NEW KNOWLEDGE IN THE CHEMISTRY OF ANTIMETABOLITES

F. ŠORM

Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, Prague 6

ABSTRACT

Exotoxin from *B. thuringiensis* was purified by chromatography and its structure was determined by means of chemical and physical methods. The formula of the substance as given above indicates that the exotoxin possesses between ribose and glucose an unusual ether bond which up to now has not been found in Nature. The study of the mechanism of action of the exotoxin shows that it competes with ATP in the RNA polymerase reaction. The anomalous structure of the exotoxin has been confirmed by the synthesis of its substantial moiety, i.e. of glucosyladenosine.

In the Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, Prague, we have been interested for a long time in antimetabolites of nucleic acids as potential cancerostatics and virostatics. Most attention has been paid to synthetic analogues of bases and nucleosides. Of these analogues, 6-azauridine is now being used clinically. Recently it was suggested on the basis of chemical speculations another aza-analogue of pyrimidine nucleosides, namely 5-azacytidine, might be of value and it was therefore prepared synthetically in this Institute[1]. We never expected that 5-azacytidine might occur in Nature. Surprisingly enough, this substance was identified as the cancerostatically active antibiotic and naturally occurring antimetabolite of *Streptoverticillium ladacanus*[2]. This finding drew our attention to substances which interfere with the biosynthesis of nucleic acids, and to investigations on the exotoxin of *Bacillus thuringiensis* which is known to exhibit insecticidal activity. It has been found in the biochemical laboratories of our Institute that this exotoxin is a specific inhibitor of DNA-dependent ribonucleic acid polymerase and that it competes specifically with adenosine triphosphate[3].

The structure of exotoxin was established in this Institute[4]. In this lecture, I should like to present the complex data relating to the proposal of the unusual structure of exotoxin as well as to the synthesis of a portion of its molecule.

Nothing was known about the chemical structure of exotoxin until recently. Its heat stability indicated a low molecular weight. The first not quite pure preparations of exotoxin were obtained by de Barjac and Dedonder[5] of the Pasteur Institute in Paris. These workers analysed it chemically and found it to

contain adenine, ribose and phosphorus in equimolar ratios. This and the acid character of the exotoxin suggested a nucleotide of anomalous composition.

In Czechoslovakia, research on *B. thuringiensis* was for a long time in the hands of entomologists, in particular Dr Vaňková who developed from it an insecticidal preparation. In cooperation with her we set out in 1966 to isolate and identify the exotoxin of *B. thuringiensis*. The credit here goes mainly to Dr Šebesta (a biochemist) and to Dr Farkaš (an organic chemist).

For producing the exotoxin on a large scale we used a variety of *B. thuringiensis*, the so-called *B. gelechiae*, which produces no endotoxin. We know now that the exotoxin of this variety is identical with the substance used by the French workers.

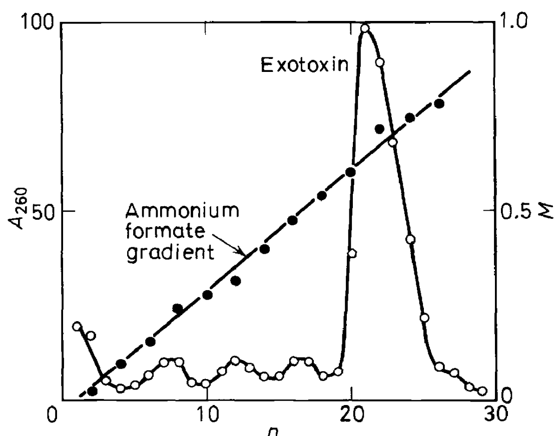


Figure 1. Chromatography of exotoxin on Dowex 1 × 2 (formate)

In the first part of the work we developed a procedure for the preparation of about 10 g amounts of exotoxin. We grew *Bacillus gelechiae* in a simple synthetic medium and, after filtration, adsorbed the exotoxin on charcoal. After desorption of the exotoxin with ethanol, and extraction of ballast substances into phenol, the principal part of the purification procedure was chromatography on Dowex 1. Gradient elution with ammonium formate yielded a uniform fraction absorbing in the u.v. region (*Figure 1*). After sublimation of ammonium formate, the purified toxin appeared as a yellowish amorphous powder. The product was fully active biologically as was shown by a standard test on the caterpillars of *Galleria mellonella*.

The compound was readily soluble in water and displayed a spectrum characteristic of adenine nucleotides. The i.r. spectrum showed a particularly characteristic absorption band 1691 cm^{-1}, corresponding to a carboxyl. From the content of nitrogen (8.7 per cent) and phosphorus (4.2 per cent) we concluded that the molecular weight of the compound would be about 750.

The electrophoretic behaviour of the exotoxin showed that we were dealing with an acidic compound. Its titration curve was determined (*Figure 2*), from which it followed that the exotoxin molecule contains four dissociable protons; three of these with a pK of 3.8 apparently belong to two carboxyl groups and to

the more acid hydrogen atom of an ester-linked phosphoric acid, the fourth with a pK of 6.6 to the less dissociable hydrogen atom of the phosphate residue. Naturally we tried first to split off the ester-linked phosphoric acid of the exotoxin which was readily achieved by applying calf intestinal alkaline phosphatase. The product was again purified by gradient elution on a column of Dowex 1. The fraction absorbing u.v. light was freeze-dried and a white amorphous, biologically completely inactive compound was obtained. The i.r. spectrum showed a principal band belonging to a carboxyl group. From the content of nitrogen, assuming that all nitrogen was present in the adenine moiety (10.4 per cent), we concluded that the dephosphorylated exotoxin has a molecular weight of about 650. In agreement with this assumption the titration curve of dephosphorylated exotoxin corresponds to the presence of only two carboxyl groups.

