BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Biomolecular systems

Edited by Peter H. Seeberger

Imprint

Beilstein Journal of Organic Chemistry
www.bjoc.org
ISSN 1860-5397
Email: journals-support@beilstein-institut.de

The *Beilstein Journal of Organic Chemistry* is published by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften.

Beilstein-Institut zur Förderung der Chemischen Wissenschaften
Trakehner Straße 7–9
60487 Frankfurt am Main
Germany
www.beilstein-institut.de

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Synthesis of D-*manno*-heptulose via a cascade aldol/hemiketalization reaction

Yan Chen, Xiaoman Wang, Junchang Wang and You Yang*

Letter

Open Access

Address:
Shanghai Key Laboratory of New Drug Design, School of Pharmacy, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China

Email:
You Yang* - yangyou@ecust.edu.cn

* Corresponding author

Keywords:
aldol reaction; cascade reaction; D-*manno*-heptulose; higher-carbon sugar; ketoheptose

Beilstein J. Org. Chem. **2017,** *13,* 795–799.
doi:10.3762/bjoc.13.79

Received: 15 February 2017
Accepted: 12 April 2017
Published: 28 April 2017

This article is part of the Thematic Series "Biomolecular systems".

Guest Editor: P. H. Seeberger

Abstract

A [4 + 3] synthesis of D-*manno*-heptulose is described. The cascade aldol/hemiketalization reaction of a C_4 aldehyde with a C_3 ketone provides the differentially protected ketoheptose building block, which can be further reacted to furnish target D-*manno*-heptulose.

Introduction

D-*manno*-Heptulose is a rare naturally occurring seven-carbon sugar first isolated from avocado [1], which exhibited promising diabetogenic effects through suppression of the glucose metabolism and insulin secretion via competitive inhibition of the glucokinase pathway [2-6]. Accordingly, ketoheptoses and fluorinated ketoheptoses were considered to be potential therapeutic agents for hypoglycemia and cancer as well as diagnostic tools for diabetes [7-12]. Amino- and azido-group-containing ketoheptoses were also synthesized for the development of novel antibiotics and the evaluation of carbohydrate–lectin interactions by conjugation with fluorescent quantum dots via click chemistry [13,14]. Besides, differentially protected D-*manno*-heptulose building blocks could serve as valuable precursors for the synthesis of *C*-glycosides [15,16].

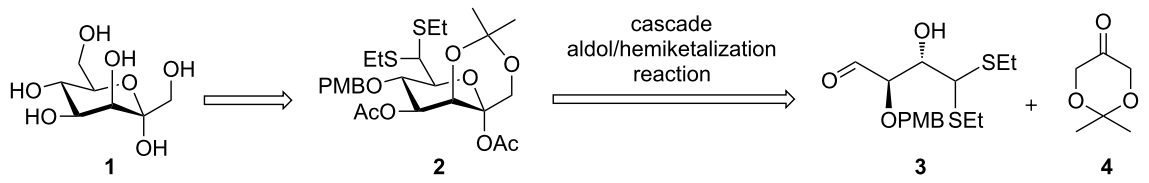
The known synthesis of D-*manno*-heptulose mainly rely on the use of rearrangements and chain elongation reactions [17]. Rearrangement reactions such as the Lobry de Bruyn rearrangement and the Bilik rearrangement employ unprotected aldoses as substrates, usually yielding an equilibrium mixture of aldoses and ketoses [18,19]. In addition to chain elongations of aldoses employing the Henry reaction, the aldol reaction, and the Wittig reaction for the preparation of ketoheptoses [20-22], sugar lactones were also often utilized for the synthesis of D-*manno*-heptulose via reactions with *C*-nucleophiles or conversion into exocyclic glycals followed by dihydroxylation [10-13,23-27]. Remarkably, Thiem et al. reported the highly efficient synthesis of D-*manno*-heptulose from D-mannose in 59% overall yield over five steps [26]. However, the synthesis of D-*manno*-heptu-

lose and its derivatives from the common differentially protected ketoheptose building block is still attractive due to the versatile functionalization possibilities of the building block into various derivatives of D-*manno*-heptulose. A de novo synthesis has proved to be an attractive strategy to produce orthogonally protected carbohydrate building blocks from simple precursors [28-39]. Here, we report a [4 + 3] approach to access differentially protected ketoheptose building blocks, which enables the synthesis of D-*manno*-heptulose. As depicted in Scheme 1, D-*manno*-heptulose (**1**) could be obtained by global deprotection of the differentially protected ketoheptose building block **2**. The ketoheptose **2** can be further divided into C_4 aldehyde **3** and C_3 ketone **4** via a cascade aldol/hemiketalization pathway.

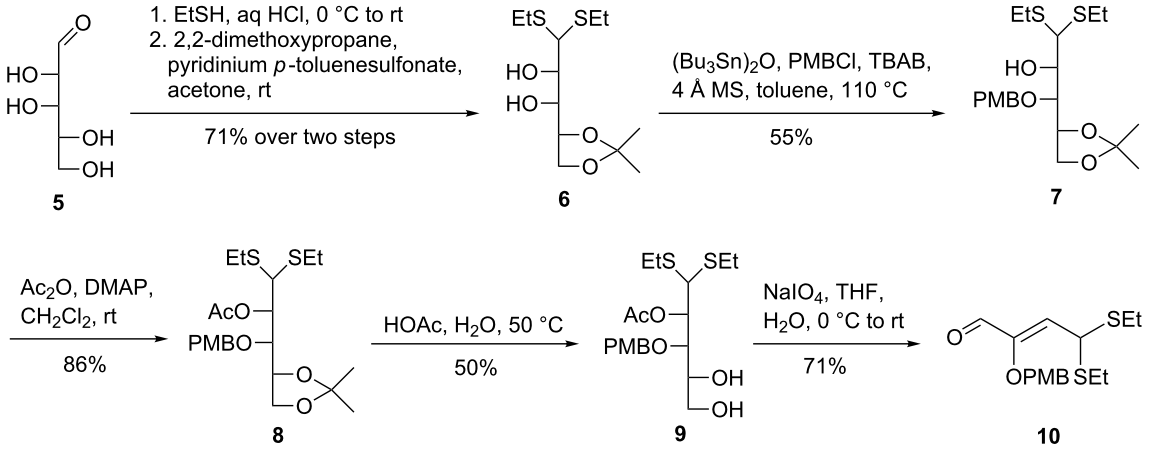
Results and Discussion

The synthesis of the C_4 aldehyde commenced with commercially available D-lyxose (**5**, Scheme 2). The reaction of **5** with ethanethiol in the presence of hydrochloric acid followed by selective protection of the 4,5-diol with 2,2-dimethoxypropane using pyridinium *p*-toluenesulfonate as the promoter gave the 4,5-*O*-isopropylidene derivative **6** in 71% yield over two steps [40]. Treatment of diol **6** with bis(tributyltin) oxide and subsequent exposure to *p*-methoxybenzyl (PMB) chloride in the presence of tetra-*n*-butylammonium bromide (TBAB) at 110 °C led to regioselective protection of the 3-OH with the PMB group, affording the 3-*O*-PMB protected alcohol **7** (55%) [41]. At this stage, we initially planned to synthesize the 2-OH-protected C_4 aldehyde for the assembly of the seven-carbon skeleton. Thus, acetylation of the 2-OH group in **7** with acetic anhydride and DMAP in dichloromethane provided ester **8** in 86% yield. The positions of the 2-acetyl and 3-PMB groups were determined by ^{1}H, ^{13}C and 2D NMR spectra of **8** (see Supporting Information File 1 for details). Cleavage of the isopropylidene acetal group in **8** under acidic conditions gave diol **9** (50%). However, oxidative cleavage of diol **9** with sodium periodate resulted in the unexpected formation of α,β-unsaturated aldehyde **10** in 71% yield, indicating that the 2-acetyl group might be prone to initiate the elimination reaction. The double bond of **10** was assigned to have *Z*-configuration based on the analysis of the NOEs between the olefinic hydrogen and the aldehyde hydrogen (see Supporting Information File 1 for details). In addition, when alcohol **7** was subjected to benzoyl chloride and DMAP in dichloromethane at room temperature or *tert*-butyldimethylsilyl chloride and imidazole in DMF at room temperature, no reaction occurred probably because of the steric hindrance between the 2-OH group and the surrounding functional groups.

To overcome the difficulties in the synthesis of the 2-OH-protected C_4 aldehyde and to improve the synthetic efficiency in



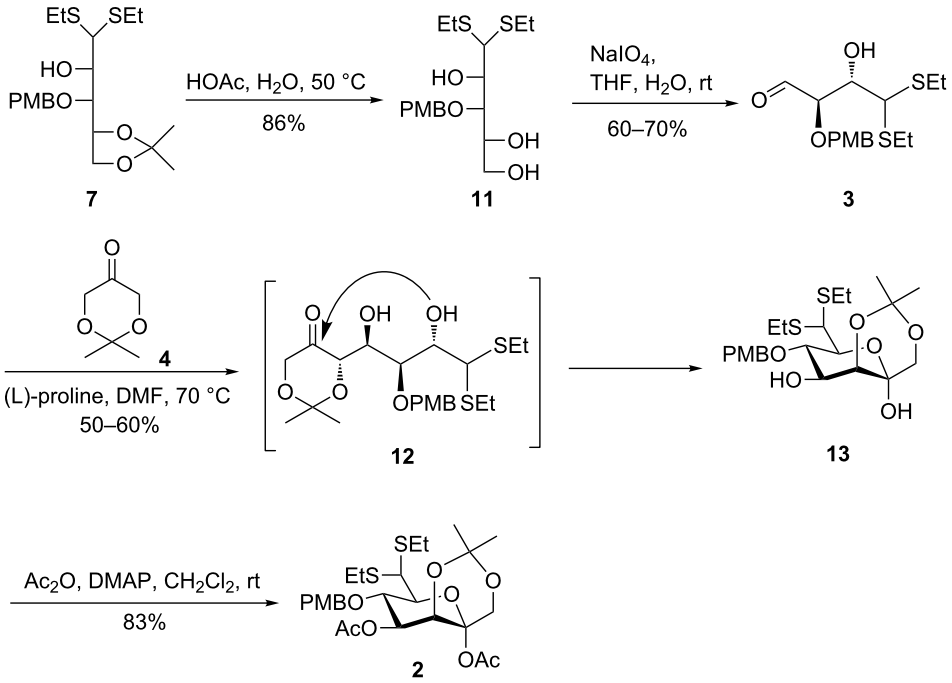
Scheme 1: Retrosynthetic analysis of D-*manno*-heptulose.



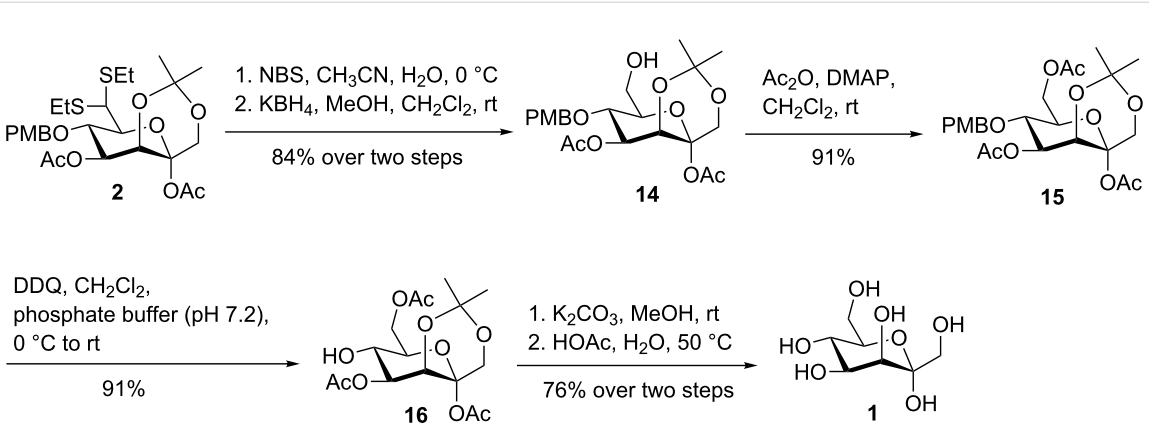
Scheme 2: Initial attempt on the synthesis of the C_4 aldehyde from D-lyxose (**5**).

the assembly of ketoheptose skeletons, we envisioned ketoheptoses could be assembled by a cascade aldol/hemiketalization reaction between 2-OH-unprotected C_4 aldehyde **3** and C_3 ketone **4**. As such, the isopropylidene acetal group in **7** was cleaved under acidic conditions to produce triol **11** in 86% yield (Scheme 3). Cleavage of the resulting vicinal diol in **11** with sodium periodate led to the C_4 aldehyde **3** in nearly 60–70% yield. In this oxidative cleavage reaction, almost no elimination product was found based on TLC monitoring. Given that the C_4 aldehyde **3** was unstable upon purification by silica gel column chromatography, it was immediately used for the subsequent coupling after the extraction procedure. The aldol reaction of aldehyde **3** with the readily available ketone **4** [42,43] under the catalysis of L-proline at room temperature for three days proceeded sluggishly, leading to the desired product in a very low yield. Gratifyingly, when the L-proline-catalyzed aldol reaction was performed at 70 °C for one day, the TLC indicated the complete consumption of aldehyde **3**, and the generated 4,5-*anti*-selective coupling intermediate **12** underwent in situ cyclization to provide hemiketal **13** as the major product in about 50–60% yield (35% overall yield from compound **11**). Notably, trace amounts of a stereoisomer and a minor highly polar unknown byproduct were also observed in this cascade reaction. The excellent *anti*-selectivity for the L-proline-catalyzed aldol reaction can be explained by the Houk–List transition state model [43-45]. Compound **13** was then acetylated to afford differentially protected ketoheptose building block **2** in 83% yield. The structure of **2** was unambiguously confirmed by ^{1}H, ^{13}C, and 2D NMR spectra (see Supporting Information File 1 for details). The anomeric α-configuration of compound **2** was confirmed by analysis of the NOE effects between the C-1 hydrogen and the C-5 hydrogen.

With the ketoheptose building block **2** in hand, we turned our attention to the synthesis of D-*manno*-heptulose (**1**). Upon exposure to NBS in acetonitrile and water, the dithioacetal in **2** was cleaved to give the corresponding aldehyde [46,47], which was then reduced by potassium borohydride in a methanol and dichloromethane solvent mixture to produce alcohol **14** as the predominant product (84% over two steps, Scheme 4). In addition, a trace amount of the deacetylated product was also detected . DDQ-mediated oxidative cleavage of the PMB group in alcohol **14** produced only a moderate yield (≈50%) of the 5,7-diol probably due to the presence of the free 7-hydroxy group. We envisaged that protection of the free 7-hydroxy group in **14** followed by treatment with DDQ could yield the desired 5-hydroxy product in high yield. Indeed, acetylation of alcohol **14** with acetic anhydride delivered ester **15** in 91% yield. Removal of the PMB group in **15** with DDQ resulted in a very clean reaction, affording alcohol **16** in an excellent yield (91%). Saponification of all esters in **16** with potassium carbonate followed by acidic cleavage of the isopropylidene



Scheme 3: Synthesis of differentially protected ketoheptose building block **2**.



Scheme 4: Synthesis of D-*manno*-heptulose (**1**).

acetal group with aqueous acetic acid furnished D-*manno*-heptulose (**1**, 76% over two steps). The structure of **1** was found to be in good agreement with those reported for α-D-*manno*-heptulose (**1**) by comparison of the NMR spectra (see Supporting Information File 1 for details) [26].

Conclusion

In summary, we have described a [4 + 3] approach for the synthesis of D-*manno*-heptulose (**1**) starting from D-lyxose (**5**). The key step is a cascade aldol/hemiketalization reaction for the construction of the differentially protected ketoheptose building block, which was finally converted into D-*manno*-heptulose for subsequent biological evaluation. Although the synthesis of D-*manno*-heptulose (5% overall yield, 13 steps) is not so efficient as the Thiem's method (59% overall yield, 5 steps), the reported differentially protected ketoheptose building blocks may find further application in the preparation of structurally diverse D-*manno*-heptulose derivatives.

Supporting Information

Supporting Information File 1

Experimental details, characterization data, and NMR spectra of all new compounds.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-13-79-S1.pdf]

Acknowledgements

Financial support from the National Thousand Young Talents Program (YC0130518, YC0140103), the Shanghai Pujiang Program (15PJ1401500), the Fundamental Research Funds for the Central Universities (WY1514052), and the Opening Project of State Key Laboratory of Bio-organic and Natural Products Chemistry (Y100-D-1501) is gratefully acknowledged.

References

1. La Forge, F. B. *J. Biol. Chem.* **1917,** *28,* 511–522.
2. Roe, J. H.; Hudson, C. S. *J. Biol. Chem.* **1936,** *112,* 443–449.
3. Simon, E.; Kraicer, P. F. *Arch. Biochem. Biophys.* **1957,** *69,* 592–601. doi:10.1016/0003-9861(57)90523-4
4. Paulsen, E. P.; Richenderfer, L.; Winick, P. *Nature* **1967,** *214,* 276–277. doi:10.1038/214276b0
5. Coore, H. G.; Randle, P. J. *Biochem. J.* **1964,** *91,* 56–59. doi:10.1042/bj0910056
6. Zelent, D.; Najafi, H.; Odili, S.; Buettger, C.; Weik-Collins, H.; Li, C.; Doliba, N.; Grimsby, J.; Matschinsky, F. M. *Biochem. Soc. Trans.* **2005,** *33,* 306–310. doi:10.1042/BST0330306
7. Paulsen, E. P. *Ann. N. Y. Acad. Sci.* **1968,** *150,* 455–456. doi:10.1111/j.1749-6632.1968.tb19069.x
8. Board, M.; Colquhoun, A.; Newsholme, E. A. *Cancer Res.* **1995,** *55,* 3278–3285.
9. Malaisse, W. J. *Diabetologia* **2001,** *44,* 393–406. doi:10.1007/s001250051635
10. Leshch, Y.; Waschke, D.; Thimm, J.; Thiem, J. *Synthesis* **2011,** 3871–3877. doi:10.1055/s-0031-1289598
11. Waschke, D.; Leshch, Y.; Thimm, J.; Himmelreich, U.; Thiem, J. *Eur. J. Org. Chem.* **2012,** 948–959. doi:10.1002/ejoc.201101238
12. Malaisse, W. J.; Zhang, Y.; Louchami, K.; Sharma, S.; Dresselaers, T.; Himmelreich, U.; Novotny, G. W.; Mandrup-Poulsen, T.; Waschke, D.; Leshch, Y.; Thimm, J.; Thiem, J.; Sener, A. *Arch. Biochem. Biophys.* **2012,** *517,* 138–143. doi:10.1016/j.abb.2011.11.014
13. Leshch, Y.; Jacobsen, A.; Thimm, A.; Thiem, J. *Org. Lett.* **2013,** *15,* 4948–4951. doi:10.1021/ol4021699
14. Schmidtke, C.; Kreuziger, A.-M.; Alpers, D.; Jacobsen, A.; Leshch, Y.; Eggers, R.; Kloust, H.; Tran, H.; Ostermann, J.; Schotten, T.; Thiem, J.; Thimm, J.; Weller, H. *Langmuir* **2013,** *29,* 12593–12600. doi:10.1021/la402826f
15. Levy, D. E.; Tang, C. *The chemistry of C-glycosides;* Pergamon: Oxford, 1995.
16. Du, Y.; Linhardt, R. J.; Vlahov, I. R. *Tetrahedron* **1998,** *54,* 9913–9959. doi:10.1016/S0040-4020(98)00405-0
17. Jacobsen, A.; Thiem, J. *Curr. Org. Chem.* **2014,** *18,* 2833–2841. doi:10.2174/1385272819666141016215205
18. Montgomery, E. M.; Hudson, C. S. *J. Am. Chem. Soc.* **1939,** *61,* 1654–1658. doi:10.1021/ja01876a007

19. Hricoviniova, Z.; Hricovini, M.; Petrusoa, M.; Matulova, M.; Petrus, L. *Chem. Pap.* **1998,** *52,* 238–243.
20. Sowden, J. C. *J. Am. Chem. Soc.* **1950,** *72,* 3325. doi:10.1021/ja01163a558
21. Schaffner, R.; Isbell, H. S. *J. Org. Chem.* **1962,** *27,* 3268–3270. doi:10.1021/jo01056a069
22. Cheng, J.; Fang, Z.; Li, S.; Zheng, B.; Jiang, Y. *Carbohydr. Res.* **2009,** *344,* 2093–2095. doi:10.1016/j.carres.2009.06.020
23. Kampf, A.; Dimant, E. *Carbohydr. Res.* **1974,** *32,* 380–382. doi:10.1016/S0008-6215(00)82116-3
24. Bessières, B.; Morin, C. *J. Org. Chem.* **2003,** *68,* 4100–4103. doi:10.1021/jo0342166
25. Liu, X.; Yin, Q.; Yin, J.; Chen, G.; Wang, X.; You, Q.-D.; Chen, Y.-L.; Xiong, B.; Shen, J. *Eur. J. Org. Chem.* **2014,** 6150–6154. doi:10.1002/ejoc.201402757
26. Waschke, D.; Thimm, J.; Thiem, J. *Org. Lett.* **2011,** *13,* 3628–3631. doi:10.1021/ol2012764
27. Li, X.; Takahashi, H.; Ohtake, H.; Shiro, M.; Ikegami, S. *Tetrahedron* **2001,** *57,* 8053–8066. doi:10.1016/S0040-4020(01)00775-X
28. Northrup, A. B.; MacMillan, D. W. C. *Science* **2004,** *305,* 1752–1755. doi:10.1126/science.1101710
29. Timmer, M. S. M.; Adibekian, A.; Seeberger, P. H. *Angew. Chem., Int. Ed.* **2005,** *44,* 7605–7607. doi:10.1002/anie.200502742
30. Ahmed, M. M.; Berry, B. P.; Hunter, T. J.; Tomcik, D. J.; O'Doherty, G. A. *Org. Lett.* **2005,** *7,* 745–748. doi:10.1021/ol050044i
31. Adibekian, A.; Timmer, M. S. M.; Stallforth, P.; van Rijn, J.; Werz, D. B.; Seeberger, P. H. *Chem. Commun.* **2008,** 3549–3551. doi:10.1039/B805159C
32. Stallforth, P.; Adibekian, A.; Seeberger, P. H. *Org. Lett.* **2008,** *10,* 1573–1576. doi:10.1021/ol800227b
33. Shan, M.; Xing, Y.; O'Doherty, G. A. *J. Org. Chem.* **2009,** *74,* 5961–5966. doi:10.1021/jo9009722
34. Ohara, T.; Adibekian, A.; Esposito, D.; Stallforth, P.; Seeberger, P. H. *Chem. Commun.* **2010,** *46,* 4106–4108. doi:10.1039/c000784f
35. Calin, O.; Pragani, R.; Seeberger, P. H. *J. Org. Chem.* **2012,** *77,* 870–877. doi:10.1021/jo201883k
36. Babu, R. S.; Chen, Q.; Kang, S.-W.; Zhou, M.; O'Doherty, G. A. *J. Am. Chem. Soc.* **2012,** *134,* 11952–11955. doi:10.1021/ja305321e
37. Gati, W.; Rammah, M. M.; Rammah, M. B.; Couty, F.; Evano, G. *J. Am. Chem. Soc.* **2012,** *134,* 9078–9081. doi:10.1021/ja303002a
38. Mlynarski, J.; Gut, B. *Chem. Soc. Rev.* **2012,** *41,* 587–596. doi:10.1039/C1CS15144D
39. Wang, H.-Y.; Yang, K.; Yin, D.; Liu, C.; Glazier, D. A.; Tang, W. *Org. Lett.* **2015,** *17,* 5272–5275. doi:10.1021/acs.orglett.5b02641
40. van Delft, F. L.; Rob, A.; Valentijn, P. M.; van der Marel, G. A.; van Boom, J. H. *J. Carbohydr. Chem.* **1999,** *18,* 165–190. doi:10.1080/07328309908543989
41. Grindley, T. B. *Adv. Carbohydr. Chem. Biochem.* **1998,** *53,* 17–142. doi:10.1016/S0065-2318(08)60043-8
42. Suri, J. T.; Mitsumori, S.; Albertshofer, K.; Tanaka, F.; Barbas, C. F., III. *J. Org. Chem.* **2006,** *71,* 3822–3828. doi:10.1021/jo0602017
43. Grondal, C.; Enders, D. *Tetrahedron* **2006,** *62,* 329–337. doi:10.1016/j.tet.2005.09.060
44. Bahmanyar, S.; Houk, K. N.; Martin, H. J.; List, B. *J. Am. Chem. Soc.* **2003,** *125,* 2475–2479. doi:10.1021/ja028812d
45. Hoang, L.; Bahmanyar, S.; Houk, K. N.; List, B. *J. Am. Chem. Soc.* **2003,** *125,* 16–17. doi:10.1021/ja028634o
46. Corey, E. J.; Erickson, B. W. *J. Org. Chem.* **1971,** *36,* 3553–3560. doi:10.1021/jo00822a019
47. Crich, D.; de la Mora, M. A.; Cruz, R. *Tetrahedron* **2002,** *58,* 35–44. doi:10.1016/S0040-4020(01)01087-0

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Expression, purification and structural analysis of functional GABA transporter 1 using the baculovirus expression system

Jing Hu*[1,2], Chris Weise[3], Christoph Böttcher[4], Hua Fan[5] and Jian Yin*[6]

Full Research Paper

Open Access

Address:
[1]Wuxi School of Medicine, Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, Jiangnan University, Lihu Avenue 1800, 214122, Wuxi, China, [2]Institut für Biochemie und Molekularbiologie, Campus Benjamin Franklin, Charité Universitätsmedizin Berlin, Arnimallee 22, Berlin D-14195, Germany, [3]Forschungszentrum für Elektronenmikroskopie, Institut für Chemie und Biochemie, Freie Universität Berlin, Fabeckstr. 36a, Berlin D-14195, Germany, [4]Institut für Chemie und Biochemie - Biochemie, Freie Universität Berlin, Thielallee 63, Berlin D-14195, Germany, [5]Institut für Laboratoriumsmedizin, Klinische Chemie und Pathobiochemie, Charité Universitätsmedizin Berlin, Berlin D-13353, Germany and [6]Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Lihu Avenue1800, Wuxi 214122, China

Email:
Jing Hu* - hujing@jiangnan.edu.cn; Jian Yin* - jianyin@jiangnan.edu.cn

* Corresponding author

Keywords:
baculovirus expression system; chromatography; γ-aminobutyric acid (GABA); GABA transporter 1 (GAT1)

Beilstein J. Org. Chem. **2017,** *13,* 874–882.
doi:10.3762/bjoc.13.88

Received: 27 February 2017
Accepted: 26 April 2017
Published: 11 May 2017

This article is part of the Thematic Series "Biomolecular systems".

Guest Editor: P. H. Seeberger

Abstract

The γ-aminobutyric acid (GABA) transporter 1 (GAT1) belongs to a family of Na^+ and Cl^--coupled transport proteins and possesses 12 putative transmembrane domains. To perform structural analyses of the GAT1 protein, the GAT1/green fluorescent protein (GFP) fusion protein was functionally expressed in insect *Sf*9 cells by the BAC-TO-BAC® baculovirus expression system. A two-step procedure to purify the GAT1/GFP fusion protein from insect *Sf*9 cells has been established and involves immunoaffinity chromatography using self-prepared anti-GFP antibodies and size-exclusion fast protein liquid chromatography (SE-FPLC). A yield of 200–300 μg of the GAT1/GFP protein could be purified from 400–600 mL of infected *Sf*9 cells. The purified protein was analyzed by transmission electron microscopy (TEM), which revealed that the GAT1/GFP fusion protein was isolated in its monomeric form.

