New aspects of glycoside bond formation*

Richard R. Schmidt,† Julio C. Castro-Palomino and Oliver Retz

Fakultät für Chemie, Universität Konstanz, M 725, D-78457 Konstanz, Germany

Abstract: Glycoside bond formation generally requires activation of the sugar at the anomeric center. To this end, anomeric oxygen exchange reactions, resulting in the *Koenigs–Knorr method* and variations or, alternatively, activation through retention of the anomeric oxygen, resulting in the *trichloroacetimidate method* and in the *phosphite method*, have been proposed. The successful application of the *trichloroacetimidate method* to the total synthesis of GPI anchors is particularly worth mentioning. $\alpha[2\text{-}3]$-Sialylation can be based on sialyl phosphites as glycosyl donors and on the nitrile effect for anomeric stereocontrol. This is exhibited for a preparative synthesis of ganglioside GM_2 which is required for tumor vaccine studies. For the generation of $\alpha[2\text{-}8]$-linkage between neuraminic acid residues anchimeric assistance by a 3-thiocarbonyloxy group is introduced. The required sialyl donor can be efficiently prepared and α-linkage to 8-*O*-unprotected neuraminic acid derivatives is almost quantitative. The limitations of chemical glycopeptide synthesis encourage to employ the protein biosynthesis machinery in combination with an expanded genetic code. This is exhibited for the synthesis of glycosylated hARF-1 protein.

GLYCOSIDE BOND FORMATION—GENERAL ASPECTS

Glycoside bond formation in order to gain chemically defined oligosaccharides and glycoconjugates remains an important task, though generation of the anomeric linkage has seen years of dynamic progress, mainly based on highly reactive glycosyl donors, on versatile building block strategies, and on advanced protective-group design [1–3]. The endeavour was stimulated by the eminent role of carbohydrates and especially of glycoconjugates in various fields of modern biology [4,5]. A particularly prominent area in this regard is glycoconjugate synthesis.
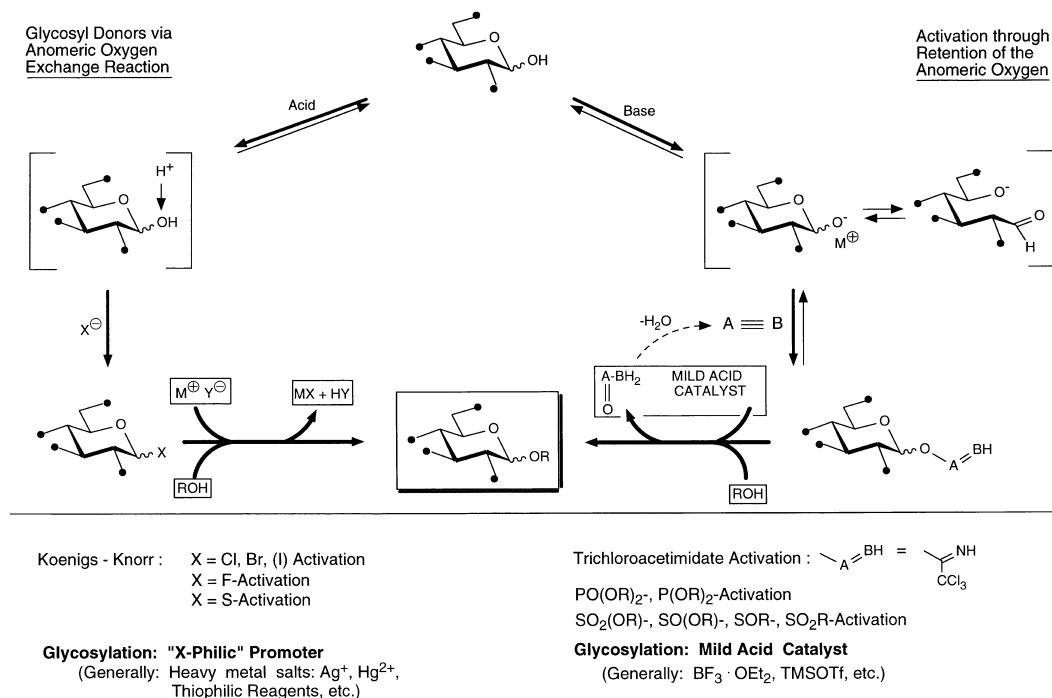
Activation of the sugar residue through **exchange of the anomeric hydroxy group** by bromine and chlorine, respectively, has led to the well-known *Koenigs–Knorr method* (Scheme 1). Thus an α-haloether is generated that can be readily activated in the *glycosylation step* by halophilic promoters, typically heavy-metal salts, thus resulting in an irreversible glycosyl transfer to the acceptor. This valuable method has been extensively reviewed [1.3]. It has been continually developed and widely applied. In spite of its generality, the requirement of at least an equimolar amount (often up to 4 eq) of metal salt as promoter (often incorrectly termed 'catalyst') and the problems concerning disposal of waste material (e.g. mercury salts) could be limiting factors for large-scale preparations. Therefore, alternative methods are of great interest.

Other *anomeric oxygen-exchange reactions* in the activation step have been quite extensively investigated. Closely related to the classic *Koenigs–Knorr method* is the introduction of fluorine as leaving group (see Scheme 1, X = F) [6] which, owing to the stability of the C–F bond, leads to stable glycosyl donors. Because of the generally lower glycosyl donor properties and because also at least equimolar amounts of promoter are required, the fluorides exhibit no real advantages over the corresponding glycosyl bromides or chlorides.

---

†Corresponding author.

Glycosyl Donors via Anomeric Oxygen Exchange Reaction

Activation through Retention of the Anomeric Oxygen

Acid

Base

H+

$X^{\ominus}$

$M^{\oplus} Y^{\ominus}$

MX + HY

ROH

$-H_2O$ → A ≡ B

A-BH₂ ‖ O

MILD ACID CATALYST

ROH

OR

Koenigs - Knorr : X = Cl, Br, (I) Activation
X = F-Activation
X = S-Activation

Glycosylation: "X-Philic" Promoter
(Generally: Heavy metal salts: Ag⁺, Hg²⁺, Thiophilic Reagents, etc.)

Trichloroacetimidate Activation :

PO(OR)₂-, P(OR)₂-Activation
SO₂(OR)-, SO(OR)-, SOR-, SO₂R-Activation

Glycosylation: Mild Acid Catalyst
(Generally: BF₃ · OEt₂, TMSOTf, etc.)

Scheme 1

Anomeric oxygen-exchange reactions by thio groups (see Scheme 1, X = SR) have recently attracted considerable attention for the generation of glycosyl donors [1,7]. Thioglycosides offer sufficient temporary protection of the anomeric center, thereby enabling various ensuing chemical modifications of the glycosyl donor without affecting the anomeric center. Additionally, they present several alternative possibilities for the generation of glycosyl donor properties: besides various thiophilic heavy-metal salts, also iodonium, bromonium, and chloronium ions are highly thiophilic; yet, at least equimolar amounts are required and with counterions, such as bromide and chloride, halogenoses are formed. Therefore, a poor nucleophile is required as counterion, for instance use of *N*-iodo- or *N*-bromosuccinimide or trifluoro-methanesulfonic acid, in order to enable a reaction with the acceptor as nucleophile. Obviously, the basic drawbacks of the activation through anomeric oxygen exchange reactions are also associated with this promoter system.

Activation through retention of the anomeric oxygen. The requirement for glycoside synthesis—high chemical and stereochemical yield, applicability to large-scale preparations, with avoidance of large amounts of waste materials, by having a glycosyl transfer from the activated intermediate through a catalytic process—were not effectively met by any of the methods described in the foregoing for the synthesis of complex oligosaccharides and glycoconjugates. However, the general strategy for glycoside bond formation seems to be correct:

1 The first step should be activation of the anomeric center under formation of a stable glycosyl donor (*activation step*).

2 The second step (*glycosylation step*) should consist of a sterically uniform, high yielding glycosyl transfer to the acceptor; it should be a catalytic process, where diastereocontrol is derived from the glycosyl donors' anomeric configuration (inversion or retention), or anchimeric assistance, influence of the solvent, thermodynamics, or any other effects.

This led us to the concept of sugar *activation through retention of the anomeric oxygen* (Scheme 1). This concept is based on simple base treatment of sugars, thereby generating from a pyranose or a furanose at first an anomeric oxide structure. Immediate alkylation (or arylation) leads to the **anomeric O-alkylation (O-arylation) method** for glycoside bond formation, which turned out to be highly

valuable for the synthesis of various glycosides [1,8]. However, extension of this methodology to the synthesis of glycoconjugates is hampered by the limited generality.

As an alternative the anomeric oxide can be used to generate glycosyl donors by addition to appropriate triple bond systems A≡B (or cumulenes A=B=C or by condensation with Z–A=BH systems, where Z represents a leaving group). The most successful methods developed thus far using these types of reactions are trichloroacetimidate (see Scheme 1 D, –A=BH = –C(CCl$_3$)=NH [1–3,8] and phosphate and phosphite formation (A=BH = PO(OR)$_2$, P(OR)$_2$) [3,9], also sulfoxide formation has been investigated [10]. The analogous glycosyl sulfates, sulfonates, and sulfites have not yet been as successful and not as extensively investigated. All these methods are particularly tempting because nature has a similar approach for generating glycosyl donors, namely *glycosyl phosphate* formation in the activation step and (*Lewis*) *acid catalysis* in the glycosylation step.

