Chapter 7

Identification of formaldehyde-induced modifications in proteins: reactions with diphtheria toxin

Bernard Metz, Gideon F. A. Kersten, Ad de Jong, Hugo Meiring, Jan ten Hove, Wim E. Hennink, Daan J. A. Crommelin, Wim Jiskoot

Abstract

Diphtheria toxoid, the principle component of diphtheria vaccines, is prepared by inactivating diphtheria toxin with formaldehyde and glycine. The treatment introduces intramolecular cross-links and intermolecular formaldehyde/glycine adducts in diphtheria toxin. The purpose of the present study was to elucidate the nature and location of formaldehyde-induced modifications at two functional sites of diphtheria toxin: the NAD^+-binding cavity and the receptor-binding site. Therefore, diphtheria toxin was chemically modified using five different reactions: (*1*) dimethylation by formaldehyde and $NaCNBH_3$, (*2*) acetylation by acetic acid N-hydroxy succinimide ester, (*3*) formaldehyde treatment, (*4*) the standard detoxification by formaldehyde and glycine, and (*5*) dimethylation followed by formaldehyde and glycine treatment. The modifications in these experimental diphtheria toxoids were studied by SDS-PAGE, primary amino group determinations (TNBS assay), and/or by liquid chromatography-mass spectrometry (LC-MS) analysis of enzymatically cleaved toxoids. LC-MS analysis confirmed that all but one lysine residue in diphtheria toxin were dimethylated by the reaction with formaldehyde and $NaCNBH_3$ (reaction *1*). According to the TNBS assay, four amino groups per toxin molecule were on average not dimethylated. The reaction of formaldehyde with lysine residues is the first step in the formation of a stable intramolecular methylene bridge. Thus, the formation of intramolecular cross-links only depends on the close proximity of a second reactive residue. Indeed, LC-MS analysis revealed nine intramolecular cross-links between lysine and a nearby reactive residue in formaldehyde-treated toxin (reaction *3*). Two masses were ascribed to peptide fragments with an intramolecular cross-link originating from the NAD^+-binding cavity, and two masses for the receptor-binding site. It was assumed that formaldehyde-glycine adducts are connected mainly to the reactive residues at the surface of the protein. As a simplified model, acetylation of lysine residues of diphtheria toxin was investigated (reaction *2*). LC-MS analysis demonstrated that buried residues are less reactive to the acetic acid N-hydroxy succinimide ester than the residues at the surface of the toxin. This result indicates that formaldehyde-glycine adducts will be only be attached to the more accessible residues. The TNBS assay also demonstrated partial modification of lysine residues (36% unmodified). Finally, the presence of formaldehyde-glycine attachments was studied at the NAD^+-binding cavity and the receptor-binding site (reaction *5*). Five peptide fragments with formaldehyde-glycine modifications were observed from the NAD^+-binding cavity, and three fragments from the receptor-binding site. In conclusion, the functional sites of diphtheria toxin are affected by the formaldehyde and glycine treatment.

Introduction

Diphtheria toxoid-containing vaccines are included in many national immunization programs. Vaccination has drastically reduced the incidence (and severeness) of diphtheria. Diphtheria vaccines are prepared from diphtheria toxin, which causes the clinical manifestations of the disease. In the 1920, Ramon and Glenny developed independently a successful method for the inactivation of diphtheria toxin, i.e. treatment with formaldehyde (1, 2). The current production process of diphtheria vaccines is essentially based on their method.

In general, the reactions of formaldehyde with amino acid residues are rather well understood. Extensive model studies have revealed the reactivity of amino acid residues and the nature of the modifications (3-8). The reaction of formaldehyde with protein starts with the formation of reversible methylol adducts on amino groups. The methylol groups partially dehydrate, yielding labile Schiff-bases, which can form cross-links with several amino acid residues. The formaldehyde treatment has a great effect on the toxicity, antigenicity and immunogenicity of diphtheria toxin (9, 10). Formaldehyde converts diphtheria toxin into a non-toxic product, called diphtheria toxoid (9), probably by destroying active sites in the molecule, e.g. the NAD^+-binding cavity and the receptor-binding site. However, the exact location and the nature of the modifications at the functional sites are unknown. Furthermore, detoxification causes complete or partial loss of epitopes, as demonstrated with anti-diphtheria toxin monoclonal antibodies (10). In spite of this, the toxoid remains very immunogenic and induces a protecting immune response by the generation of toxin-neutralizing antibodies.

The purpose of this study was to elucidate the chemical modifications in diphtheria toxin as a result of the detoxification. Because mapping of all structural modifications in the entire molecule is very laborious, attention was focused on two functional areas, i.e., the NAD^+-binding cavity and the receptor-binding site. The NAD^+-binding groove is located in the catalytic domain of diphtheria toxin that transfers the ADP-ribose moiety of NAD^+ to elongation factor-2 (EF-2) (11). The modification of EF-2 irreversibly inhibits the protein synthesis in the host cell leading to cell death. Three short peptide sequences in diphtheria toxin form the NAD^+-binding cavity, a loop from the residues 17 – 23, a β-strand followed by an α-helix from residues 50 – 67, and a β-strand from residues 147 – 150 (Figure 1A). The participation of amino acid residues His 21, Tyr 54, Tyr 65 and Glu 148 for the binding of NAD^+ has been described in the literature (12, 13). Another important area in the toxin molecule is the receptor-binding site, which is formed by a loop of amino acid residues between 511 – 530. This part of the receptor domain binds to the heparin-binding epidermal growth factor-like precursor (14, 15). The residues Tyr 514, Lys 516 Val 523, Asn 524, Lys 526 and Phe 530 participate in binding to the host cell receptor (14). The crystal structure of a receptor-bound diphtheria toxin complex is known (16) (Figure 1B). Both the NAD^+-binding cavity and the receptor-binding site contain formaldehyde reactive residues.