Introduction

γ-Aminobutyric acid (GABA) is the most abundant inhibitory neurotransmitter in the central nervous system (CNS) of vertebrate species. GABAergic neurotransmission is efficiently terminated through the quick removal of GABA from the synaptic cleft by GABA transporters (GATs). The activities of GATs are important for controlling the concentration and dwell

time of GABA in the synaptic cleft and tightly regulating the synaptic inhibition of the GABA receptor [1]. GATs, which are located in the plasma membranes of neurons and glia cells, belong to the solute carrier 6 (SLC6) family in mammals, which is subdivided into four groups based on sequence composition: GATs, GABA, osmolyte and creatine transporters, neurotransmitter amino acid, monoamine and nutrient amino acid/orphan transporters [2,3]. Four subtypes of GATs (GAT1–4) have been identified thus far [4].

GAT1, the first neurotransmitter transporter to be cloned, is abundantly but restrictively expressed throughout rat, mouse, and human CNSs [4-8]. The transmembrane (TM) topology of GAT1 demonstrates that this single polypeptide contains twelve TM domains connected by hydrophilic loops with the amino and carboxy-termini residing in the cytoplasm [9,10]. Additionally, the GAT1 protein contains three conserved N-glycosylation sites [9]. The role of N-linked glycans in the GABA uptake activity of GAT1 has been extensively clarified [11,12]. The topological model of GAT1 is in agreement with the high-resolution crystal structure of $LeuT_{Aa}$, a homologue of the SLC/neurotransmitter sodium symporter (NSS) transporters from the bacterium *Aquifexaeolicus* [13]. The $LeuT_{Aa}$ structure has provided a good model for the study of GAT1 and other SLC6 members [3]. The activities of GATs are driven by electrochemical gradients of Na^+ and Cl^- ions [14], and the binding pocket for the substrate and two sodium ions has been clearly documented in the crystal structure of $LeuT_{Aa}$. The chloride dependence has also been elucidated by identifying crucial structural elements for chloride binding [15,16]. In addition, a similar core architecture was observed based on several crystal structures of sodium symporters from different families; this observation strongly supports an alternating access mechanism as a common transport mechanism for the GAT1 protein [13,17-20].

Since the GABAergic system has been implicated in many nervous system diseases [21-24], the regulation of GABA activity is of considerable medical interest [25,26] and the predominant GABAergic nerve ending protein, GAT1, is a potential drug target. The exact three-dimensional structure of GAT1 protein could provide more information for pharmaceutical research.

The structural analysis of most membrane proteins is challenging since significant protein yields are required and because eukaryotic membrane proteins should be in a native and functional conformation during expression and purification. The natural abundance of most membrane proteins is typically not high enough for the isolation of sufficient quantities for functional and structural studies. In this work, the baculovirus expression system was selected for GAT1 protein expression. Mass spectrometry, gel electrophoresis and GABA uptake assays were employed to characterize the product during expression and purification. In our work, the GAT1 protein with a green fluorescent protein (GFP) tag at its C-terminus, which does not affect the relevant functions of GAT1 [27], could be purified in a functional form. In addition, the GFP tag has several advantages for further work. For example, it provides a powerful means for affinity purification with specific anti-GFP antibodies that have already been produced, generates green fluorescence for the efficient observation of the expression of the protein of interest, and has a known crystal structure [28,29] for future structural studies.

Results and Discussion

Expression and characterization of the GAT1/GFP fusion protein in insect cells

Construction and analysis of the GAT1/GFP recombinant baculovirus

The purification of a protein of interest for structural studies requires an abundant source of the protein from heterologous overexpression. In our work, *Escherichia coli* was not a suitable system for the expression of the GAT1/GFP recombinant protein due to its twelve TM domains and N-glycosylation status. After 30 years of development, the baculovirus expression system has become a widely applied technology for producing recombinant proteins [30,31]. Since the production of adequate quantities of a homogenous protein is a rate-limiting step, in this study, we chose the baculovirus expression system for GAT1 protein expression instead of the mammalian cell expression system.

The GAT1/GFP recombinant cDNA was cloned into the pFastBac1 vector (Figure S1A, Supporting Information File 1) and prepared in bacmids (Figure S1B, Supporting Information File 1) for the baculovirus expression system. GAT1/GFP recombinant proteins were expressed in *Sf*9 cells by viral infection for 72 h. The expression of the GAT1/GFP protein was further characterized. After 72 h post-infection, *Sf*9 cells with green fluorescence were controlled by flow cytometry (Figure 1A) and analyzed by fluorescence microscopy (Figure 1B). The expression of GAT1/GFP in insect cells was further determined by sodium dodecyl sulfate polyacrylamide gel electrophoresis SDS-PAGE, followed by Western blotting with either anti-GFP pAb (Figure 1C) or anti-GAT1 pAb (Figure 1D) and silver staining (Figure 1E) after immunoprecipitation with anti-GFP mAb (showing the band at approximately 50 kDa). In infected *Sf*9 cells, the GAT1/GFP-fusion protein shows two main bands in SDS-PAGE (7.5%) (indicated by arrows), and the monomeric form appears as a main band of approximately 75 kDa. This band can bind strongly only with *Galanthusnivalis* agglutinin (GNA, digoxigenin-conjugated

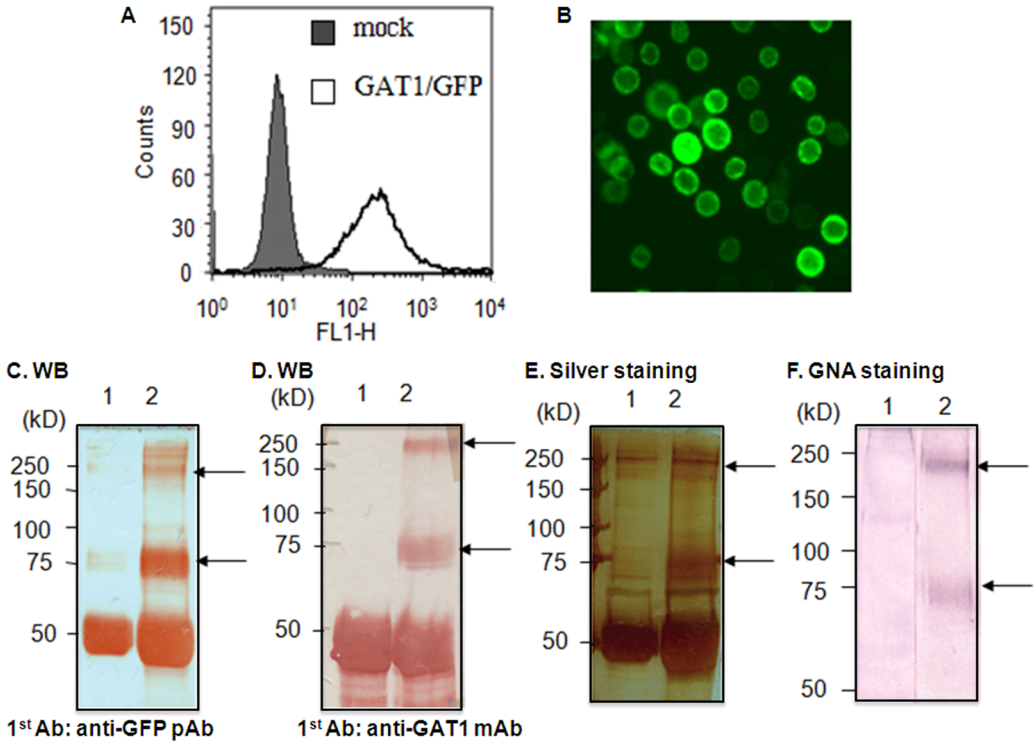


Figure 1: Expression of GFP-tagged GAT1 in infected insect cells. (A) Flow cytometry analysis of GAT1/GFP in *Sf*9 cells. (B) Fluorescence microscopy of infected *Sf*9 cells. The GFP fluorescence in the GAT1/GFP fusion protein was detected. The solubilized proteins of infected *Sf*9 cells were subjected to immunoprecipitation with anti-GFP mAb IgG and then analyzed by SDS-PAGE (7.5%) and immunoblotting. (C) and (D) Western blotting analysis of GAT1/GFP from uninfected and infected *Sf*9 cells with anti-GFP pAb or anti-GAT1 pAb. (E) Silver staining of GAT1/GFP from infected *Sf*9 cells. (F) GNA staining of GAT1/GFP from infected *Sf*9 cells. The main bands corresponding to the GAT1/GFP fusion protein are indicated with arrows; *Sf*9 cells (lane 1), GAT1/GFP infected *Sf*9 cells (lane 2).

lectin) (Figure 1F), indicating the predominance of the paucimannose structure in insect cells. In contrast, mammalian *N*-glycans have terminal sialic acid residues with more antennal diversity. The band at approximately 250 kDa corresponds to an oligomeric form or protein aggregate identified through further characterization.

The two main protein bands after immunoprecipitation and SDS-PAGE with Coomassie Blue staining (a and b in Figure 2A) were extracted for protein fingerprinting analysis. The extracted proteins were processed by trypsin treatment, and tryptic peptides were analyzed either directly by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI–TOF MS) or guanidinated and then analyzed by MALDI. The labeled peptide peaks (Figure 2B) in the mass spectrum were identified as either GAT or GFP by matching the peptides from primary sequence databases with Mascot (http://www.matrixscience.com/).

The biological function of recombinant proteins of mammalian origin expressed in insect cells may be altered by different *N*-glycan status. We observed that the terminal sialic acid residues are essential for the GABA uptake activity of GAT1 because they affect the ionic affinity for Na^+ and the conformational change of the GAT1 protein during its uptake process [12]. After 72 h post-infection, the GABA uptake activity of *Sf*9 cells was measured to determine whether the GAT1/GFP-fusion protein was functionally expressed in the baculovirus expression system. The results showed that infected *Sf*9 cells (0.15 pmol/10^6 cells) have only slightly higher GABA uptake activity than mock cells (0.1 pmol/10^6 cells) (Figure S3, Supporting Information File 1). A previous work demonstrated that co-translational *N*-glycosylation but not the terminal trimming of *N*-glycans is involved in the regulation of the correct membrane glycoprotein folding since the inhibition of *N*-glycosylation processing by 1-deoxymannojirimycin (dMM) results in a mannose-rich type of *N*-glycan that does not affect either the protein stability or intracellular trafficking [11]. Therefore, the correct folding of GAT1/GFP protein in insect cells should not be affected by the lack of terminal trimming, including the sialylation of *N*-glycans. The low GABA uptake activity of GAT1/GFP fusion protein in insect cells should result from the terminal mannose structure on the *N*-glycans, which is consistent with previous findings [12]. Therefore, the GAT1/GFP pro-

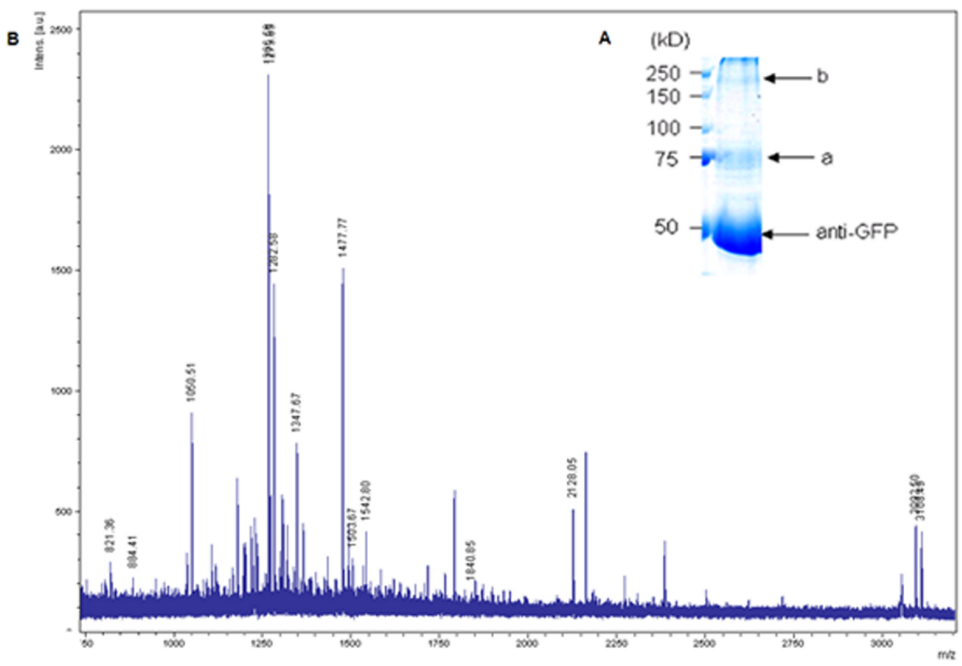


Figure 2: Protein fingerprinting. (A) Coomassie Blue staining of the GAT1/GFP protein from infected *Sf*9 cells after immunoprecipitation. Bands a and b (indicated by arrows) were extracted for MALDI–TOF analysis. (B) Mass spectrum of the trypsin digest of band a. The labeled peaks were assigned to either GAT or GFP.

tein produced in the baculovirus system can be suitable for further structural analysis with correct folding and more uniform, less complex *N*-glycans.

Purification of the GAT1/GFP fusion protein from insect cells by immunoaffinity chromatography and size-exclusion (SE) chromatography

A two-step purification procedure for the GAT1/GFP-fusion protein from *Sf*9 cells was established. The GAT1/GFP expressed in *Sf*9 cells was first isolated with mAb-GFP antibody-conjugated affinity chromatography. Subsequently, eluted fractions containing the GAT1/GFP protein were pooled and subjected to fast protein liquid chromatography based on SE (SE-FPLC) with a Superdex 200^{TM} column to obtain purified homogeneous GAT1/GFP fusion protein.

To isolate the GAT1/GFP protein from the monoclonal anti-GFP antibody (mAb-GFP)-conjugated affinity column, different elution buffers (with different pH values and ionic strengths) were tested to obtain an effective and appropriate elution condition for the GAT1/GFP protein without irreversibly denaturing or inactivating it. Since no easy and effective method to control the activity of GAT1 protein during purification exists, a near-neutral high-salt buffer containing 4 M $MgCl_2$ (pH 6) (Figure 3) was selected to avoid the irreversible aggregation of GAT1, which may be caused by high pH values.

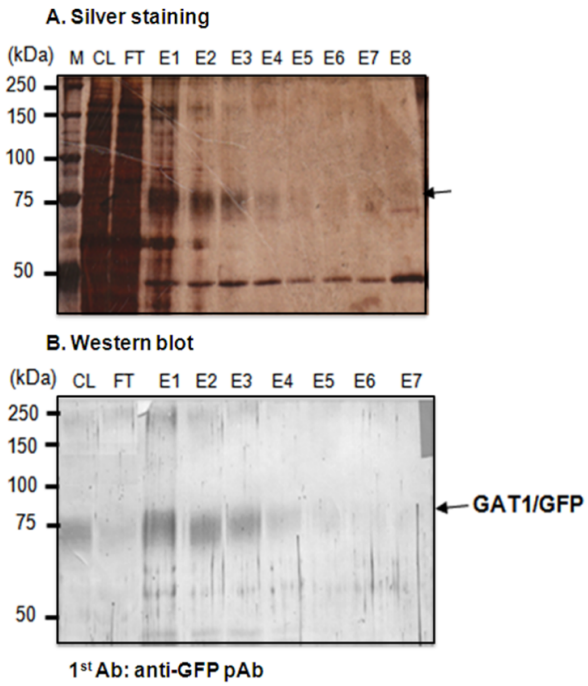


Figure 3: Isolation of the GAT1/GFP fusion protein from *Sf*9 cells by mAb-GFP-conjugated affinity column chromatography with analysis of the eluted fractions from the mAb-GFP-conjugated affinity column with 4 M $MgCl_2$ (pH 6) by SDS-PAGE with silver staining and Western blotting. The arrow denotes the fragment of mAb-GFP. M: standard marker; CL: cell lysate; FT: flow-through; E1–E8: eluted fractions 1–8.

A second purification step with SE-FPLC was performed to remove the eluted antibodies and other impurities from the eluates of the immunoaffinity column. The elution profile of the GAT1/GFP protein is shown in Figure 4A. Compared with the standard proteins (Figure 4B), two main peaks (1 and 2) appear, corresponding to M_r of 320 and 162 kDa, respectively. The fractions (300 µL per fraction) from 8.8 to 11.8 mL after SE-FPLC were further analyzed by SDS-PAGE (7.5%), followed with silver staining and Western blotting (Figure 4C and 4D). The results indicated that peak 1, which should correspond to a tetrameric GAT1/GFP with a molecular weight of 320 kDa, appeared at 8.8–10.6 mL (lanes 2–8). Peak 2 should correspond to a dimeric form with a molecular weight of 162 kDa, but despite its strong ultraviolet (UV) absorption, it contained very little GAT1/GFP protein according to the results of silver staining, Western blot and bicinchoninic acid (BCA) assay. To prevent the protein from forming oligomers during purification, different detergents were tested, and *n*-dodecyl-β-D-maltoside (DDM, 0.05%) was found to efficiently maintain the protein in its monomeric form (Figure 5A). The SE-FPLC elution profile is shown in Figure 5A. Peak 1, which appears at 13.5–14 mL, corresponds to a molecular weight of approximately 70 kDa, and the overlapping peak 2 (14.7 mL) corresponds to a molecular weight of 45 kDa, which was determined by comparison with the protein standards (Figure 4B).

Transmission electron microscopy (TEM) analysis

Cryogenic TEM (cryo-TEM) was employed to analyze both peaks from SE-FPLC to avoid the putative influences of sample drying and staining salts. The cryofixation by sample nitrifications is known to preserve the sample in the native state of the buffer environment and correspondingly the cryo-microscopy allows a direct visualization of the protein in the fully hydrated state. After a BCA control, the fraction of peak 2 containing the larger amount of protein was characterized by cryo-TEM and showed a very monodisperse distribution of particles (Figure 5B) with a diameter in the 5–6 nm range, which might correspond to protein monomers. The peak 1 fraction contains

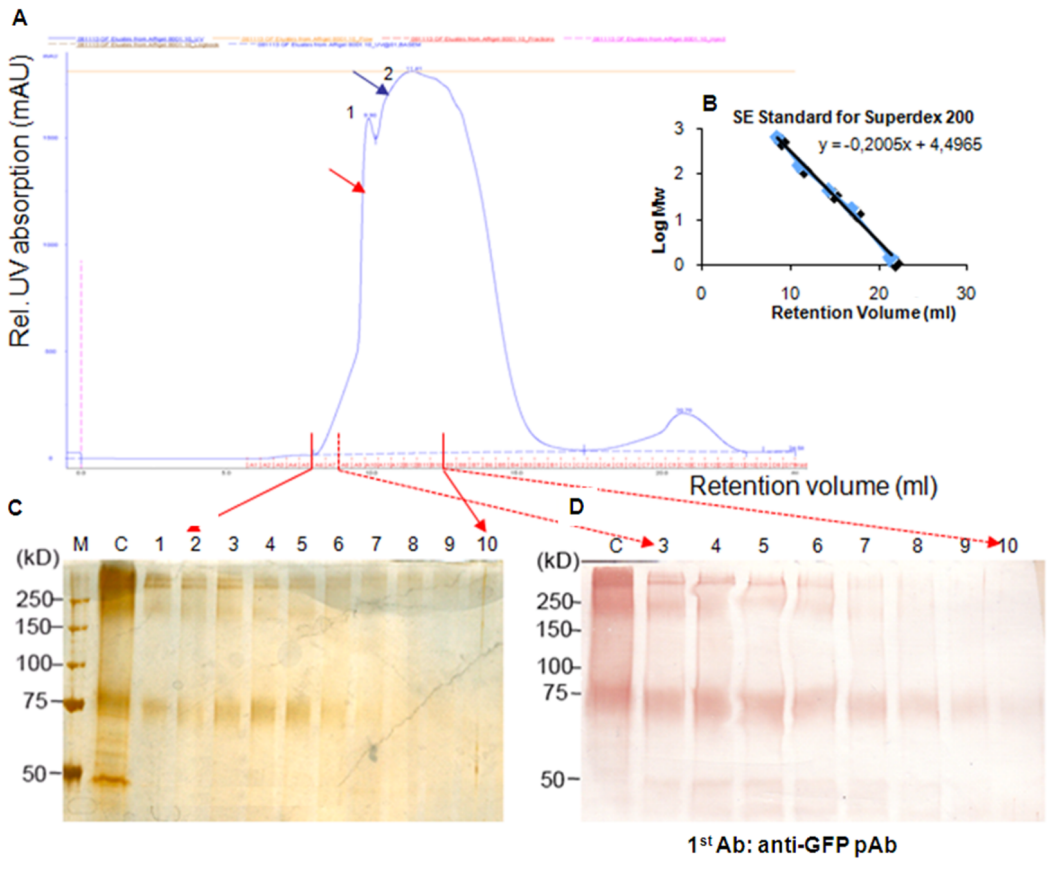


Figure 4: Purification of the GAT1/GFP fusion protein from *Sf*9 cells by SE-FPLC. (A) Elution profile of GAT1/GFP after SE-FPLC on a Superdex 200™ column. (B) The standard linear regression curve of the Superdex 200™ column was generated by plotting the log of the molecular masses of different calibration proteins against their elution volumes. (C), (D) Silver staining and Western blotting of the SDS-PAGE (7.5%) results of eluted fractions (from 8.8 to 11.8 mL) after SE-FPLC. M: standard marker; C: concentrated sample from immunoaffinity column; 1–10: SE-FPLC fractions from 8.8 to 10.8 mL.

only a low amount of the GAT1/GFP protein that exits partly in an aggregated form (Figure S4, Supporting Information File 1).

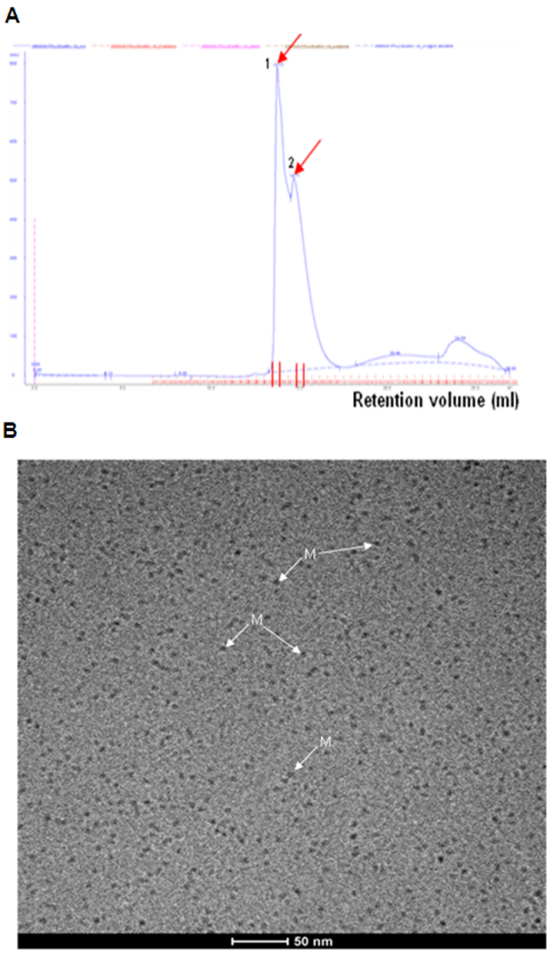


Figure 5: Characterization of the GAT1/GFP fusion protein after SE-FPLC. (A) Elution profile of GAT1/GFP protein after SE-FPLC on a Superdex 200^{TM} column. (B) Cryo-TEM image of fraction peak 2. Assumed monomeric GAT1/GFP fusion proteins (M, diameter = 5–6 nm) are indicated with white arrows. The control experiment revealed no significant population of spherical micelles.

Some noticeable small particles with high contrast were observed. The radius of DDM micelles has been reported to be approximately 2.6–3.5 nm [32], however, the observed high contrast is not typical for detergent micelles. As it was not clear whether the particles could simply be attributed to spherical micelles control cryo-TEM experiments were performed. Using a Tris-buffered saline (TBS) solution containing 0.05% DDM (which is far above the critical micellar concentration (CMC) of DDM:0.009% or 0.18 mM), no significant population of micelles could be found. These results suggest that the observed particles most likely correspond to a monodisperse preparation of GAT1/GFP. Several studies have demonstrated that GAT1 expresses in an oligomeric formation in the plasma membrane. GAT1 dimers were expressed and examined as a distinct population of 9 nm freeze-fracture particles in the plasma membrane of *Xenopuslaevis oocytes* by freeze-fracture and electron microscopy [33]. The GAT1 protein monomer was determined to be the functional unit since each monomer functions independently [33,34]. A similar example is the bacteria homologue LeuT, which was also crystallized as a dimer [13], however, each monomer has its own binding pocket, indicating that the monomers are the functional units. Therefore, a GAT1/GFP fusion protein monomer could be suitable for further structural analysis. A yield of approximately 200–300 µg of GAT1/GFP protein in this fraction was obtained from 400–600 mL of infected *Sf*9 cells, quantified by the BCA assay.

Conclusion

The aim of this project was to establish an expression and purification protocol for the production of high yields of GAT1/GFP fusion protein. In this work, the baculovirus expression system was used for the expression of the protein of interest. The full-length GAT1 protein is composed of twelve highly hydrophobic TM domains, which promote strong aggregation behavior of the protein, if it is isolated from the membrane. The presented protocol allows for the efficient production of the GAT1/GFP fusion protein in its monomeric form. We demonstrated that pure monomeric GAT1/GFP protein can be obtained with yields of approximately 200–300 µg from 400–600 mL of infected *Sf*9 cell culture. Moreover, considering the effect of *N*-linked glycans on the activity of the GAT1 protein, the glycol-engineered insect cells coupled with the baculovirus system may be further applied to produce a GAT1/GFP protein with complex, terminally sialylated *N*-glycans. Further structural analysis of the GAT1/GFP fusion protein is possible using crystallography, thereby improving our understanding of the three-dimensional structure of the GAT1 protein.

Experimental

Immunoprecipitation and Western blotting

The immunoprecipitation and Western blotting procedures were similar to those described previously [12]. Briefly, the GAT1/GFP protein was solubilized at 4 °C or on ice. After infection, insect cells were collected, washed once with phosphate-buffered saline (PBS), and resuspended in TBS buffer (50 mM Tris, pH 7.3, 150 mM NaCl). The suspended cells were sonicated at 4 °C for 15 min and cell debris was removed by centrifugation at 6,000*g* for 10 min at 4 °C. A turbid supernatant solution containing the cell membranes was obtained. After centrifugation at 100,000*g* at 4 °C for 30 min, the crude membrane fractions were solubilized in TBS containing 1% DDM and stirred for at least 4 h at 4 °C. The lysate was centrifuged at 18,000*g* at 4 °C for 1 h. Total protein concentrations of the supernatant were measured with BCA™ Protein

Assay Kit (Thermo). Quantified aliquots of the supernatants were incubated with protein-G-Sepharose-bound anti-GFP IgG for 12 h at 4 °C. After intensive washing, the immunoprecipitates were eluted by boiling for 3 min in SDS sample buffer. The supernatant aliquots were divided in half and then subjected to SDS-PAGE; the separated proteins were transferred to a nitrocellulose membrane (Millipore) by Western blotting. One blot membrane was used for the immunostaining of the GAT1 protein with the anti-GAT1 or anti-GFP polyclonal antiserum. Subsequently, the blots were incubated with horseradish peroxidase-conjugated anti-rabbit antibody (IgG) (Dako Cytomation) and then visualized using amino ethylcarbazole (AEC) and substrate buffer (Calbiochem).