For the **trichloroacetimidate method**, the *activation step* consists of a simple base-catalyzed addition of the anomeric hydroxy group to trichloroacetonitrile, and the *glycosylation step* requires only catalytic amounts of a simple (Lewis) acid for the generation of strong glycosyl donor properties, thus leading to the desired glycosides in an irreversible manner. The water liberated on glycoside bond formation is thereby transferred in two separate steps to the activating species A≡B = CCl$_3$CN under formation of stable, nonbasic trichloroacetamide (see Scheme 1) providing the driving force for the glycosylation reaction. Thus, a very economic and efficient glycosylation procedure is available. Because of the very low basicity of the liberated trichloroacetamide, the (Lewis) acid required for activation of the basic *O*-glycosyl trichloroacetimidate is released and is ready for further activation of unreacted glycosyl donors.

This is exhibited in Scheme 2 for the **formation of *N*-acetyllactosamine** from protected *O*-galactosyl trichloroacetimidate as donor and 4-*O*-unprotected *N*-acetylglucosamine as acceptor [11]. The process very much resembles enzymatic *N*-acetyllactosamine generation [12]; however, the trichloroacetimidate-based process is obviously more simple, although protection and deprotection steps have to be taken into account.
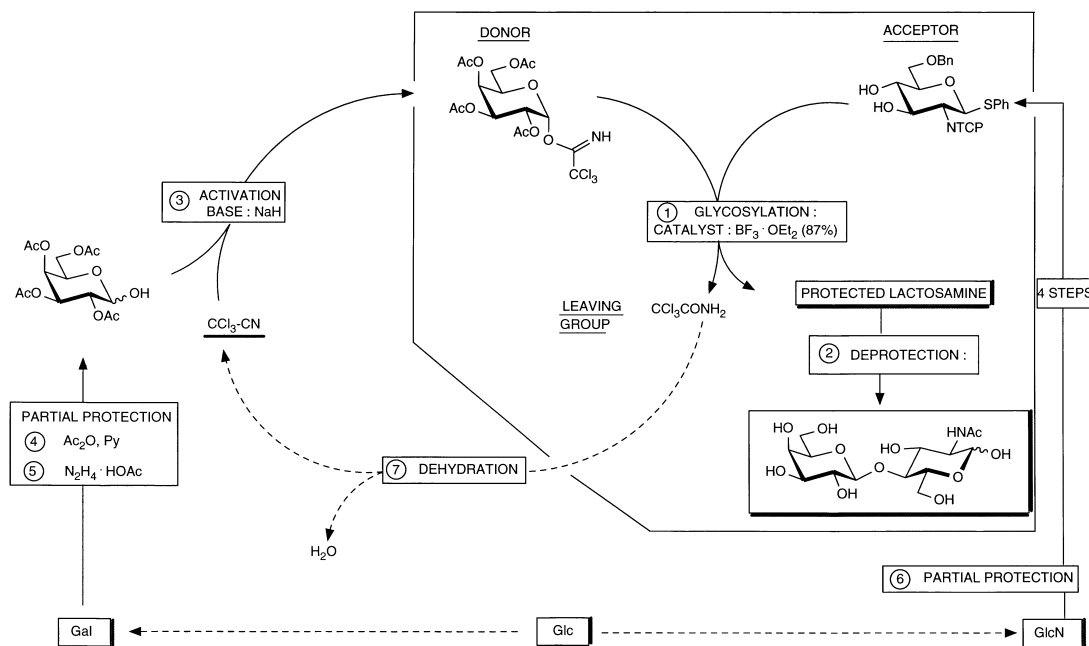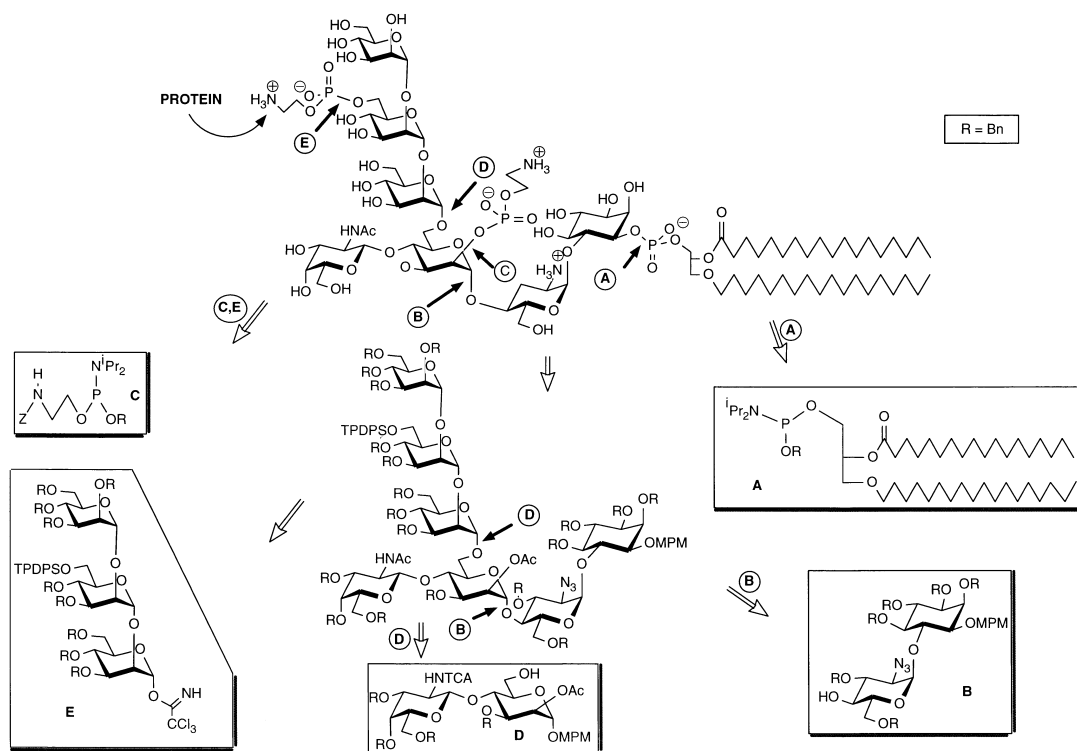

Scheme 2

RECENT APPLICATIONS OF THE TRICHLOROACETIMIDATE METHOD—GPI ANCHOR SYNTHESIS

The general significance of *O*-gycosyl trichloroacetimidates lies in their ability to act as strong glycosyl donors under relatively mild acid catalysis. This has been overwhelmingly confirmed in various

laboratories, and the scope and limitations of this method can be readily derived from these investigations [1–3,13,14]. Some representative examples of the successful application of the trichloroacetimidate method to various important glycoside bond formations from our group are the synthesis of:

- glycosphingolipids of the *lacto-*, the *lactoneo-*, the *globo-*, and the *ganglio-*series,
- various glycopeptides,
- glycosyl phosphatidyl inositol (GPI) anchors,
- saponins,
- macrocyclic glycolipids, and
- various glycosyl phosphates [3,9,15].

The last example is particularly interesting, because exchange of the trichloroacetimidate group at the anomeric position by phosphorous acid esters (for instance, phospholipids, nucleoside-monophosphoric acids, etc.) leads without catalysis directly to glycosyl phosphates which constitute interesting glycoconjugates used by nature [15]. The phospholipid exchange reaction could lead to interesting membrane constituents; however, due to the lability of the glycosidic bond in glycosyl phosphates, nature did not use this type of compounds for this purpose [16]. To overcome the stability problem, a pseudosugar, namely inositol, was employed instead. Thus, the GPI anchors are generated which possess great importance in anchoring proteins and glycoproteins to membranes [17,18]. Various structural types have been identified and the first total syntheses have been reported [18–21]. We have concentrated our efforts on the total synthesis of a ceramide-1-phosphate residue containing a GPI anchor which appears in *Saccharomyces cerevisiae* [18], and on the glycerol-1-phosphate residue containing **GPI anchor Thy-1** which appears in rat brain [21]. The strategy for the GPI anchor synthesis of rat brain Thy-1 is exhibited in Scheme 3.
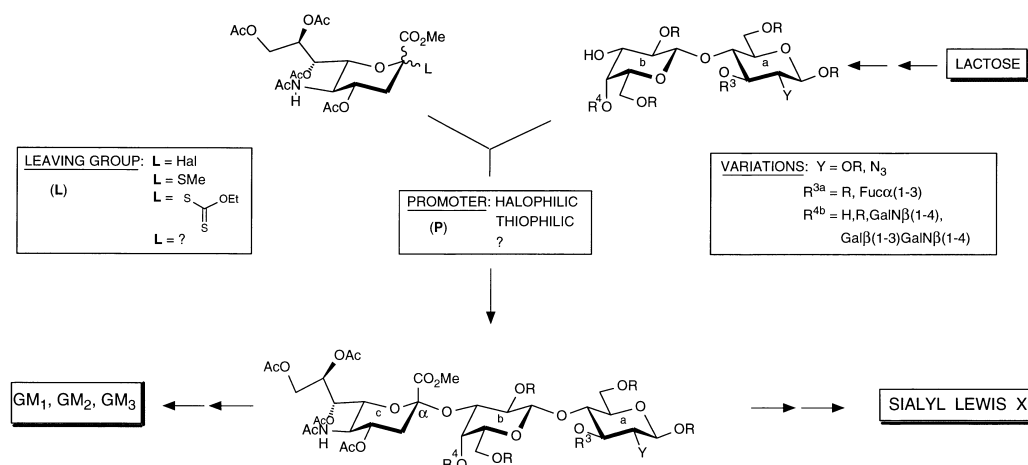

Scheme 3

The desintegrations A–E lead to building blocks **A–E**. All glycosidic linkages of building blocks **B**, **D**, and **E** and also their linkage to the (pseudo)heptasaccharide could be successfully performed by the trichloroacetimidate method, thus leading finally to the desired target molecule [21]. The versatility of the building blocks, the high regio- and diastereoselectivities, and the high yields obtained in all reaction steps provide an excellent basis for further successful syntheses in this field.

