Bioconjugate Chemistry

Article

pubs.acs.org/bc

Site-Specific Derivatization of Avidin Using Microbial Transglutaminase

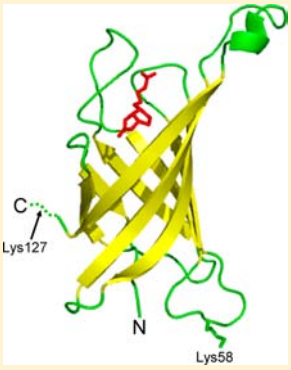
Barbara Spolaore,*[†,‡] Nunzio Damiano,[†] Samanta Raboni,[†] and Angelo Fontana[†]

[†]CRIBI Biotechnology Centre, University of Padua, Viale G. Colombo 3, 35121 Padua, Italy

[‡]Department of Pharmaceutical and Pharmacological Sciences, University of Padua, Via F. Marzolo 5, 35131 Padua, Italy

Supporting Information

ABSTRACT: Avidin conjugates have several important applications in biotechnology and medicine. In this work, we investigated the possibility to produce site-specific derivatives of avidin using microbial transglutaminase (TGase). TGase allows the modification of proteins at the level of Gln or Lys residues using as substrate an alkyl-amine or a Gln-mimicking moiety, respectively. The reaction is site-specific, since Gln and Lys derivatization occurs preferentially at residues embedded in flexible regions of protein substrates. An analysis of the X-ray structure of avidin allowed us to predict Gln126 and Lys127 as potential sites of TGase's attack, because these residues are located in the flexible/unfolded C-terminal region of the protein. Surprisingly, incubation of avidin with TGase in the presence of alkylamine containing substrates (dansylcadaverine, 5-hydroxytryptamine) revealed a very low level of derivatization of the Gln126 residue. Analysis of the TGase reaction on synthetic peptide analogues of the C-terminal portion of avidin indicated that the lack of reactivity of Gln126 was likely due to the fact that this residue is proximal to negatively charged carboxylate groups, thus hampering the interaction of the substrate at the negatively charged active site of TGase. On the other hand, incubation of avidin with TGase in the presence of carbobenzoxy-L-glutaminyl-glycine in order to derivatize Lys residue(s) resulted in a clean and high yield production of an avidin derivative, retaining the biotin binding properties and the quaternary structure of the native protein. Proteolytic digestion of the modified protein, followed by mass spectrometry, allowed us to identify Lys127 as the major site of reaction, together with a minor modification of Lys58. By using TGase, avidin was also conjugated via a Lys-Gln isopeptide bond to a protein containing a single reactive Gln residue, namely, Gln126 of granulocyte-macrophage colony-stimulating factor. TGase can thus be exploited for the site-specific derivatization of avidin with small molecules or proteins.



INTRODUCTION

Avidin is a homotetrameric glycoprotein of 128 amino acids that binds biotin with the highest affinity known in nature for a protein−ligand interaction (K_d ~10^{-15} M).[1,2] The X-ray structures of both avidin depleted of biotin (apo-avidin)[3−5] and the protein−biotin complex (holo-avidin)[3,6] are known at high resolution. The 3D crystal structure of the avidin monomer is characterized by the presence of eight antiparallel successive β-strands and the overall shape is that of a β-barrel (Figure 1). The biotin binding site is located at one end of the avidin barrel (Figure 1B). The tight interaction between avidin and biotin has been exploited for several applications in life sciences, since biotin can be easily coupled to molecules without compromising either its interaction with avidin or the function of the partner molecule.[7] Besides affinity chromatography,[8] the avidin−biotin interaction can be effectively used for protein detection, for example, with biotinylated antibodies being detected with avidin conjugated to either fluorescent or enzymatic reagents.[9] Avidin has also potential clinical applications in pretargeted radioimmunotherapy[10] and as a delivery system of drugs, radionuclides, or fluorophores for therapeutic or diagnostic purposes.[11−13]

The conjugation of avidin to small molecules or proteins is usually achieved by chemical methods and mostly by derivatization of its amino groups.[14,15] Since avidin contains nine Lys residues, their conjugation can yield heterogeneous products modified at different positions, leading in some cases to a decreased biotin binding.[16] Avidin-fusion proteins have also been produced by recombinant methods, as, for example, IgG-avidin fusion proteins that enable the delivery to brain of biotinylated drugs.[13]

