Supporting Information

Exploration of the active site of β4GalT7: Modifications of the aglycon of aromatic xylosides

Anna Siegbahn,[a#] Karin Thorsheim,[a#] Jonas Ståhle,[b] Sophie Manner,[a] Christoffer Hamark,[b] Andrea Persson,[c] Emil Tykesson,[c] Katrin Mani,[c] Gunilla Westergren-Thorsson,[c] Göran Widmalm,[b] Ulf Ellervik[a]*

[a] Center for Analysis and Synthesis, Center for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden

[b] Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, SE-106 91 Stockholm, Sweden

[c] Department of Experimental Medical Science, Lund University, BMC, SE-221 00 Lund, Sweden

[#] These authors contributed equally to this work

Table of Contents

Mass analysis

Table S1. HRMS for all galactosylated products.

Acceptor substrate	**Galactosylation product** Chemical formula	**HRMS** calcd for $[M+H]^+$	**HRMS** found
1	$C_{21}H_{26}O_{10}$	439.1599	439.1598
2	$C_{12}H_{22}O_{10}$	327.1286	327.1287
3	$C_{17}H_{24}O_{10}$	389.1442	389.1441
4	$C_{21}H_{26}O_{10}$	439.1599	439.1598
5	$C_{23}H_{28}O_{10}$	465.1755	465.1755
6	$C_{26}H_{30}O_{10}$	520.2177[a]	520.2179
7	$C_{21}H_{26}O_9S$	455.1370	455.1370
8	$C_{21}H_{26}O_{11}S$	487.1269	487.1270
9	$C_{22}H_{28}O_9$	437.1806	437.1804
12	$C_{22}H_{28}O_{10}$	453.1755	453.1755
13	$C_{23}H_{30}O_{11}$	483.1861	483.1858
14	$C_{25}H_{34}O_{12}$	527.2123	527.2123
15	$C_{27}H_{38}O_{13}$	571.2385	571.2388
16	$C_{22}H_{28}O_{11}$	486.1973[a]	486.1972
17	$C_{21}H_{26}O_{11}$	455.1548	455.1549
18	$C_{23}H_{28}O_{12}$	497.1654	497.1655
19	$C_{22}H_{25}NO_{10}$	464.1551	464.1550

[a] HRMS for $[M+NH_4]^+$

Cloning, expression and purification of β4GALT7

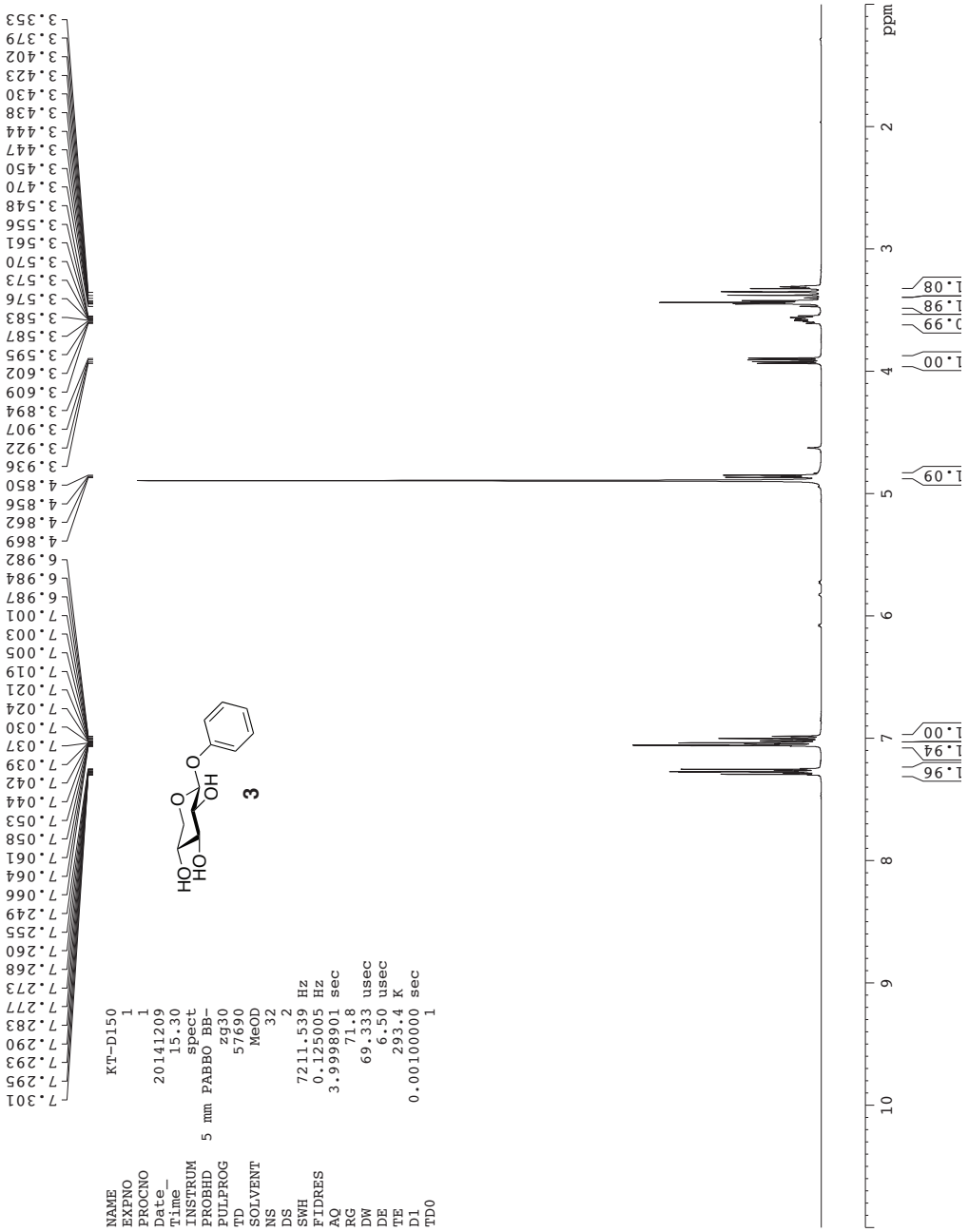
The part of the b4GALT7 open reading frame (sequence harmonized for expression in *E. Coli*, Genewiz) corresponding to aa 82 to aa 327 was subcloned into the NcoI and NotI sites of a pET-41a(+) (Merck Millipore) expression vector using the following primers (Sigma): Forward (NcoI restriction site in bold letters)

aa 82: 5'-GCATCTC**CCATGG**AGCATTGGGAGGAAGATGCCAGTT-3'

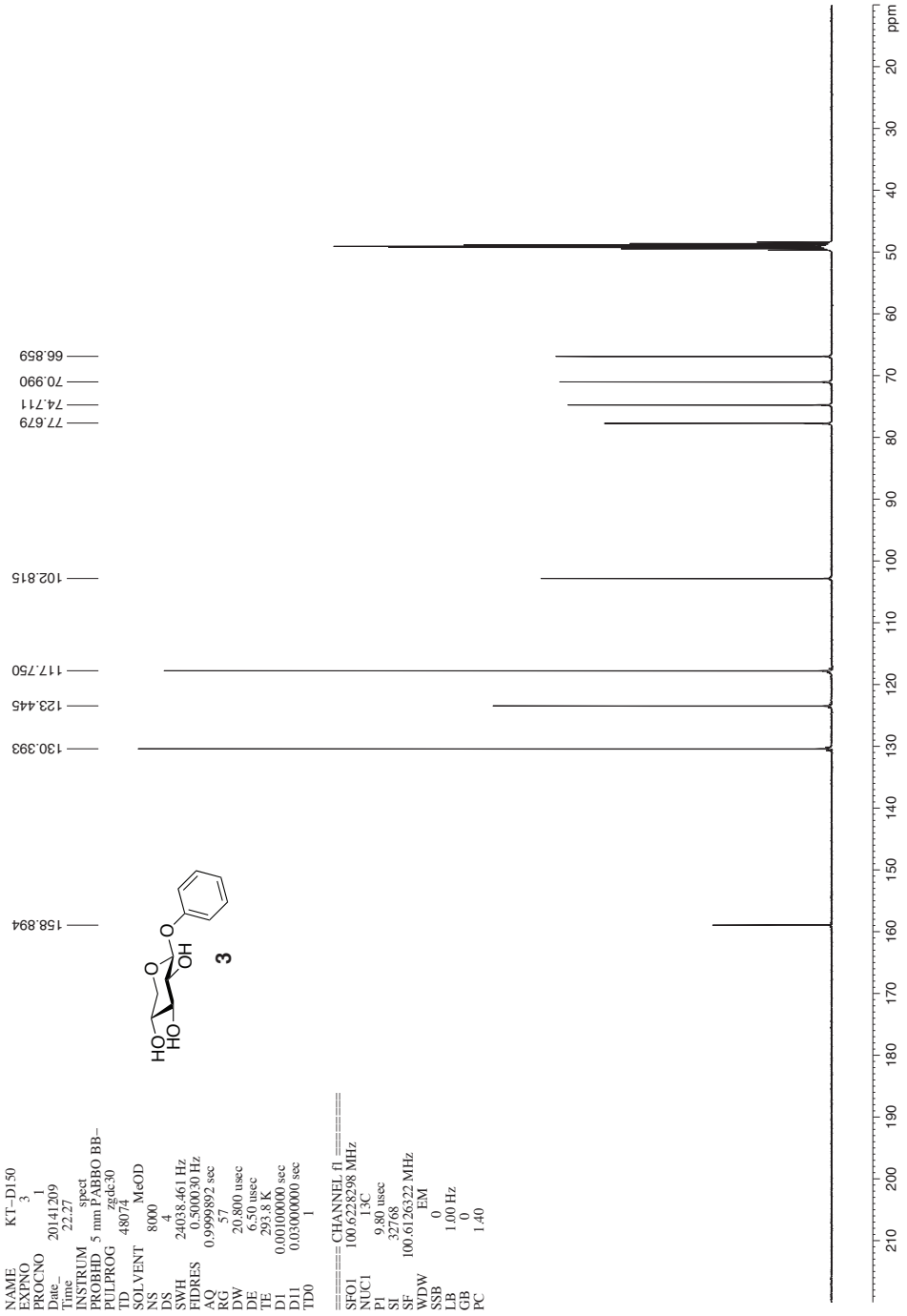
Reverse (NotI restriction site in bold letters) aa 327: 5'-GCATCT**GCGGCCGC**ATTAGCTAAATGTACACCAAGGGGTTGC-3'

The plasmid product, pET-41a(+)-β4GALT7$_{aa82}$, was amplified (Library Efficiency DH5α bacteria, Invitrogen), purified (HiSpeed Plasmid Midi Kit, Qiagen) and sequenced (Eurofins MWG Operon). BL21(DE3) *E. Coli* bacteria (Merck Millipore) were transformed with pET-41a(+)-β4GALT7$_{aa82}$ and grown on LB agar containing kanamycin (30 mg/ml). A resistant colony was transferred to selective LB medium and grown at 37 °C with shaking at 250 rpm to OD_{600}≈0.5. Expression was induced at 32 °C with IPTG (0.56 mM, Sigma) and culture was left shaking at 250 rpm over night. Cells were pelleted and then lysed with Cellytic (1X, Sigma) in Tris (20 mM, pH 7.6), NaCl (500 mM), Benzonase (50 U/ml, Sigma), lysozyme (0.2 mg/ml, Sigma) and PMSF (1 mM) for 20 min at 22 °C. Clarified lysate was applied to an equilibrated HisTrap HP column (GE Healthcare) at a flowrate of 1 ml/min. The column was washed with a phosphate buffer (20 mM, pH 7.4) containing NaCl (500 mM) and imidazole (40 mM), after which the protein was eluted by an imidazole gradient (40-250 mM). Protein purity for each eluted fraction was analyzed by SDS-PAGE on a Bis-Tris 4-12 % gel in MOPS buffer (Invitrogen), visualized by Brilliant Blue G colloidal (Sigma). Fractions containing pure protein were pooled and buffer exchanged into TBS (20 mM, pH 7.9) on PD-10 columns (GE Healthcare).

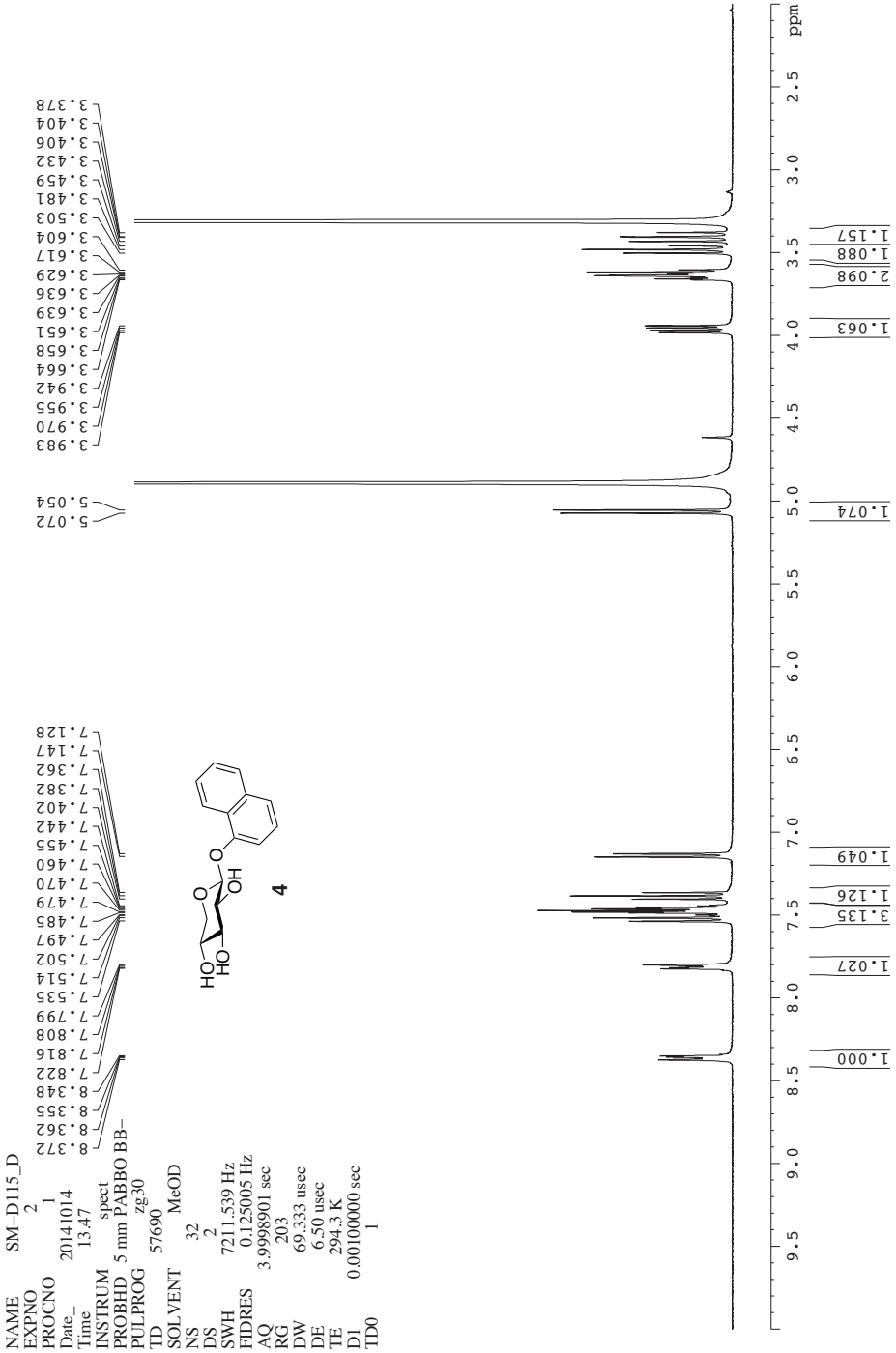
[1]H-NMR Phenyl β-D-xylopyranoside **(3)**



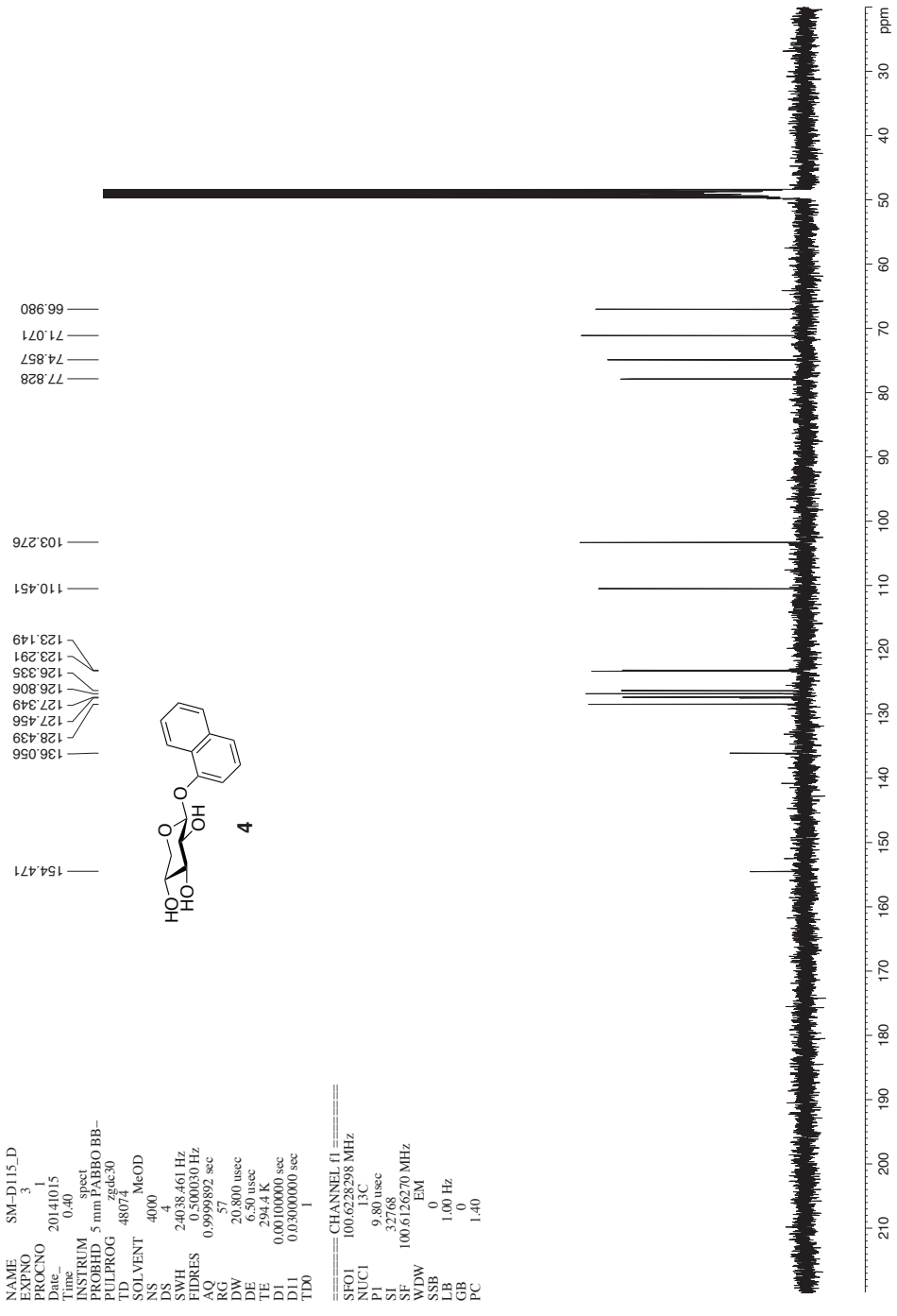
^{13}C-NMR Phenyl β-D-xylopyranoside **(3)**



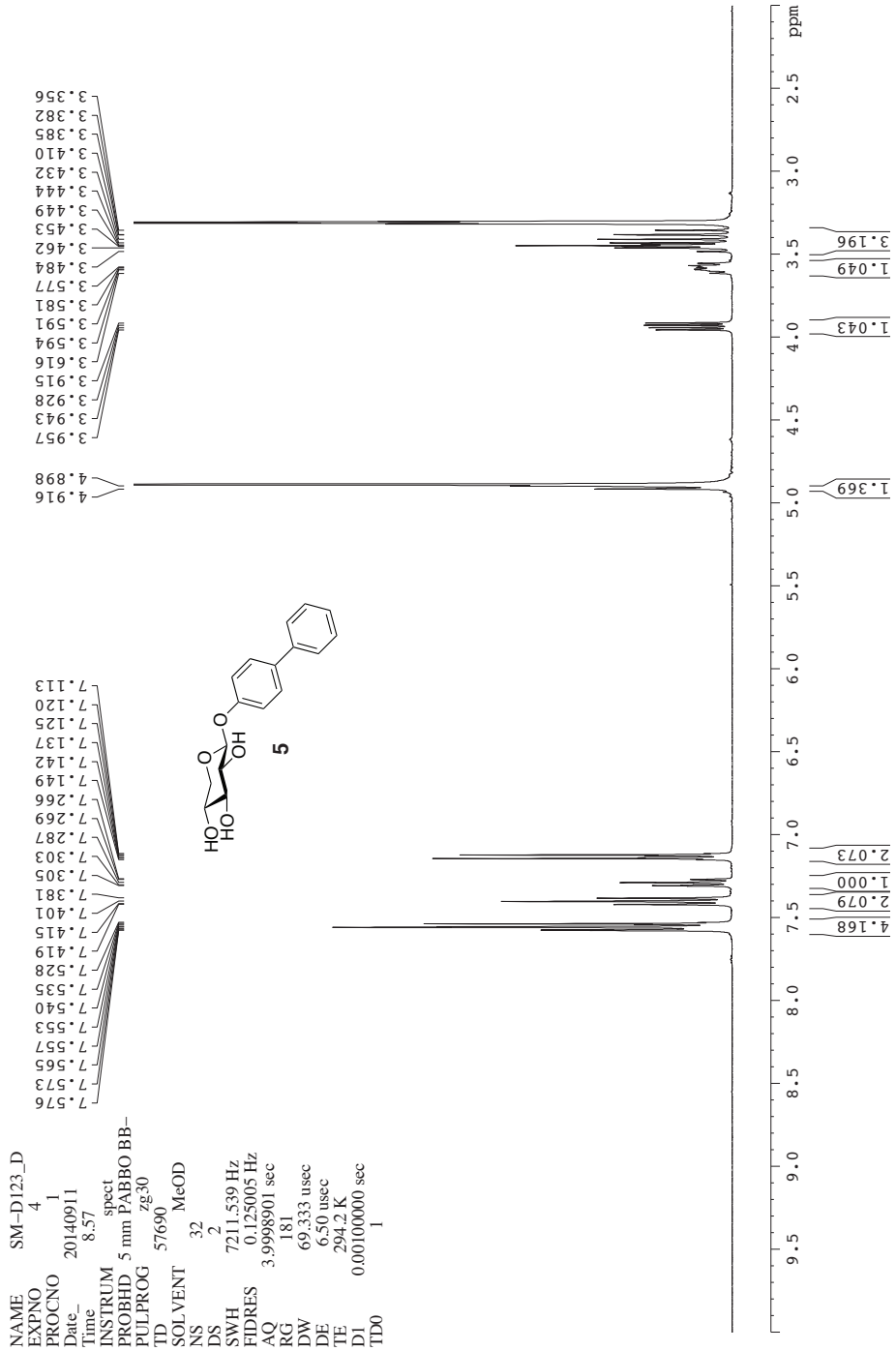
^{1}H-NMR 1-Naphthyl β-D-xylopyranoside **(4)**



