Biochem. J. (1992) 285, 957–964 (Printed in Great Britain)

Photolabelling of *Salmonella typhimurium* LT2 sialidase

Identification of a peptide with a predicted structural similarity to the active sites of influenza-virus sialidases

Thomas G. WARNER,* Reed HARRIS,* Robert McDOWELL* and Eric R. VIMR†
*Genentech, Inc., 460 Point San Bruno Blvd., South San Francisco, CA 94070, and
†Department of Pathobiology, College of Veterinary Medicine, University of Illinois, Champaign-Urbana, Urbana, IL 61801, U.S.A.

The sialidase from *Salmonella typhimurium* LT2 was characterized by using photoaffinity-labelling techniques. The well-known sialidase inhibitor 5-acetamido-2,6-anhydro-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-enonic acid (Neu5Ac2en) was modified to contain an amino group at C-9, which permitted the incorporation of 4-azidosalicylic acid in amide linkage at this position. Labelling of the purified protein with the radioactive (^{125}I) photoprobe was determined to be highly specific for a region within the active-site cavity. This conclusion was based on the observation that the competitive inhibitor Neu5Ac2en in the photolysis mixture prevented labelling of the protein. In contrast, compounds with structural and chemical features similar to the probe and Neu5Ac2en, but which were not competitive enzyme inhibitors, did not affect the photolabelling of the protein. The peptide interacting with the probe was identified by CNBr treatment of the labelled protein, followed by *N*-terminal sequence analysis. Inspection of the primary structure of the protein, predicted from the cloned structural gene for the sialidase [Hoyer, Hamilton, Steenbergen & Vimr (1992) Mol. Microbiol. **6**, 873–884] revealed that the label was incorporated into a 9.6 kDa fragment situated within the terminal third of the molecule near the *C*-terminal end. Secondary-structural predictions using the Garnier–Robson algorithm [Garnier, Osguthorpe & Robson (1978) J. Mol. Biol. **120**, 97–120] of the labelled peptide revealed a structural similarity to the active site of influenza-A- and Sendai-HN-virus sialidases with a repetitive series of alternating β-sheets connected with loops.

INTRODUCTION

Sialidases (neuraminidases, EC 3.2.2.18) are a family of enzymes which cleave sialic acid residues from the oligosaccharide side chains of both glycoproteins and gangliosides. These *N*-acylneuraminosyl-glycohydrolases are widely distributed in nature and serve a variety of important metabolic functions (Corfield & Schauer, 1982). For example, the influenza-virus enzyme resides on the viral coat surface and has at least two critical roles, facilitating the process of infection and promoting the release of progeny virus from the host cell (Air & Laver, 1989). Pathogenic bacteria may also exploit sialidases for a similar adaptive advantage for infection, and many strains actively secrete large amounts of the enzyme. Of equal importance are the mammalian sialidases located in several organelles, including the plasma membrane, cytosol and lysosomes (Tettamanti *et al.*, 1981). The lysosomal form is of particular interest in humans because deficiencies of the enzyme result in the neurodegenerative disorder, sialidosis (Lowden & O'Brien, 1979).

Of the microbial and mammalian sialidases, the influenza-A-virus enzyme is the best understood and characterized. Protein crystals have been obtained and the three-dimensional structure at 0.29 nm (2.9 Å) resolution has been determined, revealing a putative substrate-binding site containing several amino acid residues which are highly conserved between many influenza-virus strains (Colman *et al.*, 1983). Site-specific mutagenesis of selected amino acids lining the active site has further implicated the important role of these residues in the catalytic process (Lentz *et al.*, 1987). Bacterial sialidases have not been investigated at this level and, thus far, amino acid residues directly involved in substrate recognition and catalysis have not been identified. Similarly, topological and structural features of the active-site region common to the bacterial proteins remain to be revealed. A 12-amino-acid repeat containing five conserved residues centred about aspartic acid has been noted in sialidases from *Clostridium sordellii*, *Salmonella typhimurium*, *Vibrio cholerae* (Roggentin *et al.*, 1989) and *Bacteroide fragilis* (Russo *et al.*, 1990). The function of these 'Asp boxes' is unknown at the present time. In general, bacterial sialidases share up to 35% primary sequence identity, with an overall similarity of about 50% (Hoyer *et al.*, 1992).

In an attempt to identify secondary- and tertiary-structural similarities between the bacterial enzymes, particularly at the substrate-binding site, we have employed affinity-labelling techniques using the transition-state analogue of the sialidase reaction, 5-acetamido-2,6-anhydro-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-enonic acid (Neu5Ac2en), modified to contain a photoreactive functional group. These studies are based on our earlier discovery that the attachment of aryl azide units at the C-9 position of the Neu5Ac2en molecule does not alter recognition of the inhibitor molecule by the plasma membrane or the lysosomal sialidase (Warner, 1987; Warner *et al.*, 1991). We report here photolabelling of the sialidase from *S. typhimurium* LT2 and the identification of protein-derived peptide modified by the affinity reagent. A comparative analysis was made of the predicted secondary structure of the peptide with the analogous active-site regions of viral sialidases. In addition, in order to gain insight into the relative orientation between the amino acids labelled by the probe and the catalytic residues of the enzyme, we estimated the optimal conformational states of the probe molecule and calculated the distance between the anomeric centre of the Neu5Ac2en moiety at C-2 and the aryl nitrene.

Abbreviations used: Neu5Ac2en, 5-acetamido-2,6-anhydro-3,5-dideoxy-D-*glycero*-D-*galacto*-non-2-enonic acid; TFA, trifluoroacetic acid; 9-*N*-ASA-Neu5Ac2en, 5-acetamido-2,6-anhydro-9-(4′-azido-2′-hydroxy-benzamido)-3,5,9-trideoxy-D-*glycero*-D-*galacto*-nonuloso-2-enonic acid; ASA, 4-azidosalicylic acid; DOPC, 3,4-dihydro-2,2-dimethyl-4-oxo-2*H*-pyran-6-carboxylic acid; BME, β-mercaptoethanol; PVDF, poly(vinylidene difluoride).