Flow cytometry and fluorescence microscopy

*Sf*9 cells were observed after 3-day infection by flow cytometry and fluorescence microscopy to determine the cell surface expression of the GAT1/GFP fusion proteins.

MALDI mass fingerprinting

Using MALDI–TOF MS, protein fragments with blocked N-termini or that are available in limited concentrations can be easily analyzed. The MALDI mass fingerprinting was performed by Dr. Chris Weise (Free University Berlin, Germany) with a Bruker-Biflex Reflex Mass spectrometer (Bruker Daltonics) in the reflector-mode with alpha-cyano-4-hydroxycinnamic acid as matrix. Ionization was enhanced with the 337 nm-ray of a nitrogen laser. The peptide masses were determined by calibration with the PAC peptide calibrant standard. All solutions and buffers used were prepared with sterile pure high-performance LC (HPLC)-grade water (Milli-Q® Water filter apparatus, Millipore) in order to avoid contaminations. Evaluation and identification of the mass spectra were performed using the Internet search software Mascot [35].

Purification of GAT1 and protein identification

Isolation of GAT1/GFP by the immunoaffinity column

The immunoaffinity column was prepared as followed: 200 µL Affi-gel 10 protein-A-sepharose (Amersham Pharmacia, Sweden) were loaded on a 10 mL disposable chromatography column and washed twice with 1 mL ice cold double-distilled (dd)H_2O. Approximately 6–8 mg of purified antibody in 0.1 M MOPS buffer, pH 7.5, was added to the resin, and the columns were incubated overnight at 4 °C to allow covalent binding. Subsequently, unbound antibodies were collected as the flow-through, and the sepharose column was washed once with 1 mL of 1 M ethanolamine (pH 8.0). Blocking of the sepharose was performed with 3 mL of 1 M ethanolamine (pH 8.0) for 2 h at room temperature (rt). Finally, the antibody-coupled sepharose was washed three times with 15 mM sodium phosphate buffer (pH 8.0). Prior to subsequent use of the affinity columns, 1 mL of PBS with 0.02% NaN_3 was added to the sepharose, which was then stored at 4 °C.

Cells were lysed in solubilization buffer (10 mM Tris, pH 7.8, 150 mM NaCl, 1 mM $CaCl_2$, 2% DDM) containing 1 mM dithiothreitol (DTT) and protease inhibitor cocktail (1:500) (Merck). After a 4 × 30 s sonication step, the cells were incubated overnight at 4 °C with agitation. The solubilized protein was fractionated by 45 min of centrifugation at 18,000 rpm.

The cell lysate was added directly to the immunoaffinity chromatography resin. Proteins were coupled on the column by overnight agitation at 4 °C. Unbound proteins were collected as flow-through and the resin was washed three times with radio immunoprecipitation assay buffer (RIPA) and twice with prewash. The elution of the GAT1/GFP was achieved with 8 × 250 µL of elution buffer (50 mM diethylamine, pH 11.4). The eluates were neutralized immediately by adding 100 µL of 0.5 M NaH_2PO_4 to each tube in which the eluates were collected.

FPLC based on size exclusion (SE-FPLC)

The immunopurified GAT1/GFP isolated from the immunoaffinity column was concentrated to 250–300 µL with a Vivaspin column (Vivasciences) at 4 °C and 3,000*g*. The concentrate was further purified by SE-FPLC on a Superdex 200 column (GE Healthcare) that had been equilibrated with equilibrium buffer (10 mM Tris, pH 7.8, 150 mM NaCl, and 0.05% DDM). Elution of the protein was performed with equilibrium buffer at a flow rate of 0.3 mL/min. The molecular weight of the obtained protein was determined based on the elution profile of proteins standards obtained under the same buffer conditions.

TEM analysis

Negative staining preparation

Sample droplets (5 µL) of the sample were placed onto hydrophilized (glow discharged for 60 s at 8 W in a BALTEC MED 020 device (Baltec, Liechtenstein) carbon-covered microscopical copper grids (400 mesh), and the supernatant fluid was removed with a filter paper to create an ultrathin layer of the sample. A droplet of contrasting material (1% uranyl acetate, 2% phosphotungstic acid or 2% ammonium molybdate in the presence of 0.1% trehalose) was added, blotted again and air-dried. Imaging was performed using a Tecnai F20 FEG (FEI Company, Oregon) at an accelerating voltage of 160 kV under low-dose conditions. Micrographs were recorded according to the low-dose protocol of the microscope at a primary magnification of 62,000×. The defocus value was chosen to correspond to the first zero of the contrast transfer function (CTF) at ≈15 Å.

Cryo-TEM

Similarly as described previously [36], vitrified samples were transferred into a Tecnai F20 FEG using a Gatan cryo-holder and -stage (Model 626). Samples were constantly cooled by LN2 during imaging to maintain a sample temperature of $T = 93$ K. Imaging was performed at an accelerating voltage of 160 kV with a defocus value of 600 nm, which corresponds to the first zero of the CTF at 13 Å (Cs = 2.0 mm). Micrographs were recorded according to the low-dose protocol of the microscope at a primary magnification of 62, 000×.

Supporting Information

Supporting Information File 1

Additional information.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-13-88-S1.pdf]

Acknowledgements

We thank the SFB 449 (Project B4) from the German Research Foundation (DFG) and the Natural Science Foundation of Jiangsu Province, China (BK20150140) and the Public Health Research Center at Jiangnan University (No. JUPH201502).

References

1. Kanner, B. I. *J. Membr. Biol.* **2006,** *213,* 89–100. doi:10.1007/s00232-006-0877-5
2. Nelson, N. *J. Neurochem.* **1998,** *71,* 1785–1803. doi:10.1046/j.1471-4159.1998.71051785.x
3. Scimemi, A. *Front. Cell. Neurosci.* **2014,** *8,* No. 161. doi:10.3389/fncel.2014.00161
4. Liu, Q. R.; Lopez-Corcuera, B.; Mandiyan, S.; Nelson, H.; Nelson, N. *J. Biol. Chem.* **1993,** *268,* 2106–2112.
5. Borden, L. A.; Smith, K. E.; Hartig, P. R.; Branchek, T. A.; Weinshank, R. L. *J. Biol. Chem.* **1992,** *267,* 21098–21104.
6. Jursky, F.; Tamura, S.; Tamura, A.; Mandiyan, S.; Nelson, H.; Nelson, N. *J. Exp. Biol.* **1994,** *196,* 283–295.
7. Borden, L. A. *Neurochem. Int.* **1996,** *29,* 335–356. doi:10.1016/0197-0186(95)00158-1
8. Conti, F.; Minelli, A.; Melone, M. *Brain Res. Rev.* **2004,** *45,* 196–212. doi:10.1016/j.brainresrev.2004.03.003
9. Guastella, J.; Nelson, N.; Nelson, H.; Czyzyk, L.; Keynan, S.; Miedel, M. C.; Davidson, N.; Lester, H. A.; Kanner, B. I. *Science* **1990,** *249,* 1303–1306. doi:10.1126/science.1975955
10. Bennett, E. R.; Kanner, B. I. *J. Biol. Chem.* **1997,** *272,* 1203–1210. doi:10.1074/jbc.272.2.1203
11. Cai, G.; Salonikidis, P. S.; Fei, J.; Schwarz, W.; Schülein, R.; Reutter, W.; Fan, H. *FEBS J.* **2005,** *272,* 1625–1638. doi:10.1111/j.1742-4658.2005.04595.x
12. Hu, J.; Fei, J.; Reutter, W.; Fan, H. *Glycobiology* **2011,** *21,* 329–339. doi:10.1093/glycob/cwq166
13. Yamashita, A.; Singh, S. K.; Kawate, T.; Jin, Y.; Gouaux, E. *Nature* **2005,** *437,* 215–223. doi:10.1038/nature03978
14. Radian, R.; Kanner, B. I. *Biochemistry* **1983,** *22,* 1236–1241. doi:10.1021/bi00274a038
15. Zomot, E.; Bendahan, A.; Quick, M.; Zhao, Y.; Javitch, J. A.; Kanner, B. I. *Nature* **2007,** *449,* 726–730. doi:10.1038/nature06133
16. Forrest, L. R.; Tavoulari, S.; Zhang, Y.-W.; Rudnick, G.; Honig, B. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104,* 12761–12766. doi:10.1073/pnas.0705600104
17. Faham, S.; Watanabe, A.; Besserer, G. M.; Cascio, D.; Specht, A.; Hirayama, B. A.; Wright, E. M.; Abramson, J. *Science* **2008,** *321,* 810–814. doi:10.1126/science.1160406
18. Weyand, S.; Shimamura, T.; Yajima, S.; Suzuki, S.; Mirza, O.; Krusong, K.; Carpenter, E. P.; Rutherford, N. G.; Hadden, J. M.; O'Reilly, J.; Ma, P.; Saidijam, M.; Patching, S. G.; Hope, R. J.; Norbertczak, H. T.; Roach, P. C. J.; Iwata, S.; Henderson, P. J. F.; Cameron, A. D. *Science* **2008,** *322,* 709–713. doi:10.1126/science.1164440
19. Abramson, J.; Wright, E. M. *Curr. Opin. Struct. Biol.* **2009,** *19,* 425–432. doi:10.1016/j.sbi.2009.06.002
20. Ressl, S.; Terwisscha van Scheltinga, A. C.; Vonrhein, C.; Ott, V.; Ziegler, C. *Nature* **2009,** *458,* 47–52. doi:10.1038/nature07819
21. Möhler, H. *J. Recept. Signal Transduction* **2006,** *26,* 731–740. doi:10.1080/10799890600920035
22. Cope, D. W.; Di Giovanni, G.; Fyson, S. J.; Orbán, G.; Errington, A. C.; Lőrincz, M. L.; Gould, T. M.; Carter, D. A.; Crunelli, V. *Nat. Med.* **2009,** *15,* 1392–1398. doi:10.1038/nm.2058
23. Krystal, J. H.; Sanacora, G.; Blumberg, H.; Anand, A.; Charney, D. S.; Marek, G.; Epperson, C. N.; Goddard, A.; Mason, G. F. *Mol. Psychiatry* **2002,** *7* (Suppl. 1), S71–S80. doi:10.1038/sj.mp.4001021
24. Garcia-Alloza, M.; Tsang, S. W.; Gil-Bea, F. J.; Francis, P. T.; Lai, M. K.; Marcos, B.; Chen, C. P.; Ramirez, M. J. *Neurobiol. Aging* **2006,** *27,* 1110–1117. doi:10.1016/j.neurobiolaging.2005.06.003
25. Bröer, S. *Neurochem. Int.* **2006,** *48,* 559–567. doi:10.1016/j.neuint.2005.11.021
26. Madsen, K. K.; Clausen, R. P.; Larsson, O. M.; Krogsgaard-Larsen, P.; Schousboe, A.; White, H. S. *J. Neurochem.* **2009,** *109* (Suppl. 1), 139–144. doi:10.1111/j.1471-4159.2009.05982.x
27. Chiu, C. S.; Jensen, K.; Sokolova, I.; Wang, D.; Li, M.; Deshpande, P.; Davidson, N.; Mody, I.; Quick, M. W.; Quake, S. R.; Lester, H. A. *J. Neurosci.* **2002,** *22,* 10251–10266.
28. Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996,** *273,* 1392–1395. doi:10.1126/science.273.5280.1392
29. Yang, F.; Moss, L. G.; Phillips, G. N., Jr. *Nat. Biotechnol.* **1996,** *14,* 1246–1251. doi:10.1038/nbt1096-1246
30. van Oers, M. M.; Pijlman, G. P.; Vlak, J. M. *J. Gen. Virol.* **2015,** *96,* 6–23. doi:10.1099/vir.0.067108-0
31. Contreras-Gómez, A.; Sánchez-Mirón, A.; García-Camacho, F.; Molina-Grima, E.; Chisti, Y. *Biotechnol. Prog.* **2014,** *30,* 1–18. doi:10.1002/btpr.1842
32. Lipfert, J.; Columbus, L.; Chu, V. B.; Lesley, S. A.; Doniach, S. *J. Phys. Chem. B* **2007,** *111,* 12427–12438. doi:10.1021/jp073016l
33. Gonzales, A. L.; Lee, W.; Spencer, S. R.; Oropeza, R. A.; Chapman, J. V.; Ku, J. Y.; Eskandari, S. *J. Membr. Biol.* **2007,** *220,* 33–51. doi:10.1007/s00232-007-9073-5
34. Soragna, A.; Bossi, E.; Giovannardi, S.; Pisani, R.; Peres, A. *Cell. Mol. Life Sci.* **2005,** *62,* 2877–2885. doi:10.1007/s00018-005-5322-x

35. Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* **1999,** *20,* 3551–3567. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2
36. Ludwig, K.; Yan, S.; Fan, H.; Reutter, W.; Böttcher, C. *Biochem. Biophys. Res. Commun.* **2003,** *304,* 73–77. doi:10.1016/S0006-291X(03)00539-4

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Total synthesis of TMG-chitotriomycin based on an automated electrochemical assembly of a disaccharide building block

Yuta Isoda[1], Norihiko Sasaki[1], Kei Kitamura[1], Shuji Takahashi[1], Sujit Manmode[1], Naoko Takeda-Okuda[2], Jun-ichi Tamura[1,2,3], Toshiki Nokami[*1,3] and Toshiyuki Itoh[*1,3]

Full Research Paper

Open Access

Address:
[1]Department of Chemistry and Biotechnology, Graduate School of Engineering, Tottori University, 4-101 Koyama-minami, Tottori 680-8552, Japan, [2]Department of Regional Environment, Faculty of Regional Sciences, Tottori University, 4-101 Koyama-minami, Tottori 680-8551, Japan and [3]Center for Research on Green Sustainable Chemistry, Faculty of Engineering, Tottori University, 4-101 Koyama-minami, Tottori 680-8552, Japan

Email:
Toshiki Nokami* - tnokami@chem.tottori-u.ac.jp; Toshiyuki Itoh* - titoh@chem.tottori-u.ac.jp

* Corresponding author

Keywords:
automated synthesis; electrochemical oxidation; glycosylation; glucosamine; total synthesis

Beilstein J. Org. Chem. **2017,** *13,* 919–924.
doi:10.3762/bjoc.13.93

Received: 28 February 2017
Accepted: 04 May 2017
Published: 16 May 2017

This article is part of the Thematic Series "Biomolecular systems" and is dedicated to Professor Peter H. Seeberger on the occasion of his 50th birthday.

Guest Editor: P. H. Seeberger

Abstract

The total synthesis of TMG-chitotriomycin using an automated electrochemical synthesizer for the assembly of carbohydrate building blocks is demonstrated. We have successfully prepared a precursor of TMG-chitotriomycin, which is a structurally-pure tetrasaccharide with typical protecting groups, through the methodology of automated electrochemical solution-phase synthesis developed by us. The synthesis of structurally well-defined TMG-chitotriomycin has been accomplished in 10-steps from a disaccharide building block.

Introduction

Degradation of chitin into oligoglucosamines and glucosamine is an important biological process and at least several enzymes such as chitinases and glucosaminidases are involved. Various types of inhibitors such as PUGNAc [1], nagastatin [2], NAG-thiazoline (NGT) [3], and pochonicine [4], have already been developed. These compounds have exhibited strong inhibition activity; however, they have a broad spectrum toward enzymes of various species including animals (Figure 1). *N*,*N*,*N*-Trimethyl-D-glucosaminyl (TMG)-chitotriomycin (**1**) was isolated from *Streptomyces anulatus* by Kanzaki [5-7] and the first total synthesis was completed by Yu [8,9]. Although the activity of TMG-chitotriomycin (**1**) was moderate, it selectively

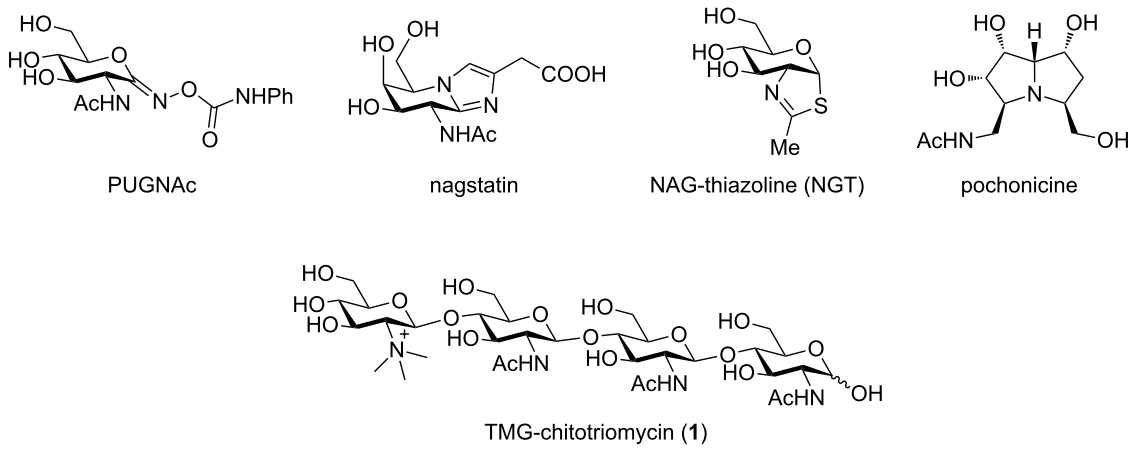


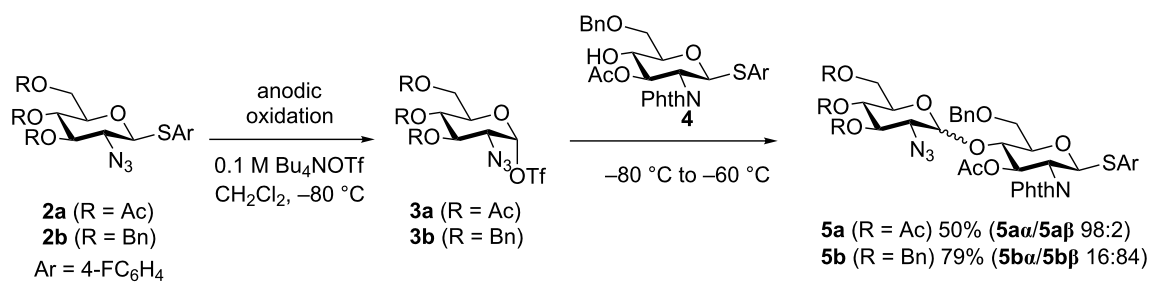
Figure 1: Inhibitors of glucosaminidases.

inhibits glucosaminidases derived from insects and fungi. Therefore, TMG-chitotriomycin (**1**) has a potential as a lead compound for safe insecticides and pesticides. Recently, the total synthesis of TMG-chitotriomycin (**1**) initiated from building blocks that were prepared by the degradation of chitin has been reported by Beau [10]; however, practical synthetic methods to provide TMG-chitotriomycin (**1**) and its derivative in preparative scale are still highly desirable.

The automated synthesis of oligosaccharides is a powerful tool for the rapid synthesis of complex oligosaccharides [11-23]. We are interested in the automated synthesis of oligosaccharides based on the concept of "reaction integration" [24] and developed an automated synthesizer for the automated electrochemical assembly of carbohydrate building blocks [25]. During the course of our study, we have achieved the synthesis of a potential precursor of TMG-chitotriomycin (**1**); however, we obtained the tetrasaccharide as a mixture of α- and β-isomers in the terminal glycosidic linkage [26]. Here we report the total synthesis of TMG-chitotriomycin (**1**) as a single stereoisomer, which was prepared by automated electrochemical assembly started from a disaccharide building block.

Results and Discussion

To synthesize the potential precursor **7** of TMG-chitotriomycin (**1**) stereoselectively, we initiated our study by optimization of the reaction conditions of the first glycosylation using 2-deoxy-2-azidothioglycoside **2** as a glycosyl donor. The azido group at the C2-position is a well-known substituent, which facilitates the formation of an α-glycosidic linkage selectively due to the lack of neighboring group participation [27]. 4-Fluorophenyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-azido-β-D-thioglucoside (**2a**) afforded the corresponding disaccharide α-isomer **5aα** exclusively by the reaction with building block **4** via the glycosyl triflate intermediate **3a** (Scheme 1). On the other hand, 4-fluorophenyl 3,4,6-tri-*O*- benzyl-2-deoxy-2-azido-β-D-thioglucoside (**2b**) gave the disaccharide β-isomer **5bβ** as the major product. Although the disaccharide α- and β-isomers of **5b** (**5bα**/**5bβ**) have the same retention factor (R_f) in thin-layer chromatography (TLC, R_f = 0.40, hexane/EtOAc 5:2 as an



Scheme 1: Synthesis of disaccharide donors.

eluent), the pure β-disaccharide isomer **5bβ** was obtained as colorless crystals after careful separation by silica gel chromatography.

It is still not clear why the disaccharide β-isomer **5bβ** was obtained as a major product from glycosyl triflate **3**. We now assume that the difference in the reactivity of the α- and the β-glycosyl triflate intermediates **3** might determine the observed selectivity (Figure 2). We previously established that glycosyl triflate **3a** was derived from thioglycoside **2a** by an NMR study under low-temperature conditions, in which the glycosyl triflate α-isomer **3aα** was confirmed as an exclusive chemical species [28,29]. Taking the stability of the glycosyl triflate α-isomer **3aα** into consideration, we propose a reaction mechanism involving α/β isomerization of glycosyl triflate **3a** as shown in Figure 2. The more reactive β-isomer **3aβ** might give the disaccharide α-isomer **5aα** exclusively if there is an equilibrium between the α-isomer and the β-isomer of **3a**. To the contrary, glycosyl triflate **3b**, derived from thioglycoside **2b**, might be more reactive and affords the β-product **5bβ** before isomerization from the α-isomer **3bα** to the β-isomer **3bβ**. In this case, glycosylation via **3bα** becomes the major pathway [30]. Although it is hard to exclude another reaction mechanism involving oxocarbenium ions as reactive intermediates, the commonly accepted reactivity difference between α- and β-isomers of glycosyl triflate **3** seems to explain the observed selectivity well.

Next, we attempted to synthesize the potential precursor **7** of TMG-chitotriomycin (**1**) using disaccharide **5bβ** as a building block as illustrated in Figure 3. The automated electrochemical assembly of building blocks was initiated by the anodic oxidation of **5bβ** and the subsequent coupling with thioglycoside **4** afforded the corresponding trisaccharide **6** as an intermediate after the 1st cycle. The same process was repeated automatically in the 2nd cycle and target tetrasaccharide **7** was obtained in 41% yield after purification by preparative gel permeation chromatography (GPC).

Deprotection and introduction of the TMG part to tetrasaccharide **7** were achieved by following the procedure reported by Yu and co-workers (Figure 4) [8]. Phthaloyl groups and acetyl groups of **7** are removed by the reaction with ethylenediamine followed by applying the conventional acetylation protocol with acetic anhydride (Ac_2O) in the presence of *N*,*N*-dimethylaminopyridine (DMAP) to convert the amino groups into acetamide

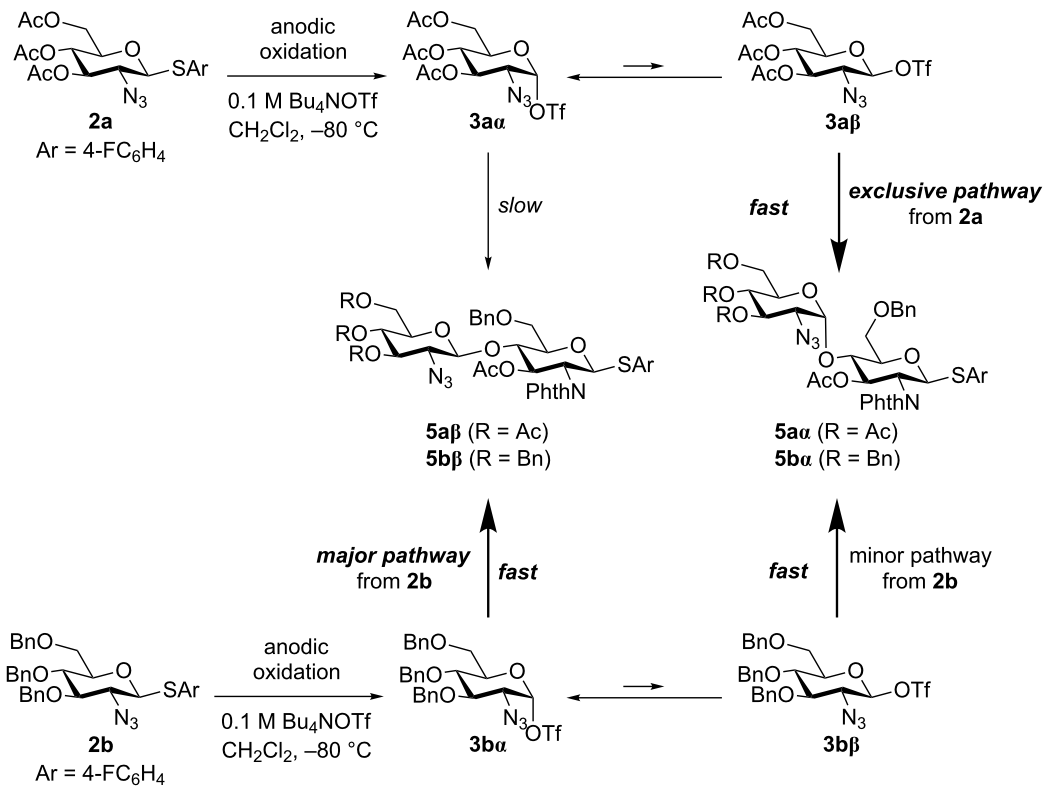


Figure 2: Proposed mechanism and origin of the selectivity.

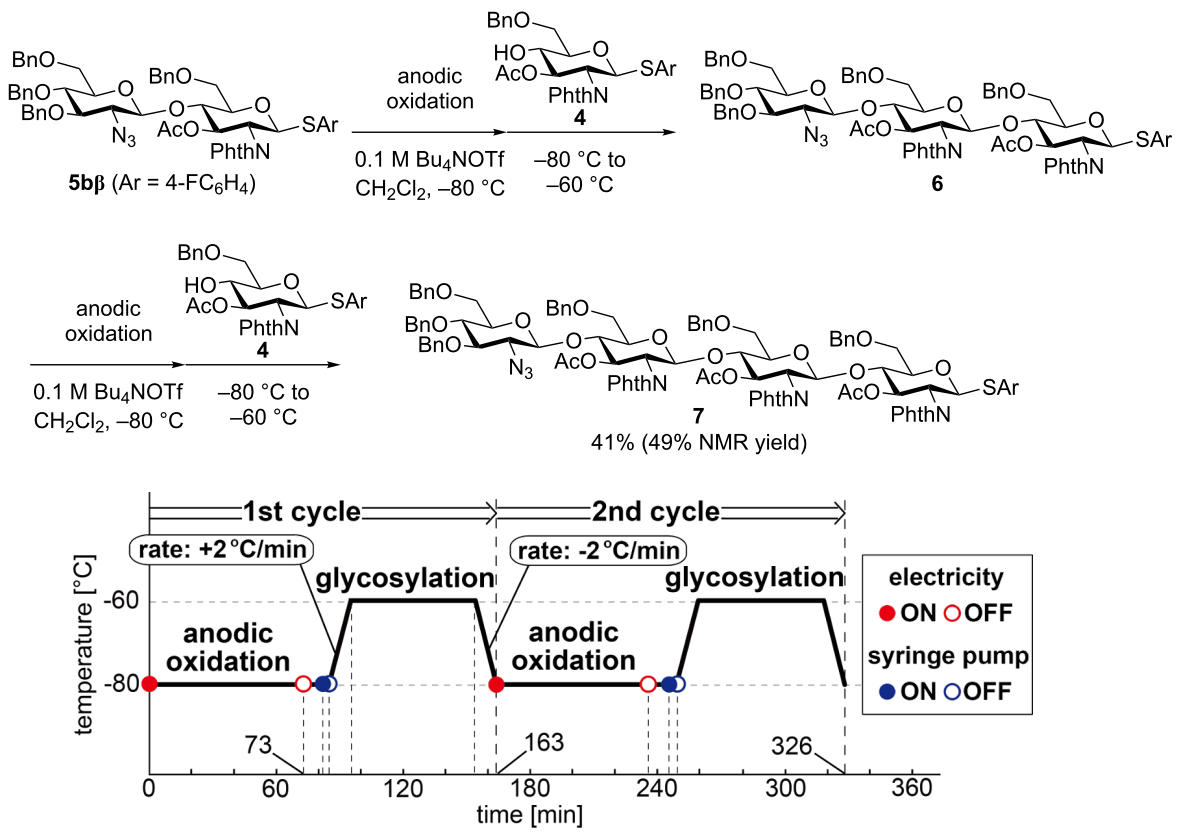


Figure 3: Synthesis of TMG-chitotriomycin precursor **7**.

