Chemical Synthesis of GPIs and GPI-Anchored Glycopeptides

Zhongwu Guo*[a] and Lee Bishop[a]

Dedicated to Professor Xikui Jiang

Keywords: Carbohydrates / GPI anchors / Glycolipids / Natural products / Glycopeptides

This microreview summarizes some recent progress in the study of glycosylphosphatidylinositols (GPI), focusing on the strategies established for the synthesis of optically pure *myo*-inositol derivatives and complex GPIs and GPI conjugates. Although in the past decade many GPIs have been achieved by linear, convergent and solid-phase synthesis, the GPIs with 2-*O*-acylated inositol residues and the GPI conjugates of peptides and glycopeptides have not been accomplished until recently. The successful coupling of peptides and glycopeptides to GPIs brings new promises to this area, but the scope of the relevant designs, in particular their application to more complex structures such as GPI-anchored glycoproteins, remains untested.

1. Introduction

Glycosylphosphatidylinositols (GPIs) are a large family of glycolipids that are expressed by all eukaryotic cells and have in common the core structure: Manα1→4GlcNH$_2$α1→6-*myo*-Ino1-OPO$_3$-lipid (**1**, Figure 1).[1] An essential function of GPIs is to anchor various extracellular molecules onto cell membranes, thus GPIs are involved in controlling the endocytosis and turnover of anchored structures and the transduction of transmembrane signals through their association with membrane microdomains on the cell surface.[2−4] Meanwhile, the high lateral mobility of GPIs and GPI-anchored structures significantly facilitates the selective release of molecules from cell surfaces as well as the exchange of membrane proteins between cells. It has also been suggested that GPI metabolites can act as secondary messengers in some hormonal pathways.[5,6] The functions of GPI-anchored molecules are even more diverse, ranging from coat proteins to receptors and enzymes.[3,7] Therefore, GPIs and GPI-anchored structures play a pivotal role in numerous biological and pathological events.

[a] Department of Chemistry, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106−7078, USA
Fax: (internat.) +1-216-368-3006
E-mail: zxg5@case.edu

Zhongwu Guo was born (1964) in China. After receiving his bachelor's degree in pharmacy (1984) and a MS in medicinal chemistry (1987) from the Second Military Medical University, he joined Professor Xikui Jiang at Shanghai Institute of Organic Chemistry (SIOC) to study physical organic chemistry. In 1989, he was selected to attend a joint training program sponsored by the Chinese and Polish Academies of Sciences and obtained his PhD in organometallic chemistry from the Institute of Organic Chemistry, Warsaw under the direction of Professor Alexander Zamojski. Dr. Guo returned to SIOC in 1991 to pursue a postdoctoral position and was appointed to the faculty there in 1994. He joined Professor Tomoya Ogawa at RIKEN, Japan, on a prestigious RIKEN Fellowship from 1996 to 1997 and then moved to Canada to serve at the National Research Council as an Assistant Research Officer. He has been a faculty member of chemistry and oncology at Case Western Reserve University since 1999 and received a Research Innovation Award from the Research Corporation in 2000. His research interests are mainly focused on glycoscience, medicinal chemistry and the synthetic studies of natural products.

Lee Bishop is in his seventh semester at Case Western Reserve University as an undergraduate BS Chemistry major with a minor in Biology. Lee is a President's Scholar and has received Dean's High Honors every semester in attendance. Lee is also a member of the Golden Key International Honour Society and is a 2004 recipient of the W. R. Veazey award for academic achievement in physical chemistry.

MICROREVIEWS: This feature introduces the readers to the author's research through a concise overview of the selected topic. Reference to important work from others in the field is included.

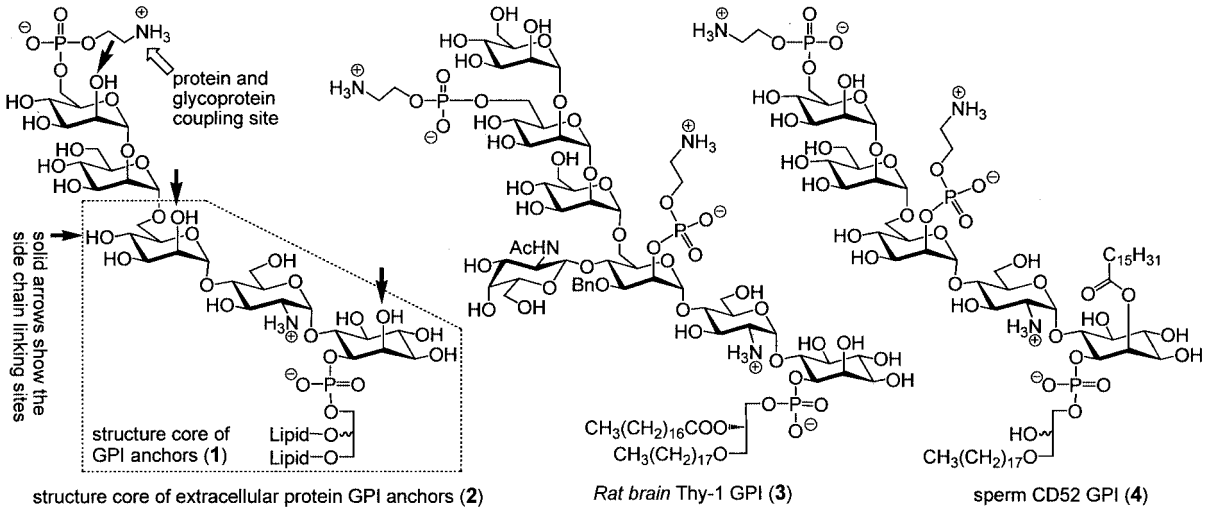


Figure 1. The structures of GPI cores and some typical membrane protein GPI anchors

The first GPI structural identification, achieved by Ferguson and co-workers in 1988, was of the GPI anchor of *T. brucei* variant surface glycoprotein (VSG).[8,9] To date, two classes of more than 50 GPIs have been elucidated. One class of GPIs, which are usually identified in protozoal parasites, share the basic core **1**.[10,11] They attach nonprotein bound extracellular glycoconjugates, for example, lipophosphoglycans (LPGs) and glycoinositol phospholipids (GIPL), onto parasite cell membranes. The second class of GPIs are membrane protein anchors, which have been identified in mammals and lower eukaryotes.[1,12,13] These GPIs have the conserved core: NH_2EtOPO_3-6Manα1→2Manα1→6Manα1→4GlcNH_2α1→6-*myo*-Ino1-OPO_3-lipid (**2**). The *C*-termini of extracellular proteins or glycoproteins are linked to the phosphoethanolamine group at the GPI nonreducing end via an amide bond.

In most cases, the GPI glycan core is further modified by one or more side chains that can be monosaccharides, oligosaccharides and phosphoethanolamine groups, as shown by the *Rat brain* Thy-1 GPI (**3**, Figure 1).[8,9] A number of GPIs, such as that of sperm CD52 (**4**),[14,15] contain an additional fatty acid attached to the inositol 2-O position to further help their association with membranes. It was observed that the 2-*O*-acylation of inositol renders the GPIs resistant to phosphatidylinositol (PI)-specific phospholipase C,[16,17] but its biological relevance is not yet clear. Nevertheless, it has been established that 2-*O*-acylation is important for the biosynthesis of GPIs.[2,18]

As with all types of natural glycoconjugates, each GPI anchor may be comprised of heterogeneous glycoforms. This heterogeneity can be further enhanced by variations of the lipid moieties. Thus, it is very difficult, if not impossible, to obtain homogeneous GPIs and GPI-linked proteins or glycoproteins from living cells. As a result, their chemical synthesis has attracted significant attention.