In this paper, we studied the reactivity of avidin toward transglutaminase (TGase) with the aim to develop an enzymatic method for the site-specific modification of this protein. Microbial transglutaminase is an enzyme extracted from *Streptomyces mobaraensis*, easily available from a commercial source and well characterized in structural terms.[17] TGase catalyzes the modification of proteins at the level of glutamine (Gln) or lysine (Lys) residues. In the first case, the reaction occurs between the γ-amido group of a protein-bound Gln residue (−$CONH_2$, the acceptor) and an amino group (−NH_2, the donor) of an alkyl-amine with the formation of an isopeptide bond (Figure 2).[18,19] The

Received: August 13, 2013
Revised: January 21, 2014
Published: February 11, 2014

dx.doi.org/10.1021/bc400378h | *Bioconjugate Chem.* 2014, 25, 470−480

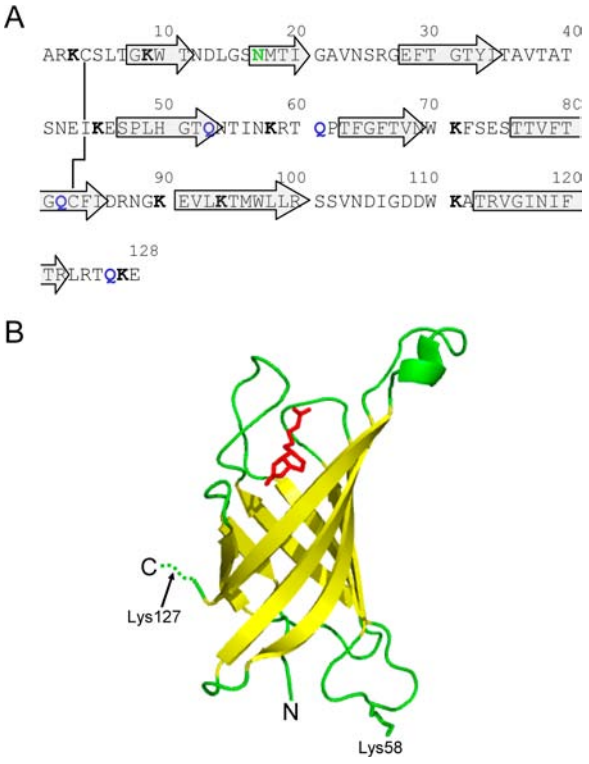


Figure 1. Amino acid sequence and 3D structure of chicken avidin. (A) Amino acid sequence of chicken avidin (Swiss-Prot entry P02701). Residue potential sites of the TGase-mediated derivatization are shown in bold black for Lys residues and in bold blue for Gln residues, while the glycosylated residue Asn17 is indicated in green. The location of the disulfide bond Cys4-Cys83 is shown by a solid line. The eight β-strands along the polypeptide chain are indicated by gray arrows. (B) 3D structure of avidin (PDB file 2AVI).[3] The biotin molecule is shown in red, the localization of residue Lys58 is shown by the lateral chain displayed in a stick style, while Lys127 is indicated by an arrow at the C-terminus of the protein since the region Arg124–Glu128 is not resolved in the X-ray structure. The structure is drawn by using PyMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC) and the polypeptide chain is colored according to the secondary structure with β-strands in yellow and connecting regions in green.

modification of Lys residues requires a ligand derivatized with a Gln-residue or a Gln analogue.[20,21]

Microbial TGase has been recently regarded as one of the promising enzymes for the site-specific protein derivatization. Indeed, different chemical moieties, ranging from poly(ethylene glycol) (PEG) polymers,[22−24] fluorescent probes to radionuclide chelating agents have been linked to proteins by using TGase.[25,26] The TGase-mediated reaction has also been exploited as a tool to affect the biological properties of specific proteins, leading to products with potentially useful applications in life science research and in biotechnology.[27,28] The main advantage of the TGase-mediated modification of proteins with respect to chemical conjugation is the possibility to achieve site-specific protein modification at the level of one or a few Gln and Lys residues of the protein substrate.[29] We recently demonstrated that reactive Gln and Lys residues are located in flexible regions of the protein substrate, indicating that the specificity of TGase is primarily determined by the conformational features of the substrate and to a lesser extent by the sequences near the Gln and Lys residues.[23,30]