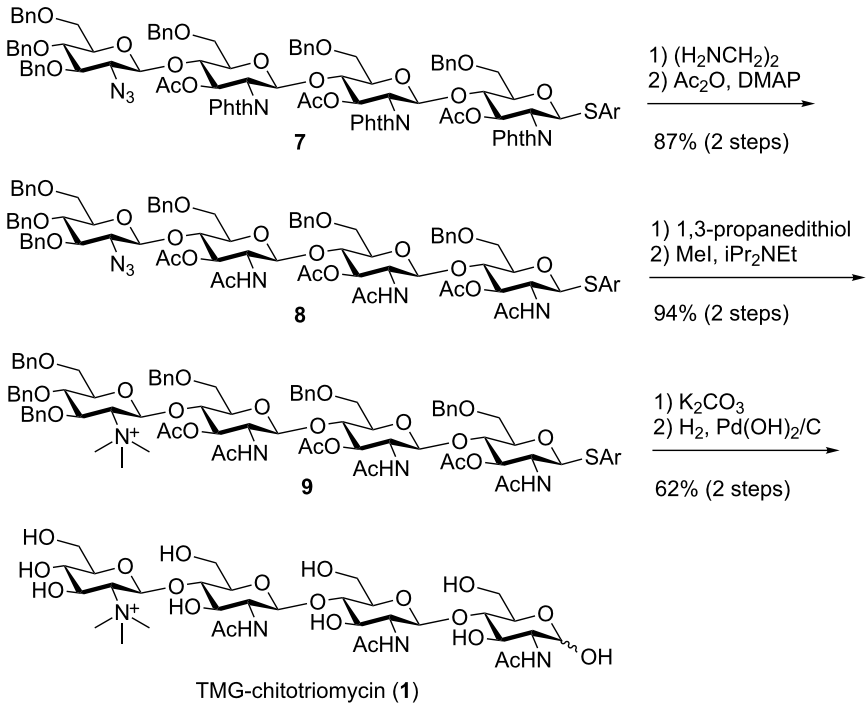


Figure 4: Synthess of TMG-chitotriomycin (**1**).

groups and to protect the hydroxy groups as acetyl groups. The 2-azido group of tetrasaccharide **8** was then reduced to a 2-amino group with 1,3-propanedithiol. Thus-obtained tetrasaccharide was treated with iodomethane (MeI) and *N*,*N*-diisopropylethylamine (iPr_2NEt) to prepare the TMG part of tetrasaccharide **9**. Deprotection of acetyl groups at the 3-*O*-positions and the subsequent global deprotection of the benzyl groups and the anomeric thioaryl group of the tetrasaccharide by hydrogenation with hydrogen gas in the presence of a palladium catalyst afforded TMG-chitotriomycin (**1**) in 21% yield (10 steps from disaccharide **5bβ**) [31].

Conclusion

In conclusion, we have achieved the stereoselective synthesis of TMG-chitotriomycin (**1**) based on the automated electrochemical assembly of disaccharide and monosaccharide building blocks. Thus-obtained structurally well-defined tetrasaccharide gave TMG-chitotriomycin after manipulations of the amino groups and global deprotection. Further investigations to improve the β-selectivity in the disaccharide synthesis and a large scale synthesis are in progress in our laboratory.

Experimental

General

^1H and ^{13}C NMR spectra were recorded in $CDCl_3$ on a Bruker AVANCE II 600 spectrometer (^1H 600 MHz, ^{13}C 150 MHz) with Me_4Si as an internal standard unless otherwise noted. Mass spectra were obtained on a Thermo Scientific Exactive mass spectrometer. Thin-layer chromatography (TLC) was carried out by using Merck precoated silica gel F254 plates (thickness 0.25 mm). Flash chromatography was carried out on a column of silica gel (Kanto Chem. Co., Silica Gel N, spherical, neutral). Gel permeation chromatography (GPC) was carried out on a Japan Analytical Industry LC-918 equipped with JAIGEL-2H using $CHCl_3$ as eluent. All reactions were performed under an Ar atmosphere unless otherwise noted.

Materials

All materials including solvents were purchased and used without further purification. Carbohydrate building blocks **2a**, **2b**, and **4** were prepared according to an previous report [11].

Automated synthesis of TMG-chitotriomycin precursor 7

The H-type glass cell equipped with glass filter was dried under vacuum and then filled with argon gas. Disaccharide building block **5bβ** (0.43 mmol, 436 mg) was placed in the anodic chamber together with Bu_4NOTf (1.7 mmol, 672 mg) and anhydrous dichloromethane (16 mL). In the cathodic chamber Bu_4NOTf (1.7 mmol, 676 mg) and anhydrous dichloromethane (16 mL) was placed with TfOH (1 mmol, 90 μL). The automated synthesis was started immediately after the cooling bath temperature reached −80 °C. The anodic oxidation (1.05 F/mol, 10 mA) takes 73 minutes and then a dichloromethane solution containing building block **4** (0.43 mmol, 1.0 mL) was added by a syringe pump. After addition of the building block, the temperature of the cooling bath was raised to −60 °C and the reaction mixture was stirred for 1 h at this temperature. Then the temperature of the cooling bath was cooled again at −80 °C to start the 2nd cycle. It takes about 326 minutes (ca. 5 h 30 min) to complete the automated assembly of building blocks. The reaction was quenched with Et_3N (0.5 mL) at −80 °C and the cell was taken from the cooling bath. The reaction mixture of the anodic chamber was evaporated and the thus-obtained crude product was purified by silica gel chromatography (hexane/EtOAc 1:1 as eluent). The precursor **7** was obtained in 49% yield (411 mg) together with building block **4** as an impurity. Further purification by GPC ($CHCl_3$ as eluent) afforded **7** in 41% yield (328 mg, 0.178 mmol).

Supporting Information

Supporting Information File 1

Experimental details of electrochemical glycosylation, global deprotection, and NMR spectra of unknown compounds.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-13-93-S1.pdf]

Acknowledgements

This work is financially supported by JSPS Grant-in-Aid for Scientific Research on Innovative Areas (No. 2707) “Middle Molecular Strategy” Grant Number JP15H05844. T. N. thanks discretionary funds of the president of Tottori University.

References

1. Horsch, M.; Hoesch, L.; Vasella, A.; Rast, D. M. *Eur. J. Biochem.* **1991,** *197,* 815–818. doi:10.1111/j.1432-1033.1991.tb15976.x
2. Aoyagi, T.; Suda, H.; Uotani, K.; Kojima, F.; Aoyama, T.; Horiguchi, K.; Hamada, M.; Takeuchi, T. *J. Antibiot.* **1992,** *45,* 1404–1408. doi:10.7164/antibiotics.45.1404
3. Knapp, S.; Vocadlo, D.; Gao, Z.; Kirk, B.; Lou, J.; Withers, S. G. *J. Am. Chem. Soc.* **1996,** *118,* 6804–6805. doi:10.1021/ja960826u
4. Usuki, H.; Toyo-oka, M.; Kanzaki, H.; Okuda, T.; Nitoda, T. *Bioorg. Med. Chem.* **2009,** *17,* 7248–7253. doi:10.1016/j.bmc.2009.08.052
5. Usuki, H.; Nitoda, T.; Ichikawa, M.; Yamaji, N.; Iwashita, T.; Komura, H.; Kanzaki, H. *J. Am. Chem. Soc.* **2008,** *130,* 4146–4152. doi:10.1021/ja077641f
6. Usuki, H.; Yamamoto, Y.; Kumagai, Y.; Nitoda, T.; Kanzaki, H.; Hatanaka, T. *Org. Biomol. Chem.* **2011,** *9,* 2943–2951. doi:10.1039/C0OB01090A

7. Shiota, H.; Kanzaki, H.; Hatanaka, T.; Nitoda, T. *Carbohydr. Res.* **2013,** *375,* 29–34. doi:10.1016/j.carres.2013.04.024
8. Yang, Y.; Li, Y.; Yu, B. *J. Am. Chem. Soc.* **2009,** *131,* 12076–12077. doi:10.1021/ja9055245
9. Yang, Y.; Liu, T.; Yang, Y.; Wu, Q.; Yang, Q.; Yu, B. *ChemBioChem* **2011,** *12,* 457–467. doi:10.1002/cbic.201000561
10. Despras, G.; Alix, A.; Urban, D.; Vauzeilles, B.; Beau, J.-M. *Angew. Chem., Int. Ed.* **2014,** *53,* 11912–11916. doi:10.1002/anie.201406802
11. Plante, O. J.; Palmacci, E. R.; Seeberger, P. H. *Science* **2001,** *291,* 1523–1527. doi:10.1126/science.1057324
12. Werz, D. B.; Castagner, B.; Seeberger, P. H. *J. Am. Chem. Soc.* **2007,** *129,* 2770–2771. doi:10.1021/ja069218x
13. Walvoort, M. T. C.; Volbeda, A. G.; Reintjens, N. R. M.; van den Elst, H.; Plante, O. J.; Overkleeft, H. S.; van der Marel, G. A.; Codée, J. D. C. *Org. Lett.* **2012,** *14,* 3776–3779. doi:10.1021/ol301666n
14. Kröck, L.; Esposito, D.; Castagner, B.; Wang, C.-C.; Bindschädler, P.; Seeberger, P. H. *Chem. Sci.* **2012,** *3,* 1617–1622. doi:10.1039/C2SC00940D
15. Ganesh, N. V.; Fujikawa, K.; Tan, Y. H.; Stine, K. J.; Demchenko, A. V. *Org. Lett.* **2012,** *14,* 3036–3039. doi:10.1021/ol301105y
16. Calin, O.; Eller, S.; Seeberger, P. H. *Angew. Chem., Int. Ed.* **2013,** *52,* 5862–5865. doi:10.1002/anie.201210176
17. Eller, S.; Collot, M.; Yin, J.; Hahm, H. S.; Seeberger, P. H. *Angew. Chem., Int. Ed.* **2013,** *52,* 5858–5861. doi:10.1002/anie.201210132
18. Tang, S.-L.; Pohl, N. L. B. *Org. Lett.* **2015,** *17,* 2642–2645. doi:10.1021/acs.orglett.5b01013
19. Schmidt, D.; Schuhmacher, F.; Geissner, A.; Seeberger, P. H.; Pfrengle, F. *Chem. – Eur. J.* **2015,** *21,* 5709–5713. doi:10.1002/chem.201500065
20. Fair, R. J.; Hahm, H. S.; Seeberger, P. H. *Chem. Commun.* **2015,** *51,* 6183–6185. doi:10.1039/C5CC01368B
21. Pistorio, S. G.; Nigudkar, S. S.; Stine, K. J.; Demchenko, A. V. *J. Org. Chem.* **2016,** *81,* 8796–8805. doi:10.1021/acs.joc.6b01439
22. Hahm, H. S.; Hurevich, M.; Seeberger, P. H. *Nat. Commun.* **2016,** *7,* No. 12482. doi:10.1038/ncomms12482
23. Hahm, H. S.; Broecker, F.; Kawasaki, F.; Mietzsch, M.; Heilbronn, R.; Fukuda, M.; Seeberger, P. H. *Chem* **2017,** *2,* 114–124. doi:10.1016/j.chempr.2016.12.004
24. Yoshida, J.-i.; Shimizu, A.; Ashikari, Y.; Morofuji, T.; Hayashi, R.; Nokami, T.; Nagaki, A. *Bull. Chem. Soc. Jpn.* **2015,** *88,* 763–775. doi:10.1246/bcsj.20150100
25. Nokami, T.; Hayashi, R.; Saigusa, Y.; Shimizu, A.; Liu, C.-Y.; Mong, K.-K. T.; Yoshida, J.-i. *Org. Lett.* **2013,** *15,* 4520–4523. doi:10.1021/ol402034g
26. Nokami, T.; Isoda, Y.; Sasaki, N.; Takaiso, A.; Hayase, S.; Itoh, T.; Hayashi, R.; Shimizu, A.; Yoshida, J.-i. *Org. Lett.* **2015,** *17,* 1525–1528. doi:10.1021/acs.orglett.5b00406
27. Baeschin, D. K.; Charperon, A. R.; Charbonneau, V.; Green, L. G.; Ley, S. V.; Lücking, U.; Walter, E. *Angew. Chem., Int. Ed.* **1998,** *37,* 3423–3428. doi:10.1002/(SICI)1521-3773(19981231)37:24<3423::AID-ANIE3423>3.0.CO;2-I
28. Nokami, T.; Shibuya, A.; Tsuyama, H.; Suga, S.; Bowers, A. A.; Crich, D.; Yoshida, J.-i. *J. Am. Chem. Soc.* **2007,** *129,* 10922–10928. doi:10.1021/ja072440x
29. Nokami, T.; Shibuya, A.; Manabe, S.; Ito, Y.; Yoshida, J.-i. *Chem. – Eur. J.* **2009,** *15,* 2252–2255. doi:10.1002/chem.200802293
30. Glycosyl triflate **3ba** was observable as a major product by low temperature NMR measurement. Nokami, T.; Shibuya, A.; Manabe, S.; Ito, Y.; Yoshida, J. unpublished results.
31. When the global deprotection with a palladium catalyst with hydrogen gas was performed in methanol, we detected the molecular ion peak of the corresponding methyl glycoside of TMG-chitotriomycin.

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Glyco-gold nanoparticles: synthesis and applications

Federica Compostella[1], Olimpia Pitirollo[2], Alessandro Silvestri[2,3] and Laura Polito[*3]

Review

Open Access

Address:
[1]Department of Medical Biotechnology and Translational Medicine, University of Milan, Via Saldini 50, 20133 Milan, Italy, [2]Department of Chemistry, University of Milan, Via C. Golgi 19, 20133 Milan, Italy, and [3]CNR – ISTM, Nanotechnology Lab., Via G. Fantoli 16/15, 20138 Milan, Italy

Email:
Laura Polito[*] - laura.polito@istm.cnr.it

* Corresponding author

Keywords:
gold nanoparticles; glyco nanoparticles; glyconanomaterials; immunology; targeting

Beilstein J. Org. Chem. **2017,** *13,* 1008–1021.
doi:10.3762/bjoc.13.100

Received: 02 March 2017
Accepted: 05 May 2017
Published: 24 May 2017

This article is part of the Thematic Series "Biomolecular systems".

Guest Editor: P. H. Seeberger

Abstract

Glyco-gold nanoparticles combine in a single entity the peculiar properties of gold nanoparticles with the biological activity of carbohydrates. The result is an exciting nanosystem, able to mimic the natural multivalent presentation of saccharide moieties and to exploit the peculiar optical properties of the metallic core. In this review, we present recent advances on glyco-gold nanoparticle applications in different biological fields, highlighting the key parameters which inspire the glyco nanoparticle design.

Introduction

Carbohydrates are well recognized as crucial biomolecules for their capability to produce an enormous number of biocodes which are translated into their ability to promote or suppress a plethora of biological events [1]. Nevertheless, carbohydrates interact with their ligands in a very weak manner, so much faint that nature found in the multi-presentation (the so called "glycocalix") a solution to have efficient carbohydrate interactions. In fact, the sum of binding affinities for single and isolated interaction is ever lower than those measured when a glyco-multivalent presentation is exploited [2]. Following nature's design, carbohydrate multivalent systems are, at the present time, the most common strategy used to study weak carbohydrate–carbohydrate or carbohydrate–protein interactions and the resulting biological processes [3,4].

The increasing ability to manipulate material at the nanosize allowed the development of many glyco nanoparticles helpful in a wide range of applications, from the drug delivery to the imaging [5-7]. Among the large number of nanomaterials, gold is one of the main exploited scaffolds for producing glyco nanoparticles [3,4]. First of all, gold is an extremely inert and biocompatible material. At nanosize, gold nanoparticles (AuNPs) are characterized by unique optical features which result extremely useful in many diagnosis and affinity studies or

protocols. Moreover, gold can be easily and covalently decorated on its surface by exploiting the strong soft–soft interaction between a Au atom and sulfur [8]. Therefore, glyco-gold nanoparticles (GAuNPs) represent a smart, multimodal and versatile nanoplatform to develop carbohydrate-based nanotechnology and nanomedicine systems.

The present review focuses its attention only toward the synthesis of gold-based nanoparticles coated with carbohydrates. A first introduction related to the main protocols to produce glyco-gold nanoparticles is followed by a section where the main parameters which should be taken into account for designing the best performing nanosystem are discussed. Then, three sections summarize the achievements and ongoing researches published in the last four years, from 2013 to 2016 with a special focus on glyco-gold nanoparticle applications (in biosensing field, as drug carriers or contrast agents and as effective tool in the immunological field).

Review

Synthesis of glyco-gold nanoparticles

Despite the great inertness of this noble metal, Au can form stable bonds with sulfur-containing compounds (i.e., thiols or disulfides) [9,10] and it is possible to easily and robustly functionalize AuNPs with organic molecules. The synthetic approaches to GAuNPs can be classified into three major categories (Figure 1).

The first one consists in a one-step preparation of ultrasmall spherical gold nanoparticles coated with a thio-glucoside [4,11-16]. The process, developed in 2001 by Penadès and co-workers [15] is a modified Brust–Shiffrin method [17], based on the in situ Au^{3+} reduction by means of $NaBH_4$ in the presence of thiol-ending neoglycoconjugates. Following this approach, a rapid synthesis of AuNPs characterized by elevated glycan density on the NP surface and a 1–10 nm diameter range can be obtained [4]. Moreover, this synthetic procedure allows to introduce different types of carbohydrates and other ligands (i.e., polyethylene chains, lipids, peptides, DNA, RNA or fluorescent dyes) in controlled ratios [4]. A modification of this technique consists in the application of the "thiol for thiol" ligand exchange method, which allows the superficial diversification by introducing small amounts of different ligands after the AuNP synthesis [11,12,18]. Recently, the Seeberger group reported a straightforward one-pot method to prepare glucose-stabilized ultrasmall AuNPs, by simply mixing at room temperature Au^{3+} salts and thio-glucoside as reducing and stabilizing agents, without the addition of $NaBH_4$ [19].

The two-step synthetic procedure consists in the initial synthesis of AuNPs stabilized by temporary compounds (i.e., citrates, amines or phosphines), followed by the ligand exchange protocol to introduce the thiolated saccharide molecules [20-22]. In this frame, the Turkevich method [23] is the most common protocol to produce citrate-capped spherical nanoparticles, with diameter size between 10 and 50 nm. Noteworthy, by employing surfactant or templating agents (i.e, $AgNO_3$ or cetyltrimethylammonium bromide, CTAB), a control over the core morphology can be implemented with this methodology, as reported for the synthesis of mannose and galactose functionalized nanorods and nanostars [24] or glucosamine-coated nanostar AuNPs [22]. Respect to other methodologies, the main drawbacks of this protocol are the longer reaction times (12–24 hours), the lower glycan loading and the difficulty for a fine control of the organic shell composition.

The last category collects all the protocols based on chemical conjugations of sugar residues to ligands displayed on the metal surface of pre-formed AuNPs. Despite the longer times required for the NP synthesis and the first coating, the main

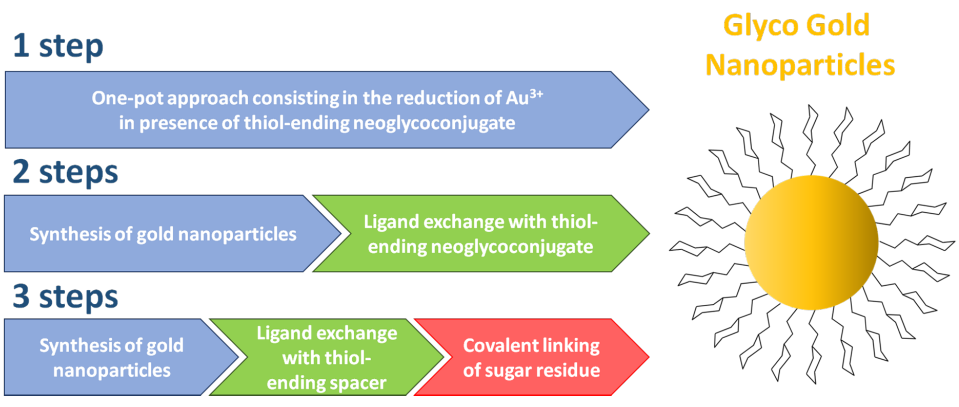


Figure 1: Schematic overview on the glyco-gold nanoparticles synthetic approaches.

advantages of this approach are the use of a lower amount of precious sugar residues and the possibility to exploit different orthogonal surface functionalizations. Reactions employed for this purpose (i.e., click reaction, amidation, conjugation via carbonyldiimidazol and perfluorophenyl azide (PFPA) photocoupling) have to be compatible with water, the common medium for AuNP preparation [25-29]. One example of the three-step approach was described by Tian and co-workers [28]. AuNPs were firstly coated by a monolayer of dithiol-cyclooctyne spacer. Then, the alkyne moiety reacted with mannosyl azide through a copper-free strain-promoted alkyne-azide cycloaddition (SPAAC), affording the GAuNPs. Alternatively, Yan and co-workers developed a three-step photocoupling procedure for the immobilization of carbohydrates on the nanoparticles surface [27]. The AuNPs were firstly coated with a PFPA-thiol monolayer via a ligand-exchange reaction, then carbohydrates were immobilized by means of a photocoupling reaction.

Glyco-gold nanoparticles: design and structural properties

GAuNPs are constituted by two main entities, the metallic core and the organic surface (i.e., mono- oligo- and polysaccharides). The features of these entities are tightly related to their application and effectiveness (Figure 2). Therefore, a careful design of GAuNP is required to obtain the best performing system.

Gold core: optical properties

AuNPs, widely employed for biomedical applications, are characterized by inert nature, resisting to air oxidation and corrosion [30]. This chemical non-reactivity and inertness make AuNPs an outstanding candidate for the development of in vitro and in vivo devices [31,32]. Furthermore, in vitro and in vivo short term reduced toxicity of AuNPs has been widely documented [10,33-38]. AuNPs own a number of peculiar optical properties, strongly dependent on the size and the morphology of the metallic core. When AuNP dimension is comparable to the gold Fermi wavelength, the nanoparticles acquire photoluminescence properties [31]. These small clusters typically present a diameter between 0.5 and 2 nm. When clusters are bigger, they do not present fluorescent emission but they show a localized surface plasmon resonance (LSPR) around 520 nm. This phenomenon consists in a collective oscillation of the free electrons of the metal, induced by the light impinging on the AuNPs [31,39]. The AuNP surface plasmon resonance is dependent upon a vast number of parameters like size, shape, morphology and environment surrounding AuNPs [40]. By controlling these parameters, it is possible to tune the LSPR peak through the visible and near IR spectra. As effect of the LSPR process, the incident light is strongly attenuated and intense electric fields are generated in the proximity of the AuNP surface. These electric fields are of great interest because they can be employed to enhance the intensity of Raman signals close to the NP surface, widely explored in biomedical diagnostics (surface enhanced Raman spectroscopy, SERS). This effect is even more intense when the nanoparticles possess an irregular shape, generating an anisotropic distribution of the magnetic field [41]. AuNP optical properties have been widely exploited in the design of GAuNPs with interesting applications such as the colorimetric biosensors. Based on the LSPR phenomenon, the controlled aggregation and de-assembling of AuNPs is reflected in a detectable variation of the colloidal solution color. Therefore, several biosensors for the detection of proteins [42], lectin [25,27,43], cancer biomarkers [28] and viruses [44,45] were developed. Photoluminescence properties of ultrasmall gold nanoclusters and nanodots have been exploited to detect *Escherichia coli* (a bacteria containing mannose-binding receptors) [46] and thyroglobulin in serum [47] and to study the

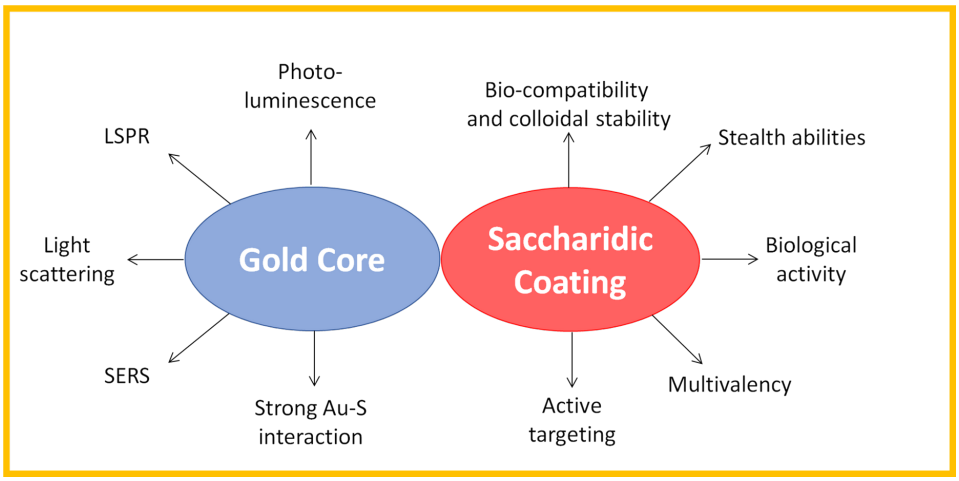


Figure 2: Glyco-gold nanoparticles: metallic core and glyco-coating contribute to the production of the versatile glyco-nanomaterials.

internalization pathway in dendritic cells, employing confocal laser microscopy (CLSM) and flow cytometry [48]. Moreover, even techniques based on light scattering, like dark field microscopy (DFM), resulted convenient for the biosensing application, considering that the light scattered by a single AuNP can be about 1 million times more intense than the fluorescence emitted by a dye [49,50].

Gold core: the size and the shape role

AuNP core dimension and shape are important parameters to take into consideration for the multivalent carbohydrate presentation [51-53]. In fact, a denser ligand packing can be obtained when NP dimensions are bigger and the curvature radius is reduced. This observation was already highlighted by Lin et al., during their studies on the interactions among Shiga toxins and multivalent GAuNPs [52]. The authors, comparing the affinities of 4, 13 and 20 nm AuNPs, demonstrated that, when the diameter grows and the surface becomes flatter, the interaction with the proteins binding sites is facilitated. Besides the AuNP size, the NP shape can affect the effectiveness of active molecules displayed on the metallic surface. In particular, an in vitro study performed by Mitragotri and co-workers reported that the endothelial targeting specificity could be enhanced by modulating the NP shape. By using a microfluidic synthetic vascular network the authors demonstrated that rod-shaped NPs exhibit elevated specific accumulation in comparison to spherical ones [54]. Recently, Kikkeri et al. published two interesting papers evaluating the impact of the AuNP shape on the carbohydrate presentation and activity [24,55]. In the first work the authors reported the employment of GAuNPs for the bacterial recognition and bacterial infection inhibition. Three different morphologies of AuNPs (sphere, rod and star-like NPs) coated with galactose and mannose derivatives were employed for the quantification of their binding affinity to *E. coli*, suggesting that each shape could induce a different bacterial adhesion. Indeed, gold nanorods presented a bacteria detection limit of 0.03 ± 0.01 μg/mL, 80-folds more sensitive than spherical and star shaped AuNPs. This discrepancy has been ascribed to the difference in the relative amount of mannose involved in the NP-bacteria interaction. The same group in 2016 investigated as well the effect of the AuNP morphology on carbohydrate–protein interactions in mammalian cells. In particular, the authors studied the shape-dependent uptake of mannose- and galactose-functionalized AuNPs in different cancer cells. By employing an enzyme-linked immunosorbent assay (ELISA), the selectivity and sensitivity of the binding were demonstrated to be dependent on the sugar nature and on the NP morphology. The cellular uptake experiments on HeLa, HepG2 and MDA-MB-231 cells showed in fact an higher clathrin-mediated internalization for rod-shaped AuNPs in comparison to spherical and star-shaped AuNPs.

