associated antigens, e. g. the Thomsen-Friedenreich antigen glycoproteins (T antigen) **1** (ref. 4, 5). These antigen components belong to the class of the O-glycoproteins. In O-glycoproteins the carbohydrate parts are O-glycosidically linked to hydroxy amino acids. Among the O-glycoproteins the mucin-type subclass, e. g. **1**, is characterized by the α-O-glycosidic bond between N-acetylgalactosamine and serine or threonine.

Fucosylation of glycoproteins also plays a crucial role in tumor development (ref. 6). In this context, N-glycoproteins fucosylated in the core region (**2**) have been found to be tumor-associated antigens (ref. 7).



1



2

The N-glycoproteins, which are most frequently found in nature, typically contain the N-glycosidic linkage between N-acetylglucosamine and the amide function of asparagine. Fucosylated core region structures **2** also occur in virus coat N-glycoproteins (ref. 8).

Owing to the decisive functions exhibited by carbohydrates of glycoproteins in regulatory processes and in the organized distribution of these biomacromolecules within multicellular organisms, the availability of these glycoconjugate structures is receiving increasing interest. As the isolation of glycoproteins from biological sources is often difficult and, furthermore, suffers from the biological microheterogeneity and because the gene-technological production of these compounds is still limited, the chemical synthesis of partial sequences of glycoproteins bearing one or many glycan side chains is demanded for biochemical, immunological, cell-biological and medicinal research.

PROBLEMS OF THE SYNTHESIS OF GLYCOPEPTIDES

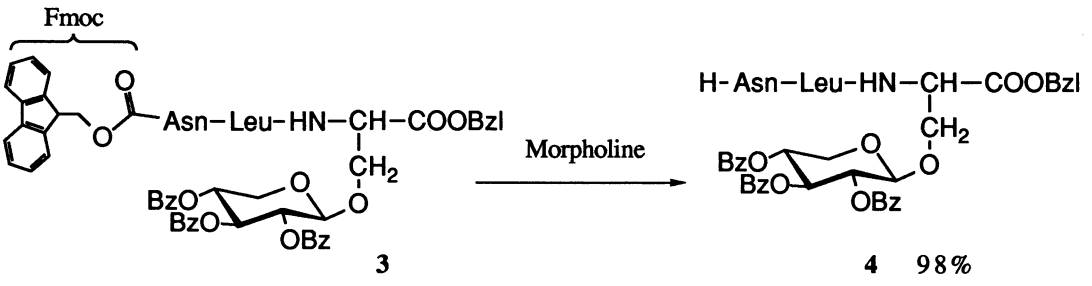
The construction of glycopeptides of exactly specified structure requires the combined application of synthetic methods of both carbohydrate and peptide chemistry (ref. 9). It begins with formation of the N- or O-glycosidic linkages between the saccharide donors and the relevant amino acids (ref. 10). These glycosidic bonds are of acetalic nature and, therefore, chemically sensitive. They can be cleaved or anomerized by acids. In addition, the O-glycosyl serine and threonine derivatives show a characteristic base-sensitivity. Moderate bases decompose these compounds via a base-catalyzed β-elimination of the entire carbohydrate (ref. 11). As a consequence, throughout the synthesis of glycopeptides, special attention has to be paid to mild methods for the removal of protecting groups (ref. 9). Furthermore, multi-step oligosaccharide synthesis (ref. 12, 13) is a precondition for the construction of glycopeptides which contain more complex saccharide portions. Since the sensitivity of particular glycosidic and intersaccharidic bonds (e. g. the fucoside bonds) of these complex glycopeptides imposes new limitations on the applicable methodology, the demands for the selectivity of the used protecting techniques arrive at an even higher level (ref. 14).

THE (9-FLUORENYL)METHOXYCARBONYL(FMOC) GROUP IN THE SYNTHESIS OF T ANTIGEN GLYCOPEPTIDES

In most syntheses of glycosylated amino acids and peptides reported up to the early eighties (ref. 11), the N-terminal benzyloxycarbonyl (Z) group was applied in combination with the benzyl ester as the protection within the peptide part. As both groups are simultaneously removed by hydrogenation, the selective deblocking at the N- or C-terminus of glycopeptides remained an unsolved problem. A reliable selective deprotection either of the carboxy or of the amino functions of glycopeptides was achieved by reductive cleavage of 2-halo-ethyl esters (ref. 15) or by the very easy base-catalyzed removal of 2-phosphonio-ethoxycarbonyl (Peoc) group from N-protected compounds (ref. 16).

The (9-fluorenyl)methoxycarbonyl (Fmoc) group (ref. 17), is frequently applied in peptide chemistry and commonly cleaved off with piperidine. It revealed to be likewise useful in the synthesis of glycopeptides, also of those containing base-sensitive glycosyl serine or threonine linkages, if its removal is carried out with the weak base morpholine (ref. 18). In this way, the N-terminal deblocking of the O-xylosyl serine tripeptide **3** representing a typical linkage structure of proteoglycans proceeded selectively and quantitatively. The numerous other blocking groups of the product **4** and the β-O-xylosyl serine bond sensitive to both base-catalyzed β–elimination and acid-catalyzed anomerization remained unaffected.

