Supplementary Information

A Bifunctional *Pasteurella multocida* β1–3-Galactosyl/N-Acetylgalactosaminyltransferase (PmNatB) for Highly Efficient Chemoenzymatic Synthesis of Disaccharides

Xiaohong Yang,[a,†] Bijoyananda Mishra,[a,†] Hai Yu,*[a] Yijun Wei,[b,c] and Xi Chen*[a]

[a]Department of Chemistry, [b]Department of Neurobiology, Physiology and Behavior; [c]Department of Statistics, University of California, One Shields Avenue, Davis, California, USA

[†]These authors contributed equally to this work.

*Corresponding authors: hyu@ucdavis.edu and xiichen@ucdavis.edu

Table of Contents

Table S1. Results for high-resolution mass spectrometry (HRMS)-based determination of the multifunctionality of His_6-PmNatB.[a]

Potential Acceptors	Potential Donors			
	UDP-Gal/UDP-Glc		UDP-GalNAc/UDP-GlcNAc	
	Expected *m*/*z*	Observed	Expected *m*/*z*	Observed
Lactose	505.1762 (+H) 527.1582 (+Na) 503.1617 (-H) 539.1384 (+Cl)	No	546.2028 (+H) 568.1848 (+Na) 544.1883 (-H) 580.1650 (+Cl)	No
GalNAcαProNHCbz GalNAcβProNHCbz GlcNAcαProNHCbz GlcNAcβProNHCbz	575.2776 (+H) 597.2266 (+Na) 573.2301 (-H) 609.2068 (+Cl)	**Yes (609.2068)**	616.2711 (+H) 638.2531 (+Na) 614.2566 (-H) 650.2333 (+Cl)	No
GlcNAcαProCl	460.1579 (+H) 482.1399 (+Na) 458.1434 (-H) 494.1201 (+Cl)	No	501.1845 (+H) 523.1665 (+Na) 499.1700 (-H) 535.1467 (+Cl)	No
Galα*p*NP Galβ*p*NP Glcα*p*NP	464.1398 (+H) 486.1218 (+Na) 462.1253 (-H) 498.1020 (+Cl)	No	505.1663 (+H) 527.1483 (+Na) 503.1518 (-H) 539.1285 (+Cl)	**Yes (539.1269)**
GlcAβ*p*NP	476.1046 (-H)	No	517.1311 (-H)	No

[a] Reactions were carried out at 30 °C with a mixture of four UDP-sugars as potential donors (5 mM each) and ten glycosides as potential acceptors (1 mM each) for 0.5 h and 20 h, respectively. Same outcomes were obtained at both reaction time points.

Table S2. ^{13}C NMR chemical shifts of **1a/b**–**4a/b**.

Residue	Carbon atom	Chemical shift (ppm)	
	C	GalNAcαProNHCbz (**1a**)	Galβ3GalNAcαProNHCbz (**3a**)
α-D-GalNAc	1	96.9	97.1
	2	49.9	48.6
	3	**68.5**	**77.4**
	4	67.7	68.8
	5	70.8	70.6
	6	61.2	61.2
	C=O	174.5	174.5
	CH_3	21.9	22.0
β-D-Gal (1-3)	1		104.7
	2		70.6
	3		72.5
	4		68.5
	5		74.9
	6		60.9
ProNHCbz	$O\underline{C}H_2CH_2CH_2NHCbz$	65.0	65.0
	$OCH_2\underline{C}H_2CH_2NHCbz$	28.4	28.5
	$OCH_2CH_2\underline{C}H_2NHCbz$	37.5	37.5
	$Ph\underline{C}H_2O$-	66.7	66.7
	C	GalNAcβProNHCbz (**1b**)	Galβ3GalNAcβProNHCbz (**3b**)
β-D-GalNAc	1	101.6	101.4
	2	52.4	51.3
	3	**71.0**	**79.9**
	4	67.8	68.0
	5	75.1	75.0
	6	60.9	61.0
	C=O	174.7	174.7
	CH_3	22.1	22.2
β-D-Gal (1-3)	1		104.8
	2		70.6
	3		72.5
	4		68.6
	5		74.7
	6		60.9
ProNHCbz	$O\underline{C}H_2CH_2CH_2NHCbz$	66.8	66.8
	$OCH_2\underline{C}H_2CH_2NHCbz$	28.8	28.8
	$OCH_2CH_2\underline{C}H_2NHCbz$	37.2	37.2
	$Ph\underline{C}H_2O$-	67.5	67.5
	C	GalαProNHCbz (**2a**)	GalNAcβ3GalαProNHCbz (**4a**)
α-D-Gal	1	98.2	98.4
	2	69.2	69.2
	3	**69.5**	**79.1**
	4	68.3	67.7
	5	70.9	70.8
	6	61.2	61.1
β-D-GalNAc(1-3)	1		103.1
	2		52.6
	3		70.4
	4		67.3
	5		74.9
	6		60.9
	C=O		175.2
	CH_3		22.2
ProNHCbz	$O\underline{C}H_2CH_2CH_2NHCbz$	65.4	65.2
	$OCH_2\underline{C}H_2CH_2NHCbz$	28.5	28.5
	$OCH_2CH_2\underline{C}H_2NHCbz$	37.7	37.6
	$Ph\underline{C}H_2O$-	66.8	66.7
	C	GalβProNHCbz (**2b**)	GalNAcβ3GalβProNHCbz (**4b**)
β-D-Gal	1	102.8	102.8
	2	70.7	70.8
	3	**72.8**	**82.0**
	4	68.6	68.4
	5	75.1	75.0
	6	60.9	60.8
β-D-GalNAc(1-3)	1		103.2
	2		52.6
	3		69.9
	4		67.8
	5		74.6
	6		61.0
	C=O		175.2
	CH_3		22.3
ProNHCbz	$O\underline{C}H_2CH_2CH_2NHCbz$	66.8	66.8
	$OCH_2\underline{C}H_2CH_2NHCbz$	28.9	28.8
	$OCH_2CH_2\underline{C}H_2NHCbz$	37.3	37.3
	$Ph\underline{C}H_2O$-	67.6	67.6

Table S3. ^{1}H NMR chemical shifts and coupling constants for anomeric hydrogens of monosaccharides in the disaccharides **3a/b** and **4a/b**.

Disaccharide	**Chemical shift (ppm) (*J* coupling constant Hz)**			
	β-D-Gal	α-D-GalNAc	β-D-GalNAc	α-D-Gal
Galβ3GalNAcαProNHCbz (**3a**)	4.43 (7.7)	4.82 (3.8)		
Galβ3GalNAcβProNHCbz (**3b**)	4.44 (7.9)		4.44 (7.9)	
GalNAcβ3GalαProNHCbz (**4a**)			4.57 (8.4)	4.85 (3.6)
GalNAcβ3GalβProNHCbz (**4b**)	4.64 (7.9)		4.64 (7.9)	

a)

```
PmNatB      MSYLDQPLVSVLICAYNADKYIEECIDAILNQTYKNLEIVVVNDGSTDTTLSKLHYFAEK      60
HiLgtD      --MENCPLVSVIVCAYNAEQYIDESISSIINQTYENLEIIVINDGSTDLTLSHLEEISKL      58
NgLgtD      ----MQPLVSVLICAYNAEKYFAQSLAAVVGQTWRNLDILIVDDGSTDGTPAIARHFQEQ      56
NmLgtA      ----MQPLVSVLICAYNVEKYFAQSLAAVVNQTWRNLDILIVDDGSTDGTLAIAQRFQEQ      56
                  *****::****.::*: :.: :::.**:.**:*:::***** * :  . : :

PmNatB      DPRIKIINNEENKGFIASLNIGISSI-----NGDYLARTDADDITKPEWIEKILGYMLSH      115
HiLgtD      DKRIKIISNKYNLGFINSLNIGLGCF-----SGKYFARMDADDIAKPSWIEKIVTYLEKN      113
NgLgtD      DGRIRIISNPRNLGFIASLNIGLDELAKSG-GGEYIARTDADDIASPGWIEKIVGEMEKD      115
NmLgtA      DGRIRILAQPRNSGLIPSLNIGLDELAKSGGGGEYIARTDADDIAAPDWIEKIVGEMEKD      116
            * **:*: :  * *:* *****:. :     .*.*:** *****: * *****:  : ..

PmNatB      PQIIAMGSYLTILSEDGNGSNLANYYEHGDEWRNPLSHREIVEAMLFRNPIHNNSMIVKS      175
HiLgtD      DHITAMGSYLEIIVEKECGI-IGSQYKTGDIWKNPLLHNDICEAMLFYNPIHNNTMIMRA      172
NgLgtD      RSIIAMGAWLEVLSEENNKSVLAAIARNGAIWDKPTRHEDIVAVFPFGNPIHNNTMIMRR      175
NmLgtA      RSIIAMGAWLEVLSEEKDGNRLARHHEHGKIWKKPTRHEDIADFFPFGNPIHNNTMIMRR      176
              * ***:* :: *.     :.   . *  * :*  *.:*   : * ******:**::

PmNatB      TVFREHGLRFDPAYQHTEDYQFWLEVSRLGELANYPESLVYYRLHNTQTSSLHNKYQNLM      235
HiLgtD      NVYREHKLIFNKDYPYAEDYKFWSEVSRLGCLANYPEALVKYRLHGNQTSSVYNHEQNET      232
NgLgtD      SVI-DGGLRFDPAYIHAEDYKFWYEAGKLGRLAYYPEALVKYRFHQDQTSSKYNLQQRRT      234
NmLgtA      SVI-DGGLRYNTERDWAEDYQFWYDVSKLGRLAYYPEALVKYRLHANQVSSKYSIRQHEI      235
            .*  :  * ::     :***:** :..:** ** ***:** **:*  *.** :.  *.

PmNatB      AKKIRKRAINYYLQDLGIIHRLGEDIFFHDIEKIQAEL--ASLSLLDNCIIKRILYDCYL      293
HiLgtD      AKKIKRENITYYLNKIGIDIKVINSVSLLEIYHVD----------KSNKVLKSILYEMYM      282
NgLgtD      AWKIKEEIRAGYWKAAGIAVGADCL-NYGLLKSTAYALYEKALSGQDIGCLRLFLYEYFL      293
NmLgtA      AQGIQKTARNDFLQSMGFKTRFDSL-EYRQIKAVAYELLEKHLPEEDFELARRFLYQCFK      294
            *  *:.     : :  *:           :               .    : :**: :

PmNatB      SLVDNKLINILYFLRDKN-NSYFNRKQKIKIIKRIIRPYKYESVL       337
HiLgtD      SLDKYTITSLLHFIKY-H-LELFDLKQNLKIIKKFIRKINVIF--       323
NgLgtD      SLEKYSLTDLLDFLTDRVMRKLFAAPQYRKILKKMLRPWKYRSY-       337
NmLgtA      RTDTLPAGAWLDFAADGRMRRLFTLRQYFGILHRLLKNR------       333
                      * *         *   *   *::::::
```

b)

% identity/similarity	PmNatB	HiLgtD	NgLgtD
PmNatB			
HiLgtD	53/69		
NgLgtD	47/63	46/62	
NmLgtA	47/65	42/62	66/74

Figure S1. Amino acid sequence alignment (**a**) and sequence identity and similarity (**b**) of PmNatB and several other CAZy GT2 family enzymes including HiLgtD (GenBank accession number AAC23227.1), NgLgtD (GenBank accession number AAA68012.1), and NmLgtA (GenBank accession number AAC44084.1).

```
   1 M  G  S  S  H  H  H  H  H  H  S  S  G  L  V  P  R  G  S  H
   1 ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCAT
  21 M  L  E  M  S  Y  L  D  Q  P  L  V  S  V  L  I  C  A  Y  N
  61 ATGCTCGAGATGAGCTATTTGGATCAACCATTAGTTTCGGTATTAATCTGTGCTTATAAT
  41 A  D  K  Y  I  E  E  C  I  D  A  I  L  N  Q  T  Y  K  N  L
 121 GCAGATAAATACATAGAAGAGTGTATTGACGCGATTTTAAATCAAACATATAAAAATCTA
  61 E  I  V  V  V  N  D  G  S  T  D  T  T  L  S  K  L  H  Y  F
 181 GAAATAGTTGTTGTGAATGATGGTTCAACAGACACTACTTTGTCAAAGCTTCATTATTTT
  81 A  E  K  D  P  R  I  K  I  I  N  N  E  E  N  K  G  F  I  A
 241 GCTGAAAAAGATCCTAGAATTAAAATTATTAATAATGAAGAAAATAAGGGTTTCATTGCT
 101 S  L  N  I  G  I  S  S  I  N  G  D  Y  L  A  R  T  D  A  D
 301 TCGCTAAATATAGGGATTTCCTCTATCAATGGTGATTATTTAGCGCGAACAGATGCTGAT
 121 D  I  T  K  P  E  W  I  E  K  I  L  G  Y  M  L  S  H  P  Q
 361 GATATTACGAAACCTGAATGGATTGAAAAAATTTTAGGATATATGTTGTCTCATCCCCAA
 141 I  I  A  M  G  S  Y  L  T  I  L  S  E  D  G  N  G  S  N  L
 421 ATTATTGCAATGGGATCGTATCTAACTATTTTGTCAGAAGATGGGAATGGGAGTAATTTA
 161 A  N  Y  Y  E  H  G  D  E  W  R  N  P  L  S  H  R  E  I  V
 481 GCTAATTATTATGAACATGGTGACGAGTGGAGAAATCCTTTAAGTCATAGAGAGATTGTT
 181 E  A  M  L  F  R  N  P  I  H  N  N  S  M  I  V  K  S  T  V
 541 GAGGCAATGTTATTCCGTAATCCTATTCATAATAACTCGATGATTGTAAAAAGTACTGTC
 201 F  R  E  H  G  L  R  F  D  P  A  Y  Q  H  T  E  D  Y  Q  F
 601 TTTAGAGAGCATGGATTACGCTTTGATCCTGCTTATCAACATACTGAAGATTACCAATTT
 221 W  L  E  V  S  R  L  G  E  L  A  N  Y  P  E  S  L  V  Y  Y
 661 TGGTTAGAAGTGAGCCGTTTGGGAGAATTGGCAAATTATCCTGAATCTTTAGTTTATTAT
 241 R  L  H  N  T  Q  T  S  S  L  H  N  K  Y  Q  N  L  M  A  K
 721 CGTCTACACAATACACAAACTTCTTCTTTACATAATAAATATCAAAATCTCATGGCAAAG
 261 K  I  R  K  R  A  I  N  Y  Y  L  Q  D  L  G  I  I  H  R  L
 781 AAAATTAGAAAAAGAGCGATCAATTATTATTTGCAAGATTTGGGTATTATTCATAGGCTA
 281 G  E  D  I  F  F  H  D  I  E  K  I  Q  A  E  L  A  S  L  S
 841 GGTGAAGATATATTTTTCCATGATATTGAAAAGATTCAGGCAGAACTGGCTAGCTTATCA
 301 L  L  D  N  C  I  I  K  R  I  L  Y  D  C  Y  L  S  L  V  D
 901 CTTTTAGATAATTGTATTATTAAAAGAATACTGTATGATTGTTATTTATCATTAGTGGAT
 321 N  K  L  I  N  I  L  Y  F  L  R  D  K  N  N  S  Y  F  N  R
 961 AATAAATTAATAAATATTCTTTATTTTTTGAGAGATAAAAACAATTCATATTTCAATAGA
 341 K  Q  K  I  K  I  I  K  R  I  I  R  P  Y  K  Y  E  S  V  L
1021 AAACAGAAAATAAAGATTATCAAAAGAATTATCCGTCCTTATAAATATGAGTCTGTATTG
 361 *
1081 TGA
```

Figure S2. DNA and amino acid sequences of His_6-PmNatB. The sequences shown in italics and underlined are from the vector plasmid. Key catalytic sites are shown in bold and underlined.

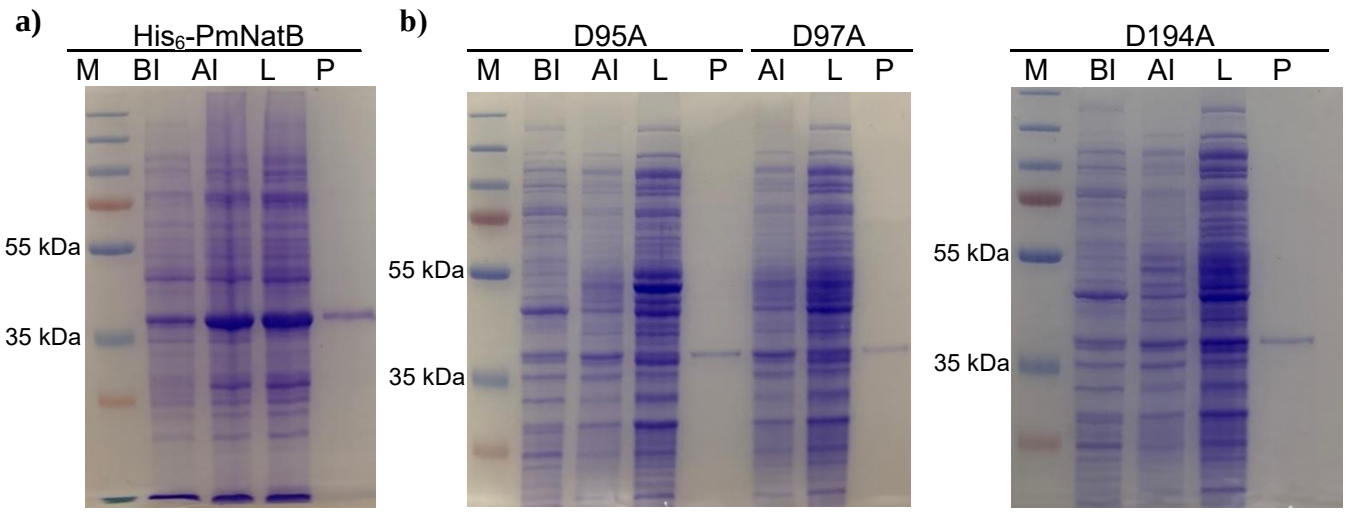


Figure S3. SDS-PAGE analyses for expression and purification of His_6-PmNatB (**a**) and mutants D95A, D97A, and D194A (**b**). Lanes: M, Thermo ScientificTM PageRulerTM Plus Prestained Protein Ladder (10–250 kDa); BI, whole cell extract before induction; AI, whole cell extract after induction; L, lysate after induction; P, Ni^{2+}-NTA column purified protein.