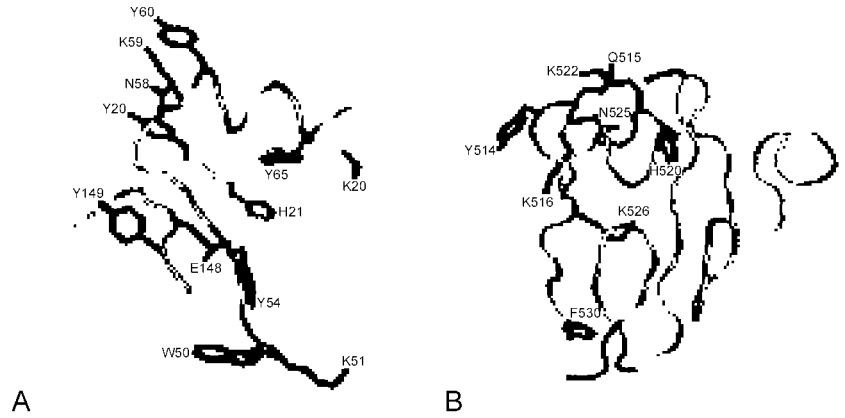


Figure 1. Images that represent two functional sites in diphtheria toxin: the NAD^+-binding cavity (A) and the receptor-binding loop (B). Picture A is based on the crystal structure of diphtheria toxin, and picture B on the crystal structure of the complex between diphtheria toxin and a fragment of cell-surface receptor (HB-EGF). Several side chains of amino acids of diphtheria toxin are represented because of their potential reactivity with formaldehyde or their participation in the protein function.

Materials and methods

Chemicals

Formaldehyde (37%), formic acid (99%), formamide, glycine, potassium dihydrogen phosphate ($KH_2PO_4{\cdot}3H_2O$) and dipotassium hydrogen phosphate ($K_2HPO_4{\cdot}3H_2O$) were purchased from Merck (Amsterdam, The Netherlands). Formaldehyde-D_2 (CD_2O) was supplied by C/D/N Isotopes Inc. (Utrecht, The Netherlands). Acetic acid N-hydroxy succinimide ester, DL-dithiotreitol (DTT) and sodium cyanoborohydride ($NaCNBH_3$) were obtained from Sigma (Zwijndrecht, The Netherlands). Chymotrypsin, endoproteinase Glu-C, trypsin and subtilisin were bought from Roche Applied Science (Almere, The Netherlands).

Chemical treatment of diphtheria toxin

Prior to reactions, diphtheria toxin (NVI, The Netherlands) was extensively against 10 mM potassium phosphate pH 7.2 (MWCO 10 kDa; CelluSept T3; Membrane Filtration Products, Inc; USA). Furthermore, formaldehyde (CH_2O), deuterium-labeled formaldehyde (CD_2O), glycine and $NaCNBH_3$ were dissolved in water to a concentration of 1.0 M. Acetic acid N-hydroxy succinimide ester was dissolved in formamide up to a final concentration of 0.8 M. Five reactions with diphtheria toxin (3mg/ml) were performed: (*1*) diphtheria toxin with formaldehyde and $NaCNBH_3$, (*2*) diphtheria toxin with acetic acid N-hydroxy succinimide ester, (*3*) diphtheria toxin with formaldehyde (CH_2O or CD_2O), (*4*) diphtheria toxin with formaldehyde (CH_2O or CD_2O) and glycine, and (*5*) diphtheria toxin with formaldehyde

(CH_2O) and $NaCNBH_3$ followed by the reaction with formaldehyde (CH_2O or CD_2O) and glycine. The composition and the reaction conditions for reaction ***4*** are most comparable to those used for vaccine preparation (10). The compositions of the reaction mixtures are given in table 1. For reaction ***1***, formaldehyde (CH_2O) and $NaCNBH_3$ were added to diphtheria toxin. The final concentrations of formaldehyde and $NaCNBH_3$ were 80 mM, and of diphtheria toxin 1.9 mg/ml. After mixing, the solution was incubated for 24 h at 35 °C. For reaction ***2***, N-hydroxy succinimide ester was added to diphtheria toxin. The final concentration of N-hydroxy succinimide ester was 13.3 mM. The mixture was incubated for 24 h at 35 °C. For reaction ***3***, diphtheria toxin was treated with 80 mM formaldehyde (CH_2O) or deuterium-labeled formaldehyde (CD_2O). Both mixtures were incubated for 1 week at 35 °C. For reaction ***4***, diphtheria toxin was treated with formaldehyde (CH_2O or CD_2O) and glycine. The final concentrations of formaldehyde and glycine were 80 mM. The solutions were incubated for 1 week at 35 °C. For reaction ***5***, 3.2 ml of the dialysed product of reaction ***1*** was subsequently incubated for 1 week at 35 °C with formaldehyde (CH_2O or CD_2O) and glycine. The final concentrations of formaldehyde and glycine were 80 mM and of diphtheria toxin 1.2 mg/ml. After each reaction, the samples were extensively dialysed against 10 mM potassium phosphate pH 7.2 (MWCO 10 kDa). For reactions ***3***, ***4*** and ***5***, after incubation and dialysis, equal volumes of the CH_2O-treated and CD_2O-treated samples were mixed. Finally, all samples were stored at 4 °C prior to the analyses.