SYNTHESIS OF SIALYLATED GLYCOCONJUGATES

In the chemical synthesis of gangliosides, sialylation has been carried out by various methodologies [22]. When halogenoses of *O,N*-acylated neuraminic acid esters are used as donors (**L** = Hal in Scheme 4) only modest yields of the desired α-products are obtained [15], especially when secondary hydroxyl groups are used as the acceptor. Therefore, 3b-*O*-unprotected lactose and lactosamine derivatives afford the desired building blocks for gangliosides of the *ganglio* (GM$_1$, GM$_2$, GM$_3$, etc.) or the *lactoneo* series (sialyl Lewis X), respectively (Scheme 4) only in low yields.
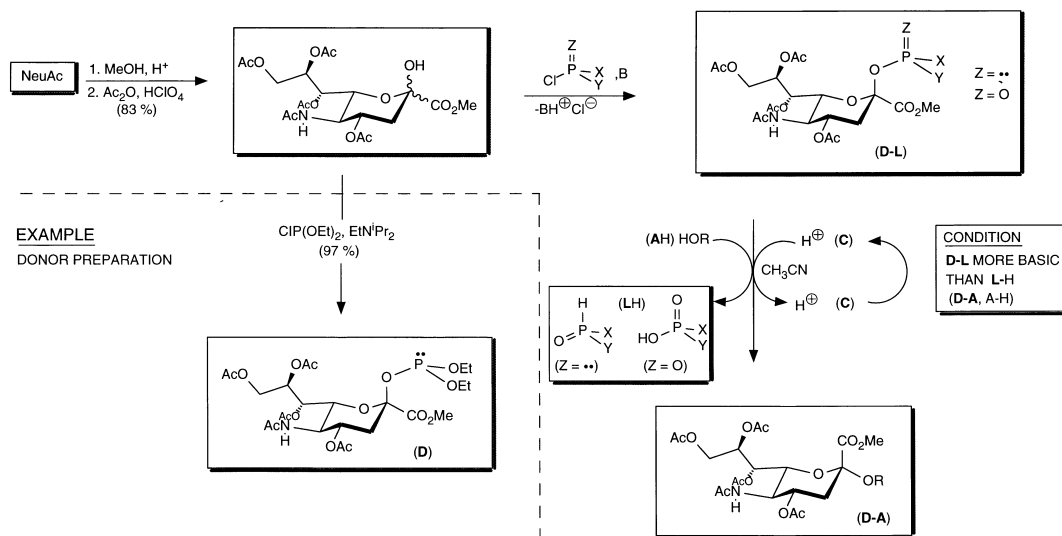

Scheme 4

Thioglycosides of neuraminic acid derivatives have been employed as sialyl donors (**L** = SR in Scheme 4) [23,24]. However, the requirement of at least equimolar amounts of thiophilic reagents as promoters **P** constitutes a major disadvantage in this approach. Therefore, improvements in the sialylation step in terms of yield of α-product and in ease of performance of the reaction were of interest.

The influence of solvents on the anomeric ratio of glycosylation reactions is well known [1–3,25]. The dramatic effect of nitriles as participating solvents has been reported by us for various glycosyl donors [26] and it has been also extended to thioglycosides of neuraminic acid derivatives as sialyl donors [27–29]. Thus, under S$_N$1-type conditions a promoter generates from the sialyl donor a (solvent separated) ion pair which in nitrile solvents will be intercepted under kinetically controlled conditions from the β-face (axial attack), thus leading to β-nitrilium-nitrile conjugates. Reaction of the β-nitrilium-nitrile conjugates with the acceptor, prior to slow transformation into α-nitrilium-nitrile species [9,27], should lead to the equatorial glycoside which is the desired α-glycosidic in neuraminic acid, as could be clearly shown [27,29].

Consideration of various leaving groups led us to phosphite (and phosphate) moieties and their derivatives [9,30] (Scheme 5). Thus, a readily available *O*-acetylated neuraminic acid derivative will be transformed into a sialyl donor (**D**) which with alcohols as acceptors (**AH**) and in nitrile solvents will provide the target molecules (**D–A**). The catalyst (**C**) due to the basicity of the leaving group will preferentially attack **D**. However, the cleavage product (= **LH**) will release the catalyst **C** because of its low basicity, with the help of the proton available from the acceptor hydroxy group. Obviously, the phosphite/phosphonate system (**Z** = ··) seems to be especially suitable because basicity and strong leaving group character in the phosphite species (= **L**) are combined with relatively low acidity and

basicity in the released phosphonate species ($=$**LH**). Additionally, phosphitylation at the anomeric position can be performed in high yields as indicated in Scheme 5 for an ethyl ester [9].
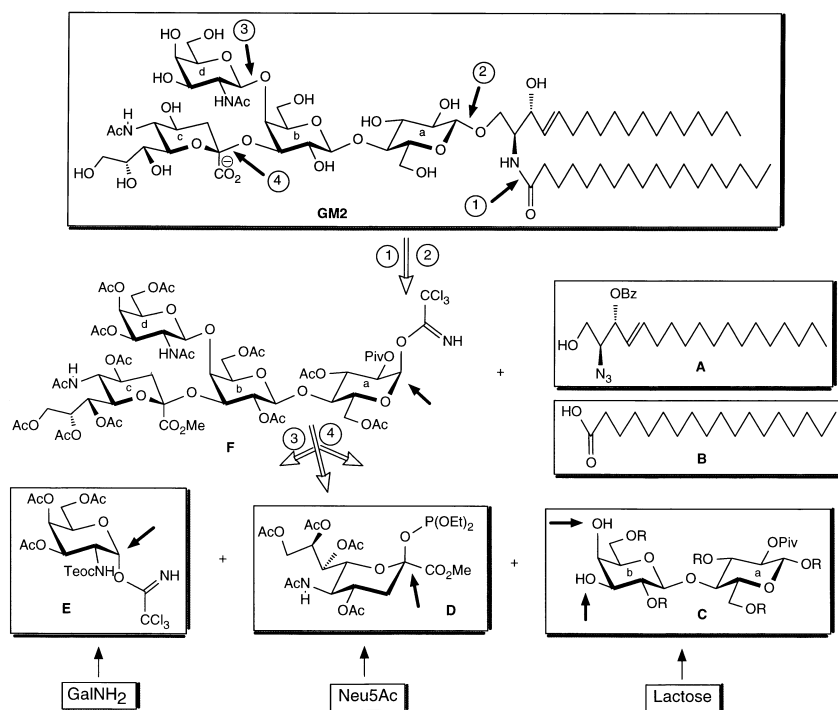

Scheme 5

Sialyl phosphites became important sialyl donors, due to the ease of their formation and their convenient activation by addition of catalytic amounts of TMSOTf. They have been employed for synthesis of gangliosides (GM$_2$, GM$_1$, GD$_{1a}$, BGM$_1$; sialyl Lewis X and Lewis^a, sialylgalactosylgloboside), and sialylated glycopeptides [15,31]. Particularly interesting is the phosphite/phosphate exchange reaction which led to convenient syntheses of CMP-Neu5Ac and derivatives having modifications in the nucleoside and the neuraminic acid residue [32]. Thus, in combination with the 'nitrile effect' a simple and highly α-selective chemical sialylation method could be developed. This chemistry led us also to enzymatic sialylations with the help of transferases and to their efficient inhibition. Transition state analogoues inhibitors could be recently developed which possess K$_i$ values in the lower nanomolar range [33].

SYNTHESIS OF GM$_2$ FOR TUMOR VACCINE STUDIES

A number of carbohydrate antigens that are expressed on human cancer cells on glycolipids and glycoproteins are considered attractive targets for immunotherapy with monoclonal antibodies and vaccines [34]. Particularly interesting is GM$_2$ because it is expressed on a number of human cancers, including melanoma, sarcoma, and renal cancer [35]. Various approaches in order to induce an immune response against GM$_2^+$ cancer cells have been pursued, however, the majority of the antibodies generated based on GM$_2$ derived from mammalian tissues failed to react with GM$_2$-expressing human cancer cells [8]. Therefore, for a systematic clinical vaccine development, a consistent source of a single, well-defined synthetic GM$_2$ is desirable. To this aim we have developed a practical method for the chemical synthesis of ample quantities of GM$_2$ which should meet the required purity standards.

