Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 15 (2005) 3619–3622
Bioorganic & Medicinal Chemistry Letters

Synthesis of multitopic neoglycopeptides displaying recognition and detection motifs

Olivier Renaudet and Pascal Dumy*

LEDSS, UMR-CNRS 5616 & ICMG FR 2607, Université Joseph Fourier, BP 53 38041 Grenoble Cedéx 9, France

Received 22 February 2005; revised 3 May 2005; accepted 11 May 2005

Abstract—We describe herein the synthesis of cyclic decapeptide template displaying clustered carbohydrate recognition motifs and detection agent on spatially separated domains. Such multitopic labeled neoglycopeptides represent attractive tools for binding assays with carbohydrate binding proteins in glycomic research.

Many biological processes including fertilization, tissue formation, cell adhesion, antigen/antibody interactions, cancer metastasis, and infection of viruses or bacteria involve polyvalent interactions between oligomeric proteins and recognition motifs.[1] The design of molecules displaying multiple binding proteins elements,[2] particularly carbohydrate-based ligands, represents an emerging challenge to investigate these fundamental recognition processes[3] and to discover new active and selective therapeutics.[4] In the past few years, considerable synthetic efforts have been expended in this direction, leading to the development of a diverse supply of glycoconjugates,[5] such as calixarenes, cyclodextrines, dendrimers, or carbohydrate-based combinatorial libraries,[6] which have shown enhancement of affinity with target proteins through multivalent interactions. However, access to well-defined multitopic structures displaying independent probes (e.g., fluorescent agent, dye, or biotin) and clustered recognition motifs remains of great interest, enabling the high-throughput screening with unlabeled target proteins by multiple binding assay formats.[7]

The crucial point of this approach focuses on controlling the molecular assembly processes, which is essential to warrant the multitopic structure of the target molecule. While many chemical or chemoenzymatic methodologies including solution and solid-phase synthesis protocols have been reported so far, the assembly of carbohydrate-based conjugates remains fastidious due to a number of manipulations of orthogonal protecting groups.[8] As a convenient alternative, we have recently reported that regioselectively addressable functionalized templates (RAFTs) provide a useful platform for the incorporation of multiple copies of peptide[9] or sugar[10] recognition motifs through an oxime-based strategy.

RAFT molecules are commonly composed of a backbone-cyclized decapeptide containing two prolines-glycines as β-turn inducers that stabilize the conformation in solution. While four lysine side chains were previously used as addressable sites for multivalent presentation of carbohydrates[10] or protein de novo design,[11] two additional lysines[9,12] can be incorporated to the sequence, providing molecules that allow functional components to be assembled and directed in well-defined and controlled spatial orientations.[13] Taking advantage of this topological feature, we report herein the preparation of well-defined RAFT molecules displaying two separated addressable domains to get multitopic labeled neoglycoconjugates. One domain of the molecule is devoted to the presentation of multiple copies of carbohydrate, which ensure the recognition properties of our system (*recognition domain*, Fig. 1). On the second domain, a molecular probe, such as fluorescein or biotin, is incorporated to characterize the interaction with the target protein (*detection domain*, Fig. 1). Both detection and carbohydrate moieties are incorporated into the molecules sequentially, by using orthogonal protecting groups Boc (lysines 3, 5, 8, and 10) and Aloc (lysines 4 and 9).[9,12]

Keywords: Oxime ligation; Template; Multivalency; Glycopeptide.
* Corresponding author. Tel.: +33 476635545; fax: +33 476514382; e-mail: pascal.dumy@ujf-grenoble.fr