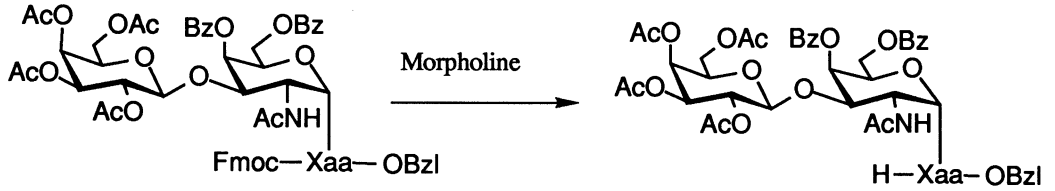
Scheme 1



3 **4** 98%

According to this procedure, the Fmoc group was also applied in the synthesis of monosaccharidic T_N antigen and disaccharidic T antigen glycopeptides (ref. 19). Treatment of the Fmoc-protected T antigen serine/threonine conjugates **5** with morpholine (or morpholine/dichloromethane 1:1) resulted in the completely selective deblocking of the amino function.

Scheme 2



5a: Xaa = Ser

5b: Xaa = Thr

6a: Xaa = Ser, > 95%

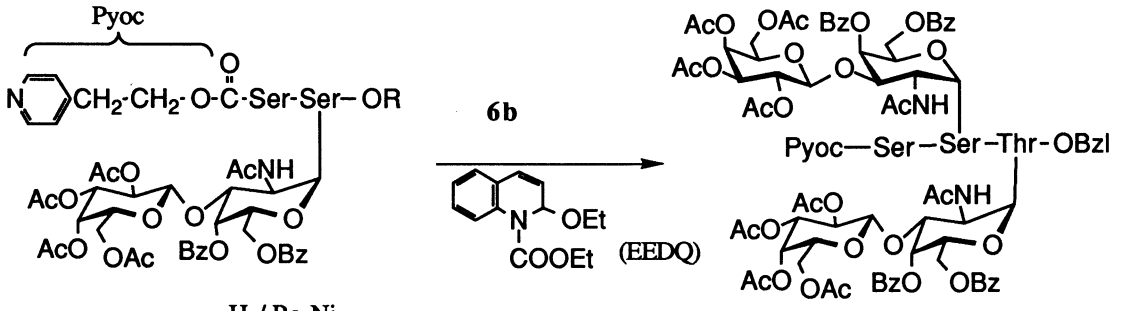
6b: Xaa = Thr, > 95%

The chain extension can then be carried out by analogous application of Fmoc-protected units (ref. 18, 20, 21). The use of the Fmoc group in this sense now provides a reliable method for the synthesis of glycopeptides. It can be applied almost routinely and has been adopted for the solid phase synthesis of glycopeptides including the sensitive O-glycopeptides (ref. 22, 23, 24).

Since the amino deblocked O-glycosyl serine/threonine derivatives, e. g. **6**, are base-sensitive bases and, therefore, unstable, the subsequent chain lengthening reactions should proceed with a sufficient rate. For this purpose, we introduced the polar 2-(pyridyl)ethoxycarbonyl (Pyoc) group (ref. 25).

The Pyoc group is quite stable to acids and bases usually applied in peptide chemistry. Its additional stability towards hydrogenolysis allowed the selective debenzylation of the glycopeptide **7**. The formed carboxy deblocked product **8** was condensed with **6b** unblocked at the amino group to yield the twofold glycosylated glycotripeptide **9** (Scheme 3).