The strategy pursued follows the GM$_3$ route [38] (Scheme 6). The most important synthetic problems in this endeavour are: (a) convenient generation of a suitable lactose building block, (b) α-selective attachment of a sialyl donor to the 3b-hydroxy group of the lactose constituent, (c) high-yielding β-selective attachment of an N-protected galactosamine residue to the low-reactive 4b-hydroxy group of the GM$_3$ trisaccharide intermediate, and (d) convenient transformation of the N-protected galactosamine residue into the N-acetyl-galactosamine constituent. Problems (a)–(d) can be solved by means of building blocks **A–E**, which constitute tetrasaccharide intermediate **F**.

After extensive investigation of various amino protective groups (N-tetrachlorophthaloyl, N,N-diacetyl,

Scheme 6

N-dimethylmaleoyl, *N*-trichloroethoxycarbonyl, [39]) we selected *N*-trichloroethoxycarbonyl (Teoc)-protected trichloroacetimidate **E** which was obtained from galactosamine via the 1-*O*-unprotected intermediate in three convenient steps in high yield (Scheme 7). This procedure avoided the tedious azidogalactose
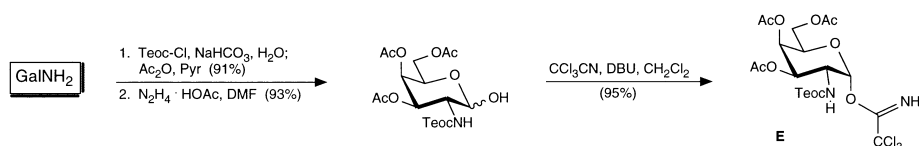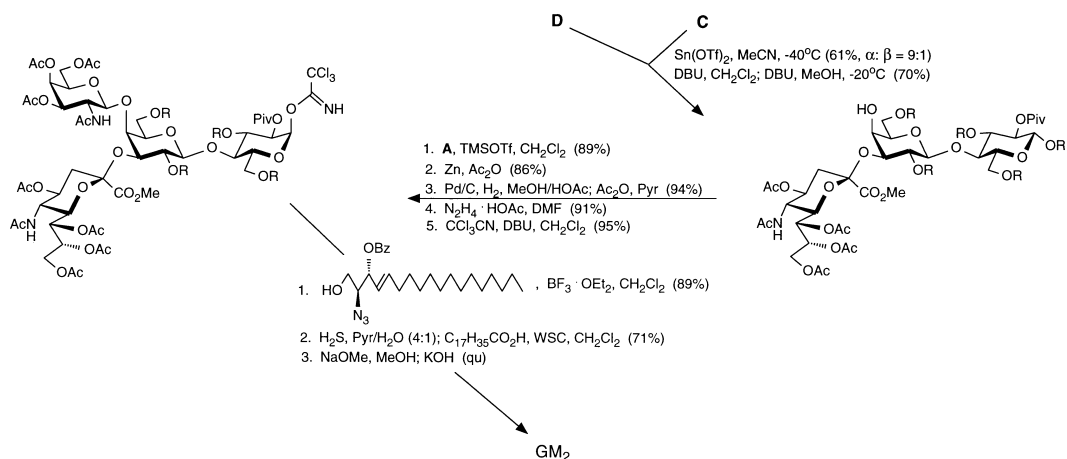

Scheme 7

production [15,38] and the difficult removal of *N*-phthaloyl protecting groups in the presence of Neu5Ac residues of previous GM$_2$ syntheses [40]. Treatment of *N*-Teoc-containing compounds with Zn/acetic anhydride leads to direct replacement of *N*-Teoc by an acetyl group [41]. The known diethyl phosphite derivative [9] was employed as Neu5Ac donor **D**.

The 3b,4b-*O*-unprotected 2a-*O*-pivaloyllactose residue **C** [42] facilitates the desired consecutive regioselective attack at the 3b-and then at the 4b-hydroxyl group, because of their different reactivities. The number of steps required for the synthesis of **C** [42] was greatly reduced (all in all 7 steps) by employing a 1a,2a-*O*-silyl-group migration [43] for the regioselective introduction of the 2a-*O*-pivaloyl group.

Reaction of sialyl donors with 2,3,4-*O*-unprotected galactosyl residues provides generally good α-selectivity. This is usually not found for 3,4-*O*-unprotected galactose derivatives [9,44]. Investigations with donor **D** and acceptor **C** and various catalysts showed that tin(II) triflate in acetonitrile at −40 °C leads to high α selectivity (α:β = 9:1) and good yields of a GM$_3$ intermediate (Scheme 8); at room temperature sialoside yields of up to 80% were obtained with a slightly higher content of the β anomer (α:β = 4:1); the α product can be separated by a simple technique. Ensuing glycosylation with donor **E** gave the tetrasaccharide in almost quantitative yield. Replacement of the Teoc group by an acetyl group with Zn/Ac$_2$O proceeded smoothly. Hydrogenolysis with Pd/C in MeOH/HOAc, followed

Scheme 8

by treatment with Ac$_2$O in pyridine led to replacement of all *O*-benzyl groups by *O*-acetyl groups. All subsequent procedures followed our previously introduced standard methods: regioselective removal of the anomeric *O*-acetyl group with N$_2$H$_4$·HOAc and then treatment with CCl$_3$CN in the presence of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) as base afforded tetrasaccharide donor **F**. Application of the 'azidosphingosine glycosylation procedure' [38], transformation of the azido into the amino group, attachment of the stearoyl residue, and then removal of all protective groups furnished the target molecule that was identical to previously synthesized material [36,40]. This material is now employed in clinical studies.

α(2-8)-LINKAGE BETWEEN NEURAMINIC ACID RESIDUES

The disaccharide sequence Neu5Acα(2-8)Neu5Ac is a principal constituent of a number of glycoconjugates including a series of gangliosides. They were found to play an important role in numerous biological phenomena being, for example, tumor-associated antigens or receptors for bacterial toxins and viruses [45]. Attempts to prepare the target disaccharide directly, using conventional nonmodified sialyl donors such as thioglycoside [46] or phosphite [47], gave thus far only very low yields.

It was assumed that the low reactivity of the 8-OH group in the derivatives of Neu5Ac possessing $^2C_5$ conformation (Scheme 9, **A**, see arrow) is caused by its interaction with the 5-acetamido group (or, alternatively, with the ring oxygen) via the formation of hydrogen bonds. In order to avoid this undesirable interaction, we have decided to apply 5-azido derivatives of Neu5Ac, derivatives of the Neu5Ac 1,7-lactone, and 1,5-lactam derivatives of Neu5Ac (Scheme 9, **B** and **C**) as sialyl acceptors [48]. In these derivatives, due to the rigid conformation, the 8-OH group and the 5-acetamido group are remote, thus preventing the interaction between them. However, though the acceptor reactivities were greatly increased, essentially only the undesired β(2-8)-linkage was obtained. Therefore, the study was focussed to the influence of the functional groups at the anomeric position. Obviously, interaction between the carboxylate group and the 8-OH group in the α-anomer (Scheme 9, **D**) could also strongly influence the acceptor properties of the 8-OH group and thus explain their low reactivity. Therefore, the reactivity of β-anomers (Scheme 9, **D**), of 2,8-*O*-unprotected (**E**), and of 2,3-dehydroneuraminic acid derivatives (**F**) was investigated [48]. Again strongly increased sialyl acceptor properties were found, yet the β(2-8)-linkage was still the major product found. Even an intramolecular approach to the synthesis of the α(2-8)-linkage was not successful, though we had recently great success with a 'rigid-spacer based approach' to intramolecular glycoside bond formation [49].