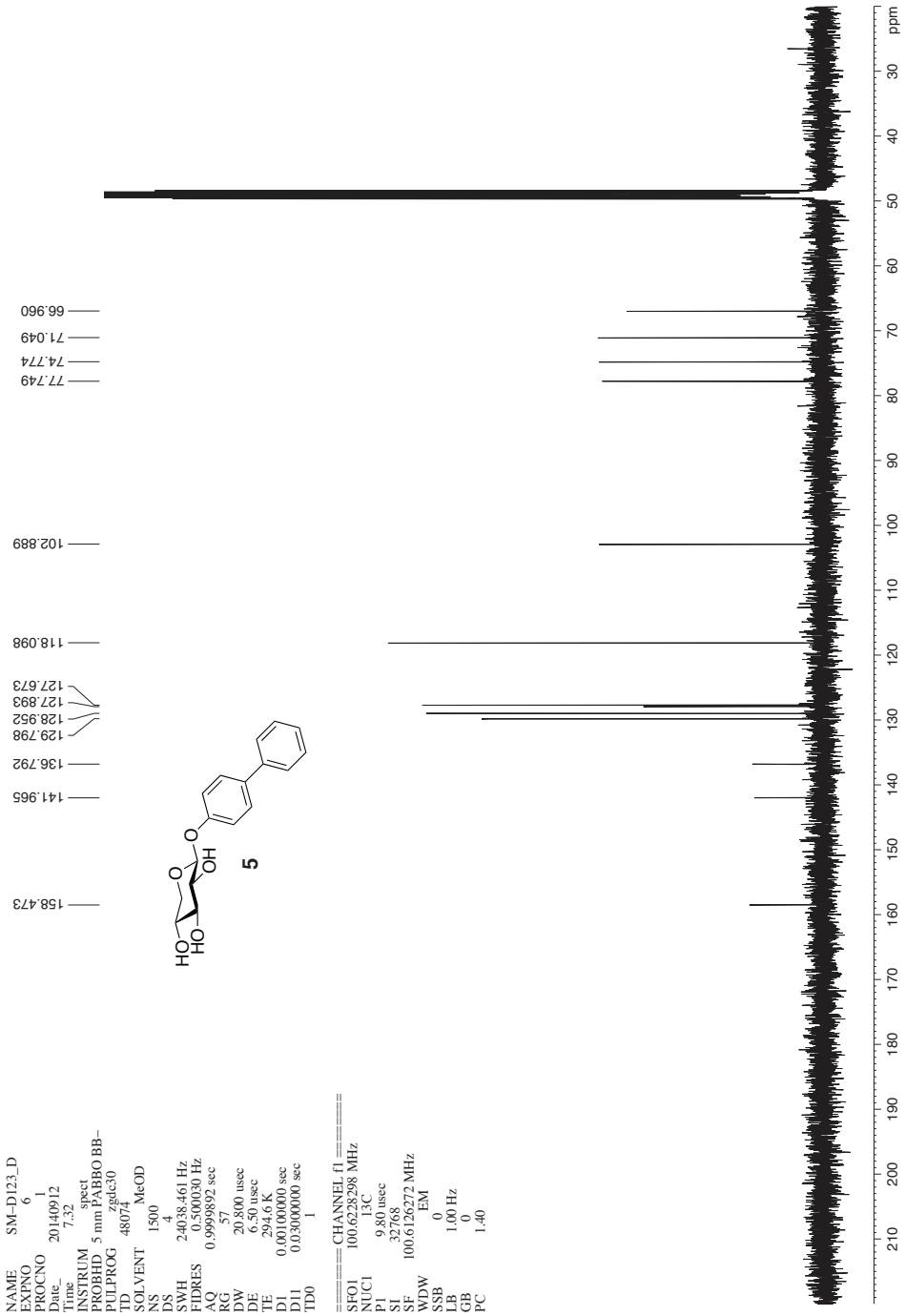
^{13}C-NMR 1-Naphthyl β-D-xylopyranoside **(4)**



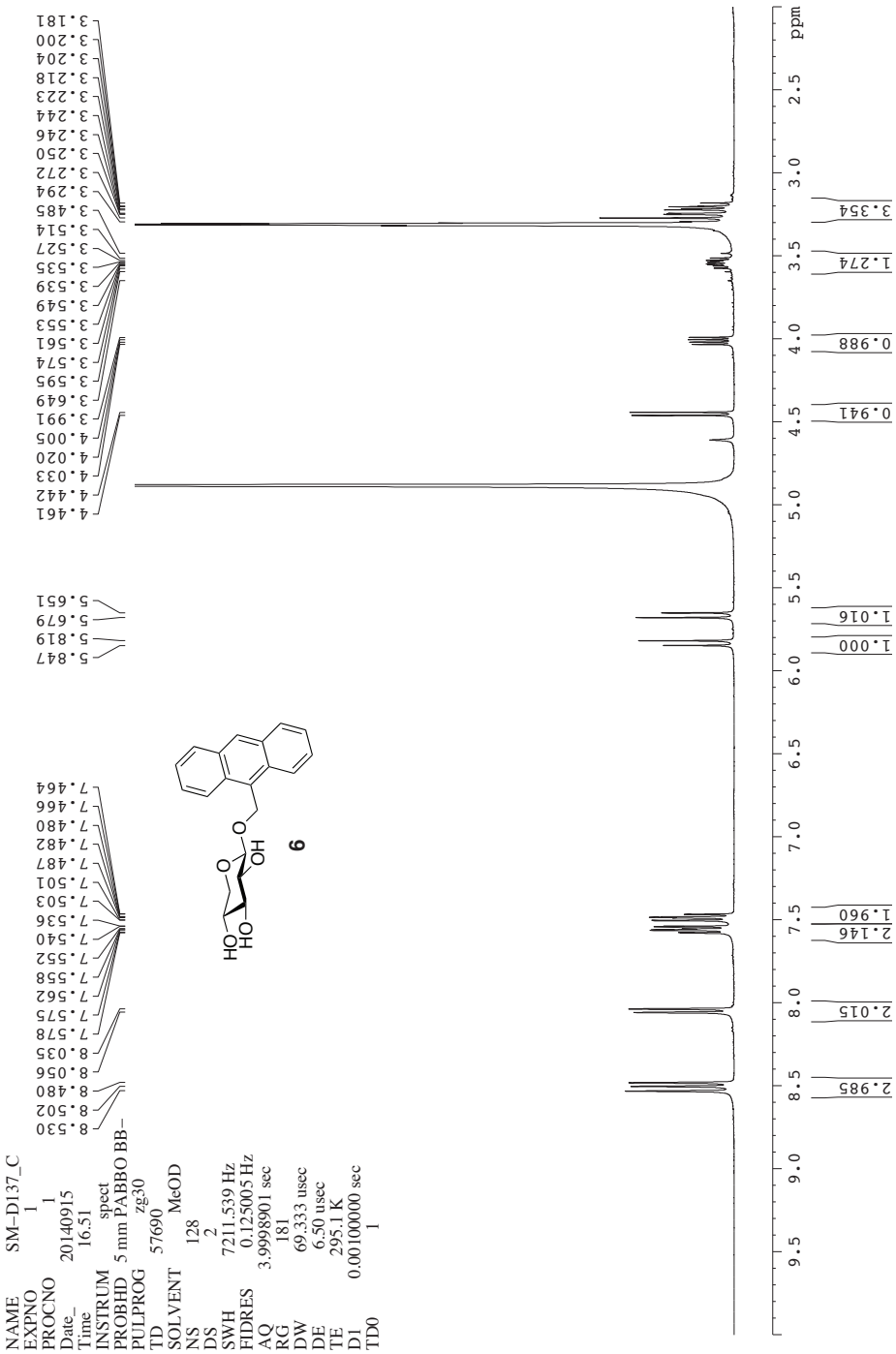
^{1}H-NMR 4-Phenylphenyl β-D-xylopyranoside **(5)**



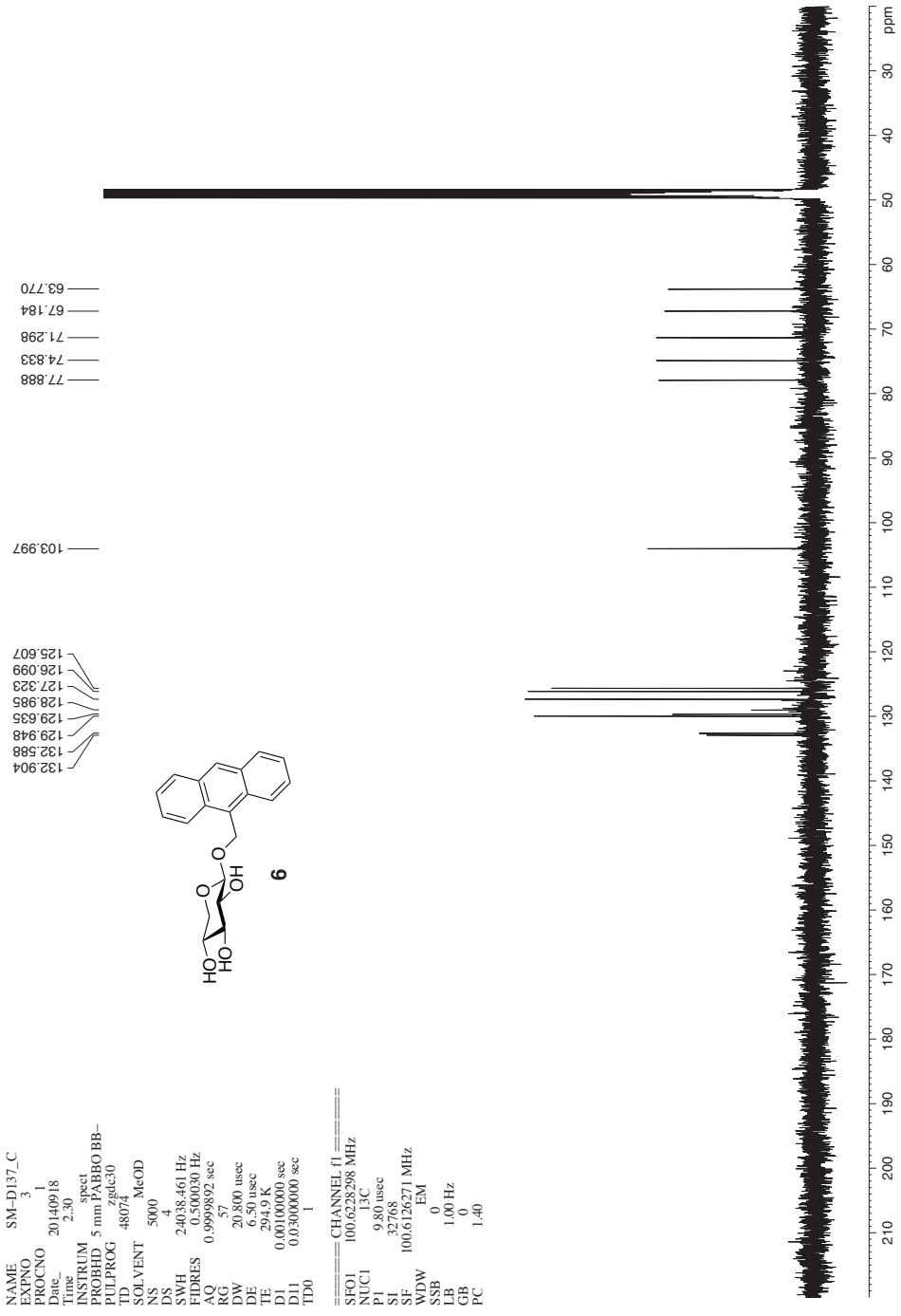
^{13}C-NMR 4-Phenylphenyl β-D-xylopyranoside **(5)**



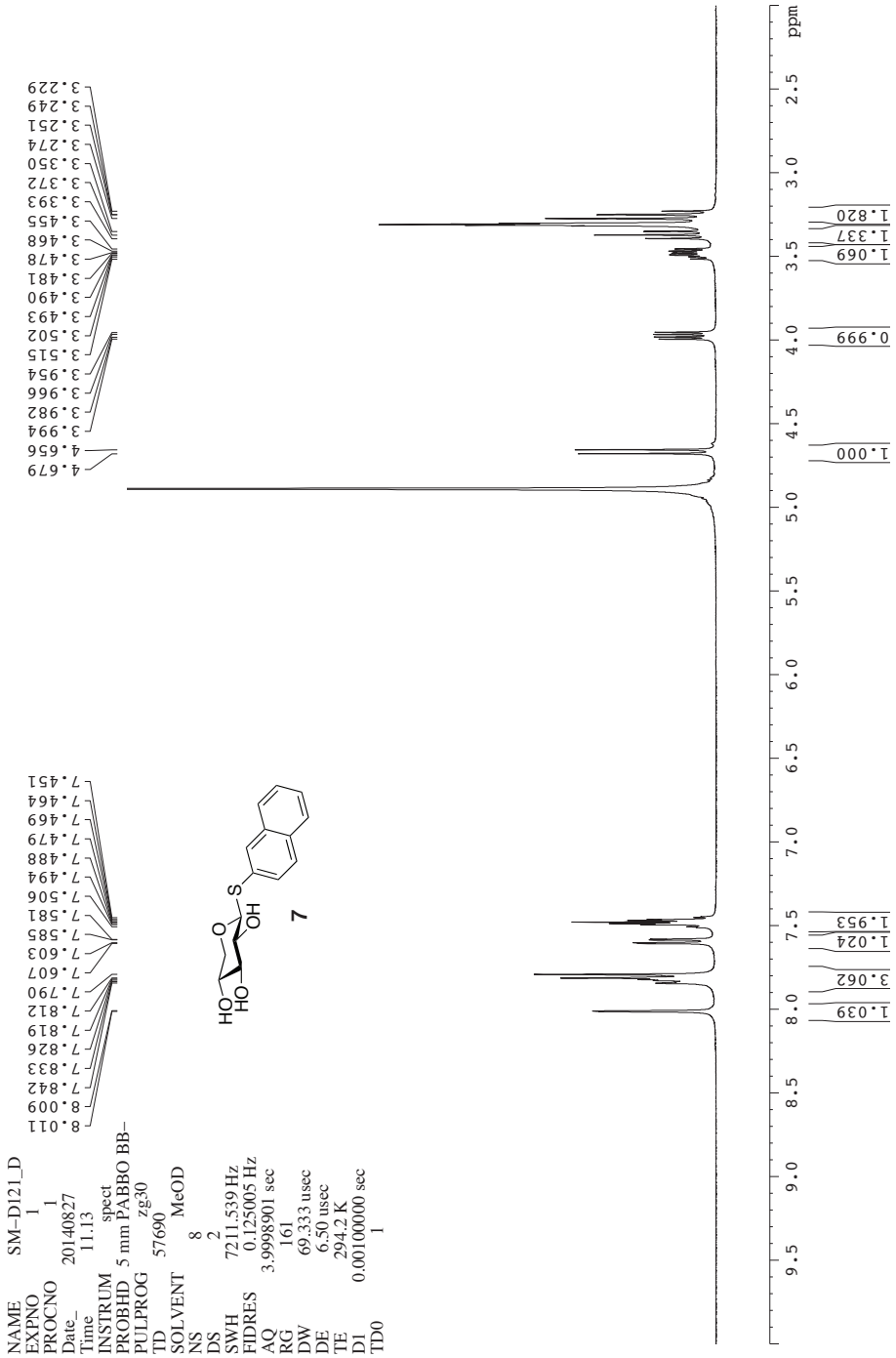
[1]H-NMR (9-Anthracene)-methyl β-D-xylopyranoside **(6)**



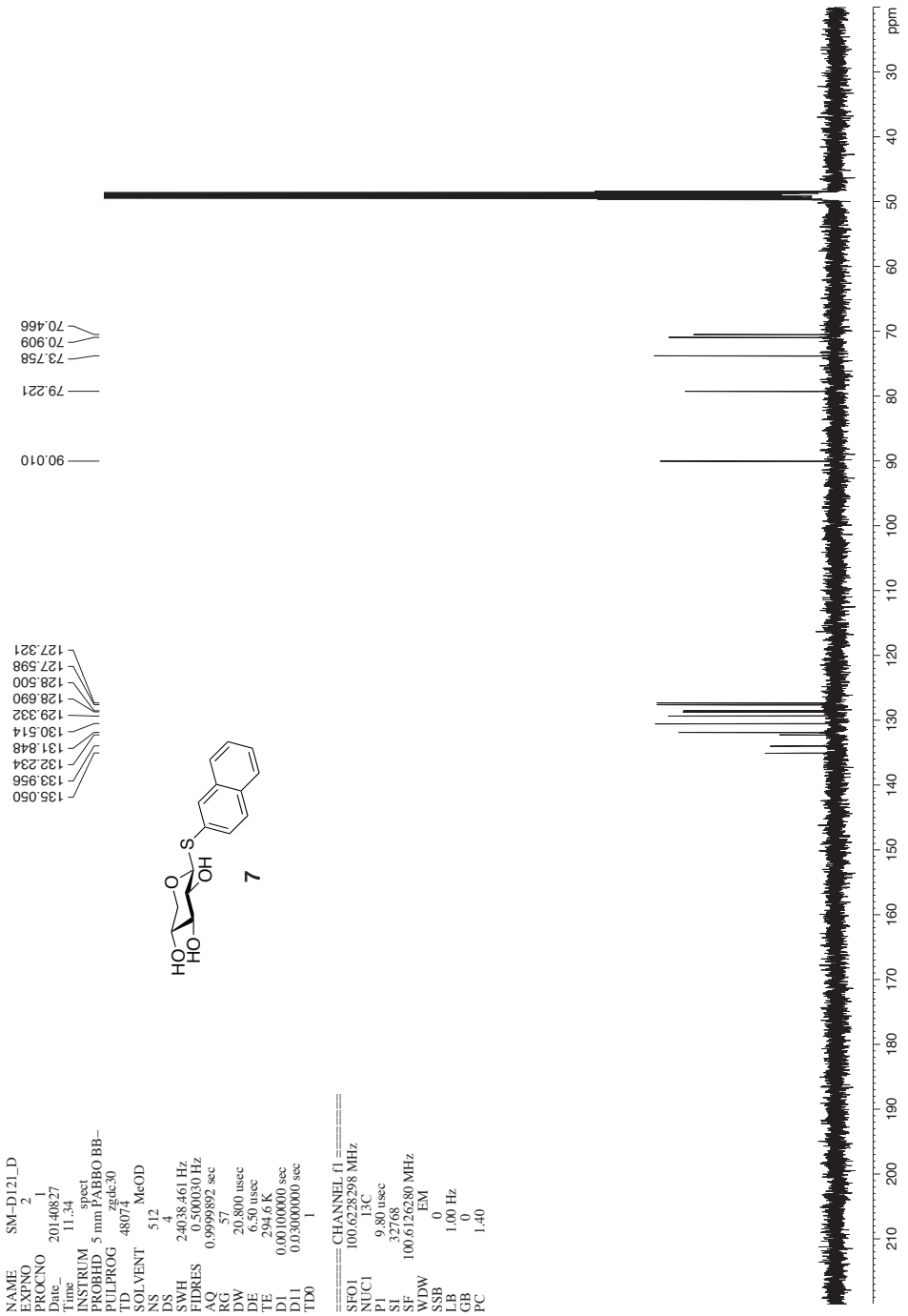
^{13}C-NMR (9-Anthracene)-methyl β-D-xylopyranoside **(6)**



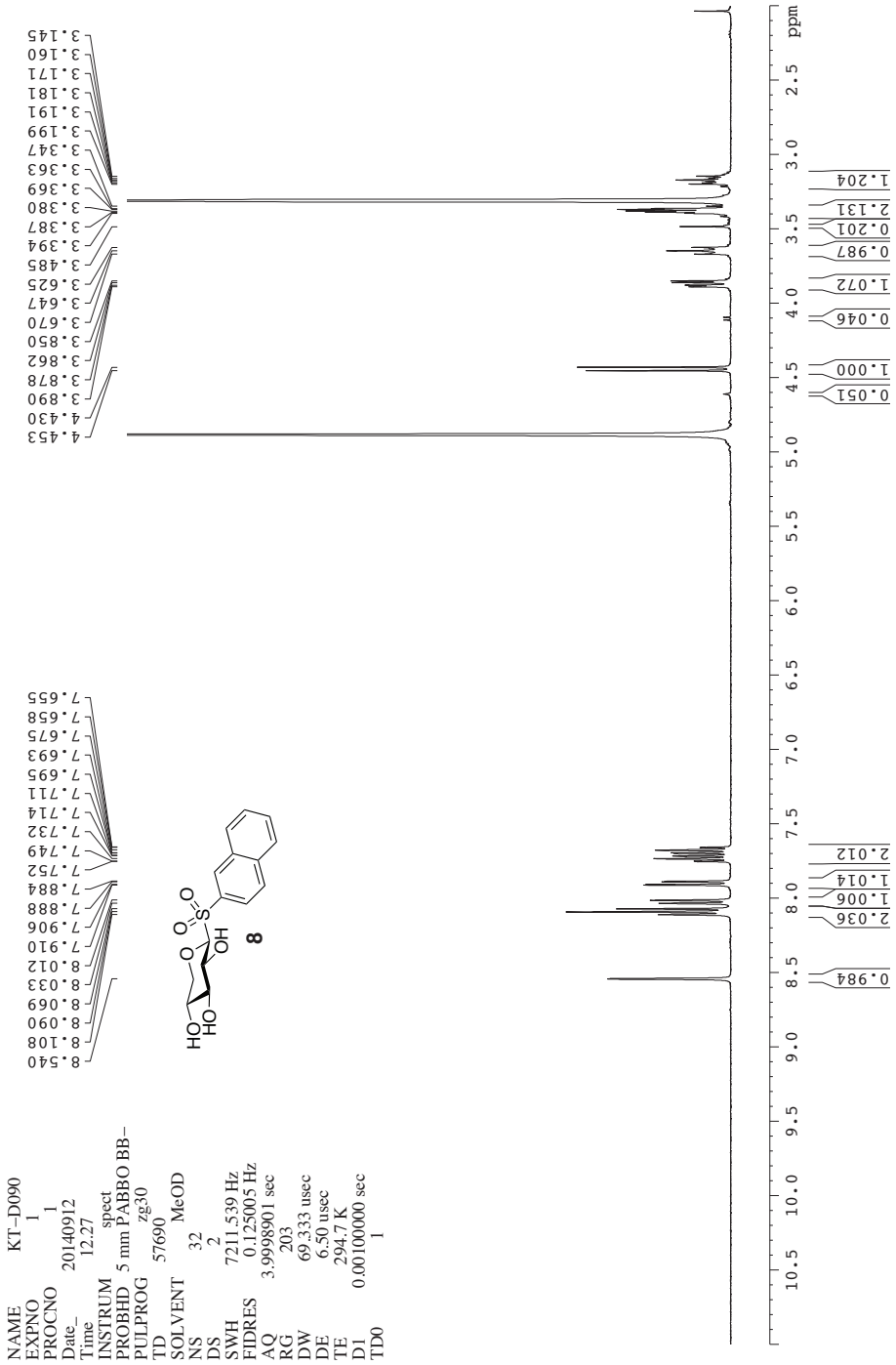
[1]H-NMR 2-Naphthyl 1-thio-β-D-xylopyranoside **(7)**



