Identification of Formaldehyde-Induced Modifications in Diphtheria Toxin

Metz, B.; Michiels, T.; Uittenbogaard, J.; Danial, M.; Tilstra, W.; Meiring, H.D.; ... ; Jiskoot, W.

Journal of Pharmaceutical Sciences 109 (2020) 543-557

Contents lists available at ScienceDirect

Journal of Pharmaceutical Sciences

journal homepage: www.jpharmsci.org

Pharmaceutical Biotechnology

Identification of Formaldehyde-Induced Modifications in Diphtheria Toxin

Bernard Metz [1, 2, *], Thomas Michiels [1], Joost Uittenbogaard [1], Maarten Danial [1], Wichard Tilstra [1], Hugo D. Meiring [1], Wim E. Hennink [2], Daan J.A. Crommelin [2], Gideon F.A. Kersten [1, 3], Wim Jiskoot [2, 3]

[1] *Intravacc, Bilthoven, The Netherlands*
[2] *Faculty of Pharmaceutical Sciences, Department of Pharmaceutics, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, The Netherlands*
[3] *Division of BioTherapeutics, Leiden Academic Centre for Drug Research (LACDR), Leiden University, Leiden, The Netherlands*

ARTICLE INFO

Article history:
Received 9 August 2019
Revised 17 October 2019
Accepted 28 October 2019
Available online 31 October 2019

Keywords:
vaccine
diphtheria toxoid
formaldehyde
chemical modification
mass spectrometry

ABSTRACT

Diphtheria toxoid is produced by detoxification of diphtheria toxin with formaldehyde. This study was performed to elucidate the chemical nature and location of formaldehyde-induced modifications in diphtheria toxoid. Diphtheria toxin was chemically modified using 4 different reactions with the following reagents: (1) formaldehyde and $NaCNBH_3$, (2) formaldehyde, (3) formaldehyde and $NaCNBH_3$ followed by formaldehyde and glycine, and (4) formaldehyde and glycine. The modifications were studied by SDS-PAGE, primary amino group determination, and liquid chromatography–electrospray mass spectrometry of chymotryptic digests. Reaction 1 resulted in quantitative dimethylation of all lysine residues. Reaction 2 caused intramolecular cross-links, including the NAD^+-binding cavity and the receptor-binding site. Moreover, A fragments and B fragments were cross-linked by formaldehyde on part of the diphtheria toxoid molecules. Reaction 3 resulted in formaldehyde-glycine attachments, including in shielded areas of the protein. The detoxification reaction typically used for vaccine preparation (reaction 4) resulted in a combination of intramolecular cross-links and formaldehyde-glycine attachments. Both the NAD^+-binding cavity and the receptor-binding site of diphtheria toxin were chemically modified. Although $CD4^+$ T-cell epitopes were affected to some extent, one universal $CD4^+$ T-cell epitope remained almost completely unaltered by the treatment with formaldehyde and glycine.

Introduction

Diphtheria and tetanus toxoids are very effective vaccine antigens, virtually eliminating associated diseases in vaccinated populations. Many countries have included these vaccines in their national immunization programs, which has drastically reduced the incidence and severity of diphtheria and tetanus.[1] Currently, a dozen companies around the world are producing diphtheria and tetanus vaccines. These toxoid vaccines were developed almost a century ago.[2-4] As a result, extensive data sets have been collected by companies and official medicines control laboratories to assure the quality of successive vaccine lots.[5,6] Much experience with toxoid vaccines has been acquired by using the traditional potency and safety tests in animals. However, although the formaldehyde-induced chemical modifications of the antigen largely determine the quality of toxoid vaccines, little is known about the molecular structure of toxoids.

Interest has arisen in structural characterization of diphtheria, tetanus, and *Clostridium difficile* toxins[7-12] and toxoids.[13-26] Three reasons for this increased attention can be deduced from the literature, including (1) the development of analytical tests for quality control of toxoid-containing vaccines, (2) the development of advanced vaccine formulations for toxoid vaccines and (3) the use of toxoids as carrier proteins in conjugate vaccines:

Abbreviations used: DTT, DL-dithiotreitol; LC-MS, Tandem liquid chromatography – electrospray ionisation mass spectrometry; MS, mass spectrometry; m/z, mass-over-charge ratio; NAD^+, nicotinamide adenine dinucleotide; TNBS, 2,4,6-trinitrobenzene sulfonic acid; CID, collision-induced dissociation; ETD, electron-transfer dissociation.

This article contains supplementary material available from the authors by request or via the Internet at https://doi.org/10.1016/j.xphs.2019.10.047.

* *Correspondence to:* Bernard Metz (Telephone: +31-30-7920-485).
E-mail address: bernard.metz@intravacc.nl (B. Metz).

https://doi.org/10.1016/j.xphs.2019.10.047

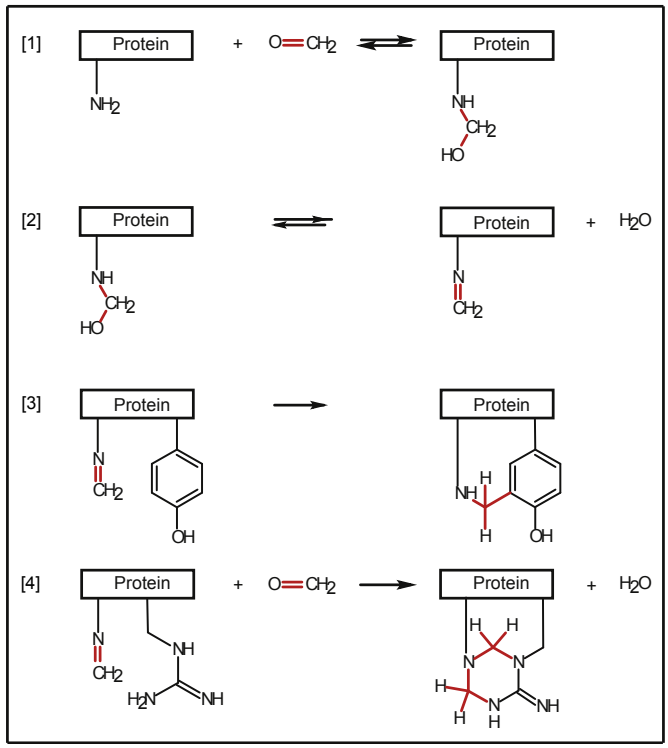
(1) Several biophysical and immunochemical tests have been developed to assess the quality of toxoid-containing vaccines.[16,18,19,21,23,25,27-37] These tests are valuable for examination of the purity, safety, potency, and stability of a product to reveal the impact of changes in a production process.[38,39]
(2) Furthermore, the development of innovative vaccine formulations with diphtheria and tetanus toxoids has demanded detailed structural characterization of the toxoids.[24,26,40-43] For example, controlled release or oral formulations with tetanus toxoid were designed by using polyester or chitosan microspheres.[20,42,44-47] Often, antigen instability in new vaccine delivery systems has been observed.[15,48-50]
(3) A third reason for renewed interest is the use of tetanus toxoid, diphtheria toxoid, and a mutant of diphtheria toxoid (CRM_{197}) as carrier proteins in the production of polysaccharide conjugate vaccines. Multiple conjugate vaccines are available on the market.[51] Furthermore, several new conjugate vaccines are under development, for example, vaccines against *Shigella flexneri* and *Salmonella typhi*.[52,53]

Altogether, insight into the molecular structure of toxoids may help to improve the quality control of vaccines, support rational development of advanced vaccine formulations, and facilitate the development of new conjugate vaccines.

Formaldehyde-mediated detoxification is an important step in the production of toxoid vaccines. To initiate the detoxification process, formaldehyde and extra glycine are usually added to diphtheria toxin—containing culture supernatant, which contains amino acids and metabolites. Formaldehyde treatment changes the toxicity, antigenicity, and immunogenicity of diphtheria and tetanus toxoid.[16,18,54] Formaldehyde treatment causes chemical modifications in toxoid molecules. Remarkably, the chemical nature of each specific modification and the exact locations of the modified residues within diphtheria toxoid molecules have not yet been identified. For diphtheria toxin, formaldehyde converts the protein into a nontoxic toxoid, probably by permanently altering critical domains in the protein, for example, the catalytic site (NAD^+-binding cavity) and the receptor-binding site. Although detoxification causes the loss of some B-cell epitopes, the toxoid remains very immunogenic and induces a protective immune response by the generation of toxin-neutralizing antibodies.[16]

The reaction of formaldehyde with a protein starts with the formation of reversible methylol adducts on amino groups (Scheme 1). The methylol groups are partially dehydrated, yielding labile Schiff-bases. These Schiff-bases generate intramolecular cross-links with accessible, reactive amino acid residues, including arginine, asparagine, glutamine, histidine, tryptophan, and tyrosine. Furthermore, formaldehyde can attach amino acids in solution to these reactive amino acid residues.[55] The conversion of reactive amino acid residues depends on their intrinsic reactivity and their accessibility for formaldehyde.[55,56]

The aim of the present study was to elucidate the chemical modifications in diphtheria toxoid as a result of the detoxification by formaldehyde and glycine. We investigated the reactivity of the individual lysine residues with formaldehyde, the formation of intramolecular cross-links and the attachment of formaldehyde-glycine moieties in diphtheria toxoid. Special emphasis was put on modifications in crucial parts of the molecule, that is, the catalytic site (NAD^+-binding cavity), the receptor-binding site, and $CD4^+$ T-cell epitope regions. The NAD^+-binding cavity is located in the catalytic domain of diphtheria toxin that transfers the ADP-ribose moiety of NAD^+ to elongation factor-2.[57] The modification of elongation factor-2 irreversibly inhibits the protein synthesis in the host cell leading to cell death. Three short peptide sequences in the diphtheria toxin molecule 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 (Fig. 1a). Amino acid residues His^{21}, Tyr^{54}, Tyr^{65}, and Glu^{148} participate in the binding of NAD^+.[58,59] Another important area in the toxin molecule is the receptor-binding site, which is formed by a loop of amino acid residues 511-530. This part of the receptor domain binds to the heparin-binding epidermal growth factor-like precursor.[60,61] The residues Tyr^{514}, Lys^{516}, Val^{523}, Asn^{524}, Lys^{526}, and Phe^{530} participate in binding to the host cell receptor.[60] The crystal structure of the receptor-bound diphtheria toxin complex is elucidated (Fig. 1b).[62] In addition, $CD4^+$ T-cell epitopes are identified by using blood from healthy subjects.[63] The $CD4^+$ T-cell epitopes are located in the B-fragment of diphtheria toxin: α-helices formed by residues 271-290, 321-340, and 331-350.



Scheme 1. The reaction of formaldehyde with proteins. The reaction starts with the formation of methylol adducts on amino groups [1]. The methylol adducts of primary amino groups are partially dehydrated, yielding labile Schiff bases [2], which can form cross-links with 6 different amino acid residues, for example, with tyrosine [3] and arginine [4].

The NAD^+-binding cavity, the receptor-binding site, and CD4+ T-cell epitopes contain formaldehyde reactive residues. The formaldehyde-induced modifications in NAD^+-binding cavity and the receptor-binding site of diphtheria toxin probably are responsible for complete detoxification. However, modifications in $CD4^+$ T-cell epitope regions of diphtheria toxin might reduce the immunogenicity of diphtheria vaccines. The chemical modifications in diphtheria toxin were analyzed by SDS-PAGE, primary amino group determination (TNBS assay) and liquid chromatography—electrospray mass spectrometry (LC-MS) after digestion with chymotrypsin. In this study, we revealed the location and chemical nature of modifications that occur in diphtheria toxoid during a detoxification process by formaldehyde and glycine.