Obviously, high α-sialylation at the 8-OH group of Neu5Ac derivatives cannot be reached just by improving the acceptor properties. Presumably, due to steric constraints and due to a stereochemically mismatched situation in the transition state between the reactands, stereocontrol by directing groups at the donor moiety is required. Therefore, anchimeric assistance by an auxiliary group at C-3 has been
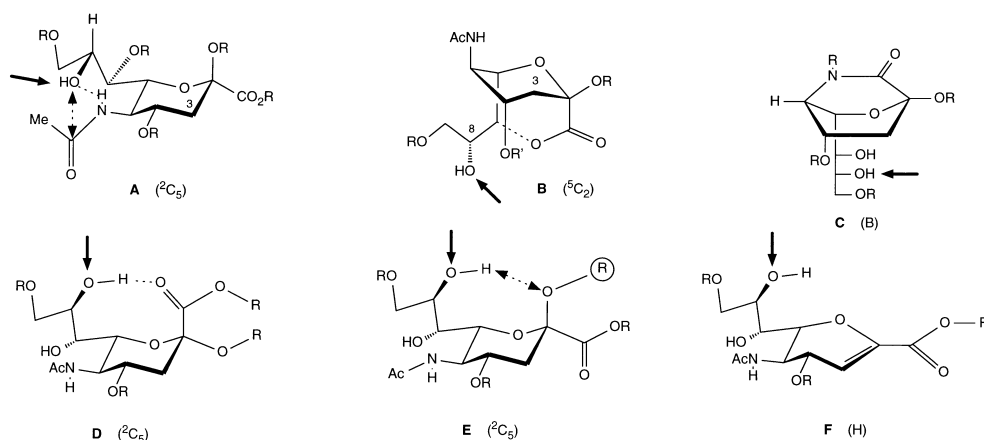
Scheme 9

proposed [50]. To this end, addition of X′Y′ reagents to 2,3-dehydroneuraminic acid **A** (Scheme 10), affording **Bα′** and **Bβ′**, and then transformation into desired **Bβ** with Y on the β-side was performed, where Y represents a group prone to neighboring group participation. Mainly HOBr and PhSCl addition to **A** (R = Ac, Bn) were investigated. HOBr addition did not only lead to the wrong regioisomers, but also α- and β-face addition was obtained; then, in a quite tedious procedure, both compounds were transformed into desired **Bβ**, respectively. PhSCl addition gave the desired regioisomers, yet again α- and β-face addition was observed; therefore, chromatographic separation and transformation into **Bβ** were required. Then, with the help of promoter systems, intermediate **C** was generated which should result in α(2-8)-sialoside formation [50]. These lengthy and not very efficient methodologies for Neu5Acα(2-8)Neu5Ac formation were even employed to the generation of Neu5Acα(2,3)Gal linkages [51].
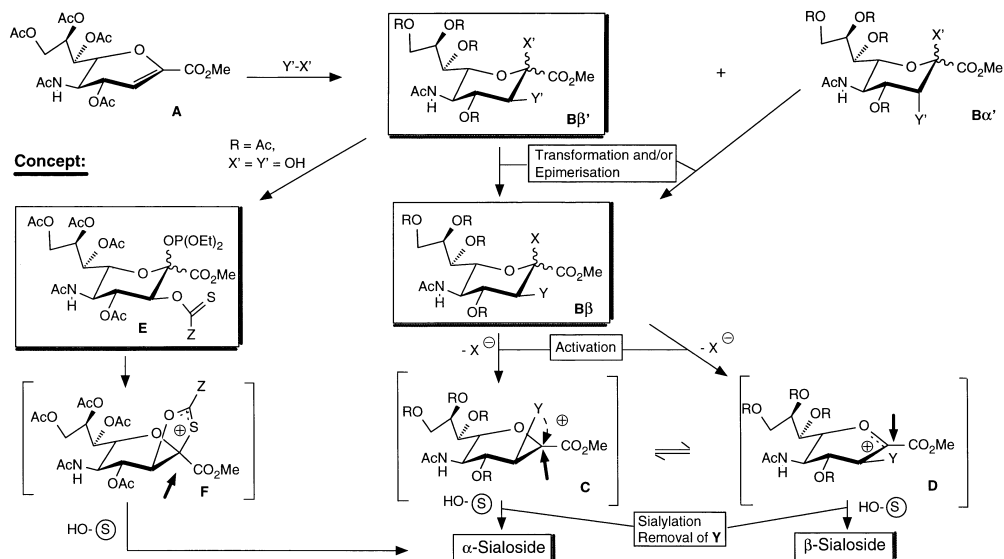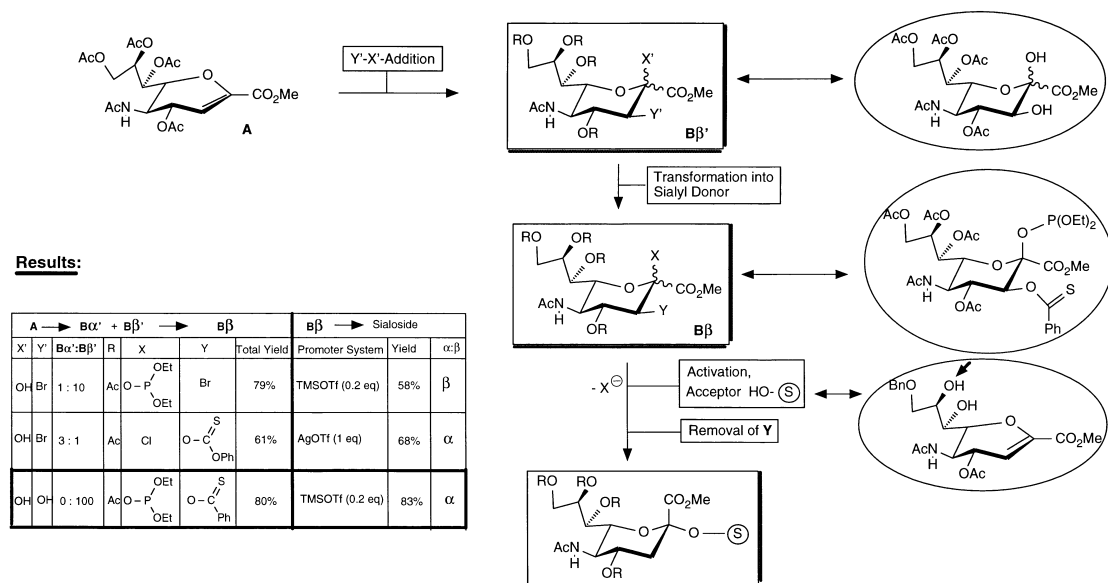

Scheme 10

In order to improve the effciency and simplicity of the Neu5Acα(2-8)Neu5Ac linkage methodology the following five demands should be fulfilled:

1 Y′-X′-Addition to **A** should lead only to **Bβ′** giving ready access to **Bβ**.

2 Activation of **Bβ** should be achieved by simple catalysis [X = OP(OR)$_2$, not SR].

3 Efficient neighboring group participation by **Y** should be provided.

4   The anchimerically assisting group **Y** should accomodate its immediate and practically quantitative removal after product formation.

5   The Neu5Ac donor and/or acceptor should be designed in such a way as to permit repetitive glycosylation yielding α(2-8)-linked homooligomers of Neu5Ac.

Obviously, a compound of the general structure **E** could accommodate demands [1–4]: acid-catalyzed removal of the phosphite moiety should lead to relatively stable five-membered oxathiolanium intermediate **F** which, on proper selection of group Z, will not result in loss of Z^+ or give with the acceptor an ortho ester derivative, but instead it will smoothly lead to the desired α-sialoside linkage in high yield.

To test this **concept**, we have first investigated the phenoxythiocarbonyloxy group as the anchimerically assisting group and chloride as the leaving group [X = Cl; Y = PhO–C-(=S)–O]. Then we investigated the usefulness of phosphite as the leaving group in the presence of bromine as the anchimerically assisting group [X = OP(OEt)$_2$; Y = Br]. Finally, the combination of the phosphite leaving group and the thiobenzoyloxy group as the anchimerically assisting group was investigated [X = OP(OEt)$_2$; Y = PhC(=S)–O]. For the results, see the table in Scheme 11. In addition, β-face selective addition of the Y′ group had to be solved. It turned out that osmium tetroxide catalyzed dihydroxylation of *O*-acetyl protected 2,3-dehydroneuraminic acid methyl ester **A** using *N*-methylmorpholine *N*-oxide as oxidizing agent affords via a β-selective dihydroxylation only the desired 2,3-dihydroxy derivative (=**Bβ′**) which could be readily transformed into 3-*O*-thiobenzoyl-2-*O*-diethylphosphite derivative **E** (=**Bβ**) as sialyl donor. With this compound the desired α(2-8)-linkage could be obtained in 83% yield (last line in the table). Thus, the demands for an efficient generation of the Neu5Acα-(2-8)Neu5Ac linkage could be well fulfilled (isolated yield from **E**, 83%; total yield from **A**, 67%). The α(2-8)-linked disaccharide offers also repetitive application of this methodology. Obviously, a convenient and highly efficient methodology for the syntheses of α-sialosides and especially of α-(2-8)-linked neuraminic acids and possibly their homooligomers could be developed.