There are several excellent reviews about the structure, biosynthesis and biological functions of GPIs and GPI-linked structures.[2,7,19−22] This microreview will therefore only focus on the chemical aspect of GPI research, especially our synthetic study of the sperm CD52 antigen, and other significant progress made since the last review by Gigg and Gigg in 1997.[23]

2. Synthesis of GPIs

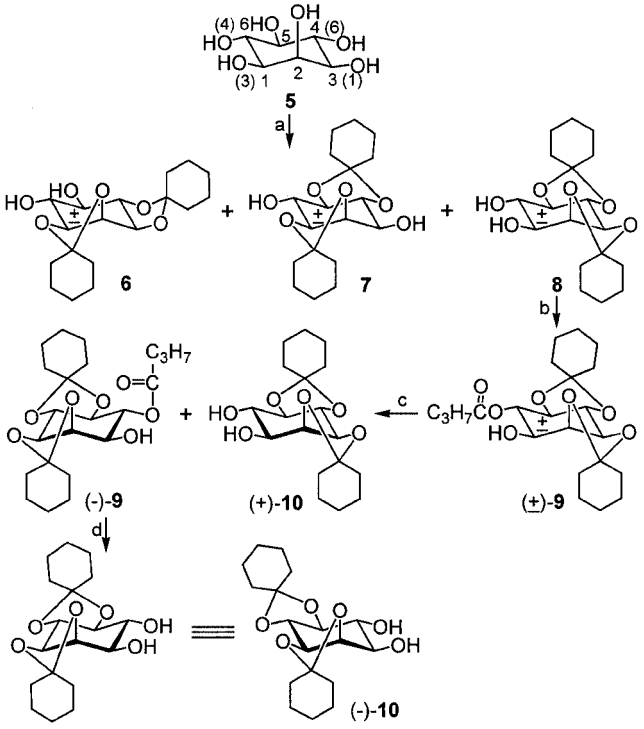
The first total synthesis of a GPI was achieved by Murakata and Ogawa in 1991,[24] who synthesized the membrane GPI anchor of *T. brucei*. Since then, a number of groups have synthesized several high-profile GPIs, such as the GPIs of *T. brucei* by Ley,[25,26] *Rat brain* Thy-1 by Fraser-Reid[27,28] and Schmidt,[29] *S. cerevisiae* and *T. gondii* by Schmidt,[30−32] *P. falciparum* by Fraser-Reid[33] and Seeberger,[34,35] and human sperm CD52 by Guo.[36,37] There have also been numerous synthetic studies of the partial and core structures of GPIs, many of which have been covered in Gigg and Gigg's review.[23]

GPI synthesis involves several different areas of chemistry, including carbohydrate, inositol, lipid and phosphate chemistry, the combination of which poses a significant challenge. Other problems that face GPI synthesis include the preparation of differentially protected and optically pure *myo*-inositol derivatives, the stereoselective assembly of the glycans, and the regioselective introduction of the side chains. The general strategies employed to tackle these problems are discussed separately below.

2.1 Synthesis of Discriminated Inositols

Owing to the fact that *myo*-inositol (**5**) is a symmetrical molecule (Scheme 1), its positions 1 and 4 are chemically identical to positions 3 and 6, respectively. However, derivatization of these sites will produce molecular chirality and form racemic mixtures. Thus, many studies have been directed at establishing methods to obtain enantiomerically pure *myo*-inositol derivatives, which have been recently reviewed by Sureshan and Shashidhar.[38] These methods can

be divided into three categories. One category is based upon the conversion of naturally occurring chiral inositol derivatives into the desired structures. The second category involves the resolution of synthetic enantiomers, which has been realized by both chemical and enzymatic approaches. The third category is de novo synthesis, that is, the preparation of enantiomerically pure *myo*-inositol derivatives from chirons. In addition, optically active inositols have also been synthesized by microbial oxidation of aromatic precursors.[39]

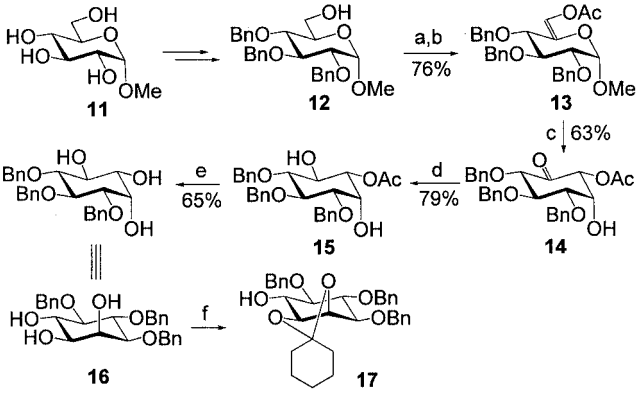


Scheme 1. Reagents and conditions: a) 1-ethoxycyclohexene or cyclohexanone, TsOH; b) Bu_2SnO, toluene, reflux; then $(C_3H_7CO)_2O$; c) cholesterol esterase or porcine pancreatic lipase; d) NaOMe, MeOH

The 1,6-*O*- and 1,2,6-*O*-differentiated *myo*-inositol derivatives used in GPI synthesis are prepared primarily from *myo*-inositol **5**, which involves enantiomeric resolution. Usually **5** is first transformed into di-*O*-cyclohexylidene derivatives **6**, **7** and **8** through the cyclohexylidenation method developed by Garegg[40] and Wyss[41] and their co-workers (Scheme 1). After separation of **6**, **7** and **8** by crystallization and chromatography, both **7** and **8** were used for further elaborations. Two strategies were established to resolve the enantiomers of the inositol derivatives. One is based on the linkage of inositol derivatives to a chiral reagent and separation of the resultant diastereoisomers, which has been successfully employed by Fraser-Reid,[27,28] Ley,[25,26] Martin-Lomas,[42,43] Ogawa,[24,44,45] Schmidt[30,46,47] and so on in their GPI syntheses. The other strategy, developed by Gou et al.,[48] makes use of enantioselective enzymatic reactions. When a racemic mixture of **9** was subjected to cholesterol esterase or porcine pancreatic lipase digestion, only one enantiomer, namely (+)-**9**, was deacylated.[48] The undigested enantiomer and the diol product could be readily separated by column chromatography to result in the optically pure (+)-**10** and (−)-**9**. Subsequent basic deacylation of (−)-**9** yielded the optically pure (−)-**10**. Our group has examined both methods and found that the chemical resolution is tedious whereas the enzymatic method can afford two enantiomers of **10** in excellent yields on large scales. Therefore, we adopted the latter method in our GPI synthesis.[36,49]

The 1,6-*O*-differentiated derivatives of inositol can be relatively easily prepared from optically pure substrates, such as (+)-**10**, because the 1,6-O positions have already been discriminated from the other positions. However, the differentiation of the remaining positions is relatively difficult, so it is more challenging to prepare 1,2,6-*O*-differentiated derivatives of inositol, which are necessary for GPIs with 2-*O*-acylated inositol residues. Several methods have been developed in this regard.

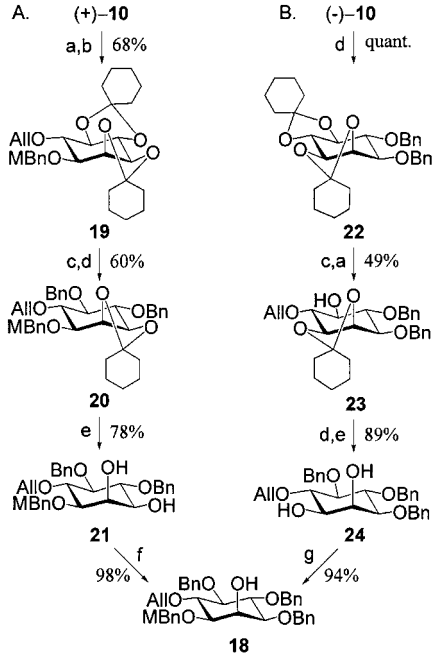
The method of Fraser-Reid and co-workers[39] involved the preparation of optically pure, differentially protected *myo*-inositols from methyl α-D-glucopyranoside (**11**, Scheme 2). Compound **11** was transformed into **12** in four steps, which was readily converted into enol acetate **13**. The key step in this synthesis was the stereocontrolled cyclization to afford the cyclic hexanone derivative **14** (63%) via the Ferrier reaction. Triol **16** was eventually formed following chelation-mediated reduction and then methanolysis to remove the acetyl group. In the synthesis of the *P. falciparum* GPI,[33] Fraser-Reid and co-workers converted **16** into cyclohexylidene derivative **17**, and after its glycosylation the protecting groups were further manipulated to achieve the differentiation of the 1-O- and 2-O positions.



Scheme 2. Reagents and conditions: a) *N*,*N*-dicyclohexylcarbodiimide (DCC), DMSO, trifluoroacetic acid (TFA), pyridine; b) Ac_2O, K_2CO_3, MeCN; c) $Hg(OAc)_2$, acetone, H_2O; d) $NaBH(OAc)_3$, AcOH, MeCN; e) NaOMe, MeOH; f) 1-ethoxycyclohexene, TsOH

Our approach for synthesizing 1,2,6-*O*-differentiated inositols started from *myo*-inositol to yield **18** on a multigram scale.[37,49] After (+)-**10** and (−)-**10** were prepared according to Scheme 1, the 1-O- and 6-O positions of (+)-**10** were differentiated by stannylene-mediated regioselective alkylation. Accordingly, an allyl group and a *p*-methoxybenzyl (MBn) group were attached to the 6-O- and 1-O positions,

respectively, to give **19** in 68% yield. The *trans*-cyclohexylidene moiety was selectively cleaved under mild acidic conditions, which was followed by benzylation to give **20**. The *cis*-cyclohexylidene group was then removed, and the resultant **21** was regioselectively benzylated under the influence of a stannylene complex to afford the desired product **18** in 32% overall yield (route A in Scheme 3). This synthesis was facilitated by repeated use of stannylene-mediated regioselective alkylations.