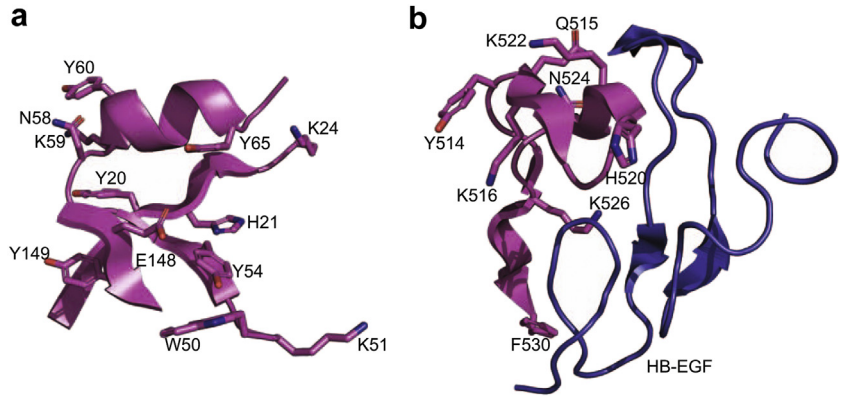


Figure 1. Catalytic and receptor binding sites in diphtheria toxin. The images represent 2 functional sites in diphtheria toxin: (a) the catalytic site (NAD^+-binding cavity; PDB: 1TOX) and (b) the receptor-binding site (PDB: 1XDT) interacting with the binding site of the cell-surface receptor (HB-EGF). Searching for PDB codes was performed on https://www.rcsb.org/. The side chains are shown of those amino acid residues of diphtheria toxin that are potentially reactive with formaldehyde.[55]

Materials and Methods

Chemicals

Formaldehyde (37%) (CH_2O), formaldehyde-D_2 (20%), (CD_2O), glycine, sodium cyanoborohydride ($NaCNBH_3$), sodium bisulphite, triethyl ammonium bicarbonate buffer pH 8.5, and dimethyl sulfoxide were obtained from Sigma-Aldrich (Schnelldorf, Germany). Formic acid (99%), sodium dihydrogen phosphate (NaH_2PO_4), disodium hydrogen phosphate (Na_2HPO_4), and sodium chloride were purchased from Merck (Darmstadt, Germany). Water was purchased from Biosolve Chimie SARL (Dieuze, France). Chymotrypsin was bought from Roche Diagnostics (Mannheim, Germany) and acetonitrile from Biosolve (Valkenswaard, the Netherlands).

Chemical Treatment of Diphtheria Toxin

Before detoxification reactions, diphtheria toxin—containing culture supernatant (Intravacc, the Netherlands) was dialyzed (MWCO 10 kDa; Slide-A-Lyzer Dialysis Cassette; Thermo Scientific; Rockford, IL) extensively against PBS (0.15 M NaCl, 7.7 mM Na_2HPO_4 and 2.3 mM NaH_2PO_4, pH 7.2). Aqueous solutions of formaldehyde (CH_2O), deuterium-labeled formaldehyde (CD_2O), glycine, and $NaCNBH_3$ were prepared at a concentration of 1.0 M in water. Four reactions (Fig. 2) with diphtheria toxin were performed: (reaction 1) diphtheria toxin with formaldehyde and $NaCNBH_3$, (reaction 2) diphtheria toxin with formaldehyde (CH_2O or CD_2O), (reaction 3) diphtheria toxin with formaldehyde (CH_2O or CD_2O) and glycine, (reaction 4) 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 conditions for reaction 2 are comparable to those used for vaccine production.[16]

Reaction 1

Formaldehyde (CH_2O) and $NaCNBH_3$ were added to diphtheria toxin (1.2 mg/mL) to final concentrations of 50 mM. After mixing, the solution was incubated for 2 h at 35°C. Then, the sample was extensively dialyzed against 10 mM PBS pH 7.2 (MWCO 10 kDa).

Reaction 2

Formaldehyde (CH_2O or CD_2O) was added to diphtheria toxin (1.2 mg/mL) to final concentrations of 80 mM. After mixing, the solutions were incubated for 1 week at 37°C. Then, the reaction was stopped by adding sodium bisulphite to a final concentration of 80 mM and subsequently extensively dialyzed against PBS (MWCO 10 kDa). Sodium bisulphite reacted with free formaldehyde in solution ($Na_2HSO_3 + CH_2O \rightleftharpoons HOCH_2SO_3Na$). As a result, the reversible methylol groups and Schiff bases on diphtheria toxoid were largely removed because the equilibrium shifted to the left (protein-NH_2 + $CH_2O \rightleftharpoons$ protein-$NHCH_2OH \rightleftharpoons$ protein-$NHCH_2 + H_2O$).

Reaction 3

Diphtheria toxin (1.2 mg/mL) was incubated for 2 h at 37°C with formaldehyde (CH_2O) and sodium cyanoborohydride. The final concentrations of formaldehyde and sodium cyanoborohydride in the samples were 50 mM. The sample was extensively dialyzed against PBS. The obtained dimethylated diphtheria toxin was incubated for 1 week at 37°C with formaldehyde (CH_2O or CD_2O) and glycine. The final concentrations of formaldehyde and glycine in the samples were 80 mM and the concentration of diphtheria toxin was 1.2 mg/mL. After the incubation, sodium bisulphite was added to a final concentration of 80 mM. Samples were extensively dialyzed against PBS (MWCO 10 kDa).

Reaction 4

Diphtheria toxin (1.2 mg/mL) was incubated for 1 week at 37°C with formaldehyde (CH_2O or CD_2O) and glycine. The final concentrations of formaldehyde and glycine in the samples were 80 mM. The reaction was stopped by adding sodium bisulphite to a final concentration of 80 mM. Both samples were extensively dialyzed against PBS (MWCO 10 kDa).

Finally, all samples were stored at 4°C before analysis by SDS-PAGE, TNBS assay, and LC-MS.

SDS-PAGE

SDS-PAGE was performed under reducing conditions, essentially as described by Sambrook et al.[64] 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 disulphide bridges. The samples were loaded onto 12% SDS-PAGE gels and electrophoretically separated. Molecular weight reference (broad range; Bio-Rad) was used for calibration. Protein bands were visualized by using Imperial Protein Stain (Pierce). The gels were scanned, and the intensity of protein bands was quantified using ImageJ 1.46r software (NIH).

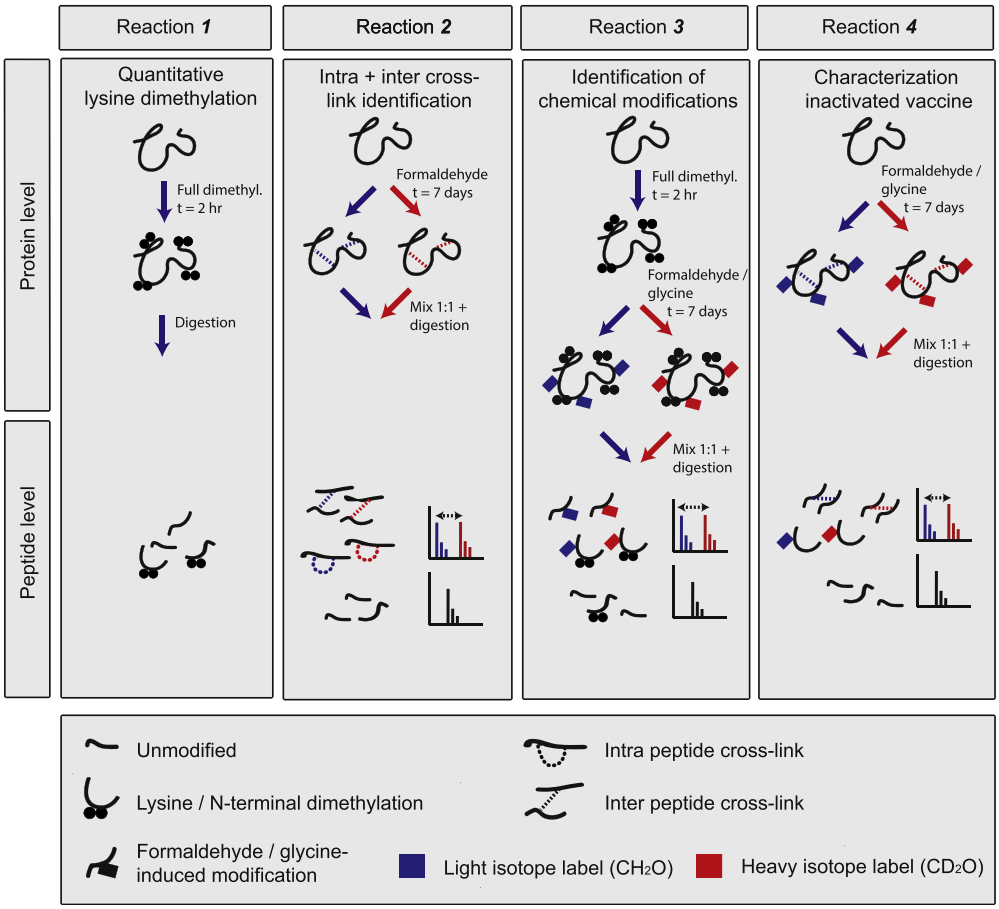


Figure 2. Characterization of formaldehyde-induced modifications in diphtheria toxoid by using stable isotope labeling strategies. Reaction 1: The accessibility of the lysine residues in diphtheria toxin after reductive dimethylation and digestion was determined by LC-MS/MS and database search analysis. Reaction 2: Target identification of cross-links in diphtheria toxoid formed on formaldehyde treatment. Treatment with light (CH_2O) and heavy (CD_2O) formaldehyde resulted in intramolecular cross-links that appear as mass spectral doublets, whereas unmodified peptides appear as singlets. After digestion, the formaldehyde-modified peptides contain a cross-link within the peptide (intrapeptide cross-link) or between 2 peptides (interpeptide cross-link). Reaction 3: Identification of modified residues by glycine attachments. First, lysine residues were blocked to prevent intrapeptide and interpeptide cross-linking. As a result of treatment with light (CH_2O) or heavy (CD_2O) formaldehyde and glycine, chemically modified peptides appear as mass spectral doublets, whereas unmodified peptides appear as singlets. Reaction 4: The standard detoxification of diphtheria toxin. The protein is treated with light (CH_2O) or heavy (CD_2O) formaldehyde and glycine, resulting in the differential labeling of both formaldehyde-induced cross-links and formaldehyde-glycine attachments.

Protein Assay

The protein concentration of dialyzed diphtheria toxin or toxoid samples was determined by using the BCA protein assay according to the manufacturer's description (Thermo Fisher Scientific, the Netherlands).

TNBS Assay

The concentration of primary amino groups present in diphtheria toxin or toxoid samples was determined by using a colorimetric assay with 2,4,6-trinitrobenzene sulfonic acid (TNBS).[65] A reference (10-100 μM) was prepared from a stock solution of 1.0 mM glycine. The number of primary amino groups in a diphtheria toxin or toxoid molecule was calculated; primary amino groups (mM)/protein concentration (mM). The molecular mass of 58.3 kDa for the diphtheria toxin was used for the calculation.

Digestion by Chymotrypsin

Diphtheria toxin and toxoids (from reaction 1, 2, 3, and 4) were individually digested by chymotrypsin (diphtheria toxin was used as a control.) To that end, 5 μL of a 1-M triethyl ammonium bicarbonate buffer, pH 8.5, and 2 μL of 1.0 mg/mL chymotrypsin were added to each sample containing 0.1 mg toxin or toxoid. Water was added to a final volume of 100 μL. Samples were incubated for 16 h at 37°C. Subsequently, the samples were stored at −20°C before LC-MS analysis.

Liquid Chromatography–Electrospray Mass Spectrometry

Protein digests were analyzed by nanoscale reversed-phase LC-MS, essentially as previously described by Meiring et al.,[66] using an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific). The digests of reaction products 1, 2, 3, and 4 were mixed in equal amounts and diluted in water containing 5% (v/v) dimethyl sulfoxide and 0.1% (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 loaded on a trapping column (Reprosil-Pur C18-AQ 5 μm, Dr. Maisch GmbH, Germany; 20 mm long × 100 μm inner diameter) with solvent A (0.1% (v/v) formic acid in water) in 10 min at 5 μL/min. The analytes were separated by reversed-phase chromatography on an analytical column (Reprosil-Pur C18-AQ 3 μm, Dr. Maisch GmbH, Germany; 27.5 cm long × 50 μm inner diameter) at a flow rate of 125 nL/min. A gradient was started with solvent B (0.1% (v/v) formic acid in

acetonitrile): 8% to 34% in 65 min and 60% for 5 min. After the gradient, the columns were equilibrated in 100% solvent A for 10 min at 125 nL/min. The peptides were measured by data-dependent acquisition, comprising a MS scan (m/z 300-1500) in the Orbitrap with a resolution of 120,000 (FWHM), followed by collision-induced dissociation (CID; ion trap) at top speed with a cycle time of 3 s. The threshold value for these precursor ions was set at 25,000 counts. The normalized collision energy was set at 35%, the isolation width at 1.6 Da, and the activation Q to 0.250. The maximum ion injection time for MS scans was set to 50 ms and for MS/MS scans to 150 ms. Precursor ions with +2 to +5 charge states were selected for MS/MS analysis. Dynamic exclusion was enabled (exclusion list with 500 entries) with repeat set to 1 and an exclusion duration of 45 s. The electron-transfer dissociation (ETD) reagent cation (202.0777) was used for internal mass calibration.

