Carbohydrate Research, 232 (1992) 227–233
Elsevier Science Publishers B.V., Amsterdam

Investigation of the active site of *Escherichia coli* β-D-galactosidase by photoaffinity labelling

Cai-Steffen Kuhn [a], Jochen Lehmann [a], Günther Jung [b] and Stefan Stevanović [b]

[a] *Institut für Organische Chemie und Biochemie der Universität Freiburg i. Br., Albertstr. 21, D-7800 Freiburg i. Br. (FRG)*

[b] *Institut für Organische Chemie der Universität Tübingen, Auf der Morgenstelle 18, D-7400 Tübingen (FRG)*

(Received October 29th, 1991; accepted January 11th, 1992)

ABSTRACT

3,7-Anhydro-2-azi-1,2-dideoxy-D-*glycero*-L-*manno*-(8-^{3}H)octitol (**1a**) and 3-azibutyl 1-thio-β-D-(6-^{3}H)galactopyranoside (**2a**) were synthesised from the unlabelled compounds by reaction with galactose oxidase, then reduction with sodium borotritide. Whereas **1a** was an efficient photoaffinity reagent for the β-D-galactosidase from *E. coli*, **2a** was ineffective. Three ^{3}H-labelled peptides were isolated after digestion of the labelled enzyme with trypsin, one of which was an octapeptide (Trp 158 to Ser 165), which is remote from the segments detected as part of the active site of the enzyme.

INTRODUCTION

Since the primary structure of the β-D-galactosidase from *E. coli* was established [1,2], many attempts have been made to elucidate the shape and functionality of the active site [3–10].

Two amino acids, Met 502 [9] and Glu 461 [10], can be modified covalently by different affinity reagents. Glu 461 and Tyr 503 were shown by site-directed mutagenesis to be involved in the catalytic process [11,12]. Stereoselective proton capture, using D-galactal [13] and (*Z*)-3,7-anhydro-1,2-dideoxy-2-deuterio-D-*galacto*-2-octenitol [14] as prochiral proton acceptors, indicated the proton-donating groups Tyr 503 and Glu 461 to be positioned on opposite sides of the bound unsaturated substrates. Glu 461 probably protonates D-galactal from below and Tyr 503 the octenitol from above the pyranosyl ring. The binding site of the β-D-galactosidase can be cross-linked with a bireactant ligand [15], but no analysis of the modified peptides has been reported.

Correspondence to: Professor Dr. J. Lehmann, Institut für Organische Chemie und Biochemie der Universität Freiburg i. Br., Albertstr. 21, D-7800 Freiburg i. Br., Germany.

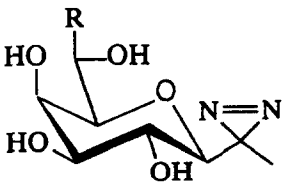
All of the probes used so far can react with the protein only if suitable functional groups are available at sufficiently close range. However, photoaffinity reagents that generate carbenes or nitrenes will chemically modify any nearby group.

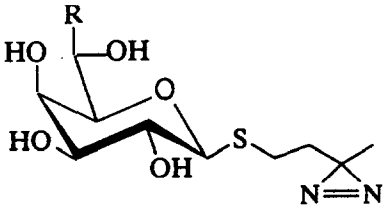
We now report on the photoaffinity labelling of the β-D-galactosidase from *E. coli*.

RESULTS AND DISCUSSION

Since the β-D-galactosidase of *E. coli* does not tolerate the introduction of photolabile groups into the glyconic part of an enzyme-resistant ligand, such a group was attached by a C–C linkage to C-1 of a β-D-galactopyranosyl ring. 3,7-Anhydro-2-azi-1,2-dideoxy-D-*glycero*-L-*manno*-octitol [16] (**1**), which is an efficient photoaffinity reagent for β-D-galactosidase [17], was prepared 8-^{3}H-labelled (**1a**). As a complementary reagent that can react further away from the glycon-binding site, 6-^{3}H-labelled (**2a**) 3-azibutyl-1-thio-β-D-galactopyranoside [18] (**2**) was prepared. Both **1** and **2** are excellent competitive inhibitors of the hydrolysis of *o*-nitrophenyl β-D-galactopyranoside by β-D-galactosidase (K_i 750 and 71 μM, respectively).

