Available online at www.sciencedirect.com
ScienceDirect
Mendeleev Commun., 2016, **26**, 279–281
Mendeleev Communications

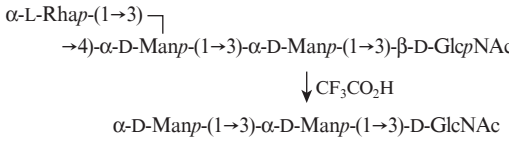
Solvolysis with trifluoroacetic acid: an efficient method for selective cleavage of polysaccharides

Vyacheslav L. L'vov,[a] Andrei V. Filatov,[b] Andrei V. Perepelov,[b] Anna M. Shpirt,[b] Alexander S. Shashkov,[b] Alexander O. Chizhov[b] and Yuriy A. Knirel*[b]

[a] *State Scientific Center 'Institute of Immunology', 115478 Moscow, Russian Federation. Fax: +7 499 618 7156; e-mail: lvovslava@gmail.com*
[b] *N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation. Fax: +7 499 135 5328; e-mail: yknirel@gmail.com*

DOI: 10.1016/j.mencom.2016.07.003

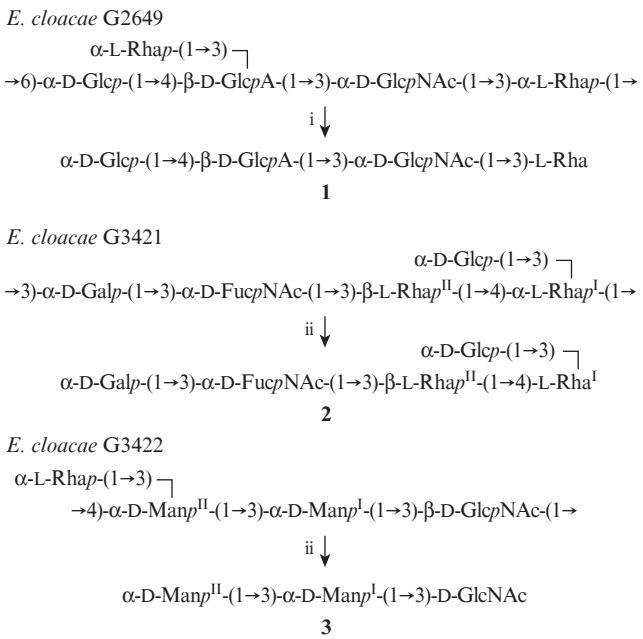
Solvolysis of polysaccharides with trifluoroacetic acid is efficient for selective cleavage of certain glycosidic linkages. This method was successfully used for structure elucidation of polysaccharides and obtaining oligosaccharides from polysaccharides for construction of glycoconjugate vaccines in cases of O-specific polysaccharides of bacteria *Enterobacter cloacae* and *Shigella flexneri*.



Partial cleavage of glycosidic linkages is a useful approach for isolation of oligosaccharide fragments in structural analysis of complex polysaccharides[1] and for construction of glycoconjugate vaccines.[2] As compared with partial hydrolysis, solvolysis with anhydrous reagents can provide higher selectivity and under certain conditions can leave amidic linkages intact that opens a good access to N-acylated amino sugars and uronic acid amides as monomers or oligosaccharide components. For last decades, two reagents were applied for solvolytic cleavage of bacterial polysaccharides, namely anhydrous hydrogen fluoride[3] and trifluoromethanesulfonic acid.[4] However, HF is difficult to handle, CF_3SO_3H is expensive, and both are highly hazardous. In search for a better solvolytic reagent, we tested anhydrous trifluoroacetic acid, which is deprived of the disadvantages of the above reagents. It proved to be highly selective in structural studies of O-specific polysaccharides (OPS) (or O-antigens) of eneteric bacteria *Escherichia coli*.[5–7] In this work, we extended CF_3CO_2H solvolysis to selective cleavage of OPS of other enteric bacteria, *Enterobacter cloacae* and *Shigella flexneri*.