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 doublets were retrieved from the mass spectra by using the software program MsXelerator (MsMetrix, Maarssen, the Netherlands). Samples were measured in triplicate. Doublets that were found with a relative intensity of 10^5 arbitrary units and present in at least 2 of the 3 replicates were selected for further evaluation. The obtained doublet lists were exported to text files and used as parent mass lists for targeted ETD and CID fragmentation. For ETD fragmentation, charge state–dependent ETD parameters were used, and for CID fragmentation, the same settings were used as described previously. Identification of the MS/MS spectra was performed with PEAKS Studio 8.5 (Bioinformatics Solutions, Waterloo, ON, Canada) against the *Corynebacterium diphtheriae* proteome (taxonomic identifier 257,309, 2267 entries) with the earlier-described formaldehyde modifications[55] as variable modifications. Mass spectra that were not automatically assigned to modified peptides were manually evaluated based on the observed mass and the number of incorporated formaldehyde molecules.

In Silico Calculation of the Solvent Accessible Surface Area

The *in-silico* calculation of the solvent accessible surface area requires knowledge on the solvent radius and assumes that the solvent molecules are spheres. As such, the connolly molecular area 1 of the solvent molecule in Chem 3D Pro v11.0 was determined followed by rearranging the equation for the area of a sphere:

$$r_{\text{solvent}} = \frac{1}{2}\sqrt{\frac{A}{\pi}} \quad (1)$$

where A is the connolly molecular area and r_{solvent} is the molecular solvent radius. As presented in Supplementary Table 1, the value for r_{solvent} depended on the type of modification performed.

The solvent accessible surface area of monomeric diphtheria toxin (open form, PDB code: 1TOX[67]) was calculated by using the surface area per residue computation (relative solvent accessibility) in Pymol v. 2.1.1 (Schrodinger LLC) by setting the solvent radius to the values corresponding to formaldehyde or methylene-modified glycine (Supplementary Table 1). The generated residue–accessibility list was then exported to GraphPad Prism in which the plots were generated.

A solvent accessibility surface area calculation was also performed on the diphtheria toxin modified via *N,N*-dimethylation at the ε-amine group of all lysine residues (reaction 1). This was done by introducing the *N,N*-dimethylation of 1TOX PDB using the PyTM python script.[68] After this modification, the surface accessibility surface area was calculated by using formaldehyde as solvent, as described previously.

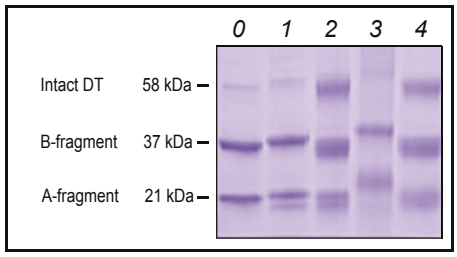


Figure 3. SDS-PAGE of diphtheria toxin (lane 0) and 4 experimental toxoids. The diphtheria toxoids were prepared by reaction 1 (lane 1), reaction 2 (lane 2), reaction 3 (lane 3), and reaction 4 (lane 4) (see Materials and Methods section for details).

Secondary Structure Information

The secondary structure information added to the plots is based on the crystal structure of diphtheria toxin (PDB code 1TOX[67]).

Results and Discussion

Production of Diphtheria Toxin and Toxoid

Diphtheria toxoid is usually prepared by adding formaldehyde and particular amino acids to the culture supernatant after the cultivation of *Corynebacterium diphtheriae*. Besides high concentrations of diphtheria toxin, the culture supernatant contains other proteins produced or secreted by the bacteria. Diphtheria toxin is synthesized as a single protein but has probably to be nicked into an A and B fragment to exert full biological activity.[69]

The dialyzed toxin batch used in this study was examined by SDS-PAGE and mass spectrometry to determine its purity. SDS-PAGE demonstrated that diphtheria toxin used in this study was almost completely nicked (Fig. 3, lane *0*). It showed a tiny band of the intact toxin at 58 Da and intense bands of the A and B fragment at 21 and 37 Da, respectively. The purity of the toxin was quantified by SDS-PAGE and appeared to be above 90%. To identify the proteins in the culture supernatant, chymotrypsin-digested material was analyzed by LC-MS. Thirty-one different proteins were identified in the concentrated culture supernatant (Supplementary Table 2). Based on the average response of the 3 most intensive peptides from each protein, the purity of the diphtheria toxin solution was estimated.[70] According to the LC-MS analysis, the culture supernatant contained 93 mole% diphtheria toxin. The contents of other proteins were between 0.05 and 1.3 mole%. The purity of diphtheria toxin was 96% based on protein weight (Supplementary Table 2). This purity was in line with the results from SDS-PAGE. The dialyzed culture supernatant was used to study in detail the chemical modifications of diphtheria toxin after formaldehyde treatment.

Formaldehyde-Reactive 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 a susceptible amino acid residue, that is, arginine, asparagine, glutamine, histidine, tryptophan, or tyrosine.[55] 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, the primary amino groups of lysine and N-terminal residues are converted to dimethylated structures with a mass increment of 28 Da.[71] Intact diphtheria toxin has 40 primary amino groups, whereas the toxin in the nicked form has one additional primary amino group.

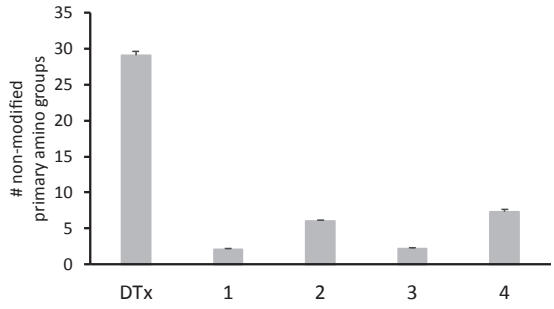


Figure 4. Average number of primary amino groups in diphtheria toxin (DTx) and several experimental toxoids (mean ± S.D.; $n = 9$). The modified toxoids were prepared by reaction 1, 2, 3, and 4 (see Materials and Methods section for details).

SDS-PAGE showed 3 protein bands of the formaldehyde-treated diphtheria toxin (reaction 2) with slightly increased masses when compared to the untreated toxin as a result of the dimethylation (Fig. 3, lane 1). Unexpectedly, a fourth protein band was observed at about 20 kDa. This 20-kDa band may represent some alternate fragmentation and not an impurity, as the diphtheria toxin preparation was highly pure (Supplementary Table 2). Another possibility might be that a formaldehyde-induced cross-link was introduced that caused a higher electrophoretic mobility due to incomplete unfolding by SDS-PAGE. However, no detailed information was collected in this study on the nature of this 20-kDa fragment.

The TNBS assay showed the presence of 29 primary amino groups in untreated diphtheria toxin (Fig. 4), which is less than the expected value (39 lysine residues and 2 N-termini). Apparently, the other primary amine groups in diphtheria toxin were not reached by TNBS or the ε-amino groups of the lysine residues were slightly less reactive with TNBS than the α-amino groups of the glycine reference. Treatment of diphtheria toxin with formaldehyde and $NaCNBH_3$ (reaction 1) resulted in a drastic reduction (93%) of the number of primary amino groups (Fig. 4). On average, 2 primary amino groups were present in diphtheria toxoid after reaction 1. This suggests that most lysine and N-terminal residues were accessible and modified by formaldehyde and $NaCNBH_3$ (reaction 1).

Furthermore, the modified lysine residues in diphtheria toxoid were identified by LC-MS analyses after digestion with chymotrypsin. Ninety-five percent (94.6%) of the total primary sequence of diphtheria toxin, including the complete NAD^+-binding and receptor-binding sites, was identified by LC-MS. Dimethylation ($\Delta M = +28$ Da) was observed for all lysine residues present in diphtheria toxin. The conversion into dimethylated lysine residues, calculated based on data of the 5 most abundant lysine-containing peptides, was on average 99.6 ± 0.3%. In conclusion, all lysine residues were accessible for formaldehyde-induced modifications, which can lead to cross-links with other reactive amino acid residues in diphtheria toxin.

Intramolecular Cross-Links in Diphtheria Toxin

During vaccine production, diphtheria toxin is usually treated with formaldehyde in the presence of amino acids. In theory, 142 of the 535 residues in diphtheria toxin can react with formaldehyde (N-terminal, Arg, Asn, Gln, His, Lys, Tyr, and Trp residues). Because many residues are partly converted, diphtheria toxoid may consist of a plethora of different reaction products. The stable formaldehyde-induced modifications can be divided into intramolecular cross-links and attachments of amino acids. In this study, 2 distinct reactions were performed to simplify the assignment of the formaldehyde-induced modifications in diphtheria toxoid, to induce either intramolecular cross-links (reaction 2) or formaldehyde-glycine attachments (reaction 3). In reaction 2, diphtheria toxin was treated with formaldehyde (CH_2O) or with deuterated formaldehyde (CD_2O) to introduce only intramolecular cross-links in the protein. The formaldehyde-glycine attachments (reaction 3) are described in the next section Formaldehyde-Glycine Attachments in Diphtheria Toxoid.

SDS-PAGE (Fig. 3; lane *2*) performed with formaldehyde-treated diphtheria toxoid (reaction 2) showed 4 broadened protein bands of the apparently intact toxoid (59 kDa), the A-fragment (21 kDa), and the B-fragment (36 kDa). Although less clear than in lane 1 (Fig. 2), a fourth protein band was observed in lane 2 at about 20 kDa. Furthermore, the protein bands were more spread out in lane 2 than in lane 0. The broad bands indicate that the formaldehyde treatment resulted in a very heterogeneous product. The broadened protein bands are probably caused by different intramolecular cross-links present in diphtheria toxoid molecules. Depending on the actually formed cross-links, diphtheria toxoid molecules were probably not completely unfolded by sodium dodecyl sulfate.[16] As a result, diphtheria toxoid and the fragments were visualized by SDS-PAGE as smeared protein bands. Furthermore, intermolecular cross-linking between toxoid molecules is unlikely, as indicated by the absence of any clear band larger than 59 kDa. However, the intensity of the apparently intact toxoid band (59 kDa) increased significantly compared to that of the toxin band (lane *0*). This observation can be explained by the formation of formaldehyde cross-links between the A-fragment and the B-fragment of diphtheria toxoid. The A-fragment and the B-fragment of diphtheria toxoid are in close proximity of each other, making the probability for formation of intramolecular cross-links by formaldehyde rather high. Intermolecular cross-links formed with small reactive molecules were not expected, because the toxin was extensively dialyzed against PBS and only formaldehyde was added in reaction 2. In conclusion, the results deduced from SDS-PAGE indicate that cross-links mainly occur intramolecularly at these formaldehyde and protein concentrations (Fig. 3; lane 2).

In addition, the number of primary amino groups in diphtheria toxoid was drastically reduced after formaldehyde treatment (reaction 2). The TNBS assay revealed, on average, the presence of 6 primary amino groups in each diphtheria toxoid molecule.