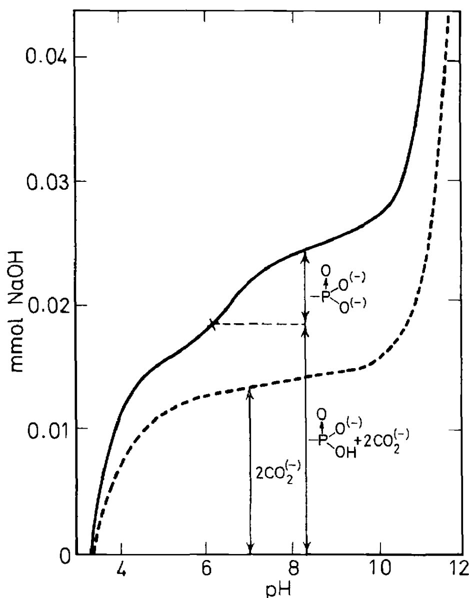


Figure 2. The titration curve of exotoxin (4.60 mg, solid line) and of dephosphorylated exotoxin (4.60 mg, dashed line)

The presence of two carboxyl groups in the molecule of exotoxin, and its dephosphorylated product, accounts for the chemical changes of the two compounds which were repeatedly observed and during which a γ-lactone with a characteristic absorption band in the i.r. region at 1 775 cm^{-1} was formed. By chromatography of older preparations of the exotoxin in the form of free acid we isolated a compound containing, in addition to a bound phosphoric acid, only one carboxyl group and one γ-lactone group in the molecule (*Figure 3*). This γ-lactone was apparently formed by an internal esterification of one of the exotoxin carboxyl groups. Similarly, the dephosphorylated toxine yielded, after short standing in a weakly acid medium, two new compounds, both of them containing one carboxyl and one γ-lactone in the molecule (*Figure 4*).

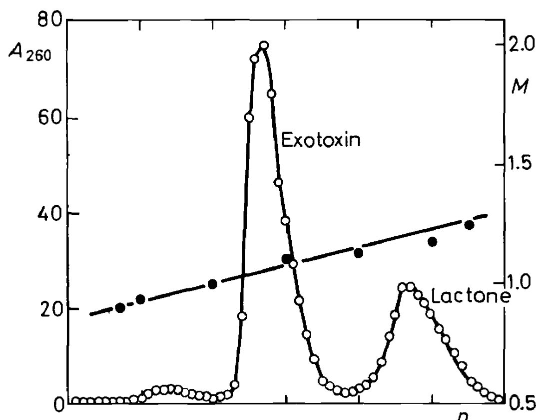


Figure 3. Rechromatography of exotoxin on Dowex 1 × 2 (formate)

One of these compounds is stable, the other is readily hydrolysed back to a dicarboxylic acid. Hence it follows that only one of the two carboxyl groups present is always lactonized.

From these preliminary data alone it is evident that exotoxin is a complicated system evidently containing sugar residues, perhaps with two carboxylic groups, in addition to adenine and bound phosphoric acid. The first step toward the elucidation of its structure was, of course, hydrolysis for which we first took the enzymatically dephosphorylated product. After it had been heated with 0.1 M HCl at 70°C, we isolated adenine from the products of hydrolysis in an amount corresponding to the supposed molecular weight of exotoxin. Among other products of hydrolysis we isolated chromatographically a sugar component, the low mobility of which suggested that it could be a disaccharide. In no case were we able to prove the presence of ribose in the hydrolysate, although the French authors describe its occurrence. However, as after hydrolysis with

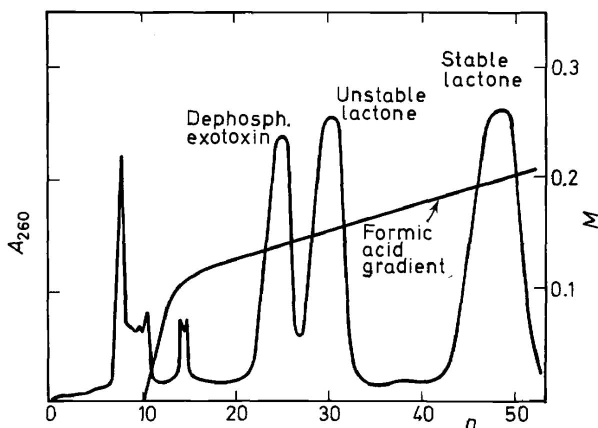


Figure 4. Separation of the mixture obtained by lactonization of the dephosphorylated exotoxin on Dowex 1 × 2 (formate)

aqueous acids the yield of the disaccharide component was very low—the yield of the product after chromatography being only about ten per cent of the weight of the starting exotoxin—we undertook the cleavage of the dephosphorylated exotoxin by methanolysis. After the separation of adenine we submitted the reaction mixture to alkaline hydrolysis with 0.1 M NaOH and on a column of Dowex 1 (acetate) we obtained a neutral fraction A, and an acid fraction B. As