Saccharidic functionalization: the coating density

Besides the inorganic core, the surrounding organic shell plays a crucial role in GAuNP chemical-physical properties. The organic functionalization, in fact, impart the stealth ability and biocompatibility of the nanomaterials. Moreover, when the metallic surface is decorated with an active target, the GAuNPs can act as a smart probe, even exploiting the multivalent effect.

Several studies were focused on the investigation of synthetic techniques to improve saccharide density on the nanomaterial surface [45,56-58]. Prosperi and co-workers coated dodecanthiol AuNPs with manno-calixarenes exploiting hydrophobic interactions, obtaining an efficient targeting against cancer cells [56]. Similarly, a reversible addition−fragmentation chain transfer (RAFT) polymerization approach has been exploited for the preparation of "multicopy–multivalent" AuNPs, decorated with complex glycan structures. The high degree of ramification resulted in a strong selective recognition [57,58] with the possibility to modulate the glycosylic loading. Even if the carbohydrate loading is important, not always a high carbohydrate density is reflected in an improved affinity and it depends on the biological event observed [18,59-64]. For example, in the case of carbohydrate−protein interactions, the carbohydrate disposition on the NP surface is more important than the loading of the active molecule. Proteins, like lectins, are characterized by different subunits possessing characteristic distances between the active pockets, so a controlled carbohydrate density can result in a better presentation of the molecules [60].

Saccharidic functionalization: the importance of the linker

The linker employed to space glycans from the Au surface is another fundamental factor to take into consideration in GAuNP design. Generally, the majority of the spacers are thiol-linkers, exploiting the strong soft–soft interaction between gold and sulfur. The spacer has the crucial role to stabilize GAuNPs in the surrounding media, which is fundamental in the case of colorimetric bioassays where the NP colloidal stability can strongly influence the state of aggregation and, consequently, the analysis results [43,65,66]. Woods et al., in a recent paper, performed a computational study to evaluate the role of the spacer in the glycan–protein interaction [67]. The authors considered different types of spacer in terms of chemical structure, length and rigidity and demonstrated that longer linkers resulted as the best performing ones, giving a better access to the protein pocket. Furthermore, they showed that rigid spacers could hinder the binding, by creating unfavorable interactions with the protein surface, while flexible spacers resulted in a better carbohydrate presentation and subsequent interaction with the target. Schlecht and co-workers experimentally demonstrated that the binding selectivity of GAuNPs toward P-selectin

was affected not only by the spacer length but also by the presence of amide bonds next to the protein pocket [51].

Saccharidic functionalization: colloidal and stealth ability

GAuNP colloidal stability and stealth abilities in biological media and in intracellular environment is a fundamental issue for the efficacy of these nanomaterials [68]. The use of stealth coatings is an effective strategy to decrease or inhibit the non-specific interaction with plasma proteins (the so called "protein corona"), that can modify GAuNP biological identity, affecting aggregation, and cellular uptake [69-71]. Moya and co-workers deeply studied the intracellular dynamics and aggregation of glucose-AuNPs by employing fluorescence correlation spectroscopy (FCS) [12]. They demonstrated that GAuNPs were ubiquitous distributed inside the cell as single NP or small aggregates, suggesting a strong intracellular stability. Liz-Marzan et al. assessed the ability of glycan ligands to reduce the protein corona formation around rod-shaped AuNPs [71]. Lactose- and polyethylene glycol (PEG)-functionalized gold nanorods demonstrated similar ability in reducing the interaction with proteins, when compared to citrate-AuNPs. Moreover, comparing the macrophage NP uptake, the authors demonstrated that the lactose shell can prevent phagocytosis more efficiently than PEG coating.

A deep knowledge of any aspect which can influence the GAuNP behavior is important to drive the correct design of the nanosystem and is strongly related to the final bioapplication.

Glyco-gold nanoparticles as nanosensors

GAuNPs are able to create supramolecular networks by interacting with proteins, enzymes or other carbohydrates and this interaction is reflected in a measurable shifting of the LSPR band, due to the aggregation of the metallic NPs. Carbohydrate–carbohydrate interactions (CCIs), generally characterized by K_D values in the millimolar range, are very difficult to detect. Updated data confirm the crucial role of these bindings in the regulation of many cellular processes [72], however, researches related to the use of carbohydrate-coated gold nanoparticles to investigate CCIs, are still in their infancy and will be not detailed in the present review [73,74]. On the contrary, a vast number of papers reported the use of glyco-gold nanoparticles as smart nanomaterial, able to recognize and interact with proteins, viruses and peptide hormones.

Detection of carbohydrate–protein interactions

Carbohydrate–protein interactions have a crucial role in many pathological and physiological cellular functions, acting as a main process correlated with a large number of biological events, such as inflammations, cancer metastasis, fertility, etc. [75-77]. Lectins are the most diffused class of proteins which recognize carbohydrate ligands and they are widely studied to develop systems able to detect lectin–carbohydrate interactions. Yan et al. exploited the use of gold nanoparticles to recognize and differentiate lectins by capping gold surface with eleven different carbohydrate ligands [78]. To achieve their goal, the authors treated GAuNPs with lectins and used the linear discriminant analysis (LDA) to analyze the LSPR shifts and recognize the lectins in a complex sample. A similar approach, based on LDA, has been employed by Gibson et al. to provide discrimination among legume lectins [79]. In their paper, the authors synthesized eleven GAuNPs, by tuning a heterogeneous coatings using only two 2-amino-2-deoxysugars (galactosamine and mannosamine). By means of this approach, and limiting the carbohydrate synthetic efforts, the GAuNPs showed an incremented lectin identification power, at low protein concentrations, enabling an easy tool for sugar–lectin recognition. The glycan–protein interaction has been studied as well employing surface-enhanced Raman spectroscopy (SERS) exploiting AuNP ability in amplifying Raman signals [80]. The greatest advantage in using SERS is the possibility to provide unique spectral signatures describing glycan–protein interactions. Boons and co-workers prepared label-free microarrays of lactose-coated AuNPs (diameter of 60 nm) to study lactose galectin (i.e., galectine 1 and galectin 3) and influenza hemagglutinin interactions. The resulting NPs were deposited and dried on a gold layer and SERS measurements performed. Finally, partial least squares discrimination analysis (PLS-DA) was performed to analyze the SERS spectra, differentiating them in different classes with statistical relevance. Long et al. [50] developed a simple procedure to analyze in real-time the carbohydrate–protein interaction at the single NP level, using for the first time an unusual technique such as the dark field microscopy, DFM (Figure 3). Briefly, the method is based on the interaction among 60 nm AuNPs capped with the protein concanavalin A (ConA, physically adsorbed on the surface) and smaller dextran-coated AuNPs. ConA-AuNPs, immobilized on a glass slide, were treated with dextran-AuNPs and when the two particles interacted, a coupling of the plasmonic oscillations was detected, resulting in a shift of the statistic peak wavelength distribution. The statistical analysis of the peak wavelength was performed starting from the RGB (red, green and blue) information obtained by DFM by analyzing them with a self-developed statistical program (nanoparticleAnalysis).

A very interesting result reports on the design of highly stable carbohydrate-coated gold nanoparticles, able to recognize at subnanomolar and low picomolar concentrations level of ConA, wheat germ agglutinin and *Ricinus communis* agglutinin [43]. In this case, a careful decoration of the gold surface has been carried out by employing a thio-amphiphilic linker to impart

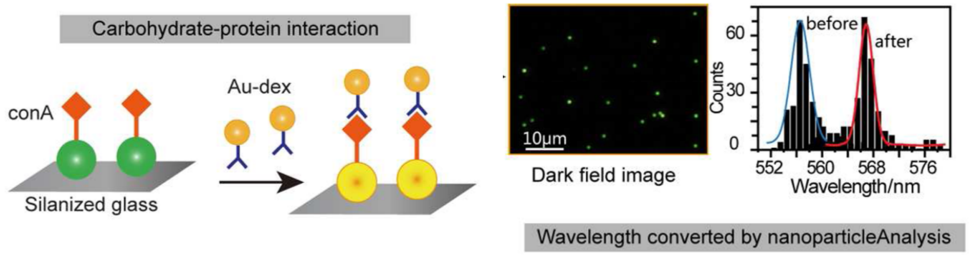


Figure 3: The protein–carbohydrate interactions analyzed by conventional dark field microscopy. The interaction among ConA-functionalized AuNPs (60 nm) and dextran-coated AuNPs (20 nm) are detected by DFM. Reprinted with permission from [50]. Copyright 2015 American Chemical Society.

improved solubility and flexibility to the glycosyl-ligand. UV–vis spectroscopy and dynamic light scattering measurements have been exploited to detect at low picomolar concentrations lactose-AuNP, mannose-AuNP and GlcNAc-AuNP interactions with their cognate lectins. Moreover, by using an immobilized antibody microarray, it was possible to design a highly sensitive lectin detection, enhanced by silver addition, by easy naked eye. Plasmon resonance spectroscopy and naked eye have been used by Long et al. (Figure 4) to study the interaction of galactoside-AuNPs with peanut agglutinin [25]. In this case, the glycosyl residue is not coated on the gold surface by means of a thiol moiety but exploiting the strain-promoted azide–alkyne cycloaddition (SPAAC) to covalently link an azido galactoside on a lipid cyclooctyne. In this manner, the amphiphilic glyco-lipid can be embedded on PEGylated-gold nanoparticles and the sugar moiety displayed on the outer shell of the nanosystem is available for interaction with lectin.

In 2016, Sakurai and co-workers [81] combined the multiple display of the carbohydrate moiety on AuNPs with the photoaffinity labeling (PAL), to concentrate and purify proteins by covalent crosslinking (Figure 5).

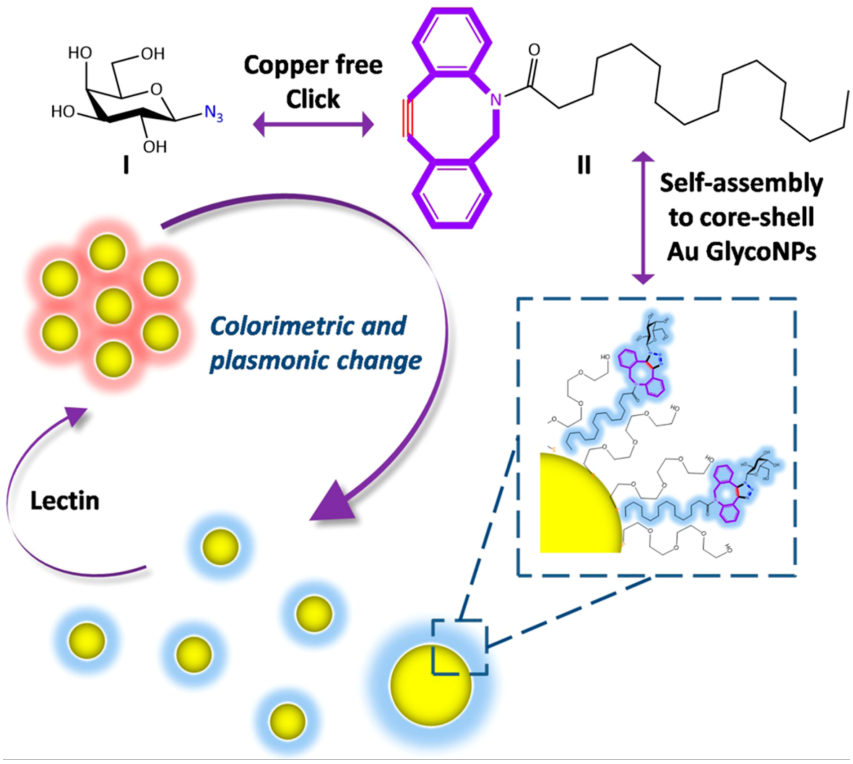


Figure 4: Copper-free cycloaddition (SPAAC) of azido galactoside on cyclooctyne and schematic depiction of Au surface coating. When galactoside-coated AuNPs interact with lectin, AuNPs aggregates, inducing a plasmonic band shift. Reprinted with permission from [25]. Copyright 2015 American Chemical Society.

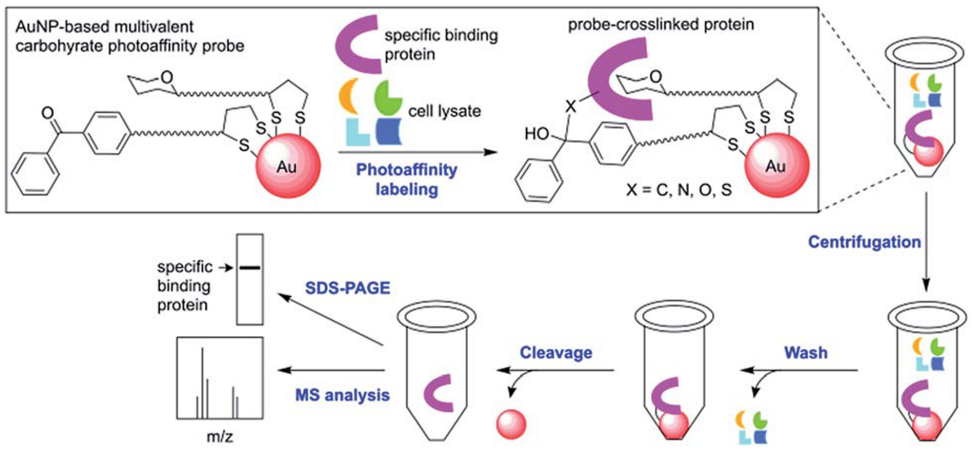


Figure 5: A photoaffinity labeling (PAL) approach based on glyco-gold nanoparticles is able to recognize and isolate only the desired protein. After washing and cleavage steps, the protein can be fully analyzed. Reprinted from [81]. Published by The Royal Society of Chemistry.

GAuNPs were synthesized by coating the metal surface via thio-chemistry with a β-D-lactose residue (recognized by a series of lectins, i.e., PNA or *Ricinus communis* agglutinin) and a benzophenone moiety as photoreactive group. GAuNPs were treated with PNA which cross-linked with the photoreactive group, affording a PNA-AuNPs complex. The latter was easily purified by the non-interacting protein via centrifugation and, eventually, the protein was quantified after an appropriate cleavage with 2-mercaptoethanol.

A very recent and interesting application showed for the first time the immobilization of AuNPs on the endface of an optical fiber by means of thio-chemistry [82]. A second layer of glycosyl residues (maltose or lactose) containing thioctic acid was covalently linked on supported AuNPs, affording the fiber-type "Sugar Chips" (Figure 6). By means of this sugar chips the carbohydrate−protein interactions have been analyzed and quantified in a very small volume with results confirmed by the conventional LSPR biosensors.

Gibson et al. [66] designed 40 nm AuNPs capped with glyco-PEG ligands to discriminate different strains of *E. coli*, on the base of different expression level of the bacterial lectin FimH with a colorimetric assay. A simple and fast bioassay has been developed by Lee et al. [83] to recognize cholera toxin (CT), a protein secreted by the *Vibrio cholerae* bacterium which is responsible for cholera disease. A thiol-modified β-galactose derivative was used to coat 10 nm AuNPs and mixed with amino-terminated quantum dots (CdTe, QD). As a consequence of the strong hydrogen bonds among QD amines and hydroxy groups of galactose, a complex of QD-AuNPs occurred as confirmed by using fluorescence resonance energy transfer (FRET). In the presence of the analyte (i.e., cholera toxin), galactose-AuNPs recognize the protein, avoiding the formation of QD complex and, consequently recovering the fluorescence. This method, al-

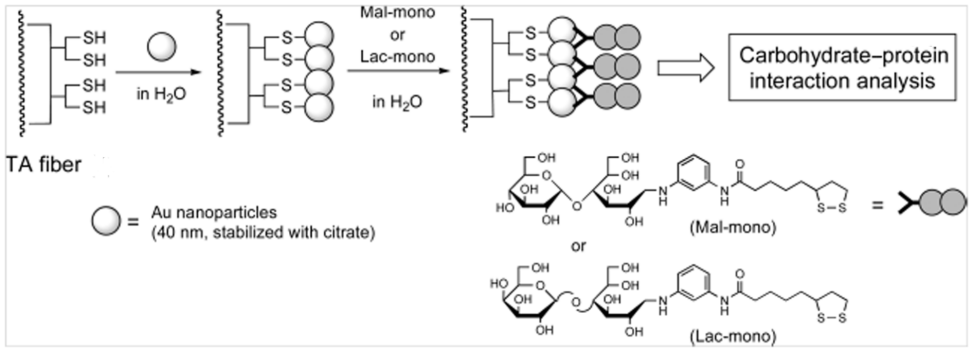


Figure 6: Design of the fiber-type "Sugar Chips". AuNPs are firstly immobilized on an optical fiber functionalized with α-lipoic acid (TA fiber). Successively, glycosil-thioctic residues were grafted on AuNPs, forming the "Sugar Chips". Reprinted with permission from [82]. Copyright 2017 American Chemical Society.

though based on a monosaccharide, showed the impressive improvement obtained for using multivalent glyco-nanosystems.

Quick serological detection of cancer biomarkers is one of the main aim in early cancer diagnosis, and GAuNPs seem to be fundamental in the development of new, cheap and easy tests. Tian et al. [28] described the design of mannosylated gold nanoparticles which create a supramolecular glycoprobe in the presence of a specific mannose-selective *Lens culinaria* lectin (LcA). This lectin is particularly interesting because it can have a strong affinity with the core-fucosylated glycoform of α-fetoprotein (AFP-L3), a recognized biomarker for the diagnosis of hepatocellular carcinoma. When the glycoprobe interact with a microplate displaying AFP-L3, LcA strongly bind to AFP-L3 releasing from the supramolecular glycoprobe GAuNPs in a concentration which can be measured via UV–vis spectroscopy. Chiodo et al. [84] gave their contribution in the development of early stage cancer or pathogen diagnosis designing an ELISA approach to detect anti-carbohydrate antibodies. The innovative procedure exploited the use of very small AuNPs (2 nm) coated with tetrasaccharide epitopes of HIV gp120 or tetrasaccharide epitopes of *Streptococcus pneumoniae* Pn14PS. The GAuNPs so obtained were directly coated on commercial ELISA plates and anti-carbohydrate antibodies were detected in the nanomolar range both using purified anti-HIV human monoclonal antibodies or serum from immunized mice against *S. pneumoniae*.

Detection of carbohydrate–influenza virus interactions

The SERS methodology was extended to the determination of the binding selectivity for human and avian influenza (H1N1 and H7N9), by using modified glycans (lactose, α(2,3)Neu5AcLacNAc and α(2,6)Neu5AcLacNAc) to introduce the active moieties on the gold monolayer. The SERS spectra registered with changes in the 1400–1600 and 1000–1200 cm^{-1} bands were indicative of the successful binding [85]. In this field, a naturally occurring sialylglycopeptide extracted from egg yolks was converted into a thiol-terminated molecule and grafted on AuNPs, to develop a gold-based sensor for influenza virus detection, using both UV–vis spectroscopy and dynamic light scattering. Moreover, a selective detection of the influenza virus can be accomplished. Viral hemagglutinin (HA) protein, in fact, can bind selectively sialic acid residues, but the human influenza virus recognizes the sialic acid α(2,6)galactose sequence while the avian virus recognizes sialic acid α(2,3)galactose chains [85]. GAuNPs functionalized with the exact sequence have been able to selectively recognize the influenza virus strain by means of a simple colorimetric assay, observing the hemagglutinin-induced aggregation of AuNPs [44,45]. Moreover, to improve the sialic residue display on metal surfaces, Russel and co-workers [45] employed trivalent α-thio-linked sialic residues as ligands of AuNPs, confirming the improving ability in binding selectively the human influenza strain.

Detection of carbohydrate–peptide hormone interactions

Dextran-coated AuNPs have been employed to investigate the binding properties of the glucose moiety displayed on a metal surface with insulin, a peptide hormone [86]. By means of a UV–vis spectroscopy assay, the authors confirmed the successful use of dextran-coated AuNPs to selectively bind low concentrations of insulin (at 1 pM), even in serum samples, and discovered that the insulin B chain is mainly responsible for the binding with glucose moiety AuNPs.

Glyco-gold nanoparticles as drug carriers and contrast agents

Drug nanocarriers play a fundamental role in the development of nano- and personalized medicine. The main advantages resulted from the protection of the drug from degradation or inactivation in vivo, the opportunity to control the drug release and reduce the systemic toxicity and the exciting chance to selectively target the damaged tissue. Despite the most common carriers are based on nano- and microparticles (i.e., albumin, poly(lactic-*co*-glycolic acid)-PLGA NPs, liposomes), many examples in literature report on the development of gold-based glyco nanoparticles as efficient system to delivery payloads. In this contest, AuNPs modified with lactose moieties and β-cyclodextrins have been reported by Vargas-Berenguel et al. [87]. The authors showed that GAuNPs were still able to recognize specific lectin (peanut agglutinin, PNA) and human galectin-3 (Gal-3), exploiting multiple lactose residues displayed on gold surface. Moreover, due to the presence of the β-cyclodextrin cavity, these nanostructures worked as site-specific delivery systems, hosting methotrexate (MTX), an anticancer drug.

The anticancer therapy based on platinum compounds, such as cisplatin, carboplatin and oxaliplatin, has many positive effects on the regression of cancer cell proliferation. Unfortunately these compounds are low tolerated by the organisms, therefore new metal-based anticancer drugs need to be developed. Recently, it has been reported the interesting synthesis of AuNPs coated with glyco-polymers and functionalized with gold(I) triphenylphosphine (Figure 7) [88]. This work showed the potentiality of these structures as novel cancer therapeutic drugs. The glucose and galactose moieties multi-displayed on gold surface acted as target for asialoglycoprotein receptors, over-expressed on liver cancer cells, delivering selectively the anticancer drug Au(I).

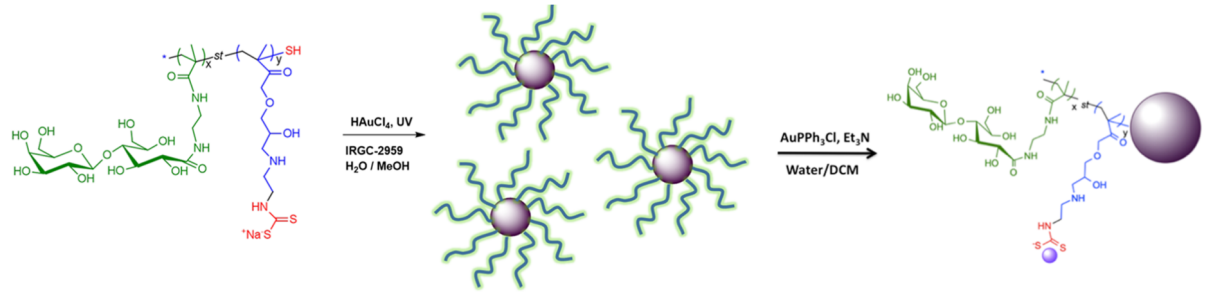


Figure 7: Design of GAuNPs, displaying both sugar residues and the anticancer Au(I)PPh_3. Reprinted with permission from [88]. Copyright 2014 American Chemical Society.

In search of novel antimicrobial antigens, chitosan-streptomycin GAuNPs (CA NPs), i.e., carbohydrate-antibiotic conjugates, were prepared and resulted able to overcome antibiotic resistance of microbial biofilms, since CA NPs render streptomycin more accessible to biofilms, thereby more available to interact with biofilm bacteria [89]. Similarly, a novel recyclable *E.coli*-specific killing GAuNP nanocomposite, carrying either a targeting agent towards *E. coli* pili and an antibacterial polymer, resulted able to bind *E. coli* specifically and to kill the attached bacteria, by enhancing the local concentration of biocidal agent on the bacterial surface [90].

GAuNPs find interesting application also in the treatment of HIV infection. β-D-Glucose coated AuNPs have been decorated with the antiviral drugs abacavir (ABC) and lamivudine (3TC) and used as delivery systems. These co-drug structures were able to release the drugs in acidic conditions to inhibit viral replication in cellular assays [91].

GAuNPs represent a promising candidate for developing theranostic agents, able to carry drugs to a selected tissue and to act as a sensitive probe. In fact, the metallic core of GAuNPs has been exploited to design and develop promising early detecting agents of cancer or inflammation. In this contest, Andresen and co-workers produced a nanogel injectable in humans to detect cancer cells. The nanogel was composed by sucrose acetate isobutyltyrate (SAIB) and poly(*N*-isopropylacrylamide) (PNIPAM)-coated AuNPs, compatible with SAIB matrix. The nanogel had good opacity for 2D X-ray visualization and was tested in vivo in immunocompetent healthy NMRI-mice and in humans [92]. The gold high X-ray absorption and a carbohydrate coating have been also exploited by Polito and co-workers to produce functional X-ray contrast agents based on glucosamine-coated AuNPs [22]. By micro-CT analysis performed on lung injured mice, the authors demonstrated that GAuNPs selectively targeted overexpressed glucose receptors (GLUT1) and accumulated in the damaged lung, while PEGylated AuNPs did not show any relevant accumulation in the inflammation site. Moreover, another study on the synthesis of glucose-functionalized AuNPs demonstrated that GLUT1-targeting GAuNPs are well compatible with blood and do not significantly interact with ovine red blood cells [93]. In the last years, numerous compounds containing Gd^{3+} have been developed and used as magnetic resonance imaging (MRI) contrast agents. Recently, gold based nanosystems containing Gd^{3+} were developed. Arus et al. described the synthesis of water dispersible GAuNPs decorated with Gd complexes [94]. The glyco-shell conferred to the nanotracer a targeting properties while Gd^{3+} acted as MRI contrast agent. The authors settled an ex vivo experiment which let them to use small amounts of contrast agent (5 nmol/animal), avoiding toxicology and in vivo test. Moreover, GAuNPs showed their extreme versatility in a very interesting paper reported by Penadés and collaborators [11]. They design GAuNPs able to pass the brain blood barrier (BBB) of healthy mice and give images of intact brain. AuNPs were decorated in a controlled ratio with glucose, neuropeptides and a chelator of gallium-68, a sensitive positron emission tomography (PET) tracer, affording the quantification of BBB permeability in healthy small animals.

Glyco-gold nanoparticles in immunology

The development of nanoparticle-based strategies for the modulation of the immunosystem represents one of the biological and biomedical applications of nanomaterials. Nanoparticle technologies offer the possibility of new approaches to tune specific immune responses for prevention or treatment of diseases. These approaches are inspired by the availability of novel nanomaterials with specific capabilities [95]. AuNPs have attracted great attention in this field due to their unique characteristics of biocompatibility and easy fabrication, already outlined previously. In the last five years, significant examples on the use of carbohydrate-coated GAuNPs in the field of immunology have been reported, addressing different possibilities of intervention on the immune system.