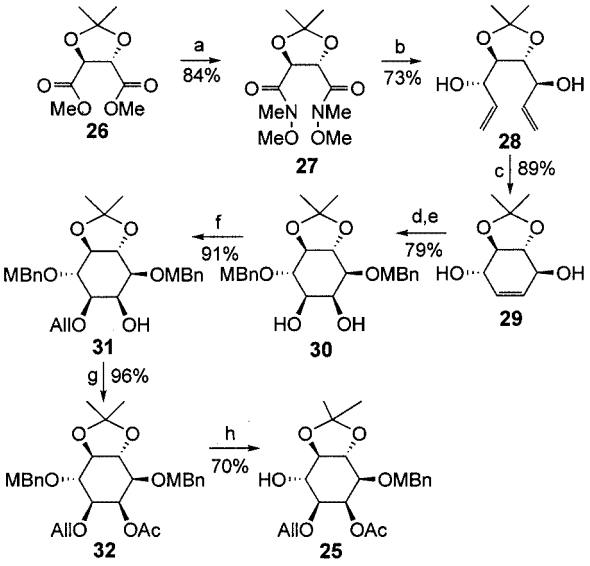


Scheme 3. Reagents and conditions: a) Bu_2SnO, toluene, reflux; then AllBr, DMF, room temp.; b) MBnCl, NaH, DMF, room temp.; c) AcCl, MeOH/CH_2Cl_2, room temp., 10 min; d) BnBr, NaH, DMF, room temp.; e) AcCl, MeOH/DCM (dichloromethane), room temp., 2 h; f) Bu_2SnO, toluene, reflux; then BnBr, DMF, room temp.; g) Bu_2SnO, toluene, reflux; then MBnCl, CsF, KI, DMF, room temp.

In most organic syntheses, after the resolution of a racemic mixture only one enantiomer is useful, and one half of the starting material is wasted. If the normally wasted enantiomer finds its use in the preparation of the same target molecule, the synthetic efficiency can be significantly improved. With this consideration in mind, we designed a route to prepare **18** from (−)-**10** (route B in Scheme 3).

Since the molecular chirality of *myo*-inositol derivatives is a result of derivatizations of various positions and the absolute stereochemistry of the inositol ring is not affected, we can expect to obtain the same synthetic target from enantiomers if the protecting groups are manipulated in the reverse order. Accordingly, (−)-**10** was converted into **18** in 42% overall yield (route B in Scheme 3) by means of similar chemistry to that used for the conversion of (+)-**10** into **18**.

Bertozzi and co-workers[50] have recently reported another elegant method to prepare optically pure 1,2,3,6-O-differentiated *myo*-inositol (**25**) from dimethyl 2,3-*O*-isopropylidene-D-tartrate (**26**, Scheme 4). It took only 8 steps and gave 25% overall yield. The key steps in this synthesis were the ring-closing of **28**, effected by Grubbs' catalyst, and subsequent stereoselective dihydroxylation of the resultant **29** to afford the *myo*-inositol derivative **30**.

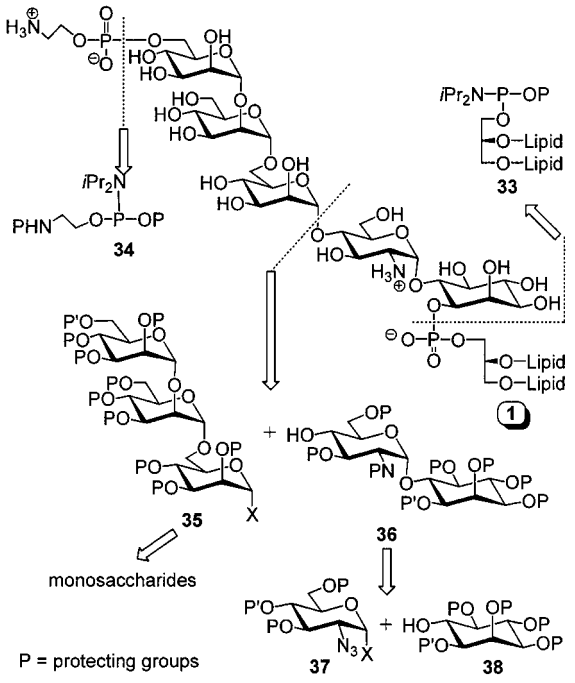


Scheme 4. Reagents and conditions: a) MeNHOMe·HCl, $AlMe_3$, DCM, 10 °C; b) CH_2=$CHMgBr$, THF; then $CeCl_3$, $NaBH_4$, MeOH; c) $Ru_2Cl_2CHPhPCy_3IMesH_2$, DCM, reflux; d) MBnCl, $BnEt_3N^+Cl^-$, 50% KOH/H_2O, toluene; e) K_2OsO_4, K_2CO_3, $K_3Fe(CN)_6$, methanesulfonamide, quinuclidine, *t*BuOH, H_2O; f) Bu_2SnO, toluene, reflux; then AllBr, $Bu_4N^+I^-$; g) Ac_2O, DMAP, *i*Pr_2NEt, DCM; h) DDQ, DCM

Of the three synthetic methods for 1,2,6-*O*-differentiated inositols discussed above, the synthesis of Fraser-Reid and co-workers is highlighted by the readily predictable regioselectivity and absolute stereochemistry of every reaction and intermediate involved, while our method is marked by the use of two enantiomers to prepare the same target. Bertozzi and co-workers' method is probably the most efficient to date. Nevertheless, because the products obtained from these syntheses possess different protecting groups, together they can meet the demands of various synthetic designs.

2.2 Synthesis of GPIs without the Additional Acyl Group on Inositol

Careful planning of protection tactics for various functional groups and strategies for assembling the residues and side chains is critical for the success of a GPI synthesis. Normally the phosphate groups are introduced at the final stage by the reactions of the appropriately deprotected oligosaccharides and various *N*,*N*-diisopropylphosphoramidites, for example, **33** and **34**, followed by oxidation using either *m*CPBA or *t*BuO_2H, as shown in Scheme 5. In the retrosynthesis of the oligosaccharide core, the α-glycosidic linkage between the mannose and glucosamine residues is usually disconnected first to give two logical building blocks **35** and **36**, safeguarding an easy and stereoselective coupling reaction at the final stage for convergent syntheses.

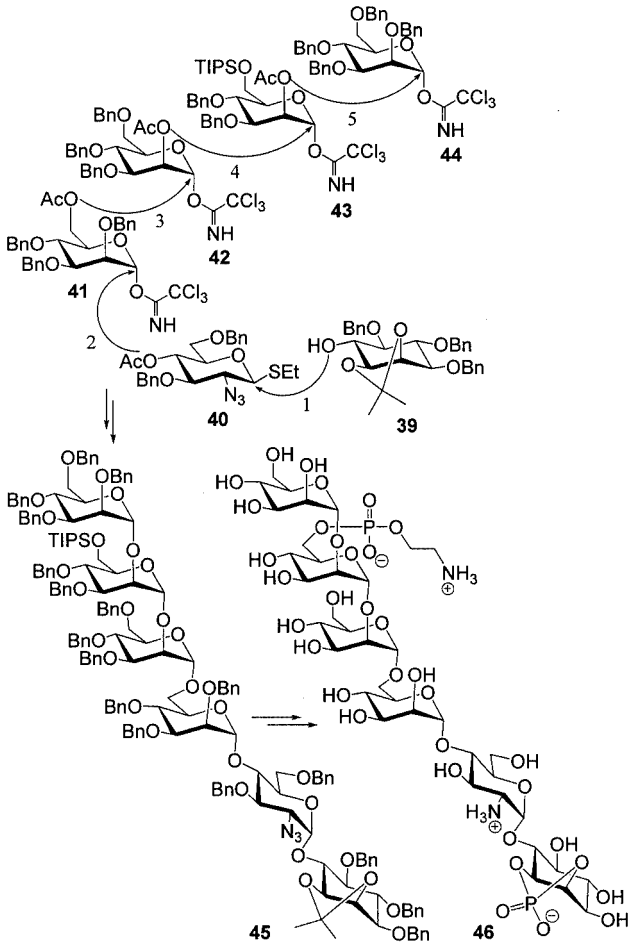


Scheme 5. Retrosynthetic scheme of GPI synthesis

The α-linked oligomannose **35** should be easily prepared, but the synthesis of pseudodisaccharide **36** is challenging. To assist the formation of the α-glycosidic linkage of glucosamine, its amino group needs to be protected by a non-participating group. Nearly all GPI syntheses employ 2-azido-2-deoxy-D-glucose (**37**) as a latent glucosamine in the glycosylation of inositol, and the azido group also serves as a "protected amino group". Various methods, including the Koenigs–Knorr method and glycosylations using glycosyl fluorides, thioglycosides, *n*-pentenyl glycosides and frequently trichloroacetimidates as the glycosyl donors, have been utilized to achieve this coupling. However, none of these methods gave particularly satisfactory yields and stereoselectivity. As a result, the coupling of glucosamine and inositol was performed at the initial stage in all syntheses with the exception of Murakata and Ogawa's synthesis of *T. brucei* GPI.[24,44,45] Fraser-Reid and co-workers[33] employed a very different strategy to deal with this problem, which will be discussed later.