Table 1. The composition of reaction mixtures.

		Reaction step 1				Reaction step 2	
Mixture[a)]	Diphtheria toxin[b)] toxin[b)]	Formaldehyde	Acetic acid ester	NaCNBH4	glycine	Formaldehyde	glycine
	(3 mg/ml)	(1M)	(0.8M)	(1M)	(1M)	(1M)	(1M)
1	3.2 ml	0.4 ml	-	0.4 ml	-	-	-
2	3.2 ml	-	83 µl		-	-	-
3	3.2 ml	0.4 ml	-	-	-	-	-
4	3.2 ml	0.4 ml	-	-	0.4 ml	-	-
5	3.2 ml	0.4 ml	-	0.4 ml	-	0.4 ml	0.4 ml

[a)] 10 mM potassium phosphate in water (pH 7.2.) was added to obtain a final volume of 5 ml.

[b)] The antigenicity of diphtheria toxin at this concentration (3 mg/ml) was 900 Lf/ml.

SDS-PAGE

SDS-PAGE was performed under reducing conditions, essentially as described by Sambrook et al. (17). Protein samples were prepared by mixing 2 µg of the toxoid in the sample buffer (60 mM Tris, 70 mM SDS, 0.1 M dithiothreitol, 0.1 mM tetrabromophenol blue and 35% glycerol diluted in water) to a volume of 20 µl and boiled for 10 min to denature the protein and to reduce disulfide bridges. The samples were loaded onto 10 % SDS-PAGE gels and electrophoretically separated. Molecular weight reference (broad range; Bio-Rad) was used for calibration. Protein bands were visualized by using

Coomassie brilliant blue. The gels were scanned and the intensity of protein bands was quantified by using the program Phoretix 1D quantifier (Phoretix International, UK).

TNBS assay

Primary amino group contents were determined using a colorimetric assay with 2,4,6-trinitrobenzenesulphonic acid (TNBS) (18).

Digestion by proteases

Diphtheria toxoids, obtained from reactions ***1*** and ***2***, were individually digested by the proteases chymotrypsin, endoproteinase Glu-C, trypsin and subtilisin. Diphtheria toxoids from reactions ***3*** and ***5*** were only digested by chymotrypsin. To that end, 10 µl of a 1.0 M buffer and 2 µl of 1.0 mg/ml protease were added to 80 µl toxoid. Water was added to a final volume of 100 µl. Each protease had a specific reaction buffer. The buffer for chymotrypsin was 1 M Tris-HCl pH 8.5 and 0.1 M $CaCl_2$; for proteinase Glu-C 1 M NH_4HCO_3 pH 8.5; for trypsin 1 M Tris-HCl pH 8.5; and for subtilisin 1 M Tris-HCl pH 8.8. Samples treated with, endoproteinase Glu-C, and trypsin were incubated for 24 h at 37 °C. Samples with subtilisin were incubated for 4 h at 37 °C. To reduce disulfide bonds, 1 µl of 0.1 M DTT was added after digestion and the samples were incubated for 1 h at 37 °C. Subsequently, the samples were stored at –20 °C before LC-MS analysis.