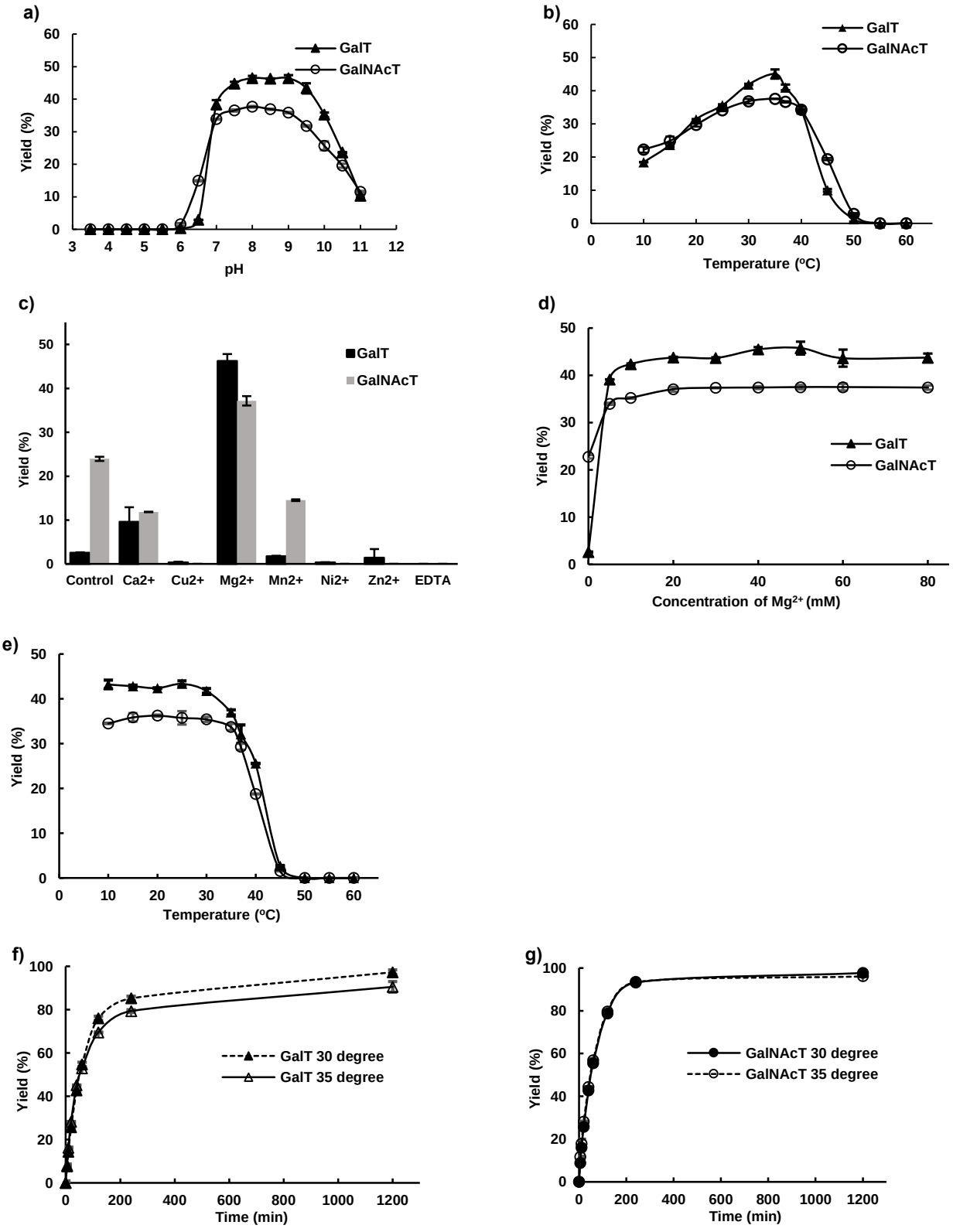
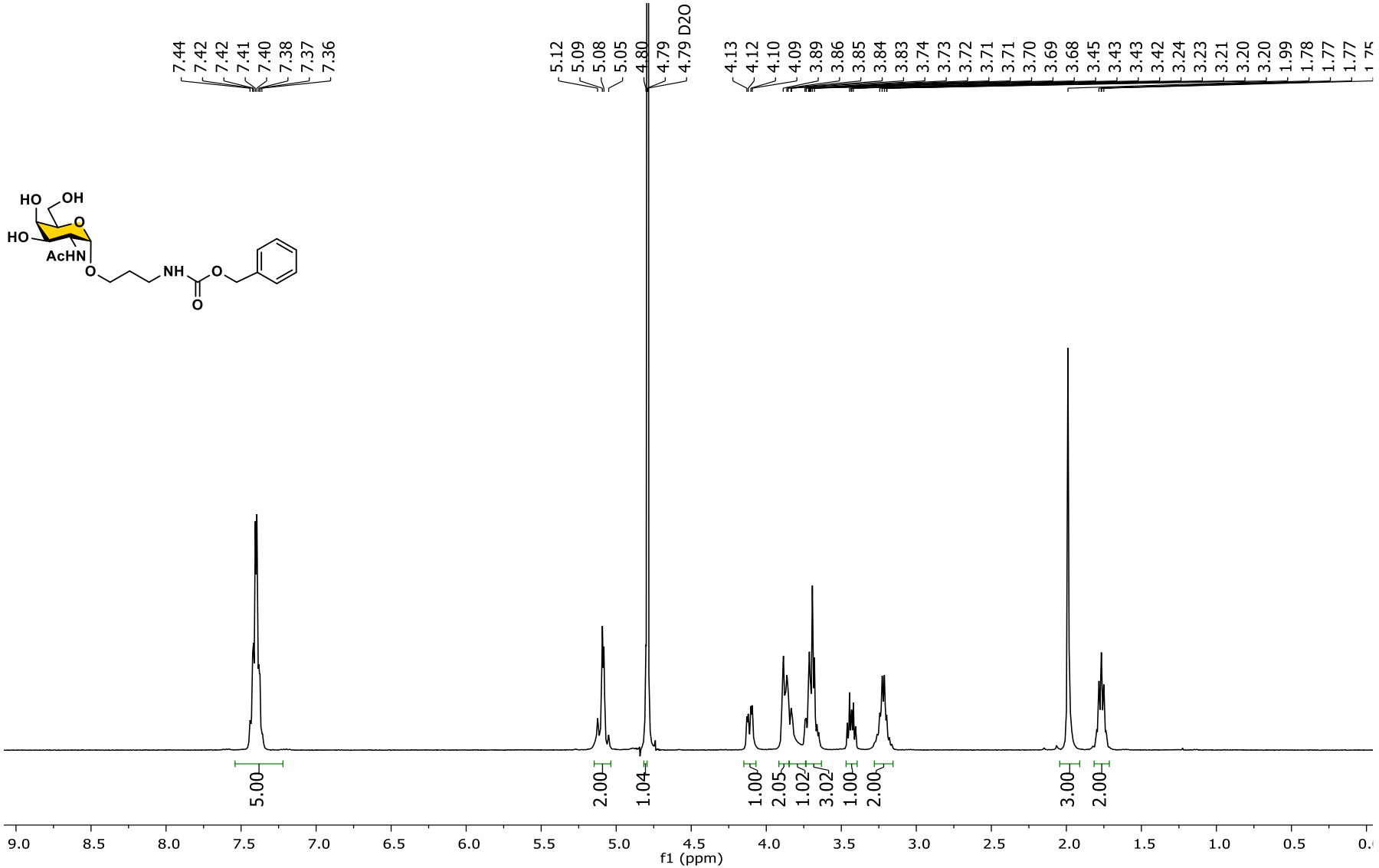
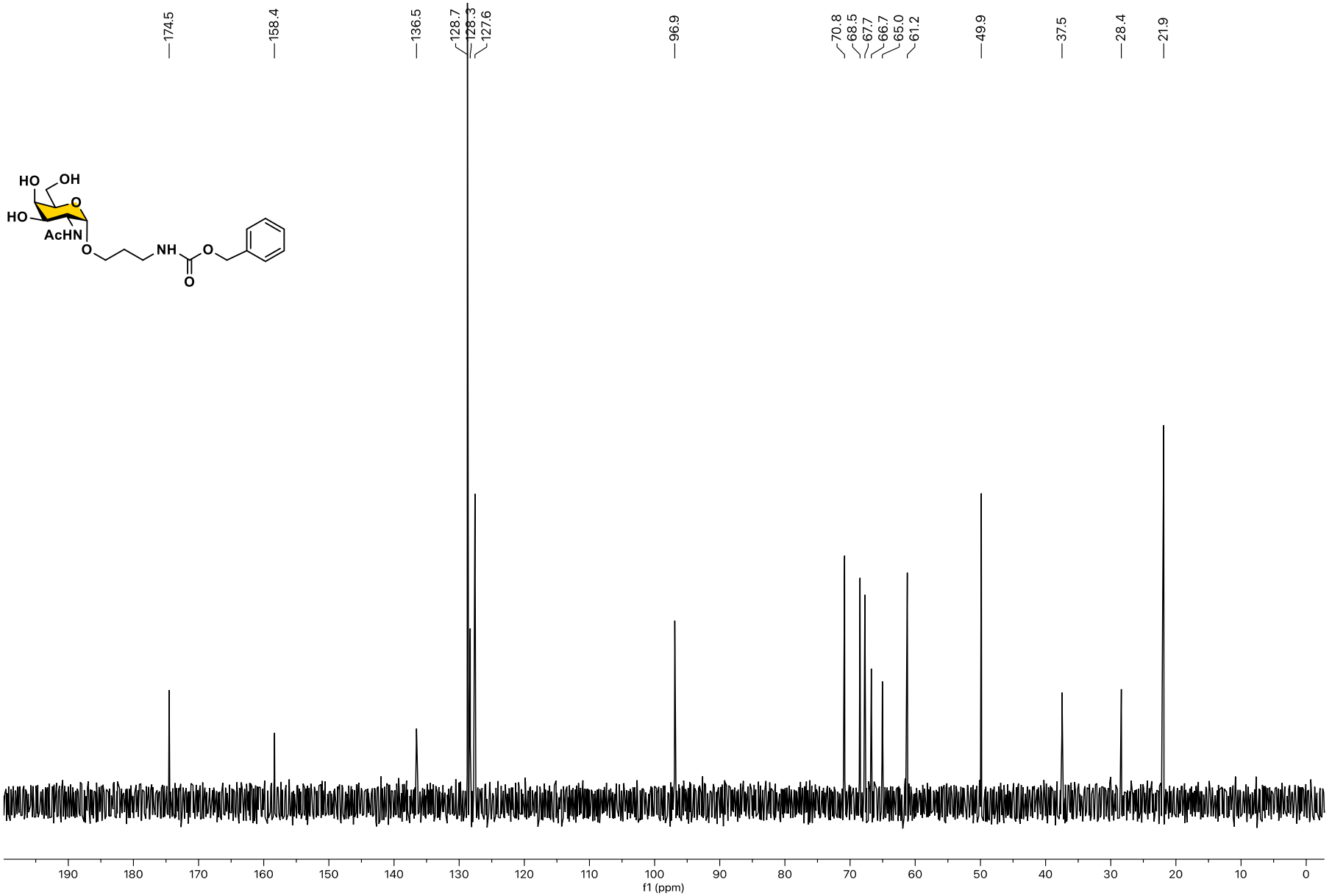


Figure S4. Biochemical characterization of His_6-PmNatB including pH profile (**a**); temperature profile (**b**); effects of divalent metal cations and EDTA (**c**); Mg^{2+} concentration effect (**d**); thermostability (**e**); and time course studies of GalT (**f**) and GalNAcT (**g**) activities at 30 °C and 35 °C.

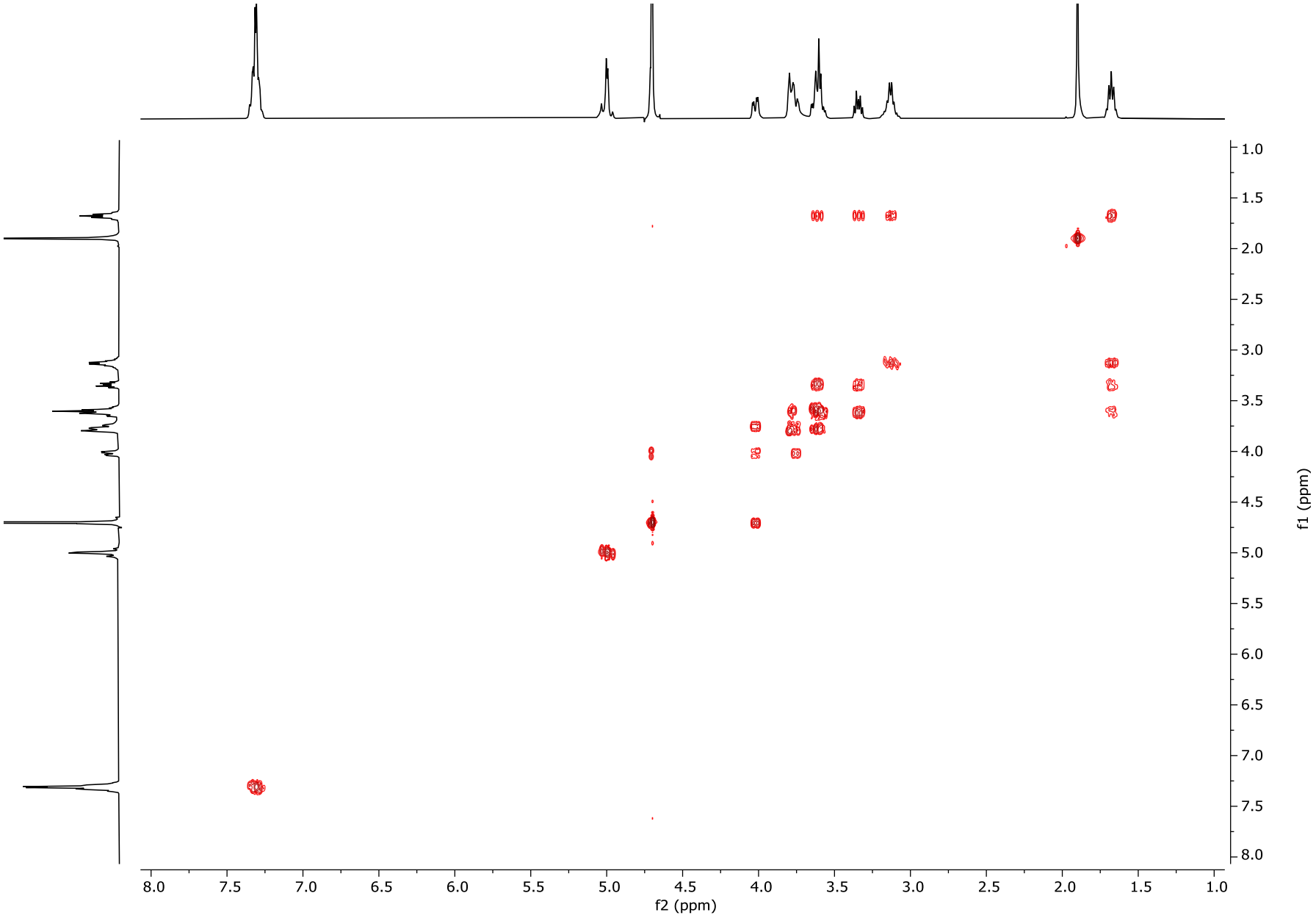
^{1}H NMR Spectra (400 MHz, D_2O) of GalNAcαProNHCbz (**1a**)



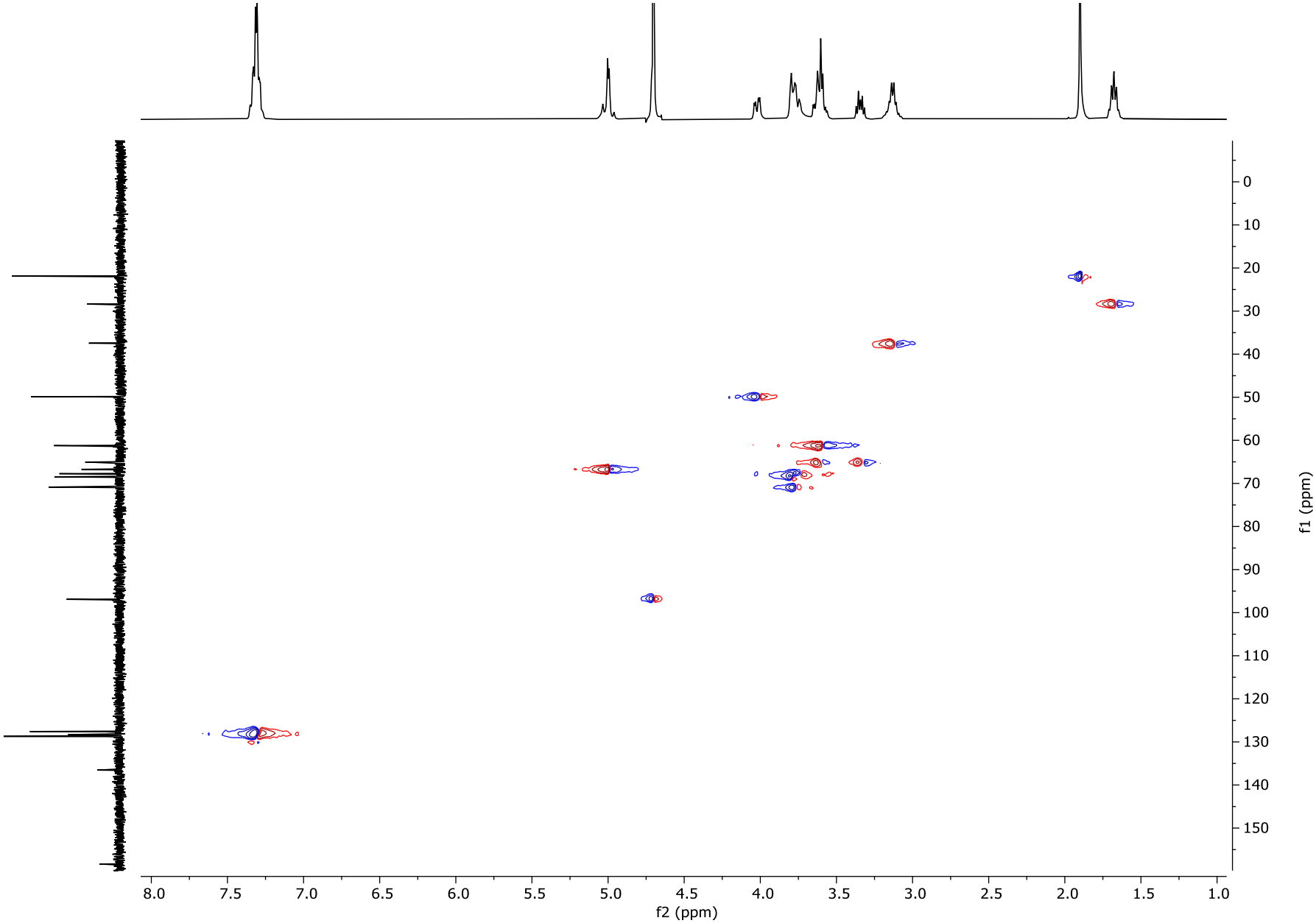
^{13}C NMR Spectra (400 MHz, D_2O) of GalNAcαProNHCbz (**1a**)