Dephosphorylated exotoxin

0.1 M HCl in methanol
6 h
(Dowex 50, H^+)

Adenine

Neutral components

0.1 M NaOH
1 h, 50°
(Dowex 1, CH_3COO^-)

Neutral fraction A

Acid fraction B

fraction A, representing probably a mixture of anomeric methyl glycosides, could be chromatographically separated only with difficulty, we transformed its sugar components into the corresponding acetonides. On thin-layer chromatography, as well as on column chromatography we obtained three well-separated fractions. The predominant fraction (70 per cent) was used for physicochemical measurements. Using acetic anhydride in pyridine we transformed it into a non-crystalline acetate and measured its n.m.r. spectrum (*Figure 5*). In the high field region of the spectrum typical signals of methyl protons are seen, two singlets of the isopropylidene grouping, three singlets of the acetyl groups, and two singlets of the methoxy groups. From the presence of the signals of these groups it can be concluded that the substance is a triacetate of an

Figure 5. The n.m.r. spectra at 100 MHz of the main methyl glycoside acetonide triacetate of exotoxin disaccharide

acetonide containing two glycosidically bound methyl groups. Other parts of the spectrum contain signals of protons of the O—CH type.

From a detailed analysis of this spectrum by frequency-swept decoupling experiments it was evident that this part of the spectrum is composed of a superposition of two isolated spin systems, one a six-spin and the other a seven-spin system. The symmetry of these spin systems suggests that one component of the disaccharide is a pentosofuranoside and the other a hexosopyranoside. The six-spin system is well resolved in the spectrum measured in benzene.

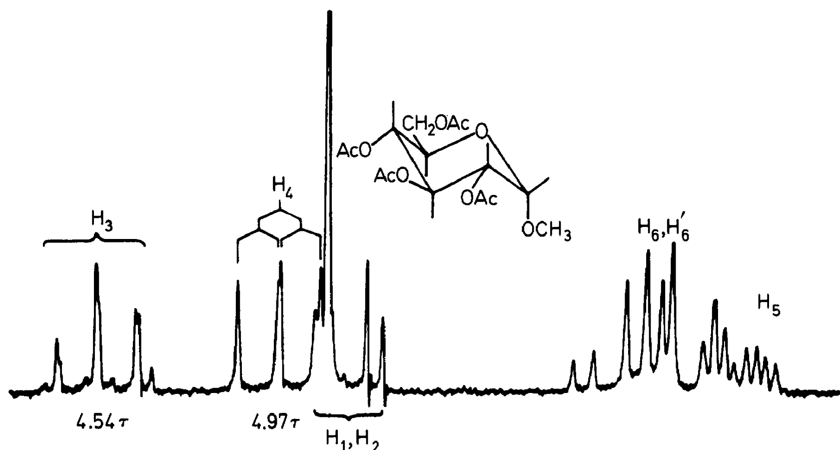


Figure 6. The n.m.r. spectrum at 100 MHz of methyl tetra-*O*-acetyl-α-D-glucopyranoside in deuterochloroform

Characteristic features of this spin system are vicinal interactions $J_{1,2}=0$ and $J_{3,4}<1$ Hz. These weak interactions show that the configurations of H_1—H_2 and H_3—H_4 protons are *trans*. The lack of appreciable solvent shifts in benzene shows that the O—CH bonds in the furanoside part of the molecule are weakly polar, and hence rather of the ether type. From this it follows that the isopropylidene group is bound in the furanoside component, which means that protons H_2—H_3 are mutually *cis* oriented and that the furanoside part of the disaccharide molecule is a β-methylribofuranoside. The similarity with the spectrum of β-methyl-2,3-isopropylidene-D-ribofuranoside also points to this conclusion. The ribose component is evidently bound by an ether-type bond to C_5, as this is also shown by the chemical shift of the proton H_5.

A characteristic feature of a seven-spin system is its large values of coupling constants $J_{2',3'}$, $J_{3'4'}$ and $J_{4',5'}$ (approximately 9–10 Hz). From this it can be judged that these protons have an alternating diaxial conformation, typical of the C1 conformation of glucopyranosides. For the sake of comparison the n.m.r. spectrum of the tetra-*O*-acetyl derivative of α-D-glucopyranoside is given in the *Figure 6*. The hexapyranoside part of the spectrum of the disaccharide corresponds to the n.m.r. spectrum of the acetylated α-methyl-D-glucoside. The spectra differ only in the symmetry of the spin system of the $H_{6'}$ and $H_{5'}$ protons and, mainly, by the position of the quartet of the proton $H_{4'}$. While in the tetraacetate of α-D-glucopyranoside the proton at $C_{4'}$ is at 5.03 p.p.m., in the