Results:

A ⟶ Bα′ + Bβ′ ⟶ Bβ							**Bβ ⟶ Sialoside**			
X′	Y′	**Bα′:Bβ′**	R	X	Y	Total Yield	Promoter System	Yield	α:β	
OH	Br	1 : 10	Ac	O–P(OEt)(OEt)	Br	79%	TMSOTf (0.2 eq)	58%	β	
OH	Br	3 : 1	Ac	Cl	O–C(=S)OPh	61%	AgOTf (1 eq)	68%	α	
OH	OH	0 : 100	Ac	O–P(OEt)(OEt)	O–C(=S)Ph	80%	TMSOTf (0.2 eq)	83%	α	

Scheme 11

Application of this concept to other sugars, for instance to glucose and mannose, leads to highly selective syntheses of α- and β-glucopyranosides, of α- and β-mannopyranosides, and of 2-deoxy-α- and -β-glucopyranosides (see Scheme 12, [52]). Therefore the usefulness of this concept is accompanied by a wide applicability.
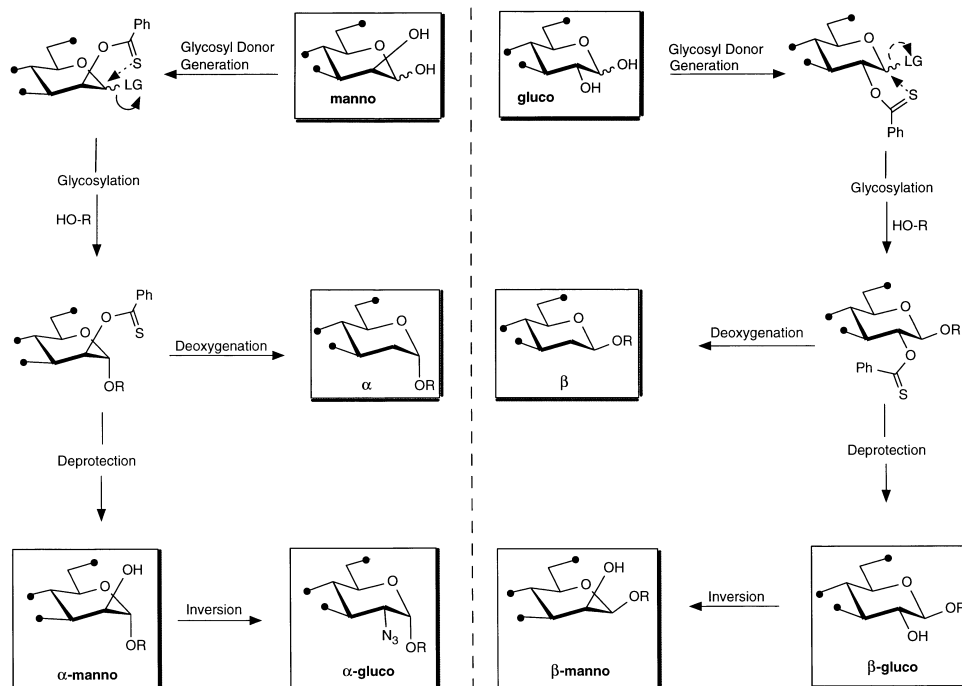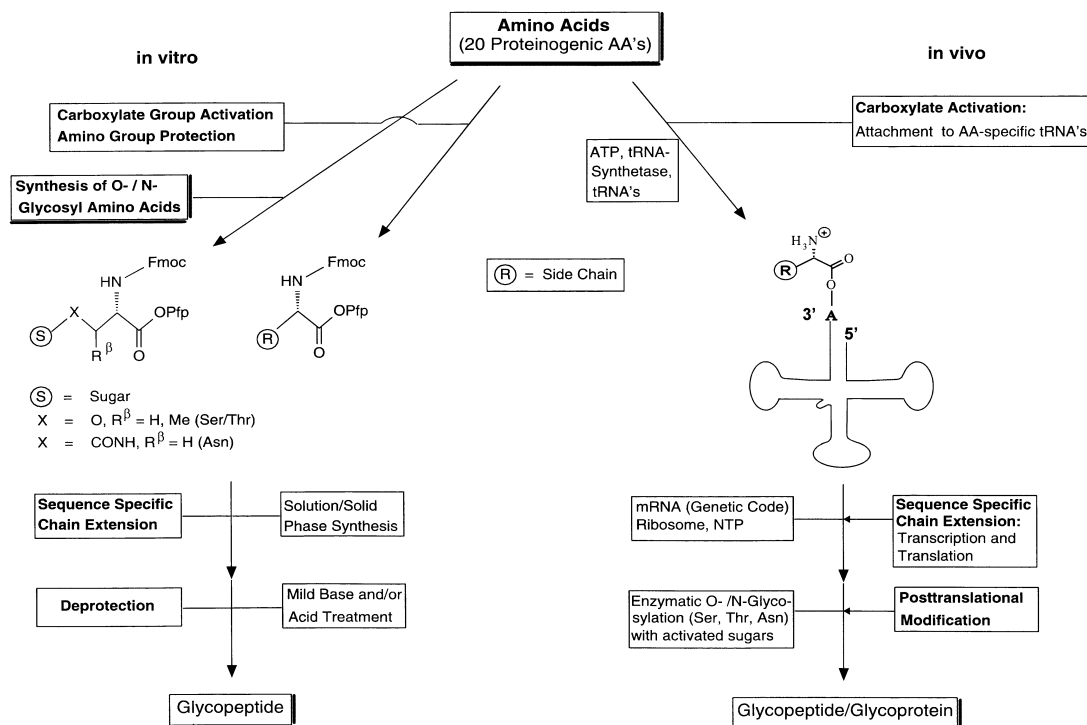
Scheme 12

GLYCOPEPTIDE AND GLYCOPROTEIN SYNTHESIS—USE OF AN EXPANDED GENETIC CODE

The ubiquitous occurrence and the importance of glycopeptides and glycoproteins has led to the investigation of different synthetic methodologies. Chemical syntheses are based on *O*-and *N*-glycosyl amino acids (for instance of Ser, Thr, Asn) which are used in solution or solid phase-based peptide synthesis (Scheme 13, left half); thus, various glycopeptides have been successfully obtained. Yet, there are limitations in terms of complexity and size of the desired glycopeptide. Conceptually the biological approach to glycoprotein synthesis is different. Transcription of DNA and translation via codon/anticodon recognition leads to proteins which are post-translationally modified with the help of glycosyltransferases (Scheme 13, right half).

However, because codon/anticodon recognition is independent of the amino acid attached to tRNA and the translation system exhibits a broad substrate acceptance, also *in vitro* approaches to the synthesis of glycoproteins using the biosynthesis machinery have been investigated [55]. Yet, site-specific incorporation of modified amino acids is limited by the genetic code, encoding for the 20 proteinogenic amino acids, which are commonly present in each protein more than once. Therefore, for site-specific variations additional codons would be required which are not available (and not required) in the protein biosynthesis machinery. However, recently such a method, which is based on an expanded genetic code, has been developed [56].

The genetic code contains three stop codons, therefore two stop codons are available for encoding modified amino acids. To this aim, site-directed mutagenesis is employed in order to introduce, for instance, stop codon UAG at the site to be modified in the protein gene, which is then used to generate the corresponding mRNA. Additionally, a suppressor tRNA is required having 5′-CUA-3′ recognizes stop codon UAG as codon; This suppressor tRNA has to be loaded by chemical means with a modified amino acid (see Scheme 14).

Obviously, the required suppressor tRNA must be recognized by the protein biosynthesis machinery; yet, it should not be a substrate for any of the aminoacyl tRNA synthetases present in the *in vitro*

Scheme 13

transcription-translation system. In order to avoid problems due to proof-reading of the amino acyl tRNA synthetases, the suppressor tRNA is derived from a different species. In this tRNA then the anticodon is replaced by 5′-CUA-3′. Chemical linkage of the modified amino acids to the 3′-/2′-end of a pdCpA residue, which is then combined with a truncated suppressor tRNA [tRNA(-CA)] with the help of T4 RNA ligase, became the method of choice to obtain modified aminoacyl suppressor tRNAs [57]. Thus,
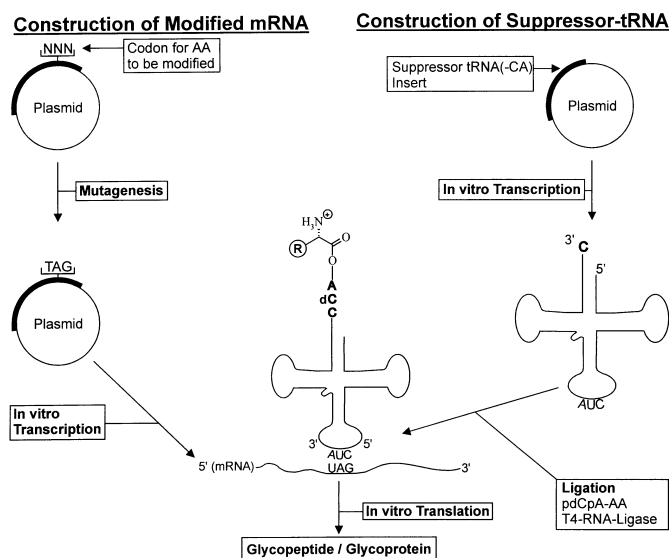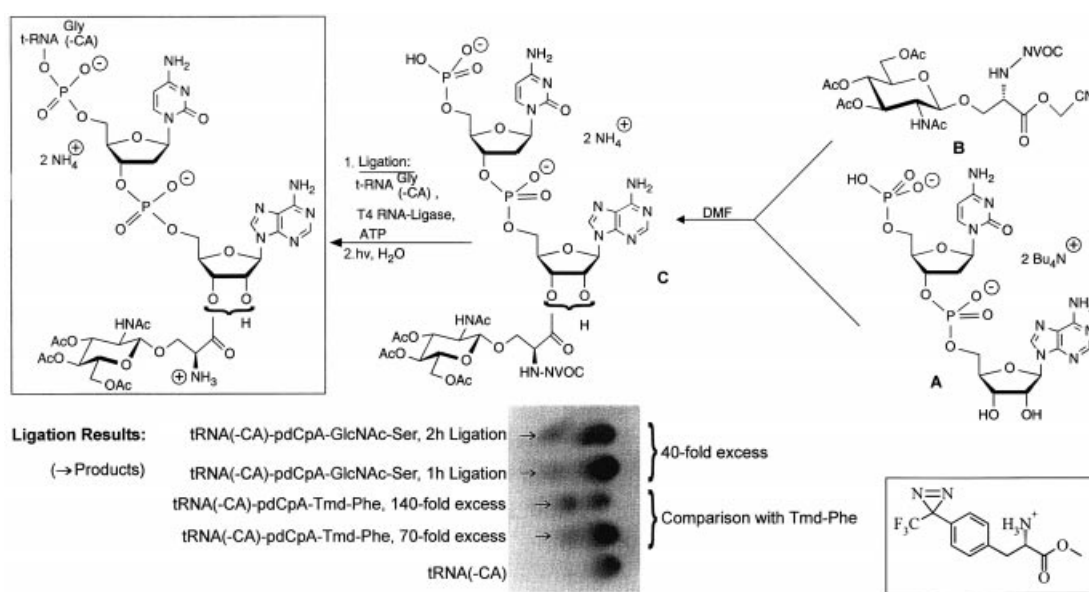