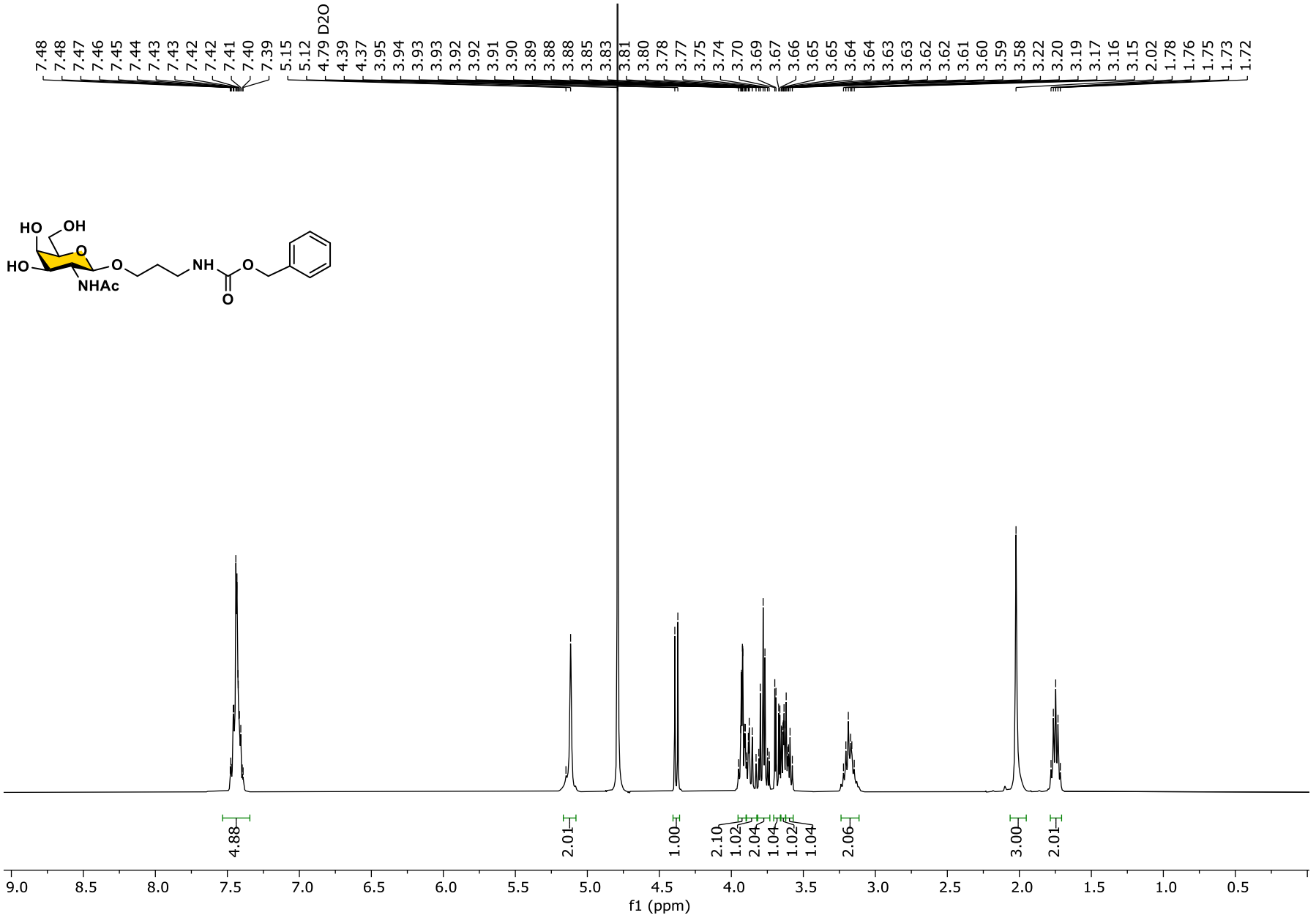
$^{1}H–^{1}H$ COSY NMR Spectra (400 MHz, D_2O) of GalNAcαProNHCbz (**1a**)



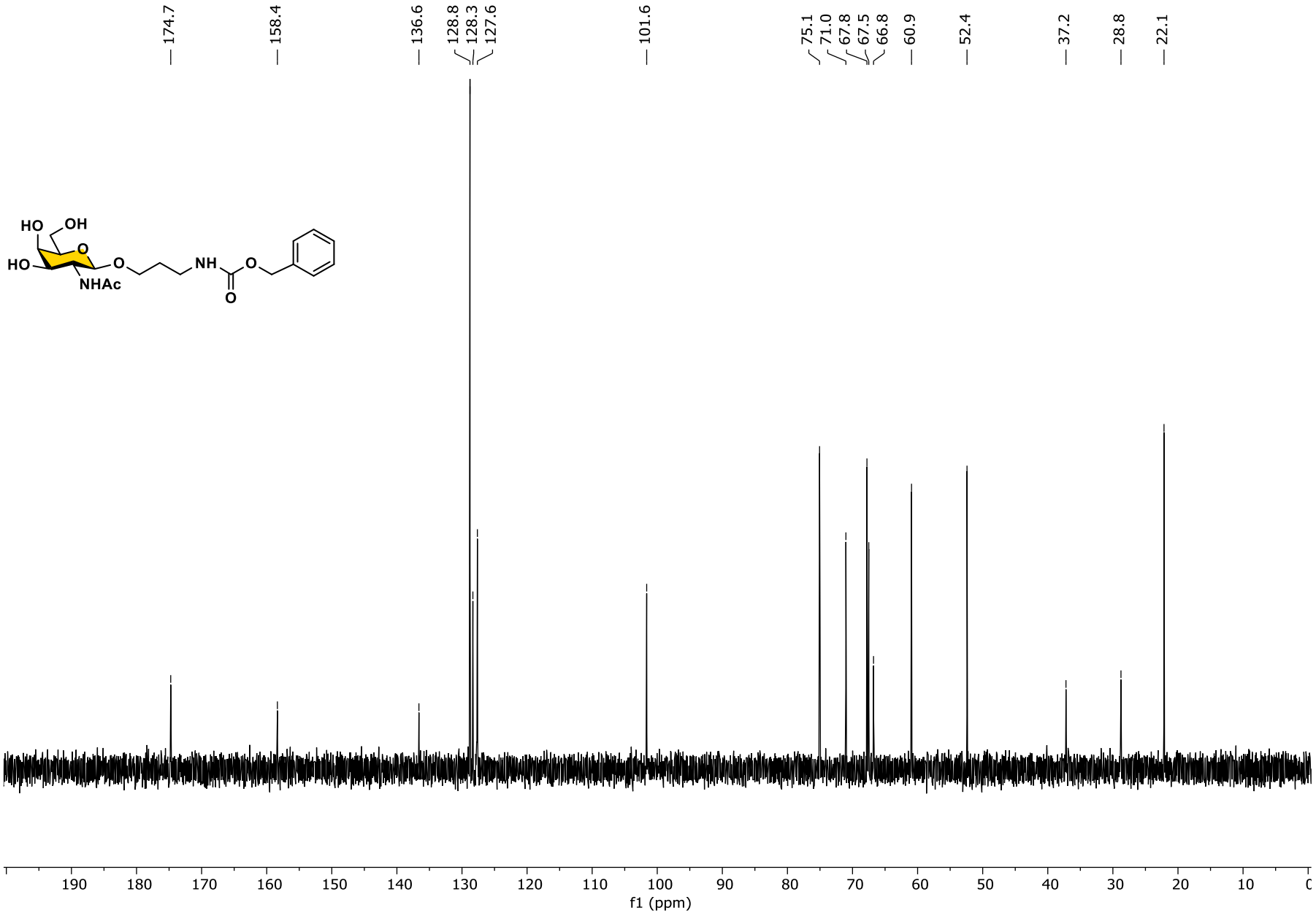
$^{1}H–^{13}C$ HSQC NMR Spectra (400 MHz, D_2O) of GalNAcαProNHCbz (**1a**)



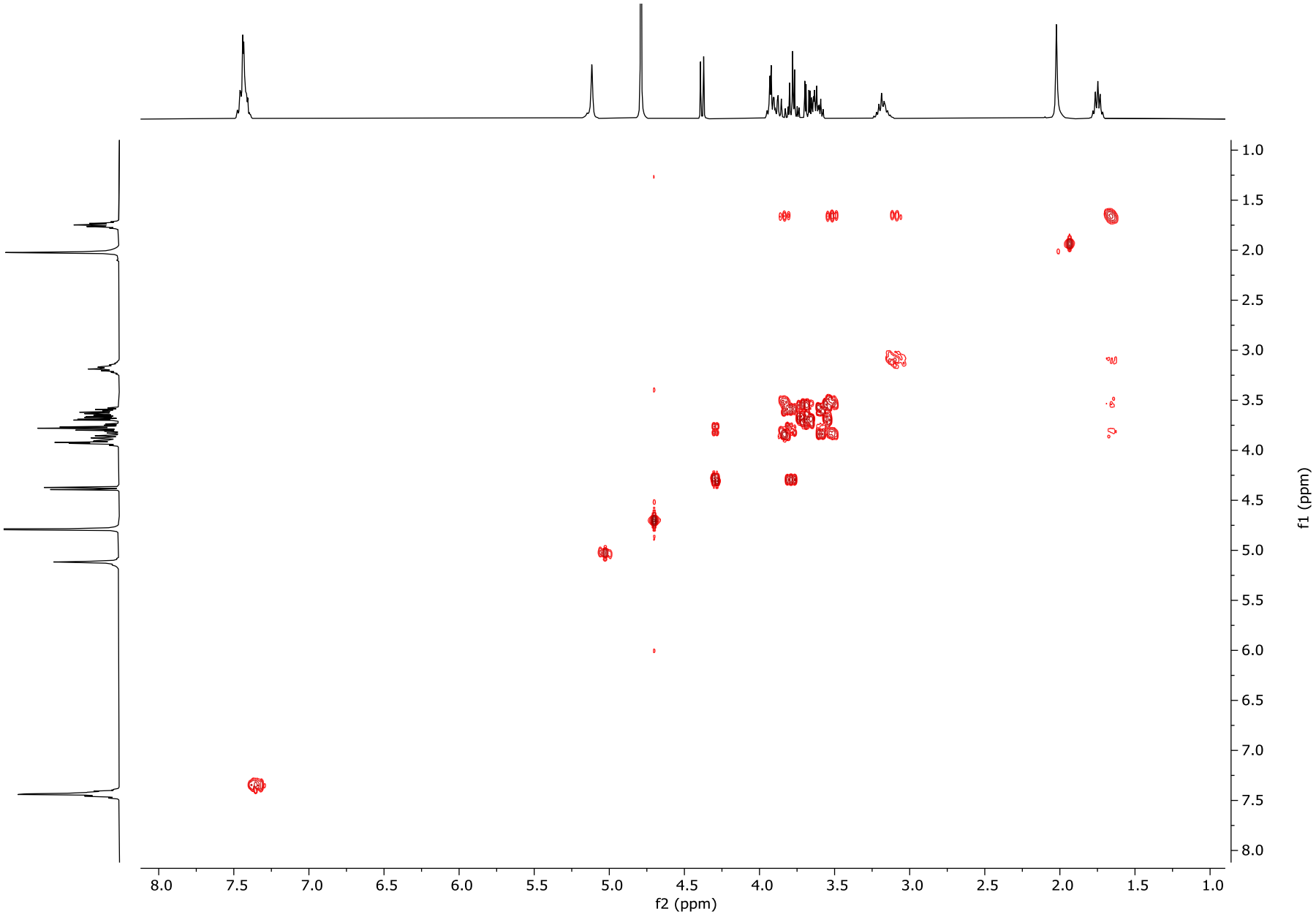
^{1}H NMR Spectra (400 MHz, D_2O) of GalNAcβProNHCbz (**1b**)



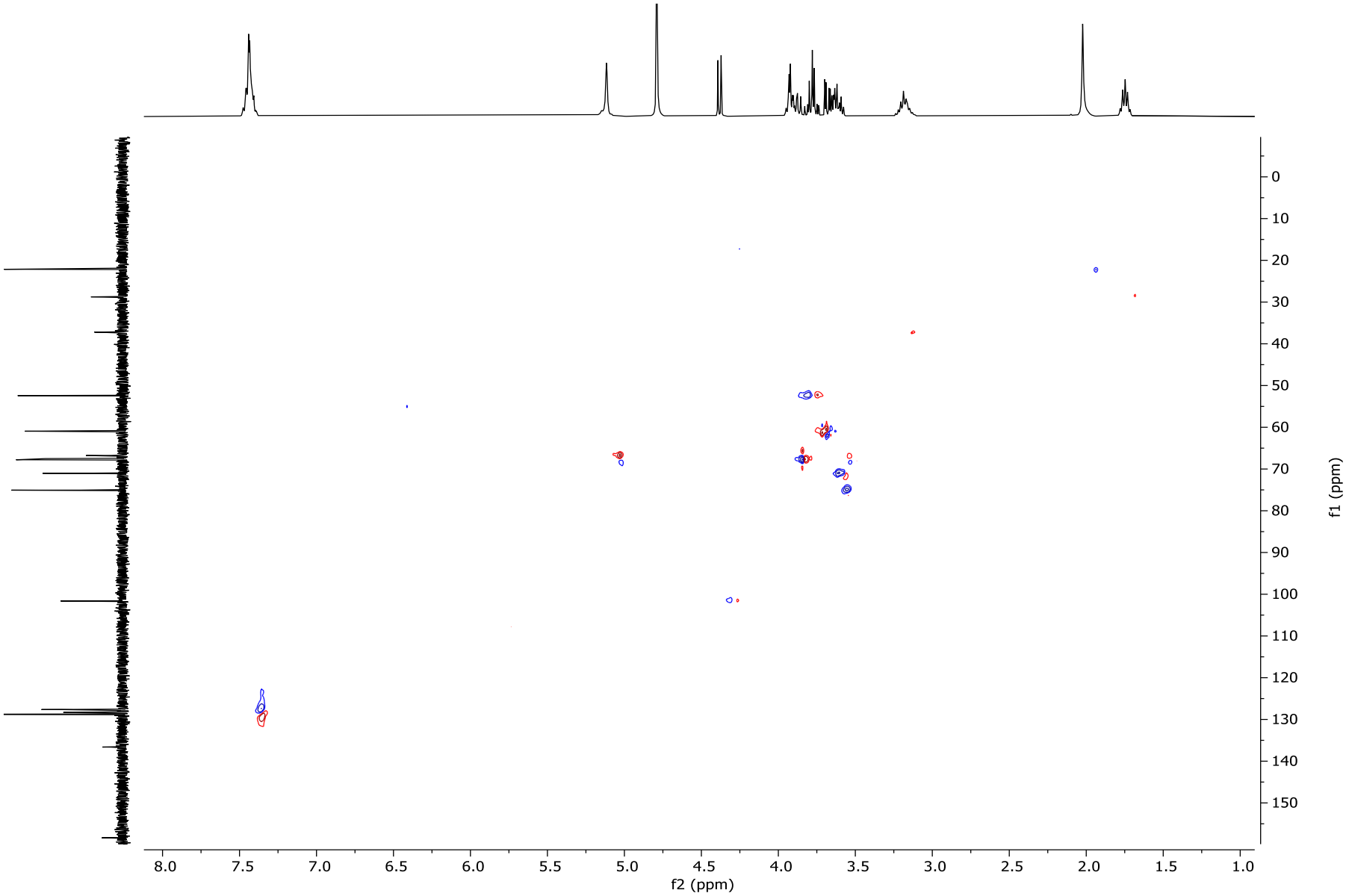
^{13}C NMR Spectra (400 MHz, D_2O) of GalNAcβProNHCbz (**1b**)