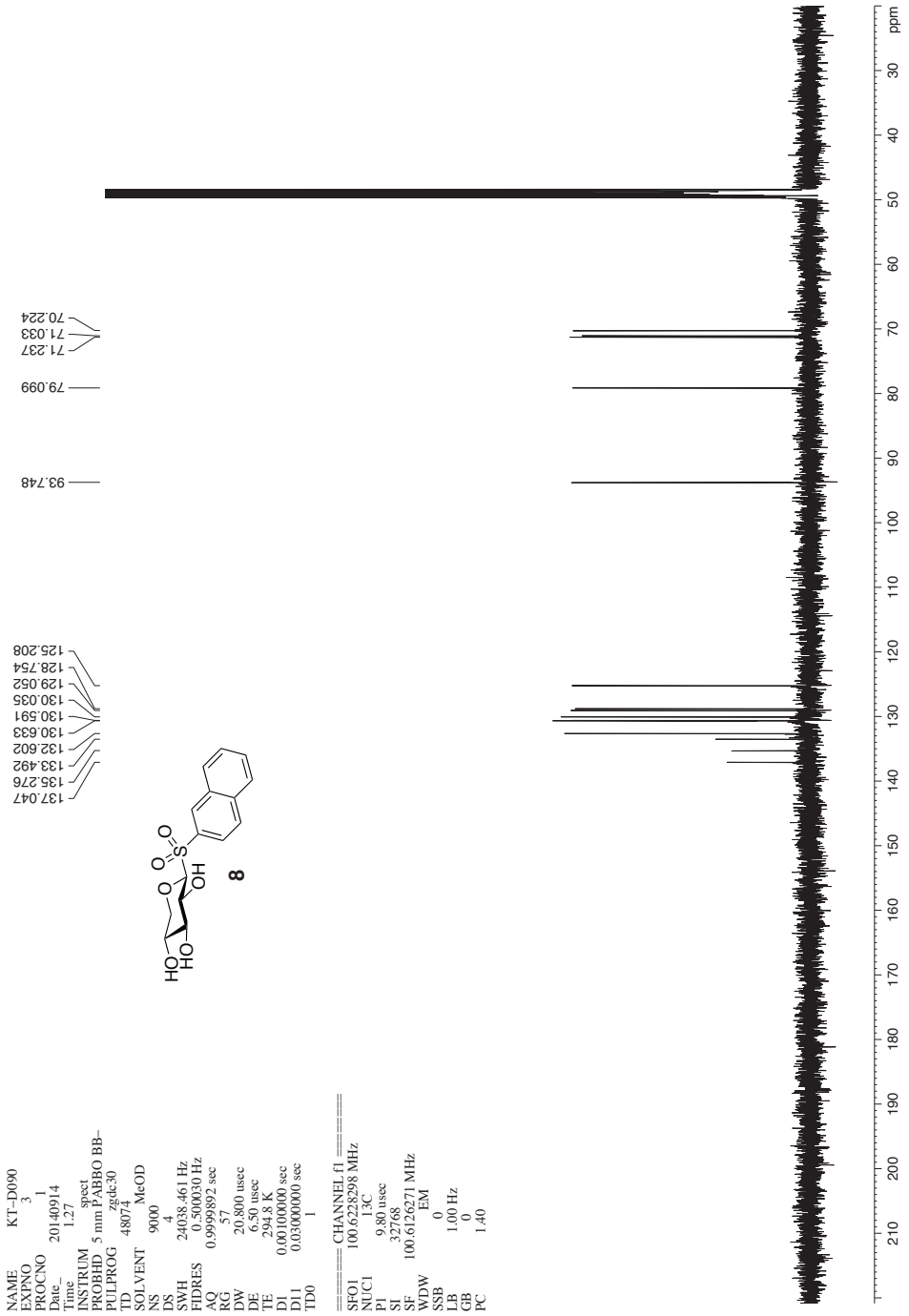
^{13}C-NMR 2-Naphthyl 1-thio-β-D-xylopyranoside **(7)**



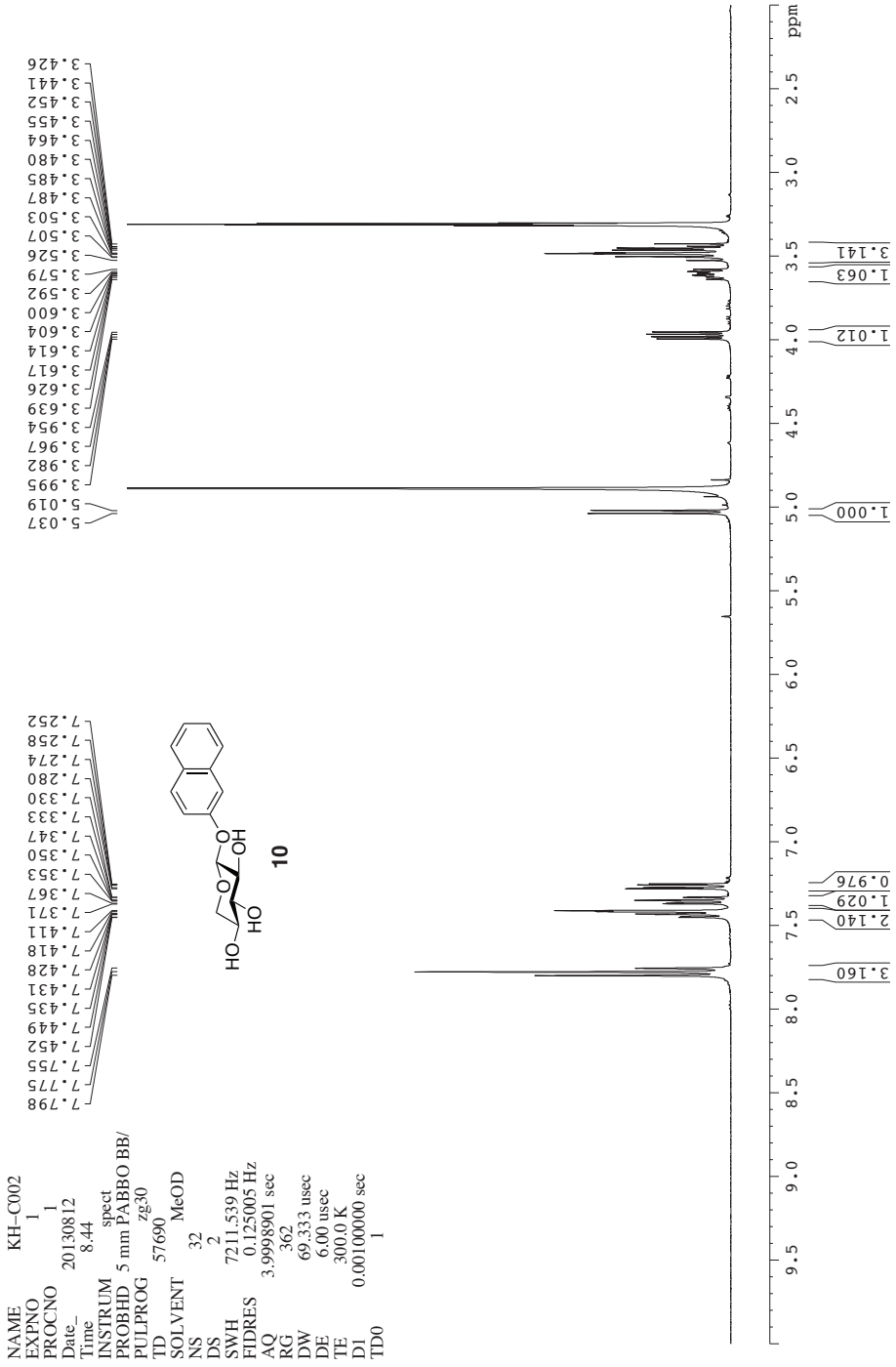
^{1}H-NMR 2-Naphthyl β-D-xylopyranosyl sulfone **(8)**



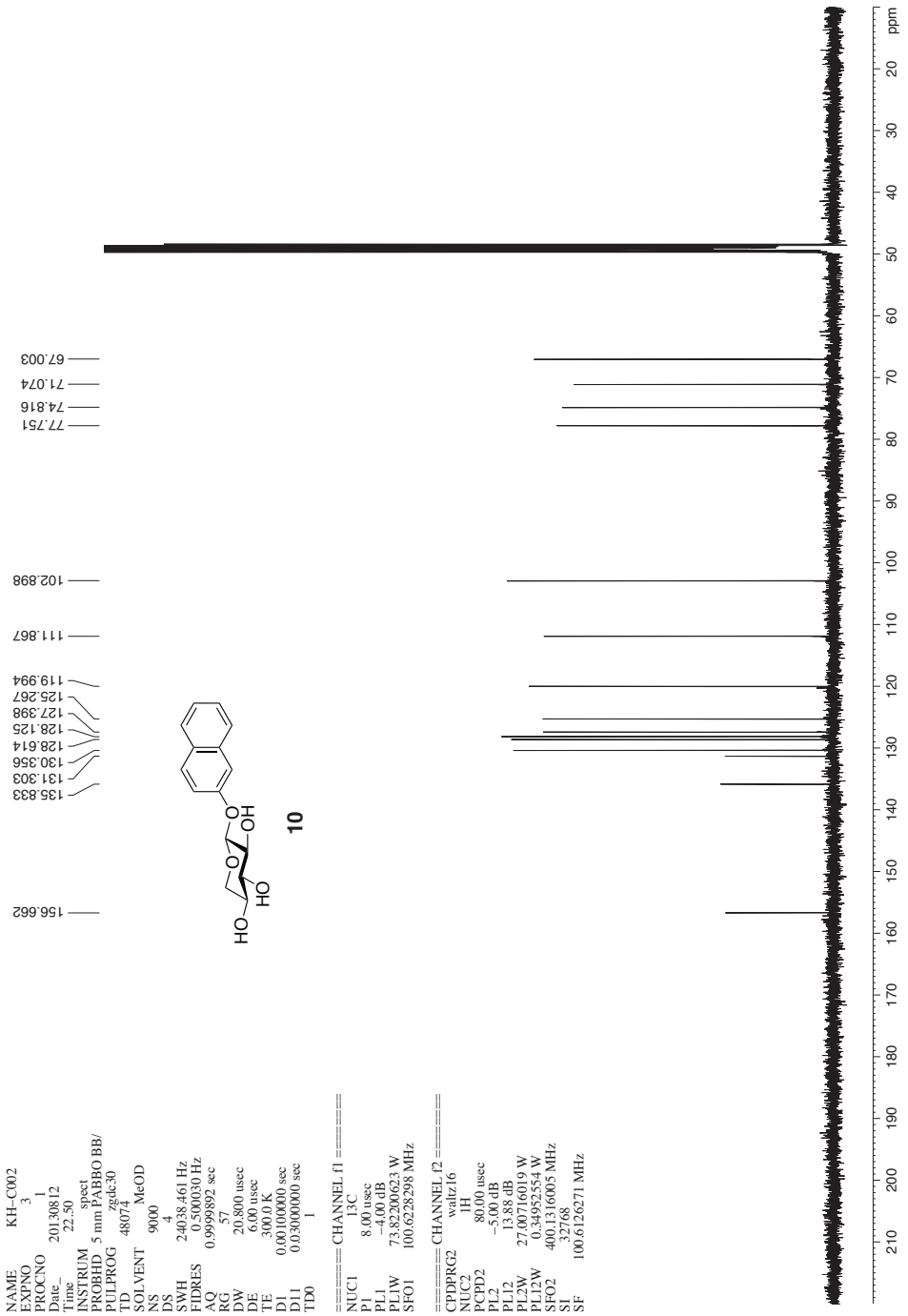
^{13}C-NMR 2-Naphthyl β-D-xylopyranosyl sulfone **(8)**



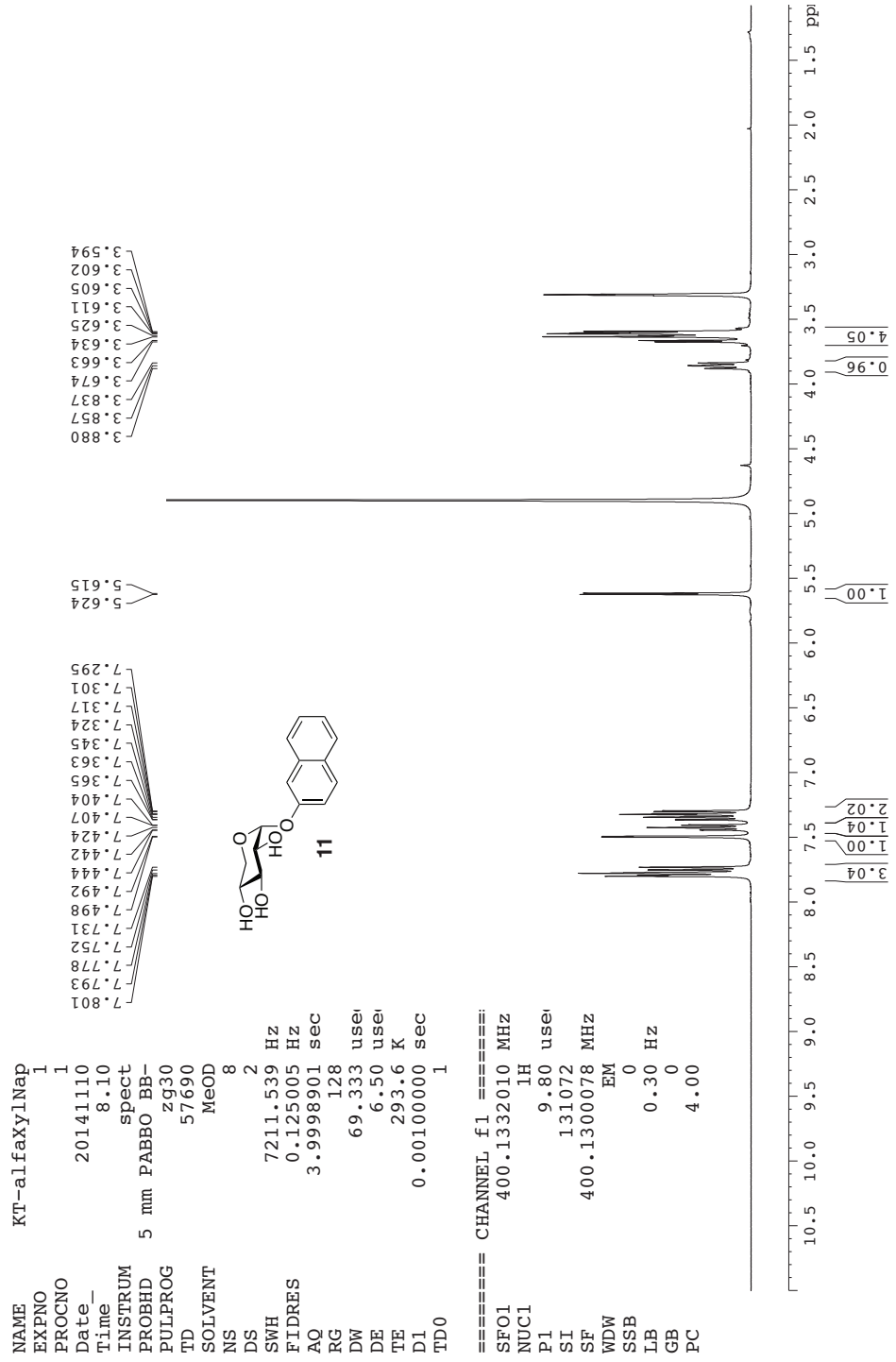
^{1}H-NMR 2-Naphthyl β-L-xylopyranoside **(10)**



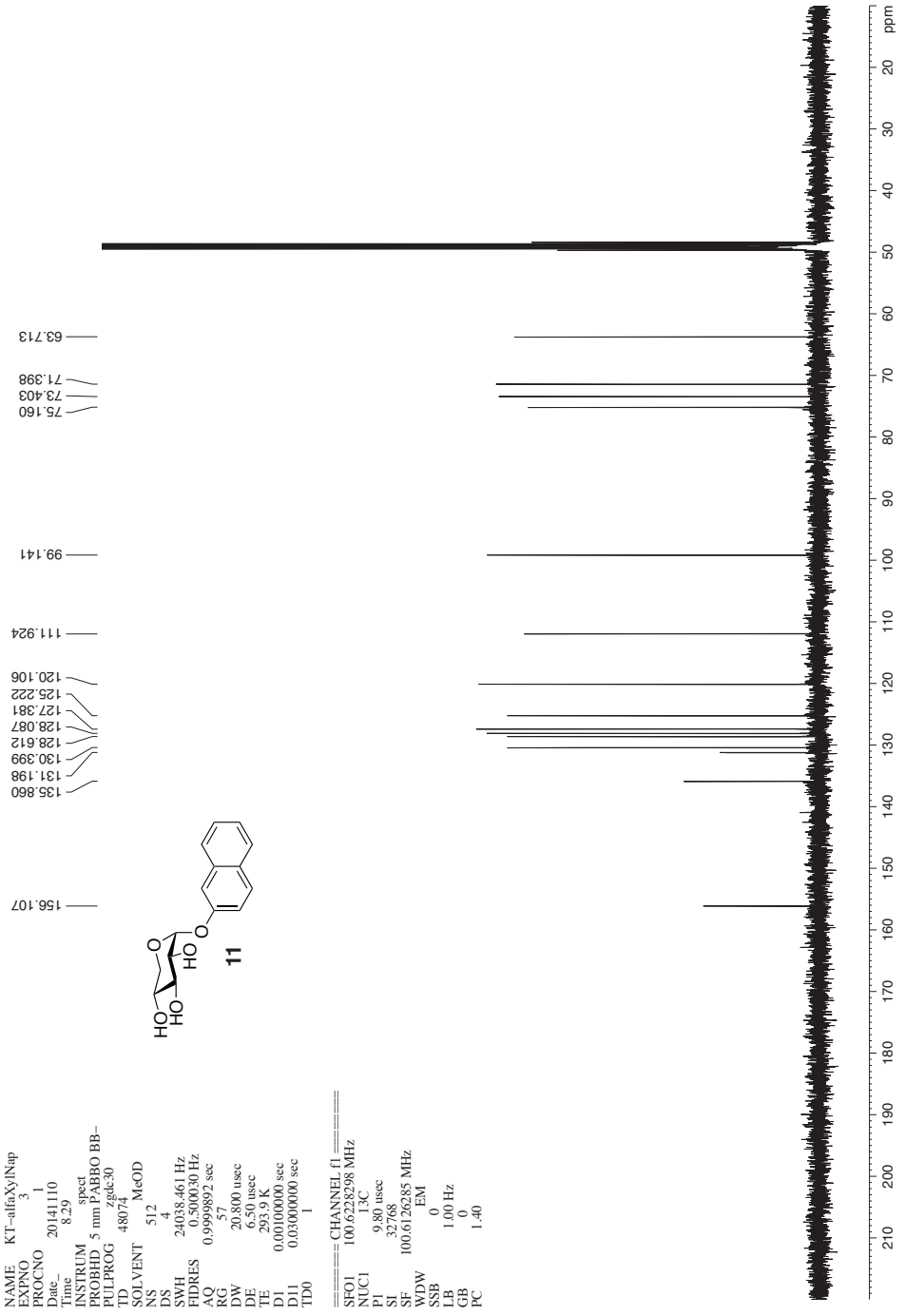
^{13}C-NMR 2-Naphthyl β-L-xylopyranoside **(10)**



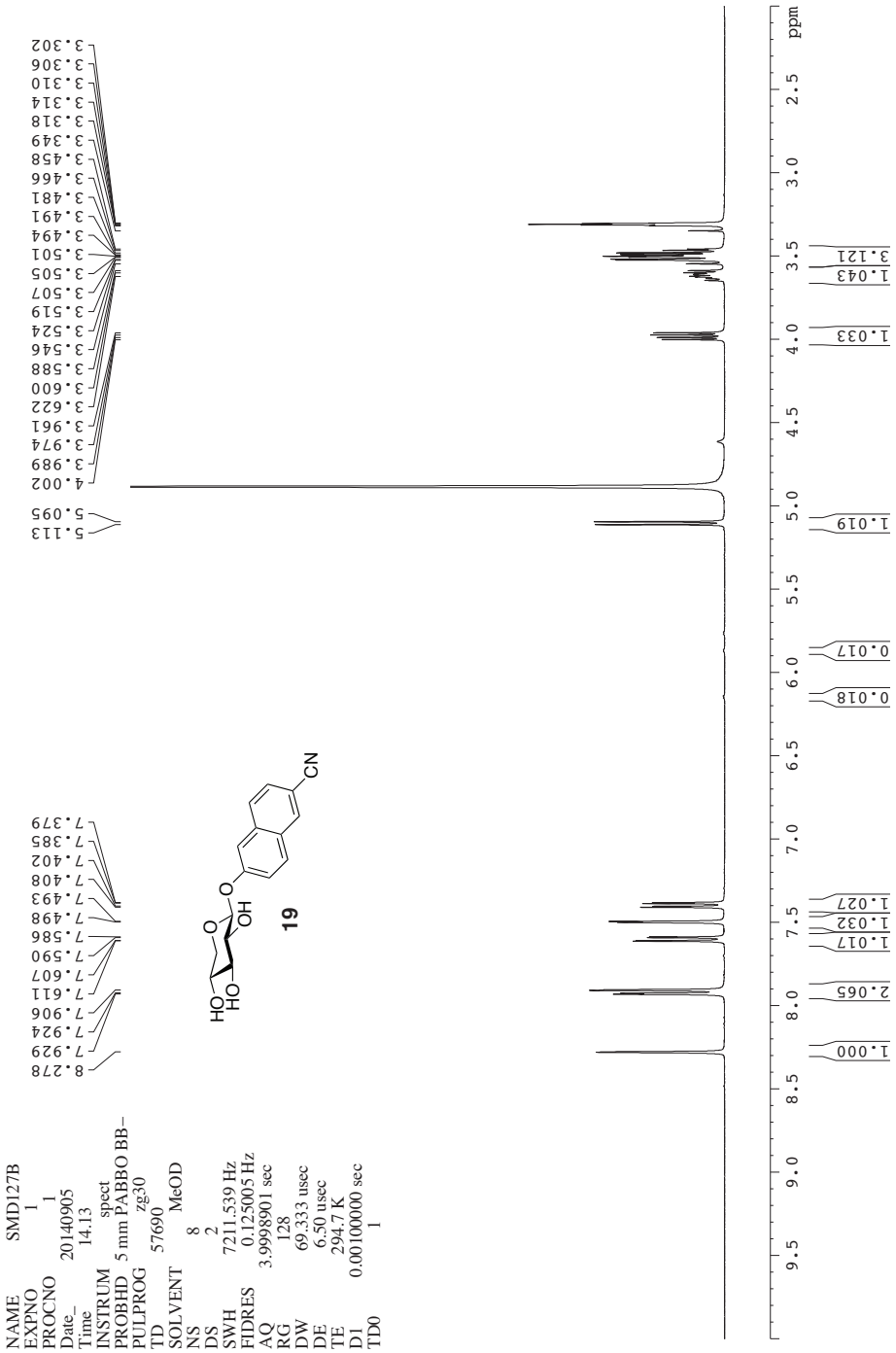
[1]H-NMR 2-Naphthyl α-D-xylopyranoside **(11)**