GAuNPs and activation of the immune system via cellular targeting

Glycan-conjugated globular shaped nanoparticles have a higher surface area to volume ratio in comparison to other glycoclusters, i.e., glycolipid micelles and glycoproteins, which allows an ordered, controlled and stable tridimensional presentation of the ligands, providing higher ligand binding capacity [6,96]. For this reason, they represent an ideal scaffold to study the activation of the immune system in response to pathogens mediated by lectins, like DC-SIGN. Mammalian cell surface lectins, expressed on the surface of dendritic cells (DC) and macrophages, recognize in a multivalent way a vast array of glycans on the exterior of pathogens. Natural carbohydrate ligands include high-mannose N-glycans, such as those found on the HIV gp120 protein, and Lewis-type glycans. Ligand recognition often initiates events which are able to trigger specific signaling and result in a robust immune activation in humans. For example, galactofuranose (Gal*f*) functionalized AuNPs were found able to elicit a pro-inflammatory response in dendritic cells via interaction with the lectin DC-SIGN, as indicated by the up-regulation of several maturation markers and increased secretion of pro-inflammatory cytokines interleukin 6 (IL-6) and tumor necrosis factor α (TNF-α) [97]. These data indicate that a multivalent system of Gal*f*, a five-membered form of galactose found in nonmammalian pathogenic species, i.e., the fungus *Aspergillus*, is able to modulate the innate immune response via dendritic cells, suggesting Gal*f*-GNPs as versatile tools in functional studies to address the role of Gal*f* in host-pathogen interaction.

DC-SIGN targeting through GAuNPs has also been exploited differently in a study on GAuNPs functionalized with α-fucosylamide, an efficacious synthetic DC-SIGN ligand, analogue of the natural fucose-containing Lewis^x trisaccharide [98]. This paper shows that GAuNPs bearing 50% of fucosylamide are able to compete with the natural gp120 ligand on DC-SIGN expressing cells, and to promote DC-SIGN internalization. On the contrary, the GAuNPs lack any immunomodulatory activity, and do not induce DC maturation or trigger cytokine responses. Thus, in this example, GAuNPs, can be considered as vehicles, neutral carriers, to target antigens to DCs via lectins, a well-known strategy explored in immunotherapy.

GAuNPs in vaccine development

AuNPs engineering is offering significant contribution to immunology also in vaccine development. The repetitive antigen display is the key point related to this nanotechnology-based approach, which aim to trigger the production of specific and functional antibodies that prevent initial infection limiting pathogen/viral dissemination. Recent publications suggest that many different aspects are becoming clear and have to be underlined.

The design of anti HIV-1 vaccines depends on the identification of sugar epitopes of the surface envelope glycoprotein of HIV-1, capable of eliciting a protective response. GAuNPs coated with high-mannose type oligosaccharide of HIV-1 gp120, were prepared as glycoconjugate systems able to mimic the natural presentation of gp-120 high-mannose glycans. It was demonstrated that they were able to inhibit DC-SIGN-mediated HIV-1 infections and to interfere with the antibody/gp120 binding process. Most of these systems proved able to elicit the production of carbohydrate-specific antibodies in animals, though the generated IgGs turned out to be unable to neutralize the virus. More recently, the $Man_9GlcNAc_2$ was identified as a conserved motif on the HIV surface: this result stemmed from the observation that partial structures of this glycan could be recognized by broadly neutralizing antibodies, found in infected patients. These data suggested the use of these self-glycans for the design of effective vaccine and inspired the preparation of new generation GAuNPs, coated with synthetic partial structures of Man_9 multimerized on the same GAuNP, which provided better binding to the anti-HIV antibody 2G12 compared to GAuNPs carrying only one individual oligomannoside [99]. This result, based on the assembly of "antennas" on the NP surface, shows that the cluster presentation of the gp120 high-mannose type glycan on GAuNPs provides different binding modes and hence more opportunities to mimic the multivalent interaction between the antibody and the saccharide. Finally, these results represent a solid support for the design of new and more effective nanoparticles based systems.

A different approach towards the development of a synthetic HIV vaccine candidate based on GAuNPs, exploit the glyconanosystem as a tool which modulate and control the biologically active conformation of a significant epitope [100]. In this respect, negatively charged GAuNPs were conjugated to the third variable region (V3 peptide) of the HIV-1 gp120. The V3 peptide is a major immunogenic domain of HIV-1, which can be exploited as a candidate for anti-HIV vaccine development, given that the immunologically active conformation is conserved. The peptide on the nanoparticle showed increased stability towards degradation as compared to the free peptide. Moreover, V3β-GAuNPs elicit antibodies in rabbits that recognize a recombinant gp120, even in the absence of highly immunogenic proteins.

Another example on the importance of the way in which glycoconjugates are presented through NPs, regards the design and construction of GAuNPs decorated with mucin-related glycans for cancer immunotherapy [57]. A monomolecular sugar

coating do not represent well the structure of mucin glycoprotein on the surface of cancer cells, which are characterized by a dense presentation of glycans attached to a protein backbone. So, multicopy–multivalent polymeric versions of the tumor-associated α-GalNAc (Tn) antigen, obtained through RAFT polymerization, were prepared and used to prepare GAuNPs with a surface that mimics much more closely the surface of cancer cells. These systems were able to generate a significant immune response in vivo, with the generation of detectable levels of antibodies specific for naturally occurring mucin glycans, without the need for a immunogenic protein component.

In a different example, Barchi and co-workers focused on the Thomsen Friedenreich tumor-associated antigen (TFag) and synthesized TFag-amino acid-coated AuNPs by means of a thiol-polyethylene glycol spacer. Among the synthesized GAuNPs, TFag-Thr-AuNPs showed, in vitro, an increased cytotoxic activity toward lymphoma cells, resulting active at a very low micromolar range [101]. With the aim to activate the human immune system against self-tumor cells, Schlecht and co-workers designed the synthesis of AuNPs functionalized with Mucin1(MUC1)-glycopeptide antigens [102]. Mucins are a family of glycosylated proteins with a high molecular weight, produced by epithelial tissues. The most studied is the membrane-bound glycoprotein MUC1, a glycoprotein with extensive O-linked glycosylation in its extracellular domain. The authors demonstrated that the multivalent presentation of MUC1-glycopeptide antigens were able to stimulate the immune system, suggesting GAuNPs are interesting platforms for developing new immunotherapeutics.

The potential of AuNPs as antigen carriers for the development of synthetic vaccines is still being investigated in the context of bacterial infections. NPs are indeed promising vector systems to explore vaccination against infectious diseases.

Recently, a study on the identification of an optimal vaccine formulation to protect against listeriosis, proposed the use of GAuNPs loaded with glucose and the LLO_{91-98} *Listeria* peptide antigen as a system to target DCs, since *Listeria* protection is largely dependent on T-cells activation [103]. The presence of glucose ensures a biocompatible and water dispersible highly efficacious DC targeting system. DC loaded in vitro with GAuNPs-LLO, co-formulated with an adjuvant, provided better protection against listeriosis than DC loaded in vitro using free LLO peptide, and induced LLO-specific T-cell immunity and protection against *Listeria*. The system was effective in mice immunization either using a DC vaccine or a standard immunization approach. The data were followed by a study on pregnant mice vaccination which demonstrated the ability of this AuNP based vaccine system to protect neonates born to vaccinated mothers from the bacteria [104].

In another study, the impact of multivalency on the ability of GAuNPs to induce immune cell responses in vitro was studied using highly pure NPs coated with non-immunoactive mono- and disaccharides related to the capsular polysaccharide of serogroup A of Neisseria meningitides [53]. The results showed that the systems were able to activate the antigen-presenting cells, to induce the releasing of effector functions of the same cells and to stimulate T cell proliferation, but highlighted that immunoactivity is strongly dependent on the size of the nanoparticle (larger NPs, 5 nm, performed better than smaller ones, 2 nm) and on the length of the saccharide fragment. There is growing evidence that the shape and size of the AuNPs affect immunological responses in vitro and in vivo [105].

In another recent study on the development of synthetic nanosystems as potential carbohydrate-based vaccine, GAuNPs were loaded with synthetic oligosaccharide fragments corresponding to the repeating units of *S. pneumoniae* (Pn) CPS type 19F and 14 [106]. This new approach has explored the effect of the presence of two different saccharide epitopes from diverse bacterial serotypes simultaneously displayed onto the NP surface on the immunological response. Mice immunization showed that the concomitant presence of Pn14 and Pn19F repeating unit fragments on the same NP critically enhance the titers of specific IgG antibodies towards type 14 polysaccharide compared with GAuNP exclusively displaying Pn14.

The GAuNP technology has also been applied in the regulation of important genes of the apoptotic pathway. New GAuNP systems displaying cMyc targeting siRNA resulted capable of inducing apoptosis via hyperactivation of cell death receptors and caspase pathways [107].

Conclusion

Glyco-gold nanoparticles represent one of the most versatile, exciting and promising hybrid nanosystems. Combining the vast properties of the noble metal at its nano-size with the fundamental biological role covered by carbohydrates, GAuNPs hold great potentiality in bio-nanomedicine. In this review, we have attempted to summarize the last results obtained in this field, anticipating the numerous applications with a detailed description and implication of each part of the hybrid system. A careful GAuNP design must take into consideration all the parameters which characterize the final glyco-nanoparticle. Well planned GAuNPs offer a unique chance to researchers to mimic the nature's glycocalix and to explore the world of weak carbohydrate interactions, opening the way to innovative diagnostic tools or new therapeutic or theranostic agents.

Acknowledgments

L.P. and F.C. acknowledge MIUR-Italy (contract 2015RNWJAM_002) for financial support.

References

1. Dwek, R. A. *Chem. Rev.* **1996,** *96,* 683–720. doi:10.1021/cr940283b
2. Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002,** *102,* 555–578. doi:10.1021/cr000418f
3. Delbianco, M.; Bharate, P.; Varela-Aramburu, S.; Seeberger, P. H. *Chem. Rev.* **2016,** *116,* 1693–1752. doi:10.1021/acs.chemrev.5b00516
4. Marradi, M.; Chiodo, F.; García, I.; Penadés, S. *Chem. Soc. Rev.* **2013,** *42,* 4728–4745. doi:10.1039/c2cs35420a
5. Jebali, A.; Nayeri, E. K.; Roohana, S.; Aghaei, S.; Ghaffari, M.; Daliri, K.; Fuente, G. *Adv. Colloid Interface Sci.* **2017,** *240,* 1–14. doi:10.1016/j.cis.2016.11.002
6. Adak, A. K.; Li, B.-Y.; Lin, C.-C. *Carbohydr. Res.* **2015,** *405,* 2–12. doi:10.1016/j.carres.2014.07.026
7. Ramella, D.; Polito, L.; Mazzini, S.; Ronchi, S.; Scaglioni, L.; Marelli, M.; Lay, L. *Eur. J. Org. Chem.* **2014,** 5915–5924. doi:10.1002/ejoc.201402701
8. Saha, K.; Agasti, S. S.; Kim, C.; Li, X.; Rotello, V. M. *Chem. Rev.* **2012,** *112,* 2739–2779. doi:10.1021/cr2001178
9. Wang, C.; Wang, J.; Liu, D.; Wang, Z. *Talanta* **2010,** *80,* 1626–1631. doi:10.1016/j.talanta.2009.09.052
10. Rayavarapu, R. G.; Petersen, W.; Hartsuiker, L.; Chin, P.; Janssen, H.; van Leeuwen, F. W. B.; Otto, C.; Manohar, S.; van Leeuwen, T. G. *Nanotechnology* **2010,** *21,* 145101–145110. doi:10.1088/0957-4484/21/14/145101
11. Frigell, J.; García, I.; Gómez-Vallejo, V.; Llop, J.; Penadés, S. *J. Am. Chem. Soc.* **2014,** *136,* 449–457. doi:10.1021/ja411096m
12. Murray, R. A.; Qiu, Y.; Chiodo, F.; Marradi, M.; Penadés, S.; Moya, S. E. *Small* **2014,** *10,* 2602–2610. doi:10.1002/smll.201303604
13. Adokoh, C. K.; Obuah, C.; Kinfe, H. H.; Zinyemba, O.; Darkwa, J. *New J. Chem.* **2015,** *39,* 5249–5258. doi:10.1039/C5NJ00293A
14. Pourceau, G.; Del Valle-Carrandi, L.; Di Gianvincenzo, P.; Michelena, O.; Penadés, S. *RSC Adv.* **2014,** *4,* 59284–59288. doi:10.1039/C4RA11741G
15. de la Fuente, J. M.; Barrientos, A. G.; Rojas, T. C.; Rojo, J.; Cañada, J.; Fernández, A.; Penadés, S. *Angew. Chem., Int. Ed.* **2001,** *40,* 2257–2261. doi:10.1002/1521-3773(20010618)40:12<2257::AID-ANIE2257>3.0.CO;2-S
16. Cambi, A.; de Lange, F.; van Maarseveen, N. M.; Nijhuis, M.; Joosten, B.; van Dijk, E. M. H. P.; de Bakker, B. I.; Fransen, J. A. M.; Bovee-Geurts, P. H. M.; van Leeuwen, F. N.; Van Hulst, N. F.; Figdor, C. G. *J. Cell Biol.* **2004,** *164,* 145–155. doi:10.1083/jcb.200306112
17. Brust, M.; Walker, M.; Bethell, D.; Schiffrin, D. J.; Whyman, R. *Chem. Commun.* **1994,** 801–802. doi:10.1039/C39940000801
18. Martínez-Ávila, O.; Hijazi, K.; Marradi, M.; Clavel, C.; Campion, C.; Kelly, C.; Penadés, S. *Chem. – Eur. J.* **2009,** *15,* 9874–9888. doi:10.1002/chem.200900923
19. Varela-Aramburu, S.; Wirth, R.; Lai, C.-H.; Orts-Gil, G.; Seeberger, P. H. *Beilstein J. Nanotechnol.* **2016,** *7,* 1278–1283. doi:10.3762/bjnano.7.118
20. Fyrner, T.; Ederth, T.; Aili, D.; Liedberg, B.; Konradsson, P. *Colloids Surf., B* **2013,** *105,* 187–193. doi:10.1016/j.colsurfb.2013.01.002
21. Combemale, S.; Assam-Evoung, J.-N.; Houaidji, S.; Bibi, R.; Barragan-Montero, V. *Molecules* **2014,** *19,* 1120–1149. doi:10.3390/molecules19011120
22. Silvestri, A.; Zambelli, V.; Ferretti, A. M.; Salerno, D.; Bellani, G.; Polito, L. *Contrast Media Mol. Imaging* **2016,** *11,* 405–414. doi:10.1002/cmmi.1704
23. Fry, F. H.; Hamilton, G. A.; Turkevich, J. *Inorg. Chem.* **1966,** *5,* 1943–1946. doi:10.1021/ic50045a024
24. Chaudhary, P. M.; Sangabathuni, S.; Murthy, R. V.; Paul, A.; Thulasiram, H. V.; Kikkeri, R. *Chem. Commun.* **2015,** *51,* 15669–15672. doi:10.1039/C5CC05238F
25. Hu, X.-L.; Jin, H.-Y.; He, X.-P.; James, T. D.; Chen, G.-R.; Long, Y.-T. *ACS Appl. Mater. Interfaces* **2015,** *7,* 1874–1878. doi:10.1021/am5076293
26. Cetin Telli, F.; Demir, B.; Barlas, F. B.; Guler, E.; Timur, S.; Salman, Y. *RSC Adv.* **2016,** *6,* 105806–105813. doi:10.1039/C6RA21976D
27. Wang, X.; Ramström, O.; Yan, M. *Anal. Chem.* **2010,** *82,* 9082–9089. doi:10.1021/ac102114z
28. He, X.-P.; Hu, X.-L.; Jin, H.-Y.; Gan, J.; Zhu, H.; Li, J.; Long, Y.-T.; Tian, H. *Anal. Chem.* **2015,** *87,* 9078–9083. doi:10.1021/acs.analchem.5b02384
29. Chen, N.; Yu, Z.-H.; Zhou, D.; Hu, X.-L.; Zang, Y.; He, X.-P.; Li, J.; Xie, J. *Chem. Commun.* **2016,** *52,* 2284–2287. doi:10.1039/c5cc09749e
30. Uson, R. *Synth. React. Inorg. Met.-Org. Chem.* **1978,** *8,* 503–504. doi:10.1080/00945717808057442
31. Yang, X.; Yang, M.; Pang, B.; Vara, M.; Xia, Y. *Chem. Rev.* **2015,** *115,* 10410–10488. doi:10.1021/acs.chemrev.5b00193
32. Cobley, C. M.; Chen, J.; Cho, E. C.; Wang, L. V.; Xia, Y. *Chem. Soc. Rev.* **2011,** *40,* 44–56. doi:10.1039/B821763G
33. Qu, Y.; Lü, X. *Biomed. Mater.* **2009,** *4,* 25007. doi:10.1088/1748-6041/4/2/025007
34. Glazer, E. S.; Massey, K. L.; Zhu, C.; Curley, S. A. *Surgery* **2010,** *148,* 319–324. doi:10.1016/j.surg.2010.04.025
35. Brandenberger, C.; Rothen-Rutishauser, B.; Mühlfeld, C.; Schmid, O.; Ferron, G. A.; Maier, K. L.; Gehr, P.; Lenz, A.-G. *Toxicol. Appl. Pharmacol.* **2010,** *242,* 56–65. doi:10.1016/j.taap.2009.09.014
36. Zhang, X.-D.; Wu, H.-Y.; Wu, D.; Wang, Y.-Y.; Chang, J.-H.; Zhai, Z.-B.; Meng, A.-M.; Liu, P.-X.; Zhang, L.-A.; Fan, F.-Y. *Int. J. Nanomed.* **2010,** *5,* 771–781. doi:10.2147/IJN.S8428
37. Simpson, C. A.; Huffman, B. J.; Gerdon, A. E.; Cliffel, D. E. *Chem. Res. Toxicol.* **2010,** *23,* 1608–1616. doi:10.1021/tx100209t
38. Fraga, S.; Brandão, A.; Soares, M. E.; Morais, T.; Duarte, J. A.; Pereira, L.; Soares, L.; Neves, C.; Pereira, E.; de Lourdes Bastos, M.; Carmo, H. *Nanomedicine: NBM* **2014,** *10 ,* 1757–1766. doi:10.1016/j.nano.2014.06.005
39. Bohren, C. F.; Huffman, D. R. Absorption and Scattering by an Arbitrary Particle. *Absorption and Scattering of Light by Small Particles;* Wiley-VCH Verlag GmbH, 2007; pp 57–81. doi:10.1002/9783527618156.ch3
40. Xia, Y.; Halas, N. J. S. *MRS Bull.* **2005,** *30,* 338–348. doi:10.1557/mrs2005.96
41. Li, N.; Zhao, P.; Astruc, D. *Angew. Chem., Int. Ed.* **2014,** *53,* 1756–1789. doi:10.1002/anie.201300441
42. Marin, M. J.; Schofield, C. L.; Field, R. A.; Russell, D. A. *Analyst* **2015,** *140,* 59–70. doi:10.1039/C4AN01466A

43. Huang, L.-D.; Adak, A. K.; Yu, C.-C.; Hsiao, W.-C.; Lin, H.-J.; Chen, M.-L.; Lin, C.-C. *Chem. – Eur. J.* **2015,** *21,* 3956–3967. doi:10.1002/chem.201405747
44. Wei, J.; Zheng, L.; Lv, X.; Bi, Y.; Chen, W.; Zhang, W.; Shi, Y.; Zhao, L.; Sun, X.; Wang, F.; Cheng, S.; Yan, J.; Liu, W.; Jiang, X.; Gao, G. F.; Li, X. *ACS Nano* **2014,** *8,* 4600–4607. doi:10.1021/nn5002485
45. Marín, M. J.; Rashid, A.; Rejzek, M.; Fairhurst, S. A.; Wharton, S. A.; Martin, S. R.; McCauley, J. W.; Wileman, T.; Field, R. A.; Russell, D. A. *Org. Biomol. Chem.* **2013,** *11,* 7101–7107. doi:10.1039/c3ob41703d
46. Chan, P.-H.; Ghosh, B.; Lai, H.-Z.; Peng, H.-L.; Mong, K. K. T.; Chen, Y.-C. *PLoS One* **2013,** *8,* No. e58064. doi:10.1371/journal.pone.0058064
47. Huang, C.-C.; Hung, Y.-L.; Shiang, Y.-C.; Lin, T.-Y.; Lin, Y.-S.; Chen, C.-T.; Chang, H.-T. *Chem. – Asian J.* **2010,** *5,* 334–341. doi:10.1002/asia.200900346
48. Le Guével, X.; Perez Perrino, M.; Fernández, T. D.; Palomares, F.; Torres, M.-J.; Blanca, M.; Rojo, J.; Mayorga, C. *ACS Appl. Mater. Interfaces* **2015,** *7,* 20945–20956. doi:10.1021/acsami.5b06541
49. Jain, P. K.; Lee, K. S.; El-Sayed, I. H.; El-Sayed, M. A. *J. Phys. Chem. B* **2006,** *110,* 7238–7248. doi:10.1021/jp057170o
50. Jin, H.-Y.; Li, D.-W.; Zhang, N.; Gu, Z.; Long, Y.-T. *ACS Appl. Mater. Interfaces* **2015,** *7,* 12249–12253. doi:10.1021/acsami.5b02744
51. Roskamp, M.; Enders, S.; Pfrengle, F.; Yekta, S.; Dekaris, V.; Dernedde, J.; Reissig, H.-U.; Schlecht, S. *Org. Biomol. Chem.* **2011,** *9,* 7448–7456. doi:10.1039/c1ob05583f
52. Chien, Y.-Y.; Jan, M.-D.; Adak, A. K.; Tzeng, H.-C.; Lin, Y.-P.; Chen, Y.-J.; Wang, K.-T.; Chen, C.-T.; Chen, C.-C.; Lin, C.-C. *ChemBioChem* **2008,** *9,* 1100–1109. doi:10.1002/cbic.200700590
53. Fallarini, S.; Paoletti, T.; Battaglini, C. O.; Ronchi, P.; Lay, L.; Bonomi, R.; Jha, S.; Mancin, F.; Scrimin, P.; Lombardi, G. *Nanoscale* **2013,** *5,* 390–400. doi:10.1039/C2NR32338A
54. Kolhar, P.; Anselmo, A. C.; Gupta, V.; Pant, K.; Prabhakarpandian, B.; Ruoslahti, E.; Mitragotri, S. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110,* 10753–10758. doi:10.1073/pnas.1308345110
55. Sangabathuni, S.; Vasudeva Murthy, R.; Chaudhary, P. M.; Surve, M.; Banerjee, A.; Kikkeri, R. *Nanoscale* **2016,** *8,* 12729–12735. doi:10.1039/C6NR03008D
56. Avvakumova, S.; Fezzardi, P.; Pandolfi, L.; Colombo, M.; Sansone, F.; Casnati, A.; Prosperi, D. *Chem. Commun.* **2014,** *50,* 11029–11032. doi:10.1039/C4CC03159H
57. Parry, A. L.; Clemson, N. A.; Ellis, J.; Bernhard, S. S. R.; Davis, B. G.; Cameron, N. R. *J. Am. Chem. Soc.* **2013,** *135,* 9362–9365. doi:10.1021/ja4046857
58. Zhang, Z.; Schepens, B.; Nuhn, L.; Saelens, X.; Schotsaert, M.; Callewaert, N.; De Rycke, R.; Zhang, Q.; Moins, S.; Benali, S.; Mespouille, L.; Hoogenboom, R.; De Geest, B. G. *Chem. Commun.* **2016,** *52,* 3352–3355. doi:10.1039/C6CC00501B
59. Reynolds, M.; Marradi, M.; Imberty, A.; Penadés, S.; Pérez, S. *Chem. – Eur. J.* **2012,** *18,* 4264–4273. doi:10.1002/chem.201102034
60. Reynolds, M.; Marradi, M.; Imberty, A.; Penadés, S.; Pérez, S. *Glycoconjugate J.* **2013,** *30,* 747–757. doi:10.1007/s10719-013-9478-6
61. Kulkarni, A. A.; Fuller, C.; Korman, H.; Weiss, A. A.; Iyer, S. S. *Bioconjugate Chem.* **2010,** *21,* 1486–1493. doi:10.1021/bc100095w
62. Marradi, M.; Di Gianvincenzo, P.; Enríquez-Navas, P. M.; Martínez-Ávila, O. M.; Chiodo, F.; Yuste, E.; Angulo, J.; Penadés, S. *J. Mol. Biol.* **2011,** *410,* 798–810. doi:10.1016/j.jmb.2011.03.042
63. Manea, F.; Bindoli, C.; Fallarini, S.; Lombardi, G.; Polito, L.; Lay, L.; Bonomi, R.; Mancin, F.; Scrimin, P. *Adv. Mater.* **2008,** *20,* 4348–4352. doi:10.1002/adma.200800737
64. Ding, Y.; Liang, J.-J.; Geng, D.-D.; Wu, D.; Dong, L.; Shen, W.-B.; Xia, X.-H.; Zhang, C. *Part. Part. Syst. Charact.* **2014,** *31,* 347–356. doi:10.1002/ppsc.201300120
65. Richards, S.-J.; Gibson, M. I. *ACS Macro Lett.* **2014,** *3,* 1004–1008. doi:10.1021/mz5004882
66. Richards, S.-J.; Fullam, E.; Besra, G. S.; Gibson, M. I. *J. Mater. Chem. B* **2014,** *2,* 1490–1498. doi:10.1039/c3tb21821j
67. Grant, O. C.; Smith, H. M. K.; Firsova, D.; Fadda, E.; Woods, R. J. *Glycobiology* **2014,** *24,* 17–25. doi:10.1093/glycob/cwt083
68. Reichardt, N.-C.; Martín-Lomas, M.; Penadés, S. *Chem. Commun.* **2016,** 13430–13439. doi:10.1039/C6CC04445J
69. Rahman, M.; Laurent, S.; Tawil, N.; Yahia, L.; Mahmoudi, M. *Protein-Nanoparticle Interactions;* Springer-Verlag: Berlin Heidelberg, 2013. doi:10.1007/978-3-642-37555-2
70. Wan, S.; Kelly, P. M.; Mahon, E.; Stöckmann, H.; Rudd, P. M.; Caruso, F.; Dawson, K. A.; Yan, Y.; Monopoli, M. P. *ACS Nano* **2015,** *2,* 2157–2166. doi:10.1021/nn506060q
71. García, I.; Sánchez-Iglesias, A.; Henriksen-Lacey, M.; Grzelczak, M.; Penadés, S.; Liz-Marzán, L. M. *J. Am. Chem. Soc.* **2015,** *137,* 3686–3692. doi:10.1021/jacs.5b01001
72. Lai, C.-H.; Hütter, J.; Hsu, C.-W.; Tanaka, H.; Varela-Aramburu, S.; De Cola, L.; Lepenies, B.; Seeberger, P. H. *Nano Lett.* **2016,** *16,* 807–811. doi:10.1021/acs.nanolett.5b04984
73. Kopitzki, S.; Dilmaghani, K. A.; Thiem, J. *Tetrahedron* **2013,** *69,* 10621–10636. doi:10.1016/j.tet.2013.10.027
74. Kopitzki, S.; Thiem, J. *Eur. J. Org. Chem.* **2013,** *19,* 4008–4016. doi:10.1002/ejoc.201201648
75. Adak, A. K.; Lin, H.-J.; Lin, C.-C. *Org. Biomol. Chem.* **2014,** *12,* 5563–5573. doi:10.1039/C4OB00827H
76. Chiricozzi, E.; Ciampa, M. G.; Brasile, G.; Compostella, F.; Prinetti, A.; Nakayama, H.; Ekyalongo, R. C.; Iwabuchi, K.; Sonnino, S.; Mauri, L. *J. Lipid Res.* **2015,** *56,* 129–141. doi:10.1194/jlr.M055319
77. Srakaew, N.; Young, C. D.; Sae-wu, A.; Xu, H.; Quesnel, K. L.; di Brisco, R.; Kongmanas, K.; Fongmoon, D.; Hommalai, G.; Weerachatyanukul, W.; Hall, S. H.; Zhang, Y.-L.; Panza, L.; Franchini, L.; Compostella, F.; Pearson, T. W.; Hancock, R. E.; Oko, R. J.; Hermo, L. S.; Tanphaichitr, N. *Hum. Reprod.* **2014,** *29,* 683–696. doi:10.1093/humrep/deu018
78. Jayawardena, H. S. N.; Wang, X.; Yan, M. *Anal. Chem.* **2013,** *85,* 10277–10281. doi:10.1021/ac402069j
79. Otten, L.; Vlachou, D.; Richards, S.-J.; Gibson, M. I. *Analyst* **2016,** *141,* 4305–4312. doi:10.1039/C6AN00549G
80. Li, X.; Martin, S. J. H.; Chinoy, Z. S.; Liu, L.; Rittgers, B.; Dluhy, R. A.; Boons, G.-J. *Chem. – Eur. J.* **2016,** *22,* 11180–11185. doi:10.1002/chem.201602706
81. Sakurai, K.; Hatai, Y.; Okada, A. *Chem. Sci.* **2016,** *7,* 702–706. doi:10.1039/C5SC03275J
82. Wakao, M.; Watanabe, S.; Kurahashi, Y.; Matsuo, T.; Takeuchi, M.; Ogawa, T.; Suzuki, K.; Yumino, T.; Myogadani, T.; Saito, A.; Muta, K.-I.; Kimura, M.; Kajikawa, K.; Suda, Y. *Anal. Chem.* **2017,** *89,* 1086–1091. doi:10.1021/acs.analchem.6b02380
83. Ahn, K.-S.; Lim, K. R.; Jeong, D.; Lee, B. Y.; Kim, K. S.; Lee, W.-Y. *Microchem. J.* **2016,** *124,* 9–14. doi:10.1016/j.microc.2015.07.007