All *E. cloacae* OPS studied included residues of L-rhamnose (6-deoxy-L-mannose) in the main chain (strains G2649 and G3421) or/and in a side chain (strains G2649 and G3422). In order to elucidate the OPS structures, they were depolymerized by treatment with CF_3CO_2H at 50 °C for 4–7 h. After evaporation of the acid in an air flow, work-up with water, and purification by size-exclusion column chromatography on TSK HW-40 (S) fractogel, oligosaccharides were obtained in yields of 20–28%. Their structures were determined by one- and two-dimensional ^{1}H and ^{13}C NMR spectroscopy[8] and negative ion mode ESI HRMS[9] essentially as described.

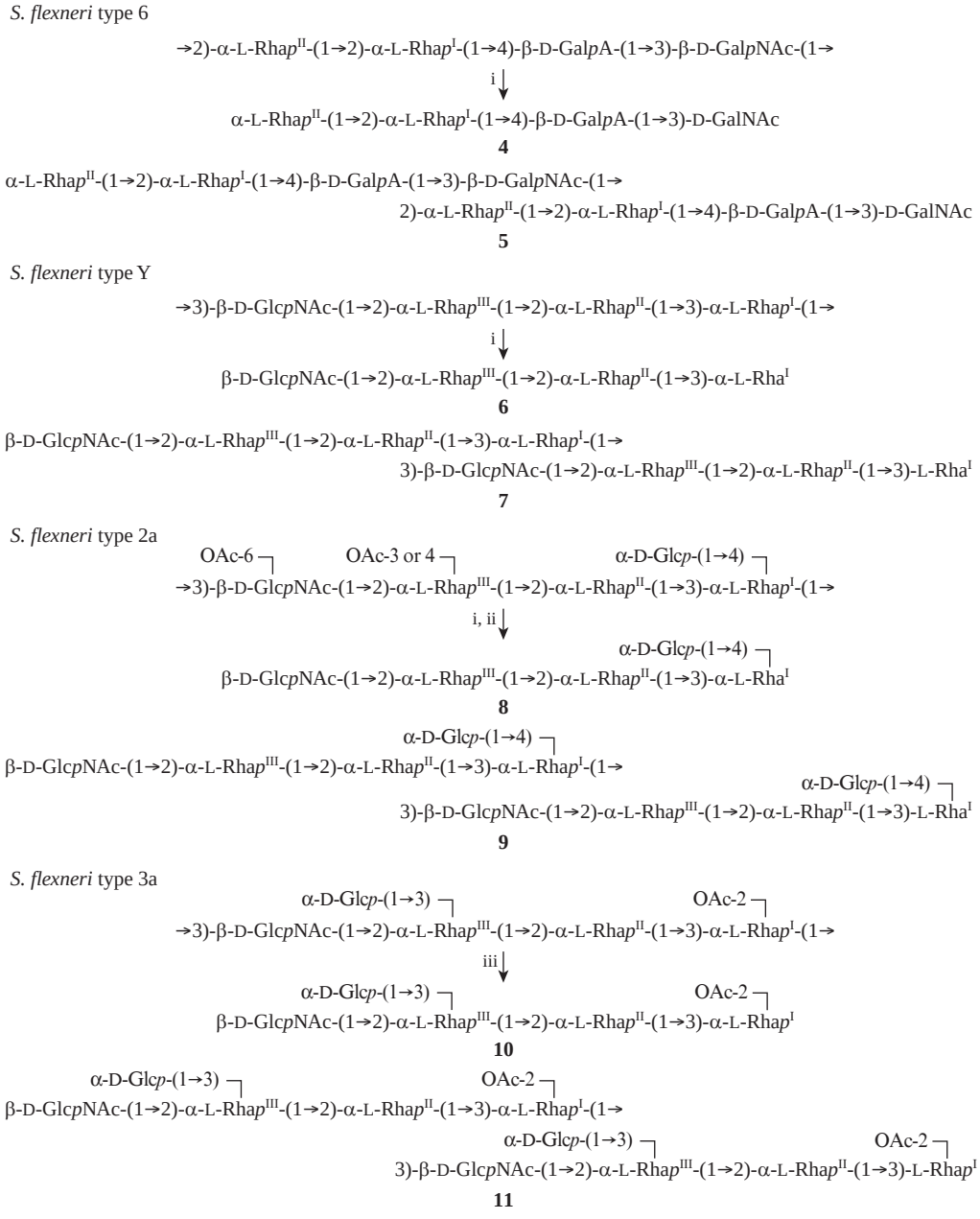
As expected based on the previous findings with L-rhamnose- and L-fucose-containing OPS,[5–7] the rhamnopyranosidic linkages were unstable towards CF_3CO_2H. As a result, oligosaccharides **1**† and **2**‡ with L-rhamnose at the reducing end were obtained from the OPS of strains G2649 and G3421, respectively (Scheme 1). Remarkably, only the α-rhamnopyranosidic linkages were cleaved, whereas the β-L-Rha*p*-(1→4)-L-Rha*p* linkage in the OPS from strain G3421 was unaffected. In addition to splitting off the side-chain L-rhamnose residue, the β-D-Glc*p*NAc-(1→4)-D-Man*p*