LC-MS

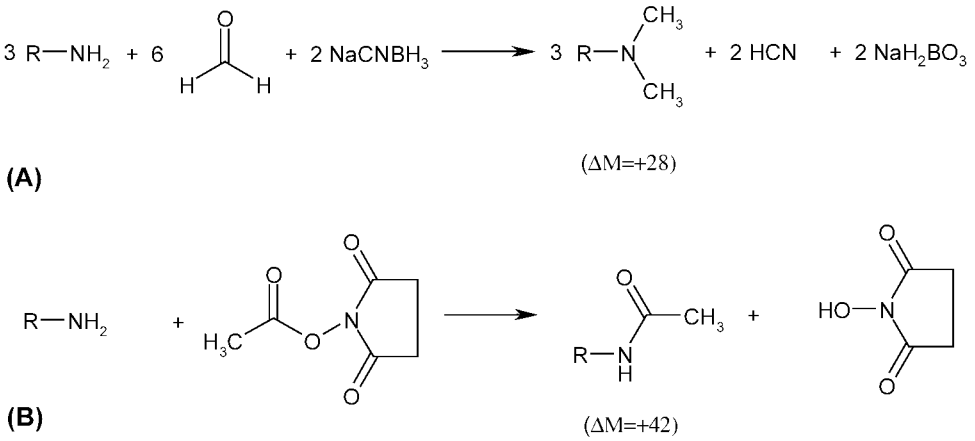
Protein digests were analyzed by nano-scale reversed-phase liquid chromatography electrospray mass spectrometry, essentially as previously described by Meiring et al. (19). The analysis was performed on two different mass spectrometers: a ThermoElectron LCQ™ Classic quadrupoe ion trap (San Jose, CA, USA) for the digests of reaction products ***1*** and ***2***, and a Waters Q-TOF Ultima API (Manchester, England) for the digests of reaction products **3** and **5**. The digests of reaction products ***1*** and ***2*** were diluted in water containing 5% (v/v) DMSO and 5% (v/v) formic acid to a concentration corresponding to 1.0 µM of the original protein concentration. An injection volume of 10 µl was used for analysis. Analytes were trapped on a 15 mm long x 100 µm inner diameter trapping column with Aqua C18 (5 µm; Phenomenex) at a flow rate of 3 µl/min of 100% solvent A (0.1 M acetic acid in water) as eluent for 10 min. Then, analytes were separated by reversed-phase chromatography by using a 25 cm long x 50 µm inner diameter. analytical column with Pepmap C18 (5 µm; Dionex) at a flow rate of 125 nl/min. A linear gradient was started with 5 % solvent B (0.1 M acetic acid in acetonitrile) to 60 % solvent B in 55 min. After the run, the columns were equilibrated in 100% solvent A for 10 min of 125 nl/min. The digested peptides were measured by data-dependent scanning comprising a MS-scan (m/z 350 – 2000) followed by collision-induced dissociation of the most abundant ion in the MS spectrum. The data were evaluated by using the TurboSequest software (ThermoElectron). The digests of reaction products ***3*** and ***5*** were diluted with water containing 5% (v/v)

DMSO and 5% (v/v) formic acid to a concentration corresponding to 2.5 nM of the original protein concentration. An injection volume of 10 µl was used for analysis. The same separation procedure was performed as described above to determine the masses of the digested peptides. Peptides containing formaldehyde modifications typically appeared as mass spectral doublets as a result of the use of "light" (CH_2O) and "heavy" (CD_2O) formaldehyde. The modified residues and the peptide sequence were assigned based upon the observed mass and the number of incorporated formaldehyde molecules.

Results

Dimethylation of lysine residues

The modifications in diphtheria toxoid after formaldehyde and glycine treatment consist of intramolecular cross-links and formaldehyde-glycine attachments. The intramolecular cross-links occur between a lysine residue and susceptible amino acid residues, including arginine, asparagine, glutamine, histidine, tryptophan and tyrosine (8). To determine the accessibility of each individual lysine residue for formaldehyde, diphtheria toxin was treated with formaldehyde and $NaCNBH_3$ (reaction *1*). In this reaction primary amino groups of lysine and N-terminal residues are converted to dimethylated structures with a mass increment of 28 Da (Scheme 1A) (8, 20). The TNBS assay showed a drastic reduction (>95%) of the number of primary amino groups in diphtheria toxin after formaldehyde/$NaCNBH_3$ treatment, which suggests that most lysine residues were modified (Figure 2). Furthermore, SDS-PAGE showed protein bands with slightly increased masses as a result of the dimethylation (Figure 3). Finally, the modified lysine residues were detected by LC-MS analyses of toxoid, which was individually digested with chymotrypsin, endoproteinase Glu-C, trypsin, or subtilisin. Eighty-eight percent of the primary sequence of diphtheria toxin, including the NAD^+ plus the receptor binding site, was detected by LC-



Scheme 1. Methylation (A) and acetylation (B) of lysine residues in proteins.

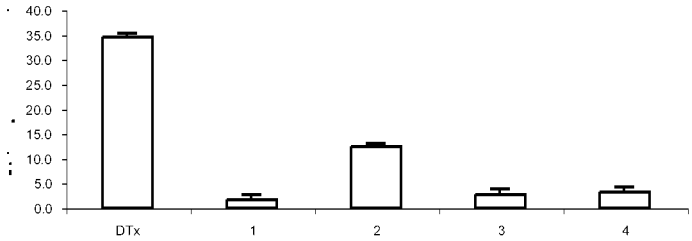


Figure 2. The number of primary amino groups in diphtheria toxin (DTx) and several toxoids (1 - 4) (mean ± S.D; n = 6). The toxoids were prepared by reaction ***1***, reaction ***2***, reaction ***3***, and reaction ***4*** (see Materials & Methods for details).

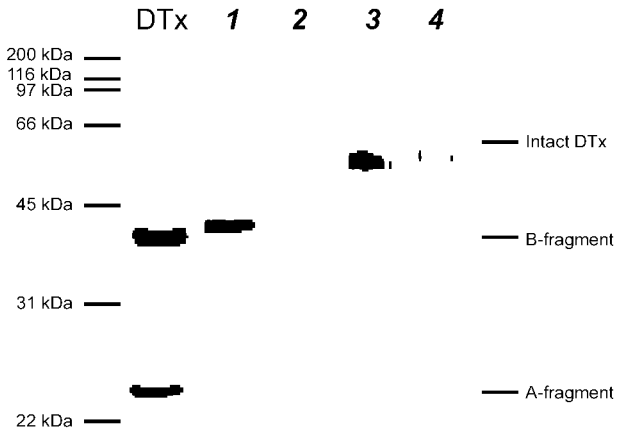


Figure 3. SDS-PAGE diphtheria toxin (DTx) and several toxoids (1 - 4). The toxoids were prepared by reaction ***1***, reaction ***2***, reaction ***3***, and reaction ***4*** (see Materials & Methods for details).