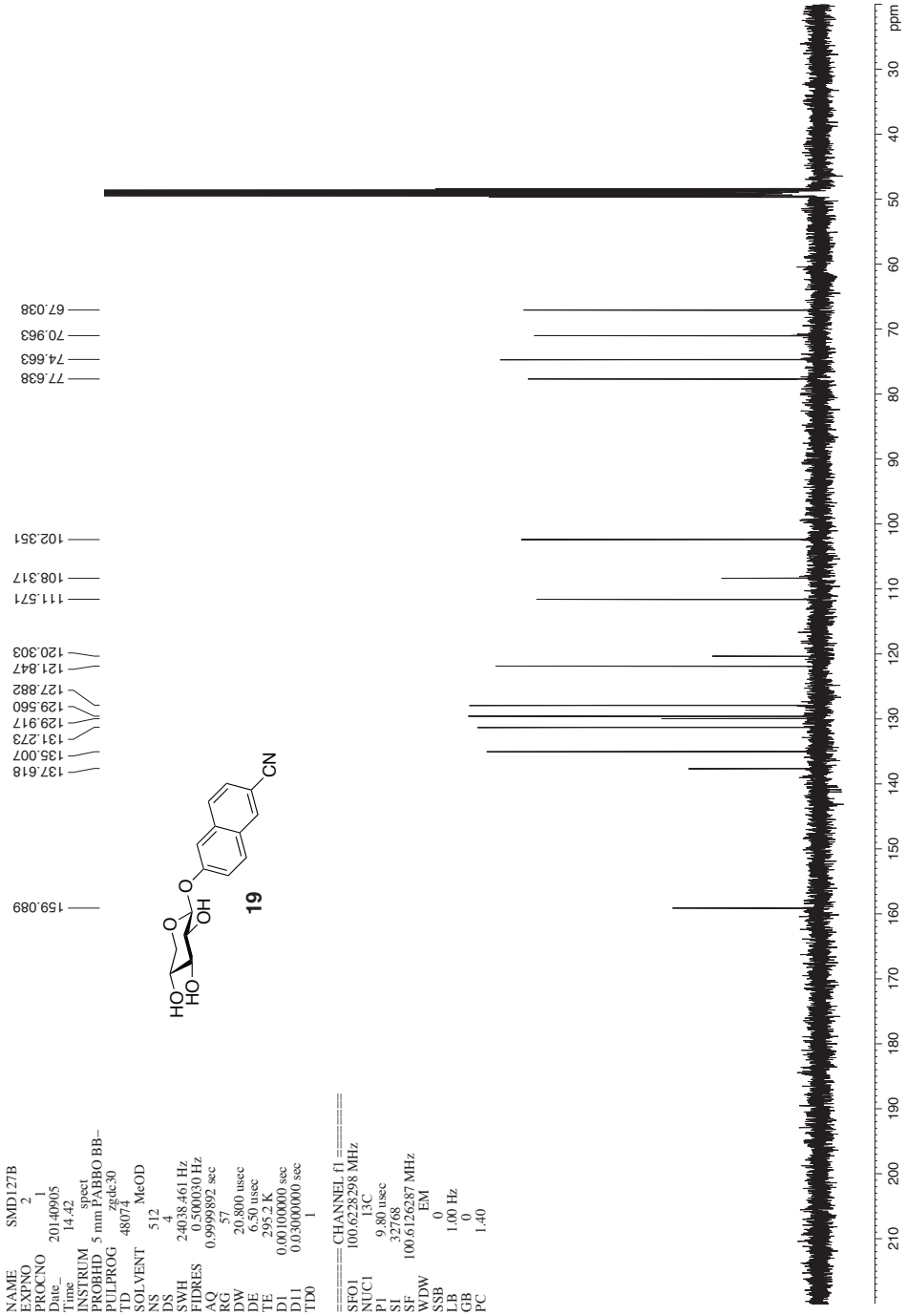
^{13}C-NMR 2-Naphthyl α-D-xylopyranoside **(11)**



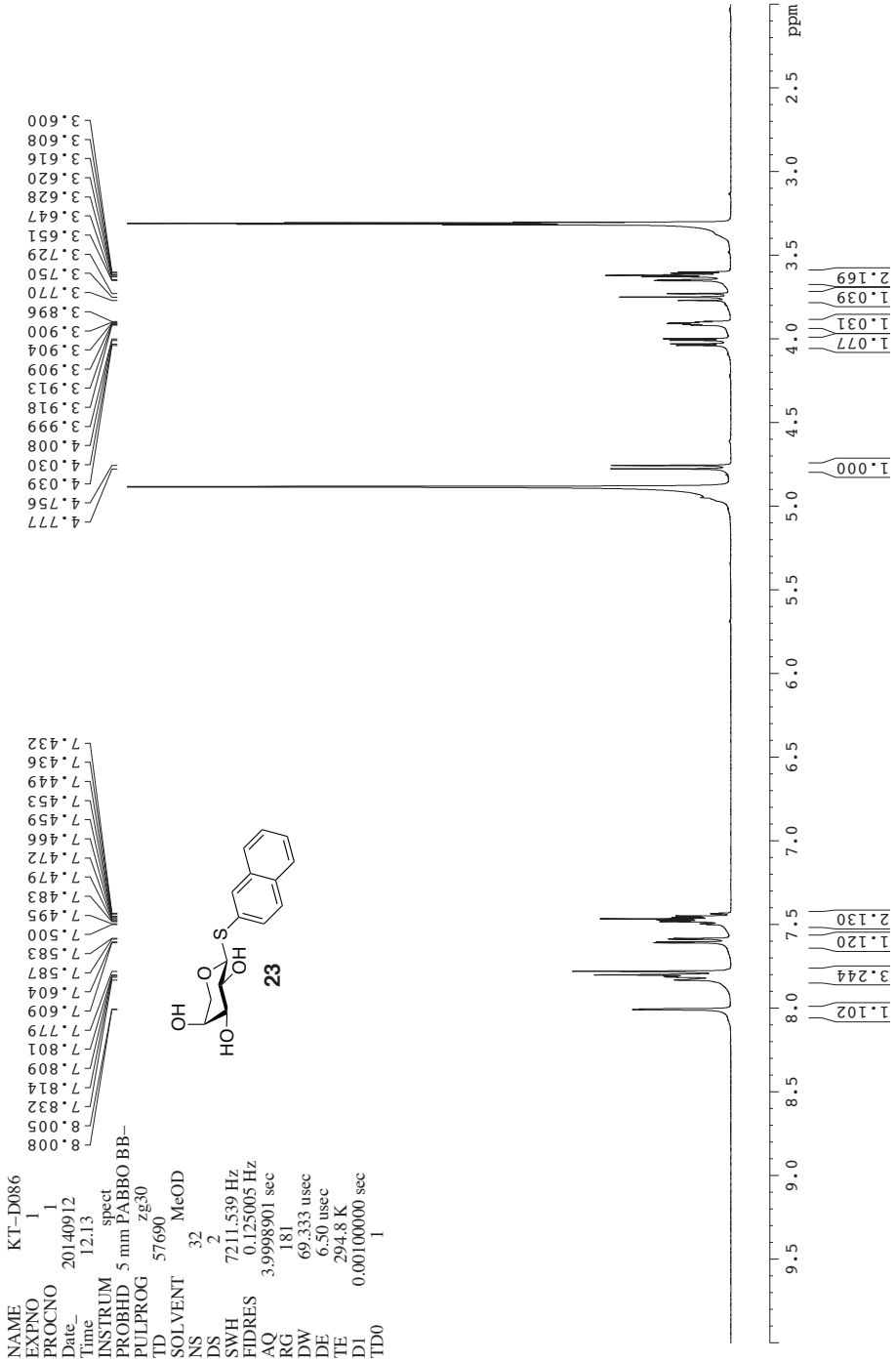
[1]H-NMR 2-(6-Cyano-naphthyl) β-D-xylopyranoside **(19)**



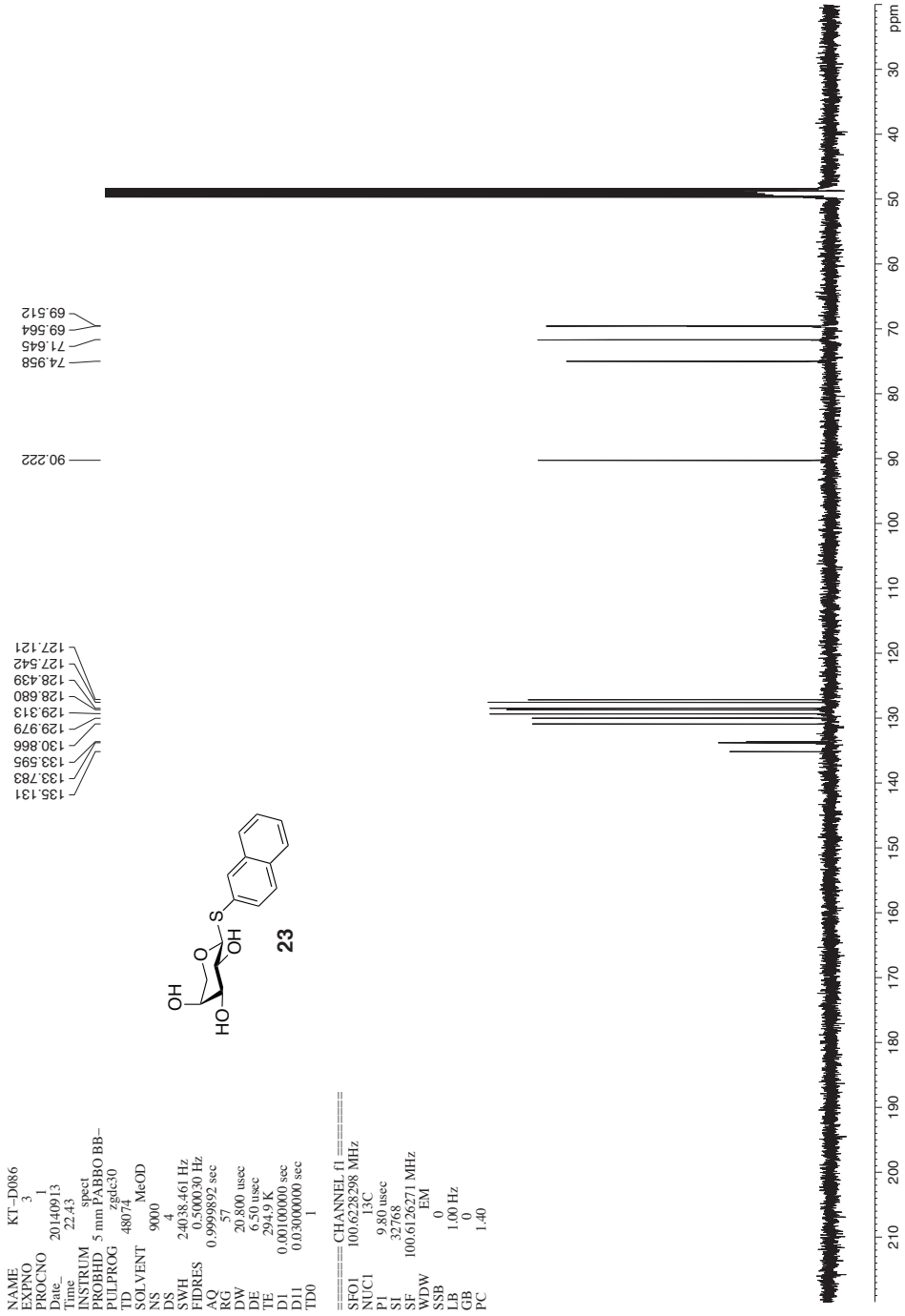
^{13}C-NMR 2-(6-Cyano-naphthyl) β-D-xylopyranoside **(19)**



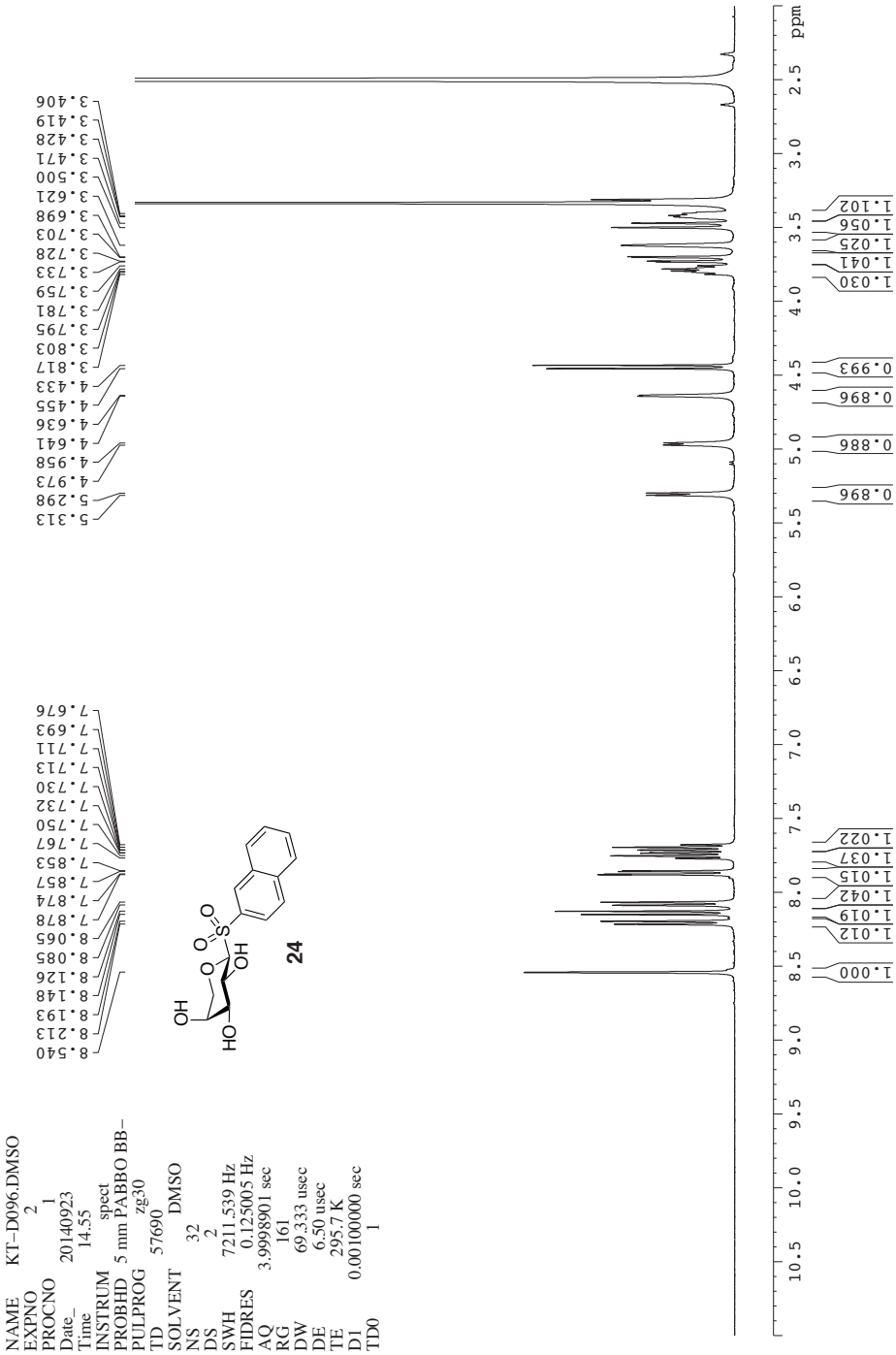
^{1}H-NMR 2-Naphthyl 1-thio-α-L-arabinopyranoside **(23)**



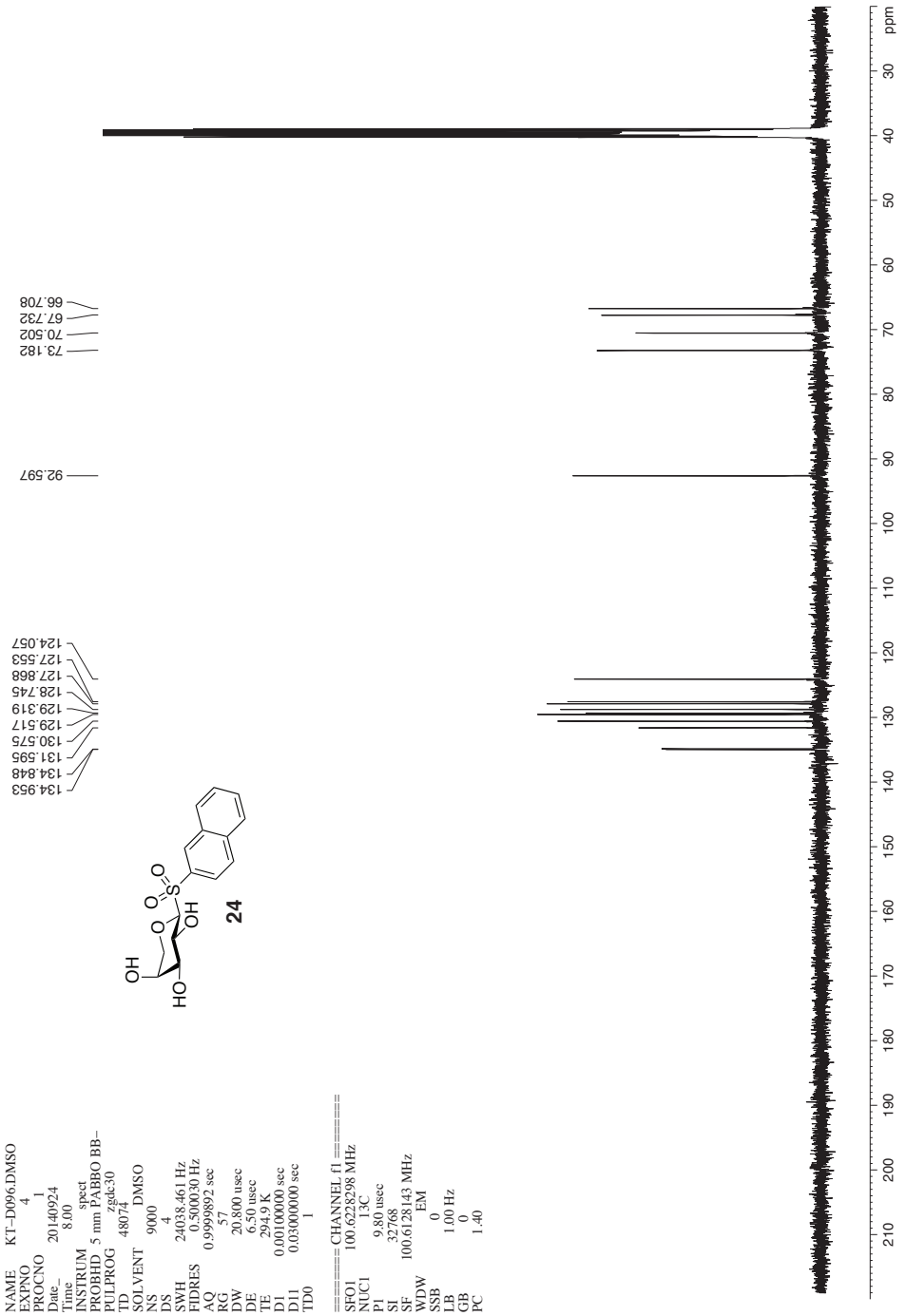
^{13}C-NMR 2-Naphthyl 1-thio-α-L-arabinopyranoside **(23)**



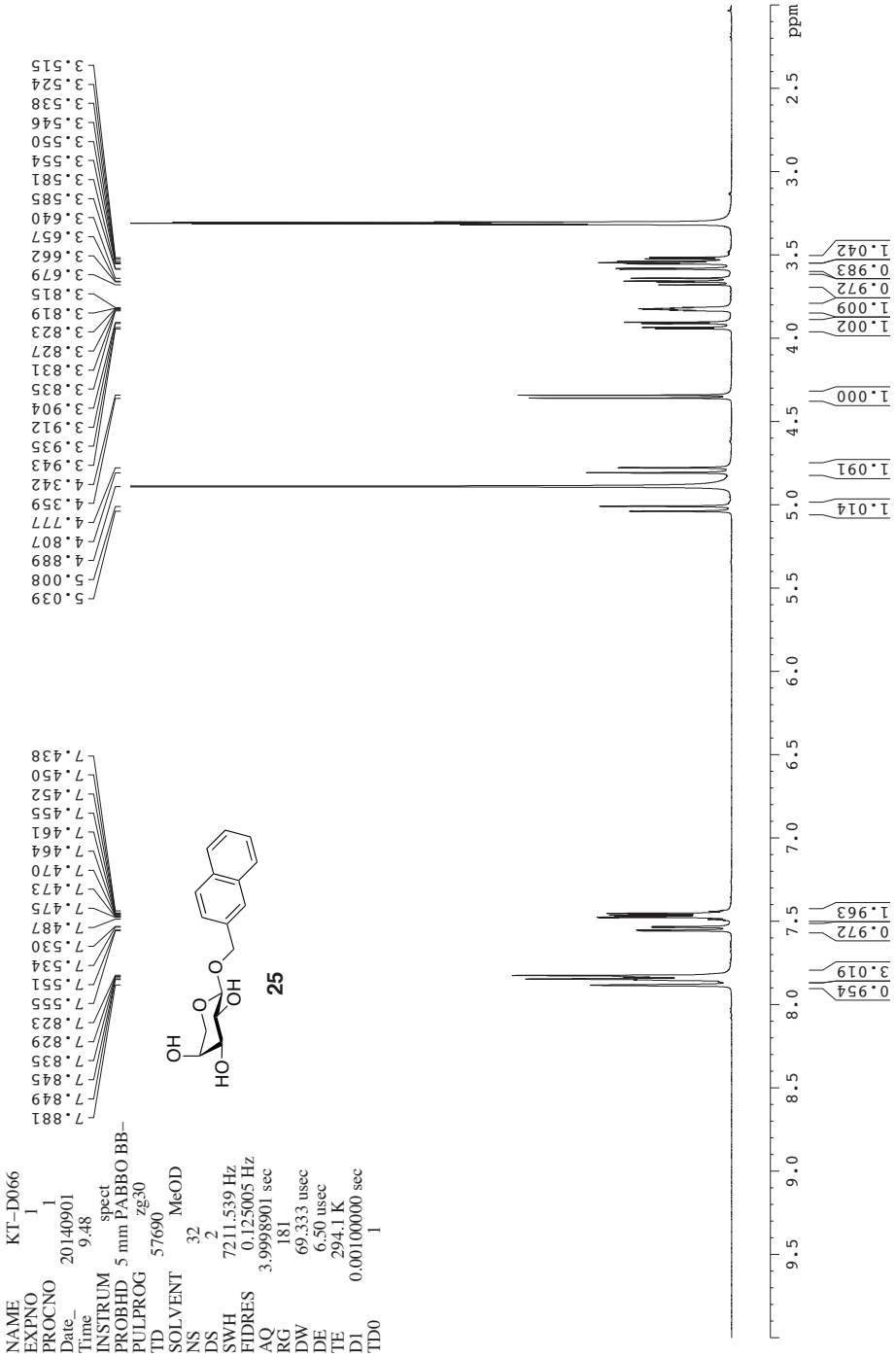
[1]H-NMR α-L-Arabinopyranoside 2-naphthyl sulfone **(24)**