On the other hand, starting from pseudodisaccharide **36**, GPI cores and intact GPIs have been prepared by linear elongation or convergent assembly of the carbohydrate chain. An example of the former is Seeberger and co-workers' synthesis of the GPI core of *P. falciparum*,[35] and an example of the latter is Ley and co-workers' synthesis of the *T. brucei* GPI.[26]

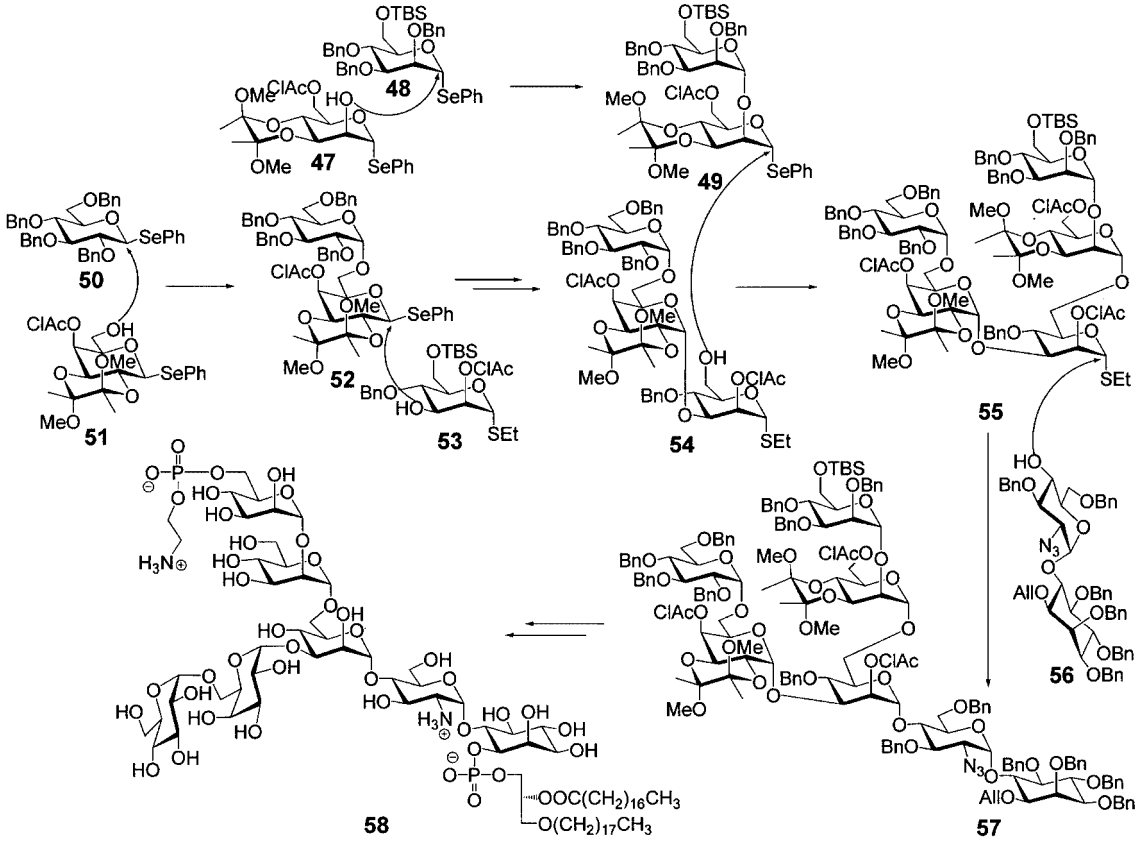
In Seeberger's synthesis of the GPI core of *P. falciparum* (Scheme 6),[35] after the glycosylation reaction between **39** and **40** to form the pseudodisaccharide, monosaccharide units **41**, **42**, **43**, and **44** were introduced sequentially by the Schmidt glycosylation method to afford the GI core **45**. The 1,2-O positions of inositol **39** were protected as an acetal, which can be replaced with a cyclic phosphate, while the 6-O position of **43** was differentially protected by a silyl group to facilitate its selective deprotection and the introduction of a phosphoethanolamine branch to this site later on. The GPI was then linked to a carrier protein to form an effective antimalaria vaccine.



Scheme 6. Linear assembly of the GPI core of *P. falciparum*

Ley's synthesis of *T. brucei* GPI was highly convergent (Scheme 7).[26] After appropriately protected monosaccharides were obtained, they were converted into short oligosaccharides **49** and **54**. The selective coupling between these two fragments afforded a pentasaccharide **55**, which was then used to glycosylate the pseudodisaccharide **56** to give the core **57**. Final stage manipulation of the protecting groups and selective introduction of the phosphoethanolamine and phospholipid groups, followed by global deprotection, eventually gave the synthetic target *T. brucei* GPI **58**.

Another important feature of this synthesis is the use of different glycosyl donors, as well as the use of 1,2-diacetal protecting groups to tune the reactivities of glycosyl donors to facilitate the "armed-disarmed" and semiorthogonal glycosylation methods for glycan construction. As a result, a few deprotection-activation steps could be avoided.



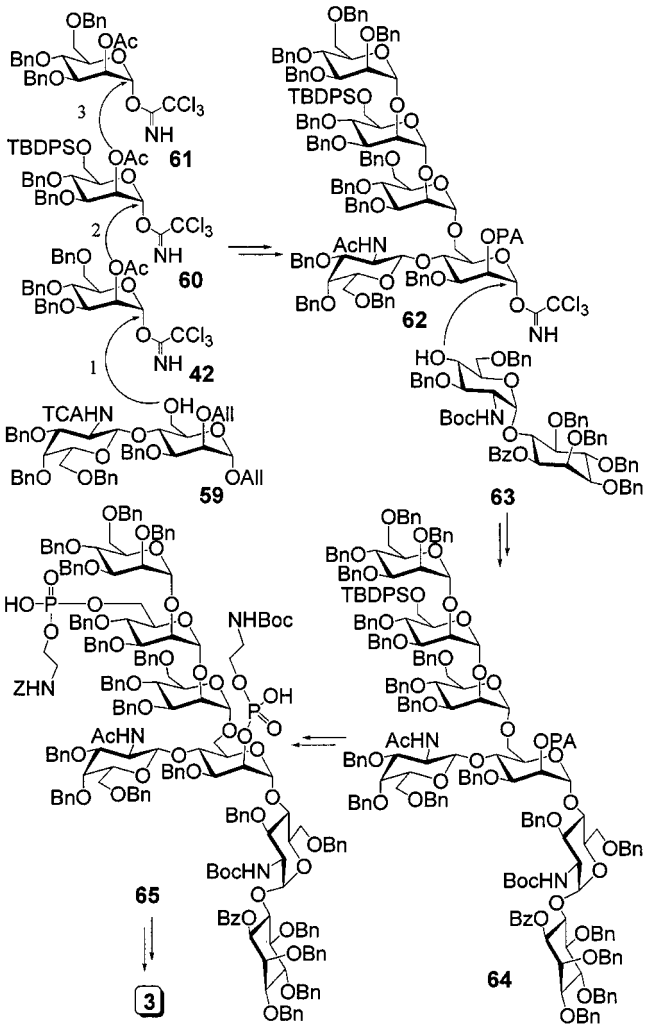
Scheme 7. Highly convergent synthesis of the intact GPI of *T. brucei*

Recently, Pekari and Schmidt[29] introduced a variable concept for the synthesis of branched GPI, which was demonstrated by the synthesis of the *Rat brain* Thy-1 GPI (Scheme 8). In this synthetic design, some branches, such as the sugar unit linked to the reducing-end mannose, were introduced at an early stage (**59**), while an array of orthogonal protecting groups were employed to block the monosaccharide units, for example, **59** and **60**, to which branches are usually attached. After sequential glycosylation of **59** by **42** and **60** to obtain a tetrasaccharide, a mannose branch and a phenoxyacetyl protecting group were attached to the 2-O position of the nonreducing- and reducing-end mannose residues to give the pentasaccharide fragment **62**. Subsequently, **62** was coupled with the pseudodisaccharide **63** to afford the branched glycan core **64**. Final stage manipulation of the protecting groups and regiospecific introduction of the phosphoethanolamine groups and the phospholipid moiety were followed by global deprotection to give the target **3**. This convergent design should be useful in the synthesis of a variety of branched GPIs. Moreover, as the two phosphoethanolamine moieties were differently protected, it is possible to selectively deprotect the one at the nonreducing end and attach a peptide chain to it.