84. Chiodo, F.; Marradi, M.; Tefsen, B.; Snippe, H.; van Die, I.; Penadés, S. *PLoS One* **2013,** *8,* No. e73027. doi:10.1371/journal.pone.0073027
85. Poonthiyil, V.; Nagesh, P. T.; Husain, M.; Golovko, V. B.; Fairbanks, A. J. *ChemistryOpen* **2015,** *4,* 662–671. doi:10.1002/open.201500182
86. Lee, K. C.; Chiang, H. L.; Chiu, W. R.; Chen, Y. C. *J. Mol. Recognit.* **2016,** *29,* 528–535. doi:10.1002/jmr.2552
87. Aykac, A.; Martos-Maldonado, M. C.; Casas-Solvas, J. M.; Quesada-Soriano, I.; García-Maroto, F.; García-Fuentes, L.; Vargas-Berenguel, A. *Langmuir* **2014,** *30,* 234–242. doi:10.1021/la403454p
88. Adokoh, C. K.; Quan, S.; Hitt, M.; Darkwa, J.; Kumar, P.; Narain, R. *Biomacromolecules* **2014,** *15,* 3802–3810. doi:10.1021/bm5010977
89. Mu, H.; Liu, Q.; Niu, H.; Sun, Y.; Duan, J. *RSC Adv.* **2016,** *6,* 8714–8721. doi:10.1039/C5RA22803D
90. Yuan, Y.; Liu, F.; Xue, L.; Wang, H.; Pan, J.; Cui, Y.; Chen, H.; Yuan, L. *ACS Appl. Mater. Interfaces* **2016,** *8,* 11309–11317. doi:10.1021/acsami.6b02074
91. Chiodo, F.; Marradi, M.; Calvo, J.; Yuste, E.; Penadés, S. *Beilstein J. Org. Chem.* **2014,** *10,* 1339–1346. doi:10.3762/bjoc.10.136
92. Jølck, R. I.; Rydhög, J. S.; Christensen, A. N.; Hansen, A. E.; Bruun, L. M.; Schaarup-Jensen, H.; von Wenck, A. S.; Børresen, B.; Kristensen, A. T.; Clausen, M. H.; Kjær, A.; Conradsen, K.; Larsen, R.; af Rosenschöld, P. M.; Andresen, T. L. *Adv. Healthcare Mater.* **2015,** *4,* 856–863. doi:10.1002/adhm.201400651
93. Wilkins, L. E.; Phillips, D. J.; Deller, R. C.; Davies, G.-L.; Gibson, M. I. *Carbohydr. Res.* **2015,** *405,* 47–54. doi:10.1016/j.carres.2014.09.009
94. Candiota, A. P.; Acosta, M.; Simões, R. V.; Delgado-Goñi, T.; Lope-Piedrafita, S.; Irure, A.; Marradi, M.; Bomati-Miguel, O.; Miguel-Sancho, N.; Abasolo, I.; Schwartz, S.; Santamaria, J.; Penadés, S.; Arús, C. *J. Nanobiotechnol.* **2014,** *12,* No. 12. doi:10.1186/1477-3155-12-12
95. Irvine, D. J.; Hanson, M. C.; Rakhra, K.; Tokatlian, T. *Chem. Rev.* **2015,** *115,* 11109–11146. doi:10.1021/acs.chemrev.5b00109
96. Reichardt, N. C.; Martín-Lomas, M.; Penadés, S. *Chem. Soc. Rev.* **2013,** *42,* 4358–4376. doi:10.1039/c2cs35427f
97. Chiodo, F.; Marradi, M.; Park, J.; Ram, A. F. J.; Penadés, S.; van Die, I.; Tefsen, B. *ACS Chem. Biol.* **2014,** *9,* 383–389. doi:10.1021/cb4008265
98. Arosio, D.; Chiodo, F.; Reina, J. J.; Marelli, M.; Penadés,, S.; van Kooyk, Y.; Garcia-Vallejo, J. J.; Bernardi, A. *Bioconjugate Chem.* **2014,** *25,* 2244–2251. doi:10.1021/bc500467u
99. Chiodo, F.; Enríquez-Navas, P. M.; Angulo, J.; Marradi, M.; Penadés, S. *Carbohydr. Res.* **2015,** *405,* 102–109. doi:10.1016/j.carres.2014.07.012
100. Di Gianvincenzo, P.; Calvo, J.; Perez, S.; Álvarez, A.; Bedoya, L. M.; Alcamí, J.; Penadés, S. *Bioconjugate Chem.* **2015,** *26,* 755–765. doi:10.1021/acs.bioconjchem.5b00077
101. Biswas, S.; Medina, S. H.; Barchi, J. J. *Carbohydr. Res.* **2015,** *405,* 93–101. doi:10.1016/j.carres.2014.11.002
102. Tavernaro, I.; Hartmann, S.; Sommer, L.; Hausmann, H.; Rohner, C.; Ruehl, M.; Hoffmann-Roeder, A.; Schlecht, S. *Org. Biomol. Chem.* **2015,** *13,* 81–97. doi:10.1039/C4OB01339E
103. Rodriguez-Del Rio, E.; Marradi, M.; Calderon-Gonzalez, R.; Frande-Cabanes, E.; Penadés, S.; Petrovsky, N.; Alvarez-Dominguez, C. *Vaccine* **2015,** *33,* 1465–1473. doi:10.1016/j.vaccine.2015.01.062
104. Calderón-González, R.; Terán-Navarro, H.; Frande-Cabanes, E.; Ferrández-Fernández, E.; Freire, J.; Penadés, S.; Marradi, M.; García, I.; Gómez-Román, J. J.; Yáñez-Díaz,, S.; Álvarez-Domínguez, C. *Nanomaterials* **2016,** *6,* 151. doi:10.3390/nano6080151
105. Niikura, K.; Matsunaga, T.; Suzuki, T.; Kobayashi, S.; Yamaguchi, H.; Orba, Y.; Kawaguchi, A.; Hasegawa, H.; Kajino, K.; Ninomiya, T.; Ijiro, K.; Sawa, H. *ACS Nano* **2013,** *7,* 3926–3938. doi:10.1021/nn3057005
106. Vetro, M.; Safari, D.; Fallarini, S.; Salsabila, K.; Lahmann, M.; Penadés, S.; Lay, L.; Marradi, M.; Compostella, F. *Nanomedicine* **2017,** *12,* 13–23. doi:10.2217/nnm-2016-0306
107. Conde, J.; Tian, F.; Hernandez, Y.; Bao, C.; Baptista, P. V.; Cui, D.; Stoeger, T.; de la Fuente, J. M. *Nanoscale* **2015,** *7,* 9083–9091. doi:10.1039/C4NR05742B

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Urea–hydrogen peroxide prompted the selective and controlled oxidation of thioglycosides into sulfoxides and sulfones

Adesh Kumar Singh, Varsha Tiwari, Kunj Bihari Mishra, Surabhi Gupta and Jeyakumar Kandasamy*§

Full Research Paper

Open Access

Address:
Department of Chemistry, Indian Institute of Technology, Banaras Hindu University, Varanasi-221005, India

Email:
Jeyakumar Kandasamy* - jeyakumar.chy@iitbhu.ac.in

* Corresponding author
§ Tel.: +91-0542-6702879; fax: +91- 0542-6702876

Keywords:
monosaccharides; oxidation; sulfones; sulfoxides; thioglycosides; urea–hydrogen peroxide

Beilstein J. Org. Chem. **2017,** *13,* 1139–1144.
doi:10.3762/bjoc.13.113

Received: 23 March 2017
Accepted: 23 May 2017
Published: 13 June 2017

This article is part of the Thematic Series "Biomolecular systems".

Guest Editor: P. H. Seeberger

Abstract

A practical method for the selective and controlled oxidation of thioglycosides to corresponding glycosyl sulfoxides and sulfones is reported using urea–hydrogen peroxide (UHP). A wide range of glycosyl sulfoxides are selectively achieved using 1.5 equiv of UHP at 60 °C while corresponding sulfones are achieved using 2.5 equiv of UHP at 80 °C in acetic acid. Remarkably, oxidation susceptible olefin functional groups were found to be stable during the oxidation of sulfide.

Introduction

Organosulfur compounds such as sulfides, sulfoxides and sulfones are useful intermediates for the construction of highly functionalized natural products [1,2]. Sulfur moieties are found in several therapeutically important molecules that possess antibacterial, antifungal, anti-ulcer, anti-atherosclerotic, antihypertensive activities, etc. [3,4]. Sulfur compounds also play an important role in carbohydrate synthesis. Thioglycosides, glycosyl sulfoxides and sulfones have been widely used as glycosyl donors in oligosaccharide synthesis which can be activated under mild reaction conditions [5-10]. Glycosyl sulfoxide donors usually provide excellent anomeric selectivity during the synthesis of various glycosyl linkages not only in solution phase but also in solid-phase oligosaccharide synthesis [6-9,11]. Glycosyl sulfones were also used as donors in the preparation of various *C*- and *O*- linked oligosaccharides and functionalized glycols [8,9,12]. In addition, glycosyl sulfones are known to be potential glycosyltransferase inhibitors [13].

Glycosyl sulfoxides and sulfones are prepared from the corresponding sulfides using various oxidizing reagents [5-7,10]. Al-

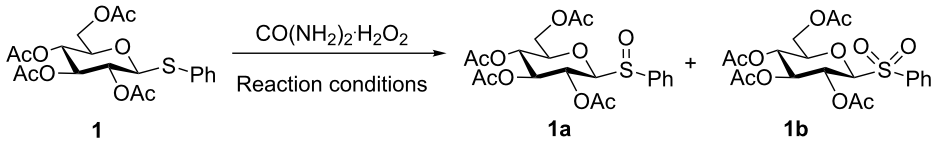
though a number of oxidation methods were developed for the oxidation of simple organic sulfides to corresponding sulfoxides and sulfones [14-16], there are only limited reports available for the preparation of glycosyl sulfoxides and sulfones from corresponding thioglycosides [5-7,17-23]. Moreover, there is no report available where a given oxidant is suitable for controlled oxidation of thioglycosides to glycosyl sulfoxides and sulfones selectively by altering the reaction conditions. It is also observed that thioglycoside oxidation suffers from low yields, poor selectivity (i.e., sulfoxide vs sulfone), use of inconvenient reaction conditions and expensive oxidants, intolerance of other oxidation susceptible functional groups, etc. Thus, developing a mild and efficient method for the controlled oxidation of sulfides to corresponding glycosyl sulfoxides and sulfones, is of great interest.

The utility of hydrogen peroxide–solid adducts in organic synthesis is well explored [24]. Most of them are found to be stable which can be easily handled and stored. One such solid adduct is urea–hydrogen peroxide (UHP) which is considered to be a safer and efficient alternative to high concentrated aqueous hydrogen peroxide solution [25]. In addition, UHP is also commercially available, inexpensive and nontoxic. The application of UHP as oxidant is well explored in various solution- as well as solid-phase organic syntheses [25-28]. In fact, we have recently reported the oxidation of arylboronic acids into corresponding phenols by using UHP as a selective oxidizing agent [29]. In continuation to our effort in developing green methodologies [29-33], here we disclose an efficient and practical method for the conversion of glycosyl sulfides into sulfoxides and sulfones in a selective and controlled manner using urea–hydrogen peroxide in acetic acid.

Results and Discussion

Initially, phenyl-2,3,4,6-tetra-*O*-acetyl-1-thio-β-D-glucopyranoside (**1**) was chosen as a substrate for the optimization study and oxidation was performed in various solvents at different temperatures in the presence of urea–hydrogen peroxide (UHP) (Table 1). Polar aprotic solvents such as dichloromethane and acetonitrile gave a negligible amount of corresponding sulfoxide (**1a**) while no sulfone (**1b**) was detected at room temperature even after 6 hours (Table 1, entries 1 and 2). However, protic solvents such as methanol, ethanol, *tert*-butanol and acetic acid were found to be relatively efficient media for the oxidation when compared with dichloromethane and acetonitrile (Table 1, entries 3–6). Among them, acetic acid gave 37% of the glycosyl sulfoxide (**1a**) after 6 h at room temperature with one equiv of UHP (Table 1, entry 6) while alcoholic

Table 1: Optimization of reaction conditions.[a]



1 **1a** **1b**

Entry	UHP (equiv)	Solvent	Temperature	Time	Yield (%)[b]	
					1a[c]	**1b**
1	1.0	DCM	rt	6 h	<5	n.d.
2	1.0	CH_3CN	rt	6 h	<10	n.d.
3	1.0	MeOH	rt	6 h	12	n.d.
4	1.0	EtOH	rt	6 h	15	n.d.
5	1.0	*t*-BuOH	rt	6 h	13	n.d.
6	1.0	AcOH	rt	6 h	37	n.d.
7	1.5	AcOH	rt	6 h	64	n.d.
8	1.5	AcOH	40 °C	6 h	90	n.d.
9	1.5	AcOH	60 °C	2 h	92	<5
10	1.5	AcOH	80 °C	2 h	87	7
11	2.0	AcOH	80 °C	3 h	65	31
12	2.5	AcOH	80 °C	10 h	<5	93

[a]Reaction conditions: Thioglycoside (0.25 mmol), solvent (2.5 mL) and urea–hydrogen peroxide (UHP) together stirred for appropriate time at different temperature. [b]Isolated Yield. [c]Obtained as *R* and *S* mixture.

solvents gave a low yield. When we increased the amount of UHP to 1.5 equiv, the reaction provides only 64% of the desired sulfoxide at room temperature (Table 1, entry 7). Therefore, the reaction was further investigated at elevated temperatures using 1.5 equiv of UHP (Table 1, entries 8 and 9) in acetic acid. Interestingly, the reaction was driven to completion with the desired sulfoxide (**1b**) in 92% yield within 2 h at 60 °C (Table 1, entry 9). It is also worth noting that less than 5% of the corresponding sulfone was detected in the crude product by ^{1}H NMR under these conditions.

Considering the importance of glycosyl sulfones, we further investigated the suitable conditions for the direct oxidation of sulfide to sulfone using UHP in acetic acid. For this, we have tried the reactions with an increased amount of UHP and elevated temperature (Table 1, entries 10–12). It was observed that with 1.5 to 2.0 equiv of UHP at 80 °C, the reaction yields a mixture of sulfoxide **1a** and sulfone **1b** (Table 1, entries 10–11) in different ratio. However, by increasing the amount of UHP to 2.5 equiv, sulfide **1** is fully converted to the corresponding sulfone **1b** in an excellent yield, i.e., 93% in 10 h at 80 °C (Table 1, entry 12).

With optimized conditions in hand (Table 1, entries 9 and 12), a controlled oxidation of various glycosyl sulfides to corresponding sulfoxides and sulfones was studied with urea-hydrogen peroxide in acetic acid (Table 2). For this study, a series of α- and β-thioglycosides, **1**–**19** were initially prepared by using literature procedures (see Supporting Information File 1). In addition, structurally diverse aglycone moieties were selected in order to study the breadth and scope of the current methodology. Initially, the oxidation of *O*-acetylated and benzoylated phenyl and *p*-tolyl thioglucopyranosides was examined (Table 2, entries 1–4). These aryl sulfides underwent oxidation very efficiently to provide the corresponding sulfoxides **1a–4a** in excellent yields, i.e., 85–93% under optimized conditions. Similarly,

Table 2: Controlled oxidation of various thioglycosides to corresponding sulfoxides and sulfones using urea–hydrogen peroxide (UHP).[a,b].

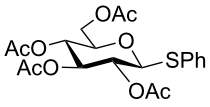
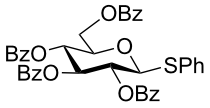
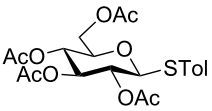
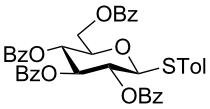
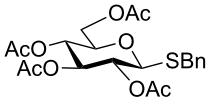
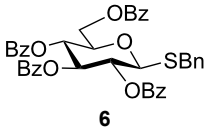
Entry	Substrate	Sulfoxide (**a**)[c]		Sulfone (**b**)	
		Time	Yield (%)[d]	Time	Yield (%)[d]
1	 **1**	2 h	92	10 h	93
2	 **2**	2 h	85	10 h	91
3	 **3**	2 h	93	10 h	94
4	 **4**	2 h	89	10 h	92
5	 **5**	1.5 h	90	8 h	94
6	 **6**	1.5 h	87	8 h	89

Table 2: Controlled oxidation of various thioglycosides to corresponding sulfoxides and sulfones using urea–hydrogen peroxide (UHP).[a,b]. (continued)

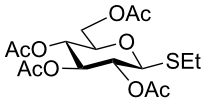
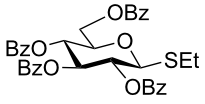
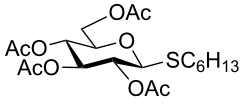
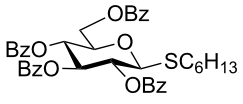
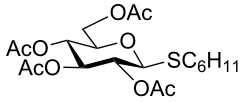
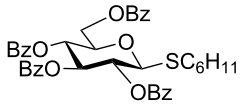
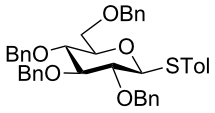
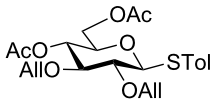
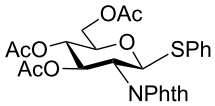
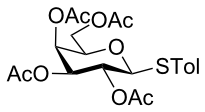
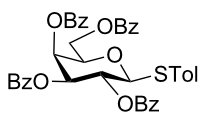
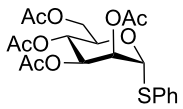
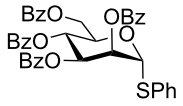
7	 **7**	1.5 h	92	8 h	94
8	 **8**	1.5 h	92	8 h	93
9	 **9**	2.0 h	87	10 h	93
10	 **10**	2.0 h	83	10 h	91
11	 **11**	2 h	90	10 h	92
12	 **12**	2 h	87	10 h	90
13	 **13**	1.5 h	82	6 h	64
14	 **14**	2 h	89	8 h	91
15	 **15**	2.5 h	91	8 h	82
16	 **16**	2.5 h	80	10 h	89
17	 **17**	2.5 h	77	11 h	87

Table 2: Controlled oxidation of various thioglycosides to corresponding sulfoxides and sulfones using urea–hydrogen peroxide (UHP).[a,b]. (continued)

18	 **18**	2 h	85	10 h	92
19	 **19**	2 h	86	11 h	89

[a]Reaction Conditions: Thioglycoside (0.25 mmol), acetic acid (2.5 mL) and UHP (1.5 equiv) stirred at 60 °C. [b]Reaction conditions: Thioglycoside (0.25 mmol), acetic acid (2.5 mL) and UHP (2.5 equiv) stirred at 80 °C. [c]Obtained as *R* and *S* mixture. [d]Isolated Yield.

corresponding glycosyl sulfones **1b**–**4b** were also achieved in 91–94% yield by simply altering the reaction conditions as described in the optimization study.

We further examined the oxidation of *O*-acetyl- (Ac) and benzoyl- (Bz) protected benzyl thioglucopyranosides, which showed a good selectivity during the controlled oxidation with UHP and provided >87% and >89% of the desired sulfoxides (**5a** and **6a**) and sulfones (**5b** and **6b**), respectively (Table 2, entries 5 and 6). Similar to phenyl and benzyl sulfides, alkyl sulfides such as ethyl, *n*-hexyl and cyclohexyl glycosyl sulfides were also successfully oxidized in a controlled and selective manner with equal efficiency (Table 2, entries 7–12).

Having studied the oxidation of electron deficient thioglucopyranosides, we further investigated the oxidation of *O*-benzyl protected 4-methylphenyl thioglycoside **13** under optimized conditions (Table 2, entry 13). This substrate was found to be more reactive than *O*-acetylated and benzoylated thioglycosides and gave the sulfoxide in a good yield within 1.5 h. However, corresponding sulfone was obtained in a moderated yield due to instability which undergoes partial amount of decomposition.

In general, olefins functional groups are known to undergo epoxidation or dihydroxylation with different oxidizing agents (e.g. *m*-CPBA, *t*-BuOOH, oxone, etc.) [34]. Therefore, the scope of this methodology was further investigated with oxidation of allyl group protected thioglycoside **14** (Table 2, entry 14). Remarkably, allyl groups were found to be very stable during the oxidation while sulfide underwent selective oxidation to corresponding sulfoxide and sulfone in 89% and 91%, respectively. Further, we have studied the oxidation of protected glucosamine thioglycoside (Table 2, entry 15) which provided 91% of sulfoxide and 82% of sulfone.

The scope of the oxidation reaction was subsequently investigated with other monosaccharides such as galacto and mannothioglycosides under optimized conditions (Table 2). Similar to glucopyranosides, galacto and mannothioglycosides **16**–**19** were successfully oxidized to corresponding sulfoxides and sulfones in good to excellent yields (Table 2, entries 16–19). Overall, sulfoxides were achieved within the time period of 1.5–2.5 h while sulfones were obtained in 6–11 h.

Conclusion

In conclusion, we have developed a practical method for the selective and controlled oxidation of thioglycosides to corresponding glycosyl sulfoxides and sulfones using the stable, inexpensive and commercially available oxidant urea–hydrogen peroxide (UHP). Glycosyl sulfoxides were achieved using 1.5 equiv of UHP at 60 °C while sulfones were achieved using 2.5 equiv of UHP at 80 °C. Remarkably, oxidation susceptible olefin functional groups were found to be stable during the sulfide oxidation.

Supporting Information

Supporting Information File 1

Experimental part and NMR spectra.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-13-113-S1.pdf]

Acknowledgements

J. K. gratefully acknowledges DST-India for young scientist start-up research grant (YSS/2014/000236). J. K. also acknowledges Max-Planck Society-Germany and DST-India (DST/INT/MPG/P-09/2016) for financial support through Indo-Max Planck partner group project. A. K. S acknowledges CSIR for junior research fellowship (JRF). S. G acknowledges IIT (BHU)

for a research fellowship. J. K. thanks to Mr. Albert Pape and Dr. K. Murugan for a helpful discussion during the manuscript preparation. J. K. thanks Central Instrumentation Facility Center (CIFC)-IIT BHU for the NMR facilities.

References

1. Rayner, C. M. *Advances in Sulfur Chemistry;* Jai Press: Stamford Connecticut, 2000.
2. Cremlyn, R. J. *An introduction to oranosulfur chemistry;* Wiley: New York, 1996.
3. Feng, M. H.; Tang, B. Q.; Liang, S. H.; Jiang, X. F. *Curr. Top. Med. Chem.* **2016,** *16,* 1200. doi:10.2174/1568026615666150915111741
4. Patani, G. A.; LaVoie, E. J. *Chem. Rev.* **1996,** *96,* 3147. doi:10.1021/cr950066q
5. Levy, D. E.; Fügedi, P. *The Organic Chemistry of Sugars;* CRC Press, Taylor & Francis Group, 2005.
6. Aversa, M. C.; Barattucci, A.; Bonaccorsi, P. *Tetrahedron* **2008,** *64,* 7659. doi:10.1016/j.tet.2008.05.051
7. Khan, S. H. *Modern Methods in Carbohydrate Synthesis;* CRC Press, Taylor & Francis Group, 1996.
8. Bennett, C. S. *Selective Glycosylations: Synthetic Methods and Catalysts;* Wiley-VCH Verlag GmbH & Co. KGaA, 2017.
9. Werz, D. B.; Vidal, S. *Modern Synthetic Methods in Carbohydrate Chemistry: From Monosaccharides to Complex Glycoconjugates;* Wiley-VCH Verlag GmbH & Co. KGaA, 2014.
10. Hanessian, S. Marcel Dekker: New York, 1997.
11. Kahne, D.; Walker, S.; Cheng, Y.; Van Engen, D. *J. Am. Chem. Soc.* **1989,** *111,* 6881. doi:10.1021/ja00199a081
12. Najera, C.; Yus, M. *Tetrahedron* **1999,** *55,* 10547. doi:10.1016/S0040-4020(99)00600-6
13. Gervay, J.; Flaherty, T. M.; Holmes, D. *Tetrahedron* **1997,** *53,* 16355. doi:10.1016/S0040-4020(97)01021-1
14. Kaczorowska, K.; Kolarska, Z.; Mitka, K.; Kowalski, P. *Tetrahedron* **2005,** *61,* 8315. doi:10.1016/j.tet.2005.05.044
15. Kowalski, P.; Mitka, K.; Ossowska, K.; Kolarska, Z. *Tetrahedron* **2005,** *61,* 1933. doi:10.1016/j.tet.2004.11.041
16. Jeyakumar, K.; Chand, D. K. *Tetrahedron Lett.* **2006,** *47,* 4573. doi:10.1016/j.tetlet.2006.04.153
17. Agnihotri, G.; Misra, A. K. *Carbohydr. Res.* **2006,** *341,* 275. doi:10.1016/j.carres.2005.11.025
18. Kakarla, R.; Dulina, R. G.; Hatzenbuhler, N. T.; Hui, Y. W.; Sofia, M. J. *J. Org. Chem.* **1996,** *61,* 8347. doi:10.1021/jo961478h
19. Wang, Q.; Wei, X.; Liao, K. J.; Li, H.; Meng, X. B.; Li, Z. J. *Tetrahedron Lett.* **2016,** *57,* 2277. doi:10.1016/j.tetlet.2016.04.045
20. Nigudkar, S. S.; Demchenko, A. V. *Chem. Sci.* **2015,** *6,* 2687. doi:10.1039/C5SC00280J
21. Agnihotri, G.; Misra, A. K. *Tetrahedron Lett.* **2005,** *46,* 8113. doi:10.1016/j.tetlet.2005.09.132
22. Chen, M.-Y.; Patkar, L. N.; Chen, H.-T.; Lin, C.-C. *Carbohydr. Res.* **2003,** *338,* 1327. doi:10.1016/S0008-6215(03)00164-2
23. Morais, G. R.; Humphrey, A. J.; Falconer, R. A. *Tetrahedron* **2008,** *64,* 7426. doi:10.1016/j.tet.2008.05.033
24. Jones, C. W. *Applications of Hydrogen Peroxide and Derivatives;* Royal Society of Chemistry: Cambridge, 1999.
25. Cooper, M. S.; Heaney, H.; Newbold, A. J.; Sanderson, W. R. *Synlett* **1990,** 533. doi:10.1055/s-1990-21156
26. Varma, R. S.; Naicker, K. P. *Org. Lett.* **1999,** *1,* 189. doi:10.1021/ol990522n
27. Caron, S.; Do, N. M.; Sieser, J. E. *Tetrahedron Lett.* **2000,** *41,* 2299. doi:10.1016/S0040-4039(00)00165-9
28. Ankudey, E. G.; Olivo, H. F.; Peeples, T. L. *Green Chem.* **2006,** *8,* 923. doi:10.1039/B604984B
29. Gupta, S.; Chaudhary, P.; Srivastava, V.; Kandasamy, J. *Tetrahedron Lett.* **2016,** *57,* 2506. doi:10.1016/j.tetlet.2016.04.099
30. Chaudhary, P.; Gupta, S.; Popuri, S.; Sabiah, S.; Kandasamy, J. *Green Chem.* **2016,** *18,* 6215. doi:10.1039/C6GC02444K
31. Gupta, S.; Chaudhary, P.; Seva, L.; Sabiah, S.; Kandasamy, J. *RSC Adv.* **2015,** *5,* 89133. doi:10.1039/C5RA18080E
32. Chaudhary, P.; Gupta, S.; Muniyappan, N.; Sabiah, S.; Kandasamy, J. *Green Chem.* **2016,** *18,* 2323. doi:10.1039/C5GC02880A
33. Gupta, S.; Sureshbabu, P.; Singh, A. K.; Sabiah, S.; Kandasamy, J. *Tetrahedron Lett.* **2017,** *58,* 909. doi:10.1016/j.tetlet.2017.01.051
34. Bäckvall, J.-E. *Modern Oxidation Methods,* 2nd ed.; Wiley-VCH: Weinheim, 2010. doi:10.1002/9783527632039

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Characterization of non-heme iron aliphatic halogenase WelO5* from *Hapalosiphon welwitschii* IC-52-3: Identification of a minimal protein sequence motif that confers enzymatic chlorination specificity in the biosynthesis of welwitindolelinones

Qin Zhu and Xinyu Liu*

Letter

Open Access

Address:
Department of Chemistry, University of Pittsburgh, 219 Parkman Avenue, Pittsburgh, PA 15260, USA

Email:
Xinyu Liu* - xinyuliu@pitt.edu

* Corresponding author

Keywords:
alkaloid biogenesis; biosynthetic divergency; C–H activation; halogenase; non-heme iron enzyme

Beilstein J. Org. Chem. **2017,** *13,* 1168–1173.
doi:10.3762/bjoc.13.115

Received: 22 March 2017
Accepted: 13 June 2017
Published: 16 June 2017

This article is part of the Thematic Series "Biomolecular systems".