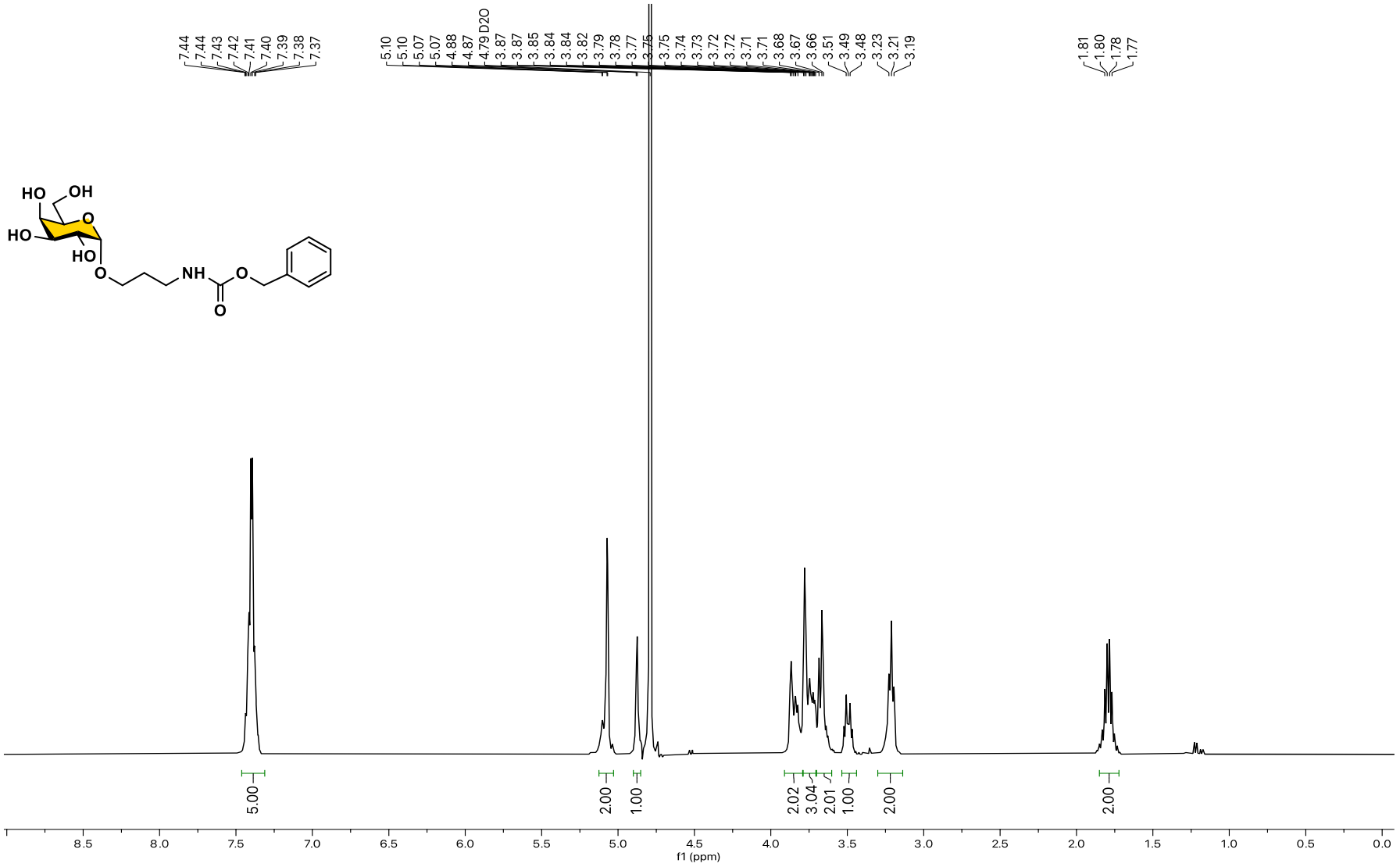
$^{1}H–^{1}H$ COSY NMR Spectra (400 MHz, D_2O) of GalNAcβProNHCbz (**1b**)



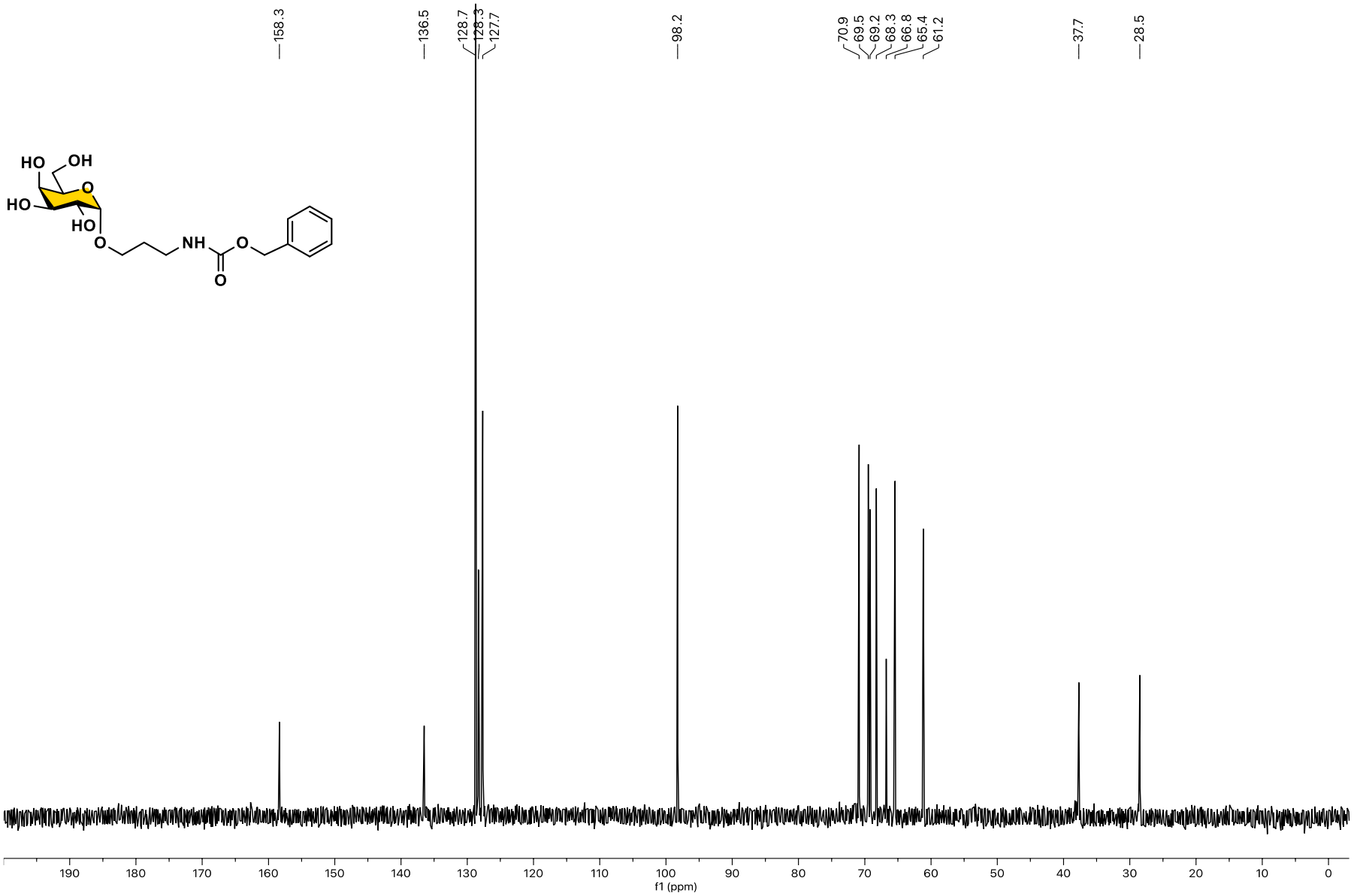
$^{1}H–^{13}C$ HSQC NMR Spectra (400 MHz, D_2O) of GalNAcβProNHCbz (**1b**)



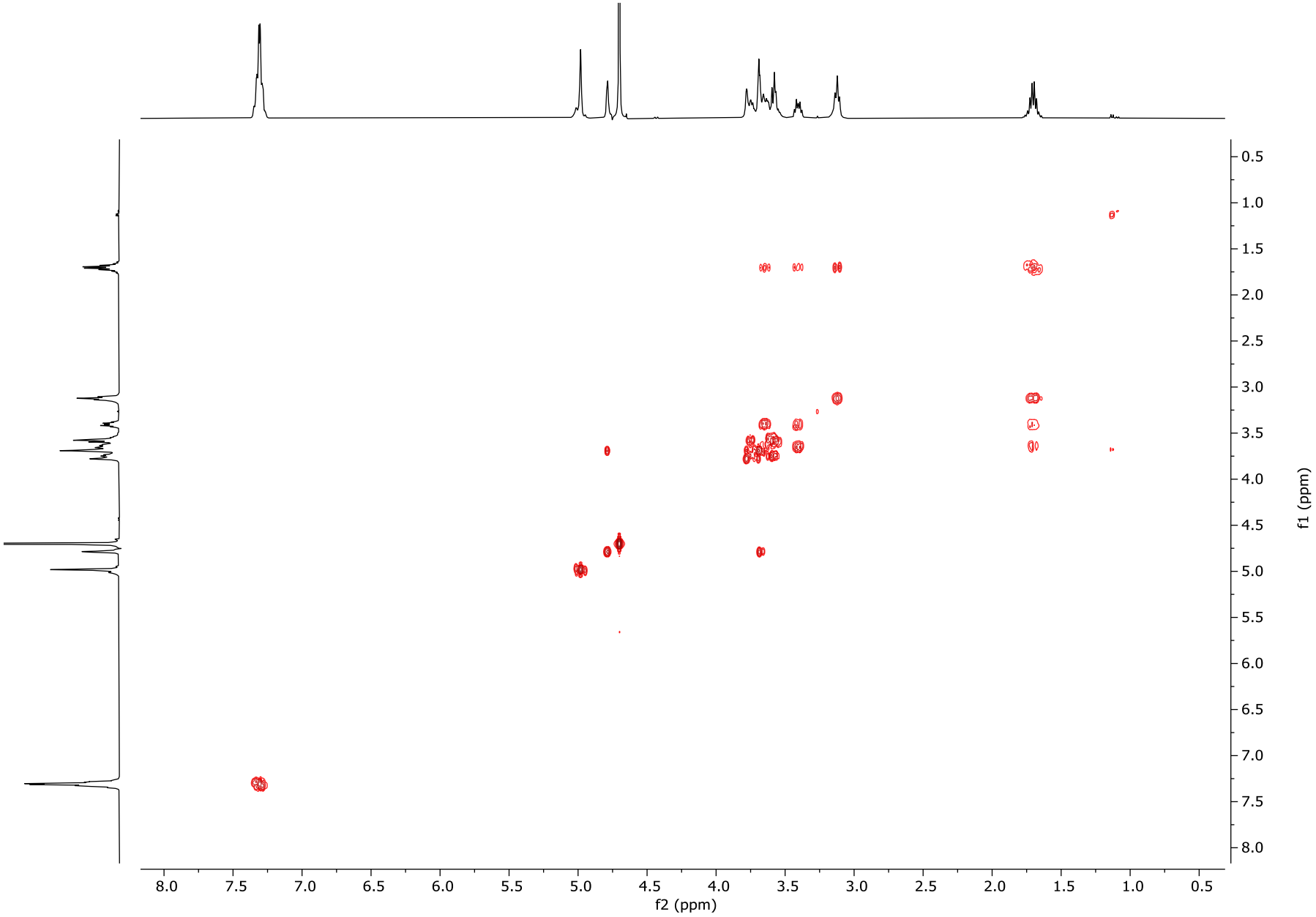
^{1}H spectra (400 MHz, D_2O) of GalαProNHCbz (**2a**)



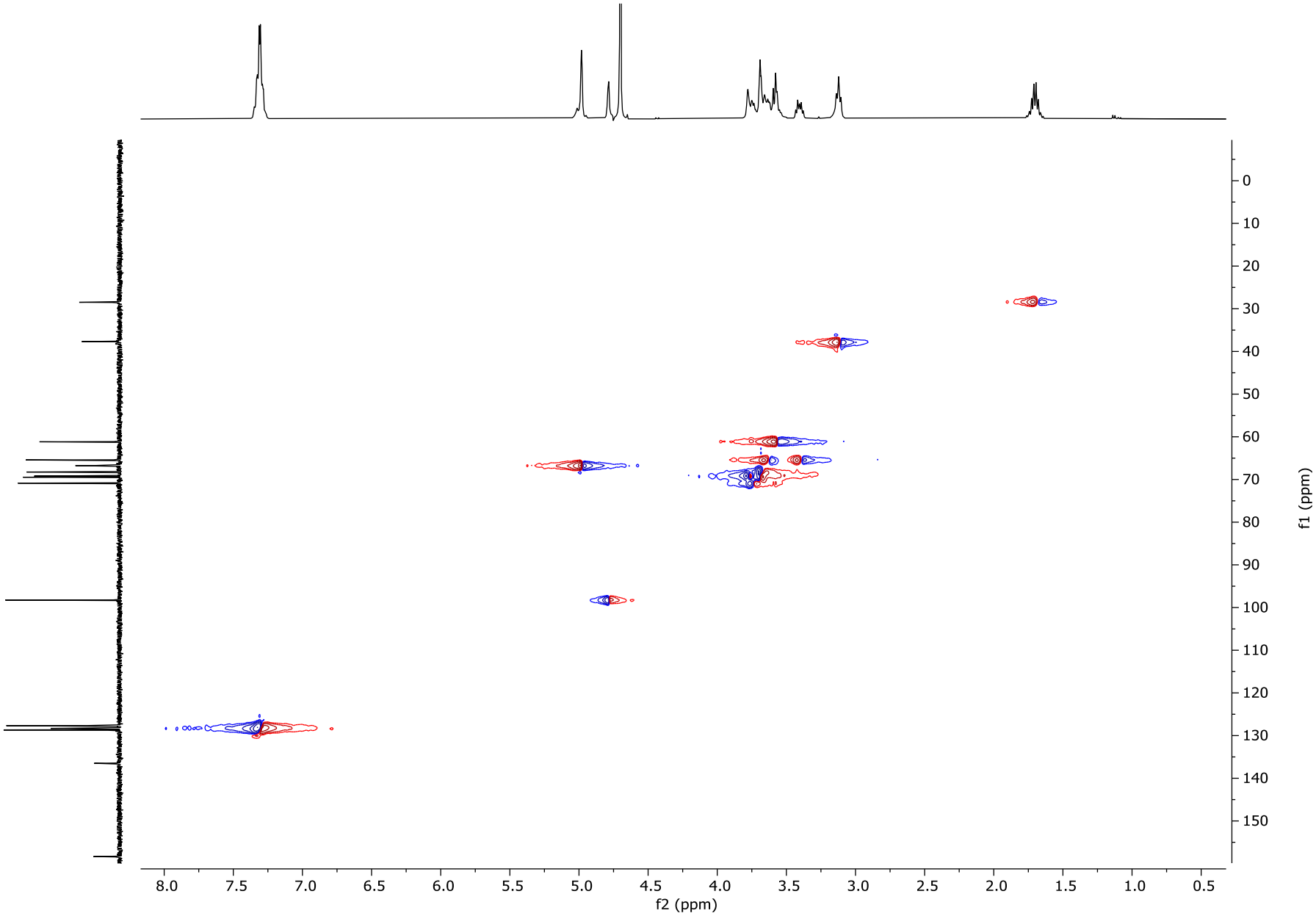
^{13}C NMR spectra (400 MHz, D_2O) of GalαProNHCbz (**2a**)



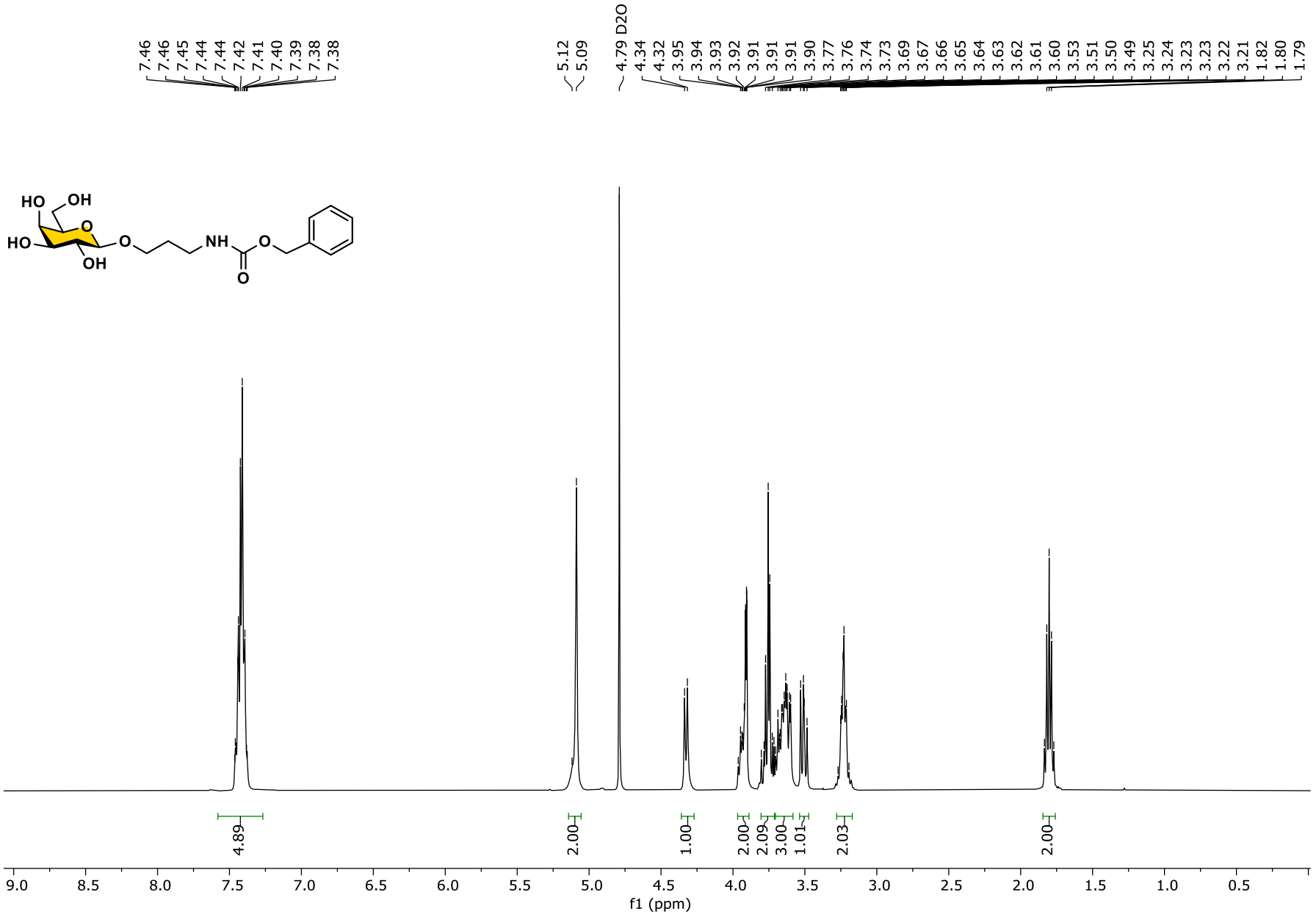
$^{1}H–^{1}H$ COSY NMR Spectra (400 MHz, D_2O) of GalαProNHCbz (**2a**)



