Available online at www.sciencedirect.com
SCIENCE DIRECT®
Tetrahedron Letters
Tetrahedron Letters 45 (2004) 8507–8510

A fused GalNAc-thiazole from a singular and unanticipated fragmentation

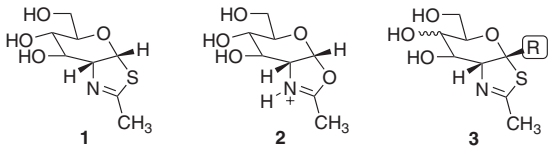
Spencer Knapp,[a,*] Benjamin Amorelli[a] and George A. Doss[b]

[a]*Department of Chemistry & Chemical Biology, Rutgers—The State University of New Jersey, 610 Taylor Road, Piscataway, NJ 08854-8087, USA*
[b]*Merck & Co., PO Box 2000, Rahway, NJ 07065-0900, USA*

Received 28 July 2004; revised 7 September 2004; accepted 13 September 2004
Available online 29 September 2004

Abstract—Attempted *O*-desilylation of an elaborated GalNAc-thiazoline *t*-butyl ester gave rise instead to a GalNAc-thiazole wherein the two-carbon acetate side chain has been severed by a process heretofore undescribed. The corresponding ether ester rearranged to a *C*-furanoside by a quite different pathway.

The GlcNAc-thiazoline **1** is a powerful inhibitor of certain *N*-acetylhexosaminidases, including the human enzyme HEX B, owing to its structural resemblance to the cyclized and highly reactive oxazolinium intermediate **2**.[1] 1-*C*-Elaborated *gluco* and *galacto* analogues (**3**) are therefore also of interest as potential inhibitors with stronger binding or greater selectivity or interesting conjugative applications.[2] We have examined, as a means to the synthesis of the latter compounds, the carbon chain extension of glyconolactones such as **7**, and have prepared the two-carbon extended GalNac-thiazoline **8** (Scheme 1). Attempted *O*-desilylation of **8** under basic conditions, however, did not afford the expected thiazoline triol, but rather a truncated and aromatized product, the GalNAc-thiazole **10**. Inasmuch as **10** is a new fused pyranose-heterocycle hybrid, and the two-carbon fragmentation (formal overall loss of *tert*-butylacetate) leading to its formation corresponds to no previously described reaction of which we are aware, we report this preliminary investigation into the mechanism of its formation.