0960-894X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2005.05.072

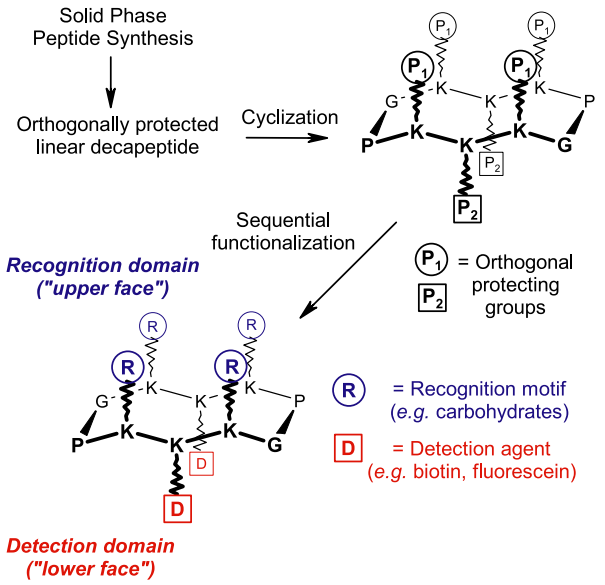
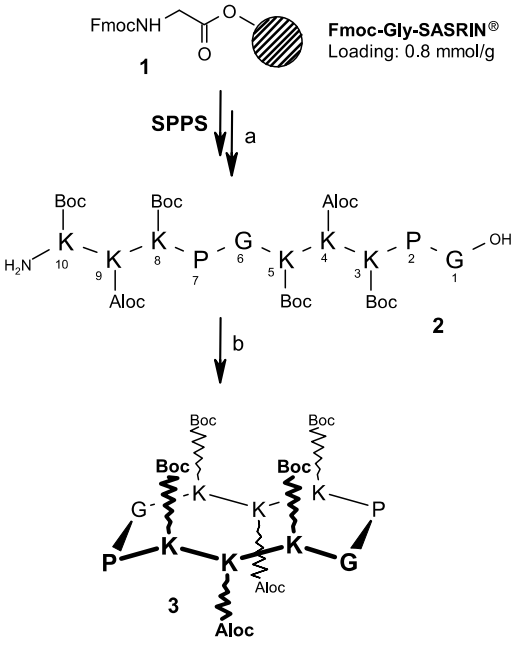


Figure 1. Solution strategy for the preparation of topological template.

The linear orthogonally protected decapeptide **2** is first prepared, following the standard Fmoc/*t*-Bu strategy from the highly acid-labile Fmoc-Gly-SASRIN® resin **1** using a parallel peptide synthesizer (Scheme 1). After the removal of N-terminal Fmoc-protecting group, the linear peptide **2** is cleaved from the support by treatment of a mixture 1% trifluoroacetic acid (TFA) in dichloromethane. Further head-to-tail cyclization of the template takes place with PyBOP[14] as the coupling reagent in the presence of diisopropylethylamine (DIEA) in dimethylformamide under high dilution (0.5 mM) to prevent oligomerization. After 30 min at room temperature, the corresponding protected cyclodecapeptide **3** is obtained in 86% yield after precipitation from diethyl ether. The sequential topological functionalization of the template is readily performed after regioselective removal of orthogonal protections (Scheme 2). Lysines at positions 4 and 9 bearing Aloc are first deprotected following a typical procedure, by treatment of a catalytic amount of $Pd(Ph_3P)_4$ and a large excess of $PhSiH_3$ in CH_2Cl_2 under nitrogen gas.[15] The cyclopeptide **4** displaying two free amino side chains on the detection face is obtained with 79% yield after semi-preparative reverse-phase HPLC purification.