EXPERIMENTAL

Reagents

4-Methylumbelliferyl-*N*-acetylneuraminic acid was from Sigma Chemical Co. (St. Louis, MO, U.S.A.). 3,4-Dihydro-2,2-dimethyl-4-oxo-2*H*-pyran-6-carboxylic acid (DOPC) was from Aldrich Chemical Co. (Milwaukee, WI, U.S.A.). T.l.c. was performed on pre-coated silicic acid glass-backed plates. Detection on the plates was made with resorcinol spray (Svennerhom, 1963) or with ninhydrin (1% in ethanol). ^{125}I as NaI was from Amersham International.

N.m.r. spectroscopy

1H-n.m.r. data were recorded on a Varian VXR-300S n.m.r. spectrometer operating at 300 MHz in the Fourier-transform mode. The samples were subjected to four freeze–thaw cycles in ^2H_2O and the spectra recorded in this solvent at 25 °C. Acetone was included as an internal standard referenced to external sodium 3-(trimethylsilyl)[$^2H_4$]propionate. For samples prepared in [^{2}H]methanol, tetramethylsilane was used as reference standard.

I.r. spectroscopy

Spectra were recorded on a Nicolet model 510 FT-IR instrument with the sample prepared in methanol and dried as a film on to NaCl plates.

M.s.

Elemental composition was determined by high-resolution fast-atom-bombardment m.s. using a JEOL HX110HF mass spectrometer operating in the two-sector mode with the sample suspended in a matrix of nitrobenzyl alcohol. The molecular ion of the solvent matrix served as the calibration standard. Molecular-mass determinations were made at low resolution using CsI for calibration.

PAGE and autoradiography

After photolysis, samples were analysed under denaturing conditions by SDS/PAGE as described by Laemmli (1970). After drying, the gels were exposed to X-ray film (usually overnight) at −70 °C using a Cronex intensifying screen (Du Pont). The amount of radioactive tracer incorporated into the protein was determined by cutting the bands from the gel and counting ^{125}I using a Apex automatic γ-radiation counter (Micro Medic, Costa Mesa, CA, U.S.A.).

Photolysis

Photolabelling was carried out in 150 mM-acetate buffer, pH 5.5, in the presence of 3–5 μCi of probe, 2 μg of protein, and 5 mM-β-mercaptoethanol, in a total volume of 30 μl. The mixture containing the protein and probe was placed on Parafilm (American National Can, Greenwich, CT, U.S.A.) maintained at 4 °C on ice, and allowed to equilibrate in the dark for several minutes. Immediately before photolysis, β-mercaptoethanol (BME) was added and the sample was positioned about 1–2 cm from the light source. Irradiation was carried out for 60 s with a UVGL-25 mineral lamp (UVP, San Gabriel, CA, U.S.A.) at 254 nM.

Amino acid sequencing

N-terminal sequencing was carried out by automated Edman degradation was a model 470 A Applied Biosystems gas phase sequencer equipped with a 120 A phenylthiohydantoin-amino acid analyser.

CNBr cleavage

After photolysis, the protein sample was denatured with guanidinium chloride, subjected to reductive carboxymethylation with iodoacetic acid, and treated with CNBr in 70% formic acid overnight (Roll *et al.*, 1978).

Tricine/SDS/PAGE

The peptide fragments released by CNBr treatment were resolved using 16.5%-(w/v)-polyacrylamide gels with Tricine buffer (Schagger & Von Jagow, 1987), developing the gel overnight at 75 V at room temperature. The peptides were electrophoretically transferred to poly(vinylidene difluoride) (PVDF) membranes for sequencing (Matsudaria, 1987). After transfer, the gel was stained for protein and subjected to autoradiography to identify the radiolabelled fragment. The size of the fragments was estimated using a commercial low-molecular-mass 2512–16949 kDa markers derived from myoglobin (LKB, Bromma, Sweden).

Salmonella sialidase

The sialidase of *S. typhimurium* LT2 was overproduced as a recombinant protein in *Escherichia coli* LE392(pSX62) (Hoyer *et al.*, 1992). A partially purified (90–95%) preparation was used for most of the initial photolysis experiments. A more highly purified preparation was used to identify the peptide fragment labelled with the probe (Hoyer *et al.*, 1991).

5-Acetamido-2,6-anhydro-9-(4′-azido-2′-hydroxy-benzamido)-3,5,9-trideoxy-D-*glycero*-D-*galacto*-nonuloso-2-enonic acid (9-*N*-ASA-Neu5Ac2en)

5-Acetamido-2,6-anhydro-9-azido-3,5,9-trideoxy-D-*glycero*-D-*galacto*-nonuloso-2-enoic acid (5.3 mg, 15 μmol), prepared as described previously (Warner *et al.*, 1991), was added to a solution of 1.0 ml of pyridine/water (7:3, v/v) and 13 μl of triethylamine (100 μmol). After purging with N_2, 10 μl of propane-1,3-dithiol (100 μmol) was added and the reaction vessel was sealed. The mixture was maintained at 50–55 °C for 6 h with mechanical stirring. The reaction progress was monitored by t.l.c. (System A: butanol/acetic acid/water, 4:2:2, by vol); the R_F values were 0.47 and 0.29 for the azide and the amine respectively, detection being with resorcinol spray and ninhydrin. After cooling the solution to room temperature, the solvent was removed under vacuum. The residue was suspended in water and applied to a column (0.5 cm × 3 cm) containing Dowex 50 (H^+ form). After washing with 5 ml of water, 5 ml of methanol and finally 5 ml of water, the product, 5-acetamido-9-amino-2,6-anhydro-3,5,9-trideoxy-D-*glycero*-D-*galacto*-nonuloso-2-enonic acid (9-amino-Neu5Ac2en) was eluted with 5 ml of 2 M-HCl and the solvent removed under vacuum.

The resulting amine was not purified further, but it was coupled with *N*-hydroxysuccinimidyl-4-azidosalicylic acid (Pierce Chemical Co.; 8.6 mg, 31 μmol) in 1.0 ml of dimethylformamide/water (1:1, v/v) adjusted to pH 8.5 with $NaHCO_3$. After 1 h at room temperature in the dark, the pH was adjusted to pH 1.0 with HCl and the solvent removed under vacuum. The product was isolated by h.p.l.c. on a column containing octadecasilane on a silica support (Ultrasphere; 4.6 cm × 25 cm; Beckman, Berkeley, CA, U.S.A.), equilibrated in 0.1% trifluoracetic acid in water.