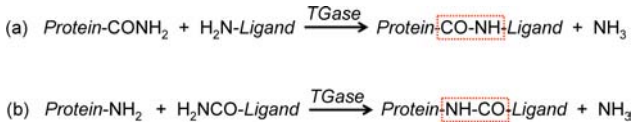


Figure 2. TGase-mediated modification of proteins. Two approaches of protein derivatization by TGase are feasible. In the first approach (a), a Lys-mimicking ligand (H_2N-*Ligand*) can be attached to a Gln-residue (*Protein*-$CONH_2$), while in the second (b), a ligand derivatized with Gln-residue (H_2NCO-*Ligand*) can be coupled to a Lys-residue of the protein (*Protein*-NH_2). The ligand can be a fluorescent probe, biotin, a drug entity, or a chelating agent.

An analysis of the X-ray structures of apo-avidin allowed us to predict two potential sites of derivatization in avidin, namely, Gln126 and Lys127 residues, located at the flexible C-terminus of the protein.[3−5] In this work, we experimentally tested the reactivity of these residues to TGase. Unexpectedly, Gln126 of avidin was derivatized at a negligible yield. Using synthetic peptide analogues covering the C-terminal sequence of avidin, we demonstrated that negatively charged carboxylate groups near Gln126 affect its derivatization. On the other hand, a quite selective and high yield TGase-mediated Lys-derivatization of avidin was achieved. We characterized this avidin bioconjugate in terms of location of the sites of modification, protein quaternary structure, and biotin binding. By using TGase, avidin was also conjugated via a Lys-Gln isopeptide bond to a protein containing a single reactive Gln residue, namely, granulocyte-macrophage colony-stimulating factor. These results indicate that TGase-mediated protein modification can be a useful tool to produce a variety of site-specific avidin bioconjugates for potential applications in modern bioscience and biotechnology.

EXPERIMENTAL PROCEDURES

Materials. Avidin was purchased from Belovo (Bastogne, Belgium) while transglutaminase (TGase) isolated from *Streptoverticillium mobaraense* was TGase ACTIVA MP from Ajinomoto Co. (Tokyo, Japan). TGase stock solutions were prepared by dissolving the powder of TGase ACTIVA MP in 0.1 M sodium phosphate buffer, pH 7.0. The concentration of the enzyme in the stock solutions was measured from the absorbance at 280 nm and typically, solutions of TGase of 0.5–1.0 mg/mL were used. The measured concentration of the enzyme was used to calculate the volume of stock solution to obtain a certain enzyme/substrate (E/S) ratio. Carbobenzoxy-L-glutaminyl-glycine (ZQG), dansylcadaverine (DC), and 5-hydroxytryptamine (HTA) were purchased from Sigma-Aldrich (Milwaukee, WI, USA). Stock solutions of these ligands were prepared by dissolving them in DMSO at concentrations of 0.06 M (DC), 0.1 M (ZQG), and 0.094 M (HTA), respectively. Porcine trypsin was from Promega (Madison, WI, USA), while endoproteinase Lys-C was from Sigma-Aldrich. All other chemicals were of analytical reagent grade and were obtained from Sigma-Aldrich. The concentrations of TGase, of avidin, and of granulocyte-macrophage colony-stimulating factor (GM-CSF) were determined from their absorbance at 280 nm using the protein's extinction coefficients generated by ProtParam (http://www.expasy.org/tools/protparam.html).

TGase-Mediated Conjugation of Avidin to DC and ZQG. Avidin was dissolved in 0.1 M phosphate buffer, pH 7.0, at a concentration of 0.8 mg/mL, and it was mixed with the stock solution of DC or HTA at a molar ratio of 1:30. TGase