Furthermore, LC-MS analysis was performed on chymotrypsin-digested mixtures of CH_2O-treated and CD_2O-treated toxoid. The use of CH_2O versus CD_2O provided a method to discriminate formaldehyde-modified peptides (mass spectral doublets) from nonunmodified ones (mass spectral singlets). The analysis gave a protein coverage of about 95% for the whole sequence (96% of A-fragment and 95% of B-fragment). Sixty mass spectral doublets were observed with at least a relative intensity above 10^5 arbitrary units. Most peptides pairs had mass differences of 2 or 4 Da, indicating the presence of 1 or 2 formaldehyde-induced modification, respectively (Supplementary Table 3). One spectral doublet had a mass difference of 6 Da. Fifty modified peptide sequences were assigned based on MS/MS data, their exact masses, and the number of incorporated formaldehyde molecules. The formaldehyde-induced modifications in diphtheria toxoid (reaction 2) are described thereafter and depicted in Figure 5:

4-Imidazolididones

In diphtheria toxoid, formaldehyde converted the N-termini of the A-fragment and B-fragment into cyclic products (Fig. 5). These ring structures are also called 4-imidazolididones.[55,72,73] The adducts were revealed in the N-terminal peptides of the A-fragment (G^1ADDVVDSSKSF^{12}) and the B-fragment (S^{194}VGSSL^{199}), resulting in a mass increment of both peptides ($\Delta M = +12$ Da).

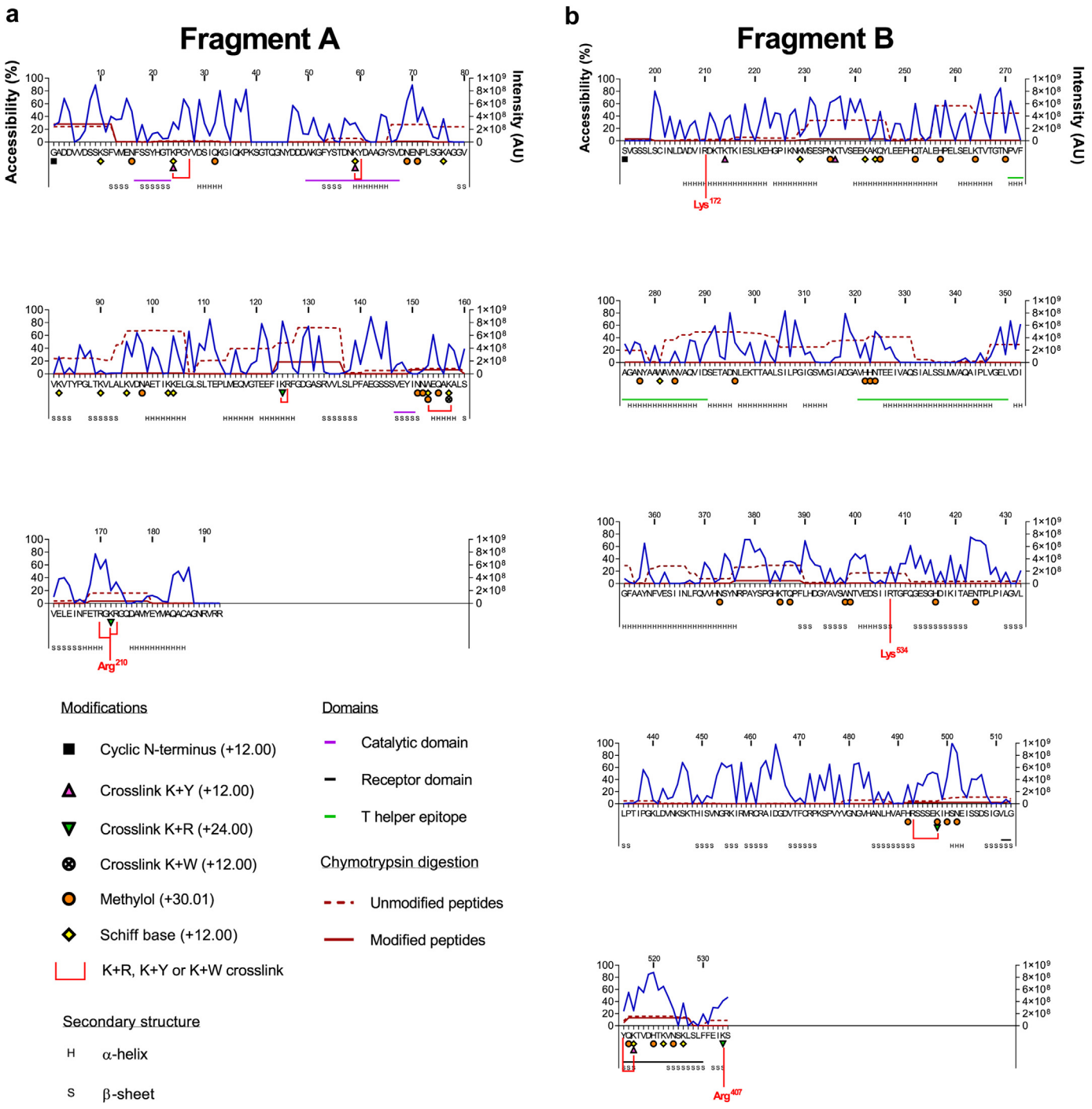


Figure 5. The modifications in diphtheria toxin caused by formaldehyde treatment (reaction 2). (a) indicates the A fragment and (b) the B fragment of diphtheria toxin. The blue line (y-axis left) indicates the theoretical solvent accessible surface area (accessibility) calculated using the molecular solvent radius of formaldehyde. The modified and unmodified peptides derived from chymotrypsin digestion are indicated with a solid and dotted dark red line, respectively, of which the intensities are plotted (y-axis right). In the case of overlapping peptides, the intensities are summed for the overlapping amino acid positions. The primary structure is indicated on the x-axis along with the amino acid position number in increments of 10 residues at the top of each plot. The different amino acid modifications identified by MS analysis along with the secondary structure elements derived from the crystal structure PDB 1TOX. The catalytic site, receptor binding site, and $CD4^+$ T-cell epitope are indicated with purple, black, and green underlines, respectively. The specific symbols are present in the figure legend. The red lines connect the residues between which a cross-link has been formed.

Methyl Group

One monomethylated lysine residue (K^{385}) was observed in diphtheria toxoid (reaction 2). Such modifications were observed previously in tetanus toxoid.[17] This particular modification cannot be caused by treatment of formaldehyde alone because a mild reducing agent is needed to convert a Schiff base into a monomethyl group. Probably, formic acid used before LC-MS analyses might act as a reducing agent (Eschweiler[74]–Clarke reaction). Therefore, this modification was not depicted in Figure 5.

Cross-Links

Seven unique peptides were identified that revealed formaldehyde-induced cross-links ($\Delta M = +24$ Da) between lysine and arginine residues in diphtheria toxoid (Fig. 5). Furthermore, 8

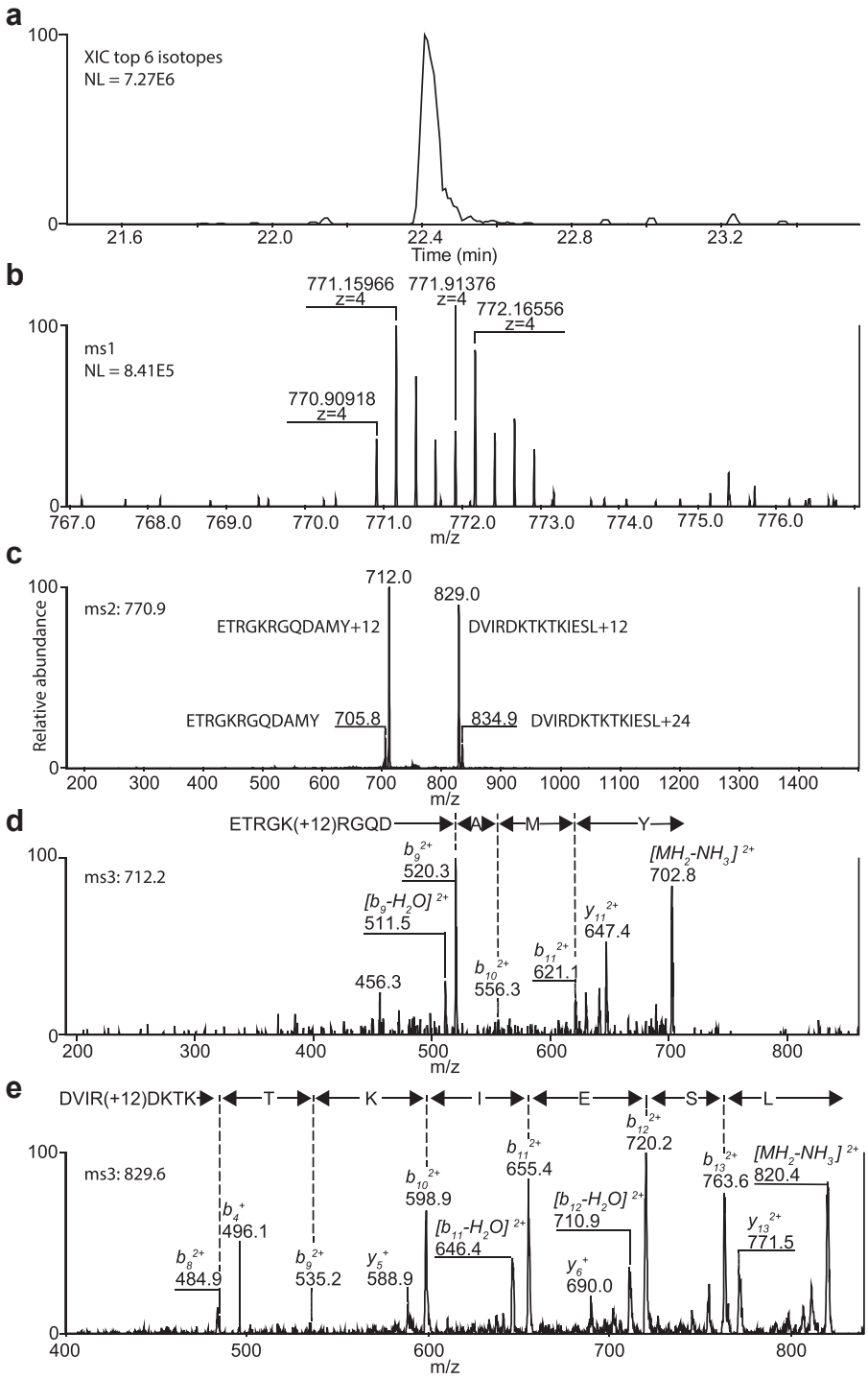


Figure 6. The intramolecular cross-link identified between A-fragment and B-fragment of diphtheria toxoid. LC/MS analyses performed on a chymotrypsin digest of diphtheria toxoid prepared by reaction 2: (a) chromatographic separation of the cross-linked dipeptide, and (b) the observed mass spectral doublet. (c) MS^2 analyses revealed the presence of a cross-link between 2 peptides: (d) ETRGKRGQDAMY (712.0 Da) and (e) DVIRDKTKTKIESL (829.0 Da). The peptide sequences of D and E were identified by MS^3 analysis.