The column was eluted with a linear gradient of increasing amounts of 0.1% TFA/acetonitrile up to 100% in 40 min at a flow rate of 3.0 ml/min, the eluate being monitored at 240 nm. The appropriate fraction was collected and the solvent removed, giving 2.2 mg or a 30% yield of the homogeneous product (R_F 0.68). ^{1}H-n.m.r. data (^2H_2O): δ 1.975 (s, 3H, N–CO–CH_3),

5.730 (d, 1H, *J* 2.1 Hz, C=CH); 6.352 (apparent s, 1 H, *m*-phenyl); 6.400 (apparent d, 1 H, *J* 9.8 Hz, *m*-phenyl); 7.825 (apparent d, 1 H, *J* 9.7 Hz, *o*-phenyl). I.r. data: film on NaCl (max. 2118 cm^{-1} N_3). Elemental composition calculated for $C_{18}H_{22}N_5O_9$: theoretical molecular mass+H^+: 452.1418 Da; found: 452.1421 Da.

Iodination of 9-*N*-ASA-Neu5Ac2en

A 1 mg portion of 1,3,4,6-tetrachloro-3α,6α-diphenylglycouril (Iodogen; Pierce) in 0.25 ml of dry dichloromethane was dried as a film on to the walls of a glass reaction vessel. After the removal of the solvent, NaI (0.32 mg, 2.2 μmol) in 40 μl of water was added to the vessel. 9-*N*-ASA-Neu5Ac2en (1.2 μg, 2.2 μM) in 0.2 ml of 10 mM-phosphate, pH 7.4, was added, followed by 200 μl of phosphate buffer wash. The mixture was maintained at room temperature with occasional agitation. After 10 min, the reaction was terminated by the addition of 0.1 ml of 0.1 M-HCl, followed by filtration through a 0.22 μm-pore-size nylon membrane (Rainin Instruments). The reaction vessel and the filter were washed with 0.2 ml of water. The filtered solution and wash were combined and the mixture was applied to a Sep-Pak C_{18} cartridge column (Waters Associates). The column was washed with 5.0 ml of water, and the iodinated products were eluted with 5 ml of acetonitrile, followed by removal of solvent under vacuum. Reaction products were purified by h.p.l.c. using a C_{18} reversed-phase column as described above for the starting material. The products were obtained as amorphous solids, giving a major band on t.l.c. analysis: $R_F = 0.7$, 0.7 and 0.74 respectively (System A). 1H-n.m.r. data [^{2}H]methanol: compound 1; δ 2.002 (s, 3 H, N–CO–CH_3), 5.950 (d, 1 H, *J* 2.2 Hz, C=CH), 6.774 (s, 1 H, C-3′ aromatic), 8.242 (s, 1 H, C-6′aromatic); compound 2; δ 2.014 (s, 3 H, N–CO–CH_3), 5.956 (d, 1 H, *J* 2.4 Hz, C=CH), 6.815 (d, 1 H, *J* 9 Hz, C-5′ aromatic), 7.826 (d, 1 H, *J* 9 Hz, C-6′ aromatic); compound 3; δ 2.026 (s, 3 H, N–CO–CH_3), 5.948 (d, 1 H, *J* 2.8 Hz, C=CH), 8.297 (s, 1 H, C–6′ aromatic). I.r. data: compound 1, 2 and 3, film on NaCl max. 2118–2120 cm^{-1}. Determined molecular masses ($M-Na^+$) for each of the three fractions as their sodium salts were 600.2, 600.2 and 726.2 Da respectively.

[^{125}I]Iodination of 9-*N*-ASA-Neu5Ac2en

The preparation of ^{125}I-labelled probe was essentially the same as that described for the unlabelled material, except that the reaction was carried out on a greatly reduced scale. In a typical preparation, 1 mCi of Na^{125}I (0.5 nmol, 2175 mCi/μmol) was added to a reaction vessel containing a film of 10 μg of iodogen and a solution of 9-*N*-ASA-Neu5Ac2en (0.5 nmol) in 60 μl of 10 mM phosphate buffer, pH 7.4. After 10 min at room temperature the reaction was terminated with 20 μl of 0.1 M-HCl and the products isolated as described for the unlabelled products. About 0.85 mCi of product was obtained, with a specific radioactivity estimated to be about 2175 mCi/μmol. A single major radiolabelled product, containing about 90% of the applied radioactivity (R_F 0.7), along with several minor labelled products, were detected by t.l.c. analysis. Radioactive products were revealed by autoradiography.

Molecular modelling of 9-ASA-Neu5Ac2en

A model of the photoprobe was constructed from the crystal structures of salicylamide (Sasada *et al.*, 1964) and Neu5Ac2en (Furuhata *et al.*, 1988) on the basis of co-ordinates deposited in the Cambridge Crystallographic Database (Allen *et al.*, 1979). The 9-azido group was appended using standard bond lengths and angles. The model was assembled with Version 5.41 of the SYBYL software system (Tripos Associates, St. Louis, MO, U.S.A.), and subjected to simple energy minimizations using the Tripos force field (Clark *et al.*, 1989); the sialic acid ring was held rigid during the minimization to avoid distortion due to possible force-field artefacts. An alternative conformation of the probe was generated on the basis of solution-phase n.m.r. studies of sialic acid (Czarniecki & Thornton, 1976); this model was also subjected to a similar minimization protocol. As a third alternative, the possible spacings between the anomeric centre at C-2 of the pyranose ring and the aryl azide were examined by systematically varying all exocylic torsions connecting the groups (excluding the salicylamide linkage) in 10° intervals. The linear distance between these two points in the molecule was noted for those conformations within 41.84 kJ (10 kcal) of the global minimum (as estimated using the Tripos force field).

Secondary-structure prediction

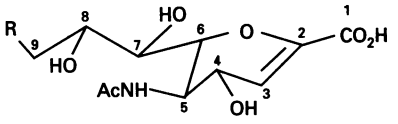
The secondary structure of the proteins were predicted from the amino acid sequence using a computer program which implemented the Garnier–Robson algorithm (Garnier *et al.*, 1978).