dx.doi.org/10.1021/bc400378h | *Bioconjugate Chem.* 2014, 25, 470−480

was added to the reaction mixtures at E/S ratios of 1/25 or 1/5 by weight, and the reactions were allowed to proceed at 37 °C. Aliquots were collected after 0, 5, and 24 h and the reactions were stopped by addition of an equal volume of an aqueous solution of 1% trifluoroacetic acid (TFA). For the derivatization with ZQG, a solution of avidin in 0.1 M phosphate buffer, pH 7.0, was mixed with the solution of ZQG at a molar ratio of 1:40. The reaction mixture was incubated at 37 °C with TGase at an E/S ratio of 1/60 by weight, and aliquots were stopped after 30 min and 5 h by addition 100 μM iodoacetamide (IAA) in 0.1 M phosphate buffer pH 7.0 at a 20 molar excess with respect to TGase. Aliquots of the reactions were analyzed by RP-HPLC using an Agilent series 1100 HPLC with an online UV-detector from Agilent Technologies (Waldbroon, Germany). RP-HPLC analyses were performed on a C18 Phenomenex column (150 × 4.6 mm, 5 μm) applying a gradient of acetonitrile (AcCN), 0.05% TFA, and water, 0.05% TFA from 0% to 80% of AcCN in 20 min and from 80% to 100% in 1 min at a flow rate of 1.0 mL/min. The effluent from the column was monitored by measuring the absorbance at 220 nm. Fractions collected from the RP-HPLC analyses were lyophilized and analyzed by MS.

Proteolytic Digestion of Avidin Derivatives. For the trypsin digestion of non-reduced avidin, protein samples were dissolved in 0.1 M NH_4HCO_3, pH 8.0, at a concentration of about 0.1 mg/mL and incubated overnight at 37 °C at an E/S ratio with trypsin of 1/100 by weight. MS analyses were performed on the lyophilized digests dissolved in a solution of 0.1% formic acid in AcCN:water (1:1). For trypsin or endoproteinase Lys-C digestion of reduced and carbamydomethylated avidin, avidin samples were dissolved in 50 mM Tris-HCl, pH 8.5, 8 M urea, 5 mM TCEP at a concentration of 2.5 mg/mL and incubated for 30 min at 37 °C. Iodacetamide was added at a final concentration of 25 mM and the solution was further incubated in the dark for 20 min at RT. The avidin samples were then diluted with 50 mM NH_4HCO_3, pH 8.0, to obtain a final urea concentration of 2 M. Endoproteinase Lys-C or trypsin were added at an E/S 1/100 and the digestion was allowed to proceed at 37 °C overnight. The proteolysis reactions were stopped by lowering the pH with an aqueous solution of 3% formic acid and desalted with Sep-Pak C18 Cartridges (Waters) prior of MS or tandem MS (MS/MS) analysis of the avidin peptides.

Mass Spectrometry Analyses. Mass spectrometry-based analyses were performed with a Micromass mass spectrometer Q-Tof Micro (Manchester, UK) equipped with an electrospray source (ESI-MS). Samples were dissolved in 0.1% formic acid in AcCN/water (1:1) and analyzed in MS and MS/MS mode. The measurements were performed at a capillary voltage of 3 kV and at a cone and extractor voltages of 35 and 1 V, respectively (positive ion mode). MS/MS analyses of the modified peptides were conducted on the Q-Tof Micro mass spectrometer at variable collision energy values and using argon as collision gas. External calibration was performed using a solution of 0.1% (v/v) phosphoric acid in 50% (v/v) aqueous acetonitrile. Instrument control, data acquisition, and processing were achieved with MassLynx v 4.1 software (Waters).

TGase-Mediated Conjugation of Avidin to GM-CSF. A solution of avidin in 0.1 M phosphate buffer, pH 7.0, was mixed with a solution of GM-CSF in the same buffer at a molar ratio of 1:2 (avidin/GM-CSF). After addition of TGase at an E/S ratio of 1/20 by weight with respect to GM-CSF, the reaction mixture was incubated at 37 °C. Aliquots were removed from the reaction mixture after 0, 1, and 5 h and diluted with an equal volume of an aqueous solution of 1% TFA. After lyophilization, aliquots of 5 μg of total protein content were analyzed by SDS-PAGE using a gel with 15% acrylamide concentration. Gels were stained with Coomassie Brilliant Blue R-250. Coomassie-stained protein bands were excised and in-gel digested, as previously described.[31] Protein digests were resuspended in 0.1% formic acid and analyzed by LC-MS/MS. LC-MS/MS analyses were performed on a Micromass CapLC unit (Waters) interfaced to a Micromass Q-Tof Micro mass spectrometer (Waters) equipped with a nanospray source. Tryptic digests were loaded at a flow rate of 20 μL/min onto an Atlantis dC18 Trap Column. After valve switching, the sample was separated on an Atlantis C_{18} column (150 × 0.075 mm, 3.5 μm particle size) (Waters) at a flow rate of 4.8 μL/min using a gradient from 5% B to 55% B in 37 min and from 55% to 85% B% in 5 min (solvent A: 95% H_2O, 5% AcCN, 0.1% formic acid; solvent B: 5% H_2O, 95% AcCN, 0.1% formic acid). Instrument control, data acquisition, and processing were achieved with MassLynx v 4.1 software (Waters). MS/MS data were analyzed by the online MASCOT software (Matrix Science, http://www.matrixscience.com).