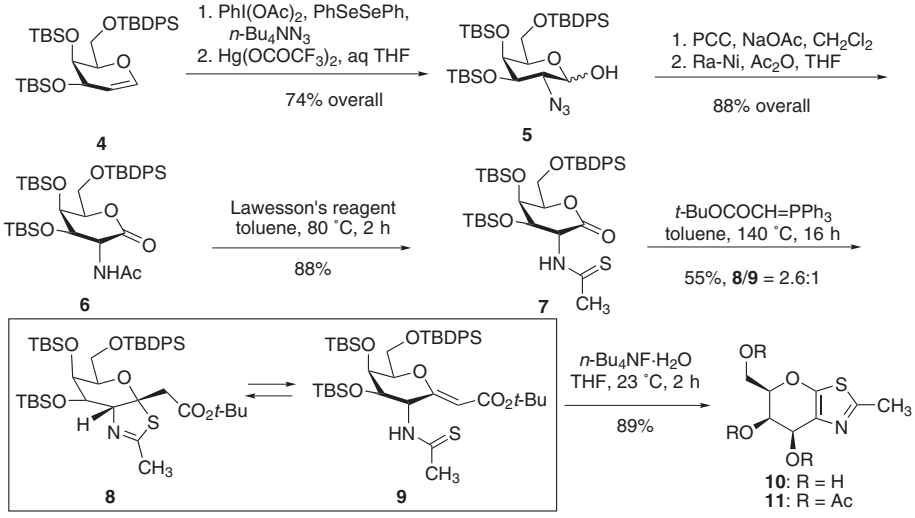


The silylated[3] D-galactal derivative **4** was converted[4] to the phenyl 2-azido-2-deoxy 1-selenogalactopyranoside, which in turn was hydrolyzed[4] at C-1 to the galactopyranose **5** (Scheme 1). Oxidation[3] to the 2-azidogalactonolactone was followed by efficient reductive acetylation[5] to provide the 2-acetamido-2-deoxy-D-galactolactone **6**, and then treatment with Lawesson's reagent gave the thioacetamido lactone **7**. This lactone reacted reluctantly upon heating with *tert*-butoxycarbonylmethylenetriphenylphosphorane,[5] but at 140 °C gave an equilibrium mixture of the chain extended vinylogous carbonate **9** and the corresponding *S*-cyclized thiazoline ester **8** (55% combined isolated yield, respective ratio 1:2.6). Although **8** and **9** could not be separated, they could be characterized by NMR analysis on the mixture and by comparison to closely related compounds.[2,6]

Attempted removal of the *O*-silyl protecting groups with 4.1 equiv of n-$Bu_4NF{\cdot}H_2O$ caused the fragmentation of the two-carbon side chain of **8/9**, and led to the isolation of the unusual thiazole triol **10** in high yield. The structure of **10** follows from spectroscopic analysis, as well as further characterization of its derived triacetate **11**. No comparable pyranose-fused thiazole has been reported,[7] to our knowledge, and no good mechanistic precedent exists for the apparent loss of the elements of *tert*-butyl acetate under such mild conditions.[8] Both **8** and **9** were converted to **10** in this reaction, because the amount of either precursor in the starting mixture is alone insufficient to account for the yield of **10**.

Keywords: Retro-aldol; Rearrangement; Aromatization.
* Corresponding author. Tel.: +1 732 445 2627; fax: +1 732 445 5312; e-mail: knapp@rutchem.rutgers.edu

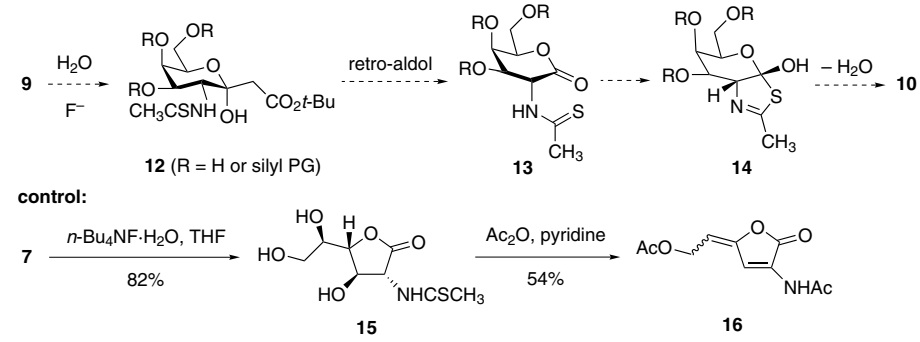
0040-4039/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2004.09.091



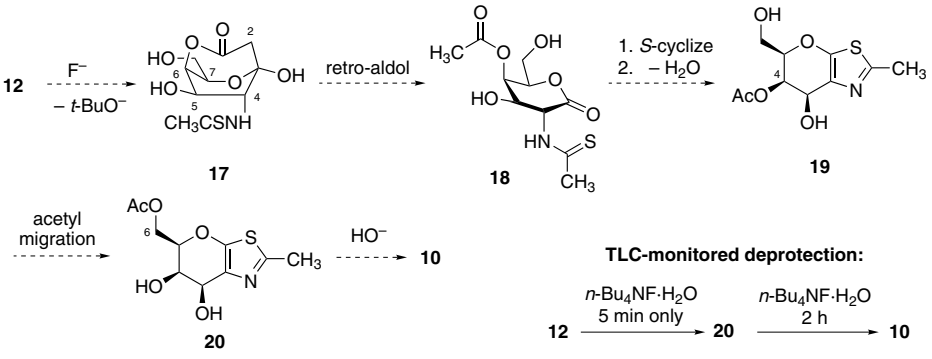
Scheme 1. Attempted synthesis of an elaborated GalNAc-thiazoline.