When the β-D-galactosidase was incubated with **1a** and irradiated at 350 nm, 20.4% of the radioactivity was incorporated into the enzyme, which corresponds with the percentage of deactivation effected by **1**. However, incubation with **2a** introduced < 1.4% of the radioactivity into the enzyme as determined by SDS-PAGE. This negative result does not reflect the fact that the photolabile diazirine in **2** is attached to a sulfur-linked aglycon. Analogues of **2** have been used to label β-hexosaminidase [19], alpha-amylase [20], and a maltose-binding protein [21]. It is possible that the reactive carbene generated when **2** is irradiated is too remote from, and does not react with, the enzyme protein. This finding could mean that the so-called aglycon-binding site of β-D-galactosidase is "open" to the outside medium, which agrees with the low aglyconic specifity of the enzyme.





1 R = H
1a R = ^{3}H

2 R = H
2a R = ^{3}H

When β-D-galactosidase (5 mg) was incubated with **1a** (69 mCi, 7.5 mM) and the mixture was irradiated with UV light of 350 nm, 20.24 μCi of the radioactivity became attached covalently to the enzyme protein and < 1.3 μCi was incorporated when the procedure was carried out in the presence of isopropyl 1-thio-β-D-galac-

TABLE I

^{3}H-labelled peptides derived from β-D-galactosidase photoaffinity labelled with **1a**

Peptide	Radioactivity incorporated (%)	Retention time in HPLC [a] (min)
A	11	41.8
B	19	42.4
C	70	21.5

[a] See Experimental.

topyranoside. Digestion of the radiolabelled and carboxymethylated protein with trypsin and reversed-phase HPLC of the resulting mixture of peptides revealed various proportions of three labelled peptides (A–C, Table I).

Since the carbene carbon in the photolysed enzyme-bound ligand **1a** was likely to be close to Tyr 503, a tryptic peptide, Ser 477 to Arg 505, was expected on the basis of the known primary structure. The first 11 cycles in the sequencing of peptide A (Fig. 1) gave the partial sequence Ser 477 to Glu 487, including Arg 482 adjacent to Pro 483, the peptide bond of which is not susceptible to trypsin. It is probable that the alkylation occurred towards the C-terminus of that peptide.

For peptide B, of which only 0.1 nmol of impure material could be isolated, only the sequence Ser–Val–Asp–Pro was determined unequivocally. This finding indi-

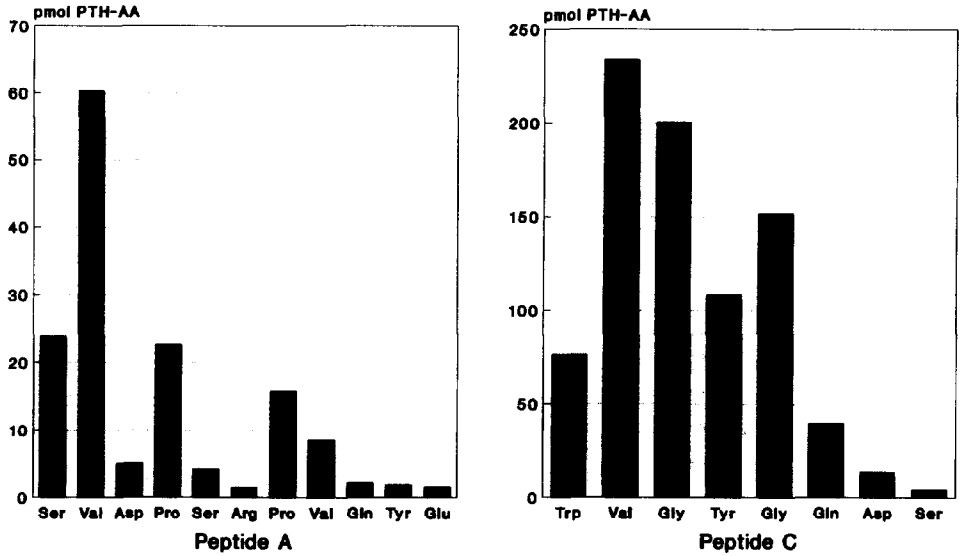


Fig. 1. Automated Edman degradation of peptides A and C: short peptides absorbed on to the glass-fibre filter were washed out more easily than proteins, and amino acids bearing functional groups in the side chain yielded lower recoveries than those with inert side chains. Peptide A contains the highly acid-labile Asp–Pro peptide bond.

449 Asn–His–Pro–Ser–Val–Ile–Ile–Trp–
Ser–Leu–Gly–Asn–*Glu*–Ser–Gly–His–Gly–
Ala–Asn–His–Asp–Ala–Leu–Tyr–Arg 473

Fig. 2. Tryptic peptide from β-D-galactosidase with the *Glu* 461-residue, labelled by conduritol C *cis*-epoxide [9], was not labelled by **1a**.

cates peptides A and B to differ only in the position of the covalently attached ligand.