RESULTS AND DISCUSSION

Characterization of the photoprobe and its radioactively labelled derivative

We have recently reported on the synthesis of a tritiated photoreactive aryl azide conjugate of Neu5Ac2en which has been invaluable for characterizing the sialidase among a complex mixture of proteins in a partially purified preparation from human placenta (Warner *et al.*, 1990). In similar photolabelling studies with bacterial sialidases, we observed that the *S. typhimurium* enzyme was labelled with a much lower efficiency than the human enzyme. This made it necessary to prepare a new probe containing a radioactive tracer having higher specific radioactivity than that used previously (> 500 μCi/μmol). 4-Azidosalicylic acid (ASA) has been widely employed as a means for incorporating a photoreactive functionality which can be made radioactive with ^{125}I (Ji & Ji, 1982; Ji *et al.*, 1985). The ASA moiety was conveniently coupled here in an amide linkage using a 9-amino derivative of Neu5Ac2en (Fig. 1). The amine was prepared from 9-azido-Neu5Ac2en (Warner *et al.*, 1991) using propane-1,3-dithiol as a potent and highly selective reductant (Bayley *et al.*, 1978). Conventional hydrogenation catalysts, such as $Pd/BaSO_4$, did not provide the specificity required, and under a variety of conditions the concomitant reduction of the 2,3 double bond was detected (Meindl & Tuppy, 1973).

Coupling of the photoreactive ASA as its activated *N*-hydroxysuccinimidyl ester to 9-amino-Neu5Ac2en proceeded in good yield. Spectral data and elemental composition of the final product were consistent with its anticipated structure. When this material was made radioactive by the incorporation of ^{125}I, several radioactive compounds were detected when the reaction mixture was examined by t.l.c. We initially suspected that the unsaturated pyran portion of the Neu5Ac2en moiety might be susceptible to the strong oxidative conditions employed for iodination, and this could account for the multiple products that were observed. Since it would be difficult to interpret clearly the results of photolabelling the protein with this complex mixture of radioactive substrates whose structures were not precisely defined, we carried out additional chemical and physical characterization of the reaction products by iodination of 9-*N*-ASA-Neu5Ac2en on large scale without the addition of Na^{125}I. When this crude mixture was examined by t.l.c., two major products were detected that presented with a pattern coincident for that observed with the radioactively labelled material. Subsequent analysis of the reaction mixture by reversed-phase h.p.l.c. revealed a total of three major iodination products (Fig. 2; fractions



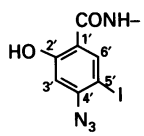
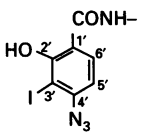
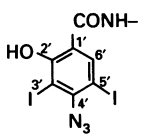
Substituent at C-9	Derivative
N_3	9-Azido-Neu5Ac2en
NH_2	9-Amino-Neu5Ac2en
	9-N-ASA-Neu5Ac2en
	Compound 1
	Compound 2
	Compound 3

Fig. 1. Structures of Neu5Ac2en derivatives modified at C-9

1, 2, and 3) and a small amount of residual starting material (fraction S). Molecular-mass determinations for the components in each of the major fractions were consistent with mono-iodination of 9-*N*-ASA-Neu5Ac2en giving rise to compounds 1 and 2 and di-iodination resulting in compound 3. ^1H n.m.r. spectral analysis of these compounds confirmed the position of iodine substitution on the salicylic amide unit. The signals for the remaining protons of the Neu5Ac2en molecule were similar among all three products and they were nearly identical with those of the starting material. Together these data confirmed that the Neu5Ac2en moiety was not significantly affected by the iodination conditions employed and that only the salicylic azido moiety was modified.

Photolabelling *Salmonella* sialidase

The partially purified recombinant sialidase used in the photolabelling studies gave a single major band (41 kDa) when analysed by PAGE under denaturing conditions (Fig. 3, lane A) (Hoyer *et al.*, 1991). In order to verify active-site-specific labelling with [*ASA*-^{125}I]9-*N*-ASA-Neu5Ac2en, photolysis experiments were carried out in the presence of various additives. These included a nitrene quenching reagent, a substrate analogue known to be competitive enzyme inhibitor, and a compound of related structure that is not a sialidase inhibitor.

The effects of BME on photolabelling of the sialidase were determined (Fig. 3, lanes B and C). It is well known that thiols decrease non-specific labelling by scavenging the solution-phase activated photoprobe molecules, which randomly interact with the protein (Bayley & Staros, 1984). Probe molecules bound to

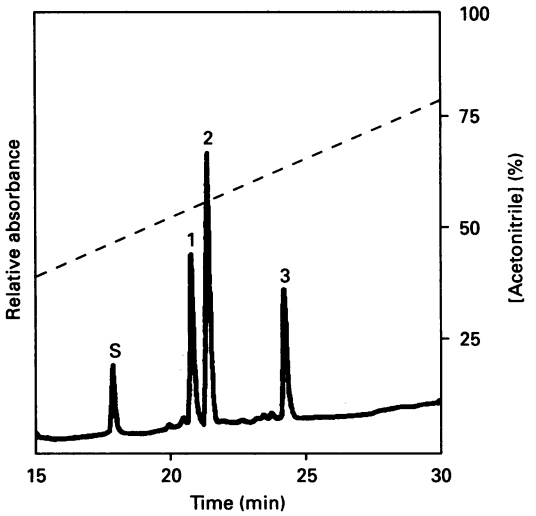


Fig. 2. H.p.l.c. chromatogram of the reaction mixture obtained by iodination of 9-*N*-ASA-Neu5Ac2en

The eluate was monitored at 240 nm. A nearly identical elution profile was obtained by monitoring at 210 nm. The components of the signals are the residual starting material (S) and the reaction products, compounds (1, 2 and 3).