Given the presence of sufficient amounts of water in the commercial n-$Bu_4NF{\cdot}H_2O$ reagent, it is possible to suggest (Scheme 2) that Michael addition of water to the C=C of **9** has occurred under the basic conditions of the reaction, followed by retro-aldol reaction to give the lactone **13** (loss of *tert*-butyl acetate), then cyclization of the thioacetamido to produce **14**, and finally loss of water and aromatization. The timing of the *O*-desilylation steps relative to other proposed processes is unknown. To help evaluate this mechanism, thioamide lactone **7** was exposed to the same desilylation conditions (Scheme 2). No thiazole **10** was detected by TLC or NMR analysis, but instead the rearranged γ-lactone **15** was formed as the major product.[9] It was further characterized by acetylation to give the known[10] *N*,*O*-diacetyl derivative **16** ($Z/E \approx 2{:}1$) with matching ^{1}H NMR spectrum. Ring contraction of the δ-lactone to the γ-lactone is well established for galactonolactones,[11] so this is not a surprising result, but it does appear to rule out **7** as an intermediate.

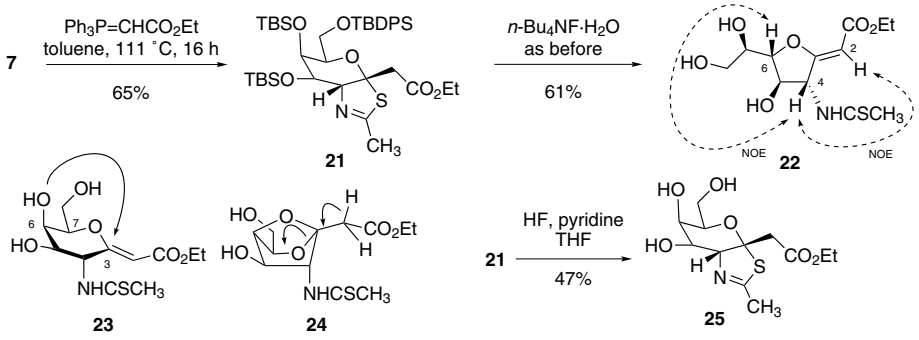
An admittedly speculative mechanism invoking participation by one of the pyranose hydroxyls is proposed in Scheme 3. In this formulation, the pyranose ring



Scheme 2. A retro-aldol mechanism for thiazole formation.



Scheme 3. A retro-aldol mechanism with hydroxyl participation.



Scheme 4. The contrasting rearrangement of the ethyl ester.

undergoes a conformational change to bring the C-6 hydroxyl close to the ester carbonyl, whereupon cyclization to the lactone **17** can occur. *n*-Bu_4NF promoted lactonizations of hydroxy esters have a precedent,[12] and alternative isomeric lactone intermediates might also be drawn invoking participation of the C-5 or C-8 hydroxyl, respectively. Retro-aldol cleavage would lead to lactone **18**, and then thioamide cyclization and loss of water can occur to produce **19**. The 4-*O*-acetyl group of **19** specifically blocks the δ-lactone to γ-lactone ring contraction pathway described above (see **15**). Hydroxide promoted hydrolysis of the acetate would lead to the thiazole **10**, possibly by acetyl migration[13] through the 6-*O*-acetyl isomer **20**.