Peptide Synthesis. The peptides were synthesized by a solid phase method on a preloaded Wang resin functionalized with the acid-labile *p*-benzyloxybenzyl alcohol resin linker (Novabiochem), using an automatized peptide synthesizer Model Syro II (Multisyntech). The fluoren-9-yl-methoxycarbonyl (Fmoc) strategy was used throughout the peptide chain assembly, with *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HATU) and *N*,*N*-ethyl-diisopropylamine as coupling reagents for N-Fmoc amino acids carrying standard side chain protective groups. Cleavage of the peptides was performed by reacting the peptidyl resins with a mixture containing TFA/H_2O/thioanisole/ethanedithiol/phenol (10 mL/0.5 mL/0.5 mL/0.25 mL/750 mg) for 2.5 h. The peptide amidated at the C-terminus was synthesized by a solid phase method on a Rink Amide resin functionalized with 4-(2′,4′-dimethoxyphenyl-Fmoc-aminomethyl)-phenoxy resin linker (Novabiochem). The functionalization of the used resin is 0.59 mmol/g.

For the methyl ester derivatization of the peptides, a solution of methanolic HCl was prepared by the dropwise addition of acetyl chloride to methanol at a volume ratio of 1/7 (v/v). The addition of acetyl chloride was performed keeping methanol under stirring and cooled on an ice bath. The peptide was solubilized in this freshly prepared solution at a concentration of 1 mg/mL. The sample was incubated at room temperature for 1 h and the excess reagent removed by drying in a Savant SpeedVac vacuum centrifuge. Peptides were purified by RP-HPLC using a semipreparative Vydac C18 (10 × 250 mm, 10 μm) using a gradient from 5% to 15% of AcCN in 3 min and from 15% to 30% in 12 min at a flow rate of 2 mL/min. The solvents used were AcCN with 0.1% TFA and water with 0.085% TFA. The identity of the purified peptides was confirmed by ESI-MS analysis.

TGase-Mediated Conjugation of the C-Terminal Avidin Peptides to HTA and ZQG. Derivatization of peptides (0.4−0.5 mg/mL) with TGase was performed in 0.1 M phosphate buffer, pH 7.0, in the presence of ZQG or HTA at a 1/30 molar ratio and upon addition of TGase at an E/S of 1/25, by weight. Reaction mixtures were incubated at 37 °C for 1 and 4 h (ZQG) or 4 h only (HTA) and stopped upon addition of a 100 μM IAA solution in 0.1 M phosphate buffer, pH 7.0.

dx.doi.org/10.1021/bc400378h | *Bioconjugate Chem.* 2014, 25, 470−480

Analysis of the Aggregation State of Avidin Modified with ZQG. The quaternary state of avidin and of its ZQG-conjugated form was determined by fast-flow liquid chromatography performed on a Superdex 75 HR 10/30 (Pharmacia) by using an AKTA FPLC instrument (Amersham Biosciences, Sweden). Samples (45 μg) of avidin and of avidin^{1,2ZQG} with or without a molar excess of biotin were loaded on the gel filtration column and chromatography was performed at a flow rate of 0.5 mL/min in 10 mM phosphate buffer, 0.65 M NaCl, pH 7.2. The column was calibrated using bovine serum albumin, ovalbumin, carbonic anhydrase, and α-lactalbumin as molecular weight markers. Heat stability experiments were performed according to Bayer et al.[32] Briefly, aliquots of 5 μg of avidin and of avidin^{1,2ZQG} were incubated with or without a molar excess of biotin and diluted in sample buffer. The protein solutions were incubated at 100 °C for 20 min before being analyzed by SDS-PAGE using a gel with 15% acrylamide concentration. The gel was stained with Coomassie Brilliant Blue R-250. The avidin^{1,2ZQG} sample used for the gel filtration analysis and the heat stability experiments was an avidin-ZQG reaction mixture after 5 h of incubation.