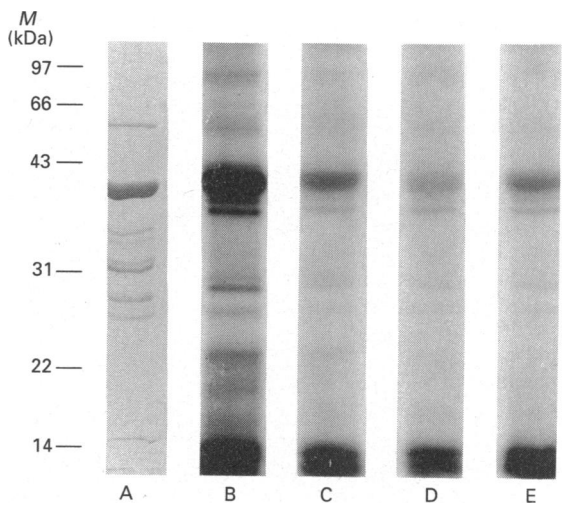


Fig. 3. Polyacrylamide gel of the recombinant *Salmonella* sialidase run under denaturing conditions as described in the Experimental section

Lane A, purified sialidase (2 μg) used in the photolabelling studies stained with Coomassie Blue; lane B, autoradiogram of the gel containing the protein (2 μg) photolysed with [*ASA*-^{125}I]9-*N*-ASA-Neu5Ac2en (see the Experimental section) without the addition of BME (15200 d.p.m. incorporated); lane C, protein photolysed with the probe and with 5 mM-BME included in the photolysis mixture (1906 d.p.m. incorporated); lane D, protein photolysed in the presence of both BME and Neu5Ac2en (2 mM) (724 d.p.m. incorporated); lane E, protein photolysed in the presence of BME and DOPC (2 mM) (2121 d.p.m. incorporated). After drying, the gel was exposed to X-ray film overnight. Molecular-mass (*M*) markers are indicated.

the active site are shielded from the actions of the thiol, and labelling of the peptide(s) buried within the substrate-binding region are unaffected. As expected, non-specific labelling of the sialidase was reduced about 7-fold by the presence of low concentrations (5 mM) of BME (Fig. 3, lane C).

When the photolysis was carried out in the presence of the

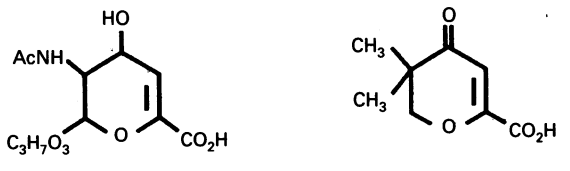


Fig. 4. Structures of the compounds used as photoprotective reagents during photolysis of the *Salmonella* sialidase

A, Neu5Ac2en; B, DOPC.

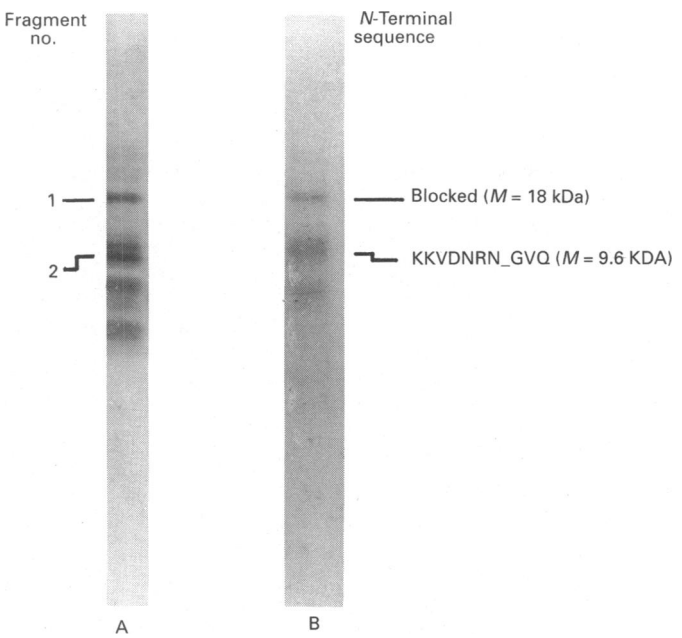


Fig. 5. PVDF membrane containing CNBr-derived peptide fragments

See the Experimental section for development of the acrylamide gel and transfer to the membrane. The protein was reduced and carboxymethylated before CNBr cleavage. Lane A, peptide fragments stained with Coomassie Blue; lane B, autoradiogram of the peptide fragments derived from the protein photolysed with the labelled probe. As indicated, the labelled fragments were subjected to *N*-terminal sequencing for identification. Shown is the sequence as obtained from the analyser. Size estimations of fragments 1 and 2 were made with molecular-mass (*M*) markers (not shown).

sialidase inhibitor Neu5Ac2en and BME, the labelling was substantially lowered, although it was not down to background levels (Fig. 3, lane D). We attribute the residual labelling of the protein under these conditions to the incomplete elimination of non-specific labelling at this concentration of BME. It was not practical to increase the concentration of the thiol above 5 mM as this completely prevented labelling of the protein. Neu5Ac2en competes with the photoprobe for the enzyme active site, displacing it from the protein and decreasing the labelling of the peptides located specifically in this region.

As a further measure of the specificity of labelling, photolysis experiments were carried out in the presence of DOPC and BME. Levels of the radiolabelled probe incorporated into the protein were nearly identical with that found when only BME was present (Fig. 3, lane E). Although DOPC contains structural and chemical features similar to those of Neu5Ac2en (Fig. 4), it is not an enzyme inhibitor, does not bind at the active site and does not prevent labelling by the photoprobe. These results are important, because they demonstrate that the effects of both Neu5Ac2en and the probe cannot be explained on the basis of simple ionic or hydrophobic interactions of these molecules with the protein attributable to the presence of the carboxylate group or the unsaturated pyranose ring respectively. Also, only background labelling was observed when the samples were maintained in the dark (results not shown). Identical photolabelling results were obtained using a sialidase preparation purified to near homogeneity (Hoyer *et al.*, 1991).

These combined results are consistent with the notion that the 9-*N*-ASA-Neu5Ac2en molecule binds to the enzyme in a highly specific manner, orienting itself with the pyranose-ring portion of the molecule in the substrate-binding pocket while positioning the aryl nitrene moiety away from the catalytic amino acid residues, near the protein–water interface, where it is only partially protected from thiol quenching. Such an orientation does not seem unreasonable, since the glycerol side chain of the sialic acid molecule, although it may participate in substrate binding, is not directly involved in the catalytic process nor is it required for bond hydrolysis.