Scheme 14

recently a serine residue by an *O*-glucosyl serine residue in firefly luciferase was replaced [58]. We would like to report on the replacement of lysine 142 in hARF 1, a protein participating in vesicle formation [59], by an *O*-acetyl GlcNAcβ(1-O)Ser residue. This work was carried out in collaboration with Prof. F. Wieland and L. Zhao, Biochemie-Zentrum Heidelberg.

Site-directed mutagenesis in order to obtain the required mRNA having the 5′-UAG-3′ codon was performed as previously described [59]. For the construction of the truncated suppressor tRNA, plasmid pUC19 [tRNA(-CA)$^{Gly}_{E. coli}$] was employed [55]. Cleavage of the plasmid with restriction enzymes (*Fok I, Pst I*) led to linearized cDNA which was used as template for the transcription into tRNA(-CA)$^{Gly}_{E. coli}$].

For the ligation with the pdCpA-GlcNAcSer residue first pdCpA (**A**, Scheme 15) and GlcNAcSer derivative **B** [60] having an activated carboxylate group and *N*-protection had to be prepared. The synthesis of **A** followed a known protocol [57]. For **B**, cyanomethyl activation of the carboxylate group and protection of the amino group by the photolabile 2-nitroveratryloxycarbonyl (NVOC) [57] group was selected, thus providing the loaded pdCpA intermediate **C**.

Ligation of excess **C** (40-fold) with supp tRNA(-CA)$^{Gly}_{E. coli}$ in the presence of T4 RNA ligase led to the desired supp tRNA$^{GlcNAcSer}_{E. coli}$ derivative as indicated by the arrows in Scheme 15, however, only in low yield. Comparison with corresponding known TmdPhe derivative reveals product formation, yet, presumably due to the use of less excess of **C**, the yields were smaller. The photolabile NVOC group was removed just before use of the complete suppressor tRNA in the translation process.
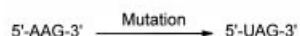

Scheme 15

The translation of the mutated mRNA of hARF 1 was carried out with rabbit reticulocyte lysate to which the mutated mRNA, the supp tRNA$^{GlcSer}_{E. coli}$ and ^{35}S-methionine were added; thus, protein formation can be identified by radiography. Scheme 16 exhibits that GlcNAcSer modified hARF 1 is formed. This is revealed by comparison with the negative and the positive control, which has known TmdPhe modification instead of lysine.

Obviously, this method for glycoprotein synthesis is only in a preliminary stage and great improvements are required in order to establish its general usefulness for the preparative synthesis of structurally defined glycoproteins. Yet, the use of the protein biosynthesis machinery in combination with an expanded genetic code and chemical modifications will presumably become the method of choice for glycoprotein synthesis.

**Construction of Modified mRNA for hARF1**

(Particpation in Vesicle Formation)

Amino Acid Sequence of hARF1

M	G	N	I	F	A	N	L	F	K	G	L	F	G	K	K	E	M	R	I	(20)
L	M	V	G	L	D	A	A	G	K	T	T	I	L	Y	K	L	K	L	G	(40)
E	I	V	T	T	I	P	T	I	G	F	N	V	E	T	V	E	Y	K	N	(60)
I	S	F	T	V	W	D	V	G	G	Q	D	K	I	R	P	L	W	R	H	(80)
Y	F	Q	N	T	Q	G	L	I	F	V	V	D	S	N	D	R	E	R	V	(100)
N	E	A	R	E	E	L	M	R	M	L	A	E	D	E	L	R	D	A	V	(120)
L	L	V	F	A	N	K	Q	D	L	P	N	A	M	N	A	A	E	I	T	(140)
D	(K)	L	G	L	H	S	L	R	H	R	N	W	Y	I	Q	A	T	C	A	(160)
T	S	G	D	G	L	Y	E	G	L	D	W	L	S	N	Q	L	R	N	Q	(180)
K																				

5'-AAG-3' →(Mutation)→ 5'-UAG-3'

Lys

Suppressor-tRNA → GlcNAc-Ser ←

Translation

Experimental	Radiography

Rabbit reticulocyte lysate

mRNA (UAG at Position 142)

supp-tRNA$^{\text{GlcNAc-Ser}}_{\text{AUC}}$

^{35}S-Met

tRNA's with all AA's

Ribosomes

Buffer

⊕ ⊕ ⊖ Ⓢ Ⓢ Ⓢ ⊖

← Background

Tmd-Phe / GlcNAc-Ser Protein
Truncated Protein
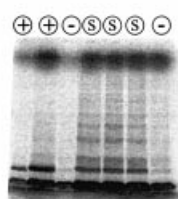
⊕ positive control with supp-tRNA$^{\text{Tmd-Phe}}_{\text{AUC}}$

⊖ negative control without supp-tRNA

Ⓢ with supp-tRNA$^{\text{GlcNAc-Ser}}_{\text{AUC}}$

Scheme 16

ACKNOWLEDGEMENTS

REFERENCES

1   R. R. Schmidt. *Angew. Chem.* **98**, 213–236 (1986); *Angew. Chem. Int. Ed. Engl.* **25**, 212–235 (1986); *Pure Appl. Chem.* **61**, 1257–1270 (1989).

2   R. R. Schmidt, W. Kinzy. *Adv. Carbohydr. Chem. Biochem.* **50**, 21–123 (1994).

3   R. R. Schmidt. *Pure Appl. Chem.* **70**, 397–402 (1998), and references therein.

4   P. Kovac. *Synthetic Oligosaccharides—Indispensable Probes for Life Sciences*, ACS Symposium Series 560. American Chemical Society, Washington, DC (1994); H. J. Gabius, S. Gabius. *Glycosciences: Status and Perspectives*, Chapman & Hall, Weinheim (1997), and papers therein.

5   J. C. Paulson. *Trends Biochem. Sci.* **14**, 272–276 (1989); A. Varki. *Glycobiology* **3**, 97–130 (1993); R. A. Dwek. *Chem. Rev.* **96**, 683–720 (1996).

6   H. Kunz. *Pure Appl. Chem.* **65**, 1223–1232 (1993); T. Mukaiyama, Y. Murai, S. Shoda. *Chem. Lett.* 431–434 (1981).

7   P. Fügedi, P. J. Garegg, H. Lönn, T. Norberg. *Glycoconjugate J.* **4**, 97–108 (1987); K. Toshina, K. Tatsuta. *Chem. Rev.* **93**, 1503–1531 (1993).

8   R. R. Schmidt. In *Modern Methods in Carbohydrate Synthesis* (S. H. Khan, R. A. O'Neill, eds), pp. 20–54. Harwood Academic Publishers GmbH, Chur (1996).

9   T. J. Martin, R. R. Schmidt. *Tetrahedron Lett.* **33**, 6123–6126 (1993); T. J. Martin, R. Brescello, A. Toepfer,

R. R. Schmidt. *Glycoconjugate J.* **10**, 16–25 (1993).

10   D. Kahne, S. Walker, Y. Cheng, D. V. Engen. *J. Am. Chem. Soc.* **111**, 6881–6882 (1989).

11   L. Lay, R. R. Schmidt. University of Konstanz, unpublished results.

12   C.-H. Wong, S. S. Hayne, G. M. Whitesides. *J. Org. Chem.* **47**, 5416–5418 (1982).

13   F. Barresi, O. Hindsgaul. *Modern Synthetic Methods* (B. Ernst, C. Leumann, eds), pp. 283–330, Verlag Helvetica Chimica Acta, Basel (1995); S. H. Khan, O. Hindsgaul. In *Molecular Biology: Frontiers in Molecular Biology* (M. Fukuda, O. Hindsgaul, eds), pp. 206–229 (1994).

14   D. M. Whitefield, S. P. Douglas. *Glycoconjugate J.* **13**, 5–17 (1996).

15   R. R. Schmidt. *Synthetic Oligosaccharides-Indispensible Probes for Life Sciences* (P. Kovac, ed.), pp. 276–296, ACS Symposium Series 560. American Chemical Society, Washington, DC (1994), and references therein.