Biotin-Binding Assays. The HABA (4-hydroxyazobenzene-2-carboxylic acid) colorimetric assay for avidin was performed according to Green et al.[33] The colorimetric assay for avidin was performed on a reaction mixture containing avidin^{1,2ZQG} (after 5 h of incubation) and on a sample of avidin in 0.1 M phosphate buffer, pH 7.0, with ZQG and IAA added at the same concentration than in the reaction mixture.

RESULTS

Site-Specific Derivatization of Avidin Mediated by TGase. Avidin purified from egg yolk is glycosylated and analysis of the carbohydrate moiety indicated an extensive glycan microheterogeneity.[34] Our ESI-MS analyses of the avidin sample used in the TGase derivatization reactions confirmed the heterogeneity, since several different molecular masses could be measured (Figure 3A). Nevertheless, we took as reference the three main isoforms of avidin (isoforms A–C), whose mass values can be assigned to a specific glycan composition in mannose and N-acetylglucosamine (Table 1).[34] No m/z signals of deglycosylated avidin were detected.

We tested the reactivity of the four Gln residues of avidin toward TGase using dansyl cadaverine (DC) and 5-hydroxytryptamine (HTA) as model substrates, since they both contain a primary amine that is a good amino donor (Figure S1). Avidin was incubated at 37 °C in the presence of 30 molar excess of DC or HTA with TGase at an E/S of 1/25 or 1/5, by weight. Aliquots of the reaction mixtures were taken after 5 and 24 h of incubation and analyzed by RP-HPLC. The RP-HPLC chromatographic profiles of native avidin and of avidin after incubation in the presence of TGase and DC or HTA showed no substantial change even after 24 h of incubation at an E/S of 1/5 (Figure S2). Indeed, in the chromatograms of the reaction mixtures, avidin eluted as a main peak with no change in retention time and in the absorbance intensity with respect to the RP-HPLC analysis of avidin after 0 h of incubation. ESI-MS analysis of avidin eluted in the RP-HPLC peaks showed identical spectra for the protein incubated with DC or HTA, with measured masses that corresponded to the different glycosylated isoforms of native avidin (Figure S2, right panels). The only difference is the relative intensities of the different glycosylated avidin species that changed upon incubation.

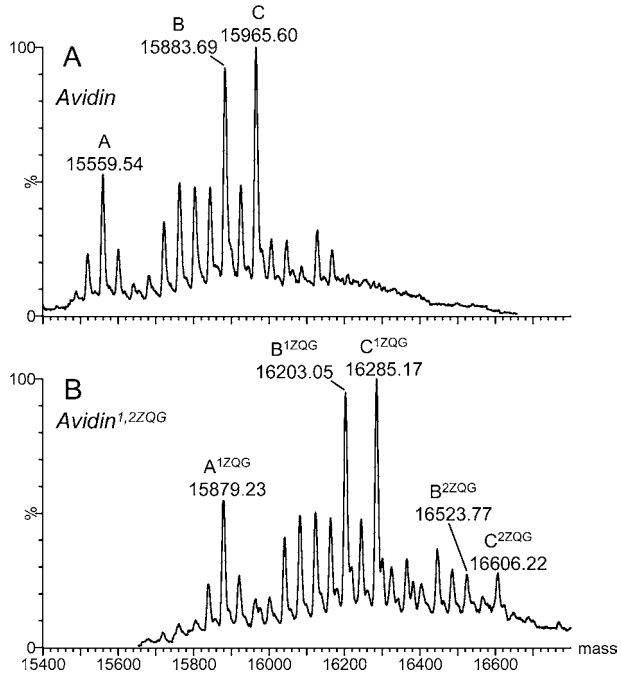


Figure 3. ESI-MS analyses of avidin and avidin conjugated to ZQG. (A) Deconvoluted ESI mass spectrum of the sample of chicken avidin used in the present study. The main components of the mass spectrum are indicated with the letters A–C (see Table 1). (B) Deconvoluted ESI mass spectrum of the ZQG-derivatives of avidin. The products of the reaction with TGase after 5 h of incubation were purified by RP-HPLC (Figure 4B) and they correspond to the protein material eluted at a retention time of 16.00 min. The main components of the mass spectrum are indicated with the letters A–C with the superscript 1ZQG to indicate the mono derivative and 2ZQG for the double derivative (see Table 1).