Considering the structure of the reactive carbene with its β-aglycon, it is understandable that the tryptic peptide (Asn 449 to Arg 473), including Glu 461 which was labelled by conduritol C *cis*-epoxide [9], was not among the labelled segments (Fig. 2).

Peptide C contained 70% of the covalently attached radioactivity and its sequence was determined as Trp–Val–Gly–Tyr–Gly–Gln–Asp–Ser (Fig. 1). This sequence (Trp 158 to Ser 165) is remote from the segments detected on the primary structure of the protein as part of the active site of the β-D-galactosidase. It is possible that this segment is folded back on the active site close to the segment 477 to 505, sandwiches the O-1β, and thereby contributes to the β-specificity of the enzyme. It is striking that the sequence of peptide C is, to a large degree, conserved in the primary structures of five other bacterial β-D-galactosidases (Fig. 3).

EXPERIMENTAL

General methods.—All reactions were monitored by TLC on Silica Gel 60 F_{254} (Merck). Flash-column chromatography [22] was carried out on Silica Gel 32-63, 60A (ICN). HPLC was performed with an LKB 2152 controller, two LKB 2150 pumps, a Rheodyne 7126 injector, an LKB 2151 variable wavelength monitor, a

	157	158	159	160	161	162	163	164	165	166	167
EC	Arg	***Trp***	***Val***	***Gly***	***Tyr***	***Gly***	***Gln***	***Asp***	***Ser***	Arg	Leu
KPN	Val	Trp	Val	Gly	Tyr	Ser	Gln	Asp	Ser	Arg	Leu
LBU	Glu	Phe	Val	Gly	Tyr	Gly	Glu	Asp	Ser	Phe	Thr
CLA	Glu	Phe	Val	Gly	Tyr	Ser	Glu	Asp	Trp	Phe	Thr
ST	Asn	Phe	Val	Gly	Tyr	Ser	Glu	Asp	Ser	Phe	Thr
EBG	Gln	Tyr	Val	Gly	Tyr	Ser	Lys	Gly	Ser	Arg	Leu

Fig. 3. ^{3}H-Labelled peptide C and homologies between the protein sequences of β-D-galactosidases from *Escherichia coli* (EC), *Klebsiella pneumoniae* (KPN), *Lactobacillus bulgaricus* (LBU), *Clostridium acetobutylicium* (CLA), *Streptococcus thermophilis* (ST), and evolved β-D-galactosidase (EBG).

Berthold LB 507 radioactivity monitor equipped with an analytic splitter (85/15), and a Z-1000/4 cell scavenged with Quickszintflow 302. Radioactive material was detected either by radioautography (Agfa-Gevaert Curix X-ray film) or with a Berthold Automatic TLC Linear Analyzer. Radioactive samples in solution were assayed with a Berthold BF-815 liquid scintillation counter. Photolabile compounds were irradiated with a Rayonet RPR 100 reactor equipped with 16 lamps (RPR 3500 A).

Discontinuous SDS-PAGE [23] was performed [24] in tubes (6 × 170 mm). The acrylamide concentration was 5% with a ratio of acrylamide:bis(acrylamide) of 97.3:2.7. The length of the gel was 100 mm. Samples were applied after heating for 1 min at 95°. The temperature during electrophoresis was maintained at 10° and the current at 1 mA/gel tube. After electrophoresis, the gel was (*a*) cut into 2-mm slices, each of which was submerged overnight in Biolute-S (0.5 mL, Zinsser), Quickszint-501 (4 mL, Zinsser) was then added, the mixture was kept for 2 h at 8°, and the radioactivity was determined; or (*b*) fixed with aq 12.5% trichloroacetic acid and stained with Serva Blue.

Enzymes.—Highly-purified β-D-galactosidase (β-D-galactoside galactohydrolase, EC 3.2.1.23) from *E. coli* was a gift from Dr. W. Littke (Freiburg) as a suspension of crystals (8 mg/mL) in M $(NH_4)_2SO_4$–0.86 M NaCl with a specific activity of 371 U/mg. Before use, the enzyme suspension was centrifuged, the precipitate was dissolved in buffer (see below), and the solution was dialysed against the buffer (3 × 500 mL) at 8° in tubing from Serva. The concentration of β-D-galactosidase was determined by the biuret procedure [25] with crystalline bovine serum albumin as the standard.

Galactose oxidase (D-galactose:oxygen 6-oxidoreductase, EC 1.1.3.9) from *Dactylium dendroides* (87 U/mg lyophilisate) was purchased from Sigma, catalase (hydrogen peroxide:hydrogen peroxide oxidoreductase, EC 1.11.1.6) from bovine pancreas from Boehringer Mannheim, and trypsin from bovine pancreas (TPCK-treated, 31 U/mg, research grade) from Serva.