Guest Editor: P. H. Seeberger

Abstract

The in vitro biochemical characterization revealed that iron/2-oxoglutarate (Fe/2OG)-dependent aliphatic halogenase WelO5* in *Hapalosiphon welwitschii* IC-52-3 has an enhanced substrate specificity towards 12-*epi*-hapalindole C (**1**) in comparison to WelO5 in *H. welwitschii* UTEX B1830. This allowed us to define the origin of the varied chlorinated versus dechlorinated alkaloid structural diversity between the two welwitindolinone producers. Furthermore, this study, along with the recent characterization of the AmbO5 protein, collectively confirmed the presence of a signature sequence motif in the C-terminus of this newly discovered halogenase enzyme family that confers substrate promiscuity and specificity. These observations may guide the rational engineering and evolution of these proteins for biocatalyst application.

Introduction

Carbon–halogen (C–X) bonds are prevalent structural motifs in modern agrochemicals, pharmaceuticals and bioactive natural products [1,2] and chlorination is the most common functionalization of this type [1,2]. Among other effects, chlorination enhances the electrophilicity of the modified carbon and alters the biological activities of drug(-like) molecules [3,4]. As aliphatic C–H groups are ubiquitous in organic molecules, synthetic transformations that allow for the selective modification of this type of functional group have been long sought after. While numerous methods have emerged that permit the late-stage

functionalizations of inert aliphatic carbons with oxygen-containing functionalities [5-8], analogous oxidative functionalizations with halogens via C–H activations remain challenging that need to be addressed [9-11]. Recently, during the systematic investigation of hapalindole-type alkaloid biogenesis [12-19], we discovered a family of Fe/2OG-dependent halogenases that can oxidatively monochlorinate aliphatic carbon centers in free-standing molecules, typified by WelO5 protein in the biogenesis of welwitindolinones [17]. This discovery provided a new opportunity to utilize protein catalysts for late-stage halogenations of unactivated carbons in bioactive small molecules.

Although the initially characterized WelO5 has a restricted substrate scope [17], we have recently shown its homolog, AmbO5, in the biogenesis of ambiguines is capable of modifying seven structurally distinct hapalindole-type alkaloids [18]. The biochemical characterizations of WelO5/AmbO5 chimera revealed that a C-terminal sequence motif plays a role in the substrate tolerance and provided insights into the origin of substrate promiscuity in this family of proteins [18]. In this work, we report the characterization of the third WelO5-type protein, WelO5*, for the biogenesis of welwitindolinones in *H. welwitschii* IC-52-3. We show that WelO5* exhibits an enhanced specificity towards 12-*epi*-hapalindole C (**1**), a substrate poorly processed by WelO5 from *H. welwitschii* UTEX B1830, while maintaining its fidelity towards 12-*epi*-fischerindole U (**2**) as WelO5. This observation provided a molecular basis for the altered structural diversity of hapalindole-type alkaloids between the two welwitindolinone producers. The extreme sequence similarity (95% identical) between WelO5* and WelO5 allowed us to trace the origin of this observed specificity difference to 11 amino acid residues at a C-terminal sequence motif, initially discovered in the comparative characterization of WelO5 and AmbO5 [18]. This further confirms the functional significance of this conserved sequence motif in this new halogenase family that may guide the rational engineering and evolution of these proteins for biocatalyst application.

Results and Discussion

H. welwitschii IC-52-3 and UTEX B1830 are two known welwitindolinone producers that were reported to produce identical sets of hapalindole-type alkaloids [20], albeit the detailed metabolite analysis from the latter was never published. During our recent effort to define the genetic and molecular basis of welwitindolinone biogenesis, we re-validated the profiles of hapalindole-type alkaloids in *H. welwitschii* UTEX B1830 as originally claimed [13]. This effort in turn led us to recognize there are two distinct differences in terms of hapalindole-type alkaloid structural diversities in these welwitindolinone producers. When the alkaloid molecules are grouped based on their biogenetic relatedness across two producing organisms (Figure 1a and Figure S1, Supporting Information File 1), the relative quantities of 12-*epi*-hapalindole C (**1**) and its biogenetic derivatives (i.e., 12-*epi*-hapalindole E (**1a**)) constitute more than 1/3 of the total hapalindole-type alkaloids isolated from *H. welwitschii* IC-52-3. This ratio is noticeably higher than what was observed for *H. welwitschii* UTEX B1830 (Figure 1b). Moreover, the ratio of chlorinated **1a** to dechlorinated **1** present in *H. welwitschii* IC-52-3 is significantly higher than that in *H. welwitschii* UTEX B1830 (Figure 1c). To understand the genetic and molecular basis for these differences in structural diversity, we compared the corresponding welwitindolinone biosynthetic gene clusters (BGCs) identified from *H. welwitschii* IC-52-3 and UTEX B1830 (Figure 1d) [13,21]. While the majority of the biosynthetic enzymes encoded in the two pathways are sequence-identical (Figure S2, Supporting Information File 1), two noticeable differences were identified (Figure 1d, red rectangle and circle highlights). First, the welwitindolinone pathway in *H. welwitschii* IC-52-3 contains a unique protein coded by *wel*U3 gene. We have recently characterized WelU3 protein by in vitro reconstitution and demonstrated it is a dedicated enzyme for the biosynthesis of **1** from 3-geranyl 3-(isocyanovinyl)indolenine [16], a common intermediate used in the biogenesis of all hapalindole-type alkaloids [15]. The presence of WelU3 in *H. welwitschii* IC-52-3 thus accounted for the observed increased quantities of **1** in this producer.

The second key difference between the two welwitindolinone BGCs resides on the halogenase coding gene *wel*O5. From the BLAST-P analysis, the protein sequence of WelO5 in *H. welwitschii* IC-52-3 is nearly (95%) identical to that in *H. welwitschii* UTEX B1830 (Figure 1d). Due to their close resemblance, we rename WelO5 in *H. welwitschii* IC-52-3 as WelO5* to facilitate the remaining discussion. Upon a close inspection of the sequence differences between WelO5* and WelO5, we realized that 11 out of the 15 varied amino acids fall into residues 215-232 (Figure 1e). We have shown the same type of C-terminal sequence motif in AmbO5, the Fe/2OG-dependent halogenase involved in the biogenesis of ambiguines, plays a role in its expanded substrate scope [18]. Moreover, our recent structural characterizations of WelO5 in the absence and presence of **2** have shown residues 215-232 of WelO5 encode an α-helical motif that helps keeping the small molecular substrate in the active site by undergoing a dramatic conformational change upon substrate binding [22]. Based on these earlier observations on the functional relevance of this C-terminal sequence motif in WelO5 and AmbO5 halogenases and the fact that the sequence of WelO5* differs near exclusively from that of WelO5 in the same region, we hypothesize that WelO5*

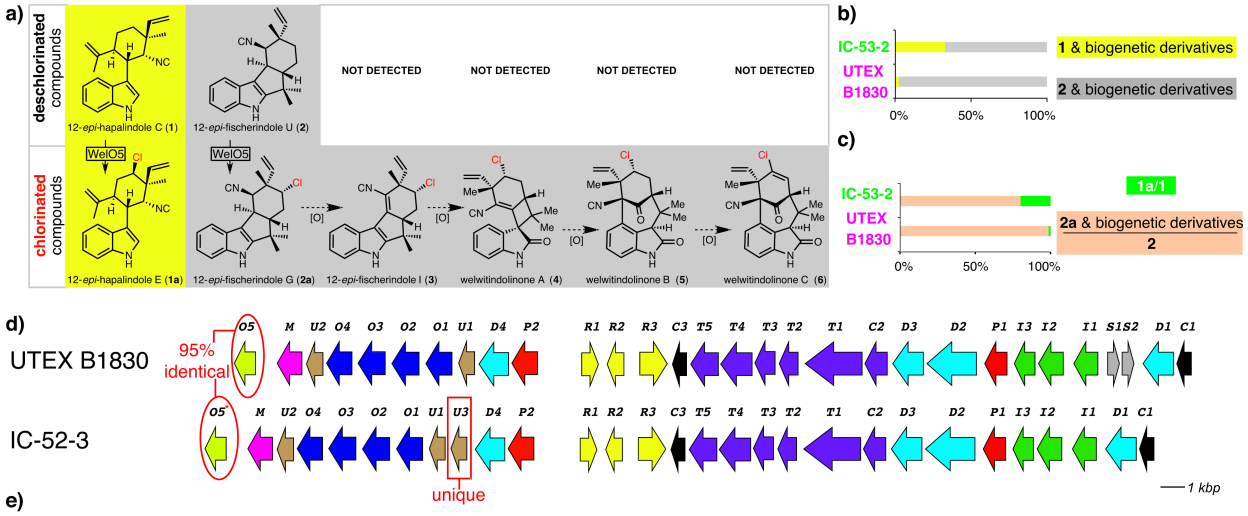


Figure 1: Comparative analysis of hapalindole-type alkaloids and their BGCs in two welwitindolinone producers implicates the functional role of Fe/2OG-dependent WelO5* halogenase in structural diversifications in *H. welwitschii* IC-52-3. a) Representative hapalindole-type alkaloids tabulated based on oxidation states in the welwitindolinone producers *H. welwitschii* UTEX B1830 and IC-52-3. b) Relative quantities of **1** and its biogenetic derivative **1a** versus **2** and its biogenetic derivatives (**2a**, **3**–**6**) in *H. welwitschii* IC-52-3 and UTEX B1830. c) Comparison of the ratio of chlorinated versus dechlorinated hapalindole-type alkaloids based on their biosynthetic origins (i.e., **1** or **2**) in *H. welwitschii* IC-52-3 and UTEX B1830. d) Comparison of welwitindolinone BGCs in *H. welwitschii* IC-52-3 and UTEX B1830. The tailoring enzyme coding genes, including *wel*U1-3, *wel*O1-5 and *wel*M, encoded in the welwitindolinone BGC from *H. welwitschii* IC-52-3, were renamed due to their extreme similarities to those from *H. welwitschii* UTEX B1830. e) Sequence comparison of WelO5 and WelO5*. Varied residues are highlighted in red with the most aggregated region 215-232 in grey rectangle.

must have an altered substrate preference to **1** and **2** to account for the observation that the ratio of chlorinated **1a** versus dechlorinated **1** present in *H. welwitschii* IC-52-3 is significantly higher than that in *H. welwitschii* UTEX B1830.

To test this hypothesis, the *wel*O5* gene was synthesized and ligated into the expression vector pQTEV. Heterologous expression in *E. coli* and purification by immobilized metal affinity chromatography (IMAC) gave the N-terminal hepta-His-tagged WelO5* in a comparable yield (20 mg/L) as for WelO5 [11]. With abundant WelO5* in hand, we proceeded on its in vitro characterization using the assay conditions established for WelO5 and AmbO5 [18]. For a 100 µL scale reaction, WelO5* (20 µM final concentration) rapidly converted circa 50% of **1** and **2** (0.5 mM final concentration) to their chlorinated derivatives **1a** and **2a** within 20 min in the presence of cosubstrate 2OG, cofactor Fe(II) and molecular oxygen (Figure 2a, bottom two lanes). Under identical conditions, WelO5 showed a comparable conversion rate of **2** to **2a** but was much more sluggish towards **1** (Figure 2a, top two lanes), consistent with our previous observation [17]. While a full steady state kinetic analysis remains challenging due to the limited substrate availability, we assessed the k_{obs} of WelO5* towards **1** and **2**, as described for WelO5 and AmbO5 [18]. WelO5* exhibits nearly identical k_{obs} towards **1** (1.8 ± 0.1 min^{-1}) and **2** (1.9 ± 0.2 min^{-1}), distinct from WelO5 which prefers **2** (k_{obs} = 1.8 ± 0.2 min^{-1}) to **1** (k_{obs} = 0.73 ± 0.08 min^{-1}) [17]. These analyses collectively confirm WelO5* does have an enhanced substrate specificity towards **1** in comparison to WelO5, while maintaining its activity towards **2**.

To gain further insights into the substrate preference of WelO5*, complementary to the k_{obs} measurement, we assessed its in vitro activity towards an equimolar amount of **1** and **2** (Figure 2b). For a 100 µL scale reaction with an equal amount of **1** and **2** (0.25 mM final concentration for each molecule), WelO5* (20 µM final concentration) was able to convert ca. 50% of **2** to **2a** and 33% of **1** to **1a** within 10 min (Figure 2b), indicating that WelO5* has a higher affinity (ca. two-fold) towards **2** than **1** under the enlisted in vitro assay conditions. This observation augments the altered substrate specificity of WelO5* versus WelO5 as the molecular basis for the varied structural diversities observed between the two welwitindolinone producers (Figure 1c).

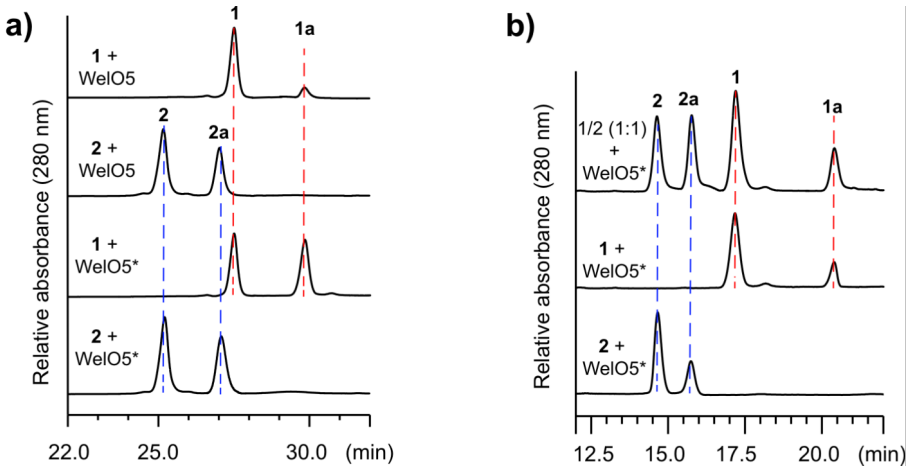


Figure 2: a) In vitro characterizations show that WelO5* has comparable activity to both **1** and **2**, distinct from WelO5. b) In vitro characterization of WelO5* substrate preference towards **1** and **2** using a substrate competition assay. For the HPLC data shown in a), assays were conducted with WelO5* or WelO5 (20 μM) with 0.5 mM of **1** or **2** for 20 min at 30 °C. HPLC was run with a C18 Luna column (Phenomenex 250 × 4.6 mm) with a solvent gradient from 60–90% methanol/water in 30 min at a flow rate 1 mL/min. For the HPLC data shown in b), assays were conducted with WelO5* (20 μM) with 0.5 mM of **1** and/or **2** for 10 min at 30 °C. HPLC was run with a XC-C18 Kinetex column (Phenomenex 150 × 2.6 mm) with an isocratic solvent 50% acetonitrile/water in 30 min at a flow rate 0.4 mL/min. These conditions were chosen to achieve a better separation of **1**/**1a**/**2**/**2a**.

As 11 out of the 15 amino acid variations between WelO5* and WelO5 fall into residues 215–232, a sequence motif that was shown previously to play a role in the expanded substrate scope of AmbO5 [7], we hypothesize the same motif in WelO5* may be important for its enhanced specificity towards **1**. To examine this hypothesis, we generated a variant of WelO5 (WelO5-*var*) by swapping its residues 215–232 to those in WelO5*. This variant was heterologously expressed and purified in an identical manner as the wild-type WelO5 and WelO5*. WelO5-*var* displayed a noticeably enhanced activity towards **1** based on the standard HPLC-based in vitro assay (Figure 3). The k_{obs} measurement (1.8 ± 0.2 min^{-1} for **1** and 1.8 ± 0.1 min^{-1} for **2**) indicates the activity of WelO5-*var* towards **1** is elevated to a comparable level as the wild type WelO5* while retaining its fidelity towards **2**. This result provides evidence that WelO5 can be rendered as specific as WelO5* towards **1** by replacing 11 varied amino acids between residues 215–232 with those of WelO5* and validates our hypothesis that the sequence variations between WelO5* and WelO5 at residues 215–232 are directly responsible for their enhanced or diminished specificity towards **1**.

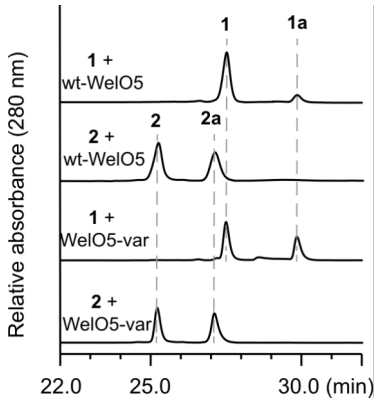


Figure 3: In vitro characterization of a WelO5 variant with enhanced specificity towards **1**. All of the HPLC data shown are in vitro assays conducted with WelO5* or its variants (20 μM) with 0.5 mM of **1** or **2** for 20 min at 30 °C. HPLC was run with a C18 Luna column (Phenomenex 250 × 4.6 mm) with a solvent gradient from 60–90% methanol/water in 30 min at a flow rate 1 mL/min.

Conclusion

In summary, intrigued by the hapalindole-type alkaloid structural diversity difference between *H. welwitschii* IC-52-3 and UTEX B1830, we examined the enzymatic activity of Fe/2OG-dependent WelO5* halogenase. Although WelO5* is nearly sequence-identical to the previously characterized WelO5, it showed enhanced chlorination activity towards **1**, distinct from WelO5. This study, along with the recent characterizations of WelU1 and WelU3 enzymes in *H. welwitschii* IC-52-3 [16], collectively provides the molecular basis for the altered structural diversity between the two welwitindolinone producers. Furthermore, the close sequence similarity between WelO5* and WelO5 allowed us to reveal a C-terminal sequence motif (residues 215–232) that harbors 11 varied amino acids between the two proteins plays the most critical role on the observed enhanced activity of WelO5* towards **1**. While the mechanism underlying how this sequence motif controls the substrate tolerance and specificity, as seen in AmbO5 previously and WelO5* in this study, is a subject for future studies. However its pres-

ence in this newly discovered halogenase family provides an entry point for the rational engineering of these enzymes for tailoring small molecules beyond hapalindole-type alkaloids.

Experimental

Protein expression and purification: Genes coding WelO5*, WelO5-*var* proteins were synthesized and ligated into pQTEV vector by BioBasic Inc. Heterologous expressions of WelO5* and WelO5-*var* in *E. coli* and subsequent purifications by IMAC were conducted in an identical fashion as described for WelO5 [17]. Protein homogeneities were assessed by SDS-PAGE analysis (Figure S3, Supporting Information File 1). The approximate yield for each protein is 20 mg/L.

In vitro assay: Substrates **1** and **2** were procured by isolation as previously described [17]. In vitro assays with a single or double small molecular substrate(s) (**1** and/or **2**) were conducted on a 100 µL scale, with 20 µM of the enzyme, 0.5 mM of the small molecular substrate(s) and appropriate cosubstrate/cofactors as described exactly for WelO5 and AmbO5 [18]. Reactions were stopped at 2 min (for k_{obs} estimation), 10 min (for Figure 2b) or 20 min (for Figure 2a/3) and extracted with EtOAc before subjecting to HPLC analysis as previously described [18]. Analytical reversed-phase HPLC was performed using a Dionex UHPLC with a photodiode array UV–vis detector (Thermo Fisher Scientific) using either a C18 Luna column (Phenomenex 250 × 4.6 mm) or a XC-C18 Kinetex column (Phenomenex 150 × 2.6 mm). The conversion ratio for each enzymatic assay from **1** to **1a** or **2** to **2a** was quantified by comparing the corresponding HPLC peak areas for compounds **1**, **1a**, **2** or **2a**, assuming their extinction coefficients at 280 nm are identical.

Supporting Information

Supporting Information File 1

Additional figures.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-13-115-S1.pdf]

Acknowledgements

The authors acknowledge University of Pittsburgh and National Institute of Health for supporting this work.

References

1. Gribble, G. W. Naturally Occurring Organohalogen Compounds - A Comprehensive Update. *Progress in the Chemistry of Organic Natural Products;* Springer Verlag: Vienna, Austria, 2010; Vol. 91, pp 1–613. doi:10.1007/978-3-211-99323-1
2. Smith, B. R.; Eastman, C. M.; Njardarson, J. T. *J. Med. Chem.* **2014,** *57,* 9764–9773. doi:10.1021/jm501105n
3. Xu, Z.; Yang, Z.; Liu, Y.; Lu, Y.; Chen, K.; Zhu, W. *J. Chem. Inf. Model.* **2014,** *54,* 69–78. doi:10.1021/ci400539q
4. Williams, P. G.; Buchanan, G. O.; Feling, R. H.; Kauffman, C. A.; Jensen, P. R.; Fenical, W. *J. Org. Chem.* **2005,** *70,* 6196–6203. doi:10.1021/jo050511+
5. Newhouse, T.; Baran, P. S. *Angew. Chem., Int. Ed.* **2011,** *50,* 3362–3374. doi:10.1002/anie.201006368
6. White, M. C. *Science* **2012,** *335,* 807–809. doi:10.1126/science.1207661
7. McIntosh, J. A.; Farwell, C. C.; Arnold, F. H. *Curr. Opin. Chem. Biol.* **2014,** *19,* 126–134. doi:10.1016/j.cbpa.2014.02.001
8. Cernak, T.; Dykstra, K. D.; Tyagarajan, S.; Vachal, P.; Krska, S. W. *Chem. Soc. Rev.* **2016,** *45,* 546–576. doi:10.1039/C5CS00628G
9. Liu, W.; Groves, J. T. *J. Am. Chem. Soc.* **2010,** *132,* 12847–12849. doi:10.1021/ja105548x
10. Schmidt, V. A.; Quinn, R. K.; Brusoe, A. T.; Alexanian, E. J. *J. Am. Chem. Soc.* **2014,** *136,* 14389–14392. doi:10.1021/ja508469u
11. Quinn, R. K.; Könst, Z. A.; Michalak, S. E.; Schmidt, Y.; Szklarski, A. R.; Flores, A. R.; Nam, S.; Horne, D. A.; Vanderwal, C. D.; Alexanian, E. J. *J. Am. Chem. Soc.* **2016,** *138,* 696–702. doi:10.1021/jacs.5b12308
12. Hillwig, M. L.; Zhu, Q.; Liu, X. *ACS Chem. Biol.* **2014,** *9,* 372–377. doi:10.1021/cb400681n
13. Hillwig, M. L.; Fuhrman, H. A.; Ittiamornkul, K.; Sevco, T. J.; Kwak, D. H.; Liu, X. *ChemBioChem* **2014,** *15,* 665–669. doi:10.1002/cbic.201300794
14. Ittiamornkul, K.; Zhu, Q.; Gkotsi, D. S.; Smith, D. R. M.; Hillwig, M. L.; Nightingale, N.; Goss, R. J. M.; Liu, X. *Chem. Sci.* **2015,** *6,* 6836–6840. doi:10.1039/C5SC02919H
15. Liu, X.; Hillwig, M. L.; Koharudin, L. M. I.; Gronenborn, A. M. *Chem. Commun.* **2016,** *52,* 1737–1740. doi:10.1039/C5CC10060G
16. Zhu, Q.; Liu, X. *Chem. Commun.* **2017,** *53,* 2826–2829. doi:10.1039/C7CC00782E
17. Hillwig, M. L.; Liu, X. *Nat. Chem. Biol.* **2014,** *10,* 921–923. doi:10.1038/nchembio.1625
18. Hillwig, M. L.; Zhu, Q.; Ittiamornkul, K.; Liu, X. *Angew. Chem., Int. Ed.* **2016,** *55,* 5780–5784. doi:10.1002/anie.201601447
19. Zhu, Q.; Liu, X. *Angew. Chem., Int. Ed.* **2017,** *56*. doi:10.1002/anie.201703932
20. Stratmann, K.; Moore, R. E.; Bonjouklian, R.; Deeter, J. B.; Patterson, G. M. L.; Shaffer, S.; Smith, C. D.; Smitka, T. A. *J. Am. Chem. Soc.* **1994,** *116,* 9935–9942. doi:10.1021/ja00101a015
21. Micallef, M. L.; Sharma, D.; Bunn, B. M.; Gerwick, L.; Viswanathan, R.; Moffitt, M. C. *BMC Microbiol.* **2014,** *14,* No. 213. doi:10.1186/s12866-014-0213-7
22. Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. *Nat. Chem. Biol.* **2016,** *12,* 636–640. doi:10.1038/nchembio.2112