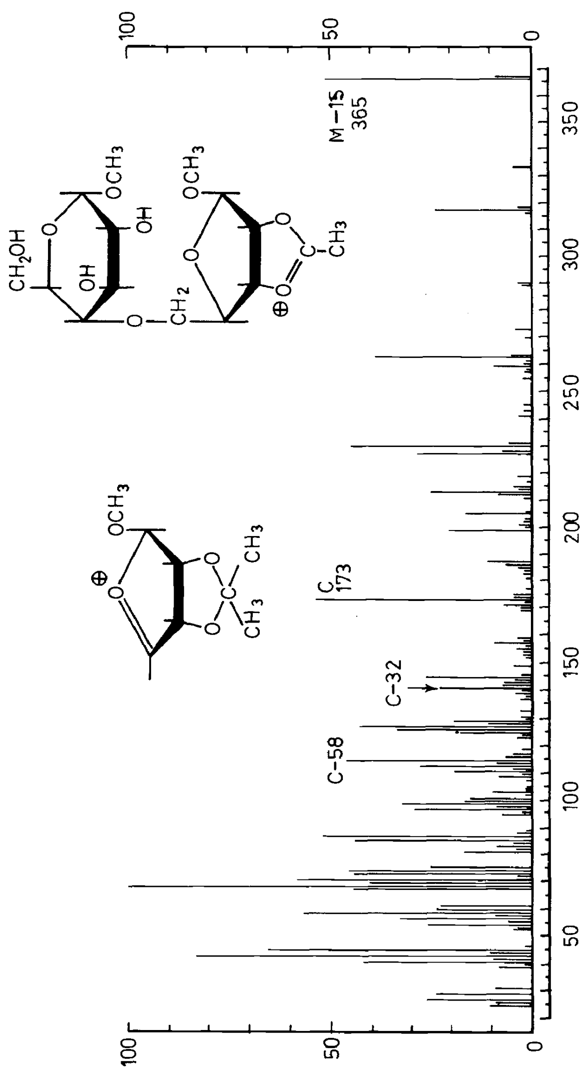
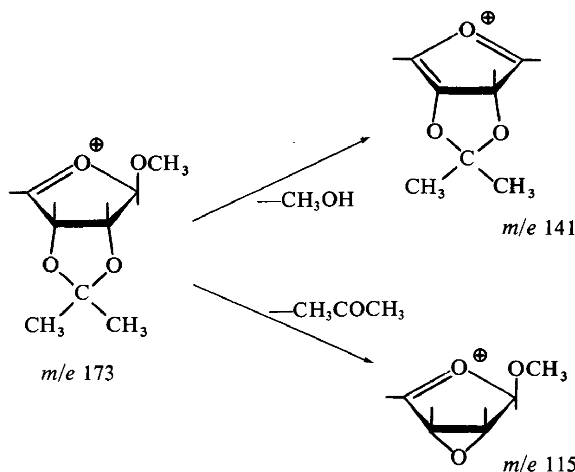


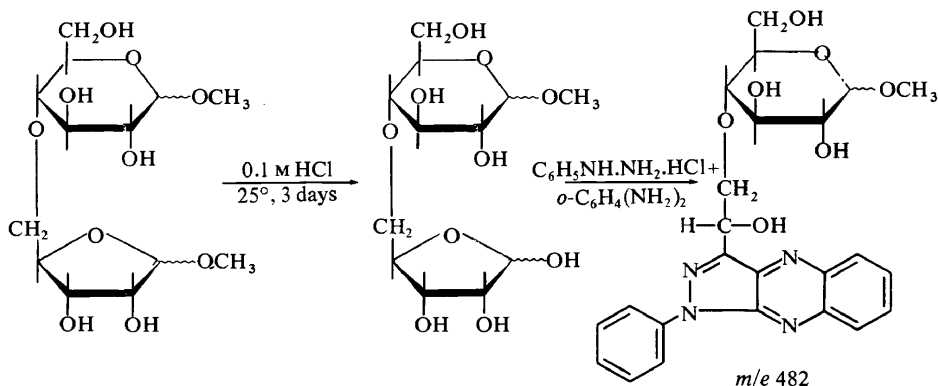
Figure 7. The mass spectrum of the main methyl glycoside acetonide of exotoxin disaccharide

spectrum of disaccharide it appears at 3.46 p.p.m. This difference of chemical shifts shows that in the disaccharide an ether-type bond is present at $C_{4'}$.

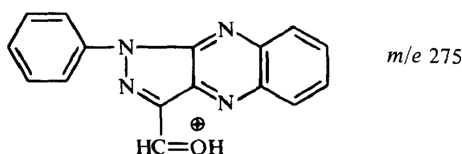
From the analysis of the n.m.r. spectrum of the acetonide triacetate it follows that the disaccharide, isolated from the exotoxin, is composed of ribose and glucose bound together by an unusual ether bond between the carbon atom C_5 of ribose and the carbon atom $C_{4'}$ of glucose.

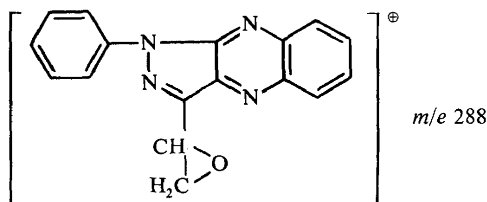
We tried to obtain support for the above structure by mass spectroscopic studies on the free acetonide (*Figure 7*). In its mass spectrum the highest peak appears at *m/e* 365. By analogy with the mass spectra of other isopropylidene derivatives we consider that this peak belongs to the ion M-15 formed by the splitting off of the methyl group from the isopropylidene group. The occurrence of an ion of mass 173 is also indicative of the presence of a ribose residue with a propylidene grouping, which we believe is further fragmentated to ion 141 by the loss of methanol and to ion 115 by the loss of acetone.