Scheme 1 *Reagents and conditions*: i, CF_3CO_2H, 50 °C, 7 h; ii, CF_3CO_2H, 50 °C, 4 h.

† *Compound* **1**. ^{1}H NMR (600 MHz, D_2O) δ: 5.40 (H-1, α-Rha), 4.97 (H-1, β-Rha), 1.29–1.30 (3H, H-6, Rha), 5.46 (H-1, Glc), 5.03, 5.00 (both H-1, GlcNAc), 4.52 (H-1, GlcA), 2.02 (3H, NAc). ^{13}C NMR (150 MHz, D_2O) δ: 95.1 (C-1, α-Rha), 94.9 (C-1, β-Rha), 18.2 (C-6, Rha), 99.6 (C-1, Glc), 95.8, 95.4 (both C-1, GlcNAc), 104.1 (C-1, GlcA), 23.3 (Me, NAc).

‡ *Compound* **2** (after reduction of Rha^{I} into rhamnitol with $NaBH_4$ in water at 20 °C). ^{1}H NMR (600 MHz, D_2O) δ: 5.17 (H-1, Gal), 5.14 (H-1, Glc), 5.02 (H-1, FucN), 4.79 (H-1, Rha), 1.36 (3H, H-6, Rha), 1.28 (3H, H-6, rhamnitol), 2.07 (3H, NAc). ^{13}C NMR (150 MHz, D_2O) δ: 100.1 (C-1, Gal), 96.4 (C-1, Glc), 95.9 (C-1, FucN), 102.0 (C-1, Rha), 18.2 (C-6, Rha), 20.1 (C-6, rhamnitol), 23.5 (Me, NAc).



Scheme 2 *Reagents and conditions*: i, CF_3CO_2H, 40 °C, 1 h; ii, 12.5% NH_4OH, 20 °C, 16 h; iii, CF_3CO_2H, 40 °C, 15 h.

linkage in the OPS from strain G3422 was cleaved affording trisaccharide **3**[§] containing D-GlcNAc at the reducing end (Scheme 1). Using these data, structures of the corresponding OPS shown in Scheme 1 were established (details of the oligosaccharide and OPS structure elucidation will be reported elsewhere).

Then, we demonstrated the applicability of CF_3CO_2H solvolysis for production of oligosaccharides from the OPS of *S. flexneri* types Y, 2a, 3a, and 6, the latter three types being among the most widespread causes of shigellosis (bacillary dysentery). These oligosaccharides are necessary for conjugation with a protein or another carrier to give preparations that would confer immunity against shigellosis and be useful for characterization of specific antibodies[10,11] and development of vaccines.[12,13] Structures of the OPS of *S. flexneri* have already been known[14] (Scheme 2). All OPSs contained α-L-Rha*p* and β-D-Glc*p*NAc or β-D-Gal*p*NAc, whose linkages are expected to be susceptible to CF_3CO_2H cleavage. The OPS were subjected to solvolysis at 40 °C for 1 h, the products were fractionated and analyzed (for the products from type 2a after O-deacetylation) as described above for the oligosaccharides from *E. cloacae*.