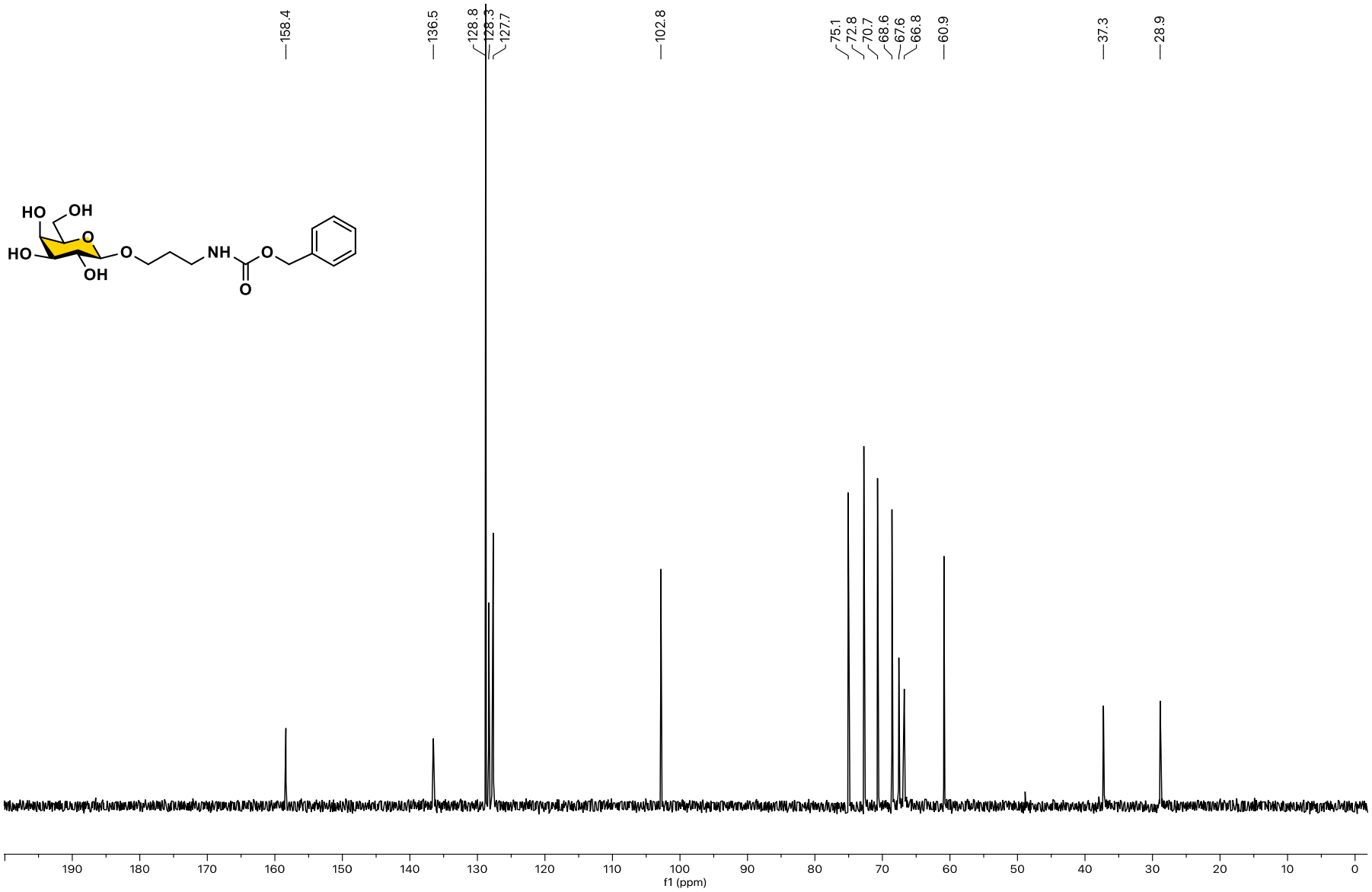
$^{1}H–^{13}C$ HSQC NMR Spectra (400 MHz, D_2O) of GalαProNHCbz (**2a**)



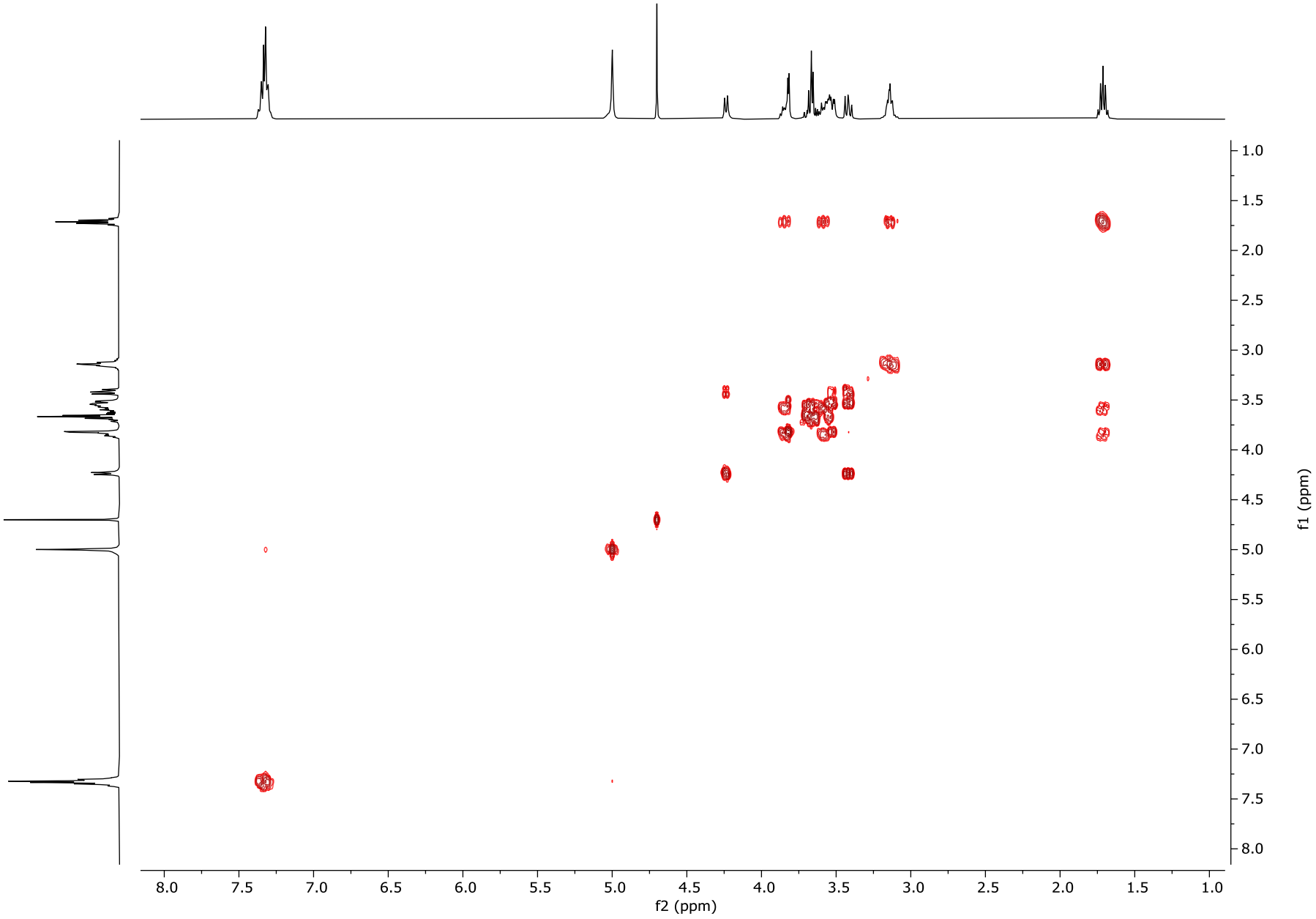
^{1}H NMR Spectra (400 MHz, D_2O) of GalβProNHCbz (**2b**)



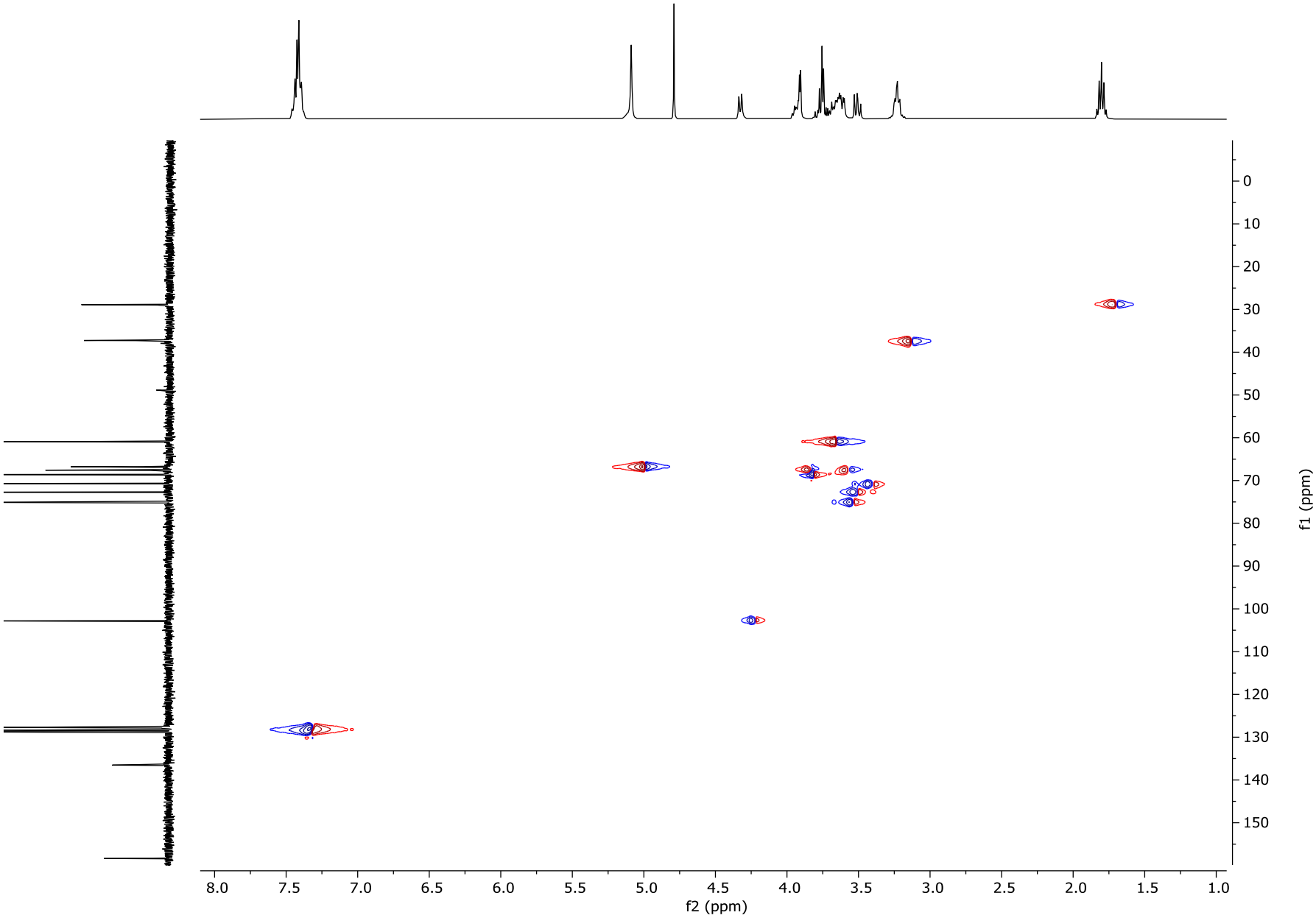
^{13}C NMR Spectra (400 MHz, D_2O) of GalβProNHCbz (**2b**)



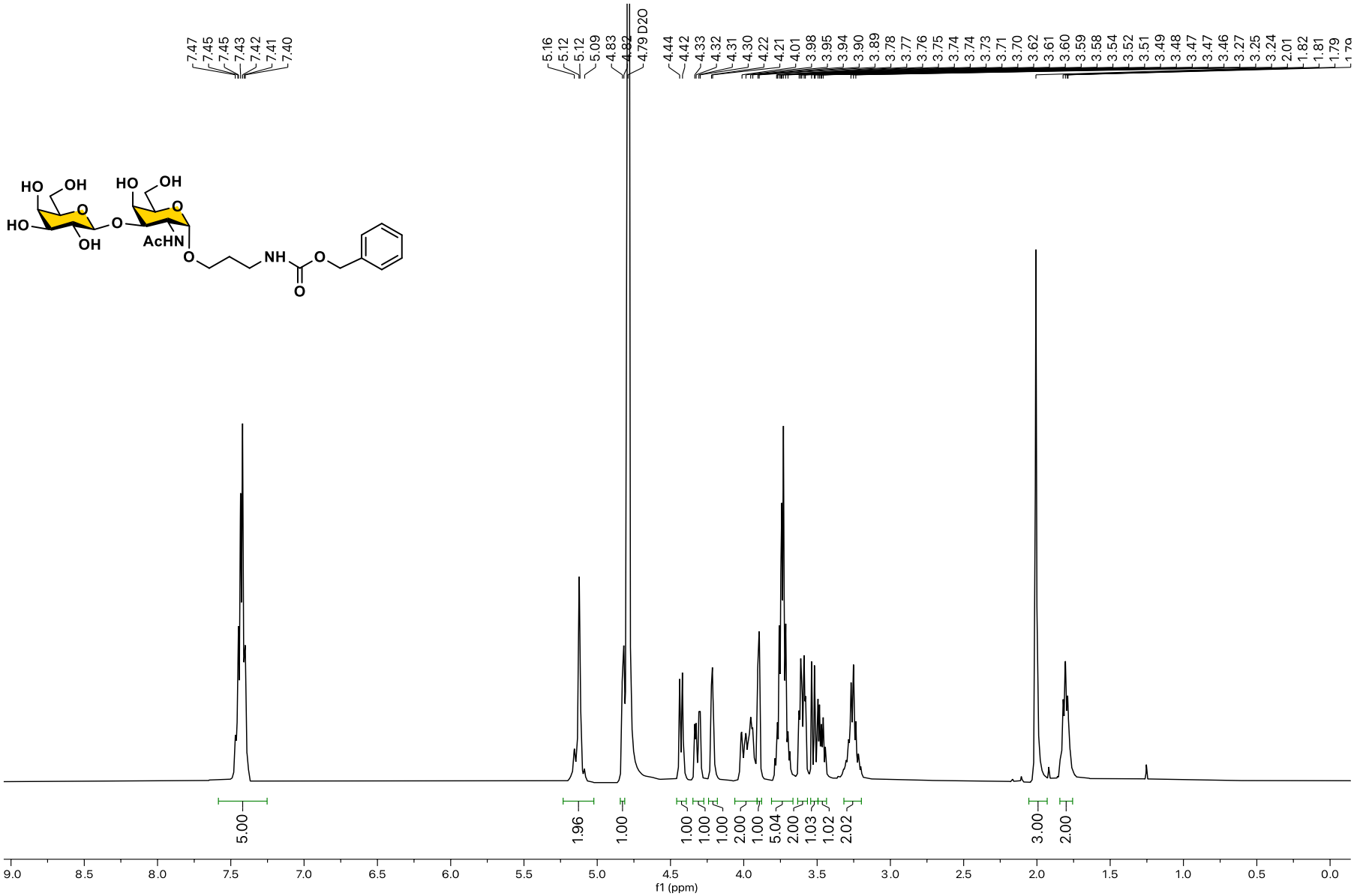
$^{1}H–^{1}H$ COSY NMR Spectra (400 MHz, D_2O) of GalβProNHCbz (**2b**)