In addition to the solution-phase synthesis of GPIs and related structures by the strategies discussed above, GPI glycans have also been prepared by solid-phase synthesis. For example, Seeberger and co-workers[34] assembled a tetramannose with an octenediol-functionalized Merrifield resin as the solid-phase support on an automatic oligosaccharide synthesizer. The oligosaccharide was released from the polymer support by Grubbs' metathesis and was obtained in 44% overall yield. The tetrasaccharide was then linked to the pseudodisaccharide in solution to afford the GPI core in 32% yield.

In a recent report on the solid-phase synthesis of the GPI core **70** by Reichardt and Martin-Lomas[51] (Scheme 9), the pseudodisaccharide was prepared first in solution and was then attached to a polyethylene glycol (PEG)-grafted polystyrene resin. This approach was employed to avoid solid-phase formation of the most problematic glycosidic linkage and to facilitate the final-stage purification. The carbohydrate chain of **66** was elongated by sequential introduction of monosaccharides **67**, **68**, and **69** by Schmidt's glycosylation method. The resultant **70** was released from the resin and isolated in 20% overall yield.

The convergent synthetic strategy has been much more commonly used in the preparation of GPIs and related structures, and Schmidt's variable synthetic design can be useful for the synthesis of a variety of branched GPIs.[29] Alternatively, as the linear assembly of the carbohydrate chain is straightforward and its protection tactics are simpler, it is more easily adopted in solid-phase synthesis, which is currently of great interest and potential. However, unless the glycosylation reactions give especially high yields



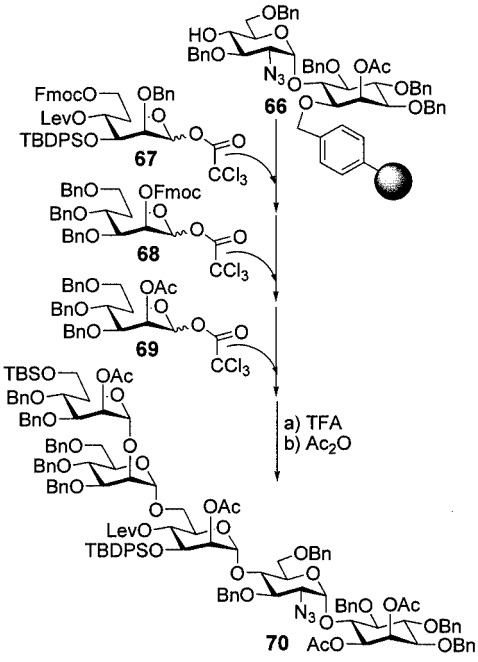
Scheme 8. A variable concept for the synthesis of the intact GPI of *Rat brain* Thy-1

and stereoselectivity, the purification of the target molecules can be difficult in solid-phase synthesis.

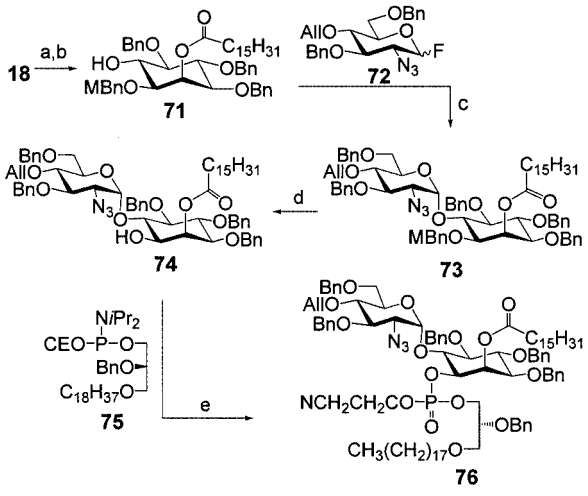
2.3 Synthesis of GPIs with 2-*O*-Acylated Inositol

The synthesis of GPIs with 2-*O*-acylated inositols represents another significant challenge. The presence of a large acyl group at the inositol 2-O position makes the already crowded inositol ring more sterically hindered, which may give rise to additional problems in the synthesis. There had been no report on the synthesis of this kind of GPI before our synthesis of sperm CD52 GPI.[36]

In the synthesis of sperm CD52 GPI, one of our concerns was whether the phospholipid moiety could be introduced at the inositol 1-O position if we decided to adopt the traditional synthetic approach, that is, by introducing the phospholipid at the final stage. A model reaction was thus performed on **74** to explore this possibility (Scheme 10).[49] After pseudodisaccharide **73** was obtained by coupling **72** to **71**, the 1-O position was selectively deprotected. The reaction of **75** and **74** gave the expected phospholipidation product **76** in good yield (80%).



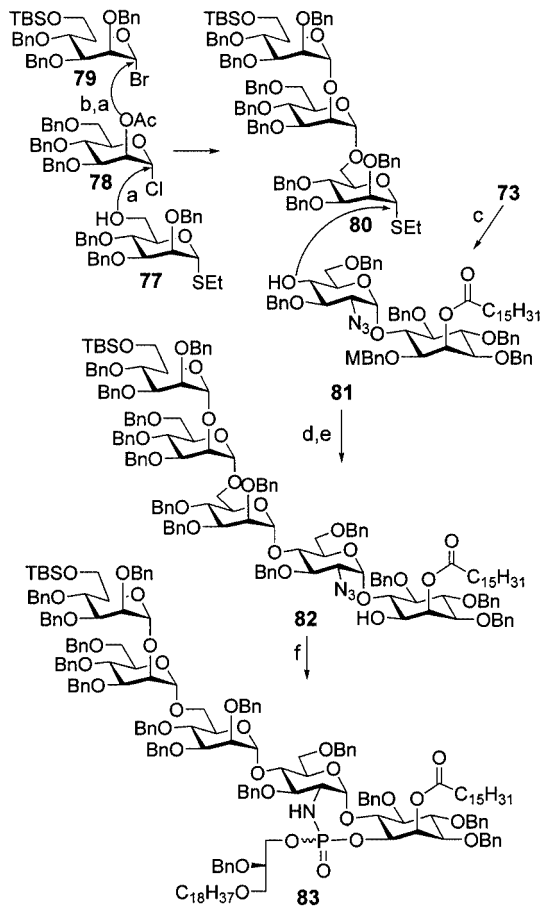
Scheme 9. Solid-phase assembly of a GPI core



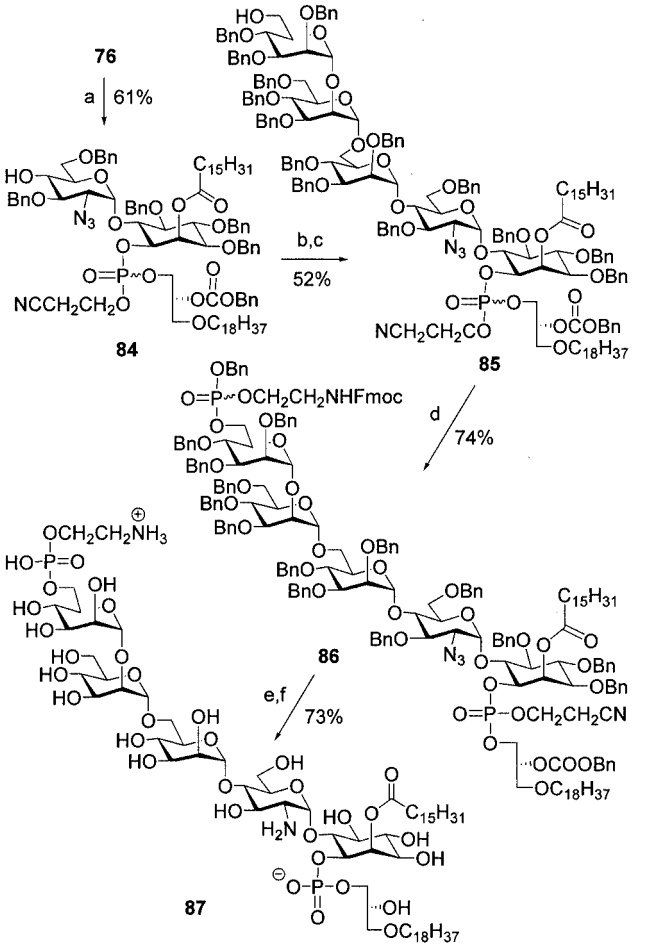
Scheme 10. Reagents and conditions: a) DCC, DMAP, $C_{15}H_{31}COOH$, DCM; b) $PdCl_2$, AcOH; c) Cp_2HfCl_2, AgOTf, DCM, TBSCl, Pyr; d) CAN, benzene/H_2O; e) tetrazole; then tBuO$_2$H

After finding out that the acyl group had little affect on the phosphorylation of **74**, we intended to synthesize the sperm CD52 GPI by the procedure shown in Scheme 11.[36] The trimannose fragment **80** was thus coupled to pseudodisaccharide **81** to obtain the glycan core. However, upon selective deprotection of its 1-O position (→**82**) and treatment with phosphorylation reagent **75**, instead of the phosphate, an unexpected cyclic phosphoramidate **83** was isolated in very good yield, indicating the involvement of the azido group. This result was surprising, as it had never been observed in any other GPI synthesis. Because this side reac-

tion was not observed in the model reaction (Scheme 10) either, the trimannose segment must have played a role in causing the problem.