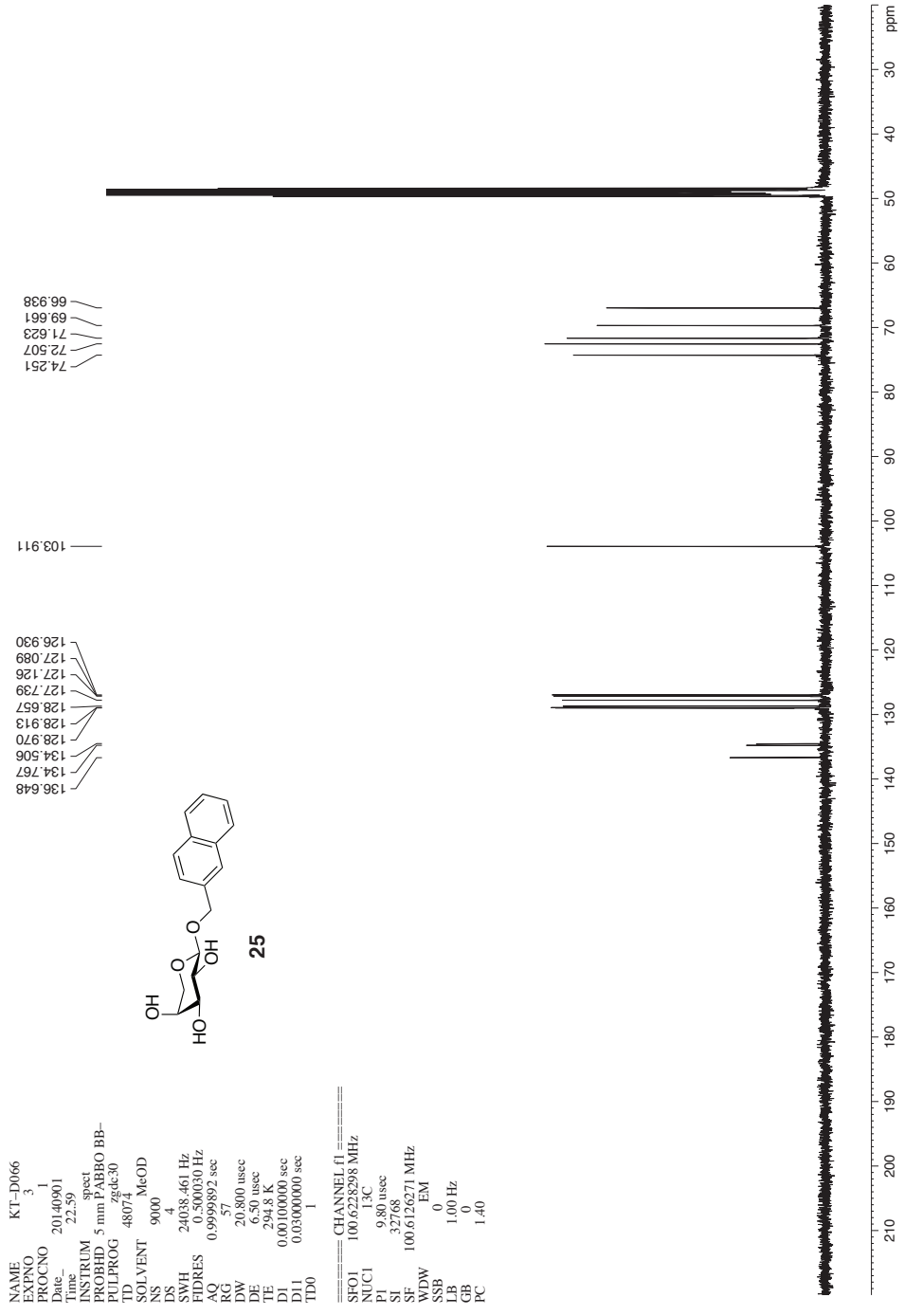
^{13}C-NMR α-L-Arabinopyranoside 2-naphthyl sulfone **(24)**



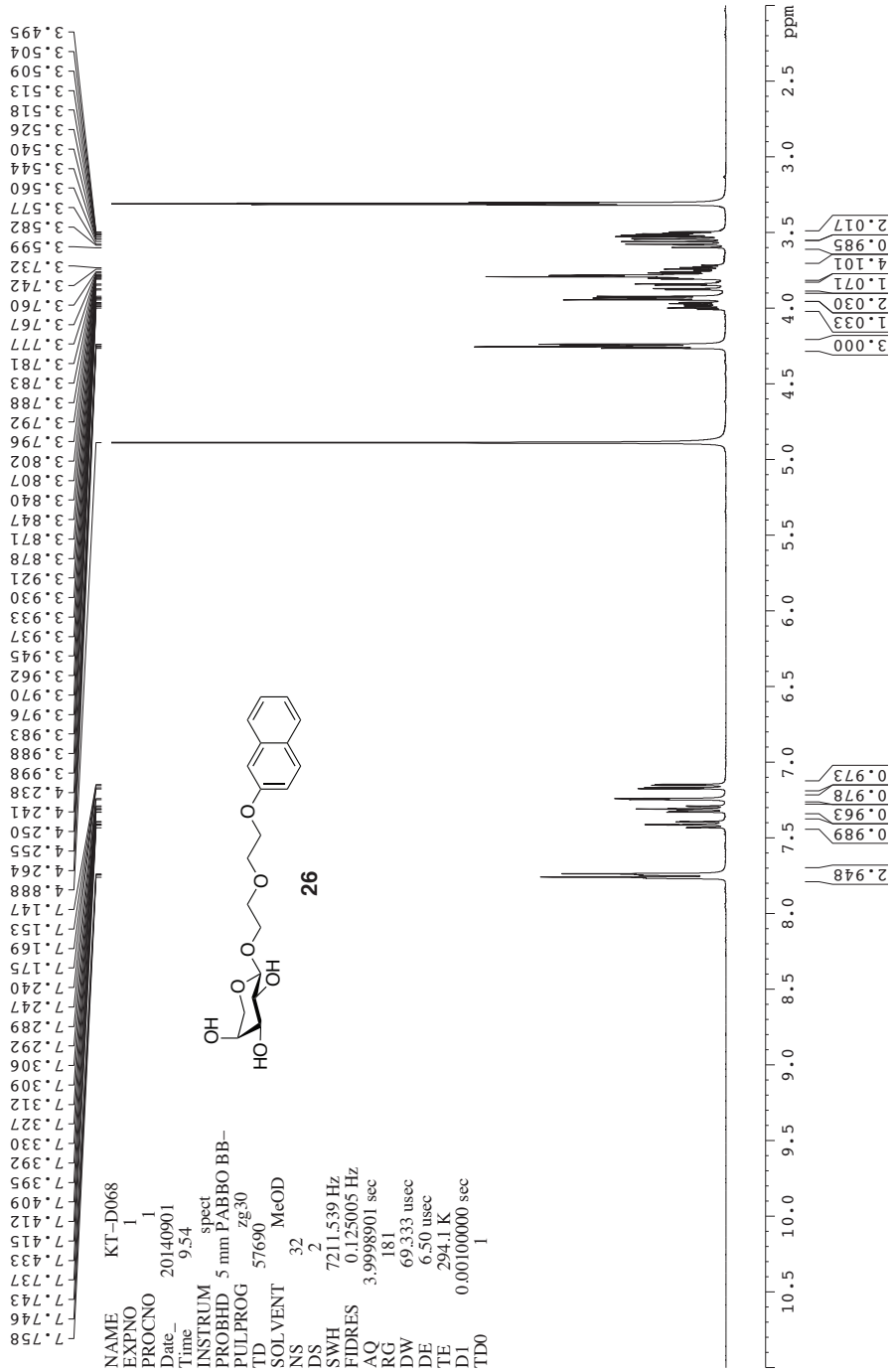
^{1}H-NMR (2-Naphthyl)-methyl α-L-arabinopyranoside **(25)**



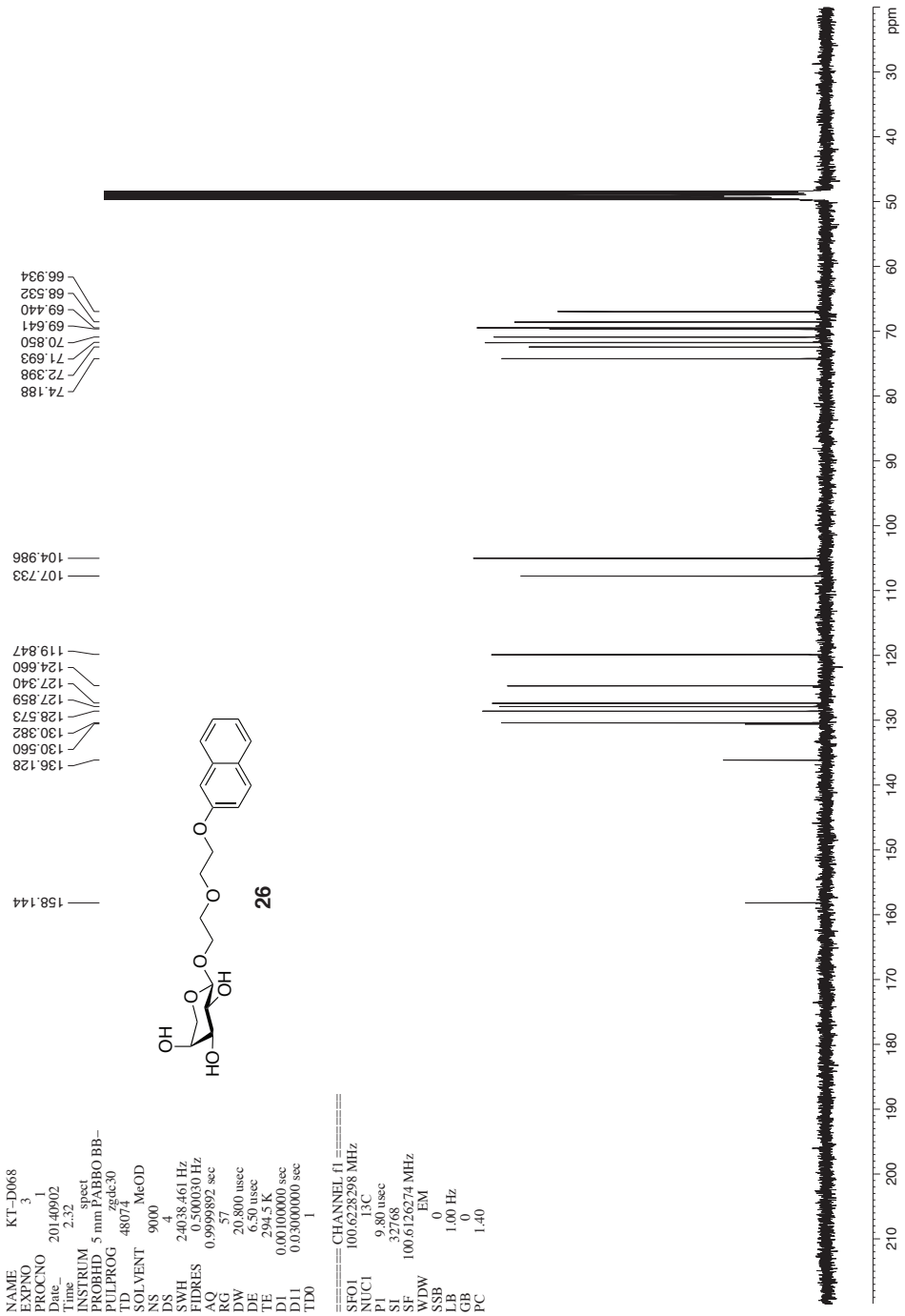
^{13}C-NMR (2-Naphthyl)-methyl α-L-arabinopyranoside **(25)**



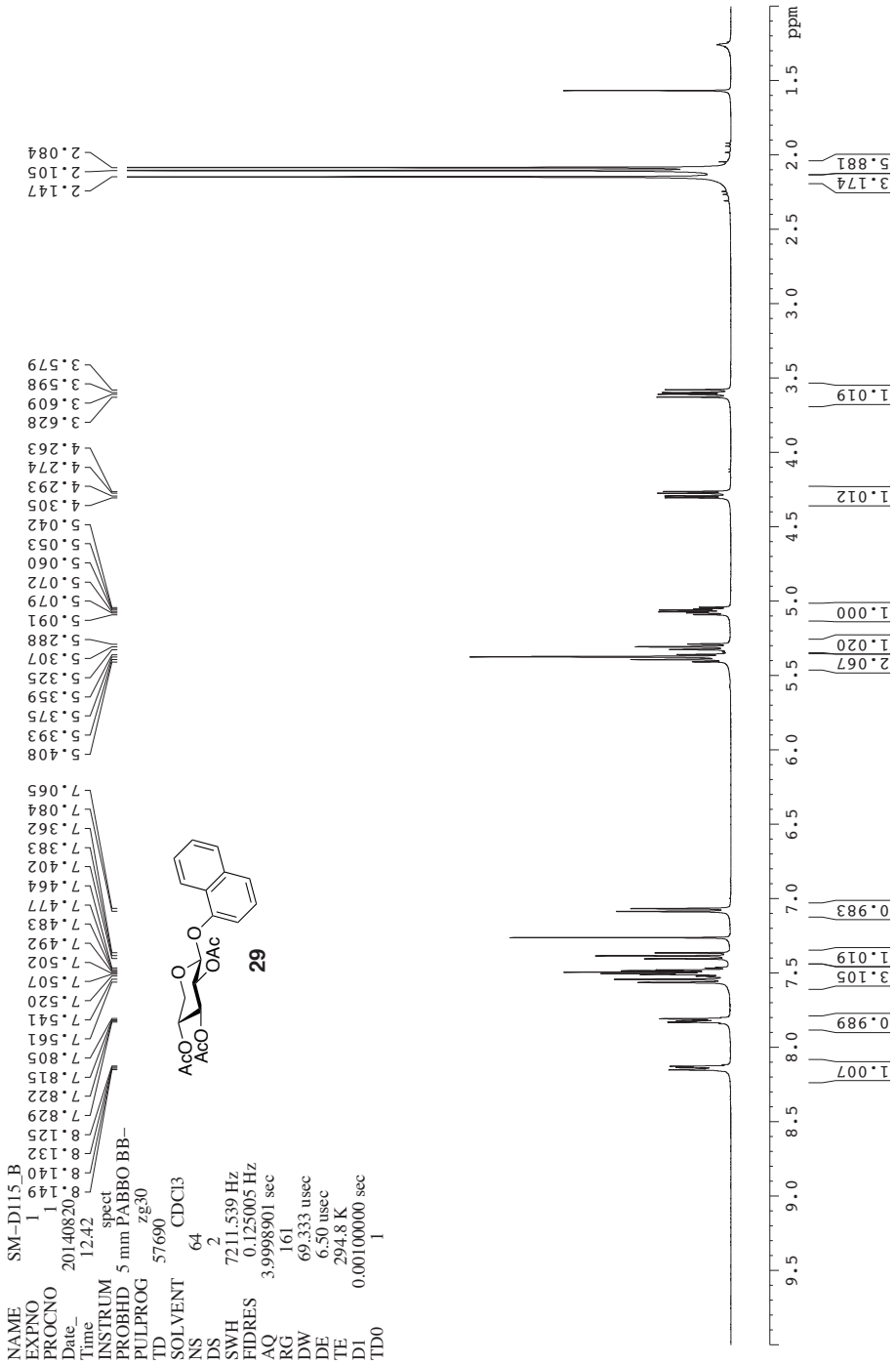
^{1}H-NMR 5-(2-Naphthoxy)-3-oxo-1-pentyl α-L-arabinopyranoside **(26)**



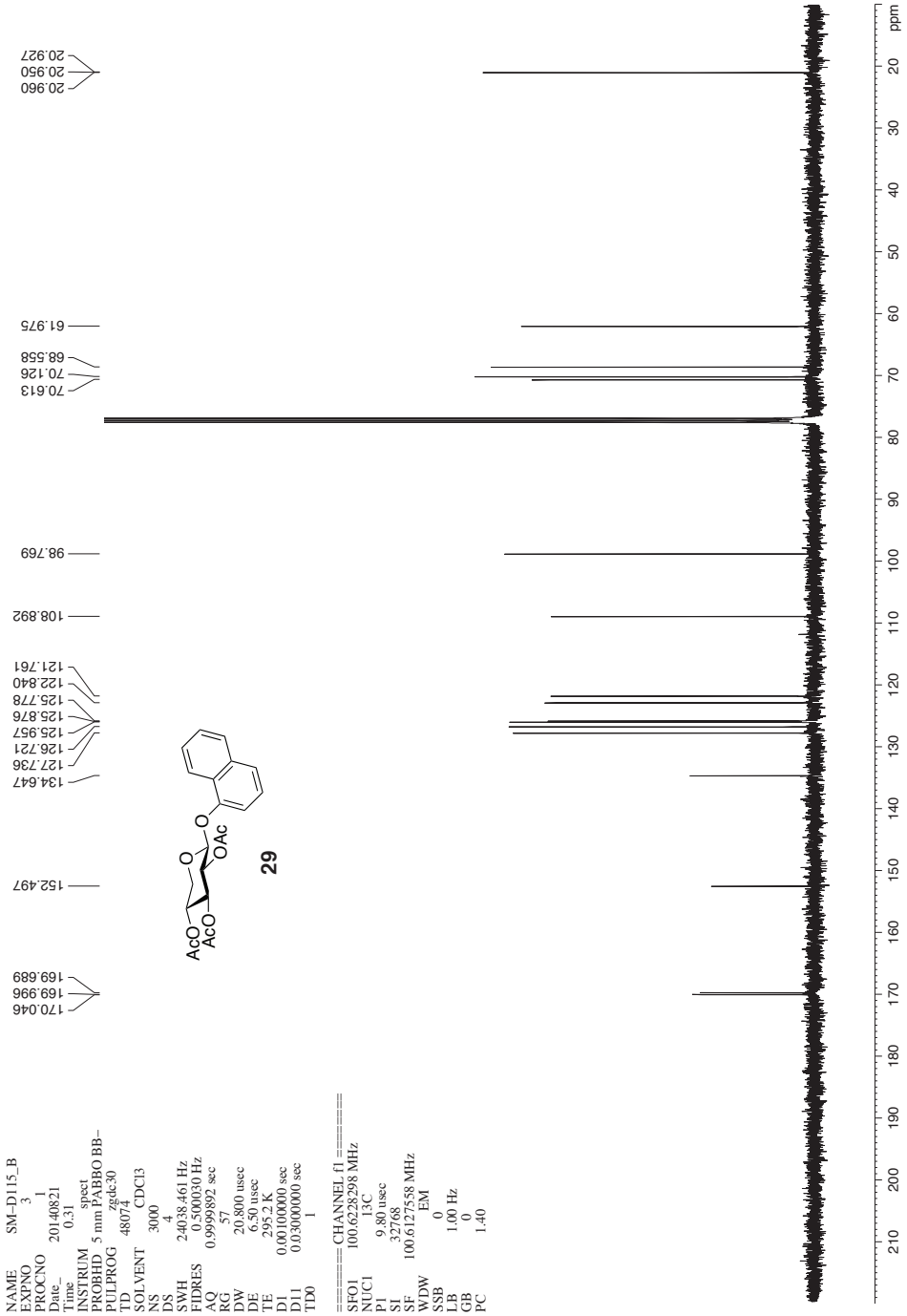
^{13}C-NMR 5-(2-Naphthoxy)-3-oxo-1-pentyl α-L-arabinopyranoside **(26)**



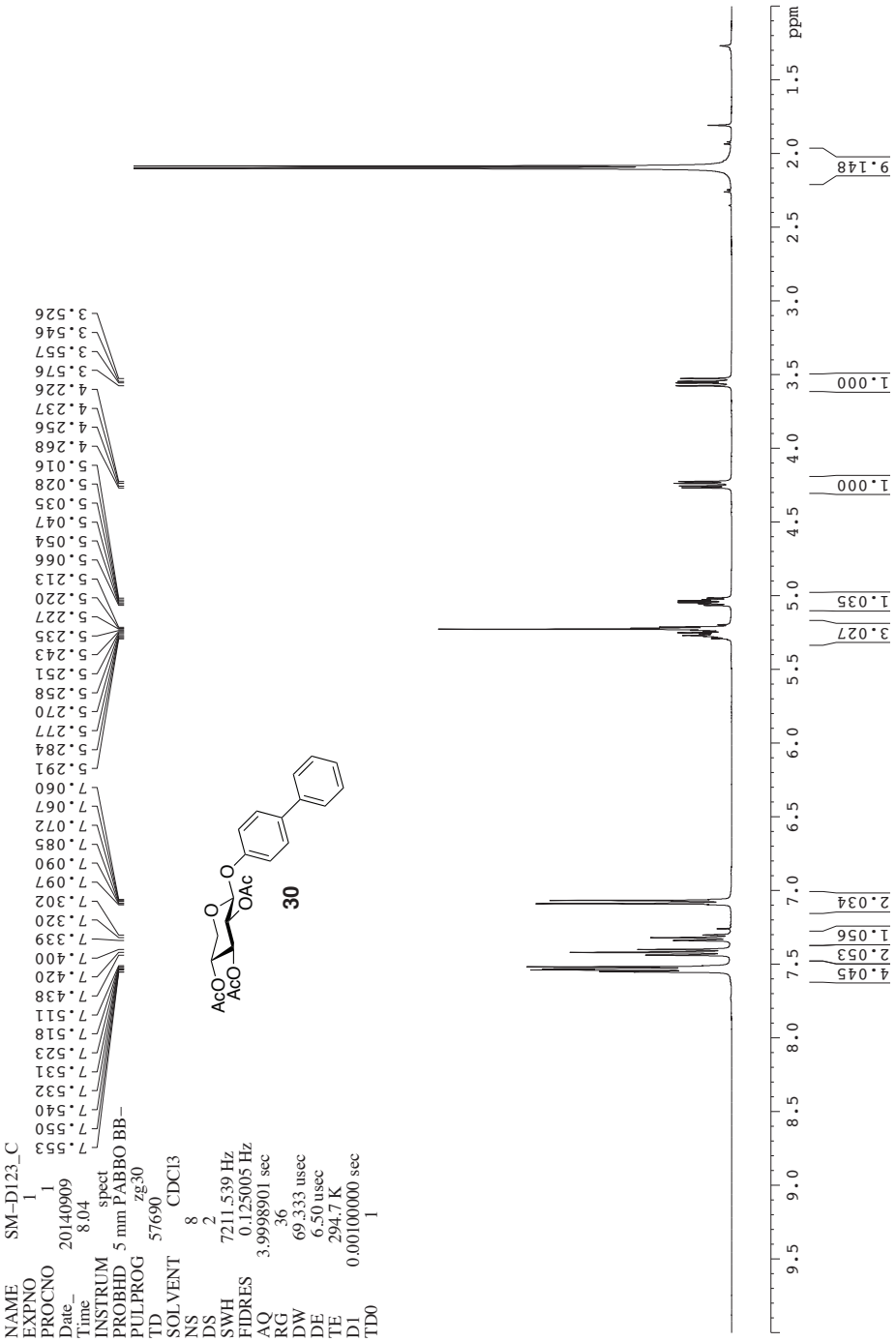
[1]H-NMR 1-Naphthyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranoside **(29)**