MS/MS spectra indicated cross-links between lysine and a second amino acid residue, that is, a tyrosine and tryptophan residue. However, these cross-links could not be confirmed unambiguously because Schiff bases cause the same mass increment as an intramolecular cross-link (ΔM = +12 Da). Therefore, the digested peptides were treated with formaldehyde and $NaCNBH_3$ to confirm the presence of an intramolecular cross-link between lysine and tyrosine or between lysine and tryptophan residues. This reaction results in monomethylation of lysine residues in case of an intramolecular cross-link (ΔM = +12 Da (cross-link) +14 Da (methyl group) = +26 Da) and dimethylation in case of a Schiff base (ΔM = +28 Da). The LC-MS analysis showed the existence of 3 intramolecular cross-links between a lysine and a tyrosine residue, and one between a lysine and a tryptophan residue (Supplementary Fig. 1). In diphtheria toxoid, the formation of formaldehyde-induced cross-links between the A-fragment and the B-fragment was

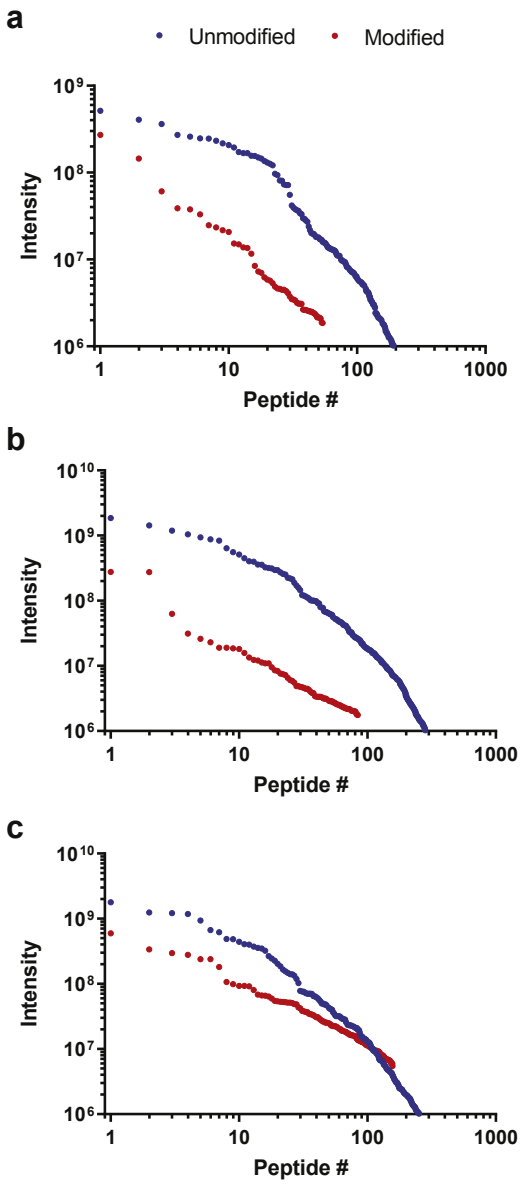


Figure 7. Relative peak intensities of mass spectral singlets (blue) and doublets (red) of peptides are presented and ranked in descending order. The singlets originate from unmodified peptides, whereas doublets are formed by the treatment with light (CH_2O) and heavy (CD_2O) formaldehyde. The relative peak intensities of peptides were determined with LC-MS (between 10^5 and 10^{10} arbitrary units) after treatment of diphtheria toxin with (a) formaldehyde (reaction 2), (b) formaldehyde and glycine after dimethylation (reaction 3) or (c) formaldehyde and glycine (reaction 4) and after digestion with chymotrypsin.

confirmed by SDS-PAGE. As such, we attempted to identify the amino acid residues that are likely to form these formaldehyde-induced cross-links. Based on the crystal structure of diphtheria toxin,[7] 2 reactive residues of the A-fragment are in close proximity of reactive residues in the B-fragment of diphtheria toxin, that is, Lys^{172}–Arg^{210} and $Arg^{190,192,193}$–Ser^{194}. The cross-link between residues Lys^{172} and Arg^{210} was confirmed by LC-MS (Fig. 6). Contrarily, the presence of a cross-link between 1 of 3 arginine residues ($Arg^{190,192,193}$) and the N-terminal amino group of B-fragment (Ser^{194}) was not found.

Cross-links were formed effectively when the distance between reactive residues was between 4.6 and 10.7 Å (Supplementary Fig. 1). Within these distances, however, several other formaldehyde-induced cross-links could be expected, but were not identified. However, incomplete protein coverage, low concentrations, or ionization intensities of these modified peptides or a poor peptide fragmentation (CID or ETD), resulting in inconclusive MS/MS spectra could lead to overseeing cross-link identifications. As such, the presence of expected cross-links cannot be ruled out.

Schiff-Bases

Nineteen peptides from diphtheria toxoid containing a Schiff-base were identified. The Schiff-bases were located on 15 lysine residues, 2 tryptophan residues and 1 histidine residue. Schiff-bases were found despite bisulphite treatment and extensive dialysis, performed to revert the reversible Schiff-bases and methylol moieties in diphtheria toxoid. The Schiff bases and methylol groups were more stable than expected.

Methylol Groups

Data analysis revealed the presence of reversible 26 methylol groups present in diphtheria toxoid ($\Delta M = +30$ Da). The methylol groups were located on 17 distinct asparagine, 6 glutamine, 4 histidine, 2 lysine residues, and 1 tryptophan residue.

Next to formaldehyde-induced modifications, oxidation of methionine and deamidation of asparagine residues were observed by LC-MS. However, the extent of oxidation and deamidation was not increased on exposure to 37°C for 7 days in reaction 2 (results not shown).

After formaldehyde treatment, the reactive residues in diphtheria toxoid are only partially converted. To get an impression of the degree of modifications, the intensities of the mass spectral singlets and doublets detected by LC-MS were plotted independently in descending order (Fig. 7a). The figure revealed that 17% of all hits were mass spectral doublets, that is, 17% of the peptide hits contained a formaldehyde-induced modification. Furthermore, the intensities of all singlets and doublets were added up, respectively. Based on the total sum of intensities of the singlets and the doublets, an average conversion by formaldehyde was calculated (Σ intensities of doublets/Σ intensities of singlets + doublets). The average conversion of the amino acid residues in diphtheria toxoid by reaction 2 was 11.6%.

In conclusion, the formaldehyde-treatment (reaction 2) resulted in 5 different types of modifications, including intramolecular cross-links in diphtheria toxoid.

Formaldehyde-Glycine Attachments in Diphtheria Toxoid

Diphtheria toxin was treated with formaldehyde and $NaCNBH_3$ (to prevent intramolecular cross-linking) and subsequently with formaldehyde (CH_2O) and glycine or with deuterium-labeled formaldehyde (CD_2O) and glycine to introduce formaldehyde-glycine attachments to receptive amino acid residues, that is, arginine, asparagine, glutamine, histidine, tryptophan, or tyrosine (reaction 3). Reaction 3 was performed to simplify the assignment of formaldehyde-glycine attachments.

Diphtheria toxoid with only formaldehyde-glycine attachments (reaction 3) revealed that the A-fragment and B-fragment of diphtheria toxoid had an apparently increased mass when compared to the untreated diphtheria toxin (Fig. 3; lane 3 vs. lane 0, respectively). Moreover, SDS-PAGE revealed that the protein bands were somewhat broadened as compared to those of diphtheria toxin (lane 0), especially the bands of the A-fragment and apparently intact diphtheria toxoid. This is probably due to the incorporation of different numbers of formaldehyde-glycine moieties in diphtheria toxoid molecules.

In addition, the average number of primary amino groups in diphtheria toxoid was considerably reduced (down to 8% relative to

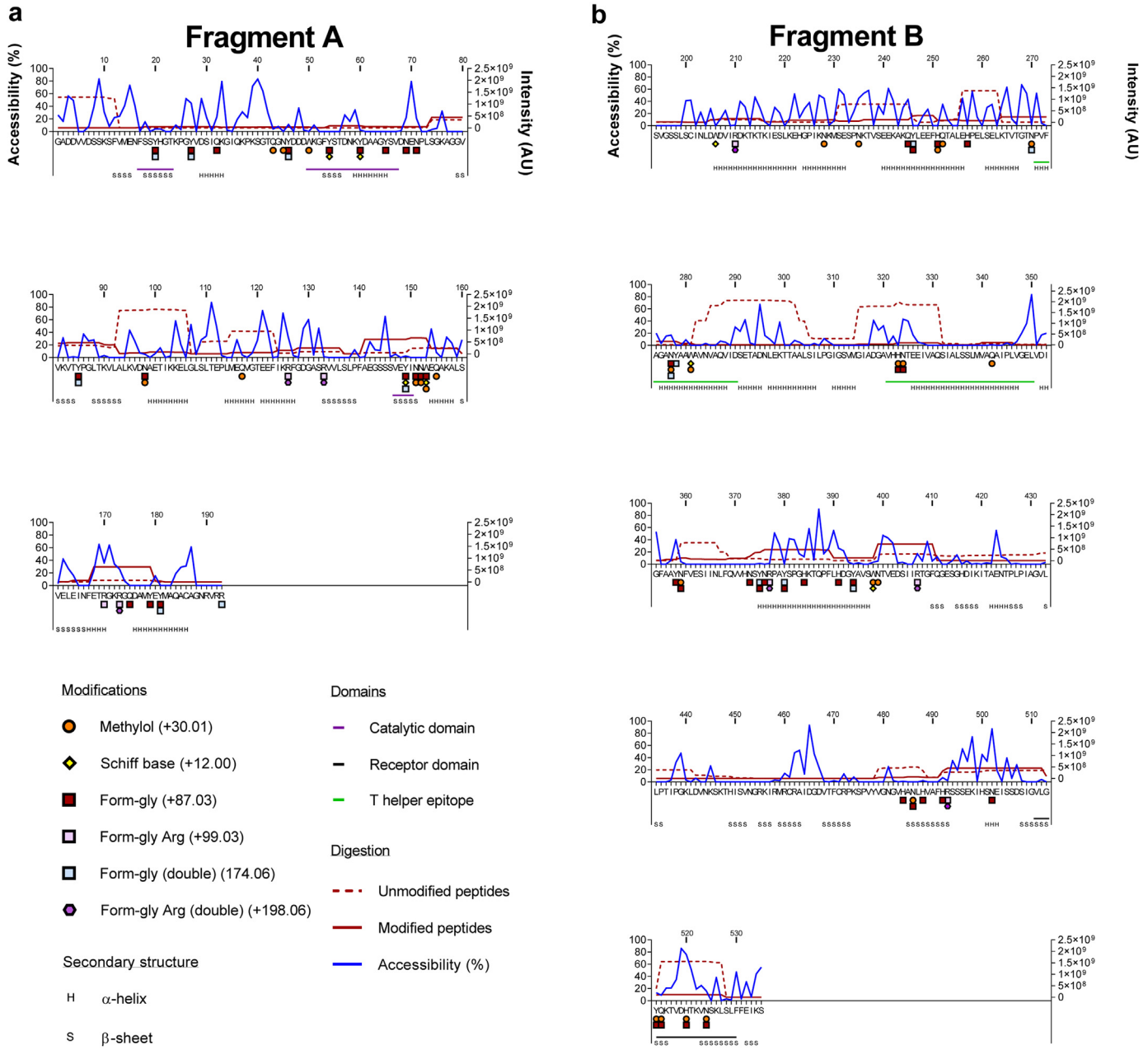


Figure 8. The modifications in diphtheria toxin caused by formaldehyde-glycine treatment after dimethylation of lysine residues (reaction 3). (a) indicates the A fragment and (b) the B fragment of diphtheria toxin. The blue line (y-axis left) indicates the theoretical solvent accessible surface area (accessibility) calculated using the molecular solvent radius of the formaldehyde-glycine Schiff base after dimethylation of the 1TOX crystal structure. The modified and unmodified peptides derived from chymotrypsin digestion are indicated with a solid and dotted dark red line, respectively, of which the intensities are plotted (y-axis right). In the case of overlapping peptides, the intensities are summed for the overlapping amino-acid positions. The primary structure is indicated on the x-axis along with the amino acid position number in increments of 10 residues at the top of each plot. The different amino acid modifications identified by MS analysis along with the secondary structure elements derived from the crystal structure PDB 1TOX. The catalytic site, receptor binding site and $CD4^+$ T-cell epitope are indicated with purple, black, and green underlines, respectively. The specific symbols are present in the figure legend.

the number in diphtheria toxin) after chemical treatment (reaction 3). The TNBS assay revealed the presence of 2 remaining primary amino groups on average in a diphtheria toxoid molecule. The number of amino groups detected in diphtheria toxoid (reaction 3) was comparable with the number of primary amino groups after dimethylation (reaction 1).

Furthermore, LC-MS analysis of the chymotrypsin-digested product of reaction 3 gave a protein coverage of 88% (72% of A-fragment and 96% of B-fragment). The analysis revealed 156 peptides with formaldehyde-glycine attachments (Supplementary Table 4). These formaldehyde-modified peptides were recognized by their mass spectral doublets by the use of CH_2O and CD_2O. Sequences of these modified peptides were assigned and several peptides were detected with multiple modifications. The modifications were identified by MS/MS analysis or allocated based on the observed masses and the number of incorporated formaldehyde molecules. Based on MS/MS analysis, 52 residues in diphtheria toxoid were identified containing a formaldehyde-glycine attachment (Fig. 8). Most of the moieties were attached to arginine and tyrosine residues, but also modifications of asparagine, glutamine, histidine, and tryptophan residues were observed. In theory, 103 of 535 amino acid residues in dimethylated diphtheria toxoid (reaction

3) can react with formaldehyde and glycine. However, arginine and tyrosine residues can attach 2 formaldehyde-glycine moieties.[55] Therefore, 137 formaldehyde-glycine moieties can be attached to one dimethylated diphtheria toxoid molecule. In addition to formaldehyde-glycine attachments, methylol groups and Schiff bases were observed on particular amino acid residues, that is, on asparagine, glutamine, histidine, and tryptophan residues.