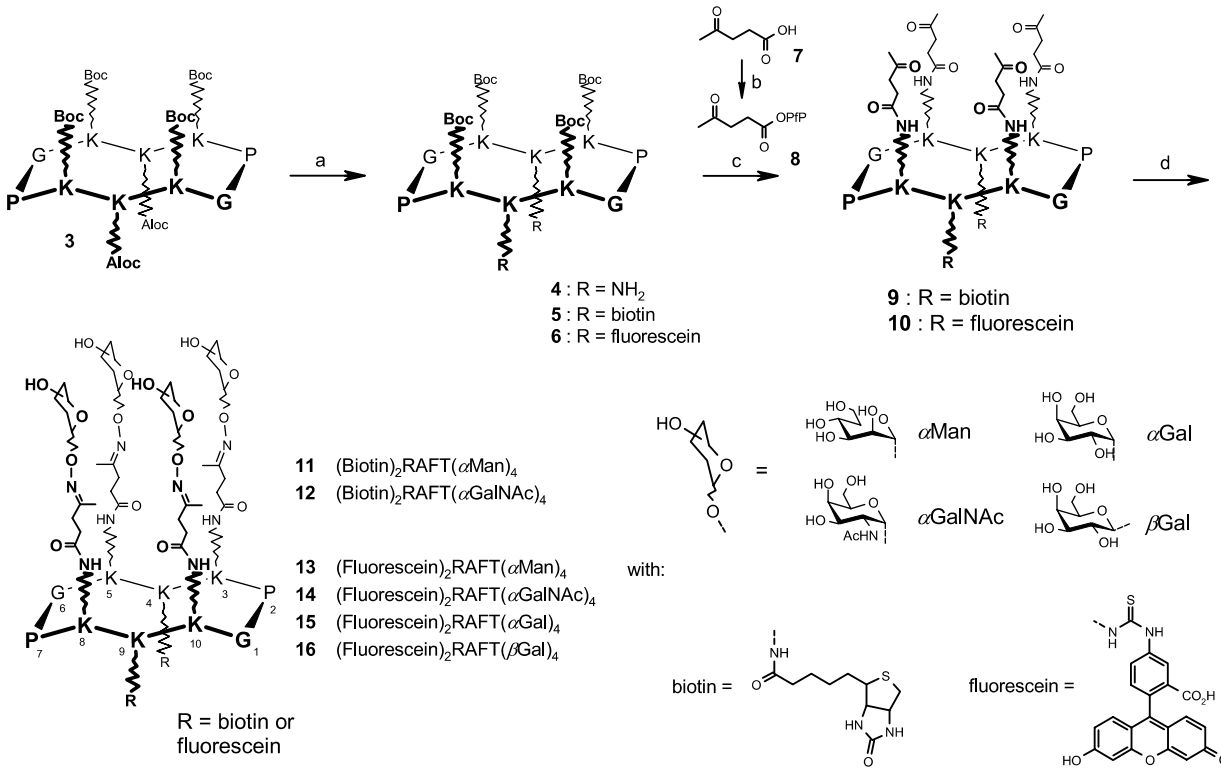
Scheme 1. Reagents and conditions: (a) successive cycles of (i) 20% piperidine/DMF, 3 × 10 min; (ii) Fmoc-aa-OH (4 equiv), PyBOP (4 equiv), DIEA (8 equiv), DMF, 30 min, 87%; (b) PyBOP, DIEA, DMF, 86%.

At this stage, we introduce two different detection agents (Scheme 2). Biotin is first chosen, as it can easily be used for binding studies with a target protein as an immobilization agent on avidin/streptavidin-coated surfaces, as well as to detect the formation of a ligand/receptor complex with streptavidin coupled with peroxidase or alkaline phosphatase.[7a,b] The coupling reaction between **4** and biotin is thus performed with PyBOP and DIEA in DMF and the progress of the reaction is monitored by an RP-HPLC analysis. The reaction cleanly proceeds in 1 h and gives a white powder corresponding to the bis-biotinylated RAFT molecule **5** after precipitation from a mixture of dichloromethane and diethyl ether (87% yield).

As a complementary alternative, we incorporate a fluorescent agent to dispose of detectable molecules suitable for carbohydrate microarrays and for recognition assays in solution using fluorescence technologies.[16] Indeed, previous studies have shown that spatial proximity of two fluorescent probes pointing to the lower face of a template has no effect on the fluorescence of the molecule since no self-quenching was clearly observed.[9] For this purpose, we introduce fluorescein at both free amino positions on the detection domain of the template **4** by treatment of fluorescein isothiocyanate (FITC) in DMF under basic conditions. After 1 h, the conversion is complete and the fluorescent RAFT compound **6** was obtained in 75% yield.

The subsequent incorporation of carbohydrate recognition motifs is realized, following an oxime-based strategy from aminooxylated modified sugar derivatives that were prepared according to the protocol that we have described previously.[17] Moreover, to prevent any risk of side reactions during the oxidative cleavage commonly required to generate aldehyde functions from serine,[18] we decide to employ levulinic acid **7** as a more direct route to introduce the ketone functions needed for the oxime ligation with aminooxy sugars, providing at the same time a suitable spacer between the anchoring site and the template.