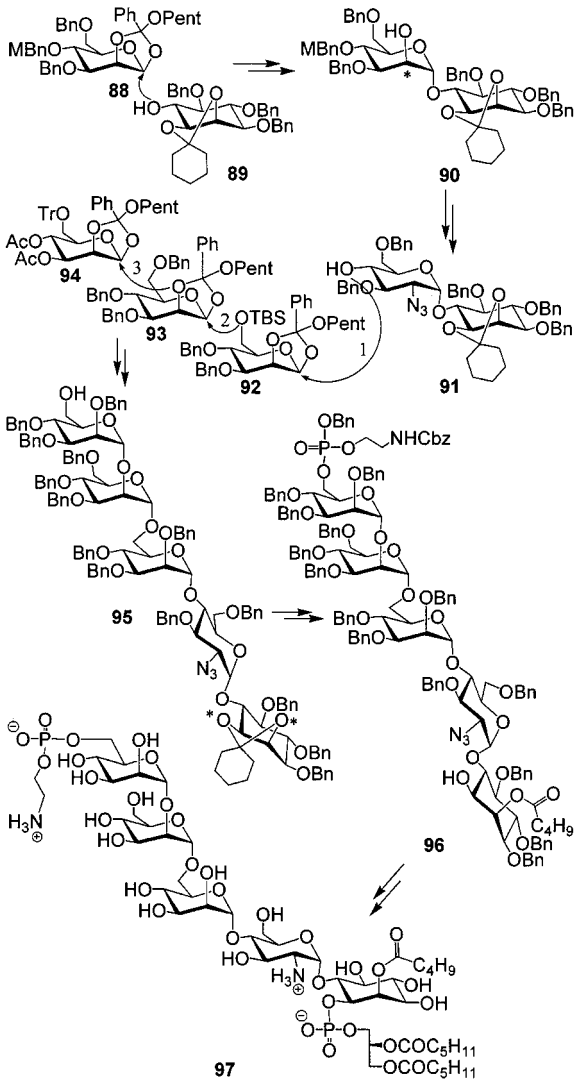
Scheme 11. Reagents and conditions: a) AgOTf, DCM; b) NaOMe, MeOH; c) $PdCl_2$, AcOH; d) NIS, TfOH, DCM; e) CAN, benzene/H_2O; f) **75**, tetrazole; then tBuO$_2$H.



Scheme 12. Reagents and conditions: a) $PdCl_2$, AcOH; b) **80**, NIS, TfOH; c) $Et_2O{\cdot}BF_3$; d) FmocNHCH$_2$CH$_2$OP(OBn)NiPr$_2$, tetrazole; then tBuO$_2$H; e) DBU; f) 10% Pd/C, H_2

The synthesis of sperm CD52 GPI was finally achieved by following the redesigned strategy shown in Scheme 12. A focal point of this design was to introduce the phospholipid at an early stage, that is, by using **76** as a key building block.[36] After the allyl group of **76** had been selectively removed, the trimannose segment **80** was attached to give the phospholipidated pentasaccharide, which was then treated with trifluoroborane to remove the silyl group to afford **85**. The reaction to introduce a phosphoethanolamine group to the free hydroxy group of **85** was easily carried out, and subsequent global deprotection in two steps eventually produced the synthetic target **87**, a GPI of sperm CD52 without the optional phosphoethanolamine group on the reducing-end mannose.

Fraser-Reid and co-workers[33] have recently synthesized another 2-*O*-acylated GPI that is a prototype membrane GPI anchor of *P. falciparum* (Scheme 13). The GI core **95** was constructed by the linear elongation of the glycan of the inositol derivative **89**, and all glycosylations were achieved by using *n*-pentenylglycosides as glycosyl donors. This synthesis has several intriguing features. Firstly, a novel strategy was utilized to tackle the problem associated with the creation of an α-glycosidic linkage between the inositol and glucosamine residues. To secure selective α-glycosylation, a mannose derivative **88** was employed as the glycosyl donor. The mannose residue was then converted into an azido derivative of glucosamine through an S_N2 reaction at C-2′. Secondly, after deprotection of the inositol's 1,2-O positions, its 2-O position was selectively acylated by the discriminatory ring opening of an ortho ester intermediate. Remarkably, the phosphorylation reactions did not interfere with the azido group of **95** and **96**, in contrast with our observations.[36] The major structural difference between the intermediates involved in Scheme 13 and the intermediates involved in our synthesis (Schemes 11 and 12) is that longer lipids were used in the latter. However, the reason why short lipids were chosen for this synthesis is not evident from the literature.[33]



Scheme 13. Synthesis of a fully lipidated and phosphorylated prototype GPI of *P. falciparum*

3. Synthesis of GPI-Anchored Glycopeptides and Proteins

The availability of GPI-linked conjugates is currently of primary interest, but there has been very little research in this area. The first chemical synthesis of a GPI-anchored structure was reported by Seeberger and co-workers[35] in his study of a GPI-based antimalaria vaccine. After the malarial GPI toxin was prepared according to the methods described in Scheme 6, it was linked to a carrier protein to form a glycoconjugate that has been proved to induce protective immune responses in animals. However, the linkage between the GPI and the carrier protein was artificial.

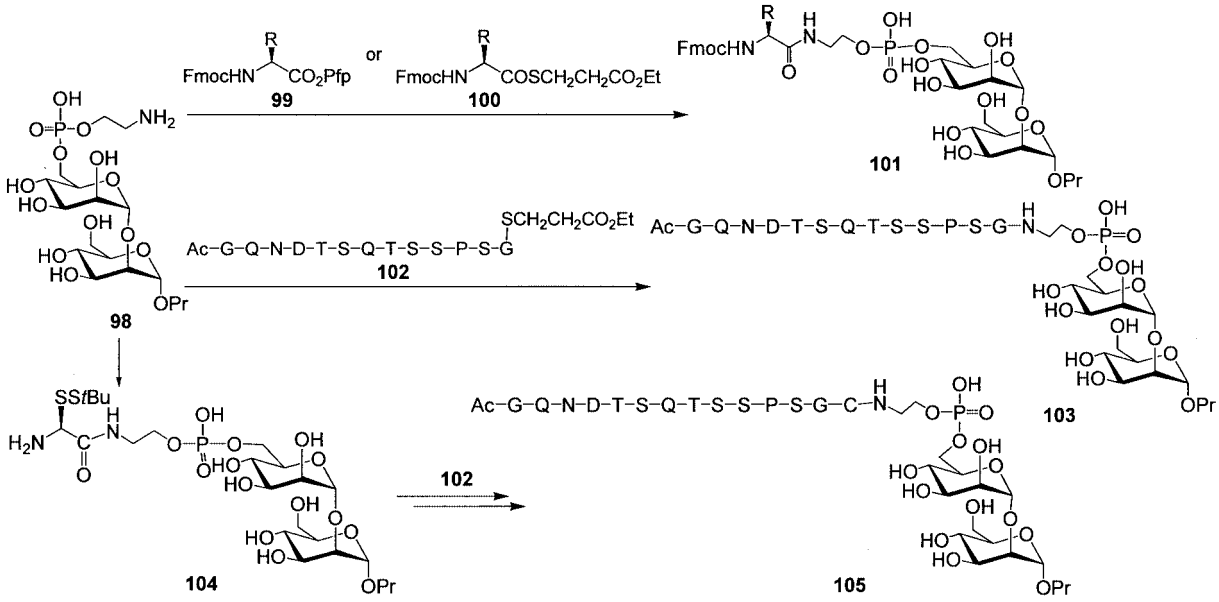
Nakahara and co-workers[52] have recently studied the attachment of free amino acids and peptides to GPIs through a native amide bond using disaccharide **98** as a significantly simplified GPI model (Scheme 14). The reaction between **98** and amino acid pentafluorophenyl esters **99** gave about 60% yields of the desired products **101**, while *O*-acylation was also observed as a side reaction in most cases. The reaction of amino acid thioesters **100** using $AgNO_3$ as a promoter gave **101** in good yields (76%) without the side reaction. However, when the method was applied to a free peptide **102**, the reaction was less effective, though relatively clean, affording the coupling product in a much reduced yield (34%). This work also involved a study of the native ligation. For this purpose, **98** was converted into the cysteine conjugate **104**, which was then treated with the thioester **102**. This reaction was much more efficient and gave the expected product in 78% yield. These results indicate that the direct coupling of free peptides or proteins to GPIs may not be feasible, but the native ligation method is potentially useful for addressing this problem.[57,58]

We have recently accomplished the first chemical synthesis of a GPI-linked peptide[37] and a natively linked GPI-glycopeptide conjugate **114**,[53] which is the skeleton of the sperm CD52 antigen[15] (Scheme 15). The first report was focused on establishing the proper protection tactics and coupling methods for the GPI-linked glycopeptide synthesis.[37] The reaction between a fully protected GPI and a short free peptide gave excellent results by using *N*-hydroxybenzotriazole (HOBt) and DCC or 1-[3-(dimethylamino)-propyl]-3-ethylcarbodiimide hydrochloride (EDC) as the coupling reagents. Therefore, a fully protected GPI and a fully protected glycopeptide were used in the synthesis of **114**.[53] To achieve the fully protected glycopeptide **109**, a 2-chlorotrityl resin with an extremely acid-sensitive linker was used in the solid-phase glycopeptide synthesis (**106** → **108**). The glycopeptide was retrieved from the resin by treatment with 10% acetic acid, which did not affect either the glycosidic linkages or the protecting groups, including the amino acid side-chain protections. The GPI was prepared by the convergent assembly method discussed above to afford **113**, which has an unprotected phosphoethanolamine group linked to the glycan nonreducing end.