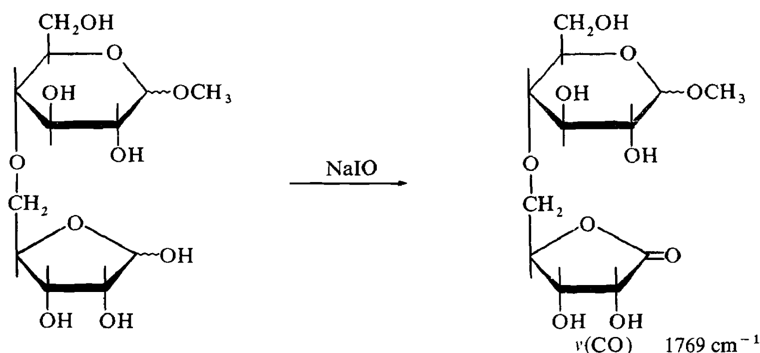


In order to decide whether ribose and glucose, representing the components of the toxine disaccharide, belong to the D- or the L-series, we tried to transform the ribose part into the corresponding 1,4-lactone. For this reason we first submitted the mixture of anomeric glycosides obtained by methanolysis to partial hydrolysis with 0.1 M HCl at room temperature. The reducing monomethylglycoside of the disaccharide thus formed was a syrupy material and therefore we tried to characterize it by conversion into a crystalline derivative. A derivative of flavazole was found to be the most suitable for this purpose. As we had already seen earlier with a series of model compounds, flavazole derivatives give a molecular ion in the mass spectrum. In agreement with the proposed structure the disaccharide-flavazole derivative gave a molecular peak at *m/e* 482. The spectrum also contained strong peaks at *m/e* 275 and 288 which are typical of a flavazole nucleus. The reducing disaccharide mono-





methylglycoside was then transformed by hypoiodite oxidation into a lactone. This was characterized both by its i.r. spectrum, which contained a distinct band at 1 769 cm^{-1} corresponding to a five-membered lactone, and by the n.m.r. spectrum of its penta-acetyl derivative. The ORD spectrum of this lactone is



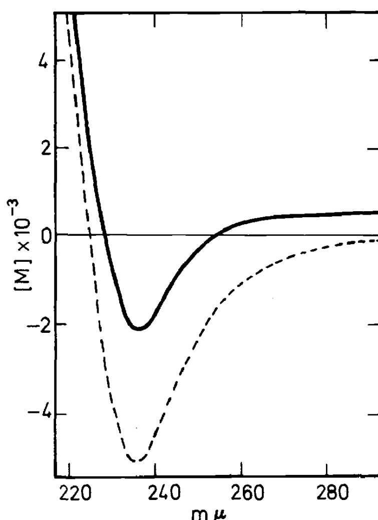
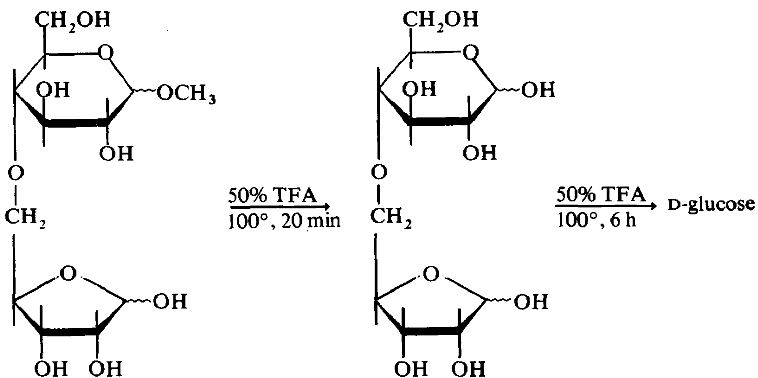


Figure 8. ORD spectra in water of 'disaccharidelactone' (solid line) and D-ribono-1,4-lactone (dashed line)

practically identical with the ORD spectrum of D-ribonolactone (*Figure 8*). On the basis of a study of ORD curves of five-membered sugar lactones, undertaken recently by Japanese authors, it can be judged that the disaccharide lactone has in its ribose part the same configuration as D-ribonolactone. Hence, the pentose component of the disaccharide is D-ribose. For determination of the stereochemistry of the glucose part of the molecule we tried to obtain the sugar itself by a vigorous hydrolysis. By reaction with 50 per cent trifluoroacetic acid at 100°C for 20 minutes we succeeded in obtaining a disaccharide with two free aldehydic groups. This disaccharide can be obtained in low yield, as I mentioned at the beginning, by direct acid hydrolysis of exotoxin. If the hydrolysis was prolonged to six hours, the disaccharide was further hydrolysed and a monosaccharide was obtained as the only product. This was identified by paper chromatography, and by gas chromatography of its silyl derivative, as glucose. The assignment of this sugar to the D-series was carried out by an enzymatic test, using glucose oxidase.



In this way we have determined the structure of the disaccharide component of exotoxin, and I should now like to devote attention to the acid component, which we called fraction B, isolated after methanolysis. This component crystallized spontaneously. The n.m.r. spectrum of this substance in dimethyl sulphoxide-d_6 is very simple. It contains only two broadened singlets of equal height at 4.17 p.p.m. and 3.95 p.p.m. This simple spectrum indicated that it must be due to a dicarboxylic acid, completely symmetric, having six carbon atoms in the molecule. These conditions are fulfilled by only two substances, galactaric and allaric acids. The i.r. spectrum of compound B measured in KBr is different from that of galactaric acid, but identical with that of allaric acid. The identity of both compounds was further corroborated by the i.r. spectra of their methyl esters. From the study of the products of hydrolysis it followed that dephosphorylated exotoxin is composed of adenine, an etherically bound disaccharide composed of D-ribose and D-glucose, and also allaric acid. What remained to be determined was the way in which these components are joined together.