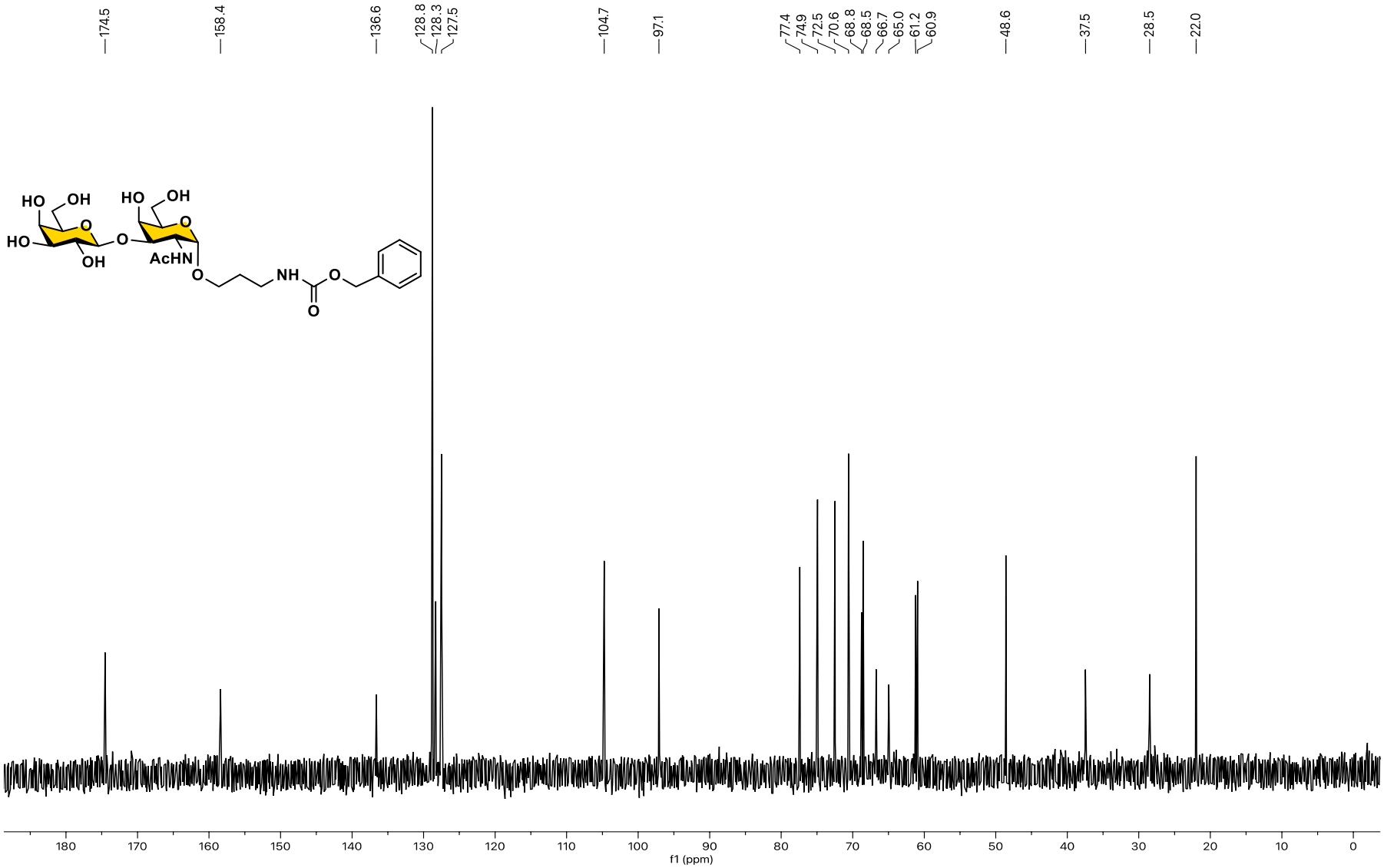
$^{1}H–^{13}C$ HSQC NMR Spectra (400 MHz, D_2O) of GalβProNHCbz (**2b**)



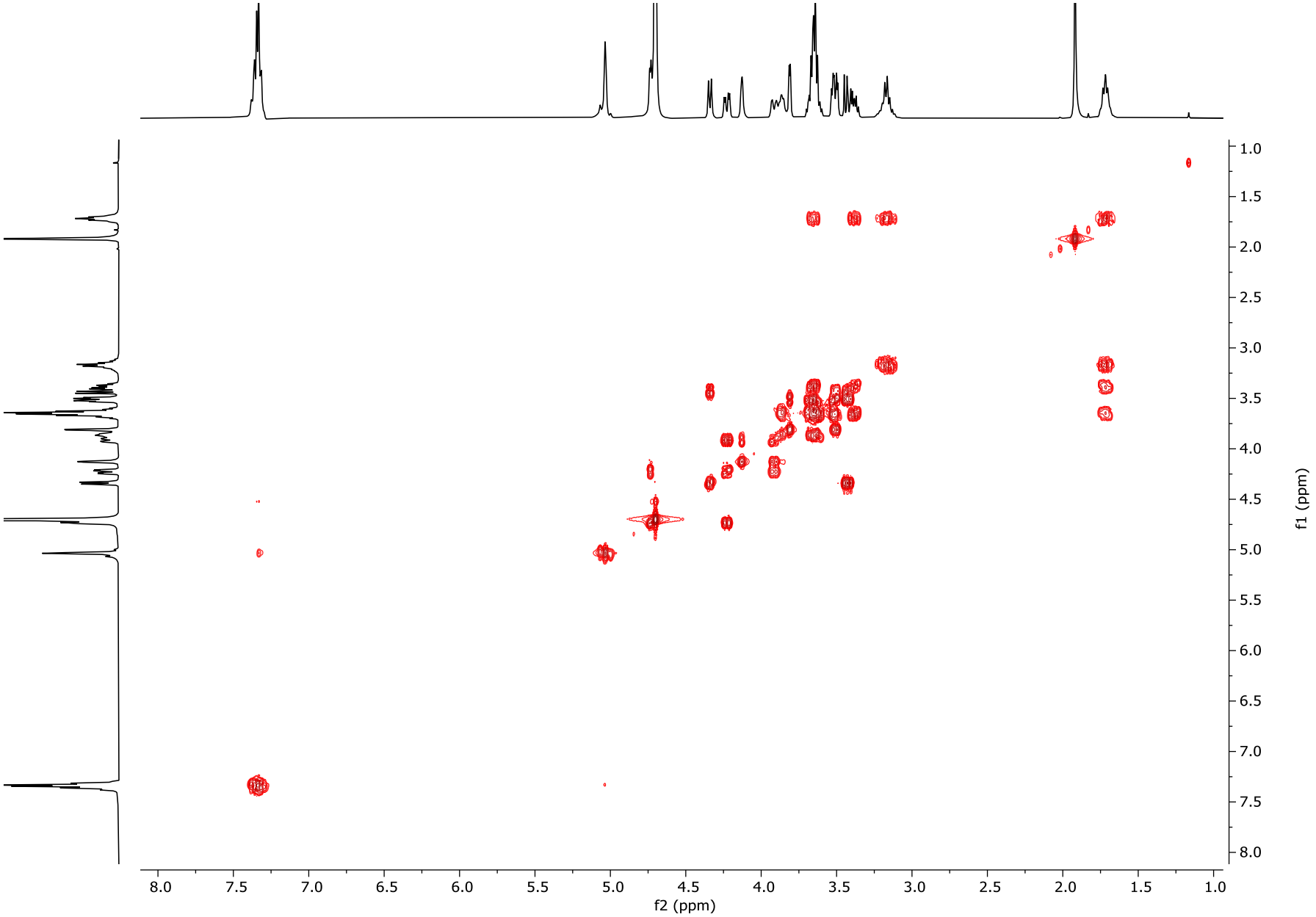
^{1}H Spectra (400 MHz, D_2O) of Galβ3GalNAcαProNHCbz (**3a**)



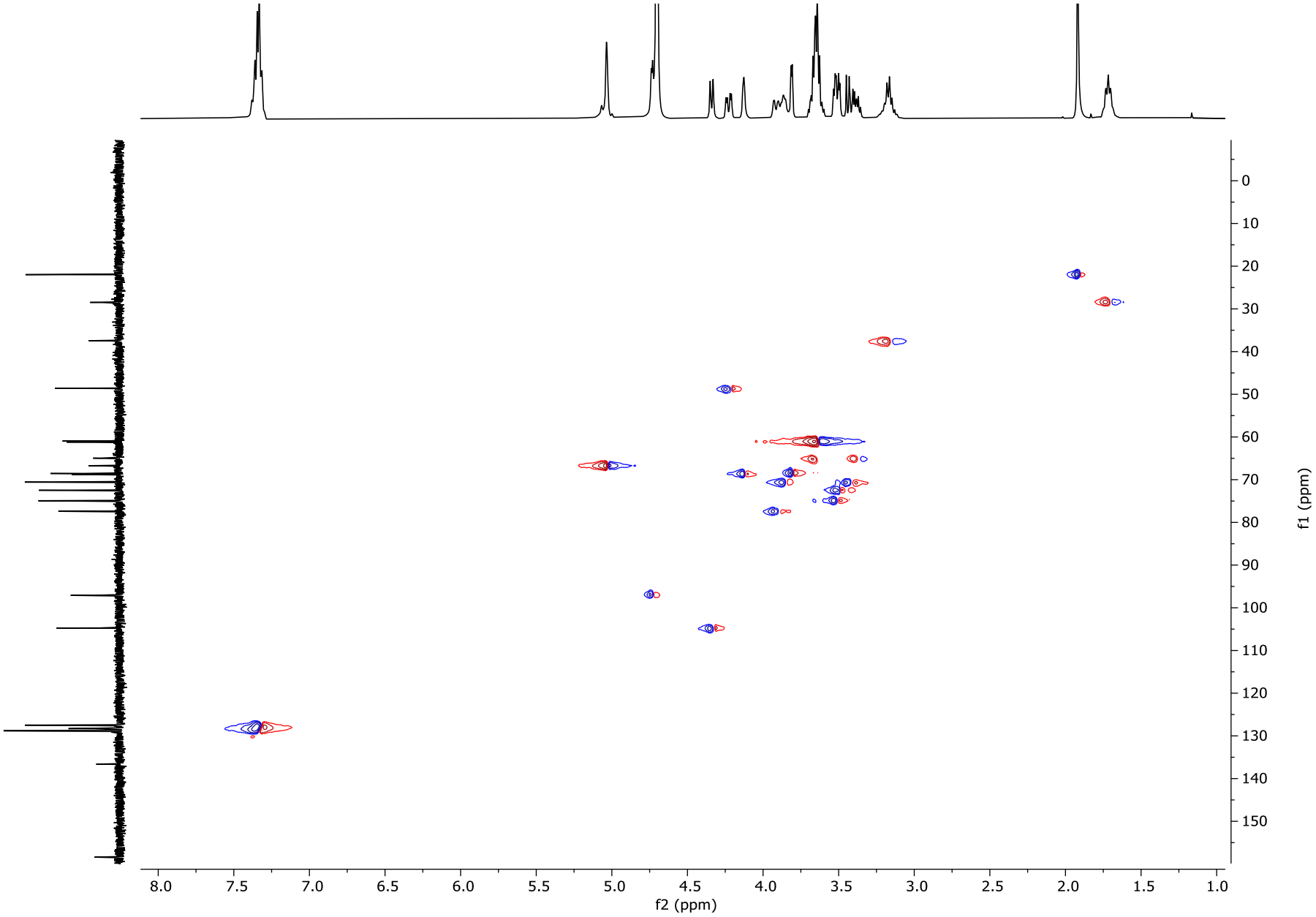
^{13}C NMR Spectra (400 MHz, D_2O) of Galβ3GalNAcαProNHCbz (**3a**)



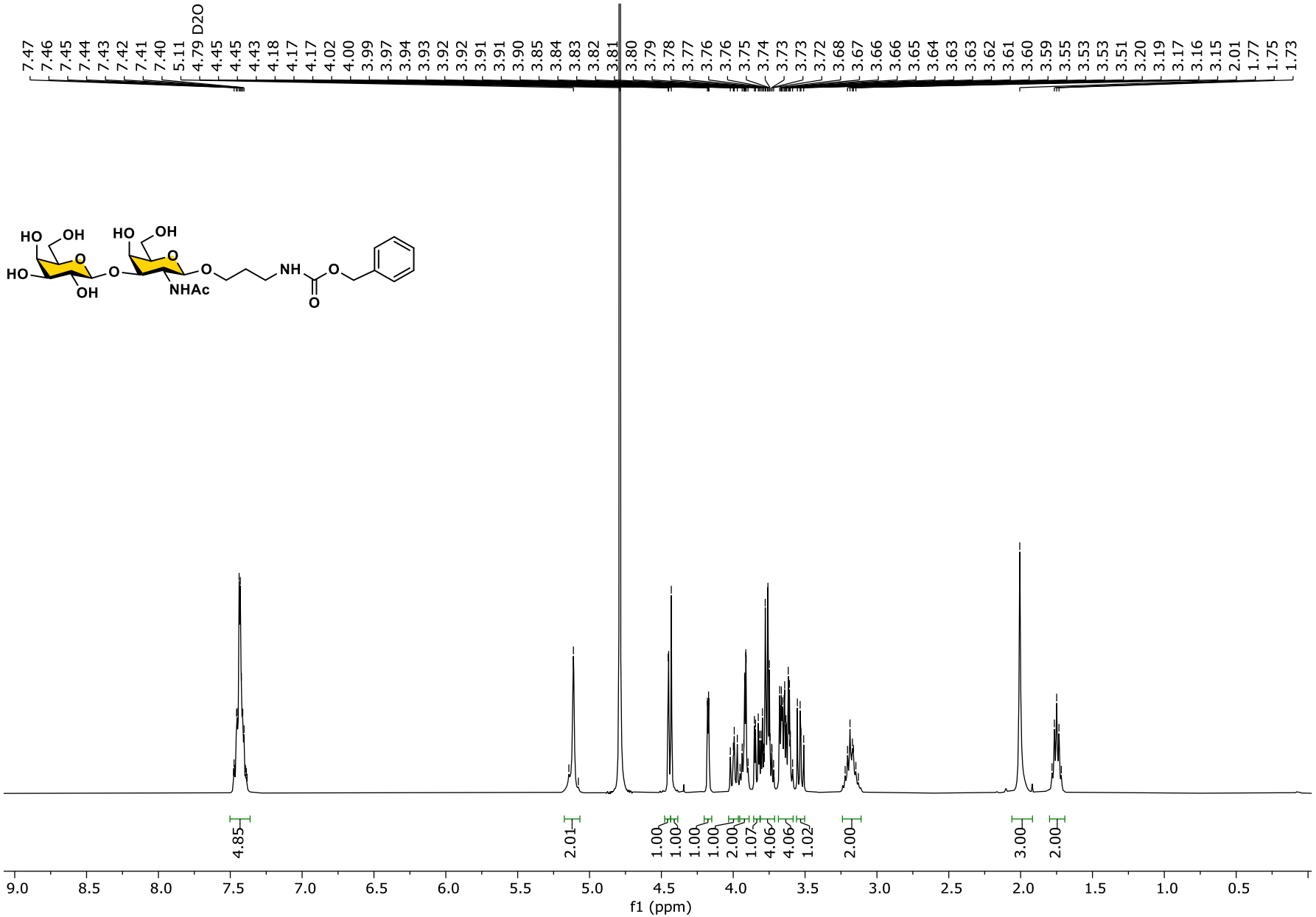
$^{1}H–^{1}H$ COSY NMR Spectra (400 MHz, D_2O) of Galβ3GalNAcαProNHCbz (**3a**)



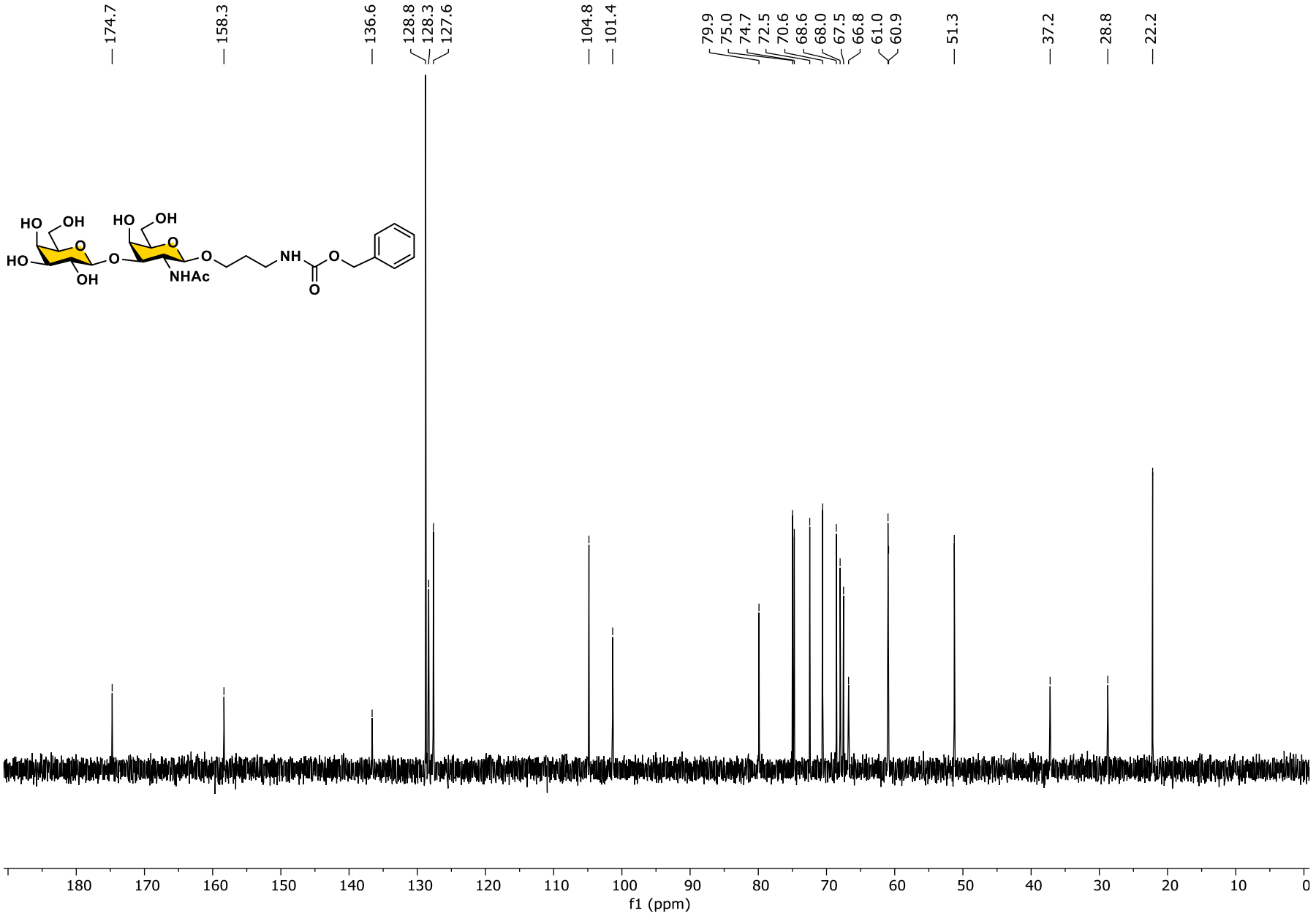
$^{1}H–^{13}C$ HSQC NMR Spectra (400 MHz, D_2O) of Galβ3GalNAcαProNHCbz (**3a**)