3,7-Anhydro-2-azi-1,2-dideoxy-D-glycero-L-manno(*8-*^{3}*H*)*octitol* (**1a**).—To a solution of **1** (30 mg, 13.7 μmol) in sodium phosphate buffer (1.5 mL, pH 7.2, 10 mM) were added D-galactose oxidase (1 mg, 87 U/mg) and catalase (35 μL, 1300 U/mL) at 20°. Formation of the 8-aldehydo compound was indicated by TLC (R_F 0.23, 7:2:1 EtOAc–MeOH–H_2O). After 8 h, the mixture was passed through a column (0.5 × 3 cm) of silica gel by elution with H_2O. The eluate was concentrated under diminished pressure to 100 μL and added to a fresh solution of NaB^3H_4 (100 mCi, 62.9 Ci/mmol) in M NaOH (50 μL). After 12 h, the mixture was neutralised (acetic acid) and concentrated. Flash-column (0.5 × 10 cm) chromatography (17:2:1 EtOAc–MeOH–H_2O) of the residue gave 85 mCi of **1a**, 9.2 Ci/mmol determined by co-crystallisation of **1a** (0.1 mCi) with **1** (150 mg) from EtOH.

*3-Azibutyl 1-thio-β-*D*-(6-*^{3}*H*)*galactopyranoside* (**2a**).—Compound **2** (35 mg, 126 μmol) was treated with galactose oxidase (1.15 mg) and catalase (50 μL) as

described above. After 10 h, TLC (17:2:1 EtOAc–MeOH–H_2O) showed the absence of **2** (R_F 0.3), and the mixture was concentrated to 100 μL by lyophilisation and added to a solution of NaB^3H_4 (100 mCi, 8.7 Ci/mmol) in M NaOH (100 μL). After 12 h, the mixture was neutralised with acetic acid and concentrated to dryness. Flash-column chromatography (R_F 0.27, 5:1 EtOAc–MeOH) of the residue gave **2a** (32 mCi, 2.17 Ci/mmol). The radioactive product co-chromatographed on 2D TLC (7:2:1 EtOAc–MeOH–H_2O) with **2**.

Labelling of β-D-galactosidase.—(a) With **1a**. A solution of β-D-galactosidase (5 mg) in 50 mM sodium phosphate buffer (1 mL, pH 6.8, mM $MgCl_2$) and **1a** (69 mCi, 0.75 μmol) was deoxygenated with a stream of N_2 for 2 min, then irradiated at 350 nm for 20 min, and dialysed against buffer (3 × 500 mL) and H_2O (2 × 500 mL) at 8° until no more radioactivity was released into the diffusate. The radioactivity of the protein solution, determined by liquid scintillation counting of an aliquot with Quickszint 1 (Zinsser), corresponded to 2.19 nmol of **1a** bound to 10.74 nmol of the tetrameric enzyme.

(b) With **2a**. The β-D-galactosidase (0.5 mg, 1.07 nmol) was labelled with **2a** (1.55 mCi, 2.17 Ci/mmol) as described in (*a*). SDS-PAGE of 50 μg of the dialysed protein revealed that only 1.36% (3.2 nCi) of the tetrameric enzyme had been labelled.

Reduction and carboxymethylation of ^{3}H-labelled β-D-galactosidase.—A solution of freeze-dried, ^{3}H-labelled β-D-galactosidase (5 mg) in 0.5 M Tris-HCl buffer (1.5 mL, pH 7.5; 6 M guanidinium HCl, 5 mM EDTA, and 7 mM dithioerythritol) was kept for 5 h at room temperature. Sodium iodoacetate (5 mg, 24 μmol) was added in the dark, followed, after 20 min, by 2-mercaptoethanol (70 μL). The protein was precipitated by dialysis against H_2O (2 × 400 mL) and collected by centrifugation.

Treatment of ^{3}H-labelled β-D-galactosidase with trypsin.—A suspension of the reduced and carboxymethylated enzyme in 0.2 M NH_4HCO_3 buffer (1.5 mL) was treated with trypsin (2%) for 3 h at 37°. A similar amount of trypsin was then added and the digestion was continued for 2 h and monitored by reversed-phase HPLC on a column (4.6 × 250 mm) of C-18 (VYDAC 218 TP) with MeCN–H_2O containing 5 mM trifluoroacetic acid at 1 mL/min. The MeCN gradient was as follows: 0–10 min, 10%; 10–80 min, 10–40%; 80–85 min, 40–100%. The peptides were monitored at 220 nm and the radioactivity was detected simultaneously.