Scheme 3



H_2/ Ra-Ni

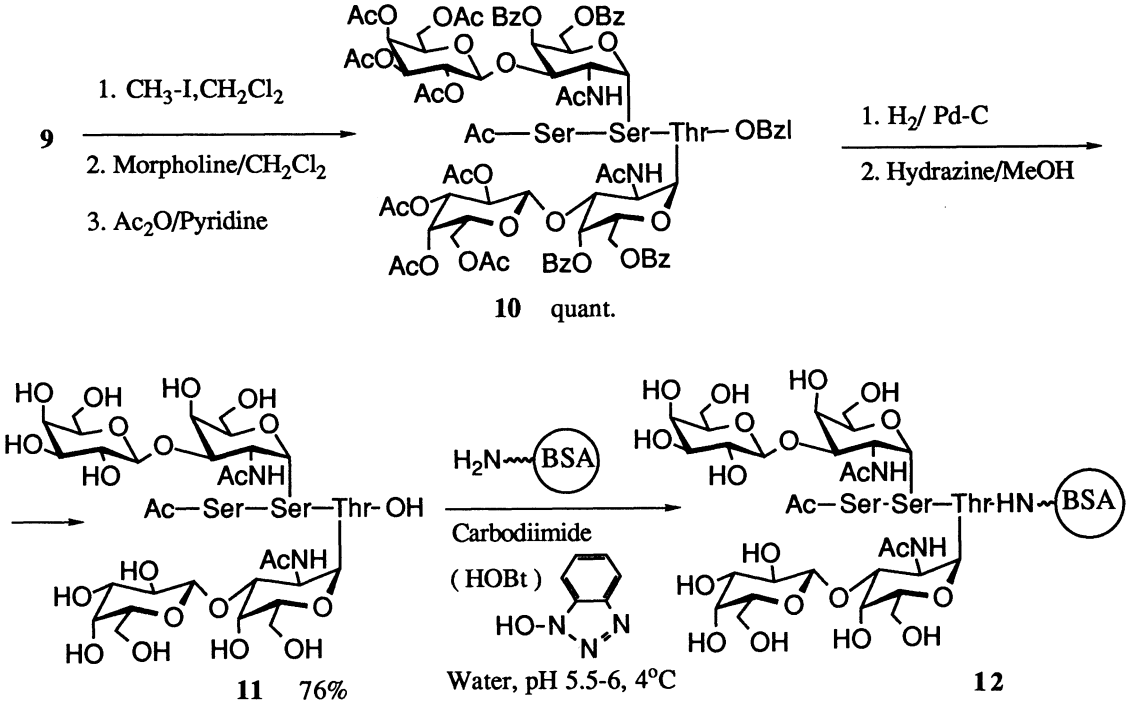
7: R = Bzl ⟶ **8**: R = H, 90%

9 44%

In spite of its stablity the Pyoc group can selectively be removed from the complex sensitive glycopeptide **9**, when it is transformed to the CH-acidic pyridinium form which is cleavable by morpholine/dichloromethane via an E1cB pathway. The resulting product was N-acetyled to give **10** which was subjected to hydrogenolytic debenzylation. Removal of the ester protecting groups from the carbohydrate part delivered the free T antigen glycopeptide **11**, the N-terminal sequence of human glycophorin with M blood group specifity (ref. 19).

Using a modified carbodiimide procedure, the T antigen glycopeptide **11** was coupled to bovine serum albumin (BSA) to furnish the synthetic glycoprotein antigen **12**.

Scheme 4

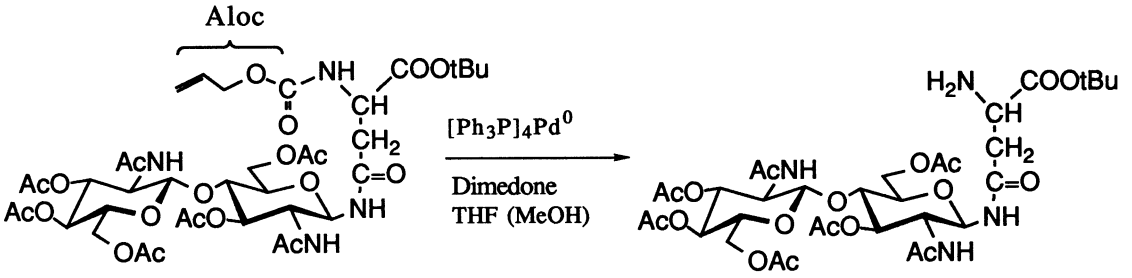


According to the carbohydrate analysis, the glycoconjugate **12** on average contained 38 synthetic glycopeptide units **11** per macromolecule (ref. 19). It was supplied for immunological investigations of this T antigen structure.

ALLYLIC PROTECTING GROUPS AND THEIR PALLADIUM(0)-CATALYZED REMOVAL

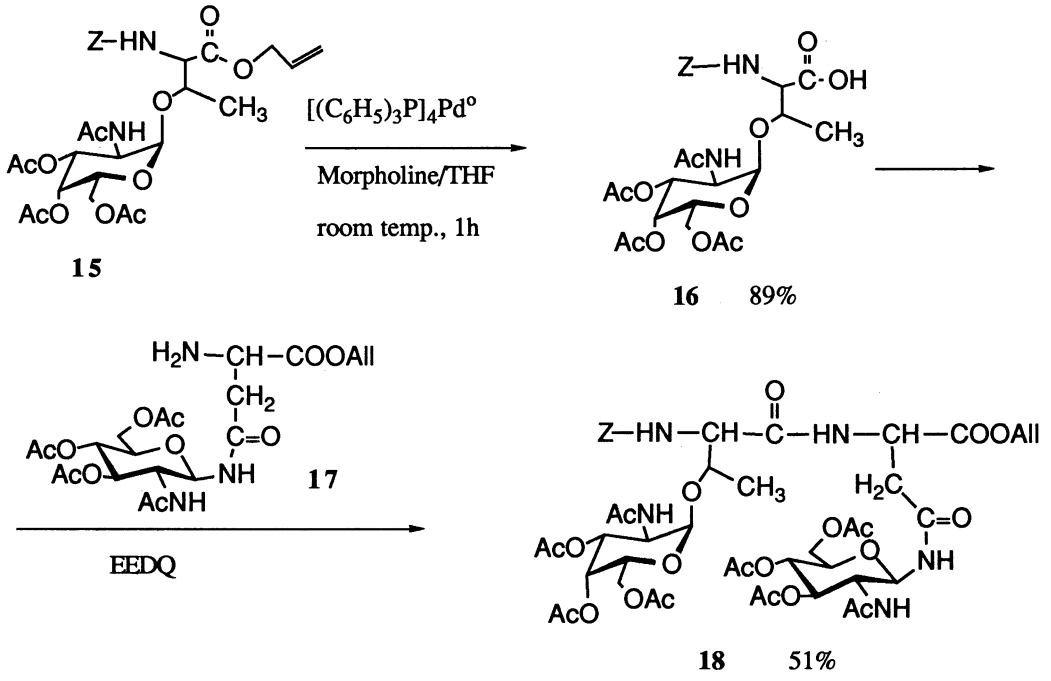
The allyl ester (ref. 26) and the allyloxycarbonyl (Aloc) group (ref. 27) constitute particularly valuable protecting groups for the synthesis of glycopeptides. They are stable to acids and bases commonly applied in peptide synthesis. However, they can be selectively removed under practically neutral conditions by application of Pd(0)-catalyzed allyl transfer to weak bases, e. g. morpholine, or neutral nucleophiles like CH-acidic compounds, preferentially N,N´-dimethyl barbituric acid (ref. 28). These nucleophiles prevent the "self-trapping" of the allyl moiety by nucleophilic functions of the glycopeptide, e. g. the liberated amino group. Treatment of the allyloxycarbonyl protected N-chitobiosyl asparagine tert-butyl ester **13**, for example, with catalytic amounts of tetrakis-triphenylphosphine palladium(0) in tetrahydrofuran in the prescence of dimedone resulted in the completely selective cleavage of the Aloc group to yield **14**. As the other protecting groups and the O- and N-glycosidic bonds are untouched, **14** can be subjected to N-terminal chain extension (ref. 29).