16   T. J. Martin, G. Dufner, B. Kratzer, R. R. Schmidt. *Glycoconjugate J.* **13**, 547–553 (1996).

17   M. A. J. Ferguson, A. F. Williams. *Annu. Rev. Biochem.* **57**, 285–320 (1988), and references therein.

18   T. G. Mayer, B. Kratzer, R. R. Schmidt. *Angew. Chem.* **106**, 2289–2293 (1994); T. G. Mayer, B. Kratzer, R. R. Schmidt. *Angew. Chem. Int. Ed. Engl.* **33**, 2177–2181 (1994); T. G. Mayer, R. R. Schmidt. *Eur. J. Org. Chem.* (submitted).

19   C. Murakata, T. Ogawa. *Carbohydr. Res.* **234**, 75–91 (1992), in this approach the enantiomer of the naturally occurring inositol moiety was used.

20   A. S. Campbell, B. Fraser-Reid. *J. Am. Chem. Soc.* **117**, 10 387–10 388 (1995).

21   D. Tailler, V. Ferrieres, K. Pekari, R. R. Schmidt. *Tetrahedron Lett.* **40**, 679–682 (1999).

22   M. P. deNinnio. *Synthesis* 583–593 (1991); K. Okamoto, T. Goto. *Tetrahedron* **46**, 5835–5957 (1990).

23   H. Prabhanjan, K. Aoyama, K. Kiso, A. Hasegawa. *Carbohydr. Res.* **233**, 87–99 (1992), and references therein.

24   A. Marra, P. Sinaÿ. *Carbohydr. Res.* **195**, 303–308 (1990).

25   H. Paulsen. *Angew. Chem.* **94**, 184–201 (1982); *Angew. Chem. Int. Ed. Engl.* **21**, 155–197 (1982).

26   R. R. Schmidt, E. Rücker. *Tetrahedron Lett.* **21**, 1421–1424 (1980); R. R. Schmidt, J. Michel. *J. Carbohydr. Chem.* **4**, 141–169 (1985); R. R. Schmidt, M. Behrendt, A. Toepfer. *Synlett* 694–696 (1990).

27   R. R. Schmidt. Lecture on *'Fuji '90, Post Symposium of the XVth International Carbohydrate Symposium*, August 1990; see also R. R. Schmidt. In *Carbohydrates—Synthetic Methods and Applications in Medicinal Chemistry* (H. Ogura, A. Hasegawa, T. Suami, eds), pp. 66–88. Kodanasha Ltd, Tokyo (1992).

28   A. Hasegawa, T. Nagahama, H. Ohki, K. Kotta, H. Ishida, M. Kiso. *J. Carbohydr. Chem.* **10**, 493–498 (1991).

29   A. Toepfer. Dissertation, University of Konstanz (1992).

30   M. M. Sinn, H. Kondo, C.-H. Wong. *J. Am. Chem. Soc.* **115**, 2260–2267 (1993).

31   J. M. Lassaletta, K. Carlsson, P. J. Garegg, R. R. Schmidt. *J. Org. Chem.* **61**, 6873–6880 (1996); G. Hummel, R. R. Schmidt. *Tetrahedron Lett.* **38**, 1173–1176 (1997), and references therein.

32   T. J. Martin, R. R. Schmidt. *Tetrahedron Lett.* **34**, 1765–1768 (1993); T. J. Martin, H. Braun, R. R. Schmidt. *Bioorg. Med. Chem.* **2**, 1203–1208 (1994); G. Dufner. Dissertation, University Konstanz (1997).

33   This work was presented; due to space limitations it is not compiled here; for ref. see: C. Schaub, B. Müller, R. R. Schmidt. *Glycoconjugate J.* **15**, 345–354 (1998); F. Amann, C. Schaub, B. Müller, R. R. Schmidt. *Chem. Eur. J.* **4**, 1106–1115 (1998); B. Müller, C. Schaub, R. R. Schmidt. *Angew. Chem.*, in print.

34   G. Ritter, P. O. Livingston. *Seminars Cancer Biol.* **2**, 401–409 (1991).

35   W. B. Hamilton, F. Helling, K. O. Lloyd, P. O. Livingston. *Int. J. Cancer* **53**, 566–573 (1993).

36   K. Kitamura, P. O. Livingston, S. R. Fortunato, E. Stockert, F. Helling, G. Ritter, H. F. Oettgen, L. J. Old. *Proc. Natl Acad. Sci. USA* **92**, 2805–2809 (1995).

37   J. C. Castro-Palomino, G. Ritter, S. R. Fortunato, S. Reinhardt, L. J. Old, R. R. Schmidt. *Angew. Chem.* **109**, 2081–2085 (1997); *Angew. Chem. Int. Ed. Engl.* **36**, 1998–2001 (1997).

38   T. Stauch, U. Greilich, R. R. Schmidt. *Liebigs Ann.* 2101–2111 (1995); U. Greilich, R. Brescello, K.-H. Jung, R. R. Schmidt. *Liebigs Ann.* 663–672 (1996), and references therein.

39   M. R. E. Aly, J. C. Castro-Palomino, E. I. Ibrahim, E. H. El, Ashry, R. R. Schmidt. *Eur. J. Org. Chem.*, 2305–2316 (1998), and references therein.

40   M. Sugimoto, M. Numata, K. Koike, Y. Nakahara, T. Ogawa. *Carbohydr. Res.* **156**, C1–C5 (1986); A. Hasegawa, T. Nagahama, H. Ohki, M. Kiso. *J. Carbohydr. Chem.* **11**, 699–714 (1992).

41 W. Dullenkopf, J. C. Castro-Palomino, L. Manzoni, R. R. Schmidt. *Carbohydr. Res.* **296**, 135–147 (1996).

42 Y. Ito, M. Numata, M. Sugimoto, T. Ogawa. *J. Am. Chem. Soc.* **111**, 8508–8510 (1989); M. Numata, M. Sugimoto, Y. Ito, T. Ogawa. *Carbohydr. Res.* **203**, 205–217 (1990).

43 J. M. Lassaletta, R. R. Schmidt. *Synlett* 925–927 (1995); J. M. Lassaletta, M. Meichle, S. Weiler, R. R. Schmidt. *J. Carbohydr. Chem.* **15**, 241–254 (1996).

44 T. Murase, H. Ishida, M. Kiso, A. Hasegawa. *Carbohydr. Res.* **184**, C1–C4 (1988).

45 K.-A. Karlsson. *Annu. Rev. Biochem.* **58**, 309–350 (1989).

46 O. Kanie, M. Kiso, A. Hasegawa. *J. Carbohydr. Chem.* **7**, 501–506 (1988).

47 T. J. Martin, Dissertation, University of Konstanz (1994). Y. E. Tsvetkov, R. R. Schmidt, unpublished results.

48 Y. E. Tsvetkov, R. R. Schmidt. *Tetrahedron Lett.* **35**, 8583–8586 (1994); Y. E. Tsvetkov, R. R. Schmidt. *Carbohydr. Lett.* **2**, 149–156 (1996); J. C. Castro–Palomino, Y. E. Tsvetkov, R. Schneider, R. R. Schmidt. *Tetrahedron Lett.* **38**, 6837–6840 (1997).

49 This work was presented; due to space limitations it is not compiled; for ref. see: U. Huchel, R. R. Schmidt. *Tetrahedron Lett.*, **39**, 7693–7694 (1998).

50 K. Okamoto, T. Kondo, T. Gotto. *Tetrahedron* **44**, 1291–1298 (1988); Y. Ito, M. Numata, M. Sugimoto, T. Ogawa. *J. Am. Chem. Soc.* **111**, 8508–8510 (1989); T. Ercégovic, G. Magnusson. *J. Org. Chem.* **60**, 3378–3384 (1995); T. Kondo, T. Tomoo, H. Abe, M. Isobe, T. Goto. *J. Carbohydr. Chem.* **15**, 857–878 (1996).

51 J. C. Castro-Palomino, Y. E. Tsvetkov, R. R. Schmidt. *J. Am. Chem. Soc.* **120**, 5434–5440 (1998), and references therein.

52 J. C. Castro-Palomino, R. R. Schmidt. *Synlett*, 501–503 (1998).

53 H. G. von Garg, K. dem Bruch, H. Kunz. *Adv. Carbohydr. Chem. Biochem.* **50**, 277–310 (1994).

54 U. K. Saha, L. S. Griffith, J. Rademann, A. Geyer, R. R. Schmidt. *Carbohydr. Res.* **304**, 21–28 (1997).

55 J. Brunner. *Annu. Rev. Biochem.* **62**, 483–514 (1993).

56 D. Mendel, V. W. Cornish, P. G. Schultz. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 435–462 (1995).

57 S. A. Robertson, J. A. Ellman, P. G. Schultz. *J. Am. Chem. Soc.* **113**, 2722–2729 (1991).

58 T. Arslan, S. V. Mamaev, N. V. Mamaeva, S. M. Hecht. *J. Am. Chem. Soc.* **119**, 10 877–10 887 (1997).

59 F. Wieland, private communication.

60 U. K. Saha, R. R. Schmidt. *J. Chem. Soc., Perkin Trans* **1**, 1855–1860 (1997).