The coupling of **109** to **113** was performed in DCM and *N*-methylpyrrolidinone (NMP) (2:1) under a nitrogen atmosphere using HOBt and EDC as condensation reagents.[54] The reaction was very efficient (70%) with the product easily purified by column chromatography. The global deprotection was achieved in two steps, namely, catalytic hydrogenolysis to remove benzyl groups and treatment with 15% TFA/DCM to remove the peptide side-chain protections. It was important to follow this deprotection sequence, as the glycosidic linkage of free fucose is stable to dilute TFA but that of the benzylated fucose is not.[55,56] The synthetic target **114** was finally purified by reversed-phase HPLC.

4. Conclusions

Since the first chemical synthesis of a GPI by Murakata and Ogawa in 1991,[24] this area has witnessed great progress. For example, a number of GPIs and several GPI-anchored structures have been synthesized, and many useful strategies related to these syntheses have been established, which in-



Scheme 14. Coupling between free amino acids or peptides and a GPI model

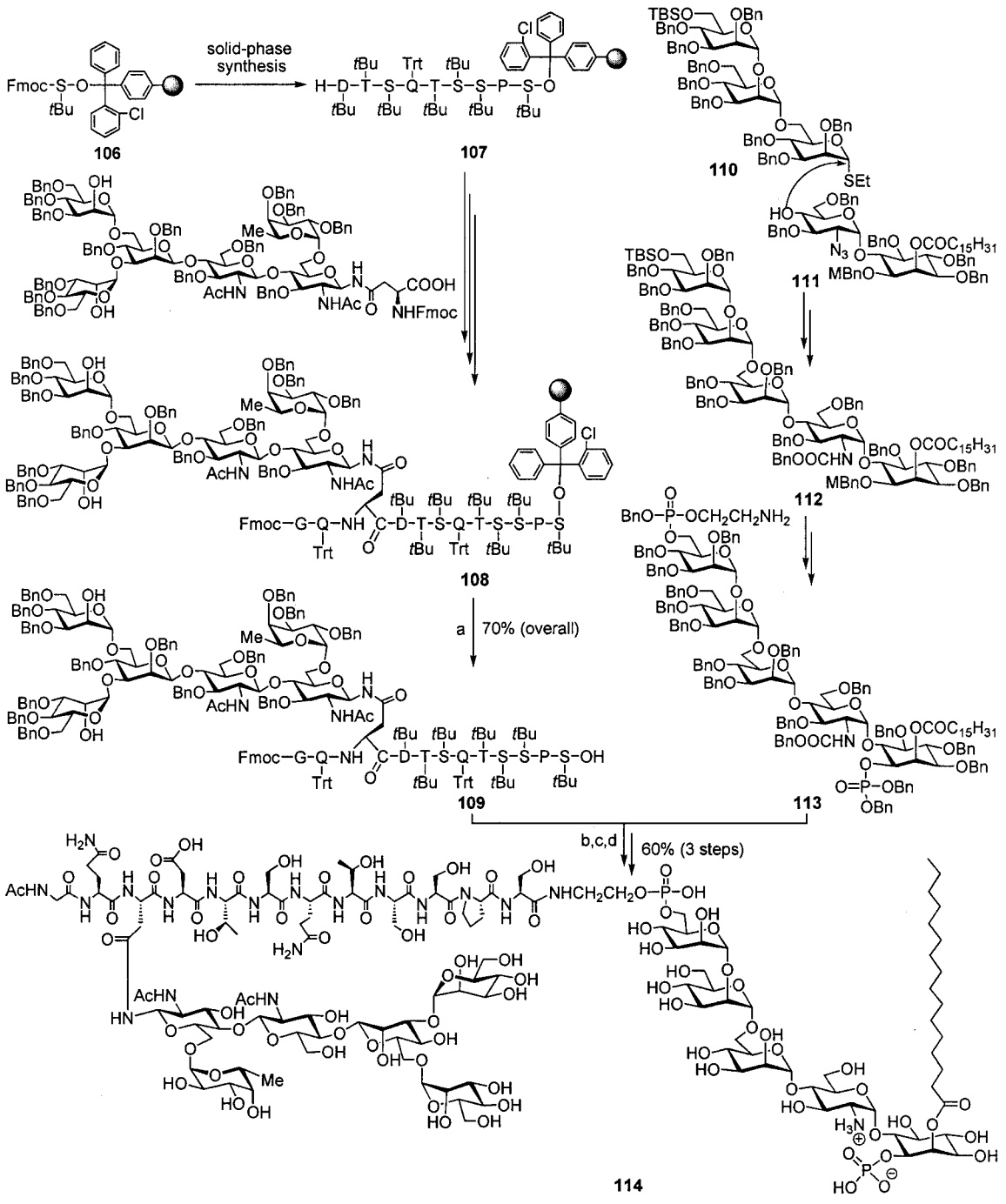
clude tactics for the positional discrimination of GPI glycans to enable the regioselective introduction of side chains, various strategies for GPI assembly, and strategic designs to facilitate the attachment of peptides or glycopeptides to GPIs. If the structural features involved are highly conserved, for example, the amide linkage between peptides and GPIs, the relevant strategies may be of general application. However, this by no means implies that problems associated with the chemical synthesis of GPIs and GPI-anchored structures have been solved or that the syntheses are routine practice. In fact, each additional side chain or the positional change of a side chain, even a difference in the size of a side chain, may result in the redesign of protection tactics and synthetic strategies. This is clearly illustrated by the problems encountered in the synthesis of 2-*O*-acylated GPIs,[36] as well as by the different behaviors of intermediates involved in the preparation of **76**, **87** and **97**.[33,36]

With regard to GPI research we believe that several unresolved issues shown below need special attention in the future. One significant challenge is the effective and stereoselective construction of the α-glycosidic linkage between the glucosamine and the inositol residues. Another challenge is the chemical synthesis of GPIs having unsaturated lipid chains. This is largely defined by the difficulty of establishing appropriate protection tactics. For example, acetyl and benzyl groups are among the most commonly used permanent protecting groups, but they are not useful in the synthesis of this kind of GPI because the usual conditions employed to deprotect acetyl and benzyl groups may affect the ester bonds and the carbon−carbon double bonds of the lipid chains, respectively. The third challenge is the synthesis of natively linked GPI-anchored proteins and glycoproteins. Although conjugates with GPIs linked to a short peptide or glycopeptide have been achieved synthetically, attaching a large protein or glycoprotein to GPIs is much more complex. A hopeful solution to this problem may be the native ligation method,[57,58] as indicated by the work of Nakahara and co-workers.[52] However, there is no established method to prepare GPI-linked peptides with a cysteine at the peptide N terminus. The fourth challenge is the structural study of GPIs. It is almost impossible to obtain single crystals of GPIs or their conjugates. Meanwhile, it is extremely difficult to study GPIs by NMR methods because of their solvation behavior. GPIs are almost insoluble in most organic solvents, and even if some GPIs are soluble in solvents such as DMSO, the samples usually give broad signals, making their analysis difficult. In water, GPIs tend to form micelles, which can further broaden the NMR signals. To circumvent this problem, in the structural elucidation of GPIs the target molecules are usually treated with PI phospholipase to release the lipid components, so that the lipids and the glycans can be analyzed separately. However, these studies cannot give much information about the conformations of natural GPIs or GPI-anchored structures.

Nevertheless, with the increase in the number of identified GPIs and GPI-anchored proteins and glycoproteins and with the accumulation of knowledge about their biological functions, it is expected that in future years there will be tremendous growth in the research of GPIs, especially with regard to their chemical synthesis, structural studies and biomedical applications.