To get an impression of the conversion of reactive residues by formaldehyde and glycine (reaction 3), the intensities of the modified (mass spectral doublets) and nonmodified peptides (mass spectral singlets) were plotted separately in descending order (Fig. 7b). The intensities of doublets relative to singlets give an impression on the conversion of the reactive amino acid residues in diphtheria toxoid molecules by formaldehyde and glycine. It revealed that 30% of all hits were mass spectral doublets, demonstrating the presence of a formaldehyde-induced modification in the peptide. The observations indicate that the average conversion of amino acid residues by formaldehyde and glycine was 21.6% (Σ intensities of doublets/Σ intensities of singlets + doublets).

The conversion of individual reactive residues in diphtheria toxoid by formaldehyde and glycine (reaction 3) depends on multiple factors, for example, concentrations of reagents (diphtheria toxin, formaldehyde, glycine), reaction time, and pH. Moreover, the intrinsic reactivity and accessibility of residues in diphtheria toxin probably determine the conversion degree. The reactivity depends on the type of amino acid residue, as demonstrated before.[55] The present study revealed that several rather inaccessible but reactive residues (<25% accessible) contained formaldehyde/glycine-induced modifications (Fig. 8). The accessibility of the modified residues was calculated based on the crystal structure of diphtheria toxin.[67] Unexpectedly, no relationship ($R^2 = 0.11$) was observed between the conversion of a particular type of residue (i.e., arginine and tyrosine) and its accessibility (Supplementary Fig. 2). Contrarily, the results of the present study do not match our previous study with insulin in which the conversion by formaldehyde and glycine was depending on the accessibility.[56] However, the correlation "conversion versus accessibility" was based on only 4 tyrosine residues present in insulin.

In conclusion, glycine molecules were attached to diphtheria toxin by formaldehyde. Most formaldehyde-glycine attachments were located on the arginine and tyrosine residues in diphtheria toxoid, irrespective of their position in the molecule.

Detoxification of Diphtheria Toxin by Formaldehyde and Glycine

The detoxification of diphtheria toxin with a mixture of formaldehyde and glycine resulted in a very heterogeneous toxoid (reaction 4). The heterogeneity of diphtheria toxoid could be visualized by SDS-PAGE (Fig. 3; lane 4), showing 3 broadened protein bands as compared to those of diphtheria toxin (lane 0), corresponding to the entire diphtheria toxoid (approximate molecular weight of 58 kDa), A-fragment (22 kDa), and B-fragment (37 kDa). After chemical treatment (reaction 4), the intensity of the apparent intact diphtheria toxoid band was increased when compared to diphtheria toxin (lane 0). The broadened protein bands were probably caused by intramolecular cross-links, as was shown for formaldehyde-treated diphtheria toxoid (reaction 2) by SDS-PAGE (Fig. 3; lane 2). Moreover, different numbers of formaldehyde-glycine moieties (i.e., intermolecular cross-links) may have been attached to diphtheria toxin molecules as well, which might have contributed to the broadening of the protein bands. In addition, the TNBS assay showed 75% reduction in the number of primary amino groups, with on average only 7 primary amino groups per toxoid molecule.

Multiple formaldehyde-induced modifications (reaction 4) in diphtheria toxoid were determined by using LC-MS analysis (Supplementary Table 5). The chemical nature and the location of formaldehyde/glycine-induced modifications were identified in diphtheria toxoid (Fig. 9).

The stable formaldehyde-induced modifications could be divided into intramolecular cross-links (reaction 2) and glycine attachments (reaction 3). Moreover, reversible methylol groups and Schiff bases were identified. LC-MS analysis provided a protein coverage of 92% (86% of A-fragment and 95% of B-fragment).

We focused predominantly on the formaldehyde-induced modifications in 3 areas of diphtheria toxin which are important for its toxicity and immunogenicity: (1) the NAD^+-binding cavity in the catalytic domain, (2) the loop in the receptor domain involved in receptor binding, and (3) T-cell epitopes. Chemical modifications in the NAD^+-binding cavity and receptor binding site contribute to complete detoxification of the diphtheria toxin. However, other modifications might contribute to the reduction of toxicity as well, for example, owing to conformational changes and cross-linking of A-fragment and B-fragment.

Catalytic Site

The NAD^+-binding cavity of the diphtheria toxin molecule consists of 3 short peptide sequences which are folded together (residues: 17-23, 50-67, and 147-150). In theory, 10 amino acid residues of the NAD^+-binding cavity are reactive with formaldehyde alone or with formaldehyde and glycine (Tyr^{20}, His^{21}, Trp^{50}, Lys^{51}, Tyr^{54}, Asn^{58}, Lys^{59}, Tyr^{60}, Tyr^{65}, Tyr^{149}). Amino acid residues His^{21}, Tyr^{54}, and Tyr^{65} are involved in the binding of NAD^+ (Fig. 1). Three formaldehyde-glycine attachments were found on tyrosine residues (Tyr^{54}, Tyr^{60}, and Tyr^{149}). The expected cross-link observed previously between residues Lys^{59} and Tyr^{60} (in reaction 3) was not identified after treatment with formaldehyde and glycine (reaction 4). Based on these data, we can conclude that formaldehyde-glycine attachments are the most common modifications present in the NAD^+-binding cavity, which will contribute to the inhibition of NAD^+-binding by diphtheria toxin.

Receptor-Binding Site

A second location studied in more detail for formaldehyde-induced modifications was the receptor-binding site of diphtheria toxin. With LC-MS analyses, modifications were identified at the receptor-binding site of diphtheria toxin (Fig. 9). The peptides contain an intramolecular cross-link between Lys^{516} and Tyr^{514}. Furthermore, formaldehyde-glycine attachments were 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.[60] Moreover, residue Lys^{522} was cross-linked by formaldehyde to another residue (e.g., Tyr^{514} and His^{520}) in the peptide (Gly^{513}—Leu^{527}) based on the mass increment of 12 Da found on this lysine residue.

T-Cell Epitopes

Diphtheria toxoid contains 3 $CD4^+$ T-cell epitopes which are recognized by 70%-82% of the diphtheria-vaccinated humans.[63] All $CD4^+$ T-cell epitopes are located in the B-fragment of diphtheria toxin residues 271-290, 321-340, and 331-350, with 2 overlapping epitopes. Several formaldehyde/glycine-induced modifications were found in epitope 271-290 and epitope 321-340. However, the epitope 331-350 contains only one reactive residue (Gln^{331}). A formaldehyde-glycine attachment to this glutamine residue occurred, but in the vast majority of the diphtheria toxoid molecules this epitope remained unchanged (Fig. 9). The latter is probably important for vaccine potency, as this particular epitope is recognized by 82% of the human population ($n = 100$).[63]

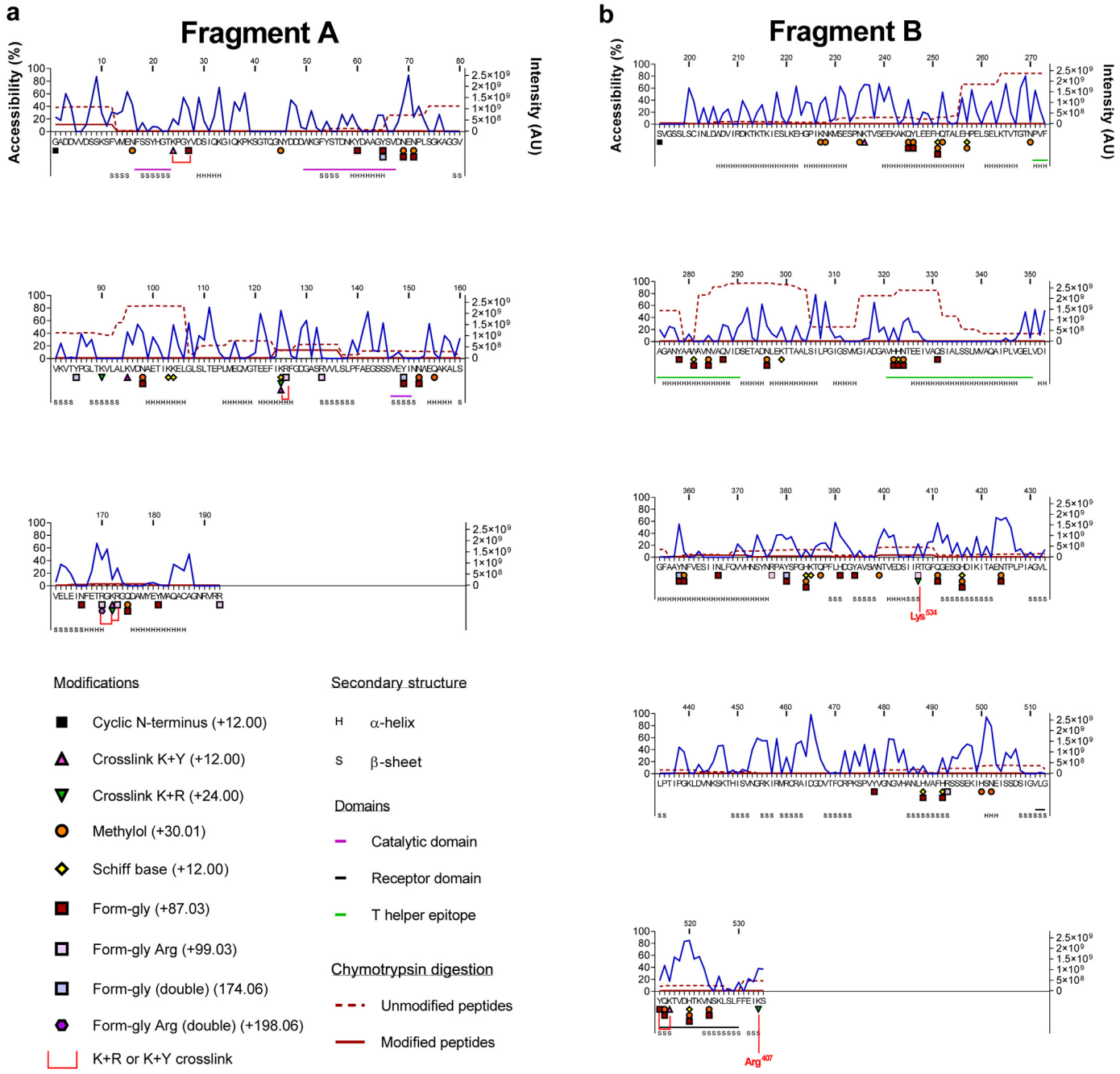


Figure 9. The modifications in diphtheria toxin caused by formaldehyde-glycine treatment (reaction 4). (a) indicates the A fragment and (b) the B fragment of diphtheria toxin. The blue line (y-axis left) indicates the theoretical solvent accessible surface area (accessibility) calculated using the molecular solvent radius of the Schiff base of formaldehyde and glycine. The modified and unmodified peptides derived from chymotrypsin digestion are indicated with a solid and dotted dark red line, respectively, of which the intensities are plotted (y-axis right). In the case of overlapping peptides, the intensities are summed for the overlapping amino-acid positions. The primary structure is indicated on the x-axis along with the amino acid position number in increments of 10 residues at the top of each plot. The different amino acid modifications identified by MS analysis along with the secondary structure elements derived from the crystal structure PDB 1TOX. The catalytic site, receptor binding site, and $CD4^+$ T-cell epitope are indicated with purple, black, and green underlines, respectively. The specific symbols are present in the figure legend. The red lines are connecting the residues between which a cross-link has been formed.