A subsequent deprotection experiment established the 6-*O*-acetyl thiazole **20** as an intermediate. When the desilylation reaction was analyzed by TLC after only 5 min (Scheme 3), the major product observed was a higher R_f (with respect to **10**) thiazole, **20**. Upon further exposure to the reaction conditions, **20** was slowly converted to **10**. The structure of **20** was secured by its isolation and spectroscopic characterization (the ~0.5 ppm downfield shift of the H-6's versus those of **10** indicates the 6-*O*-acetyl); furthermore, independent acetylation of **20** gave the previously described thiazole triacetate **11**.

Lactone **7** also undergoes chain extension when treated with (carbethoxymethylene)-triphenylphosphorane (Scheme 4). In this case, no alkene analogous to **9** is present in the product; cyclization to thiazoline **21** is complete. Treatment of **21** with *n*-Bu_4NF under the same conditions as for **8** led to no thiazole **10** whatsoever. Instead, the only isolated product proved to be the deprotected and rearranged thioamide ester **22**. Its structure and stereochemistry are based on fully assigned ^1H and ^{13}C NMR spectra, and NOESY, HSQC and HMBC analysis. In particular, NOE cross-peaks were observed for H-2/H-4 and H-4/H-6. The source of **22** might be intramolecular conjugate addition of O-6 at C-3 of the unsaturated ester **23** to give a bicyclic intermediate **24**, followed by cleavage of the O-7 ether bond by β-elimination. This may be thought of as a process vinylogous to that which formed **15**. Successful desilylation of **21** to **25** without rearrangement was later achieved under acidic conditions (HF, pyridine).

The subtleties of these polyfunctional substrates, particularly in regard to the reactivity differences between *tert*-butyl and ethyl esters (**8** and **21**, respectively), cannot in our view be accounted for by a single simple explanation. However, the apparent ease of the retro-Michael thiazoline ring opening reaction of **8** relative to that of **21** (and perhaps relative to *O*-desilylation) might set the stage for the emergence of the distinctive subsequent pathways that have been proposed in Scheme 3.[14]

Acknowledgements

We are grateful to NIH (AI055760) for financial support. We thank Christian Ventocilla for assisting with the preparation of starting materials and Etzer Darout for helpful discussions.

References and notes

1. Mark, B. L.; Mahuran, D. J.; Cherney, M. M.; Zhao, D.; Knapp, S.; James, M. N. G. *J. Mol. Biol.* **2003**, *327*, 1093–1109.
2. Knapp, S.; Yang, C.; Haimowitz, T. *Tetrahedron Lett.* **2002**, *43*, 7101–7104.
3. Koviach, J. L.; Chappell, M. D.; Halcomb, R. L. *J. Org. Chem.* **2001**, *66*, 2318–2326.
4. Czernecki, S.; Ayadi, E. *Can. J. Chem.* **1995**, *73*, 343–350.
5. Xie, J.; Molina, A.; Czernecki, S. *J. Carbohydr. Chem.* **1999**, *18*, 481–498.
6. The ^1H NMR signal for the thiazoline methyl of **8** at 2.17 ppm (br s, long range coupled to H-2) is diagnostic: Knapp, S.; Myers, D. S. *J. Org. Chem.* **2002**, *67*, 2995–2999. The thioacetamide methyl singlet of **9**, observed at 2.47 ppm, is also diagnostic, and integration gives their ratio. Complete spectra are given in Ref. 14.
7. The closest analogous non-aromatized structure is the corresponding GalNac-thiazoline (Ref. 6). Analogous *benzo*-fused glycals have also been synthesized: Pachamuthu, K.; Schmidt, R. R. *Synlett* **2003**, 1355–1357.
8. The thermal reversion of glyconolactone-thiazole adducts back to the glyconolactones has been observed: Dondoni, A.; Marra, A. *Heterocycles* **1999**, *50*, 419–426.
9. The corresponding acetamido lactone has been reported: Zsoldos-Mady, V.; Pinter, I.; Nesmelyi, A.; Messmer, A. *Carbohydr. Res.* **1994**, *252*, 85–95.
10. Horton, D.; Thomson, J. K.; Varela, O.; Nin, A.; de Lederkremer, R. M. *Carbohydr. Res.* **1989**, *193*, 49–60.