In the OPS from type 6, the least stable was the β-D-Gal*p*NAc-(1→2)-L-Rha*p* linkage, which was cleaved selectively (~80%) to give approximately equal amounts of tetrasaccharide **4**[¶] and octasaccharides **5**[††] corresponding to a monomer and a dimer

[§] *Compound* **3**. ^{1}H NMR (600 MHz, D_2O) δ: 5.26, 5.28 (both H-1, Man^I), 5.11 (H-1, Man^{II}), 5.16 (H-1, α-GlcN), 4.76 (H-1, β-GlcN), 2.06 (3H, NAc). ^{13}C NMR (150 MHz, D_2O) δ: 102.1, 102.3 (both C-1, Man^I), 103.5 (C-1, Man^{II}), 92.3 (C-1, α-GlcN), 95.9 (C-1, β-GlcN), 23.4 (Me, NAc).

[¶] *Compound* **4**. ^{1}H NMR (600 MHz, D_2O) δ: 5.40 (H-1, Rha^I), 4.97 (H-1, Rha^{II}), 1.27–1.28 (6H, H-6, Rha^{I,II}), 4.52, 4.57 (both H-1, GalA), 5.24 (H-1, α-GalN), 4.73 (H-1, β-GalN), 2.05 (3H, NAc). ^{13}C NMR (150 MHz, D_2O) δ: 100.9 (C-1, Rha^I), 103.4 (C-1, Rha^{II}), 18.0 (C-6, Rha^{I,II}), 92.6 (C-1, α-GalN), 96.5 (C-1, β-GalN), 105.6, 105.8 (both C-1, GalA), 23.7 (NAc), 23.8 (Me). HRMS (ESI), *m*/*z*: 688.2302 (calc. for $C_{26}H_{43}NO_{20}$, *m*/*z*: 688.2306 $[M-H]^-$).

[††] HRMS (ESI). Compound **5**, *m*/*z*: 1359.4571 (calc. for $C_{52}H_{84}N_2O_{39}$, *m*/*z*: 1359.4578 $[M-H]^-$). Compound **7**, *m*/*z*: 1299.5025 (calc. for $C_{52}H_{88}N_2O_{35}$, *m*/*z*: 1299.5095 $[M-H]^-$). Compound **9**, *m*/*z*: 1623.6142 (calc. for $C_{64}H_{108}N_2O_{45}$, *m*/*z*: 1623.6151 $[M-H]^-$). Compound **10**, *m*/*z*: 862.3189 (calc. for $C_{34}H_{57}NO_{24}$, *m*/*z*: 862.3198 $[M-H]^-$). Compound **11**, *m*/*z*: 1707.6370 (calc. for $C_{68}H_{112}N_2O_{47}$, *m*/*z*: 1707.6363 $[M-H]^-$).

of the oligosaccharide repeat of the OPS (O-unit), respectively (Scheme 2). A small amount of a trimer (~15% of the total) also was isolated.

Fractionation of solvolysis products from the OPS each of types Y and 2a yielded three fractions (I–III). In both cases, the lowest molecular mass fraction III contained tetrasaccharide **6**[‡‡] (from type Y) or pentasaccharide **8**[§§] (from type 2a), which corresponded to the single O-units. The reducing end of this oligosaccharides was occupied by Rha^{I}. Hence, these oligosaccharides arose from the specific cleavage of the α-L-Rha*p*^{I}-(1→3)-D-Glc*p*NAc linkage. The same linkage was cleaved selectively by CF_3CO_2H in the OPS from *E. coli* O49 (authors' unpublished data; for the OPS structure see ref. 15).

Fractions II and I included a dimer (octasaccharide **7**[††] and decasaccharide **9**[††]) and a trimer of the O-units, respectively. They were accompanied by lower molecular mass products resulted from unspecific cleavage (4–10%) of other linkages (those of Rha^{II}, Rha^{III}, Glc, and GlcNAc). The selectivity of the cleavage of the α-L-Rha*p*-(1→3)-D-Glc*p*NAc linkage was 75–80%. An easier cleavage of this linkage was observed in the type 2a OPS, which gave the monomer as the major product, whereas the type Y OPS afforded mainly the trimer (the ratios of fractions I–III were ~1:1:2 and ~3:1.5:1 from types 2a and Y, respectively). The difference is evidently due to 4-substitution of Rha^{I} with Glc in the type 2a OPS.