MS. Dimethylation was observed for all but one lysine residue in diphtheria toxin. Only dimethylation of the residue Lys 456 could not be confirmed. Also, the unmodified residue Lys 456 was not observed. The accessible surface area of this residue is 9.1%, as calculated according to Fraczkiewicz and Braun (21). Also from the crystal structure of diphtheria toxin (22) it becomes apparent that residue Lys 456 is very inaccessible as a result of steric hindrance by surrounding residues. However, other "inaccessible" lysine residues, e.g. Lys 103 with an accessible surface of 2.6%, were modified. Thus, we assume that the residue Lys 456 residue can also react with formaldehyde. Although the TNBS assay indicates the presence of a few primary amino groups in the diphtheria toxoid (from reaction **1**), no unmodified lysine residues were observed by LC-MS.

Acetylation of lysine residues

In our previous study with model peptides, we demonstrated that under conditions commonly used for vaccine preparation, formaldehyde primarily reacts with glycine and the formaldehyde-glycine adducts are then attached to several receptive amino acid residues (8). We assumed that these adducts are connected mainly to the reactive residues at the surface of the protein. Therefore as a simplified model, the attachment of acetyl groups to lysine residues of diphtheria toxin was investigated (reaction ***2***; Scheme 1B). The acetylation by acetic acid N-hydroxy succinimide ester was intended to mimic the attachment of formaldehyde-glycine adduct. The reaction caused a mass increase of diphtheria toxin, as was visualized by SDS-PAGE (Figure 3; lane ***2***) The mass increase was more than for the dimethylated toxin. Furthermore, a strong reduction of the number of primary amino groups (64%) was observed (Figure 2), although not as high as with dimethylation (96%; reaction ***1***). Apparently, the acetic acid N-hydroxy succinimide ester reacted much slower with buried lysine residues than the more exposed lysine residues (reaction ***2***).

LC-MS was used to demonstrate the extent of acetylation. Therefore, the toxoid obtained from reaction ***2*** was individually digested with chymotrypsin, endoproteinase Glu-C, trypsin, and subtilisin. Peptide fragments were detected with acetylated (mass increment of 42 Da) and non-acetylated lysine residues. Most lysine residues (27 out of 39) were completely acetylated. However, twelve lysine residues only partially reacted with the acetic acid ester, probably as a result of reduced reactivity (Table 2). In general, the partially modified residues have small accessible surface areas as a result of shielding by surrounding residues (Table 2). However, there are exceptions to this rule: according to the crystal

Table 2. Calculated accessibility of partially acetylated lysine residues of diphtheria toxin (reaction **2**).

Lysine residue no.	Calculated accessible surface[a] %
51	28
59	36
103	3
157	64
216	5
229	18
242	39
244	53
264	49
419	64
440	21
522	72

[a] The accessibility of each amino acid residue was calculated from the crystal structure of diphtheria toxin using a described method of Fraczkiewicz and Werner (21). The accessibility varied from 0 % for completely buried residues to 100% for surface residues. The accessible surface area of fully acetylated lysine residues were on average 56 ± 18% (mean ± S.D.), whereas of the partially acetylated residues 38 ± 22%.

structure, residues Lys 90, Lys 172 and Lys 385 are rather inaccessible (accessibility of 6.4%, 15.2% and 12.4%, respectively), but these residues showed 100% conversion.

Formaldehyde-induced cross-links in diphtheria toxin

The treatment of diphtheria toxin with a mixture of formaldehyde and glycine (reaction ***4***) results in a very heterogeneous toxoid. The heterogeneity could be visualized by SDS-PAGE, showing a very diffuse protein band of diphtheria toxoid compared to diphtheria toxin (Figure 3; lane ***4***). In theory, 140 residues in diphtheria toxin can react with formaldehyde. Since many residues are partly converted, diphtheria toxoid will consist of thousands of different reaction products. As a result of the high diversity in modifications, it is very laborious to determine all modified residues in this large protein (58.3 kDa). Therefore, in the present study we have focused on the formaldehyde-induced modifications in two areas of diphtheria toxin, which are important for the toxic action: (i) the NAD^+-binding groove in the catalytic domain, and (ii) the loop in the receptor domain involved in receptor binding.

In reaction ***3***, diphtheria toxin was treated with formaldehyde (CH_2O) or with deuterium-labeled formaldehyde (CD_2O) to introduce mainly intramolecular cross-links in the protein. Intermolecular cross-linking between two toxin molecules hardly occurred, otherwise it was observed by SDS-PAGE (Figure 3; lane ***3***). LC-MS analysis of chymotrypsin-digested mixtures of CH_2O-treated and CD_2O-treated toxin revealed eleven fragments containing an intramolecular cross-link (Table 3). Cross-links can be identified by their typical mass increases ($\Delta M = +12$ or $+24$ Da) (8).