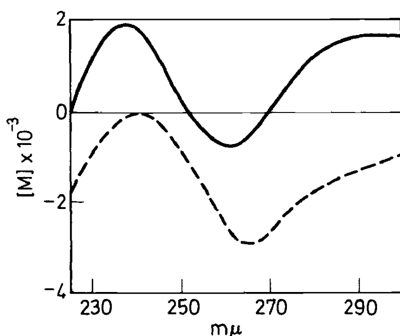
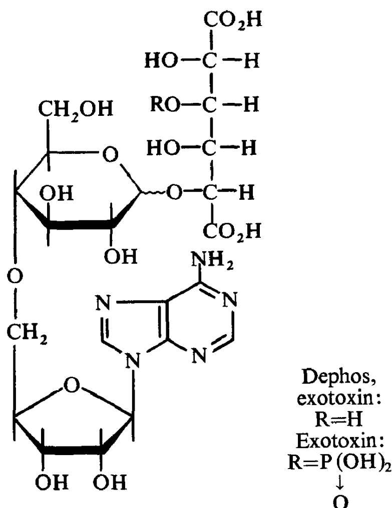


Figure 9. ORD spectra in 0.1 M HCl of exotoxin (solid line) and adenosine-5′-triphosphate (dashed line)

Adenine is substituted in position 9; positions 7 or 3 are excluded by the fact that both exotoxin and the product after dephosphorylation have λ_{max} at 260 nm, while derivatives of adenine substituted in positions 3 or 7 have maxima at 270 nm. Adenine can be connected to the anomeric centre of either D-ribose or D-glucose. The connection of adenine to D-ribose follows from the n.m.r. spectrum of the dephosphorylated product. The spectrum of this part is in good agreement with the n.m.r. spectra of 5′-substituted adenosines. The configuration of the anomeric centre in ribose, carrying the adenine residue, follows from the similarity in shape of the ORD curves of exotoxin and adenosine triphosphate (*Figure 9*). It now remained to solve the link of this anomalous nucleoside with allaric acid. In view of the relatively easy hydrolysis of this link, it can be concluded that allaric acid is bound by a glycosidic bond with C_1 of glucose. The position of the substitution in allaric acid was not proved, but we can conclude on the basis of the formation of two different five-membered mono-γ-lactones from the dephosphorylated exotoxin in acid medium, which I have already mentioned at the beginning of my talk, that the substituent is in the neighbourhood of the carboxyl group. The problem of the configuration on the

anomeric centre of glucose, as well as the configuration of the substituted carbon atom in allaric acid, is the object of our further studies.

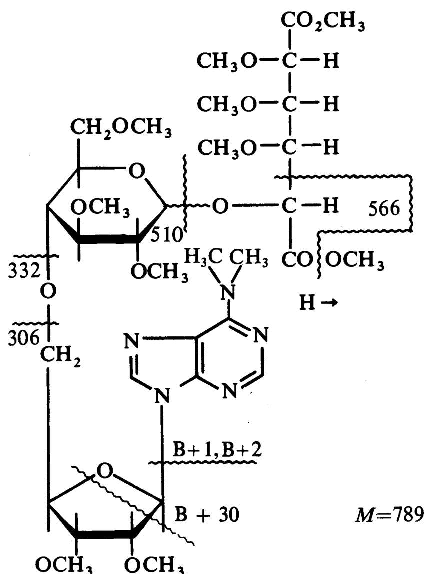
I should like now to return to the structure of exotoxin. On methanolysis it affords, in addition to adenine, a mixture of anomeric glycosides and also an acid component containing phosphoric acid bound by an ester link with the molecule of allaric acid. We deduce the position of the phosphate residue on the allaric acid from the behaviour of exotoxin on oxidation with periodic acid. Exotoxin consumes one mole of periodic acid within five minutes, which corresponds to the oxidation of a *cis*-diol grouping in the furanose part of the molecule. After dephosphorylation of exotoxin the consumption of periodic acid during the rapid phase of the oxidation is increased to two moles. This shows that allaric acid in exotoxin does not contain a 1,2-diol grouping, as this is blocked by a phosphate residue. Therefore, we believe that the phosphate residue is in postion 4 of allaric acid.



Dephos, exotoxin: R=H
Exotoxin: R=P(OH)$_2$ → O

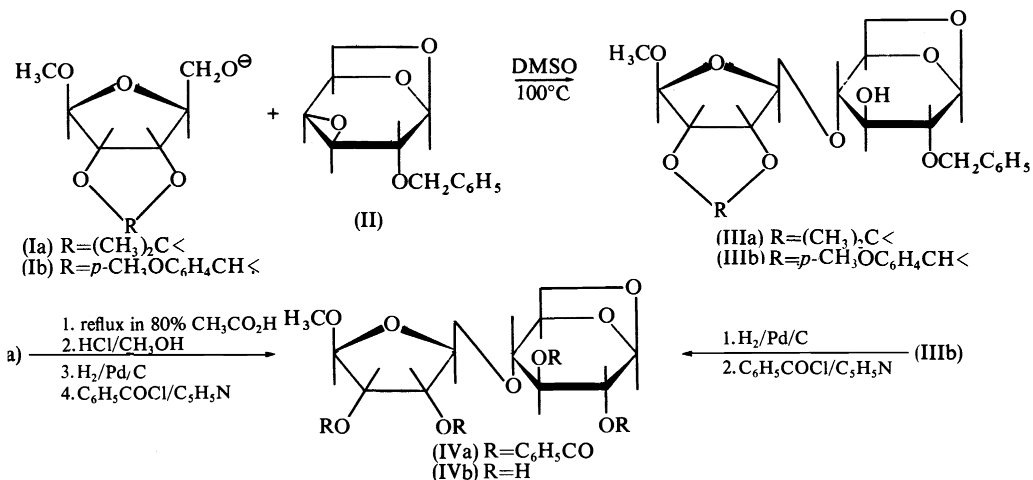
The correctness of this structure was proved by the mass spectrum of the permethyl derivative of dephosphorylated exotoxin (the permethylation was effected with methyl iodide in dimethylformamide with the use of sodium hydride as the base). The molecular ion at m/e 789 and the ionic fragments were in full accordance with the structure proposed.