The conversion of each reactive amino acid residues in a regular diphtheria toxoid, induced by formaldehyde and glycine (reaction 4), is difficult to determine. An impression is given by comparing the intensities of the mass spectral doublets with the mass spectral singlets (Fig. 7c). The study revealed that 17% of the peptides contained formaldehyde-induced modifications. An average conversion of amino acid residues in diphtheria toxoid by formaldehyde and glycine (reaction 4) was calculated based on the total sum of intensities of the singlets and the doublets (Σ intensities of doublets/Σ intensities of singlets + doublets). The average conversion of the amino acid residues in diphtheria toxoid was 6.2%. The average conversion by a standard detoxification reaction (reaction 4) was less than with the other 2 reaction conditions (reactions 2 and 3).

Based on these data, we conclude that both the NAD^+-binding cavity and the receptor-binding site are affected by intramolecular cross-links or by formaldehyde-glycine attachments during the

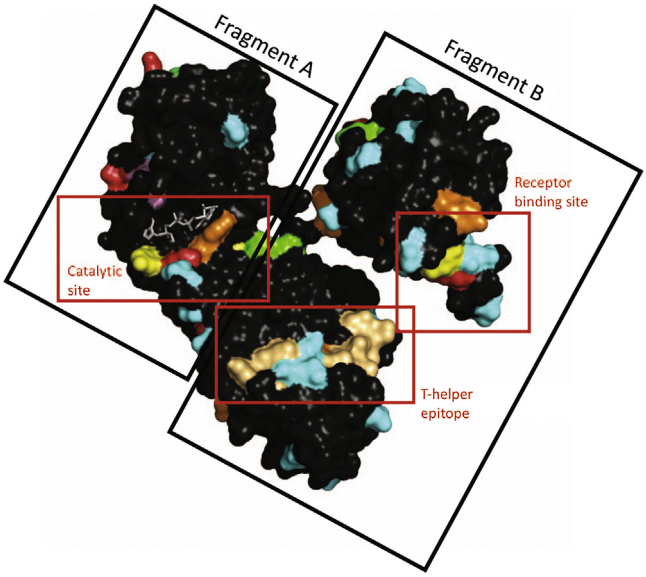


Figure 10. Illustration of formaldehyde/glycine-induced modifications on diphtheria toxin (PDB 1TOX). Amino acid residues that have cross-links being confirmed by mass spectrometric analysis are colored red (Lys), green (Arg), yellow (Tyr), and pink (Trp). Formylation-glycine modifications on tryptophan and tyrosine are colored orange. Methylol, Schiff base, or formaldehyde-glycine modifications on His, Gln, and Asn are colored cyan. Formylation-glycine modifications on Arg are colored green. Nonmodified amino acids are colored black. Pymol v2.1.1 (Schrodinger LLC) was used to create this illustration.

standard detoxification reaction (reaction 4). A combination of formaldehyde-induced modifications in diphtheria toxoid will ensure complete detoxification of each toxoid molecule. Despite the many modifications in several $CD4^+$ T-cell epitopes, one major epitope (at residue numbers 321-340) remained largely unmodified (Fig. 10).

Concluding Remarks

The formaldehyde-induced modifications in diphtheria toxoid were investigated with a detailed analysis of the NAD^+-binding cavity and receptor-binding site. In both areas of the toxoid molecule, intramolecular cross-links and formaldehyde-glycine attachments were found. Moreover, cross-links occurred between the A-fragment and B-fragment of diphtheria toxoid. The conversion of these sites contributes to the inactivation of diphtheria toxin (Fig. 10).

The study revealed that all primary amino groups in diphtheria toxin are accessible to formaldehyde (demonstrated by reaction 1). Only 11 of 39 lysine residues formed formaldehyde-induced intramolecular cross-links with neighboring amino acid residues (reaction 2). The formaldehyde-induced intramolecular cross-links (Table 1) were formed between residues in close proximity of each other (distances of α-carbon atoms 3.8-10.9 Å). However, several other intramolecular cross-links that were expected, according to the crystal structure of diphtheria toxin, could not be confirmed. Possibly, these cross-links were not constructed because these residues are very mobile in solution or because their orientation is incompatible with cross-linking. In addition to intramolecular cross-links, formaldehyde- and glycine-induced modifications were introduced during the detoxification reaction by formaldehyde and glycine. The modifications are most frequently found on arginine and tyrosine residues. Remarkably, relatively inaccessible but reactive residues were also modified by formaldehyde and glycine. In a few particular cases, the conversion of inaccessible residues was rather high (above 50%; Fig. 10). Unlike our expectation, the

Table 1
Expected Formaldehyde-Induced Cross-Links in Diphtheria Toxin

Position	Sequence	Distance Reactive Sites (Å)[a]	α-Carbon Distance (Å)[b]	Identified
K–R cross-links				
K90–R133	**K**VLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGAS**R**	13.2	6.1	no
K125–R126	**KR**	10.7	3.8	yes
R170–K172	**RGK**	8.9	5.6	yes
K172–R173	**KR**	6.8	3.8	yes
K172–R210	**K**RGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVI**R**	7.5	10.9	yes
R210–K214	**R**DKT**K**	7.0	6.6	no
R407–K534	**R**TGFQGESGHDIKITAENTPLPIAGVLLPTIPGKLDVNKSKTHI SVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHV AFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEI**K**	6.7	4.8	yes
K440–R493	**K**LDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVY VGNGVHANLHVAFH**R**	10.7	5.5	no
K445–R462	**K**SKTHISVNGRKIRMRC**R**	5.3	9.0	no
R493–K498	**R**SSSE**K**	12.7	9.7	yes
K–Y cross-links				
K24–Y27	**K**PG**Y**	8.6, 10.4	5.8	yes
K24–Y65	**K**PGYVDSIQKGIQKPKSGTQGNYDDDWKGFYST DNKYDAAG**Y**	7.1, 8.4	7.5	no
K51–Y54	**K**GF**Y**	10.0, 12.0	9.2	no
K59–Y60	**KY**	4.6, 5.3	3.8	yes
K59–Y181	**K**YDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLAL KVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGAS RVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFE TRGKRGQDAMYE**Y**	5.5, 6.6	9.4	no
K214–Y358	**K**TKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQT ALEHPELSELKTVTGTNPVFAGANYAAWAVNVAQVID SETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIV AQSIALSSLMVAQAIPLVGELVDIGFAA**Y**	6.5, 8.5	10.0	no
Y380–K385	**Y**SPGH**K**	5.7, 6.7	7.0	no
Y514–K516	**YQK**	4.6, 6.0	6.5	yes
K–W cross-links				
W153–K157	**W**EQA**K**	10.6	6.8	yes

[a] Distance measured from N^{ε} atom on Lys to one or more *N* atoms of the guanidine group on Arg, *meta* C-H bond on Tyr, or to *N* atom on the indol group of Trp.
[b] Distance measured between the α-carbon atoms between 2 cross-linked residues.

accessibility did not evidently steer the conversion of reactive residues. Probably, the formaldehyde-glycine moieties could reach rather inaccessible residues due to conformational dynamics within the diphtheria toxin molecule.

The detoxification process changes the antigenicity and immunogenicity of diphtheria toxoid.[16] The formaldehyde-induced modifications located on or near the immunodominant B-cell and T-cell epitopes of diphtheria toxin are of special interest. These modifications might affect the antigenicity and immunogenicity significantly. The immunodominant $CD4^+$ T-cell epitopes have been identified and examined.[63] Our study revealed that one dominant $CD4^+$ T-cell epitope remained almost completely unaltered on formaldehyde-glycine treatment. This particular epitope is recognized by 82% of the subjects ($n = 100$). Unfortunately, the immunodominant B-cell epitopes are unknown. Our previous study revealed that formaldehyde treatment results in the impairment of particular B-cell epitopes in diphtheria toxin. The impairment was measured by biosensor analysis using a set of monoclonal antibodies.[16] Therefore, formaldehyde-induced modifications in these immunodominant B-cell epitopes cannot be excluded. A different detoxification matrix might result in a distinct antigenicity and immunogenicity. Therefore, the detoxification process could determine the efficacy of the diphtheria vaccine in humans considerably.

The strategy followed in this study was suitable to identify the different types of formaldehyde-induced modifications in diphtheria toxoid. However, in many vaccine production processes of diphtheria toxoid, the detoxification of diphtheria toxin takes place in a matrix of culture supernatant with various amino acid compositions. The attachments of different amino acids to the reactive residues of diphtheria toxin will add to the heterogeneity of the resulting toxoid. The identification of modified residues in such a diphtheria toxoid can be very difficult and laborious, even with the sophisticated mass spectrometers and software tools available today. Nevertheless, the methods can be used to analyze such products in comparability studies, for example, to support registration of these products after process or formulation improvements. The work demonstrates that with current powerful analytical techniques, it is possible to approach classical vaccines as if they were well-defined biologicals.

Acknowledgments

The authors thank Geert Mommen for his fruitful discussions and support to this work. This project was funded by the Ministry of Agriculture, Nature and Food Quality.