Scheme 5



Analogously, the allyl ester protection is selectively removed from the α-O-glycosyl threonine derivate **15** carrying a T_N antigen side chain. The Pd(0)-catalyzed allyl transfer to morpholine in tetrahydrofuran at room temperature gave the carboxy deblocked product **16**, which was condensed with the N-terminally deblocked N-glycosyl asparagine allyl ester **17** to furnish the complex partial structure **18** of human glycophorin, in which an α-O-glycosyl threonine is immediately linked to a β-N-glycosyl asparagine (ref. 30).

Scheme 6

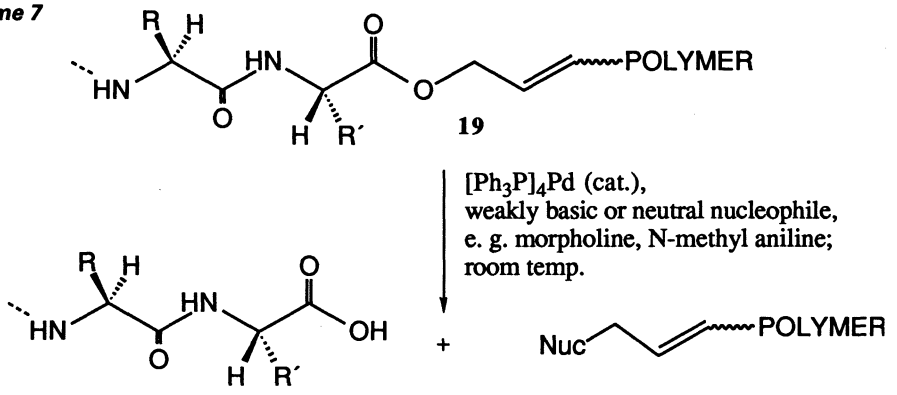


The efficiency of the allylic protecting groups has been demonstrated in syntheses of numerous N- and O-glycopeptides, e. g. partial structures of proteoglycans, of transmembrane neuraminidases from two different strains of influenza virus, other partial sequences of human glycophorin and segments of human serum glycoproteins (ref. 9, 30).

THE ALLYLIC ANCHORING PRINCIPLE IN SOLID PHASE SYNTHESIS

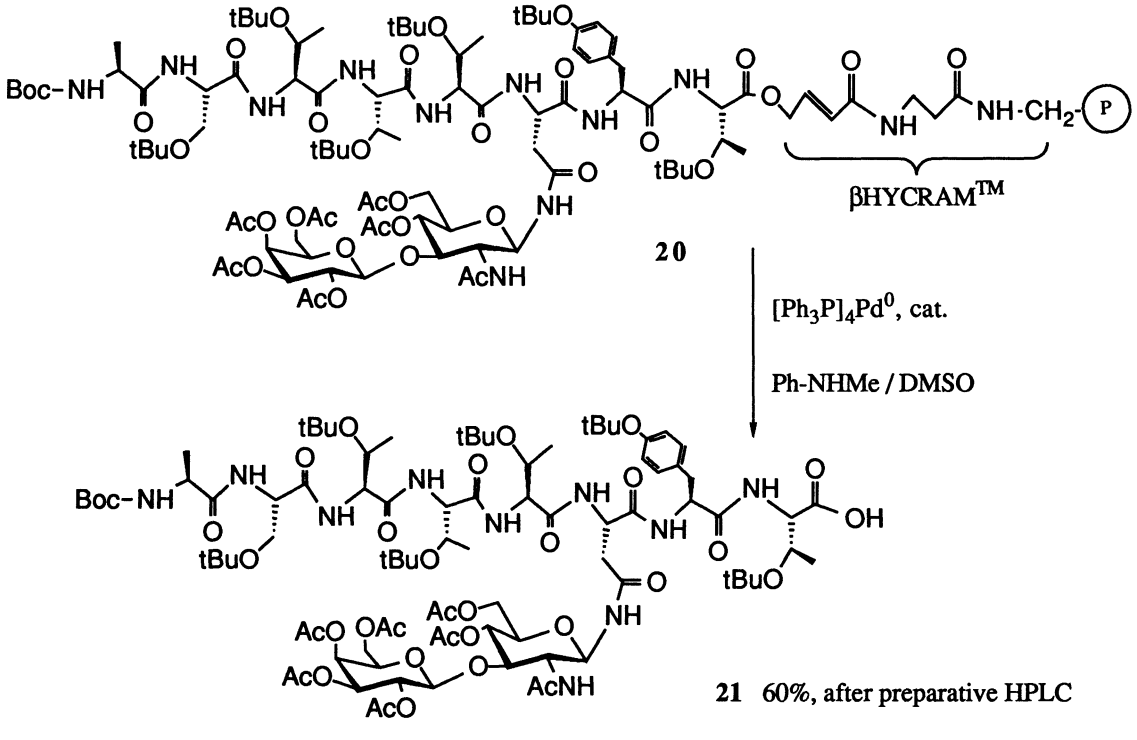
The advantages of the allyl ester being a stable carboxy protection which can be cleaved mildly and selectively were exploited for the development of a new allylic anchoring principle **19** for the solid phase synthesis of peptides and glycopeptides (ref. 31).