Table 1. Molecular Masses of the Main Avidin Glycosylated Isoforms and of Their Derivatives Conjugated to ZQG

avidin species (carbohydrate composition)	number of ZQG	molecular mass (Da) found[a]	molecular mass (Da) calculated[b]
Avidin	——	——	14341.19
A[c] ($Man_5GlcNAc_2$[d])	——	15559.54 ± 0.10	15558.29
	1ZQG	15879.23 ± 0.52	15878.59
	2ZQG	——[e]	16198.89
B ($Man_7GlcNAc_2$[d])	——	15883.69 ± 0.23	15882.58
	1ZQG	16203.05 ± 0.42	16202.88
	2ZQG	16523.77 ± 0.43	16523.18
C ($Man_5GlcNAc_4$[d])	——	15965.60 ± 0.33	15964.68
	1ZQG	16285.17 ± 0.35	16284.98
	2ZQG	16606.22 ± 0.34	16605.28

[a]Experimental molecular masses determined by ESI-MS. [b]Calculated average molecular masses. [c]The different glycosylated isoforms are labeled according to the ESI-MS analysis reported in Figure 3. [d]Carbohydrate composition of the N-glycosylated avidin isoforms. Man is mannose, while GlcNAc stands for N-acetylglucosamine. [e]The charge state series of isoform A^{2ZQG} and of isoform B^{1ZQG} cannot be distinguished in the mass spectrum at the resolution of the Q-Tof Micro mass spectrometer.

Mass spectra of avidin are complicated by glycosylation that could mask a low level of conjugation (Figure S2), in particular, in the case of derivatization with HTA, for which the molecular masses of the avidin conjugated isoforms are close to values

dx.doi.org/10.1021/bc400378h | Bioconjugate Chem. 2014, 25, 470–480

measured in the mass spectrum of native avidin. In order to unravel a low level of derivatization, avidin samples collected from the RP-HPLC analyses of the protein incubated with DC or HTA and TGase (E/S 1/5) for 24 h were digested with the endoproteinase Lys-C. This protease was used instead of trypsin because based on the avidin sequence, it allows us to produce Gln-containing peptides of ideal size to be analyzed by MS. The digests were analyzed by LC-MS/MS obtaining a 66−69% sequence coverage of the protein and mapping all the four Gln-containing peptides (Table S1). In the LC-MS/MS analysis of the digest of avidin incubated in the presence of HTA, we detected m/z signals of peptide 112−128 derivatized with HTA at the level of Gln126 (Figure S3, Table S1). However, the signal intensity of the conjugated peptide was 0.3% in respect to the non-modified peptide, indicating a negligible yield of derivatization (Figure S3). No modified peptide was instead detected in the LC-MS/MS analysis of the digest of avidin incubated in the presence of DC.

In order to identify Lys residues in the avidin sequence that can be modified by TGase, the protein was incubated at 37 °C with TGase at an E/S ratio of 1/60 by weight in the presence of carbobenzoxy-L-glutaminylglycine (ZQG) at a 40 molar excess with respect to avidin. ZQG is a Gln-containing substrate efficiently conjugated to Lys residues by the TGase-catalyzed reaction (Figure S1). The reaction mixture was analyzed by RP-HPLC after different times of incubation (Figure 4). A new RP-HPLC peak with a slightly higher retention time (RT, 16.00 min) appeared after 30 min of incubation and became prominent after 5 h of reaction. ESI-MS analysis identified this species as the ZQG-derivative of the A−C isoforms of avidin (Figure 3B, Table 1). Traces of the ZQG-biderivative were also present. After 5 h of reaction, the reaction yield of the avidin^{1,2ZQG} derivative is of 92% as estimated on the basis of the areas of the chromatographic peaks of the derivative and of unreacted avidin.