Besides intramolecular cross-links, the formaldehyde-glycine attachments in diphtheria toxin were investigated. Prior to the formaldehyde and glycine reaction, lysine residues in diphtheria toxin were dimethylated to prevent intramolecular cross-linking (reaction ***5***). LC-MS analysis of the chymotrypsin-digested product revealed 82 peptides with formaldehyde modifications, as was clear from mass spectral doublets by the use of CH_2O and CD_2O. Thirty-six peptide sequences with modifications could be assigned based on the observed masses and the number of incorporated formaldehyde molecules. The assigned peptides cover 40% of the whole diphtheria toxin sequence.

The standard inactivation reaction for vaccine preparation (reaction ***4***), results in intramolecular cross-links, as observed in toxoids from reaction ***3***, and formaldehyde-glycine attachments, found in toxoids from reaction ***5***. The individual modifications at the NAD^+-binding cavity and the receptor-binding site have not been identified in this product (reaction ***4***).

Modifications in the catalytic site

Two toxin fragments with intramolecular cross-links were observed that belong to the NAD^+-binding groove (reaction ***3***). The cross-link is probably formed between the residues Lys 59 and Tyr 60 (peptides 2 and 3; Table 3). The distance of the side chains of both

Table 3. Diphtheria toxin-derived peptides with an intramolecular cross-link.[a)]

Peptide	Sequence	Fragment	Observed Mass
1	GADDVVDSSKSF[b)]	1-12	1238.6 (0.0)[c)]
2	STDNKYDAAGY	51-60	1216.6 (-0.1)
3	DDDWKGFYSTDNKYDAAGY	47- 65	2242.2 (0.7)
4	IKRFGDGASRVVL	124-136	1441.9 (-0.1)
5	IKRFGDGASRVVLSL	124-138	1641.7 (0.2)
6	INNWEQAKAL	150-159	1198.6 (0.0)
7	DVIRDKTKTKIESL	207-220	1669.9 (0.1)
8	NRPAYSPGHKTQPFL	376-390	1611.8 (0.0)
9	HRSSSEKIHSNEISSDSIGVL	492-512	2294.0 (-0.1)
10	GYQKTVDHTKVNSKLSL, GYQKTVDHTKVNSKLSL	513-528	1929.3 (0.7)
11	QKTVDHTKVNSKLSLF	515-530	1857.8 (-0.7)

[a)] Diphtheria toxin-derived peptides were obtained from reaction mixture 3 after the digestion with chymotrypsin (see Materials & Methods for details).

[b)] Intramolecular cross-links were most likely formed between the underlined residues. Both side chains are close to each other (<5Å), according to the crystal structure (22).

[c)] Deviation from the theoretical mass.

residues is about 2 Å (22). However, no cross-links were observed between Lys 59 and two other tyrosine residues, Tyr 20 and Tyr 181, although they are in close proximity of the lysine residue (ca. 4 Å). The masses of the expected cross-linked peptide fragments are presented in table 4. However, we assume that incomplete digestion by chymotrypsin resulted in several overlapping peptides, which are not recognized as such. Furthermore, the applied LC-method might be unsuitable for such large and probably hydrophobic peptide fragments, i.e. extremely long retention times or no elution at all.

Five peptide fragments with formaldehyde-glycine attachments were found, originating from the catalytic cleft (reaction ***5***; Table 5). However, the exact location of the modified residue(s) could not be determined, because all these assigned peptides contain two or more reactive residues. Previous work with synthetic peptides enables us to propose the 'most likely' structure of the modified peptides (Table 5). The tyrosine and arginine residues for

Table 4. Expected peptide fragments from the NAD^+-binding groove containing an intermolecular cross-link.[a)]

Sequence	Cross-link between	MH+	Mass observed
STDNKYDAAGY SSYHGTKPGY STDNKY	Lys 59 and Tyr60	1216.5	yes
STDNKY KEHGPIKNKMSESPNKTVSEEKAKQYL	Tyr 20 and Lys 59	1834.8	no
	Tyr 181 and Lys 59	3838.9	no

[a)] The peptides were expected after digestion of diphtheria toxoid, prepared with reaction ***3***.

Table 5. Modified peptide fragments from the NAD^+-binding site of diphtheria toxin.[a)]

MH+ (Da)	# CH_2O incorporated	Assigned peptide sequence [c)]	Fragment
874.4 (0.0)[b)]	1	**(H+87)GT(K+28)PGY** [d)]	21-27
		HGT(K+28)PG(Y +87)	
1451.6 (0.2)	1	**DAAG(Y +87)SVDNENPL**	61-73
		DAAGYSVD(N+87)ENPL	
		DAAGYSVDNE(N+87)PL	
1538.6 (0.2)	2	**DAAG(Y +174)SVDNENPL**	61-73
		DAAG(Y +87)SVD(N+87)ENPL	
		DAAG(Y +87)SVDNE(N+87)PL	
		DAAGYSVD(N+87)E(N+87)PL	
1542.8 (0.2)	1	**AEGSSSVE(Y+87)INNW**	141-153
		AEGSSSVEYI(N+87)NW	
		AEGSSSVEYIN(N+87)W	
		AEGSSSVEYINN(W+87)	
1630.0 (0.4)	2	**AEGSSSVE(Y+174)INNW**	141-153
		AEGSSSVE(Y+87)I(N+87)NW	
		AEGSSSVE(Y+87)IN(N+87)W	
		AEGSSSVE(Y+87)INN(W+87)	
		AEGSSSVEYI(N+87)(N+87) W	
		AEGSSSVEYI(N+87)N(W+87)	
		AEGSSSVEYIN(N+87)(W+87)	

[a)] The product of mixture 5 was digested by chymotrypsin (see Materials & Methods for details).