It is logical that we have tried to support the unusual structure of exotoxin synthetically. Inspection of the exotoxin molecule reveals numerous problems. Thus, for example, the connection of two sugar residues by an ether linkage is quite novel and has not hitherto been reported in the chemistry of sugars. For this reason, the synthesis of exotoxin was started from the most difficult side, namely, by the preparation of the disaccharidic fragment (obtained from the naturally occurring material by an acid methanolysis). This work was done by Dr Prystaš.



The reaction of 2-*O*-substituted 1,6:3,4-dianhydro-β-D-galactopyranoses with substituted D-ribofuranosides possessing a free hydroxylic function at position 5 appeared to be the most promising route, especially in view of the acid- or alkali-catalysed opening of the epoxide ring, reported recently in the synthesis of 1,6-anhydro-β-D-glucopyranoses alkylated at position 4[6].

On the basis of preliminary acid- and alkali-catalysed epoxide ring opening experiments, the latter catalyst was used. 2-*O*-Benzyl-1,6:3,4-dianhydro-β-D-galactopyranose (II) and methyl-2,3-*O*-isopropylidene- (Ia) or *p*-anisylidene-β-D-ribofuranoside (Ib) served as reaction components. Their condensation in



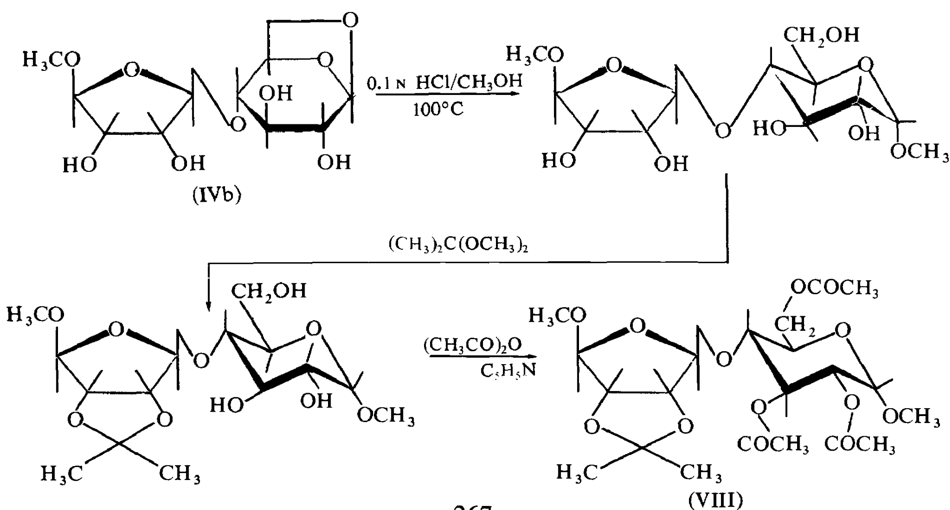
(IV) $\xrightarrow{H_2SO_4/(CH_3CO)_2O}$ (V) $R{=}C_6H_5CO$ $\xrightarrow[CH_2Cl_2]{HBr/CH_3CO_2H}$

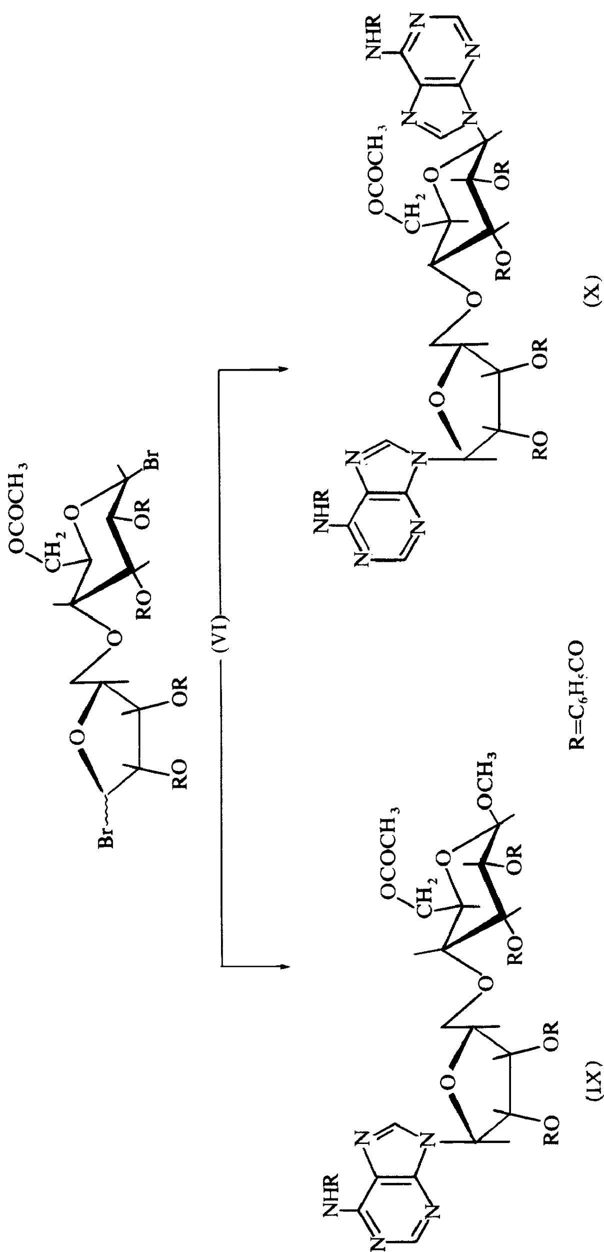
(VI) $R{=}C_6H_5CO$ $\xrightarrow{CH_3OH/Ag_2O}$ (VIIa) $R^1{=}C_6H_5CO;\ R^2{=}CH_3CO$
(VIIb) $R^1{=}R^2{=}H$
(VIIc) $R^1{=}R^2{=}CH_3CO$