Identification of the peptide comprising the substrate-binding domain

Treatment of the highly purified photolabelled sialidase with CNBr generated several peptide fragments when analysed by SDS/PAGE and protein-stained (Fig. 5, lane A). Two of the fragments were radioactively labelled at a level similar to that of the intact protein. Neither of these peptides were labelled when Neu5Ac2en was included in the photolysis mixture, although background labelling in other fragments was still observed. In order to identify the labelled peptides, *N*-terminal sequence analysis was carried out as indicated (Fig. 5, lane B). Fragment 1 (18 kDa) gave no signal in the sequencer, indicating that the *N*-terminus was blocked. Fragment 2, which was also radioactively labelled, was identified as a 9.6 kDa fragment comprising the *C*-terminal third of the protein. Fortuitously this is expected to be the largest CNBr-derived fragment on the bases of the methionine content and the predicted amino acid sequence of the protein (Hoyer *et al.*, 1992). Thus it seems reasonable to assume that peptides larger than 9.6 kDa must be due to the incomplete cleavage of the protein. A similar fragmentation pattern was obtained in duplicate experiments and under various conditions. Considering its molecular mass of 18 kDa and in the fact that fragment 1 is also radioactively labelled, we conclude that it is a composite of the 9.6 kDa peptide and the adjacent preceding 8.3 kDa peptide in the protein (Fig. 6). Other CNBr-derived peptides were labelled at much lower levels, which is consistent with the expectation that some non-specific labelling of the intact protein would occur.

Secondary-structural predictions

The secondary-structural organization of the 9.6 kDa peptide (Fig. 7*c*) was compared with a portion of the active site of the influenza-A-virus sialidase (Colman *et al.*, 1983) (Fig. 7*a*) and an analogous section of the Sendai-HN-virus sialidase–haemagglutinin (Middleton *et al.*, 1990) (Fig. 7*b*). The Garnier–Robson algorithm was used to predict those segments within these regions with the highest probability of β-sheet structure (Garnier *et al.*, 1978). As indicated, the predicted location of β-sheets for the influenza-virus sialidase was in good agreement with those determined from the crystal structure of the protein.

Molecular modelling of 9-*N*-ASA-Neu5Ac2en

Three conformational models of the photoaffinity probe were developed, based on: (i) the solution conformation of sialic acid as determined from magnetic-resonance studies (Czarniecki & Thornton, 1976), (ii) the crystal structure co-ordinates of

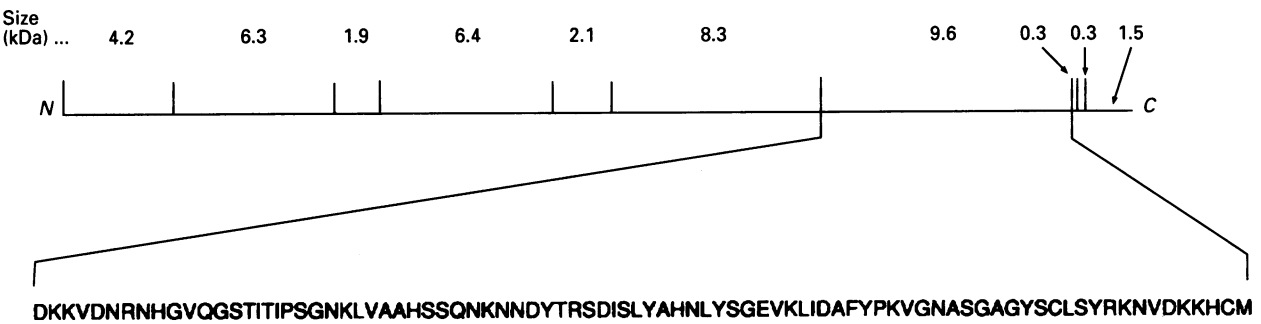


Fig. 6. Schematic representation of *Salmonella* sialidase showing the positions of methionine residues (vertical lines) and the predicted size of the CNBr-derived fragments

The expanded region shows the amino acid sequence of the 9.6 kDa peptide labelled with the photoprobe.

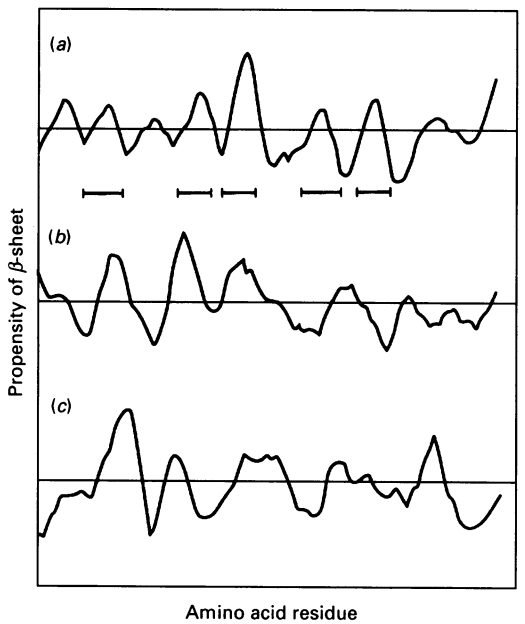


Fig. 7. Regions of β-sheets, as predicted by the Garnier–Robson algorithm (Garnier *et al.*, 1978) within the active site of influenza-A-virus sialidase (residues 210–333) (*a*), an analogous portion of the Sendai-HN-virus sialidase–haemagglutinin (residues 233–310) (*b*) and the 9.6 kDa peptide of *S. typhimurium* sialidase (residues 270–356) (*c*)

Bars show the β-sheet locations within the influenza-virus enzyme as determined from the crystal structure (Colman *et al.*, 1983).

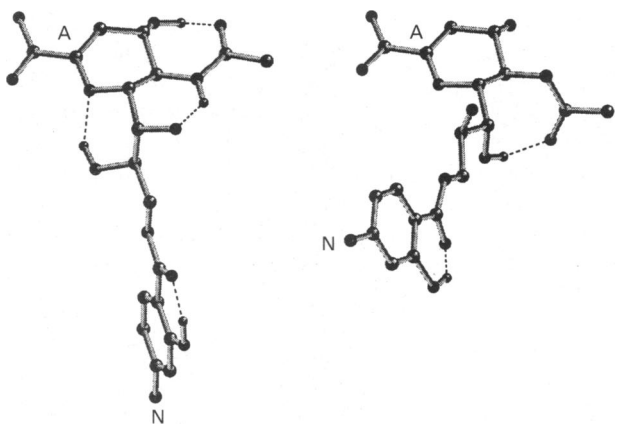


Fig. 8. Computer-generated molecular models of 9-*N*-ASA-Neu5Ac2en according to two of the optimal (minimal energy) conformations

The model of the *left* is based on the solution conformation of sialic acid determined from magnetic-resonance studies (Czarniecki & Thornton, 1976). The model on the *right* is based on the crystal structure of Neu5Ac2en (Furuhata *et al.*, 1988). Internal hydrogen bonds are shown as broken lines; only hydrogen atoms impliated in internal hydrogen-bonding are illustrated. Shown are the inclusion of the 2,3 double bond into the pyranose ring and the attachment of the ASA moiety at C-9 as its nitrene form. A, anomeric carbon assumed to be very near the catalytic amino acid residues of the protein; *N*, nitrene generated from the aryl azide. The distance between these two centres in the ^{1}H n.m.r.-derived model [1.24 nm (12.4 Å)] is essentially equivalent to the distance of 1.23 nm (12.3 Å) from the crystal structure.