[b)] Deviation from the theoretical mass.

[c)] Lysine residues were dimethylated (ΔM +28 Da) prior to formaldehyde-glycine treatment.

[d)] Peptides containing the most likely modification(s) based on results obtained with synthetic peptides (8) are indicated in bold.

instance are the most reactive amino acids. The results indicate that formaldehyde-glycine adducts are formed at the residues His 21, Tyr 65 and Tyr 149. Amino acid residues His 21 and Tyr 65 are involved in the binding of NAD^+, whereas Tyr 149 is the adjacent residue of Glu 148 that participates in the binding. In conclusion, the catalytic cleft is affected by intramolecular cross-links and formaldehyde-glycine attachments.

Modifications in the receptor-binding site

A second location for which the effect of formaldehyde treatment was studied in more detail was the receptor-binding site of diphtheria toxin. With LC-MS analyses of the toxoid obtained from reaction ***3***, two masses were found belonging to peptide fragments of the receptor-binding site of diphtheria toxin (peptides 10 and 11; Table 3). The peptides probably contain an intramolecular cross-link between the amino acid residues Lys 522 and Gln 515, or between residues Lys 516 and Tyr 514.

Furthermore, three masses were found that could be ascribed to peptides with formaldehyde-glycine adducts (reaction ***5***; Table 6). The adducts are probably formed at

amino acid residues Tyr 514, His 520 and Asn 524. For residues Tyr 514 and Asn 524 participation in receptor binding has been demonstrated (14). Thus, both the receptor-binding site and the NAD^+-binding site are modified during the detoxification reaction.

Table 6. Modified peptide fragments from the receptor-binding site of diphtheria toxin.[a)]

MH+ (Da)	# CH_2O incorporated	Assigned peptide sequence	Fragment
1669.0 (0.1)[b)]	1	**Q(K+28)TVD(H+87)T(K+28)VNS(K+28)L**[c) d)] Q(K+28)TVDHT(K+28)V(N+87)S(K+28)L	515-527
1889.2 (0.4)	1	**G(Y+87)Q(K+28)TVDHT(K+28)VNS(K+28)L** GYQ(K+28)TVD(H+87)T(K+28)VNS(K+28)L GYQ(K+28)TVDHT(K+28)V(N+87)S(K+28)L	513-527
1976.2 (0.4)	2	**G(Y+174)Q(K+28)TVDHT(K+28)VNS(K+28)L** G(Y+87)Q(K+28)TVD(H+87)T(K+28)VNS(K+28)L G(Y+87)Q(K+28)TVDHT(K+28)V(N+87)S(K+28)L GYQ(K+28)TVD(H+87)T(K+28)V(N+87)S(K+28)L	513-527

[a)] The product of mixture 5 was digested by chymotrypsin (see Materials and Methods and Table 1 for details).
[b)] Deviation from the theoretical mass.
[c)] Lysine residues were dimethylated (ΔM +28 Da) prior to formaldehyde-glycine treatment.
[d)] Peptides containing the most likely modification(s) based on results obtained with synthetic peptides are indicated in bold (8).

Discussion

In the present study, the type and extent of formaldehyde-induced modifications of diphtheria toxin were investigated, with a detailed analysis of the NAD^+-binding groove and the receptor-binding site. In both areas of the toxin molecule, intramolecular cross-links and formaldehyde-glycine attachments were found. The conversion of these sites probably contributes to the inactivation of diphtheria toxin. In principle, all primary amino groups in diphtheria toxin are accessible for formaldehyde as demonstrated by the reaction with formaldehyde and $NaCNBH_3$ (reaction *1*). It was demonstrated for a few lysine residues that they form intramolecular cross-links (reaction *3*). The observed intramolecular cross-links (Table 3) are formed between residues that are in close proximity (<5Å). Several other intramolecular cross-links were expected according to the crystal structure. When the toxin is in solution, these cross-links might not be formed because of the high local mobility of the reactive amino acid residues involved in the formation of cross-links. The distance between the residues is probably too large to be effective in cross-linking.

Steric hindrance is probably a major factor contributing to the low or absent reactivity of amino acid residues with formaldehyde-glycine. The effects of steric hindrance were studied in a model reaction of diphteria toxin with acetic acid N-hydroxy succinimide ester (reaction *2*). Indeed as was observed, acetylation of buried amino (lysine) groups is in general not quantitative, whereas the more accessible residues were fully converted under the studied conditions. It is however likely that a similar phenomenon will be oberved for

the attachments of formaldehyde and glycine adducts to diphtheria toxin.

The detoxification process changes the antigenicity and immunogenicity of diphtheria toxoid (10). Although we observed many other masses related to modified peptides belonging to other parts in diphtheria toxin than the NAD^+-binding cavity and the receptor-binding loop, their contribution in reducing the toxicity or in changing the antigenicity is unknown. Theoretically, every intramolecular cross-link in the protein may be sufficient to inactivate the toxin. The formaldehyde-induced modifications on or near the immunodominant epitopes of diphtheria toxin are of interest to know, because these might affect the potency of the vaccine. The exact locations of immunodominant epitopes of diphtheria toxin are not yet known. Identification of these epitopes will be subject to future work.