11. Bierenstiel, M.; Schlaf, M. *Eur. J. Org. Chem.* **2004**, 1474–1481; Pravdic, N.; Fletcher, H. G., Jr. *Carbohydr. Res.* **1970**, *12*, 471–474.
12. Adiyaman, M.; Lawson, J. A.; Hwang, S.-W.; Khanapure, S. P.; FitzGerald, G. A.; Rokach, J. *Tetrahedron Lett.* **1996**, *37*, 4849–4852.
13. Motawia, M. S.; Olsen, C. E.; Denyer, K.; Smith, A. M.; Moller, B. L. *Carbohydr. Res.* **2001**, *56*, 6511–6516.
14. Spectra for new compounds. Compound **5** (≈1:1 mixture of α- and β-anomers): IR (film, cm^{-1}) 2112; ^1H NMR (400 MHz, $CDCl_3$, δ in parts per million, multiplicity, integration, J in hertz) 7.67–7.61 (m, 4H), 7.46–7.35 (m, 6H), 5.23 (d, 0.5H, 2.8), 4.35 (d, 0.5H, 7.6), 4.06 (br s, 0.5H), 3.98 (dd, 1H, 10.4, 2.0), 3.86 (br t, 1H, 6.4), 3.82–3.63 (m, 5.5H), 3.49 (dd, 0.5H, 10.0, 7.6), 3.34 (dd, 0.5H, 10.0, 7.6), 3.25 (t, 0.5H, 6.8), 2.85 (br s, 1H), 1.07 (s, 9H), 0.98 (s, 4.5H), 0.97 (s, 4.5H), 0.87 (s, 9H), 0.20 (s, 1.5H), 0.18 (s, 3H), 0.16 (s, 4.5H), 0.06 (s, 1.5H), 0.05 (s, 1.5H); ^{13}C NMR (100 MHz, $CDCl_3$) 136.0, 135.6 (two peaks), 135.5, 134.8, 133.7, 133.6, 133.4, 133.1, 130.0, 129.8 (two peaks), 129.7, 129.6, 129.5, 127.9, 127.7 (two peaks), 96.7, 75.8, 74.5, 71.1, 70.2, 66.1, 62.5, 31.6, 26.9, 26.5, 26.3, 26.2, 26.0, 25.5, 22.6, 19.2, 18.5, −3.6, −3.7, −3.9, −4.1, −4.6, −4.7, −4.8 (two peaks); ESI-MS *m/z* 694 MNa^+. Compound **6**: IR 1744, 1658; ^1H NMR 7.65–7.58 (m, 4H), 7.46–7.35 (m, 6H), 6.26 (d, 6.8, 1H), 4.39–4.35 (m, 2H), 4.33 (br s, 1H), 3.87–3.79 (m, 3H), 2.02 (s, 3H), 1.07 (s, 9H), 0.94 (s, 9H), 0.82 (s, 9H), 0.16 (s, 3H), 0.11 (s, 3H), 0.08 (s, 3H), 0.07 (s, 3H); ^{13}C NMR 170.9, 168.1, 135.5 (2C's), 132.9, 132.7, 130.0, 129.9, 127.9, 127.8, 78.7, 71.6, 69.5, 60.8, 56.2, 26.9, 26.1, 25.9, 22.8, 19.1, 18.4, 18.3, −3.8, −4.1, −4.9, −5.0; ESI-MS *m/z* 708 MNa^+. Compound **7**: IR 1741; ^1H NMR 7.65–7.61 (m, 4H), 7.47–7.36 (m, 6H), 4.91 (d, 1H, 8.4), 4.37 (dd, 1H, 8.4, 5.6), 4.33 (br s, 1H), 4.00 (app t, 1H, 7.2), 3.92–3.87 (m, 2H), 3.82 (dd, 1H, 10.4, 5.6), 2.51 (s, 3H), 1.08 (s, 9H), 0.92 (s, 9H), 0.85 (s, 9H), 0.17 (s, 3H), 0.13 (s, 3H), 0.10 (s, 3H), 0.05 (s, 3H); ^{13}C NMR 201.9, 166.7, 135.5 (2C's), 132.9, 132.7, 130.2, 129.9, 127.9, 127.8, 78.0, 69.3, 60.9, 60.3, 55.6, 33.7, 26.9, 26.1, 25.9, 19.1, 18.4, 18.3, −3.7, −4.0, −4.9, −5.0; ESI-MS *m/z* 724 MNa^+. Compounds **8/9**: ^1H NMR (300 MHz, 2.6:1 respective mixture, integration given for **8**) 7.65–7.60 (m, 4H), 7.44–7.34 (m, 6H), 5.58 (br s), 5.45 (dd, 6.4, 4.8), 4.18 (br s), 4.14 (d, 1H, 8.4), 4.01 (br d, 1H, 1.8), 4.03–3.91 (m), 3.86 (dd, 5.1, 2.4), 3.82–3.71 (m, 2H), 3.65 (dd, 1H, 7.8, 3.6), 3.58 (dd, 1H, 8.7, 1.8), 2.80 and 2.68 (ABq, 2H, 15.0), 2.47 (s), 2.17 (s, 3H), 1.36 (s, 9H), 1.09 (s), 1.06 (s, 9H), 0.93 (s, 9H), 0.92 (s), 0.86 (s, 9H), 0.23 (s), 0.18 (s, 3H), 0.11 (s), 0.10 (s, 3H), 0.08 (s, 3H), 0.05 (s, 3H); ^{13}C NMR 200.4, 168.0, 163.4, 159.5, 135.6, 135.5, 133.4 (2C's), 133.1, 129.9 (2C's), 129.7 (2C's), 127.8, 127.7 (2C's), 127.6, 104.3, 80.9, 80.8, 79.8, 78.1, 75.6, 75.3, 70.9, 69.9, 69.3, 68.4, 62.0, 61.7, 58.0, 48.6, 34.3, 27.8, 26.4, 26.2, 25.9, 22.3, 19.2, 19.1, 18.5, 18.4, −3.7, −4.1, −4.5, −4.7, −5.0, −5.4; ESI-MS *m/z* 822 MNa^+. Compound **10**: ^1H NMR (400 MHz, CD_3OD) 4.79 (d, 1H, 4.8), 4.26 (ddd, 1H, 7.2, 4.4, 1.2), 4.08 (dd, 1H, 4.8, 2.0), 3.96 (dd, 1H, 12.0, 7.2), 3.87 (dd, 1H, 11.6, 4.4), 2.53 (s, 3H); ^{13}C NMR (100 MHz, CD_3OD) 154.6, 154.2, 132.4, 83.9, 67.1, 66.3, 61.8, 19.1; ESI-MS *m/z* 240 MNa^+. Compound **11**: ^1H NMR (300 MHz, $CDCl_3$)6.15 (d, 1H, 4.5), 5.63 (dd, 1H, 4.5, 2.1), 4.57 (ddd, 1H, 7.5, 4.8, 2.1), 4.41 (dd, 1H, 11.7, 7.5), 4.29 (dd, 1H, 12.0, 4.8), 2.58 (s, 3H), 2.11 (three s, 3H each); ^{13}C NMR (75 MHz, $CDCl_3$) 170.3, 170.0, 169.7, 154.3, 153.9, 126.6, 77.2, 64.5, 64.3, 61.3, 21.0, 20.8, 20.7, 19.8; ESI-MS *m/z* 366 