Scheme 7



Depending upon the character of the allyl trapping nucleophile, the medium for the releasing of the synthesized peptides from the polymeric support can be adjusted to more or less neutral conditions. N-Glycopeptides (ref. 32) and O-glycopeptides (ref. 24) can be detached by this method without destroying the glycosidic linkages. As the allylic anchor **19** is stable to acidic and basic conditions, both the Boc and the Fmoc group are applicable as the temporary amino protection in these solid phase syntheses. One structural form of the anchor **19** consists in the amino acid esters of 4-hydroxycrotonoyl aminomethyl-substituted polymeric supports (Scheme 8, HYCRAM^{TM}. HYCRAM is a Trade Mark of the Orpegen GmbH, Heidelberg, Germany).

Scheme 8



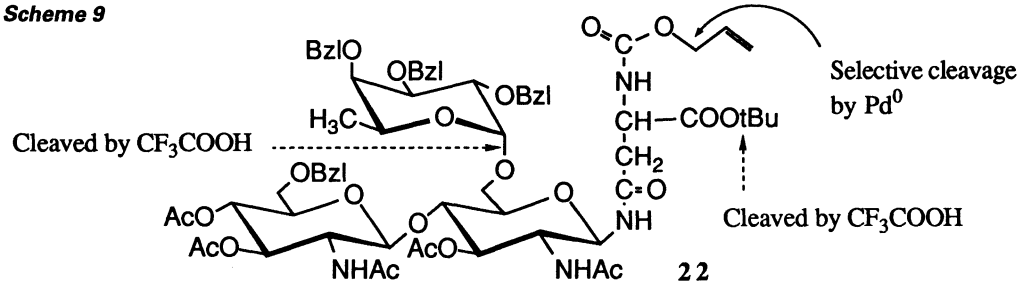
Using the improved HYCRAM version with an β-alanine insert (βHYCRAM), the polymer-linked glyco-octapeptide **20** with the lacto-lactosamine N-glycosidically bound to asparagine was synthesized by application of the Fmoc group as the temporary amino protection and a typical solid phase synthesis protocol. The release of the glyco-octapeptide **21** from the resin-linked form **20** was performed upon Pd(0)-catalyzed allyl transfer reaction to N-methyl aniline as the trapping nucleophile. The reaction is favorably carried out in DMSO/dichloromethane mixtures and yielded **21**. The numerous other protecting groups as well as the O-and

N-glycosidic bonds remained unaffected (Scheme 8). Glycopeptides selectively deblocked in only one function can be purified much easier than the completely deblocked ones. Thus, **21** was isolated in pure form after preparative HPLC in gram-scale in a yield of 60 % (ref. 33). It represents the N-glycosylated peptide T, a partial sequence of gp 120 of HIV-1, which has been reported to be the signal sequence for the binding of the HIV-1 to the CD_4 receptor of the lymphocytes (ref. 34). The efficiency of the allylic anchoring principle has also been documented in large-scale syntheses of pharmaceutical peptides (ref. 35).

N-GLYCOPEPTIDES FUCOSYLATED IN THE CORE REGION

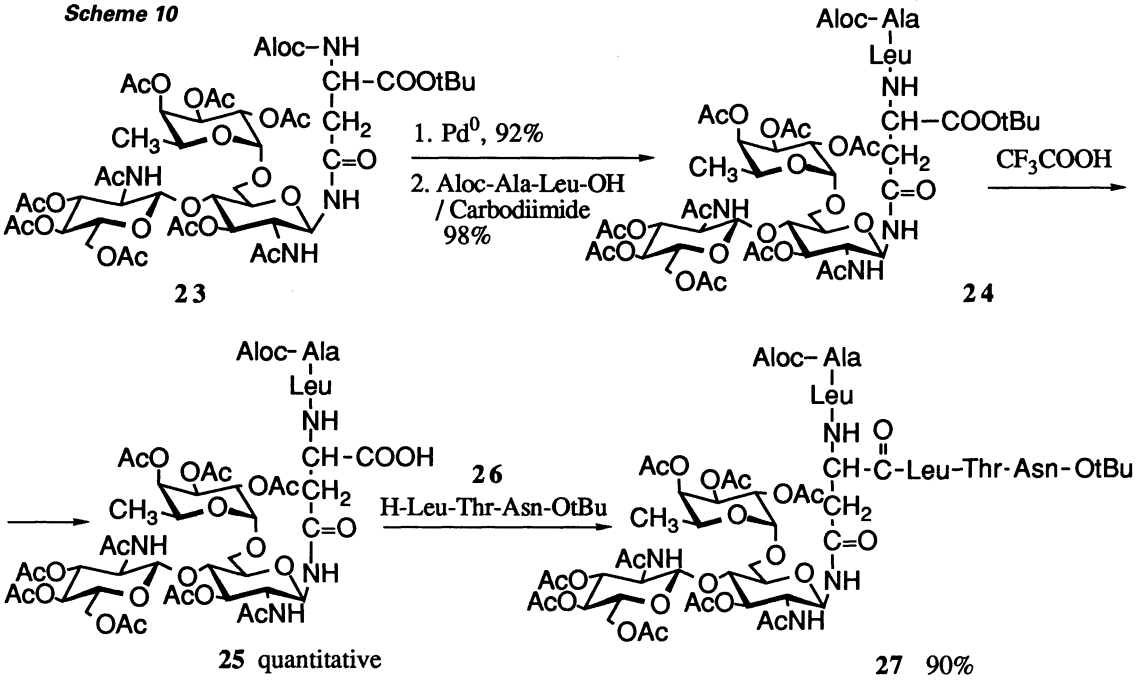
The construction of N-glycopeptides fucosylated in the core region demands the α-glycosidic binding of a fucose unit to O-6 of a chitobiose unit. As a general concept in these syntheses, we applied the azido group as the protected form of the required anomeric amine. The stereoselective introduction of the α-fucoside is only possible if ether-type-protected fucosyl donors are applied. The fucosyl chitobiose azide synthesized according to these guidelines was hydrogenolyzed to yield the trisaccharide amine which was condensed with N-allyloxycarbonyl aspartic acid α–tert-butyl ester to form the fucosyl chitobiose asparagine derivative **22** (ref. 14). The Aloc group can be removed from this conjugate with complete selectivity by applying the palladium (0)-catalyzed allyl transfer to dimedone. However, the treatment of **22** with trifluoroacetic acid, in order to remove the tert-butyl ester, resulted in the complete cleavage of the α-fucoside bond (ref. 14).