The Boc-protecting groups of biotinylated and fluorescent products **5** and **6** are first removed with a solution of 50% TFA in dichloromethane to quantitatively



Scheme 2. Reagents and conditions: (a) (i) $PhSiH_3$, $Pd(Ph_3P)_4$, DMF, 79%; (ii) biotin, PyBOP, DIEA, DMF, 87% or FITC, DIEA, DMF, 75%; (b) pentafluorophenol, DCC, CH_2Cl_2, 99%; (c) (i) TFA/CH_2Cl_2 (1/1), 95%; (ii) **8**, DIEA, DMF, 70%; (d) Aminooxy sugar (Man-α-ONH_2, GalNAc-α-ONH_2, Gal-α-ONH_2 or Gal-β-ONH_2), AcONa buffer 0.1 M pH 4.0, 40–60% after semi-preparative RP-HPLC purification.

obtain RAFT molecules presenting four free side chains at positions 3, 5, 8, and 10 of the recognition domains, which are used without further purification. The levulinic acid **7** is first activated as the pentafluorophenol ester **8** using a standard procedure by treatment with dicyclocarbodiimide (DCC) in dichloromethane. The incorporation of four copies of **8** into the deprotected RAFTs **5** and **6** takes place under basic conditions in DMF to get **9** and **10** with 70% yield after semi-preparative RP-HPLC purification.

The chemoselective functionalization of the recognition domain takes place with aminooxylated α-mannose, α-*N*-acetylgalactose (mucin-related T_N-antigen), and α/β-galactose[17], which are relevant for diagnostic or therapeutic applications. The oxime ligation occurred in sodium acetate buffer, pH 4.0, at room temperature and the progress of the reaction is monitored by RP-HPLC analysis. Whereas the conversion is found to be incomplete after 2 days compared to our previous studies,[10] six new tetravalent RAFT-ligated glycoconjugates **11**–**16** are yet to be obtained after HPLC purification in satisfactory isolated yield (≈50%). These glycosylated compounds are characterized by electrospray ionization mass spectrometry (entries 1–6, Table 1). Additionally, two-dimensional TOCSY and ROESY NMR experiments performed on product **12**[19] permit the assignment of each proton and also confirm the sequence of the peptidic backbone.

Table 1. Analytical data of glycopeptides **11**–**16**

Compound	R_t[a]	M_{calcd}	m/z[b]
11	18.9	2631.0	877.9 $[M+3H]^{3+}$
12	18.9	2795.2	932.3 $[M+3H]^{3+}$
13	20.9	2957.2	986.1 $[M+3H]^{3+}$
14	20.4	3121.4	1041.0 $[M+3H]^{3+}$
15	19.1	2957.2	986.5 $[M+3H]^{3+}$

[a] Reverse-phase HPLC retention time is given in minutes: linear gradient 95:5 A:B to 40:60 A:B in 30 min (Column: nucleosil 100 Å 5 μm C_{18} particles, 250 × 4.6 mm; solvent B: 0.09% TFA in 90% acetonitrile and solvent A: 0.09% TFA; flow rate: 1 mL/min; detection: $\lambda = 214$ nm for biotinylated conjugates, $\lambda = 214$ and 499 nm for fluorescent conjugates).

[b] Mass spectrometry analysis was performed by electrospray ionization

In summary, we report, in this paper, the preparation of multitopic neoglycopeptides presenting clustered carbohydrates as recognition motifs and molecular probes for biological characterizations with various carbohydrate binding proteins. These molecules displaying both recognition and detection properties on two separated addressable domains represent attractive tools to investigate the functional and structural features of the multivalent interactions between proteins and oligosaccharides. In addition, the easy chemical access of such neoglycoconjugates using oxime ligation strategy combined with the development of a fully solid-phase protocol for synthesis and screening

with target proteins might enable the rapid discovery of new active and selective carbohydrate-based therapeutics. This is currently explored in the laboratory and will be reported in due course.

Acknowledgments

This work was supported by the Association pour la Recherche contre le Cancer (ARC), the Centre National pour la Recherche Scientifique (CNRS) for specific action (AIP), and the Institut Universitaire de France (IUF). We thank Prof. G. Esposito from the University of Udinese, Italy, and Dr. C. Fontaine for NMR studies of compound **12**.