MNa^+. Compound **15**: IR 1773; ^1H NMR (300 MHz, CD_3OD) 5.92 (d, 1H, 5.7), 4.61 (d, 1H, 6.0), 4.59 (d, 1H, 1.5), 3.88–3.82 (m, 2H), 3.59–3.57 (m, 2H), 2.57 (s, 3H); ESI-MS *m/z* 258 MNa^+. Compound **20**: IR 1732; ^1H NMR (300 MHz, $CD_3OD/CDCl_3$) 4.80 (d, 1H, 4.2), 4.57–4.49 (m, 1H), 4.45–4.42 (m, 1H), 4.40 (dd, 1H, 10.5, 3.3), 4.08 (dd, 1H, 4.5, 10.8), 2.55 (s, 3H), 2.09 (s, 3H); ^{13}C NMR (75 MHz, $CD_3OD/CDCl_3$) 171.9, 155.6, 153.3, 131.3, 80.5, 66.5, 65.2, 63.6, 20.9, 18.9; ESI-MS *m/z* 260 MH^+. Compound **21**: IR 1740, 1615; ^1H NMR (300 MHz, $CDCl_3$) 7.66–7.60 (m, 4H), 7.46–7.33 (m, 6H), 4.19 (d, 1H, 8.7), 4.09 and 4.05 (ABq, 2H, 7.2), 3.99 (br d, 1H, 1.5), 3.80–3.64 (m, 3H), 3.59 (dd, 1H, 8.7, 1.8), 2.88 and 2.77 (ABq, 2H, 15.0), 2.17 (s, 3H), 1.18 (t, 3H, 7.2), 1.05 (s, 9H), 0.93 (s, 9H), 0.84 (s, 9H), 0.17 (s, 3H), 0.09 (s, 3H), 0.08 (s, 3H), 0.04 (s, 3H); ^{13}C NMR (50 MHz, $CDCl_3$) 169.0, 168.7, 136.1, 133.9, 130.2, 128.1, 104.8, 77.8, 76.4, 75.8, 69.1, 62.7, 60.9, 47.4, 27.3, 26.7, 26.5, 22.9, 19.6, 19.0, 18.9, 14.5, −3.2, −3.5, −4.0, −4.5; ESI-MS *m/z* 772 MH^+. Compound **22**: ^1H NMR (600 MHz, CD_3CN/D_2O spike, assignments by HSQC and HMBC) 5.86 (dd, 7.4, 1.4, H-4), 4.81 (d, 1.4, H-2), 4.39 (t, 7.4, H-5), 4.35 (dd, 7.1, 3.0, H-6), 4.09–4.03 (m, CH_3CH_2O), 3.75 (dt, 3.0, 6.2, H-7), 3.67–3.59 (m, two H-8), 2.51 (s, CH_3CSNH), 1.20 (t, CH_3CH_2O); ^{13}C NMR (150 MHz, CD_3CN/D_2O spike) 203.7 (CH_3CSNH), 168.3 (C-3), 166.2 (C-1), 88.9 (C-2), 87.7 (C-6), 72.8 (C-5), 70.3 (C-7), 64.6 (C-4), 63.0 (C-8), 59.8 (OCH_2CH_3), 32.9 (CH_3CSNH), 13.8 (OCH_2CH_3); ESI-MS *m/z* 306 MH^+. Compound **25**: ^1H NMR (300 MHz, $CDCl_3$) 4.44 (d, 1H, 8.1), 4.20 and 4.15 (ABq, 2H, 6.9), 4.00–3.80 (m, 4H), 3.62 (td, 1H, 8.1, 3.0), 3.17 (d, 1H, 17.1), 3.08 (d, 1H, 17.4), 2.81 (d, 1H, 8.4), 2.23 (s, 3H), 2.23–2.20 (m, 1H), 2.09–2.05 (m, 1H), 1.27 (t, 3H, 7.2); ESI-MS *m/z* 306 MH^+.