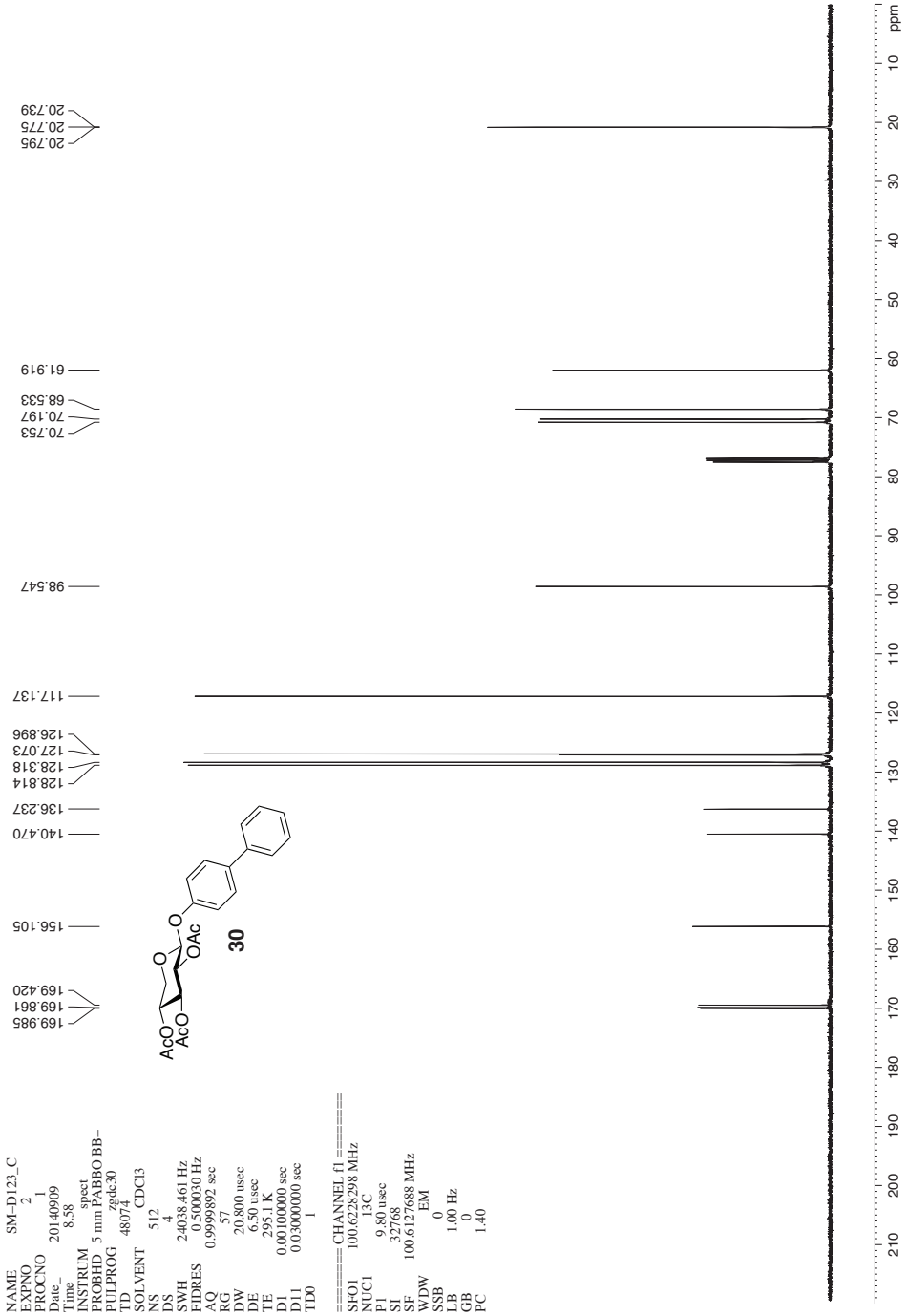
^{13}C-NMR 1-Naphthyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranoside **(29)**



[1]H-NMR 4-Phenylphenyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranoside **(30)**

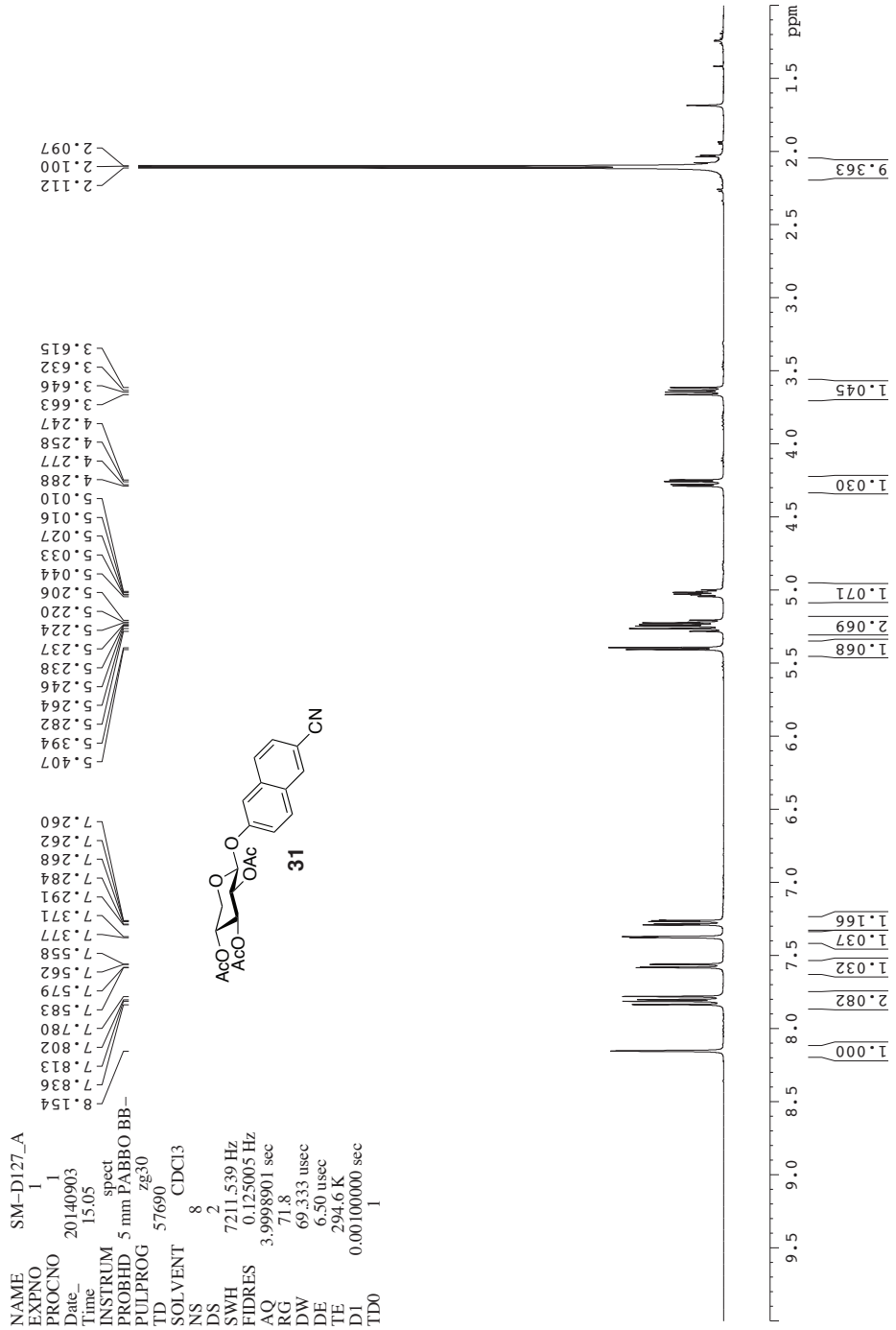


^{13}C-NMR 4-Phenylphenyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranoside **(30)**

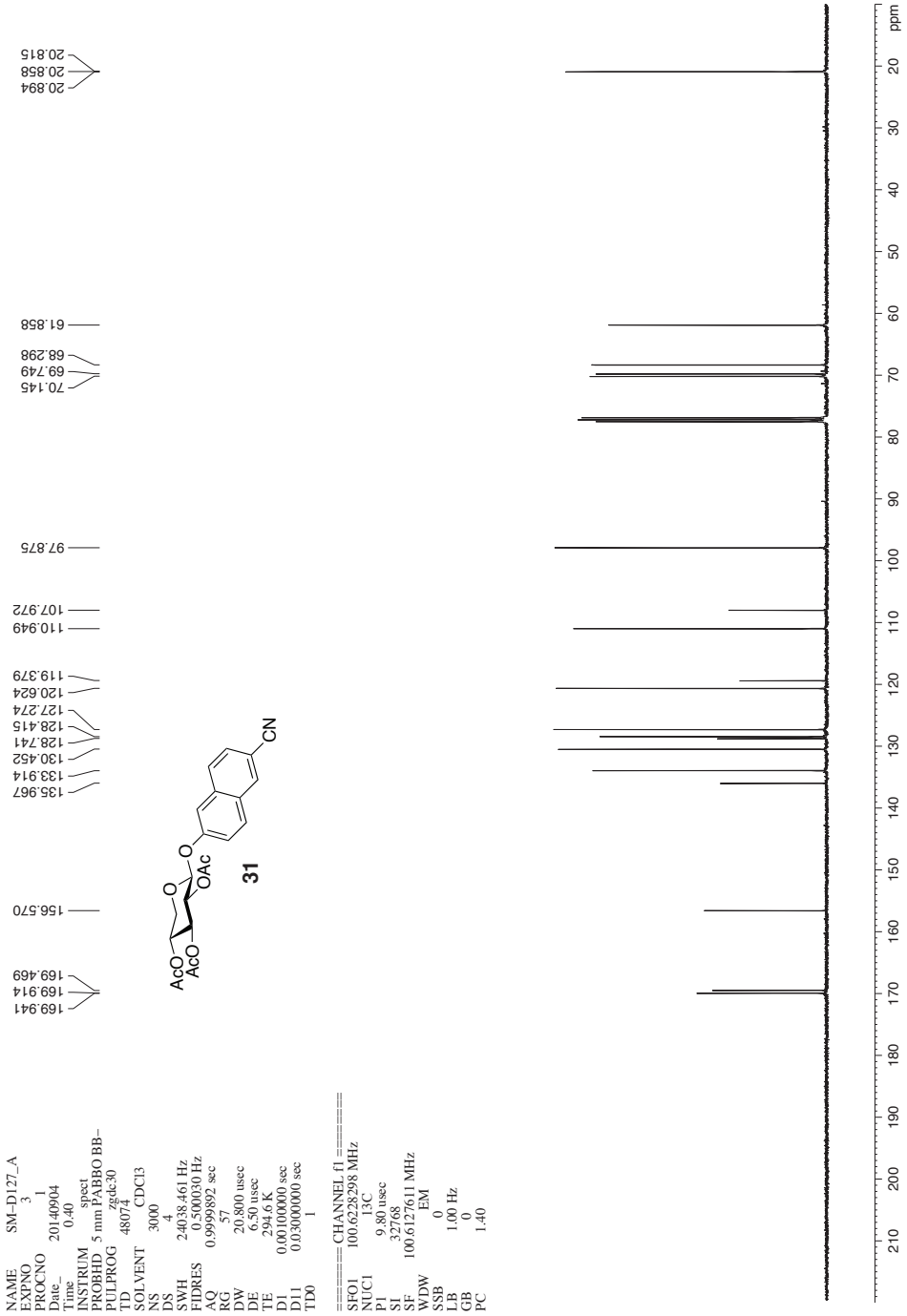


NAME SM-D123_C
EXPNO 2
PROCNO 1
Date_ 20140909
Time 8.58
INSTRUM spect
PROBHD 5 mm PABBO BB–
PULPROG zgdc30
TD 48074
SOLVENT CDCl3
NS 512
DS 4
SWH 24038.461 Hz
FIDRES 0.500030 Hz
AQ 0.9999892 sec
RG 57
DW 20.800 usec
DE 6.50 usec
TE 295.1 K
D1 0.00100000 sec
D11 0.03000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 100.6228298 MHz
NUC1 13C
P1 9.80 usec
SI 32768
SF 100.6127688 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

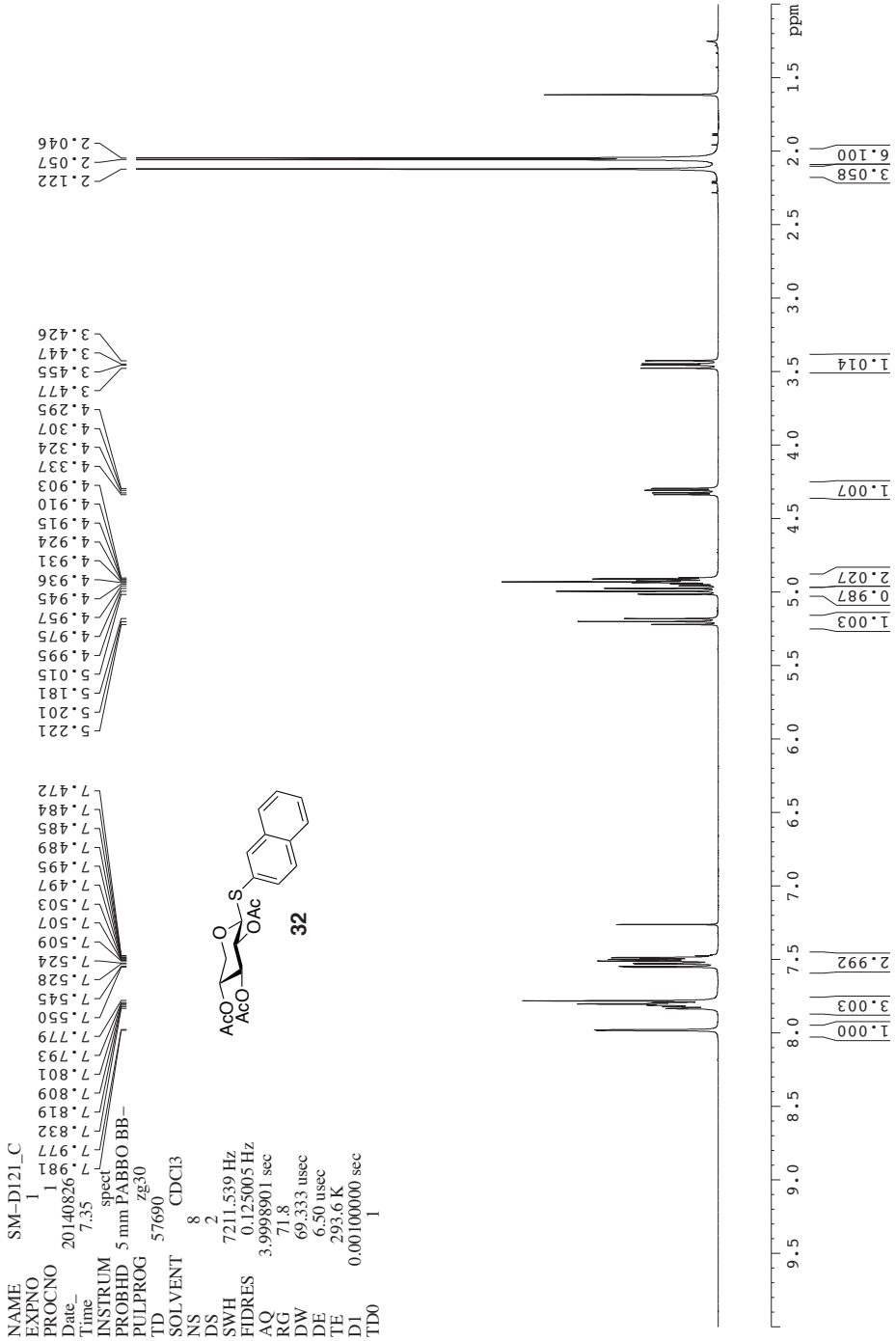
[1]H-NMR 2-(6-Cyano-naphthyl) 2,3,4-tri-*O*-acetyl-β-D-xylopyranoside **(31)**



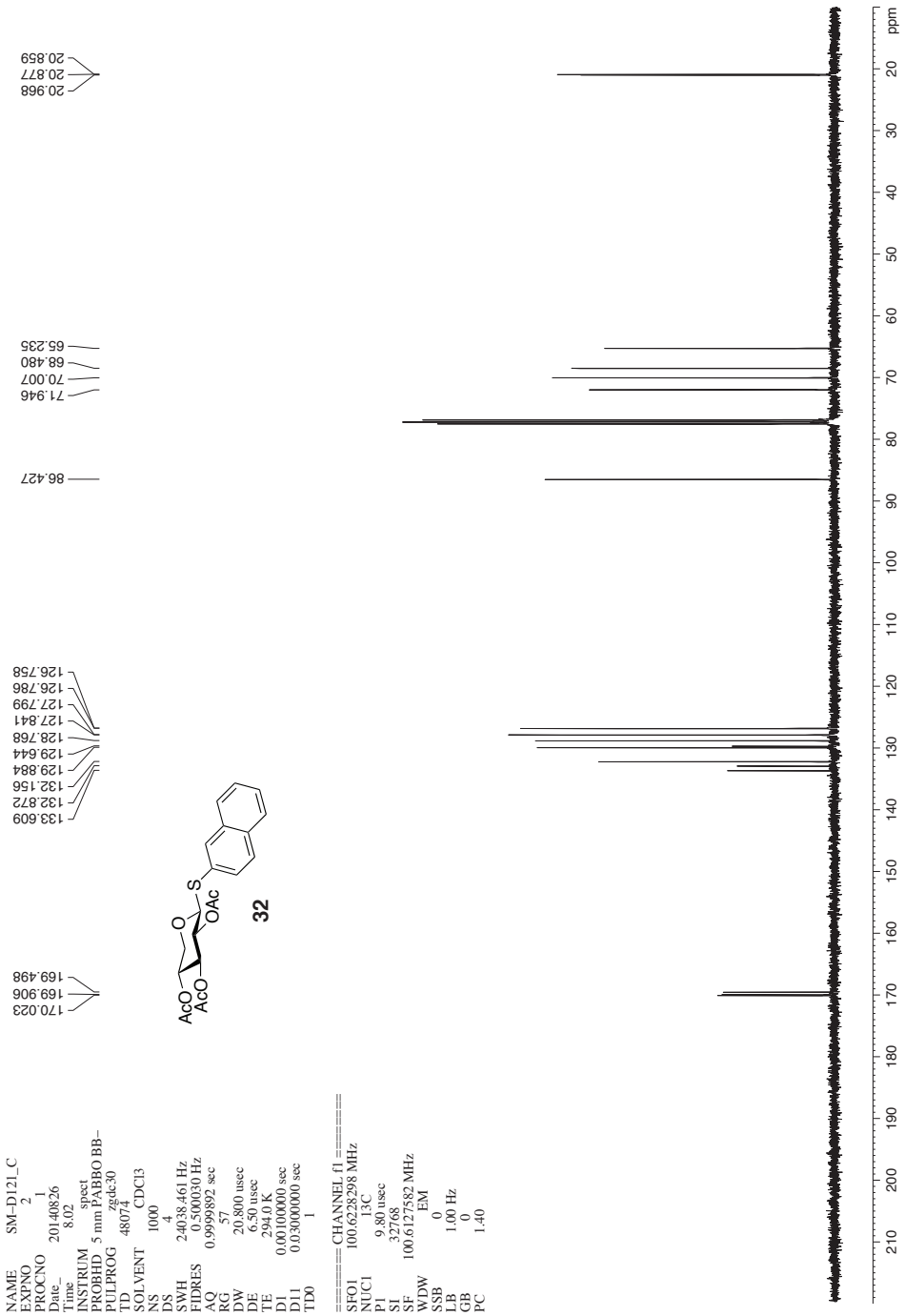
^{13}C-NMR 2-(6-Cyano-naphthyl) 2,3,4-tri-*O*-acetyl-β-D-xylopyranoside **(31)**



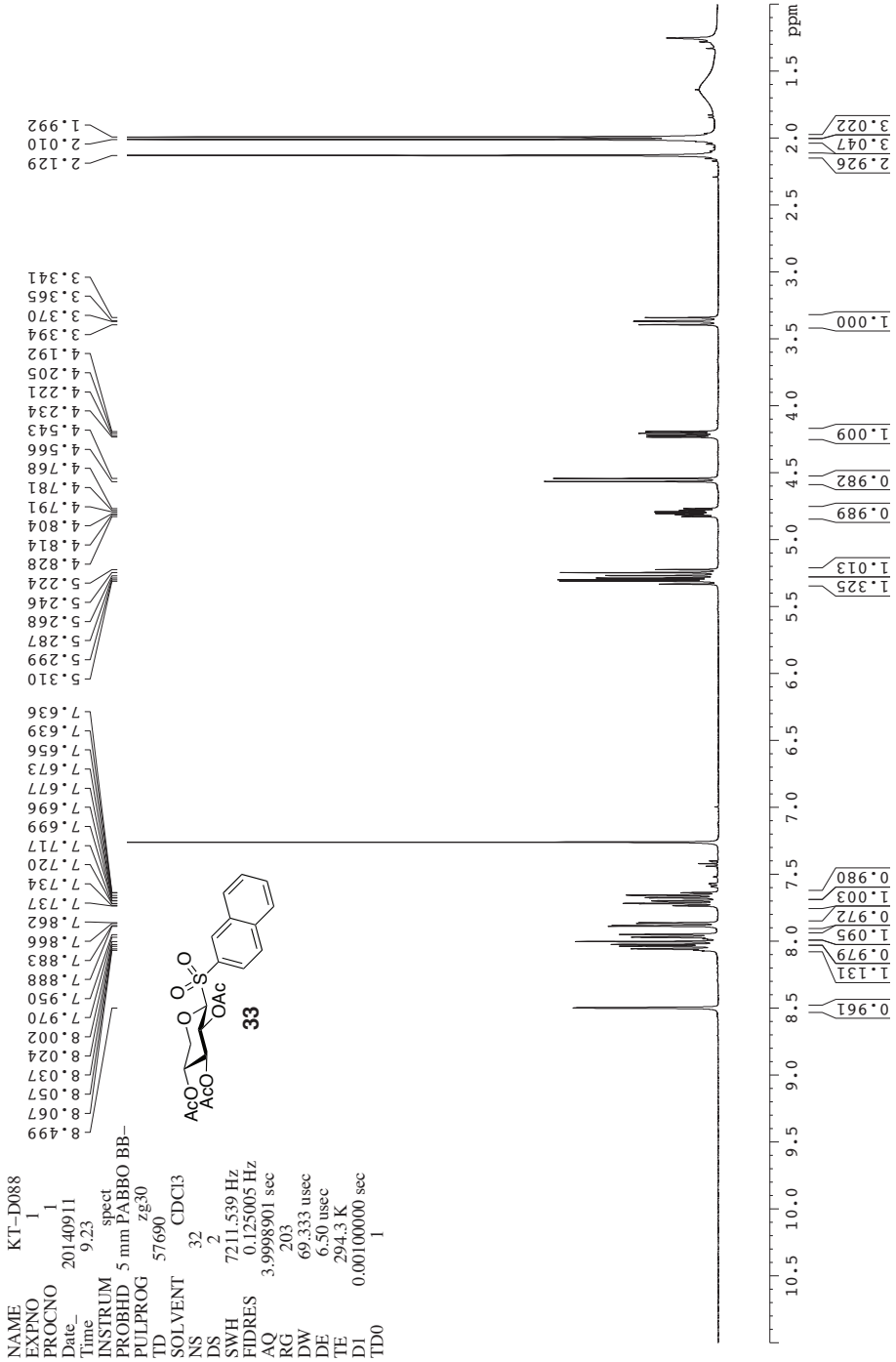
[1]H-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl-1-thio-β-D-xylopyranoside **(32)**