References and notes

1. Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2754.
2. For selected review about peptide and protein recognition by designed molecules, see: Peczuh, M. W.; Hamilton, A. D. *Chem. Rev.* **2000**, *100*, 2479.
3. (a) Lee, Y. C.; Lee, R. T. *Acc. Chem. Res.* **1995**, *28*, 321; (b) Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555; (c) Gestwicki, J. E.; Cairo, C. W.; Strong, L. E.; Oetjen, K. A.; Kiessling, L. L. *J. Am. Chem. Soc.* **2002**, *124*, 14922.
4. (a) Sears, P.; Wong, C.-H. *Angew. Chem., Int. Ed.* **1999**, *38*, 2300; (b) Yarema, K. J.; Bertozzi, C. R. *Curr. Opin. Chem. Biol.* **1998**, *2*, 49; (c) Danishefsky, S. J.; Allen, J. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 836; (d) Toyokuni, T.; Singhal, A. K. *Chem. Soc. Rev.* **1995**, 231; (e) Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357.
5. (a) For selected publications, see: Fulton, D. A.; Stoddart, J. F. *Bioconjugate Chem.* **2001**, *12*, 655; (b) Roy, R. *Curr. Opin. Struct. Biol.* **1996**, *6*, 692; (c) Sprengard, U.; Schudok, M.; Kretzschmar, G.; Kunz, H. *Angew. Chem., Int. Ed.* **1996**, *35*, 321; (d) Wittmann, V.; Seeberger, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 4348; (e) Veprek, P.; Jezek, J. *J. Peptide Sci.* **1999**, *5*, 5.
6. (a) St. Hilaire, P. M.; Meldal, M. *Angew. Chem., Int. Ed.* **2000**, *39*, 1162; (b) Osborn, H. M. I.; Khan, T. H. *Tetrahedron* **1999**, *55*, 1807.
7. (a) Love, K. R.; Seeberger, P. H. *Angew. Chem., Int. Ed.* **2002**, *41*, 3583; (b) Ortiz Mellet, C.; García Fernandez, J. M. *ChemBioChem* **2002**, *3*, 819; (c) Hirabayashi, J. *Trends Biotechnol.* **2003**, *21*, 141; (d) Jelinek, R.; Kolusheva, S. *Chem. Rev.* **2004**, *104*, 5987; (e) Shin, I.; Park, S.; Lee, M. *Chem. Eur. J.* **2005**, *11*, 2894.
8. (a) For selected reviews, see: Herzner, H.; Reipen, T.; Schultz, M.; Kunz, H. *Chem. Rev.* **2000**, *100*, 4495; (b) Seitz, O. *ChemBioChem* **2000**, *1*, 214; (c) Davis, B. G. *Chem. Rev.* **2002**, *102*, 579.
9. Boturyn, D.; Coll, J.-L.; Garanger, E.; Favrot, M.-C.; Dumy, P. *J. Am. Chem. Soc.* **2004**, *126*, 5730.
10. (a) Renaudet, O.; Dumy, P. *Org. Lett.* **2003**, *5*, 243; (b) Renaudet, O.; Dumy, P. *Tetrahedron Lett.* **2004**, *45*, 65.
11. Peluso, S.; Dumy, P.; Nkubana, C.; Yokokawa, Y.; Mutter, M. *J. Org. Chem.* **1999**, *64*, 7114, and references cited therein.
12. (a) Scheibler, L.; Dumy, P.; Boncheva, M.; Leufgen, K.; Mathieu, H.-J.; Mutter, M.; Vogel, H. *Angew. Chem., Int. Ed.* **1999**, *38*, 696; (b) Scheibler, L.; Dumy, P.; Stamou, D.; Duschl, C.; Vogel, H.; Mutter, M. *Tetrahedron* **1998**, *54*, 3725.