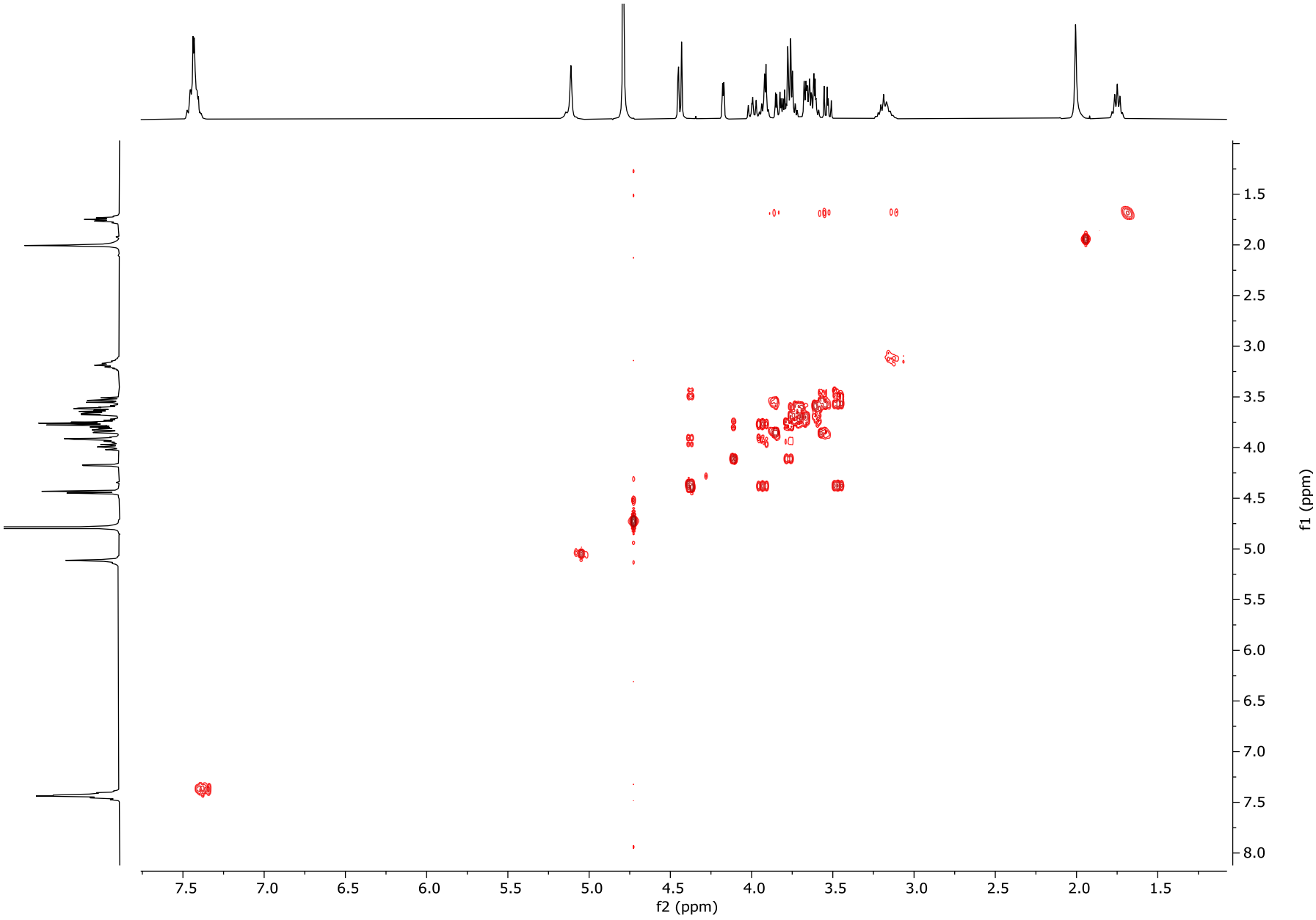
^{1}H NMR Spectra (400 MHz, D_2O) of Galβ3GalNAcβProNHCbz (**3b**)



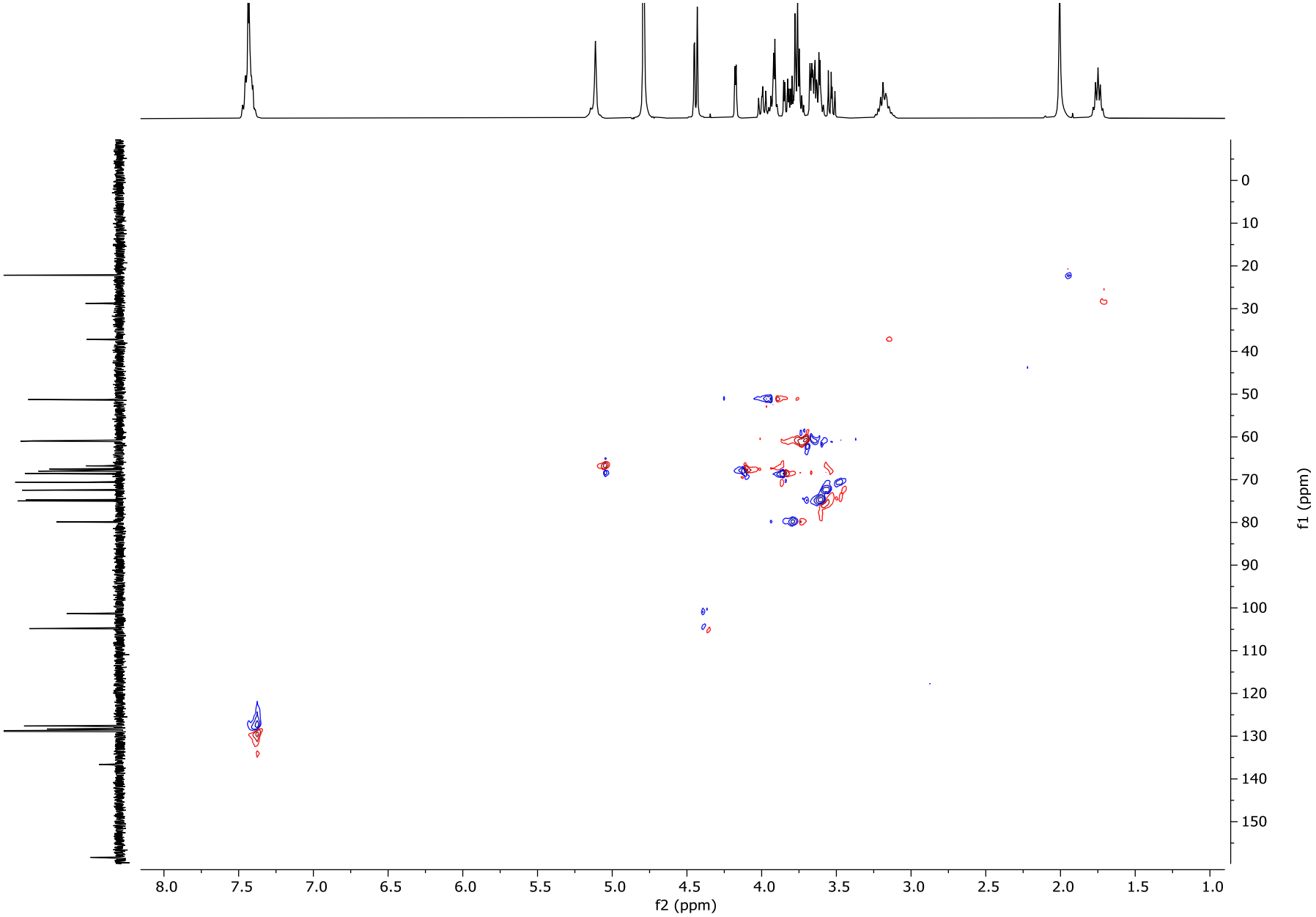
^{13}C NMR Spectra (400 MHz, D_2O) of Galβ3GalNAcβProNHCbz (**3b**)



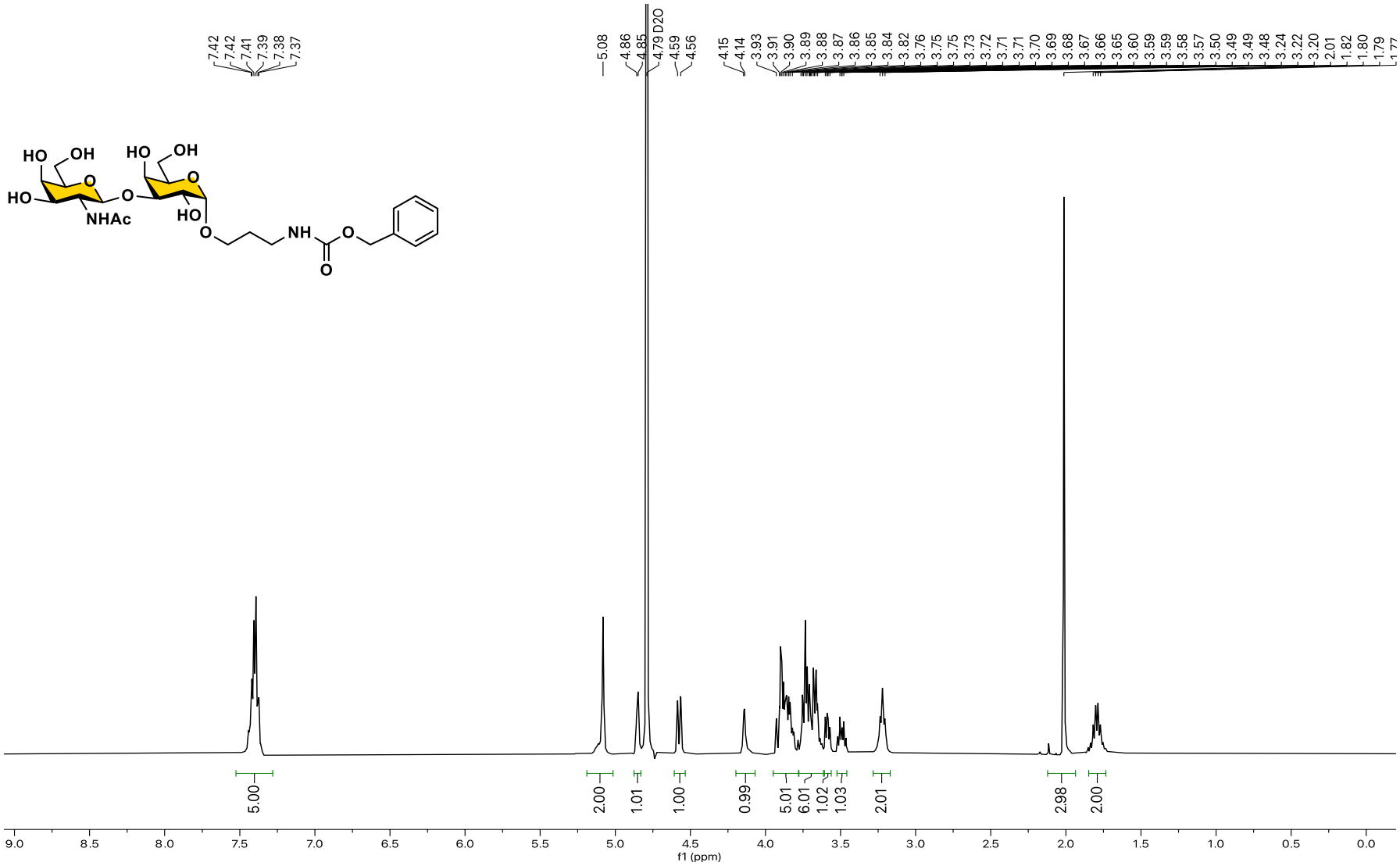
$^{1}H–^{1}H$ COSY NMR Spectra (400 MHz, D_2O) of Galβ3GalNAcβProNHCbz (**3b**)



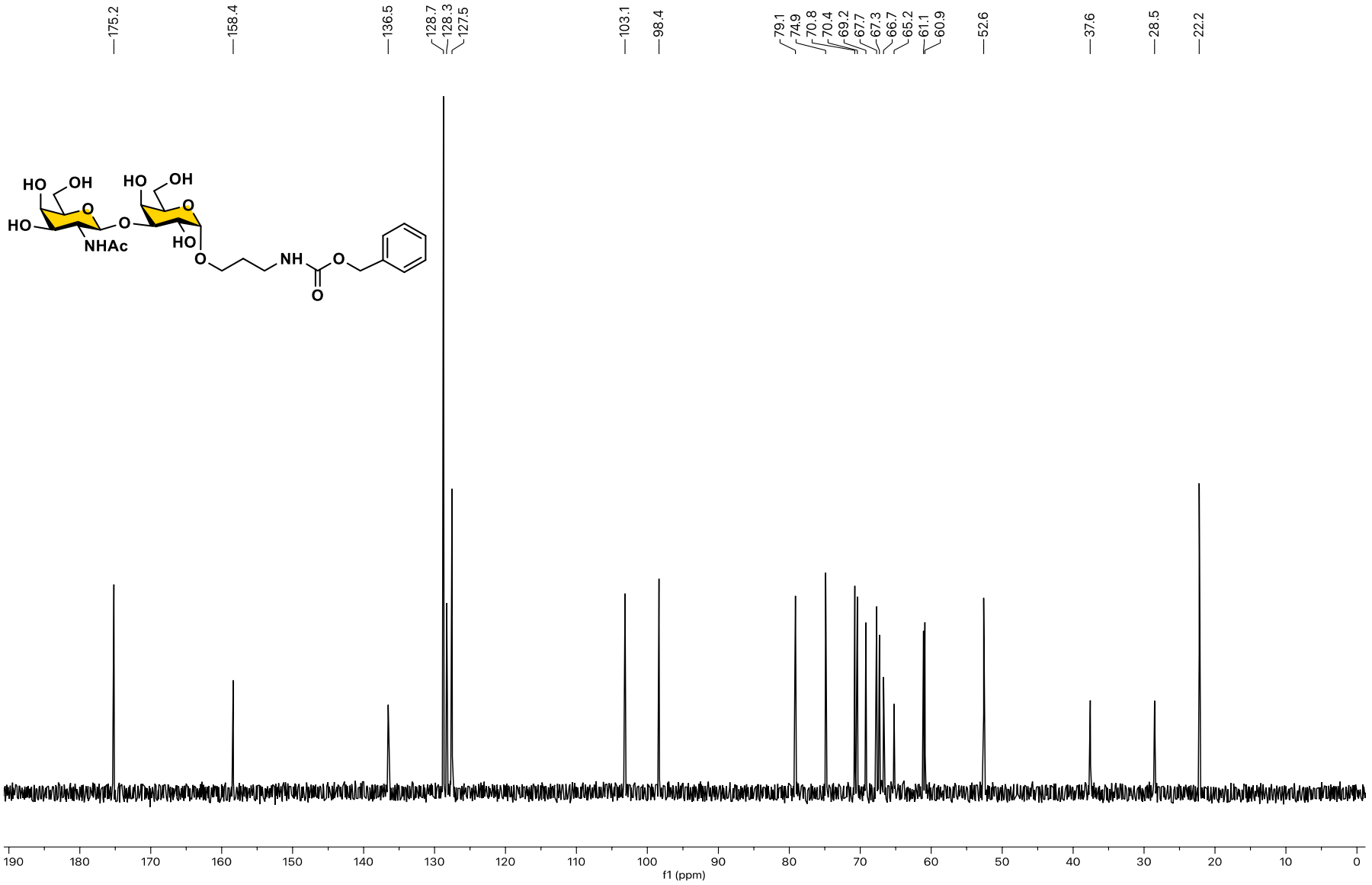
$^{1}H–^{13}C$ HSQC NMR Spectra (400 MHz, D_2O) of Galβ3GalNAcβProNHCbz (**3b**)



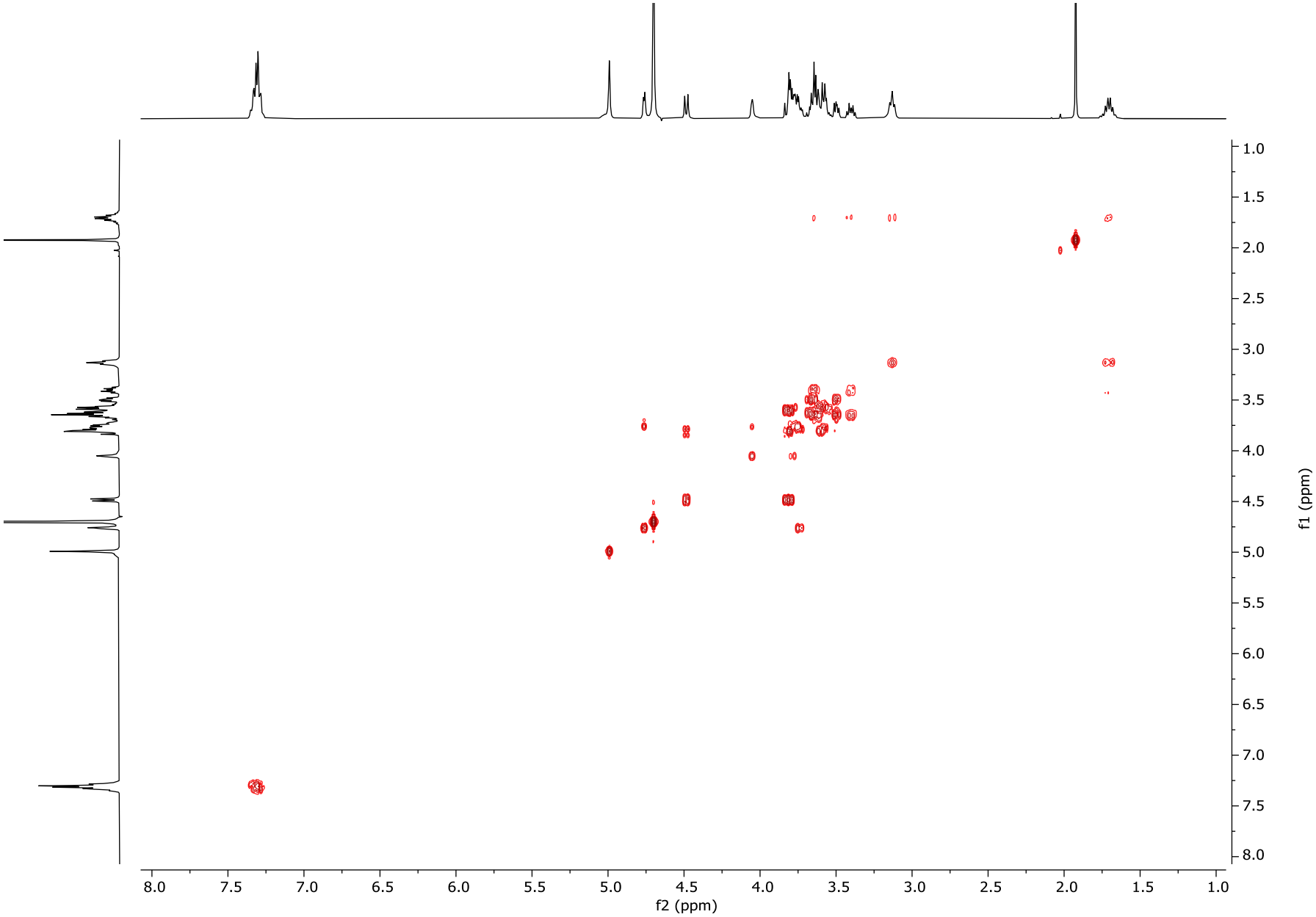
^{1}H NMR Spectra (400 MHz, D_2O) of GalNAcβ3GalαProNHCbz (**4a**)