Acknowledgments

Z. G. thanks his devoted students and research associates, especially Dr. Jie Xue and Mr. Ning Shao, who have made great contributions to the synthetic studies of sperm CD52 antigen. We appreciate the financial support of our research from the US Natural Science Foundation (CHE-0407144), the National Insti-



Scheme 15. Reagents and conditions: a) HOAc/trifluoroethanol (TFE)/DCM (1:1:8), room temp.; b) HOBt, EDC, DCM/NMP (2:1), room temp.; c) 10% Pd/C, H_2, $CHCl_3$/MeOH/H_2O (10:10:3), room temp.; d) $HSiEt_3$/TFA/DCM (1:1.5:7.5), room temp.

tute of Health (R01-CA95142) and the Petroleum Research Funds of the American Chemical Society (ACS PRF 37743-G).

[1] M. A. J. Ferguson, A. F. Williams, *Ann. Rev. Biochem.* **1988**, *57*, 285−320.
[2] J. R. Thomas, R. A. Dwek, T. W. Rademacher, *Biochemistry* **1990**, *29*, 5413−5422.
[3] P. J. Robinson, *Cell Biol. Intern. Rep.* **1991**, *15*, 761−767.
[4] D. D. Eardley, M. E. Koshland, *Science* **1991**, *251*, 78−81.
[5] M. C. Field, *Glycobiology* **1997**, *7*, 161−168.
[6] G. N. Gaulton, J. C. Pratt, *Semin. Immunol.* **1994**, *6*, 97−104.
[7] K. Hwa, *Adv. Exp. Med. Biol.* **2001**, *491*, 207−214.
[8] S. W. Homans, M. A. J. Ferguson, R. A. Dwek, T. W. Rademacher, R. Anand, A. F. Williams, *Nature* **1988**, *333*, 269−272.
[9] M. A. J. Ferguson, S. W. Homans, R. A. Dwek, T. W. Rademacher, *Science* **1988**, *239*, 753−759.
[10] M. A. Ferguson, J. S. Brimacombe, S. Cottaz, R. A. Field, L. S. Guther, S. W. Homans, M. J. McConville, A. Mehlert, K. G. Milne, J. E. Ralton, *Parasitology* **1994**, *108*, S45−54.
[11] M. J. McConville, M. A. J. Ferguson, *Biochem. J.* **1993**, *294*, 305−324.
[12] P. T. Englund, *Ann. Rev. Biochem.* **1993**, *62*, 121−138.

[13] S. Udenfriend, K. Kodukula, *Ann. Rev. Biochem.* **1995**, *64*, 563−591.
[14] A. Treumann, M. R. Lifely, P. Schneider, M. A. J. Ferguson, *J. Biol. Chem.* **1995**, *270*, 6088−6099.
[15] S. Schroter, P. Derr, H. S. Conradt, M. Nimtz, G. Hale, C. Kirchhoff, *J. Biol. Chem.* **1999**, *274*, 29862−29873.
[16] T. Kinoshita, N. Inoue, *Curr. Opin. Chem. Biol.* **2000**, *4*, 632−638.
[17] A. M. Tartakoff, N. Singh, *TiBS* **1992**, *17*, 470−473.
[18] Y. S. Morita, A. Acosta-Serrano, P. T. Englund, *Carbohydr. Chem. Biol.* **2000**, *3*, 417−433.
[19] M. A. J. Ferguson, *J. Cell Sci.* **1999**, *112*, 2799−2809.
[20] P. M. Rudd, B. P. Morgan, M. R. Wormald, D. J. Harvey, C. W. Van Den Berg, S. J. Davis, M. A. J. Ferguson, R. A. Dwek, *Biochem. Soc. Trans.* **1997**, *25*, 1177−1184.
[21] A. K. Menon, N. A. Baumann, D. M. Rancour, *Carbohydr. Chem. Biol.* **2000**, *4*, 757−769.
[22] R. N. Cole, G. W. Hart, *New Comprehensive Biochem.* **1997**, *29*, 69−88.
[23] R. Gigg, J. Gigg, in *Glycopeptides and related compounds: synthesis, analysis, and applications* (Eds.: D. G. Large, C. D. Warren), Marcel Dekker, Inc., New York, **1997**, pp. 327−392.
[24] C. Murakata, T. Ogawa, *Tetrahedron Lett.* **1991**, *32*, 671−674.
[25] D. K. Baeschlin, A. R. Chaperon, V. Charbonneau, L. G. Green, S. V. Ley, U. Lucking, E. Walther, *Angew. Chem. Int. Ed.* **1998**, *37*, 3423−3428.
[26] D. K. Baeschlin, A. R. Chaperon, L. G. Green, M. G. Hahn, S. J. Ince, S. V. Ley, *Chem. Eur. J.* **2000**, *6*, 172−186.
[27] A. S. Campbell, B. Fraser-Reid, *J. Am. Chem. Soc.* **1995**, *117*, 10387−10388.
[28] U. E. Udodong, R. Madsen, C. Roberts, B. Fraser-Reid, *J. Am. Chem. Soc.* **1993**, *115*, 7886−7887.
[29] K. Pekari, R. R. Schmidt, *J. Org. Chem.* **2003**, *68*, 1295−1308.
[30] T. G. Mayer, B. Kratzer, R. R. Schmidt, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2177−2181.
[31] T. G. Mayer, R. R. Schmidt, *Eur. J. Org. Chem.* **1999**, 1153−1165.
[32] K. Pekari, D. Tailler, R. Weingart, R. R. Schmidt, *J. Org. Chem.* **2001**, *66*, 7432−7442.
[33] J. Lu, K. N. Jayaprakash, B. Fraser-Reid, *Tetrahedron Lett.* **2004**, *45*, 879−882.
[34] M. C. Hewitt, D. A. Snyder, P. H. Seeberger, *J. Am. Chem. Soc.* **2002**, *124*, 13434−13436.
[35] L. Schofield, M. C. Hewitt, K. Evans, M.-A. Siomos, P. H. Seeberger, *Nature* **2002**, *418*, 785−789.
[36] J. Xue, Z. Guo, *J. Am. Chem. Soc.* **2003**, *125*, 16334−16339.
[37] J. Xue, N. Shao, Z. Guo, *J. Org. Chem.* **2003**, *68*, 4020−4029.
[38] K. M. Sureshan, M. S. Shashidhar, T. Praveen, T. Das, *Chem. Rev.* **2003**, 4477−4503.
[39] Z. J. Jia, L. Olsson, B. Fraser-Reid, *J. Chem. Soc., Perkin Trans. 1* **1998**, 631−632.
[40] P. J. Garegg, T. Iversen, R. Johansson, B. Lindberg, *Carbohydr. Res.* **1984**, *130*, 322−326.
[41] D. J. R. Massy, P. Wyss, *Helv. Chim. Acta* **1990**, *73*, 1037−1057.
[42] M. Martin-Lomas, N. Khiar, S. Garcia, J.-L. Koessler, P. M. Nieto, T. W. Rademache, *Chem. Eur. J.* **2000**, *6*, 3608−3621.
[43] H. Dietrich, J. F. Espinosa, J. L. Chiara, J. Jimenez-Barbero, Y. Leon, I. Varela-Nieto, J.-M. Mato, F. H. Cano, C. Foces-Foces, M. Martin-Lomas, *Chem. Eur. J.* **1999**, *5*, 320−336.
[44] C. Murakata, T. Ogawa, *Tetrahedron Lett.* **1990**, *31*, 2439−2442.
[45] C. Murakata, T. Ogawa, *Carbohydr. Res.* **1992**, *235*, 95−114.
[46] T. G. Mayer, R. Weingart, F. Münstermann, T. Kawada, T. Kurzchalia, R. R. Schmidt, *Eur. J. Org. Chem.* **1999**, 2563−2571.
[47] B. Kratzer, T. G. Mayer, R. R. Schmidt, *Eur. J. Org. Chem.* **1998**, 291−298.
[48] D. Gou, Y. Liu, C. Chen, *Carbohydr. Res.* **1992**, *234*, 51−64.
[49] J. Xue, Z. Guo, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2015−2018.
[50] R. M. Conrad, M. J. Grogan, C. R. Bertozzi, *Org. Lett.* **2002**, *4*, 1359−1361.
[51] N.-C. Reichardt, M. Martin-Lomas, *Angew. Chem. Int. Ed.* **2003**, *42*, 4674−4677.
[52] Y. Tanaka, Y. Nakahara, H. Hojo, Y. Nakahara, *Tetrahedron* **2003**, *59*, 4059−4067.
[53] N. Shao, J. Xue, Z. Guo, *Angew. Chem. Int. Ed.* **2004**, *43*, 1569−1573.
[54] L. A. Carpino, A. El-Faham, F. Albericio, *J. Org. Chem.* **1995**, *60*, 3561−3564.
[55] N. Shao, J. Xue, Z. Guo, *J. Org. Chem.* **2003**, *68*, 9003−9011.
[56] C. Unverzagt, H. Kunz, *Bioorg. Med. Chem.* **1994**, *2*, 1189−1201.
[57] R. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776−779.
[58] C. Lin, J. Tam, *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 6584−6588.

Received March 25, 2004
Early View Article
Published Online July 14, 2004