Scheme 9



To overcome these serious difficulties, a strategy for the synthesis of the fucosyl chitobiose azide was developed which includes an exchange of the ether-type protection of the fucose part. To achieve this aim, the 4-methoxy-benzyl ether (Mpm)-protected fucosylbromide was applied as the donor (ref.14, 36). After α-fucoside formation via in situ-anomerization (ref. 14), the Mpm groups were removed via oxidation with ceric

Scheme 10



ammonium nitrate (ref. 38) without affecting the azide function. In this way, acetyl-protection was introduced at all carbohydrate hydroxyl functions prior to the transformation of the anomeric azide to the amine. The condensation of the O-acetyl protected trisaccharide amine with the unblocked β-carboxy group of Aloc aspartic acid tert-butyl ester delivered the conjugate **23**, which was N-terminally deblocked via Pd(0)-catalyzed removal of the Aloc group (ref. 14). Chain extension gave the **24**, which was subjected to the cleavage of the tert-butyl ester by treatment with trifluroacetic acid to give the carboxy deblocked component **25** in almost quantitative yield (Scheme 10). Actually, the α-fucoside bond remained absolutely stable under these conditions, thus, revealing an important and interesting indirect stabilizing effect on the intersaccharide bonds exhibited by the O-acetyl protecting groups of the saccharide part (ref. 14).

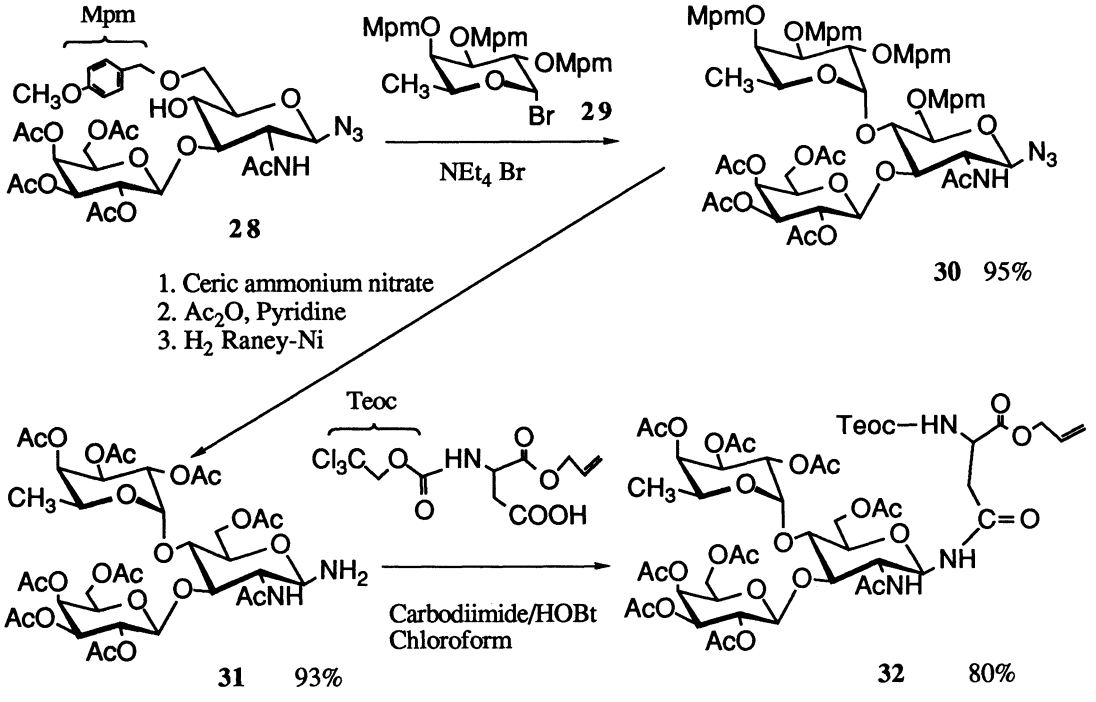
The compound **25** was condensed with a tripeptide ester **26** to form the trisaccharide hexapeptide **27** (ref. 14), a partial sequence of an envelope glycoproteine of a murine leukemia virus (ref. 8). Subsequent deblocking of the carboxy and carbohydrate functions allowed the conjugation of the synthetic glycopeptide with bovine serum albumin to furnish a glycoprotein antigen (ref. 14).