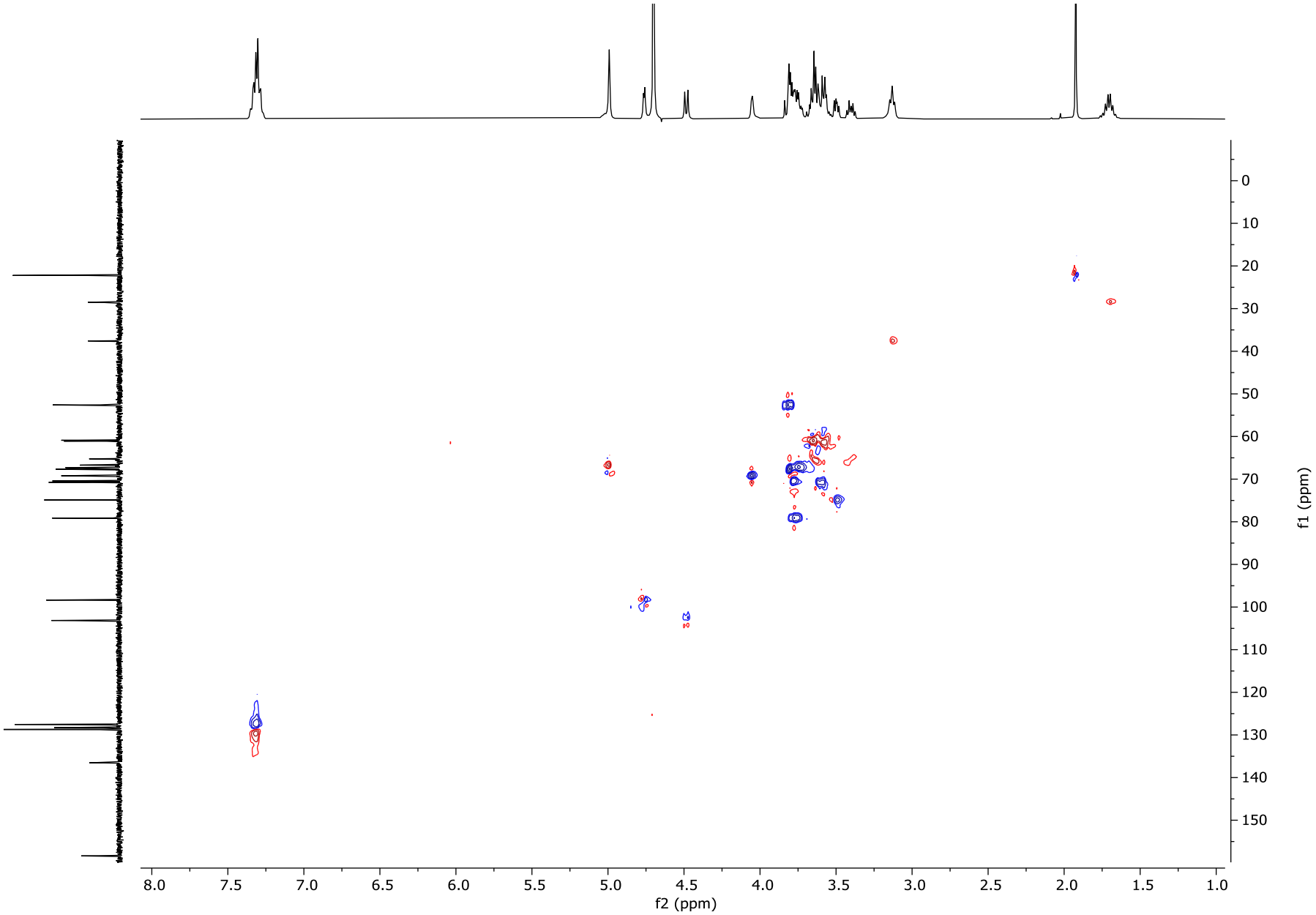
^{13}C NMR Spectra (400 MHz, D_2O) of GalNAcβ3GalαProNHCbz (**4a**)



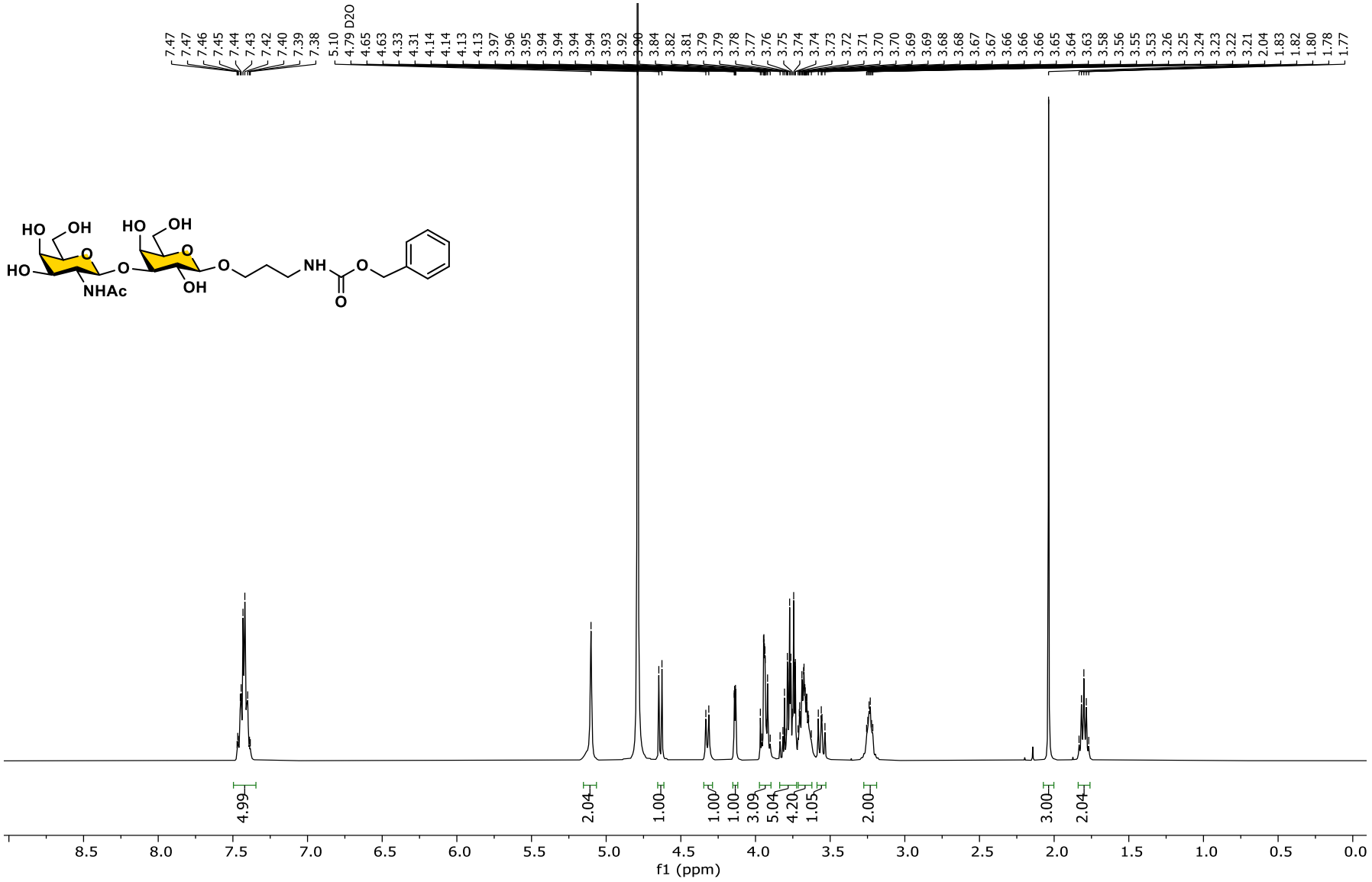
$^{1}H–^{1}H$ COSY NMR Spectra (400 MHz, D_2O) of GalNAcβ3GalαProNHCbz (**4a**)



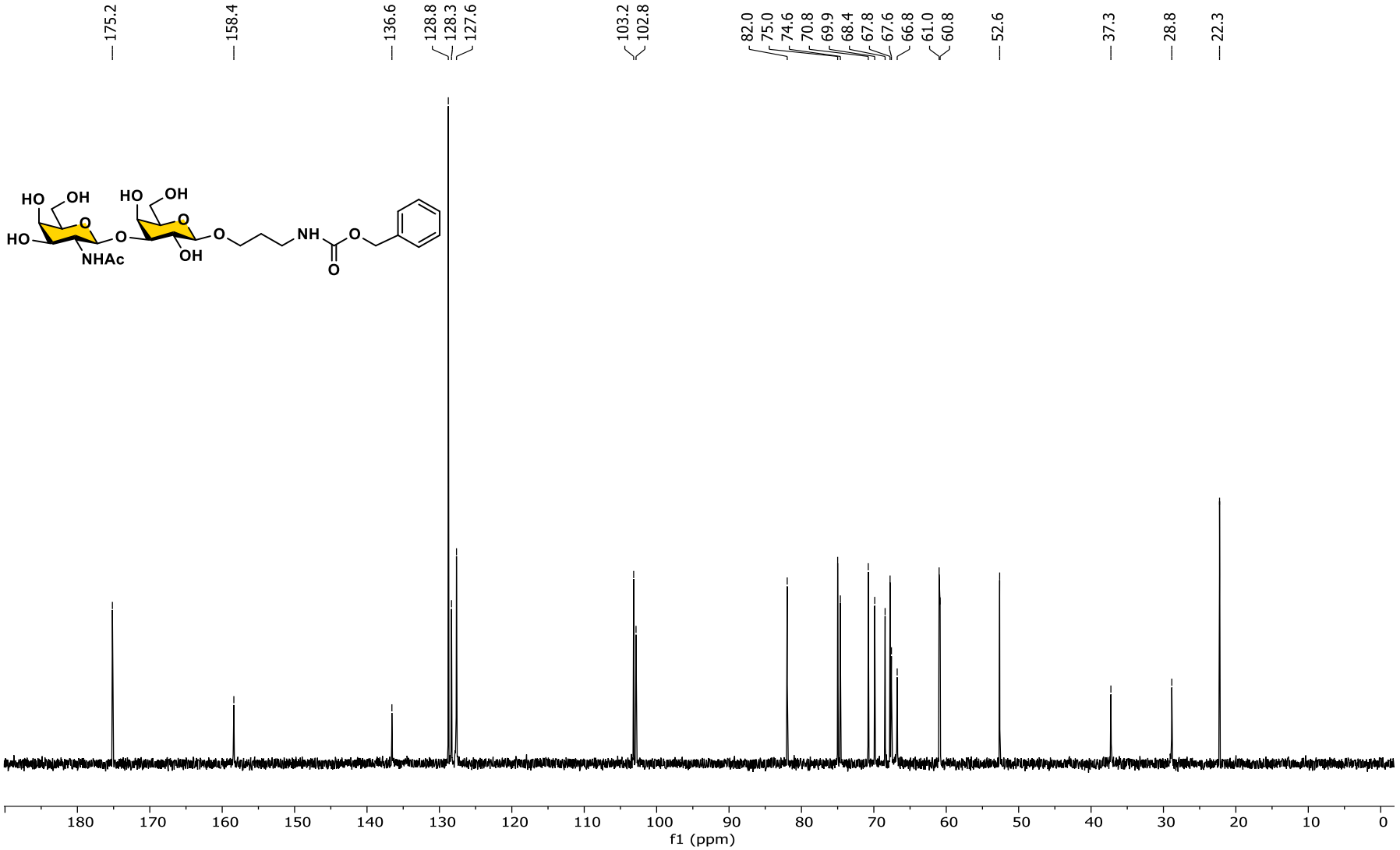
$^{1}H–^{13}C$ HSQC NMR Spectra (400 MHz, D_2O) of GalNAcβ3GalαProNHCbz (**4a**)



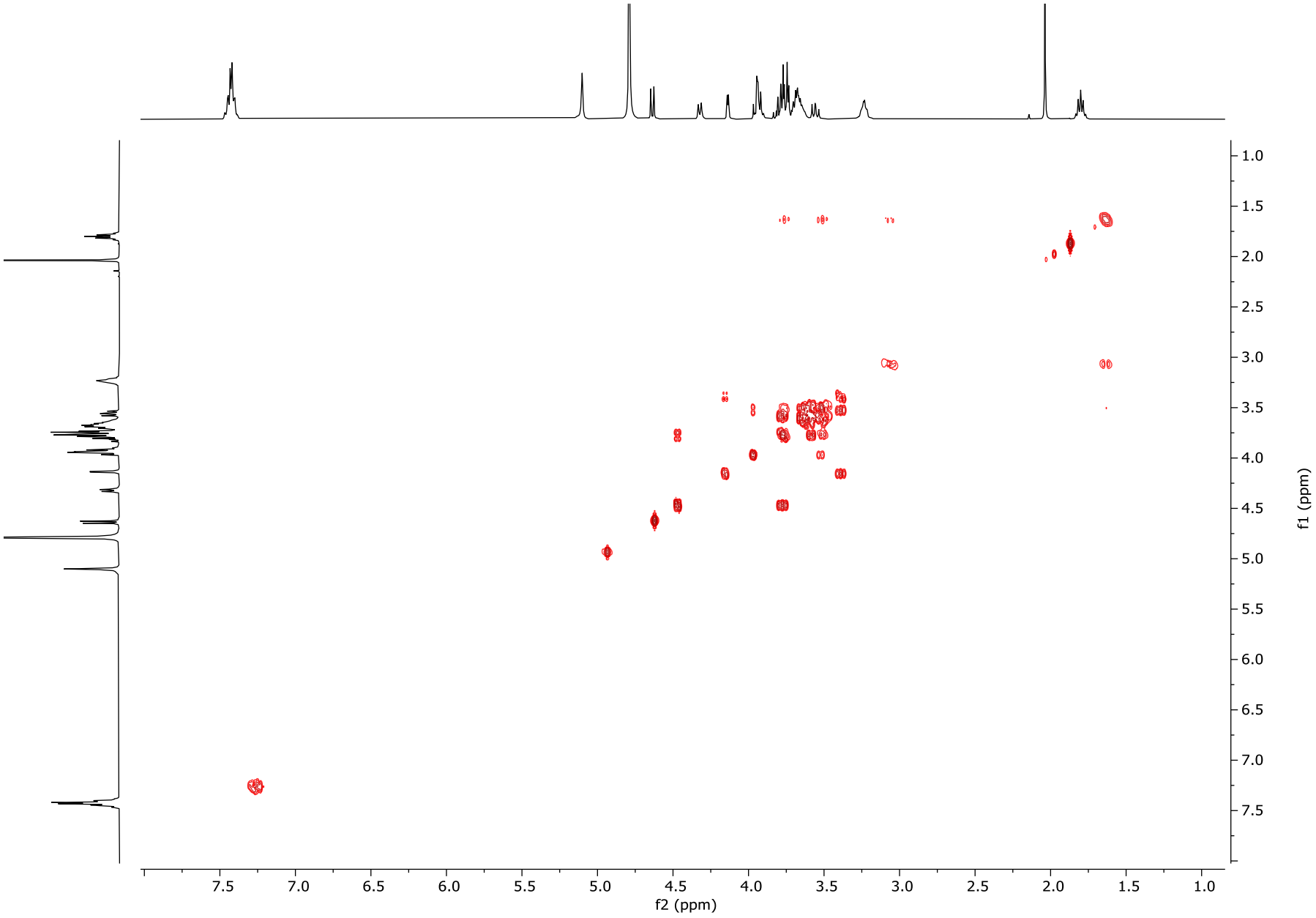
^{1}H NMR Spectra (400 MHz, D_2O) of GalNAcβ3GalβProNHCbz (**4b**)



^{13}C NMR Spectra (400 MHz, D_2O) of GalNAcβ3GalβProNHCbz (**4b**)



$^{1}H–^{1}H$ COSY NMR Spectra (400 MHz, D_2O) of GalNAcβ3GalβProNHCbz (**4b**)



$^{1}H–^{13}C$ HSQC NMR Spectra (400 MHz, D_2O) of GalNAcβ3GalβProNHCbz (**4b**)