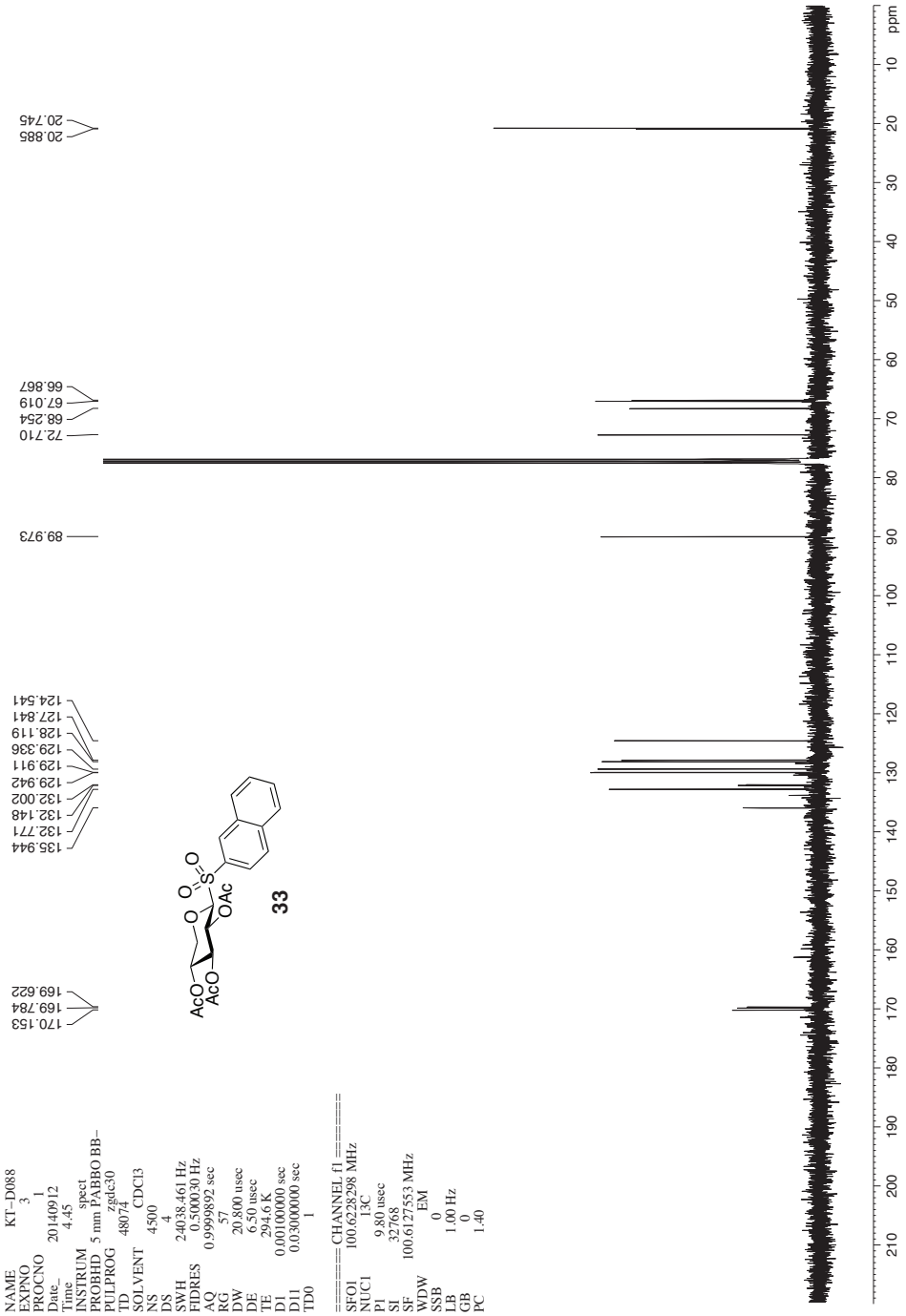
^{13}C-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl-1-thio-β-D-xylopyranoside **(32)**



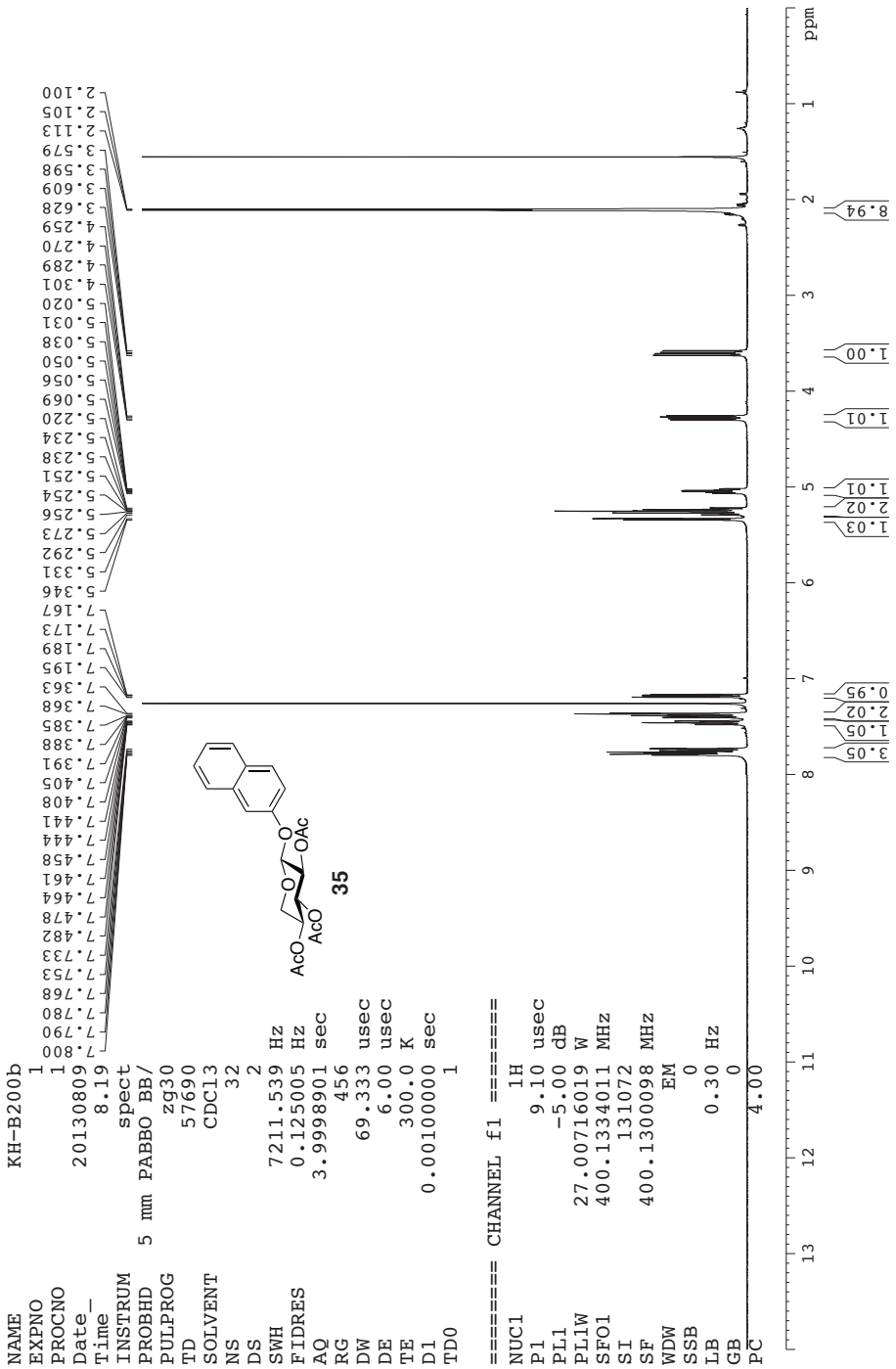
[1]H-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl β-D-xylopyranosyl sulfone **(33)**



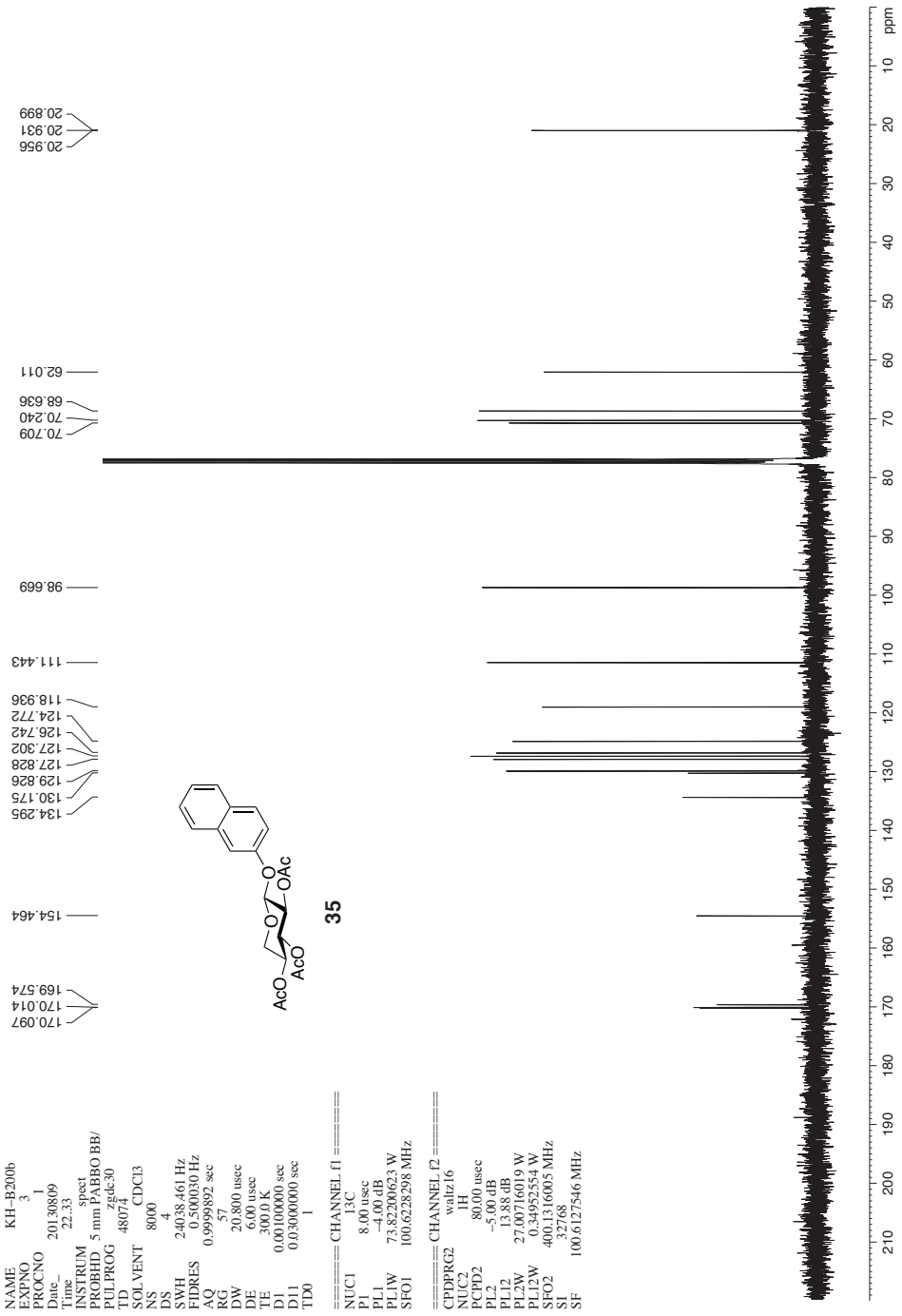
^{13}C-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl β-D-xylopyranosyl sulfone **(33)**



^{1}H-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl β-L-xylopyranoside **(35)**

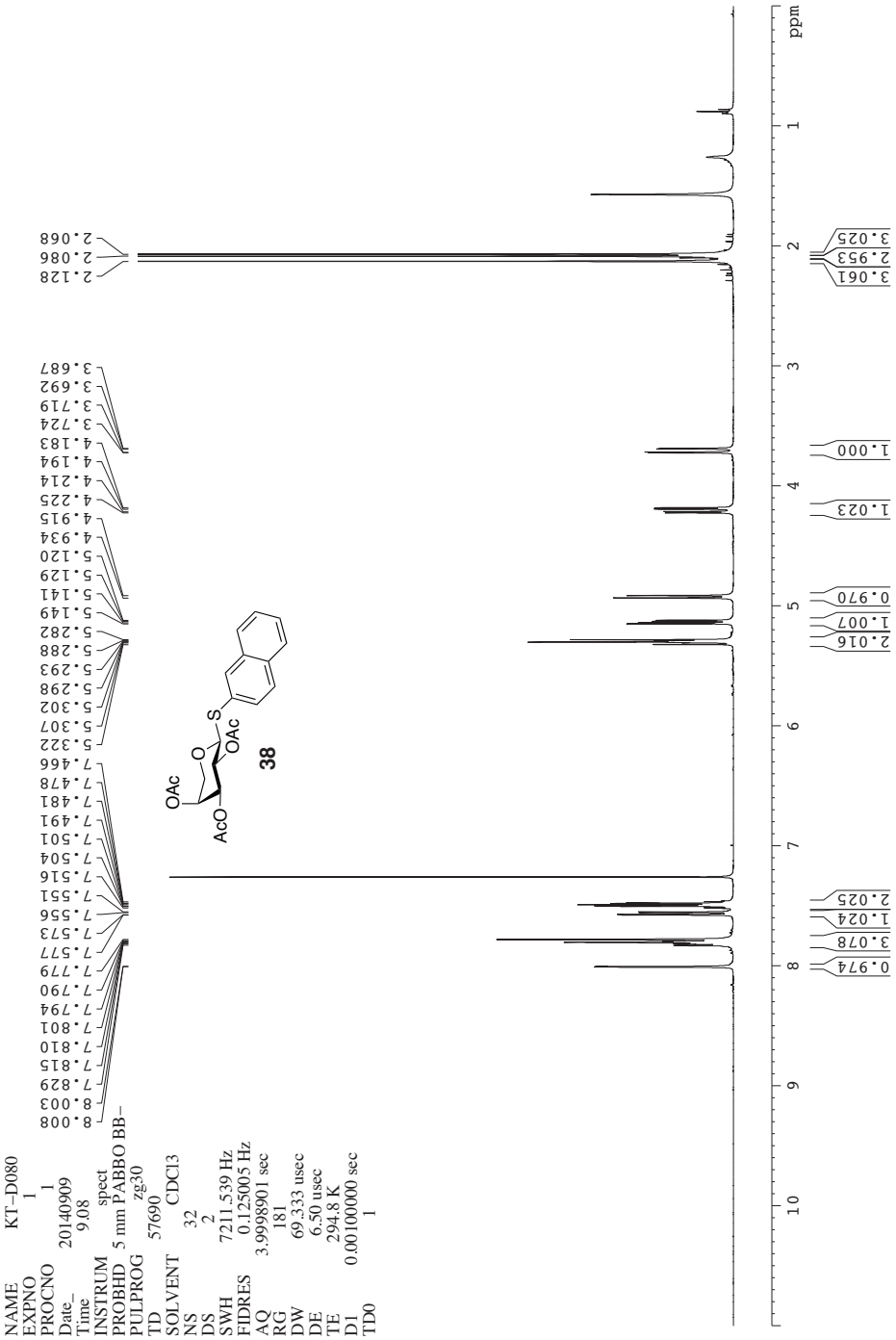


^{13}C-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl β-L-xylopyranoside **(35)**

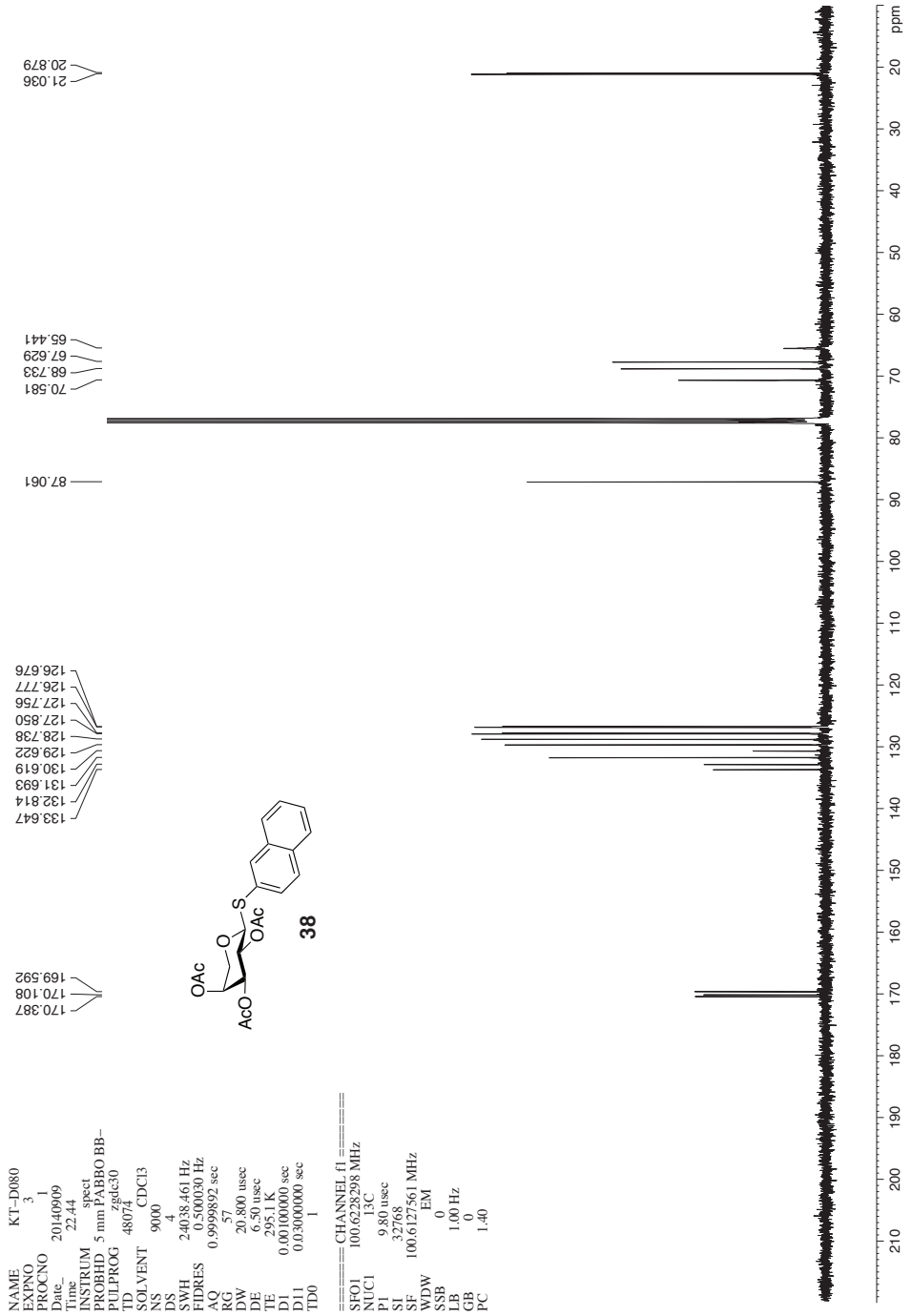


NAME KH-B200b
EXPNO 3
PROCNO 1
Date_ 20130809
Time 22.33
INSTRUM spect
PROBHD 5 mm PABBO BB/
PULPROG zgdc30
TD 48074
SOLVENT CDCl3
NS 8000
DS 4
SWH 24038.461 Hz
FIDRES 0.500030 Hz
AQ 0.9999892 sec
RG 57
DW 20.800 usec
DE 6.00 usec
TE 300.0 K
D1 0.00100000 sec
D11 0.03000000 sec
TD0 1
======== CHANNEL f1 ========
NUC1 13C
P1 8.00 usec
PL1 4.00 dB
PL1W 73.82200623 W
SFO1 100.6228298 MHz
======== CHANNEL f2 ========
CPDPRG2 waltz16
NUC2 1H
PCPD2 80.00 usec
PL2 -5.00 dB
PL12 13.88 dB
PL2W 27.00716019 W
PL12W 0.34952554 W
SFO2 400.1316005 MHz
SI 32768
SF 100.6127546 MHz

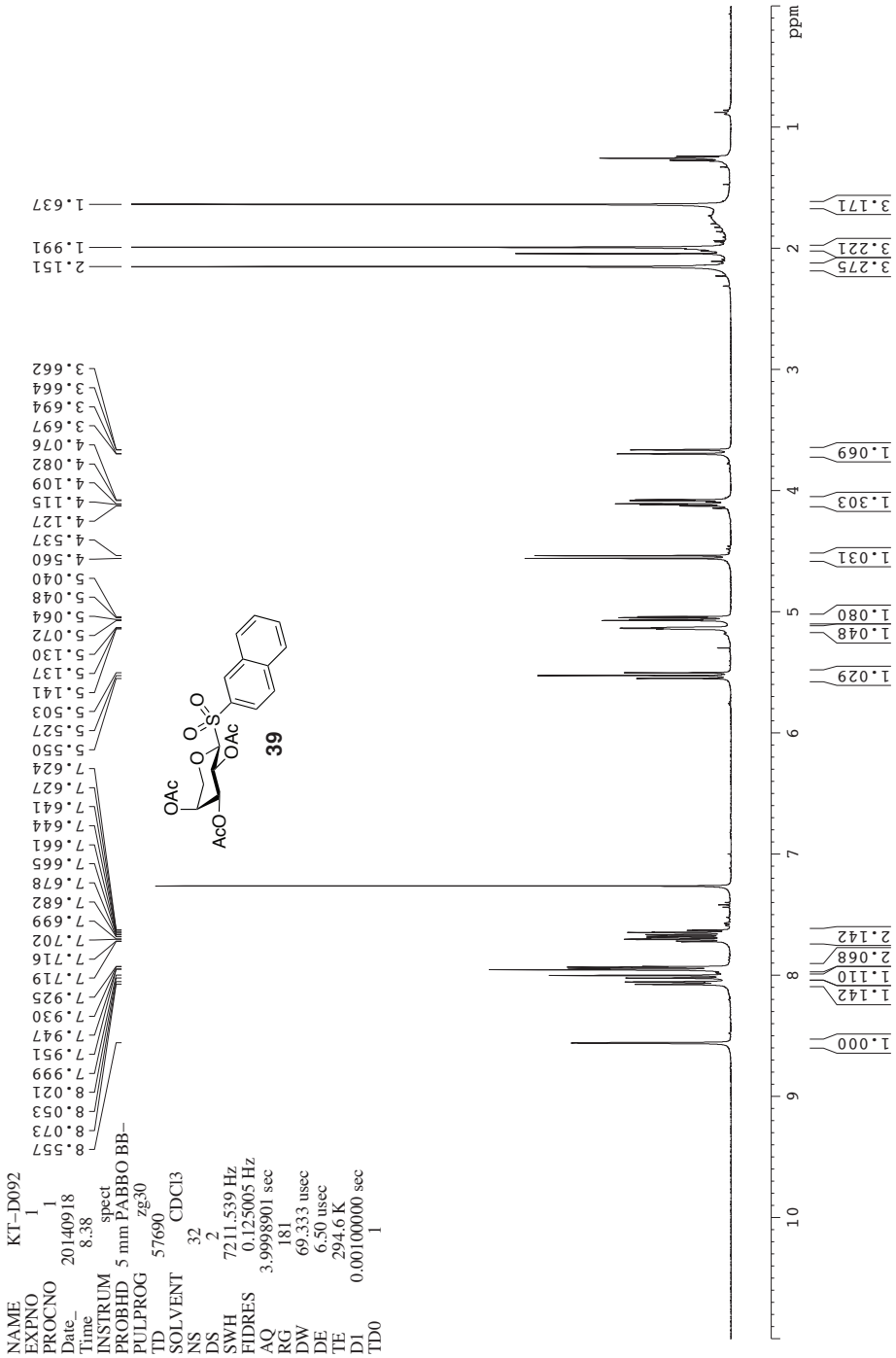
[1]H-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl 1-thio-α-L-arabinopyranoside **(38)**