GLYCOPEPTIDES WITH LEWIS ANTIGEN SIDE CHAINS

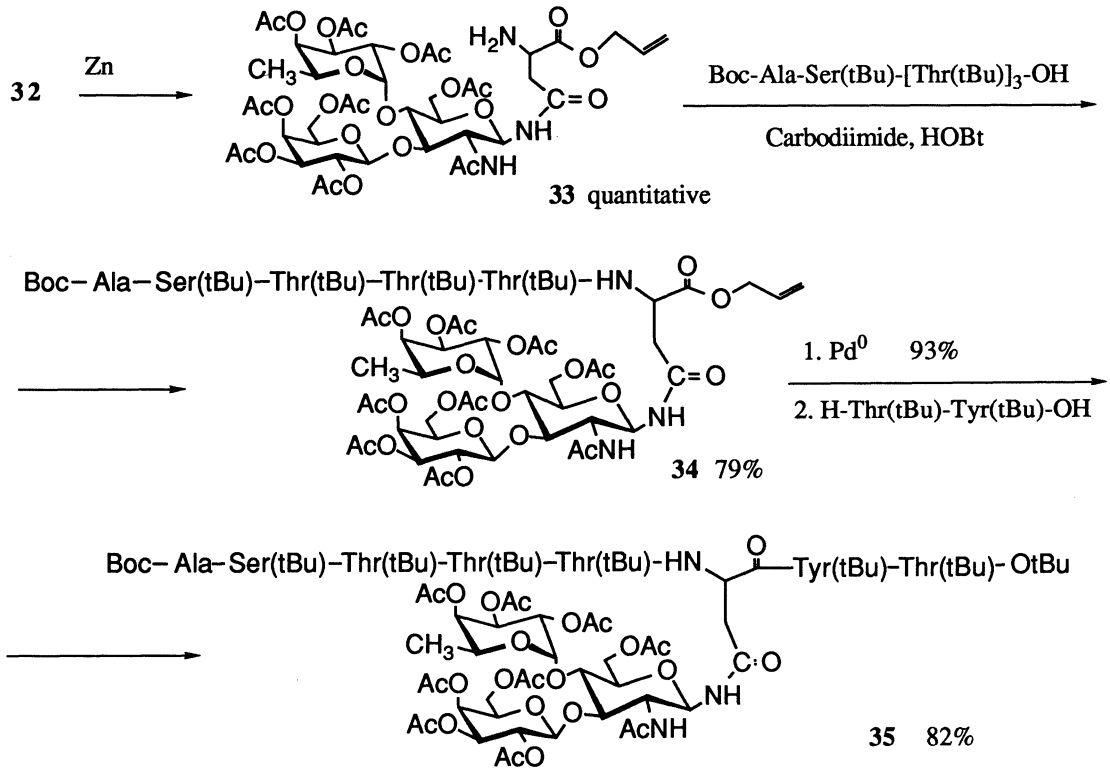
With respect to their roles as immunodeterminants, tumor-associated antigens or adhesion molecule substructures, Lewis antigen conjugates are receiving increasing attention. By application of the azide as the protected anomeric amine equivalent, in this case in the lactosamine derivative **28**, and the α-glycosidic introduction of the fucose unit by using the Mpm protected fucosyl bromide **29**, the Lewis^a antigen structure **30** was synthesized. Oxidative removal of the Mpm groups followed by O-acetylation and subsequent reduction of the azide gave the O-acetylated Lewis^a trisaccharide amine **31**. Its coupling with N-(2,2,2-trichloroethoxycarbonyl)aspartic acid allyl ester furnished the N-glycosidic Lewis^a asparagine conjugate **32** (Scheme 11, ref.39).

The protecting group pattern of **32** allows the mild and selective unblocking of both the amino and the carboxy function. Reductive elimination performed with zinc in acetic acid led to **33** in almost quantitative yield. It was condensed with a preformed tert-butyl-type protected pentapeptide to give the glycohexapeptide

Scheme 11



Scheme 12



34. The C-terminal allyl ester of **34** was removed completely selectively via Pd(0)-catalyzed allyltransfer to N,N´-dimethylbarbituric acid (yield 93 %). The obtained carboxy deblocked product was finally condensed with the O-tert-butyl protected tyrosine threonine dipeptide ester to yield the glycosylated peptide T **35** carrying the Lewis^a antigen side chain (Scheme 12, ref. 40). In the meantime, the complete deblocking of **35** was carried out and the synthetic glycopeptide was supplied for virological and immunological investigations. It should be noted that an analogous Lewis^x antigen trisaccharide and its asparagine conjugates can be constructed by application of a modified strategy including anomeric azides and Mpm protected fucosyl donors (ref. 41).

The selected examples outlined in this survey may illustrate that glycopeptides which constitute partial structures of biologically interesting glycoproteins are now accessible with high efficiency and in preparative amounts. Whereas the basic structures containing mono- or disaccharides can be synthesized more or less routinely, e. g. by using the modified Fmoc technique, the construction of the more complex glycopeptides, e. g. such ones containing fucose or β-mannoside units (ref. 42) demand more chemoselective differentiation and intercompatibility of the synthetic methods applied in the diverse transformations of these multi-step syntheses.