The OPS of type 3a was much more stable towards solvolysis, and no significant cleavage was observed at 40 °C for 1 h. Prolongation of the treatment to 3 h yielded a mixture of higher oligosaccharides with Rha^{I} at the reducing end, the lowest one being a trimer. Longer solvolysis caused a deeper OPS cleavage but the selectivity decreased significantly. Thus, after 15 h a ~1 : 3 mixture of monomer **10**[††] and dimer **11**[††] was obtained with selectivity of ~30%, and partial O-deacetylation (~30%) occurred. A higher stability of the type 3a OPS could be accounted for by 2-O-acetylation of Rha^{I}.

Therefore, CF_3CO_2H solvolysis is useful for production of dimers and trimers from *S. flexneri* types Y, 2a, and 6 or a trimer and higher oligosaccharides from *S. flexneri* type 3a. It is evident that, if necessary, higher oligosaccharides can be obtained from all types by solvolysis under milder conditions, *e.g.* at 20 or 30 °C. Importantly, solvolysis for short time did not affect O-acetyl groups, which may be associated with specific immunodeterminants within O-antigens.

In summary, the present data confirmed that CF_3CO_2H is a highly selective reagent for solvolytic cleavage of glycosidic linkages, which can be applicable both in structural studies of polysaccharides and preparation of oligosaccharide fragments suitable for conjugation. The least stable towards this reagent are linkages of 6-deoxy-α-hexopyranosides (α-Rha*p*, α-Fuc*p*) and 2-acetamido-2-deoxy-β-hexopyranosides (β-Glc*p*NAc, β-Gal*p*NAc). Other linkages (those of hexopyranoses, 2-acetamido-2-deoxy- and 2-acetamido-2,6-dideoxy-α-hexopyranoses, hexuronic acids) are essentially stable at temperatures of 50 °C or below. The odd-seeming instability of the β-Glc*p*NAc and β-Gal*p*NAc linkages under mild acidic conditions, in which other aldopyranosidic linkages are unaffected, is in fact known; *e.g.* it has been reported for cases of treatment of bacterial polysaccharides with 1% HOAc at 100 °C[16] or 48% HF at 0 °C.[17] Solvolytic cleavage with CF_3CO_2H is sensitive to neighboring monosaccharides and non-carbohydrate substituents, and a different environment may forward reaction to different directions, as was observed for the OPS of *S. flexneri* type 6 *vs*. the other *S. flexneri* types studied. Acid CF_3CO_2H is significantly milder than HF and CF_3SO_3H, which cleave α- and β-hexopyranosidic and 2-acetamido-2,6-dideoxy-α-hexopyranosidic linkages, in some cases also linkages of hexuronic acids and 2-acetamido-2-deoxy-α-hexopyranoses, even at lower temperatures.[3,4] Further studies will enable better understanding fine regularities of the CF_3CO_2H solvolysis and show whether this reagent is applicable for selective cleavage of polysaccharides by other, more stable glycosidic linkages at higher temperatures.

Work on the OPS of *E. cloacae* was supported by the Russian Foundation for Basic Research (grant no. 14-04-00663), and that of *S. flexneri* by the Russian Science Foundation (grant no. 14-14-01042).