The ^{3}H-labelled peptides were isolated from 100-μg samples diluted with 100 μL of 5 mM trifluoroacetic acid by HPLC as described above but with a flow rate of 0.8 mL/min. The ^{3}H-labelled peptides were freeze-dried and stored at −20°.

Sequence analysis of the ^{3}H-labelled peptides.—Automated Edman degradation was carried out in a pulsed-liquid protein sequencer 477A equipped with an on-line phenylthiohydantoin(PTH)-amino acid analyser 120 A (Applied Biosystems, Weitersstadt, FRG). All reagents and solvents used were from Applied Biosystems. A solution of each freeze-dried peptide in aq 25% trifluoroacetic acid (60 mL) was applied to a glass-fibre filter activated with trifluoroacetic acid. Sequencing was performed with the standard programmes Begin-1 and Normal-1,

which included treatment with aq 25% trifluoroacetic acid for 20 min at 64°. Under these conditions, ester groups are expected to be hydrolysed.

ACKNOWLEDGMENTS

We thank the Deutschen Forschungsgemeinschaft for financial support (of JL), Dr. W. Littke (Freiburg) for a generous gift of pure β-D-galactosidase, and Professor Dr. R.E. Huber (Calgary) for valuable advice.

REFERENCES

1 A.V. Fowler and I. Zabin, *J. Biol. Chem.*, 253 (1978) 5521–5525.
2 A. Kalnins, K. Otto, U. Ruther, and B. Müller-Hill, *EMBO*, 2 (1983) 593–597.
3 M. Brockhaus and J. Lehmann, *FEBS Lett.*, 62 (1976) 154–156.
4 M.L. Sinnott and P.J. Smith, *J. Chem. Soc., Chem. Commun.*, (1976) 223–224.
5 M.L. Sinnott and P.J. Smith, *Biochem. J.*, 175 (1978) 525–538.
6 M. Brockhaus and J. Lehmann, *Carbohydr. Res.*, 63 (1978) 301–306.
7 G. Kurz, J. Lehmann, and E. Vorberg, *Carbohydr. Res.*, 93 (1981) C14–C20.
8 G. Kurz, J. Lehmann, and E. Vorberg, *Carbohydr. Res.*, 110 (1982) C21–C24.
9 F. Naider, Z. Bohak, and J. Yarif, *Biochemistry*, 11 (1972) 3202–3208; A.V. Fowler, I. Zabin, M.L. Sinnott, and P.J. Smith, *J. Biol. Chem.*, 253 (1978) 5283–5285.
10 M. Herrchen and G. Legler, *Eur. J. Biochem.*, 138 (1984) 527–531.
11 D.E. Bader, M. Ring, and R.E. Huber, *Eur. J. Biochem.*, 153 (1988) 301–306.
12 M. Ring, D.E. Bader, and R.E. Huber, *Biochem. Biophys. Res. Commun.*, 152 (1988) 1050–1055.
13 J. Lehmann and B. Zieger, *Carbohydr. Res.*, 58 (1977) 73–78.
14 J. Lehmann and P. Schlesselmann, *Carbohydr. Res.*, 113 (1983) 93–98.
15 R.E. Huber, J. Lehmann, and L. Ziser, *Carbohydr. Res.*, 214 (1991) 35–41.
16 G. Kurz, J. Lehmann, and R. Thieme, *Carbohydr. Res.*, 136 (1985) 125–133.
17 C.-S. Kuhn and J. Lehmann, *Carbohydr. Res.*, 160 (1987) C6–C8.
18 C.-S. Kuhn, J. Lehmann, and J. Steck, *Tetrahedron*, 46 (1990) 3129–3134.
19 C.-S. Kuhn, J. Lehmann, and K. Sandhoff, Bioconjug. Chem., in press.
20 M. Blanc-Muesser, H. Driguez, J. Lehmann, and J. Steck, *Carbohydr. Res.*, 223 (1992) 129–136.
21 J. Lehmann, E. Schiltz, and J. Steck, *Carbohydr. Res.*, CAR 11126.
22 M.C. Still, M. Kahn, and A. Mitra, *J. Org. Chem.*, 43 (1978) 2923–2926.
23 U.K. Laemmli, *Nature (London)*, 227 (1970) 680–685.
24 P.H. O'Farrell, *J. Biol. Chem.*, 250 (1975) 4007–4021.
25 G. Beisenherz, H.J. Boltze, T. Bücher, R. Czok, K.H. Garbade, E. Meyer-Arendt, and G. Pfleiderer, *Z. Naturforsch., Teil B*, (1953) 555–557.