References

1. Liang JL, Tiwari T, Moro P, et al. Prevention of pertussis, tetanus, and diphtheria with vaccines in the United States: recommendations of the advisory committee on immunization practices (ACIP). *MMWR Recomm Rep*. 2018;67(2):1-44.
2. Glenny AT, Hopkins BE. Diphtheria toxoid as an immunising agent. *Br J Exp Pathol*. 1923;4(283):283-287.
3. Ramon G. Sur la toxine et sur l' anatoxine diphtériques. *Ann Inst Pasteur*. 1924;38:1-10.
4. Ramon G, Descombey P. Sur l'Immunisation antitétanique et sur la production de l'antitoxine tétanique. *CR Soc Biol*. 1925;93:508-509.
5. Corbel MJ, Das RG, Lei D, et al. WHO working group on revision of the manual of laboratory methods for testing DTP vaccines-report of two meetings held on 20-21 July 2006 and 28-30 March 2007, Geneva, Switzerland. *Vaccine*. 2008;26(16):1913-1921.
6. World Heath Organization. *Manual for Quality Control of Diphtheria, Tetanus and Pertussis Vaccines*. Geneva: WHO Document Production Services; 2013. Expanded Programme on Immunization (EPI) of the Department of Immunization, Vaccines and Biologicals.
7. Choe S, Bennett MJ, Fujii G, et al. The crystal structure of diphtheria toxin. *Nature*. 1992;357(6375):216-222.
8. Fotinou C, Emsley P, Black I, et al. The crystal structure of tetanus toxin Hc fragment complexed with a synthetic GT1b analogue suggests cross-linking between ganglioside receptors and the toxin. *J Biol Chem*. 2001;276(34):32274-32281.
9. Breidenbach MA, Brunger AT. 2.3 A crystal structure of tetanus neurotoxin light chain. *Biochemistry*. 2005;44(20):7450-7457.
10. Masuyer G, Conrad J, Stenmark P. The structure of the tetanus toxin reveals pH-mediated domain dynamics. *EMBO Rep*. 2017;18(8):1306-1317.
11. Sauve S, Gingras G, Aubin Y. NMR study of mutations of glycine-52 of the catalytic domain of diphtheria toxin. *J Pharm Biomed Anal*. 2018;150:72-79.
12. Mishra RPN, Yadav RSP, Jones C, et al. Structural and immunological characterization of E. coli derived recombinant CRM197 protein used as carrier in conjugate vaccines. *Biosci Rep*. 2018;38(5):1-14.
13. Paliwal R, London E. Comparison of the conformation, hydrophobicity, and model membrane interactions of diphtheria toxin to those of formaldehyde-treated toxin (diphtheria toxoid): formaldehyde stabilization of the native conformation inhibits changes that allow membrane insertion. *Biochemistry*. 1996;35(7):2374-2379.
14. Schwendeman SP, Costantino HR, Gupta RK, et al. Stabilization of tetanus and diphtheria toxoids against moisture-induced aggregation. *Proc Natl Acad Sci U S A*. 1995;92(24):11234-11238.
15. Johansen P, Merkle HP, Gander B. Physico-chemical and antigenic properties of tetanus and diphtheria toxoids and steps towards improved stability. *Biochim Biophys Acta*. 1998;23(2):425-436.
16. Metz B, Jiskoot W, Hennink WE, Crommelin DJA, Kersten GFA. Physicochemical and immunochemical techniques predict the quality of diphtheria toxoid vaccines. *Vaccine*. 2003;22(2):156-167.
17. Thaysen-Andersen M, Jorgensen SB, Wilhelmsen ES, Petersen JW, Hojrup P. Investigation of the detoxification mechanism of formaldehyde-treated tetanus toxin. *Vaccine*. 2007;25(12):2213-2227.
18. Metz B, Tilstra W, Van der Put R, et al. Physicochemical and immunochemical assays for monitoring consistent production of tetanus toxoid. *Biologicals*. 2013;41(4):231-237.
19. Regnier M, Metz B, Tilstra W, et al. Structural perturbation of diphtheria toxoid upon adsorption to aluminium hydroxide adjuvant. *Vaccine*. 2012;30(48):6783-6788.
20. Shukla A, Mishra V, Bhoop BS, Katare OP. Alginate coated chitosan microparticles mediated oral delivery of diphtheria toxoid. Part A. Systematic optimization, development and characterization. *Int J Pharm*. 2015;495(1):220-233.
21. Alsarraf H, Dedic E, Bjerrum MJ, et al. Biophysical comparison of diphtheria and tetanus toxins with the formaldehyde-detoxified toxoids, the main components of diphtheria and tetanus vaccines. *Virulence*. 2017;8(8):1880-1889.
22. Bayart C, Peronin S, Jean E, et al. The combined use of analytical tools for exploring tetanus toxin and tetanus toxoid structures. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2017;1054:80-92.
23. Kalbfleisch K, Deshmukh S, Mei C, et al. Identity, structure and compositional analysis of aluminum phosphate adsorbed pediatric quadrivalent and pentavalent vaccines. *Comput Struct Biotechnol J*. 2019;17:14-20.
24. Samra HS, He F, Bhambhani A, et al. The effects of substituted cyclodextrins on the colloidal and conformational stability of selected proteins. *J Pharm Sci*. 2010;99(6):2800-2818.
25. Salnikova MS, Joshi SB, Rytting JH, Warny M, Middaugh CR. Physical characterization of clostridium difficile toxins and toxoids: effect of the formaldehyde crosslinking on thermal stability. *J Pharm Sci*. 2008;97(9):3735-3752.
26. Salnikova MS, Joshi SB, Rytting JH, Warny M, Middaugh CR. Preformulation studies of clostridium difficile toxoids A and B. *J Pharm Sci*. 2008;97(10):4194-4207.
27. Iwaki M, Horiuchi Y, Komiya T, et al. Toxoid flocculation assay by laser light-scattering. *J Immunol Methods*. 2007;318(1−2):138-146.
28. Prieur S, Broc S, Gal M, Poirier B, Fuchs F. Development of an in vitro potency test for tetanus vaccines: an immunoassay based on Hc fragment determination. *Dev Biol (Basel)*. 2002;111:37-46.
29. Bolgiano B, Fowler S, Turner K, et al. Monitoring of diphtheria, pertussis and tetanus toxoids by circular dichroism, fluorescence spectroscopy and size-exclusion chromatography. *Dev Biol (Basel)*. 2000;103:51-59.
30. Winsnes R, Sesardic D, Daas A, Behr Gross ME. Collaborative study for the validation of serological methods for potency testing of diphtheria toxoid vaccines-part 1. *Pharmeuropa Bio*. 2004;2003(2):35-68.
31. Sesardic D, Winsnes R, Rigsby P, Behr-Gross ME. Collaborative study for the validation of serological methods for potency testing of diphtheria toxoid vaccines−extended study: correlation of serology with in vivo toxin neutralisation. *Pharmeuropa Bio*. 2004;2003(2):69-76.
32. Metz B, Jiskoot W, Mekkes D, et al. Quality control of routine, experimental and real-time aged diphtheria toxoids by in vitro analytical techniques. *Vaccine*. 2007;25(39−40):6863-6871.
33. Coombes L, Stickings P, Tierney R, Rigsby P, Sesardic D. Development and use of a novel in vitro assay for testing of diphtheria toxoid in combination vaccines. *J Immunol Methods*. 2009;350(1−2):142-149.
34. Coombes L, Tierney R, Rigsby P, Sesardic D, Stickings P. In vitro antigen ELISA for quality control of tetanus vaccines. *Biologicals*. 2012;40:466-472.
35. Westdijk J, Metz B, Spruit N, et al. Antigenic fingerprinting of diphtheria toxoid adsorbed to aluminium phosphate. *Biologicals*. 2017;47:69-75.
36. Hickey JM, Toprani VM, Kaur K, et al. Analytical comparability assessments of 5 recombinant CRM197 proteins from different manufacturers and expression systems. *J Pharm Sci*. 2018;107(7):1806-1819.
37. Moller J, Kraner M, Sonnewald U, et al. Proteomics of diphtheria toxoid vaccines reveals multiple proteins that are immunogenic and may contribute to

protection of humans against corynebacterium diphtheriae. *Vaccine*. 2019;37(23):3061-3070.
38. Chirino AJ, Mire-Sluis A. Characterizing biological products and assessing comparability following manufacturing changes. *Nat Biotechnol*. 2004;22(11):1383-1391.
39. Federici M, Lubiniecki A, Manikwar P, Volkin DB. Analytical lessons learned from selected therapeutic protein drug comparability studies. *Biologicals*. 2013;41(3):131-147.
40. Sasiak AB, Bolgiano B, Crane DT, et al. Comparison of in vitro and in vivo methods to study stability of PLGA microencapsulated tetanus toxoid vaccines. *Vaccine*. 2000;19(7–8):694-705.
41. Jaganathan KS, Rao YU, Singh P, et al. Development of a single dose tetanus toxoid formulation based on polymeric microspheres: a comparative study of poly(D,L-lactic-co-glycolic acid) versus chitosan microspheres. *Int J Pharm*. 2005;294(1–2):23-32.
42. Alonso MJ, Gupta RK, Min C, Siber GR, Langer R. Biodegradable microspheres as controlled-release tetanus toxoid delivery systems. *Vaccine*. 1994;12(4):299-306.
43. Jung T, Kamm W, Breitenbach A, Klebe G, Kissel T. Loading of tetanus toxoid to biodegradable nanoparticles from branched poly(sulfobutyl-polyvinyl alcohol)-g-(lactide-co-glycolide) nanoparticles by protein adsorption: a mechanistic study. *Pharm Res*. 2002;19(8):1105-1113.
44. Kersten GF, Donders D, Akkermans A, Beuvery EC. Single shot with tetanus toxoid in biodegradable microspheres protects mice despite acid-induced denaturation of the antigen. *Vaccine*. 1996;14(17–18):1627-1632.
45. van der Lubben IM, Kersten G, Fretz MM, et al. Chitosan microparticles for mucosal vaccination against diphtheria: oral and nasal efficacy studies in mice. *Vaccine*. 2003;21(13–14):1400-1408.
46. Ahire VJ, Sawant KK, Doshi JB, Ravetkar SD. Chitosan microparticles as oral delivery system for tetanus toxoid. *Drug Dev Ind Pharm*. 2007;33:1112-1124.
47. Arthanari S, Mani G, Peng MM, Jang HT. Chitosan-HPMC-blended microspheres as a vaccine carrier for the delivery of tetanus toxoid. *Artif Cells Nanomed Biotechnol*. 2016;44(2):517-523.
48. Xing DK, Crane DT, Bolgiano B, et al. Physicochemical and immunological studies on the stability of free and microsphere-encapsulated tetanus toxoid in vitro. *Vaccine*. 1996;14(13):1205-1213.
49. Determan AS, Wilson JH, Kipper MJ, Wannemuehler MJ, Narasimhan B. Protein stability in the presence of polymer degradation products: consequences for controlled release formulations. *Biomaterials*. 2006;27(17):3312-3320.
50. Gordon S, Young K, Wilson R, et al. Chitosan hydrogels containing liposomes and cubosomes as particulate sustained release vaccine delivery systems. *J Liposome Res*. 2012;22(3):193-204.
51. Broker M, Berti F, Schneider J, Vojtek I. Polysaccharide conjugate vaccine protein carriers as a “neglected valency”—potential and limitations. *Vaccine*. 2017;35(25):3286-3294.
52. Martin LB, Simon R, MacLennan CA, et al. Status of paratyphoid fever vaccine research and development. *Vaccine*. 2016;34(26):2900-2902.
53. Mani S, Wierzba T, Walker RI. Status of vaccine research and development for Shigella. *Vaccine*. 2016;34(26):2887-2894.
54. Rappuoli R. New and improved vaccines against diphtheria and tetanus. In: Levine MM, et al., eds. *New Generation Vaccines*. New York, NY: Marcel Dekker, Inc.; 1997:417-435.
55. Metz B, Kersten GFA, Hoogerhout P, et al. Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. *J Biol Chem*. 2004;279(8):6235-6243.
56. Metz B, Kersten GF, Baart GJ, et al. Identification of formaldehyde-induced modifications in proteins: reactions with insulin. *Bioconjug Chem*. 2006;17(3):815-822.
57. Van Ness BG, Howard JB, Bodley JW. 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(22):10717-10720.
58. Papini E, Schiavo G, Sandona D, Rappuoli R, Montecucco C. Histidine 21 is at the NAD+ binding site of diphtheria toxin. *J Biol Chem*. 1989;264(21):12385-12388.
59. Blanke SR, Huang K, Collier RJ. Active-site mutations of diphtheria toxin: role of tyrosine-65 in NAD binding and ADP-ribosylation. *Biochemistry*. 1994;33(51):15494-15500.
60. Shen WH, Choe S, Eisenberg D, Collier RJ. Participation of lysine 516 and phenylalanine 530 of diphtheria toxin in receptor recognition. *J Biol Chem*. 1994;269(46):29077-29084.
61. Naglich JG, Metherall JE, Russell DW, Eidels L. Expression cloning of a diphtheria toxin receptor: identity with a heparin-binding EGF-like growth factor precursor. *Cell*. 1992;69(6):1051-1061.
62. Louie GV, Yang W, Bowman ME, Choe S. Crystal structure of the complex of diphtheria toxin with an extracellular fragment of its receptor. *Mol Cell*. 1997;1(1):67-78.
63. Diethelm-Okita BM, Okita DK, Banaszak L, Conti-Fine BM. Universal epitopes for human CD4+ cells on tetanus and diphtheria toxins. *J Infect Dis*. 2000;181(3):1001-1009.
64. Sambrook J, Fritsch EF, Maniatis T. *Molecular Cloning*. 2nd ed. New York, NY: Cold Spring Harbor Laboratory Press; 1989.
65. Habeeb AFSA. Determination of free amino groups in proteins by trinitrobenzenesulfonic acid. *Anal Biochem*. 1966;14:328-336.
66. Meiring HD, van der Heeft E, ten Hove GJ, de Jong A. Nanoscale LC-MS$^{(n)}$: technical design and applications to peptide and protein analysis. *J Sep Sci*. 2002;25:557-568.
67. Bell CE, Eisenberg D. Crystal structure of nucleotide-free diphtheria toxin. *Biochemistry*. 1997;36(3):481-488.
68. Pytms. 2015. Available at: https://pymolwiki.org/index.php/Pytms. Accessed August 5, 2019.
69. Carroll SF, Barbieri JT, Collier RJ. Diphtheria toxin: purification and properties. *Methods Enzymol*. 1988;165:68-76.
70. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics*. 2006;5(1):144-156.
71. Jentoft N, Dearborn DG. Protein labeling by reductive alkylation. *Methods Enzymol*. 1983;91:570-579.
72. Braun KP, Cody RB, Jones DR, Peterson CM. A structural assignment for a stable acetaldehyde-lysine adduct. *J Biol Chem*. 1995;270(19):11263-11266.
73. Fowles LF, Beck E, Worrall S, Shanley BC, de Jersey J. The formation and stability of imidazolidinone adducts from acetaldehyde and model peptides. A kinetic study with implications for protein modification in alcohol abuse. *Biochem Pharmacol*. 1996;51(10):1259-1267.
74. Eschweiler W. Ersatz von an Stickstoff gebundenen Wasserstoffatomen durch die Methylgruppe mit Hülfe von Formaldehyd. *Chem Ber*. 1905;38:880-882.