13. (a) Dumy, P.; Eggleston, I. M.; Esposito, G.; Nicula, S.; Mutter, M. *Biopolymers* **1996**, *39*, 297; (b) Peluso, S.; Rückle, T.; Lehmann, C.; Mutter, M.; Peggion, C.; Crisma, M. *ChemBioChem* **2001**, *2*, 432.
14. Coste, J.; Le Nguyen, D.; Castro, B. *Tetrahedron Lett.* **1990**, *31*, 205.
15. (a) Thieriet, N.; Alsina, J.; Girald, E.; Guibé, F.; Albericio, F. *Tetrahedron Lett.* **1997**, *38*, 7275; (b) Dangles, O.; Guibé, F.; Balavoine, G.; Lavielle, S.; Marquet, A. *J. Am. Chem. Soc.* **1987**, *52*, 4984.
16. For a recent review about fluorescence technologies for analysis of protein interactions, see: Yan, Y.; Marriott, G. *Curr. Opin. Chem. Biol.* **2003**, *7*, 1.
17. Renaudet, O.; Dumy, P. *Tetrahedron Lett.* **2001**, *42*, 7575.
18. Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem.* **1992**, *3*, 138.
19. ^{1}H NMR data of glycopeptide **12** (500 MHz, H_2O/D_2O, 303 K): δ ppm 8.60 (2H, NH-$G^{1,6}$), 8.43 (2H, NH-$K^{4,9}$), 8.16 (2H, 2× NHAc GalNAc), 8.10 (2H, 2× NHAc GalNAc), 8.07 (2H, NHε-$K^{5,10}$), 7.98 (2H, NHε-$K^{3,8}$), 7.94 (2H, NHε-$K^{4,9}$), 7.93 (2H, NH-$K^{5,10}$), 7.82 (2H, NH-$K^{3,8}$), 6.47 (2H, 2× NH biotin), 6.41 (2H, 2× NH biotin), 5.46 (2H, 2× H-1′ GalNAc), 5.44 (2H, 2× H-1′ GalNAc), 4.73 (2H, Hα-$K^{3,8}$), 4.65 (2H, 2× H biotin), 4.47 (2H, 2× H biotin), 4.44 (2H, Hα-$K^{10,5}$), 4.41 (2H, Hα-$P^{2,7}$), 4.36 (2H, Hα-$K^{4,9}$), 4.35 (2H, 2× H-2′ GalNAc), 4.32 (2H, 2× H-2′ GalNAc), 4.16 (2H, Hα-$G^{1,6}$), 4.08 (4H, 4× H-4′ GalNAc), 4.06 (2H, 2× H-3′ GalNAc), 4.05 (2H, 2× H-3′ GalNAc), 3.95 (4H, 4× H-5′ GalNAc), 3.85 (2H, Hα′-$G^{1,6}$), 3.84 (2H, Hδ-$P^{2,7}$), 3.76 (8H, 8× H-6′ GalNAc), 3.69 (2H, Hδ′-$P^{2,7}$), 3.37 (2H, 2× H biotin), 3.24 (2H, Hε-$K^{5,10}$), 3.23 (4H, Hε-$K^{4,9}$), 3.20 (4H, Hε-$K^{3,8}$), 3.18 (2H, Hε′-$K^{5,10}$), 3.04 (2H, 2×H biotin), 2.87 (4H, 2×CH_2 levul.-$K^{3,8}$), 2.83 (2H, 2× H biotin), 2.82 (2H, CH_2 levul.-$K^{5,10}$), 2.71 (2H, CH_2 levul.-$K^{5,10}$), 2.53 (2H, CH_2 levul.-$K^{3,8}$), 2.51 (4H, CH_2 levul.-$K^{5,10}$), 2.48 (CH_2 levul.-$K^{3,8}$), 2.35 (2H, Hβ-$P^{2,7}$), 2.29 (4H, 4× H biotin), 2.14 (2H, Hγ-$P^{2,7}$), 2.10 (6H, 2× NHAc GalNAc), 2.07 (6H, 2× NHAc GalNAc), 2.05 (2H, Hγ′-$P^{2,7}$), 2.00 (6H, N=C–Me), 1.98 (2H, Hβ′-$P^{2,7}$), 1.94 (6H, N=C–Me), 1.88 (4H, Hβ-$K^{4,9}$ and Hγ-$K^{4,9}$), 1.82 (4H, Hβ-$K^{3,8}$, Hγ-$K^{3,8}$), 1.81 (4H, Hβ-$K^{5,10}$ and Hγ-$K^{5,10}$), 1.80 (4H, Hβ′-$K^{4,9}$ and Hγ′-$K^{4,9}$), 1.76 (2H, 2× H biotin), 1.67 (2H, 2× H biotin), 1.66 (4H, Hβ′-$K^{3,8}$ and Hγ′-$K^{3,8}$), 1.65 (2H, 2× H biotin), 1.63 (2H, 2× H biotin), 1.57 (4H, Hβ′-$K^{5,10}$ and Hγ′-$K^{5,10}$), 1.55 (2H, Hδ-$K^{4,9}$), 1.53 (2H, Hδ-$K^{3,8}$), 1.52 (2H, Hδ-$K^{5,10}$), 1.44 (2H, 2× H biotin), 1.42 (2H, 2× H biotin), 1.38 (2H, Hδ′-$K^{3,8}$), 1.37 (4H, Hδ′-$K^{4,9}$ and Hδ′-$K^{5,10}$).