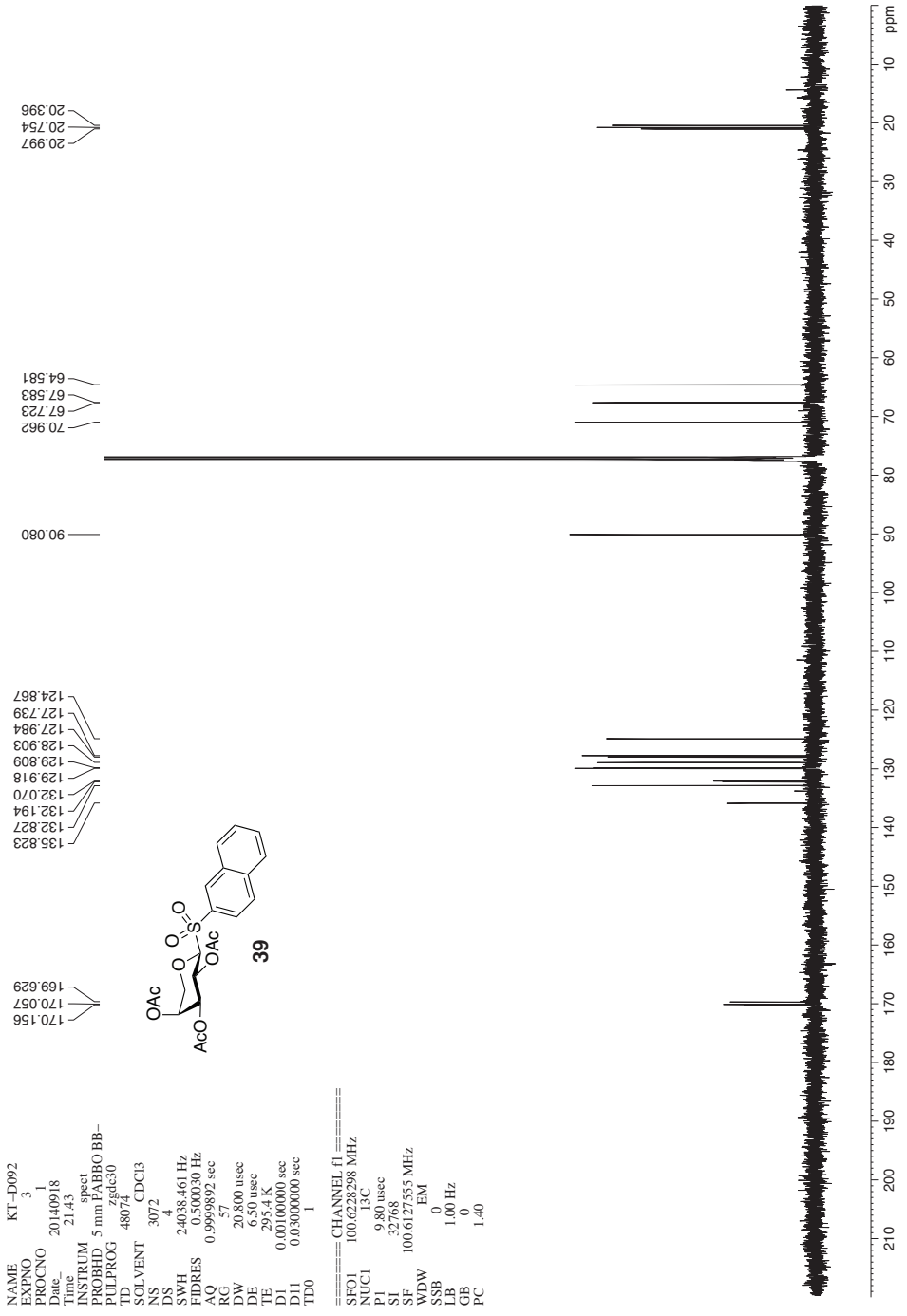
^{13}C-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl 1-thio-α-L-arabinopyranoside **(38)**



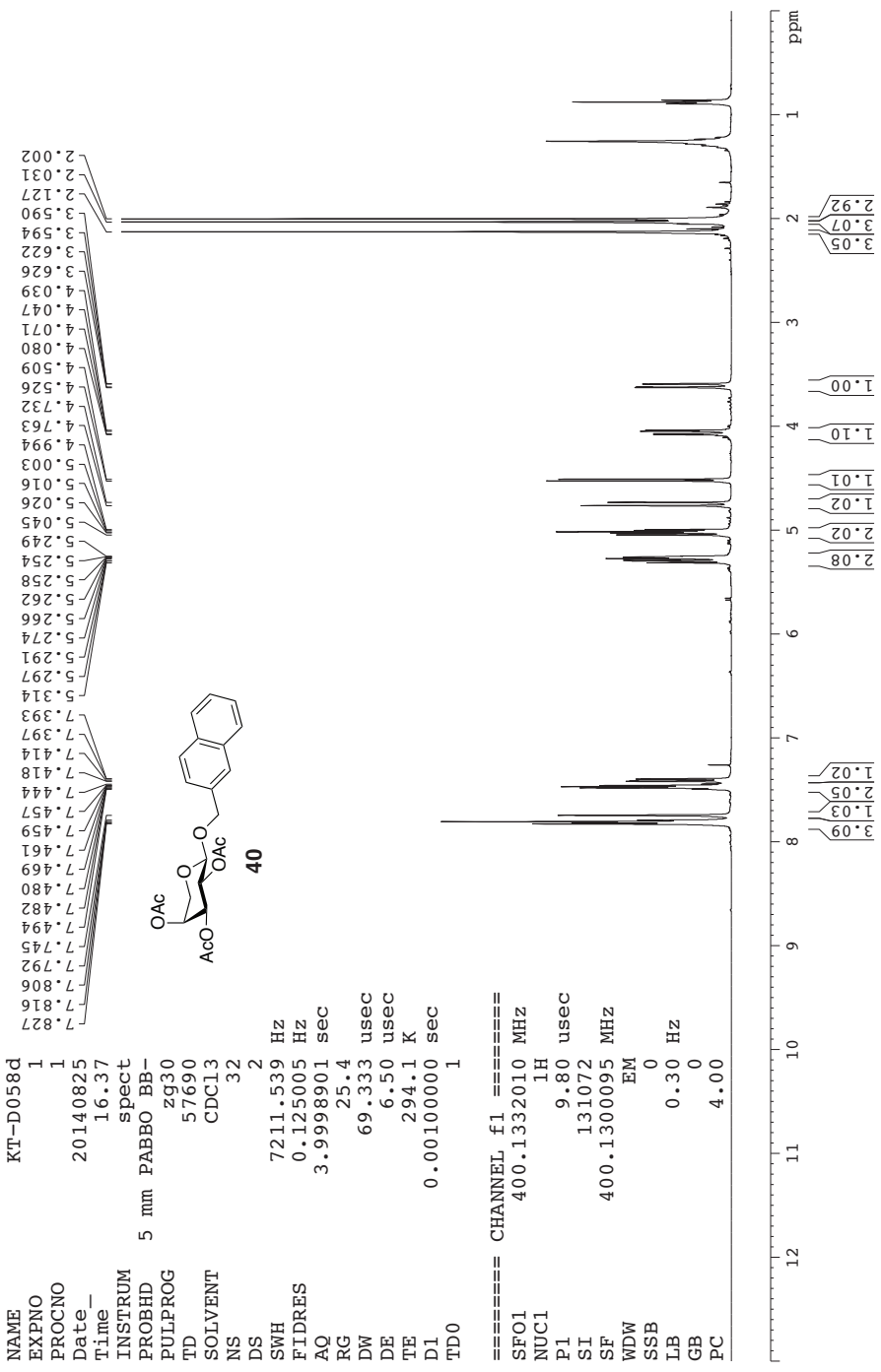
^{1}H-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl α-L-arabinopyranosyl sulfone **(39)**



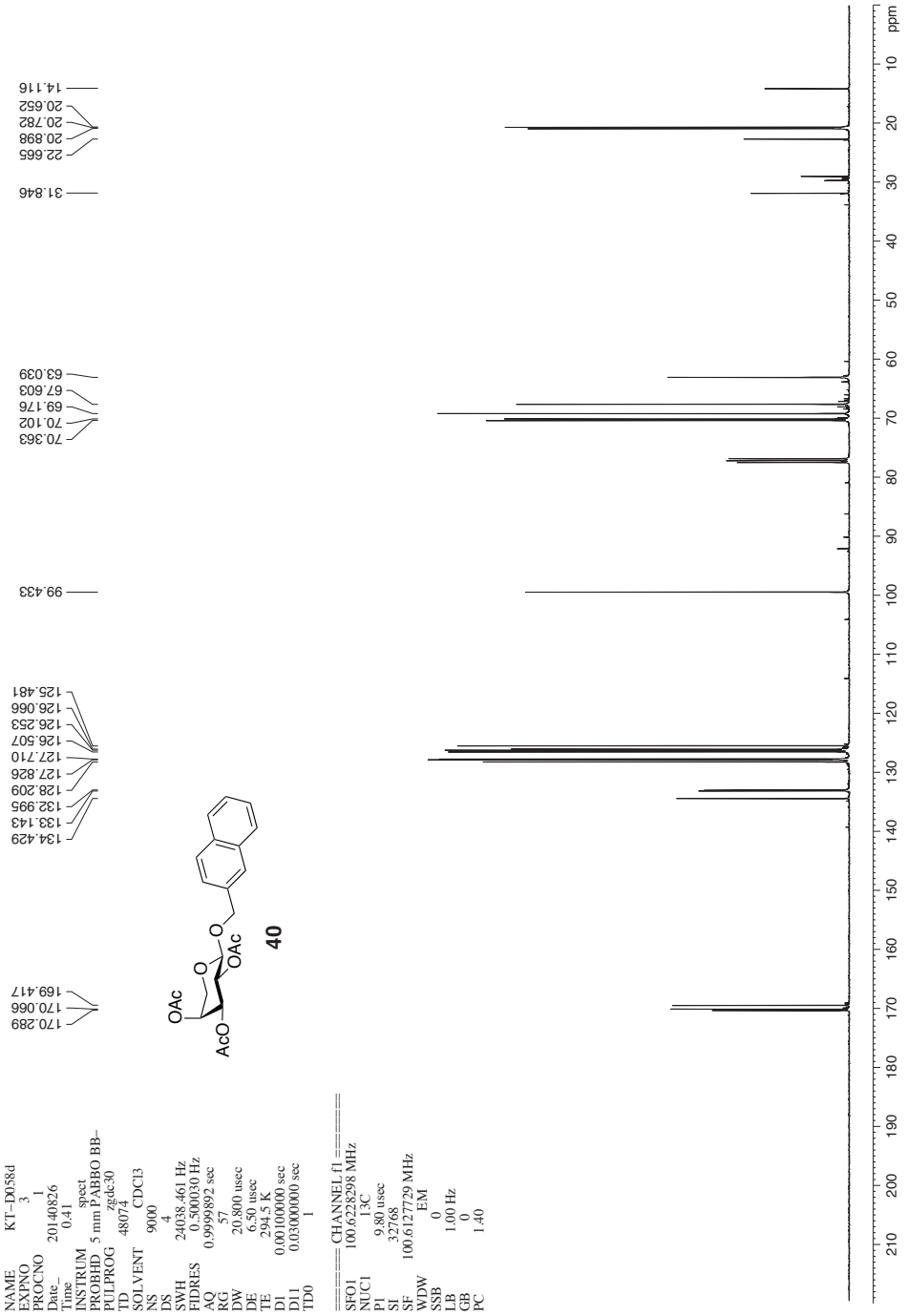
^{13}C-NMR 2-Naphthyl 2,3,4-tri-*O*-acetyl α-L-arabinopyranosyl sulfone **(39)**



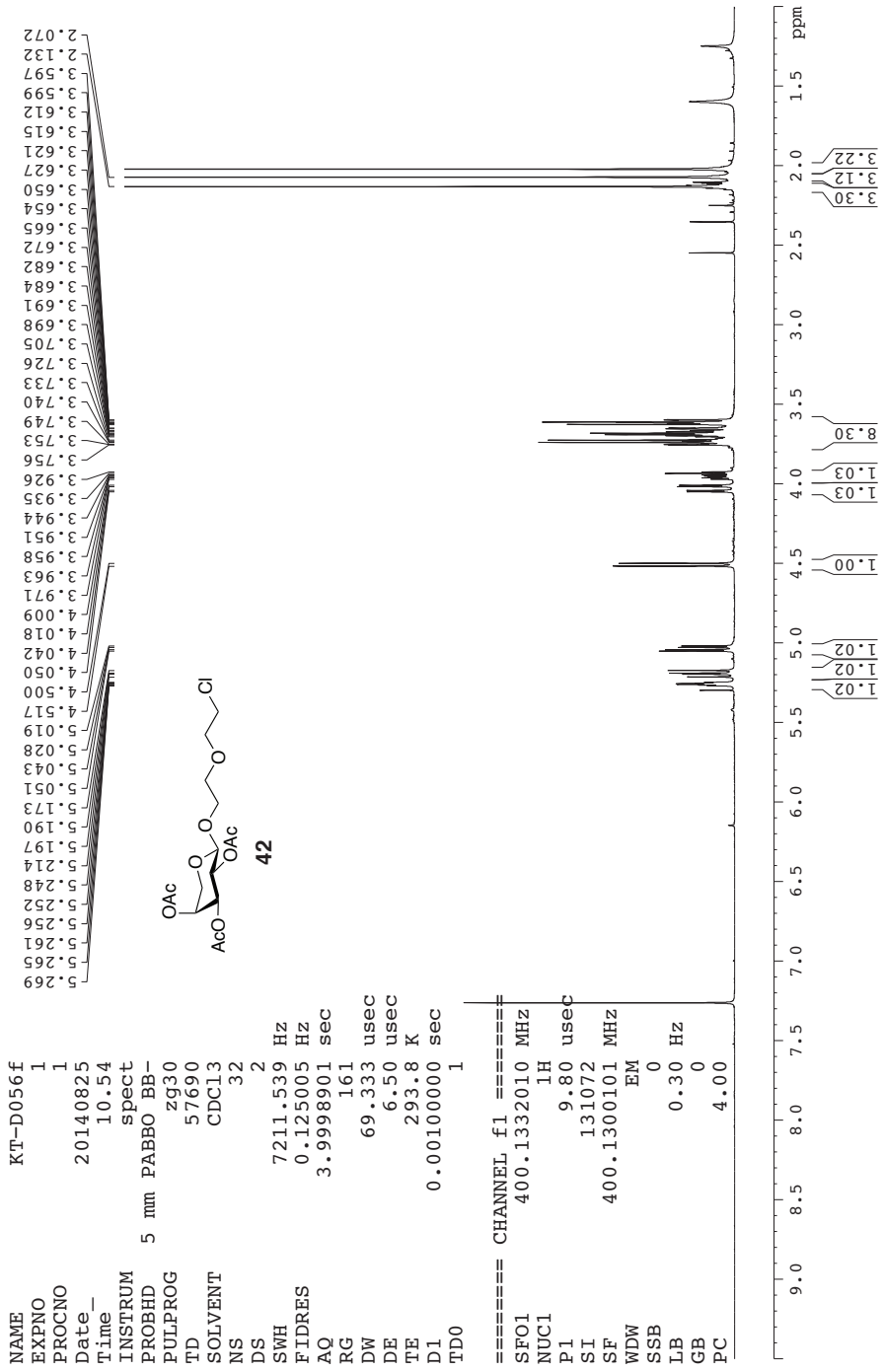
[1]H-NMR (2-Naphthyl)-methyl 2,3,4-tri-*O*-acetyl α-L-arabinopyranoside **(40)**



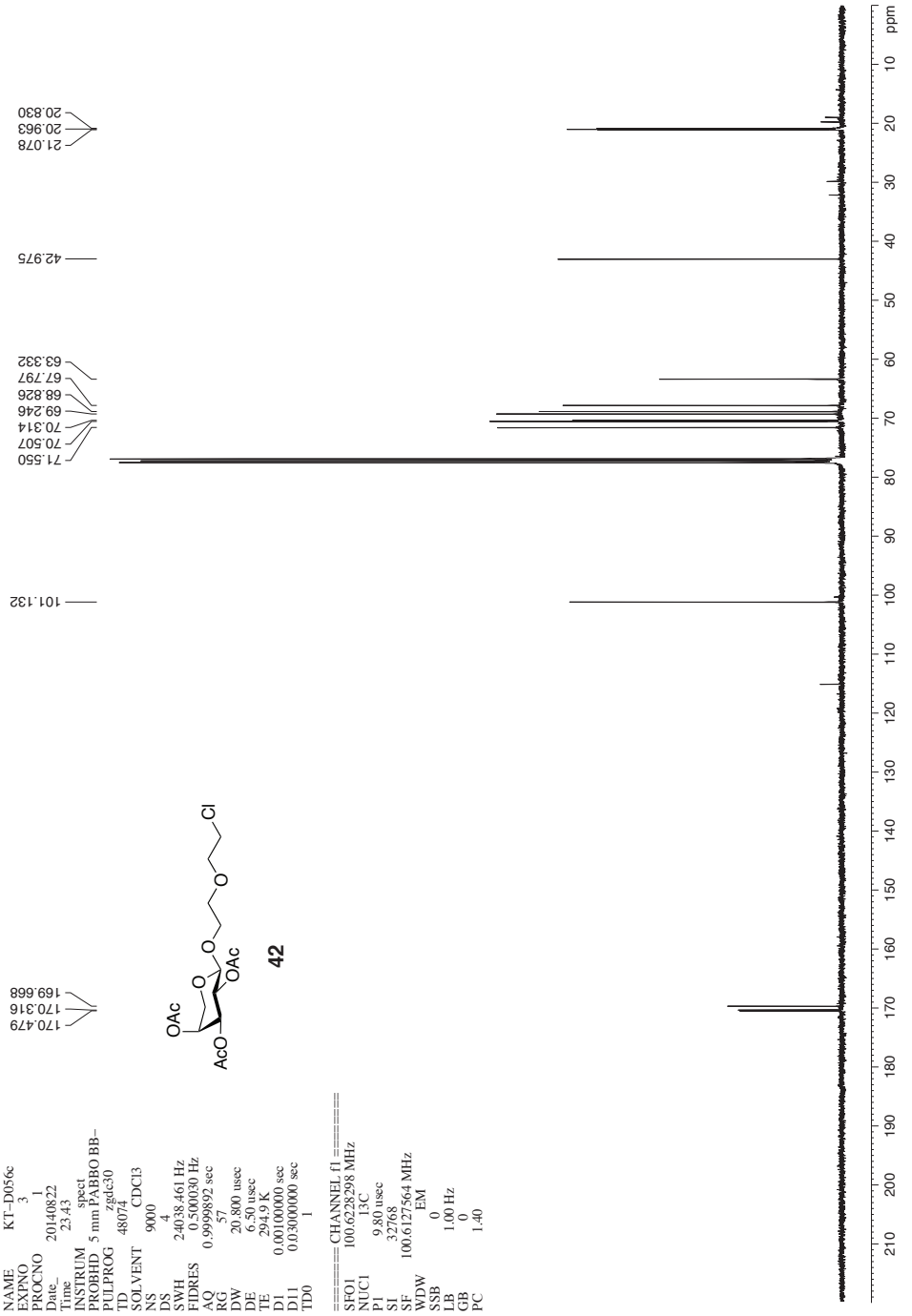
^{13}C-NMR (2-Naphthyl)-methyl 2,3,4-tri-*O*-acetyl α-L-arabinopyranoside **(40)**



[1]H-NMR 5-Chloro-3-oxo-1-pentyl 2,3,4-tri-*O*-acetyl-α-L-arabinopyranoside **(42)**

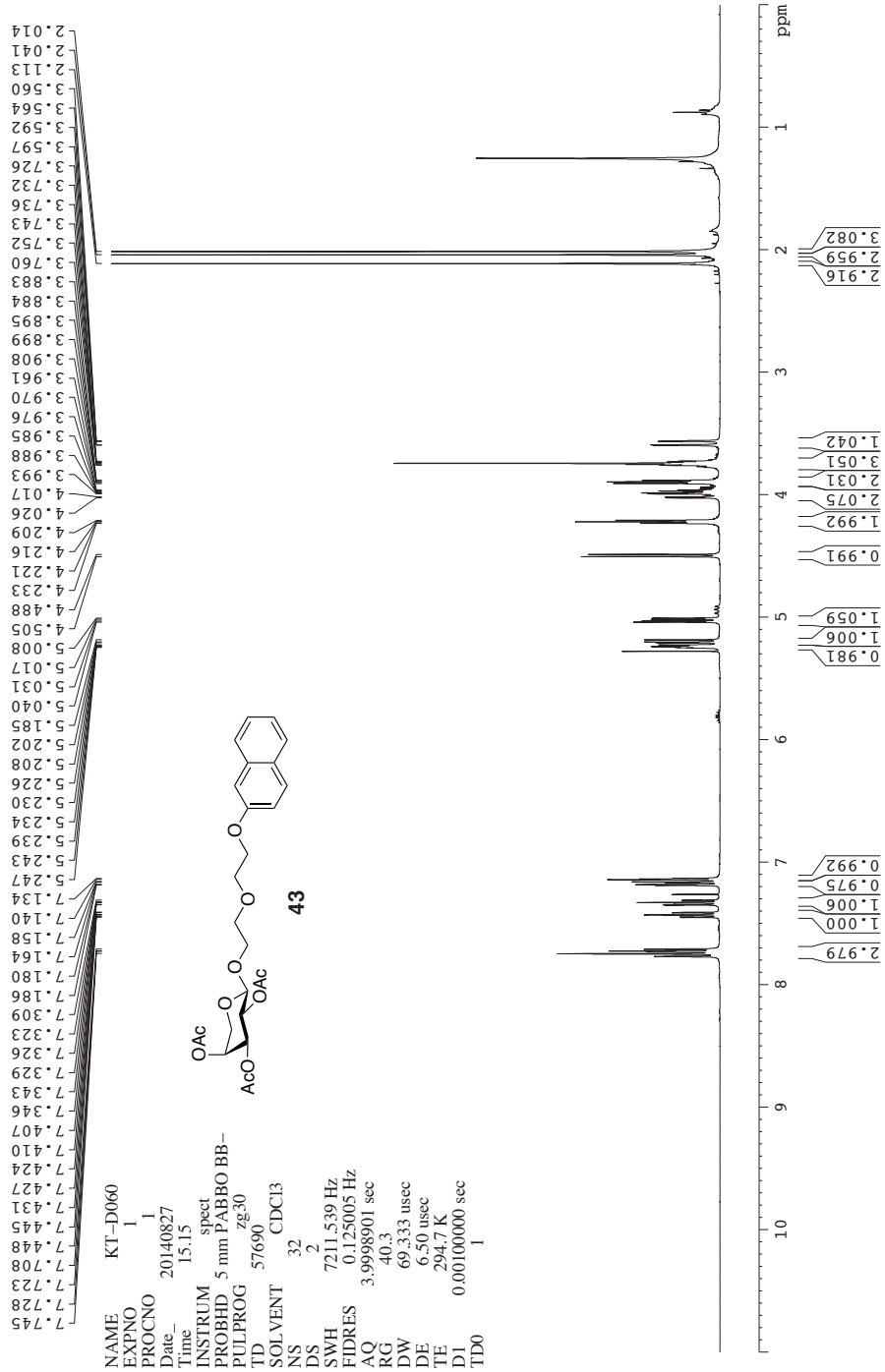


^{13}C-NMR 5-Chloro-3-oxo-1-pentyl 2,3,4-tri-*O*-acetyl-α-L-arabinopyranoside **(42)**



NAME KT–D056c
EXPNO 3
PROCNO 1
Date_ 20140822
Time 23.43
INSTRUM spect
PROBHD 5 mm PABBO BB–
PULPROG zgdc30
TD 48074
SOLVENT CDCl3
NS 9000
DS 4
SWH 24038.461 Hz
FIDRES 0.500030 Hz
AQ 0.9999892 sec
RG 57
DW 20.800 usec
DE 6.50 usec
TE 294.9 K
D1 0.00100000 sec
D11 0.03000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 100.6228298 MHz
NUC1 13C
P1 9.80 usec
SI 32768
SF 100.6127564 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

[1]H-NMR 5-(2-Naphthoxy)-3-oxo-1-pentyl 2,3,4-tri-*O*-acetyl-α-L-arabinopyranoside **(43)**



^{13}C-NMR 5-(2-Naphthoxy)-3-oxo-1-pentyl 2,3,4-tri-*O*-acetyl-α-L-arabinopyranoside **(43)**