References

1 G. O. Aspinall, *Acc. Chem. Res.*, 1987, **20**, 114.
2 J. Y. Wang, A. H. C. Chang, H.-K. Guttormsen, A. L. Rosas and D. L. Kasper, *Vaccine*, 2003, **21**, 1112.
3 Y. A. Knirel, E. V. Vinogradov and A. J. Mort, *Adv. Carbohydr. Chem. Biochem.*, 1989, **47**, 167.
4 Y. A. Knirel and A. V. Perepelov, *Aust. J. Chem*., 2002, **55**, 69.
5 A. V. Perepelov, A. V. Filatov, Q. Wang, V. L. L'vov, Y. Quan, A. S. Shashkov, L. Wang and Y. A. Knirel, *Carbohydr. Res*., 2014, **388**, 30.
6 L. Jiang, A. V. Perepelov, A. V. Filatov, B. Liu, A. S. Shashkov, S. N. Senchenkova, L. Wang and Y. A. Knirel, *Carbohydr. Res*., 2014, **397**, 27.
7 A. V. Perepelov, X. Guo, A. V. Filatov, B. Liu and Y. A. Knirel, *Carbohydr. Res*., 2015, **416**, 32.
8 A. V. Perepelov, V. L. L'vov, B. Liu, S. N. Senchenkova, M. E. Shekht, A. S. Shashkov, L. Feng, P. G. Aparin, L. Wang and Y. A. Knirel, *Carbohydr. Res*., 2009, **344**, 1588.
9 A. M. Tsedilin, A. N. Fakhrutdinov, D. B. Eremin, S. S. Zalesskiy, A. O. Chizhov, N. G. Kolotyrkina and V. P. Ananikov, *Mendeleev Commun.*, 2015, **25**, 454.
10 N. I. A. Carlin, M. A. J. Gidney, A. A. Lindberg and D. R. Bundle, *J. Immunol.*, 1986, **137**, 2361.
11 N. I. A. Carlin, D. R. Bundle and A. A. Lindberg, *J. Immunol.*, 1987, **138**, 4419.
12 A. Phalipon, M. Tanguy, C. Grandjean, C. Guerreiro, F. Belót, D. Cohen, P. J. Sansonetti and L. A. Mulard, *J. Immunol*., 2009, **182**, 2241.
13 H. F. Said, A. Phalipon, M. Tanguy, C. Guerreiro, F. Belót, B. Frisch, L. A. Mulard and F. Schuber, *Vaccine*, 2009, **27**, 5419.
14 A. V. Perepelov, M. E. Shekht, B. Liu, S. D. Shevelev, V. A. Ledov, S. N. Senchenkova, V. L. L'vov, A. S. Shashkov, L. Feng, P. G. Aparin, L. Wang and Y. A. Knirel, *FEMS Immunol. Med. Microbiol*., 2012, **66**, 201.
15 A. V. Perepelov, Q. Wang, S. N. Senchenkova, S. D. Shevelev, A. S. Shashkov, L. Feng, Y. A. Knirel and L. Wang, *Biochemistry (Moscow)*, 2008, **73**, 406 (*Biokhimiya*, 2008, **73**, 498).
16 S. N. Senchenkova, A. S. Shashkov, Y. A. Knirel, N. K. Kochetkov, K. Zych and Z. Sidorczyk, *Carbohydr. Res.*, 1996, **293**, 71.
17 N. A. Kocharova, J. E. Thomas-Oates, Y. A. Knirel, A. S. Shashkov, U. Dabrowski, N. K. Kochetkov, E. S. Stanislavsky and E. V. Kholodkova, *Eur. J. Biochem*., 1994, **219**, 65.

Received: 8th February 2016; Com. 16/4837

‡‡ *Compound* **6**. ^{1}H NMR (600 MHz, D_2O) δ: 5.09 (H-1, α-Rha^{I}), 4.88 (H-1, β-Rha^{I}), 5.17 (H-1, Rha^{II}), 5.21 (H-1, Rha^{III}), 1.28–1.31 (9H, H-6, $Rha^{I–III}$), 4.70 (H-1, GlcN), 2.05 (3H, NAc). HRMS (ESI), *m*/*z*: 658.2562 (calc. for $C_{26}H_{45}NO_{18}$, *m*/*z*: 658.2564 $[M-H]^-$).

§§ *Compound* **8**. ^{1}H NMR (600 MHz, D_2O) δ: 5.06 (H-1, α-Rha^{I}), 4.89 (H-1, β-Rha^{I}), 5.06, 5.08 (both H-1, Rha^{II}), 5.16 (H-1, Rha^{III}), 1.40–1.42 (3H, H-6, Rha^{I}), 1.29 (6H, H-6, $Rha^{II,III}$), 5.19, 5.24 (both H-1, Glc), 4.71 (H-1, GlcN), 2.06 (3H, NAc). HRMS (ESI), *m*/*z*: 820.3056 (calc. for $C_{32}H_{55}NO_{23}$, *m*/*z*: 820.3092).