REFERENCES

1. For a comprehensive treatment, see: R.J. Ivatt, Ed., *The Biology of Glycoproteins*, Plenum Press, New York (1984).
2. R. J. Wieser and F. Oesch, *J. Cell. Biol. 103*, 361-367 (1986).
3. For a review, see e. g.: M. Fukada, *Biochim. Biophys. Acta 780*, 119-150 (1985).
4. G. F. Springer, *Science 224*, 1198-1206 (1984).
5. J. Samuel, A. A. Noujaim, G. D. MacLean, M. R. Sureshand and B. M. Longenecker, *Cancer Res. 50*, 4801-4808 (1990).
6. W. Reutter and C. Bauer in: *Advances in Enzyme Regulation*, Vol. 24, p 405, Pergamon Press, Oxford (1986).
7. R. Tauber, C. Park, W. Hofmann and W. Reutter in: *Membranes in Tumor Growth* (T. Galeotti, A. Cittandi, G. Neri and S. Papa, Eds), p 97, Elsevier, Amsterdam (1982).

8. M. Schlüter, D. Linder and R. Geyer, *Carbohydr. Res. 138*, 305-312 (1985).
9. For review, see: H. Kunz, *Angew. Chem. Int. Ed. Engl. 26*, 294-308 (1987).
10. For review, see: H. Garg and R. W. Jeanloz, Adv. *Carbohydr. Chem. Biochem. 43*, 135-201 (1985).
11. B. Andersen, P. Hoffmann and K. Meyer, *Biochim. Biophys. Acta 74*, 309-310 (1963).
12. R. R. Schmidt, *Angew. Chem. Int. Ed. Engl. 25*, 212-235 (1986).
13. H. Paulsen, *Angew. Chem. Int. Ed. Engl. 21*, 155-224 (1982).
14. H. Kunz and C. Unverzagt, *Angew. Chem. Int. Ed. Engl. 27*, 1697-1699 (1988).
15. H. Kunz and M. Buchholz, *Angew. Chem. Int. Ed. Engl. 20*, 894-895 (1981).
16 H. Kunz and H. Kauth, *Angew. Chem. Int. Ed. Engl. 20*, 895-896 (1981).
17. L. A. Carpino, *Acc. Chem. Res. 20*, 401-407 (1987).
18. P. Schultheiss-Reimann and H. Kunz, *Angew. Chem. Int. Ed. Engl. 22*, 62-63 (1983).
19. H. Kunz and S. Birnbach, *Angew. Chem. Int. Ed. Engl. 25*, 360-362 (1986); H. Kunz, S. Birnbach and P. Wernig, *Carbohydr. Res. 202*, 207-223 (1990).
20. B. Ferrari and A. A. Pavia, *Tetrahedron 41*, 1939-1944 (1985)
21. H. Paulsen and M. Schultz, *Carbohydr. Res. 159*, 37-52 (1987)
22. H. Paulsen, G. Merz, S. Peters and U. Weichert, *Liebigs Ann. Chem. 1990*, 1165; S. Peters, T. Bielefeldt, M. Meldal, K. Bock and H. Paulsen, *J. Chem. Soc. Perkin Trans. I 1992*, 1163-1171.
23. B. Lüning, T. Norberg and T. Tejbrant, *Glycoconjugate J. 6*, 5-19 (1989).
24. H. Kunz, B. Dombo and W. Kosch, in *Peptides 1988* (G. Jung and E. Bayer, Eds), p. 154, W. de Gruyter, Berlin (1989).
25. H. Kunz and S. Birnbach, *Tetrahedron Lett. 25*, 3567-3570 (1984).
26. H. Kunz and H. Waldmann, *Angew. Chem. Int. Ed. Engl. 23*, 71-72 (1984).
27. H. Kunz and C. Unverzagt, *Angew. Chem. Int, Ed. Engl. 23*, 436-437 (1984).
28. H. Kunz and J. März, *Angew. Chem. Int. Ed. Engl. 27*, 1375-1377 (1988).
29. H. Kunz, H. Waldmann and C. Unverzagt, in: *Peptides 1986* (D. Theodorpoulos, Ed.), p. 615, W. de Gruyter, Berlin (1987).
30. S. Friedrich-Bochnitschek, H. Waldmann and H. Kunz, *J. Org. Chem. 54*, 751-756 (1989).
31. H. Kunz and B. Dombo, *Ger. Pat. Appl.* P 3720269.3 (19th June 1987).
32. H. Kunz and B. Dombo, *Angew. Chem. Int. Ed. Engl. 27*, 711-713 (1988).
33. H. Kunz, W. Kosch and J. März, in *Peptides - Chemistry and Biology* (J. A. Smith and J. E. Rivier, Eds) p. 502, ESCOM, Leiden (1992).
34. D. E. Brenneman, J. M. Buzy, M. R. Ruff and C. B. Pert, *Drug Develop. Res. 15*, 371-379 (1988).
35. C. Birr, G. Becker, H. Nyuyen-Trong, T. Müller, M. Schramm, H. Kunz and W. Kosch, in: *Peptides - Chemistry and Biology* (J. A. Smith and J. E. Revier, Eds), p. 674, ESCOM, Leiden (1992).
36. H. Kunz and C. Unverzagt, *J. Prakt. Chem.*, issue 7 (1992), in press.
37. R. U. Lemieux and H. Driguez, *J. Am. Chem. Soc. 97*, 4063-4069 (1975).
38. R. Johannsson and B. Samuelsson, *J. Chem. Soc. Perkin Trans I 1984*, 2371-2374.
39. J. März and H. Kunz, *Synlett 1992*, 589-590.
40. H. Kunz and J. März, *Synlett 1992*, 591-592.
41. H. Kunz and K. von dem Bruch, unpublished; Dissertation K. von dem Bruch, Universität Mainz, to be submitted.
42. W. Günther and H. Kunz, *Angew. Chem. Int. Ed. Engl. 29*, 1050-1051 (1990).