In conclusion, the approach followed in this study is suitable to identify formaldehyde-induced modifications in diphtheria toxoid. The methods described here are suitable for the characterisation of diphtheria toxoids and, probably, also for other formaldehyde-inactivated antigens, including tetanus toxoid, pertussis toxoid and inactivated polio vaccine. The methods can be used in comparability studies, e.g. to support registration of these products after process or formulation improvements. The work demonstrates that with current powerful analytical methods it is possible to approach classical antigens as if they were well-defined biologicals.

References

1. Ramon, G., *Sur la toxine et sur l'anatoxine diphtériques.* Ann Inst Pasteur., 1924. **38**: 1-10.
2. Glenny, A.T. and B.E. Hopkins, *Diphtheria toxoid as an immunising agent.* Br J Exp Pathol, 1923. **4**: 283-287.
3. Fraenkel-Conrat, H. and D.K. Mecham, *The reaction of formaldehyde with proteins VII. Demonstration of intermolecular cross-linking by means of osmotic pressure measurements.* J Biol Chem, 1948. **177**: 477-487.
4. Fraenkel-Conrat, H. and H.S. Olcott, *Reaction of formaldehyde with proteins VI. cross-linking of amino groups with phenol, imidazole, or indole groups.* J Biol Chem, 1948. **174**: 827-843.
5. Blass, J., *préparation et étude d'une base de Mannich cristallisée obtenue par action du formol sur mélange de thréonine et de dimethy-2,4-phénol.* Bulletin de la société chimique de France, 1966. **10**: 3120-3121.
6. Blass, J., B. Bizzini, and M. Raynaud, *Mechanisme de la détoxification par le formol.* C.R.Acad.Sc.Paris, 1965. **261**: 1448-1449.
7. Kelly, D.P., M.K. Dewar, R.B. Johns, S. Wei-Let, and J.F. Yates, *Cross-linking of amino acids by formaldehyde. Preparation and 13C NMR spectra of model compounds.* Adv Exp Med Biol, 1977. **86**: 641-647.
8. Metz, B., G.F.A. Kersten, P. Hoogerhout, et al., *Identification of formaldehyde-induced modifications in proteins: reactions with model peptides.* J Biol Chem, 2004. **279**: 6235-43.
9. Rappuoli, R., *New and improved vaccines against diphtheria and tetanus*, in *New generation vaccines*, M.M. Levine, et al., Editors. 1997, Marcel Dekker, inc.: New York. 417-435.
10. Metz, B., W. Jiskoot, W.E. Hennink, D.J.A. Crommelin, and G.F.A. Kersten, *Physicochemical and immunochemical techniques predict the quality of diphtheria toxoid vaccines.* Vaccine, 2003. **22**: 156-167.
11. Van Ness, B.G., J.B. Howard, and J.W. Bodley, *ADP-ribosylation of elongation factor 2 by diphtheria toxin. Isolation and properties of the novel ribosyl-amino acid and its hydrolysis products.* J Biol Chem, 1980. **255**: 10717-20.
12. Papini, E., G. Schiavo, D. Sandona, R. Rappuoli, and C. Montecucco, *Histidine 21 is at the NAD+ binding site of diphtheria toxin.* J Biol Chem, 1989. **264**: 12385-8.
13. Blanke, S.R., K. Huang, and R.J. Collier, *Active-site mutations of diphtheria toxin: role of tyrosine-65 in NAD binding and ADP-ribosylation.* Biochemistry, 1994. **33**: 15494-500.
14. Shen, W.H., S. Choe, D. Eisenberg, and R.J. Collier, *Participation of lysine 516 and phenylalanine 530 of diphtheria toxin in receptor recognition.* J Biol Chem, 1994. **269**: 29077-84.
15. Naglich, J.G., J.E. Metherall, D.W. Russell, and L. Eidels, *Expression cloning of a diphtheria toxin receptor: identity with a heparin-binding EGF-like growth factor precursor.* Cell, 1992. **69**: 1051-61.
16. Louie, G.V., W. Yang, M.E. Bowman, and S. Choe, *Crystal structure of the complex of diphtheria toxin with an extracellular fragment of its receptor.* Mol Cell, 1997. **1**: 67-78.
17. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular cloning.* second ed. 1989, New York: Cold Spring Harbor Laboratory Press.
18. Habeeb, A.F.S.A., *Determination of free amino groups in proteins by trinitrobenzenesulfonic acid.* Anal. Biochem., 1966. **14**: 328-336.
19. Meiring, H.D., E.v.d. Heeft, G.J.t. Hove, and A.d. Jong, *Nanoscale LC-MS$^{(n)}$: technical design and applications to peptide and protein analysis.* J Sep Sci, 2002. **25**: 557-568.
20. Jentoft, N. and D.G. Dearborn, *Protein labeling by reductive alkylation.* Methods Enzymol, 1983. **91**: 570-579.
21. Fraczkiewicz, R. and W. Braun, *Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules.* J Comput Chem, 1998. **19**: 319-333.
22. Choe, S., M.J. Bennett, G. Fujii, et al., *The crystal structure of diphtheria toxin.* Nature, 1992. **357**: 216-22.