Neu5Ac2en (Furuhata *et al.*, 1988) and (iii) a general examination of all feasible orientations of the probe.

Remarkably, in the conformations derived from either crystallographic or solution-phase magnetic-resonance studies, the distance between the anomeric centre at C-2, which is presumably situated very near the catalytic residues, and the aryl nitrene, was nearly identical at about 1.2 nm (12 Å) (Fig. 8). The systematic search procedure indicated the possible distance between these two groups, considering all energetically accessible conformations of the linker group, ranged from 0.9 to 1.3 nm (9 to 13 Å), with over 90% of the conformations spanning range 1.05–1.25 nm (10.5–12.5 Å).

Conclusion

We have carried out photoaffinity labelling studies of the recombinant sialidase of *S. typhimurium* in order to characterize the amino acid residues lining the active-site region of the protein. The portion of the sialidase which was identified by this technique comprises the *C*-terminal third of the molecule. Although little information is available about the tertiary structure or mechanism of action of bacterial sialidases, a comparison with the well-defined active site of the influenza-A-virus enzyme may provide insight into the structural organization of the bacterial protein.

The amino acids forming the active-site cavity of the influenza-A-virus sialidase are included in a region of the protein which spans 200 amino acids out of a total of 469 residues. Typically, for several influenza-virus sialidases, this domain begins at about 100 residues from the *N*-terminus and is made up of alternating segments of β-sheets connected with loops (Colman *et al.*, 1983; Blumberg *et al.*, 1985; Jorgensen *et al.*, 1987). This β-sheet–loop structural arrangement at the active site has been confirmed by crystallographic analysis (Colman *et al.*, 1983). Jorgensen *et al.*

(1987) have observed that the β-sheet regions of the influenza protein can be accurately predicted from the amino acid sequence by the Garnier–Robson algorithm. Using this approach they have identified similar domains in other viral sialidases and haemagglutinin proteins which are presumed to form the active-site clefts. Secondary-structure predictions of the *S. typhimurium* peptide labelled with the probe indicate that this portion of the bacterial protein is organized in a similar manner. At least four alternating regions with a high probability of β-sheet were identified. Structurally, the 9.6 kDa peptide bears a general similarity to active site of influenza-A-virus and other viral sialidases, with an apparently remarkable similarity to the enzyme of Sendai HN virus. This suggests that the 9.6 kDa peptide comprises a portion of the active-site cleft of the bacterial enzyme.

Recent analysis of the influenza-A-virus sialidase with site-specific mutagenesis techniques have led Lentz *et al.* (1987) to speculate that histidine-274 and glutamic acid-276 and -277 of the influenza-A-virus protein are directly involved in catalysis. These residues are not positioned near the *C*-terminus, but they are centrally located with the protein sequence. The apparent difference in active-site location between the two enzymes may simply be due to the fact that the influenza-virus enzyme is 93 amino acids longer than the *S. typhimurium* protein. However, other amino acids (arginine-371 and tyrosine-406) which are located at the terminal third of the influenza molecule were also found to be critical for maintaining catalytic activity and for proper protein folding. This demonstrates the importance of the *C*-terminal portion of the influenza-A-virus enzyme in contributing an integral part to the sialyl-conjugate-cleaving process. The analogous region of the bacterial sialidase may have a similar functional significance, although the specific mechanism of catalysis may be different. Adjacent glutamic acid residues, such as those at the putative catalytic site of the influenza-A-virus sialidase, were not present in the peptide labelled with the photoprobe or elsewhere in the *S. typhimurium* sequence. Other viral sialidases, such as that of the parainfluenza type 4A virus, may also employ a catalytic process different from that of the influenza-A-virus enzyme. In this case aspartic acid-300 and tyrosine-301 have been shown, by mutational analysis and by conservation between several paramyxoviral enzymes, to be vital to catalysis (Bando *et al.*, 1990).

Our data do not verify that the catalytic amino acids or the residues directly involved in binding are located on the peptide identified with the probe, although this possibility cannot be ruled out, since the protection experiments with Neu5Ac2en and DOPC are consistent with this tenet. On the basis of molecular-modelling studies, regardless of the conformation assumed by the substrate analogue when bound to the enzyme, its covalent structure imposes a 1.05–1.25 nm (10.5–12.5 Å) limit on the distance between the anomeric centre at C-2 and the aryl nitrene nitrogen. Assuming that the residues involved in catalysis are located on the same peptide contiguous to the site labelled by the probe, and, if the secondary structure of this region in the intact protein has β-sheet characteristics, then the catalytic residues would be about 3–4 amino acids distant from those labelled with the probe. However, it should be noted that the probe, as its reactive aryl nitrene form, may diffuse away from the catalytic residues before it attaches covalently to the peptide. Thus, although the conformational constraints of the probe limit the distance between its reactive centre and the site of bond cleavage at the anomeric carbon, the actual site of attachment to the protein may be farther from the catalytic residues than that expected by this analysis. In addition, such a migration may also result in multiple sites of attachment on the peptide. In the studies described here, the low efficiency of labelling precluded further characterization of the labelled peptide, and the reaction stoichiometry of the probe with the peptide could not easily be ascertained. Because of these limitations, we conclude that photolabelling of this protein with 9-ASA-Neu5Ac2en will not be useful for revealing functional roles of individual residues at the active site. However, the approach taken here can be employed for identifying a generalized region within the active site of the protein, which is likely to be near the residues involved in substrate cleavage, recognition and binding.