$\xrightarrow[CH_3CO_2H{-}(CH_3CO)_2O\ 10:1]{34\%\ HBr}$ (VIII) $R{=}CH_3CO$ $\xrightarrow[Ag_2O]{CH_3OH}$ (VIIc)

excess sodium salt of dimethyl sulphoxide afforded compounds IIIa and IIIb, respectively, in 40 to 60 per cent yields. In the former case, the isopropylidene group was removed in refluxing 80 per cent aqueous acetic acid. This treatment, however, was accompanied by a partial hydrolysis of the methyl riboside. For this reason, the product was subjected to the action of methanolic hydrogen chloride prior to the hydrogenolytical removal of the benzyl group and protection of the free hydroxylic functions. The resulting anomeric mixture of tetrabenzoates (IVa) was separated by chromatography on silica gel. The same product was obtained by hydrogenolysis of compound (IIIb) and the subsequent benzoylation of the crystalline ether (IVb). The opening of the 1,6-anhydro ring of laevoglucosan was effected by an acid-catalysed acetolysis. This reaction was accompanied by cleavage of the methyl glycoside linkages and their replacement by acetate residues. The resulting triacetate (V) was converted into the dihalogenose (VI) which then afforded the protected diglycoside ether (VIIa) with the β-configuration at the anomeric centres 1 and 1′. The penta-acetate (VIIc) was also prepared from compound (IIIa) by the action of hydrogen bromide in a mixture of acetic acid and acetic anhydride. This treatment results in a simultaneous removal of protecting groups, opening of the 1,6-anhydro ring, acetylation, and conversion into the acetylated dihalogenose. Reaction of the latter compound with methanol led to the required diglycoside ether derivative (VIIc).

In the course of our earlier investigations on the structure of exotoxin, the disaccharide with the assumed β- and α-configuration at the anomeric centres 1 and 1′, respectively, was converted by isopropylidenation and subsequent acetylation into the compound (VIII). This substance has now been prepared with the use of the above-mentioned synthetic derivatives. Thus, the acid-catalysed methanolysis of the crystalline ether (IVb) afforded an anomeric mixture of methyl ribofuranoside and methyl glucopyranoside ether predominantly with the configuration β at position 1 and α at position 1′. This mixture was converted into the isopropylidene derivative and this was acetylated after chromatography on silica gel. The n.m.r. spectrum of this product was in every




OCOCH3
CH2
Br
RO
OR
(VI)
NHR
(X)
(IX)
OCH3
R=C6H5CO

respect identical with that of the specimen obtained from the naturally occurring exotoxin. On the other hand, the n.m.r. spectrum of the anomeric ether, also prepared by Dr Prystaš, was quite different. In this manner, we have unequivocally confirmed the presence of the unusual disaccharide grouping in the molecule of exotoxin.

Furthermore, the dihalogenose (VI) was used in the synthesis of the nucleoside which contained adenine and the disaccharidic component. In this preparation, use was made of the considerably different reactivity of the protected ribofuranosyl bromide and glucopyranosyl bromide towards the chloromercuric salt of N^6-benzoyladenine. Reaction of equimolecular amounts of the chloromercuric salt and the dihalogenose (VI) followed by treatment with methanol in the presence of silver oxide afforded compound (IX). Compound (X), containing two adenine residues in the molecule, was obtained by glycosylation with the use of an excess of the chloromercuric salt. This approach opened up the route to a total synthesis of the dephosphorylated exotoxin as well as of the exotoxin itself.

REFERENCES

1. A. Pískala and F. Šorm. *Coll. Czech. Chem. Commun.* **29**, 2060 (1964); F. Šorm, A. Pískala, A. Čihák and J. Veselý, *Experientia*, **20**, 202 (1964).
2. M. E. Bergy and R. R. Herr. *Antimicrobial Agents and Chemotherapy*, 625 (1966); L. J. Hanka, J. S. Evans, D. J. Mason and A. Dietz. *Antimicrobial Agents and Chemotherapy*, 619 (1966).
3. K. Šebesta and K. Horská. *B.B.A.* **169**, 281 (1968).
4. J. Farkaš, K. Šebesta, K. Horská, Z. Samek, L. Dolejš and F. Šorm, *Coll. Czech. Chem. Commun.* **34**, 1118 (1969).
5. H. de Barjac and M. R. Dedonder, *C. R. Acad. Sci., Paris*, **260**, 7050 (1965); *Bull. Soc. Chim. Biol* **50**, 941 (1968).
6. L. J. Carlson, *J. Org. Chem.* **30**, 3953 (1965); J. E. Höök and B. Lindberg, *Acta Chem. Scand.* **20**, 2363 (1966); M. Černý, L. Kalvoda and J. Pacák, *Coll. Czech. Chem Commum*, **33**, 1143 (1968).