A more precise understanding of the bacterial enzymes and their active sites will be possible when the crystal structure of these proteins becomes available. Nevertheless, the results we have obtained will be valuable in designing mutagenesis studies of this *C*-terminal segment of the protein to further examine its role in catalysis and substrate recognition. Similar photolabelling studies with other bacterial sialidases may reveal tertiary structural similarities between the active sites of these enzymes that might not otherwise be apparent from inspection of the primary structure. It will be of great interest to determine whether a *C*-terminal peptide of other microbial sialidases is labelled with this affinity reagent as well. Toward this end we have recently shown active-site-specific photolabelling of the sialidases from *V. cholerae* and *C. perfringens* (Warner & Vimr, 1991).

Finally, while the studies here were underway, photolysis of the *Clostridium* sialidase, using an ASA derivative of Neu5Ac2en similar to the one we have developed, appeared in a preliminary communication (Van Der Horst *et al.*, 1990). However, the method of synthesis of the key intermediate, 9-amino-Neu5Ac2en, used in preparation of the probe, and amino acid sequence of the labelled peptide, were not described. For these reasons it is not possible to make a direct comparison with the results obtained with the *S. typhimurium* enzyme.

We thank Dr. Mike Spellman for his helpful comments regarding this work. This work was supported by Genentech and National Institutes of Health grants NS 22323 and Al 23039 to T.G.W. and E.R.V. respectively.

REFERENCES

Air, G. M. & Laver, G. W. (1989) Proteins: Struct. Funct. Genet. **6**, 341–356

Allen, F. H., Bellard, S., Brice, M. D., Cartwright, B. A., Doubleday, A., Higgs, H., Hummelink, T., Hummerlink-Peters, B. G. & Kennard, O. (1979) Acta Crystallogr. Sect. B: Struct. Sci. **B35**, 2331–2339

Bando, H., Kondo, K., Kawano, M., Komada, H., Tsurudome, M., Nishio, M. & Ito, Y. (1990) Virology **175**, 307–312

Bayley, H. & Staros, J. V. (1984) in Azides and Nitrenes (Scriven, E. F. V., ed.), pp. 433–490, Academic Press, New York

Bayley, H., Standring, D. N. & Knowels, J. R. (1978) Tetrahedron Lett. **39**, 3633–3634

Blumberg, B., Giorgi, C., Roux, L., Raju, R., Dowling, P., Chollet, A. & Kolakosfsky (1985) Cell (Cambridge, Mass.) **41**, 269–278

Clark, M., Cramer, R. D. & Opdenboshch, N. V. J. (1989) J. Comput. Chem. **10**, 982–1012

Colman, P. M., Varghese, J. N. & Laver, W. G. (1983) Nature (London) **303**, 41–44

Corfield, A. P. & Schauer, R. (1982) in Sialic Acids (Shauer, R., ed.), pp. 195–261, Springer-Verlag, New York

Czarniecki, M. F. & Thornton, E. R. (1976) J. Am. Chem. Soc. **98**, 1023–1025

Furuhata, K., Satao, S., Goto, M., Tatayanagi, H. & Ogura, H. (1988) Chem. Pharm. Bull. **36**, 1872–1876

Garnier, J., Osguthorpe, D. F. & Robson, B. (1978) J. Mol. Biol. **120**, 97–120

Hoyer, L., Roggentin, P., Schauer, R. & Vimr, E. R. (1991) J. Biochem. (Tokyo) **110**, 462–467

Hoyer, L. L., Hamilton, A. C., Steenbergen, S. M. & Vimr, E. R. (1992) Mol. Microbiol. **6**, 873–884

Ji, I. & Ji, T. H. (1982) Anal. Biochem. **121**, 286–289
Ji, I., Shin, J. & Ji, T. H. (1985) Anal. Biochem. **151**, 348–349
Jorgensen, E. D., Collins, P. L. & Lomedico, P. T. (1987) Virology **156**, 12–24
Laemmli, U. K. (1970) Nature (London) **227**, 680–685.
Lentz, M. R., Webster, R. G. & Air, G. M. (1987) Biochemistry **26**, 5351–5358
Lowden, J. A. & O'Brien, J. S. (1979) Am. J. Hum. Genet. **31**, 1–18
Matsudaria, P. (1987) J. Biol. Chem. **262**, 10035–10038
Meindl, P. & Tuppy, H. (1973) Monatsch. Chem. **194**, 402–414
Middleton, Y., Tashiro, M., Thai, T., Oh, J., Galen, J., Lawrisak, L., Vimr, E. R. & Schaeur, R. (1990) Virology **176**, 656–657
Roggentin, P., Rother, B., Kaper, J. B., Seymour, J., Pritzer, E., Klenk, H. D., Roh, R. & Seto, J. T. (1989) Glycoconjugate J. **6**, 349–353
Roll, D. E., Aguanno, J. J., Coffee, C. J. & Glew, R. H. (1978) J. Biol. Chem. **263**, 6992–6996
Russo, T. A., Thompson, J. S., Godoy, V. G. & Malamy, M. H. (1990) J. Bacteriol. **172**, 2594–2600
Sasada, Y., Takano, T. & Kakudo, M. (1964) Bull. Chem. Soc. Jpn. **37**, 940–946
Schagger, H. & Von Jagow, G. (1987) Anal. Biochem. **166**, 368–379
Svennerholm, L. (1963) J. Neurochem. **10**, 613–620
Tettamanti, G., Durand, P. & Di Donato, S. (eds.) (1981) Perspectives in Inherited Metabolic Diseases, Edi Ermes, Milan
Van Der Horst, G. T. J., Mancini, G. M. S., Brossmer, R., Rose, U. & Verheijen, F. W. (1990) J. Biol. Chem. **265**, 10801–10804
Warner, T. G. (1987) Biochem. Biophys. Res. Commun. **148**, 1323–1329
Warner, T. G. & Vimr, E. R. (1991) Glycoconjugate J. **8**, 273
Warner, T. G., Louie, A. & Potier, M. (1990) Biochem. Biophys. Res. Commun. **173**, 13–19
Warner, T. G., Louie, A., Potier, M. & Ribiero, A. (1991) Carbohydr. Res. **215**, 315–321

Received 28 November 1991/30 January 1992; accepted 5 February 1992