In order to identify which Lys residues are modified by TGase, native avidin and avidin^{1,2ZQG} formed after 5 h of reaction and purified by RP-HPLC were digested with trypsin. ESI-MS analysis of the two tryptic digests allowed to cover the 76.6% of the avidin sequence and in particular six of the nine Lys residues (Table 2). Analysis of the digests did not map Lys3, Lys 90, and Lys94 because trypsin hydrolysis at the level of these residues leads to very short peptides (Figure 1A). Nevertheless, derivatization with ZQG of these Lys residues inhibits trypsin hydrolysis and generates longer and more hydrophobic peptides, whose m/z signals were not observed in the digest of avidin^{1,2ZQG}. Three Lys containing peptides were instead found conjugated to ZQG in the avidin^{1,2ZQG} sample. Indeed, the m/z signals of peptides 46−59, 123−128, and 125−128 showed a mass increment corresponding to the addition of ZQG and they all contained one Lys residue (Table 2). MS/MS spectra acquired on the m/z signals of the modified peptides 46−59 and 125−128 confirmed that the ZQG-conjugated Lys residues are located at position 58 and 127, respectively (Figure S4 and S5). In the ESI-MS spectrum acquired on the digest of avidin^{1,2ZQG}, the unmodified peptide 46−59 still showed an intense signal, while m/z signals of peptides 123−128 and 125−128 without derivatization were not detected (Table 2). This result indicates that Lys127 is the preferential site of modification of avidin.

We also exploited the TGase-mediated reaction to conjugate avidin to a protein containing a Gln-reactive residue. In particular, we chose to conjugate avidin to granulocyte-macrophage colony-stimulating factor (GM-CSF), since previous experiments conducted in our laboratory demonstrated that only Gln126 in GM-CSF is reactive toward TGase-mediated conjugation (unpublished results, Figure S6). Indeed, Gln126 is located at the level of the disordered C-terminal end of the protein.[35] In addition, since GM-CSF has no reactive Lys residues, the dimer that could complicate the reaction pattern does not form in the presence of TGase. Analysis by SDS-PAGE of the TGase-catalyzed reaction between avidin and GM-CSF after 1 and 5 h of incubation (Figure 5, lanes 5 and 6, respectively) indicated the formation of only one product that migrates at a molecular weight consistent with the conjugation of avidin to GM-CSF. In-gel digestion of this protein band followed by LC-MS/MS analysis of the peptide mixture and MASCOT search of the MS/MS data allowed to identify the presence of both avidin and GM-CSF (Table S2), thus confirming the formation of the conjugation product. In order to determine the sites of cross-linking between avidin and GM-CSF, LC-MS/MS data were manually inspected and m/z signals corresponding to the formation of isopeptide bonds between peptides 112−127 of GM-CSF and 46−59 of avidin

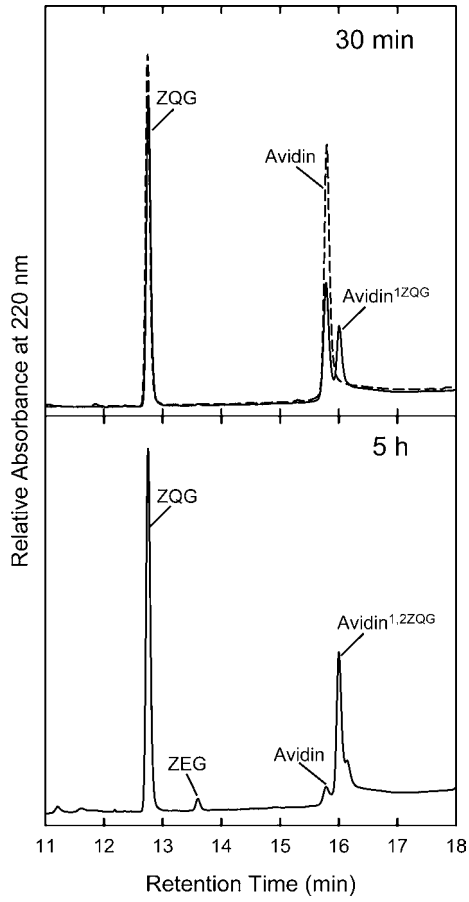


Figure 4. RP-HPLC analyses of the TGase-conjugated derivatives of avidin. Avidin was incubated with TGase at an E/S of 1/60 (by weight) in the presence of ZQG (40 molar excess in respect to avidin) and aliquots of the reaction mixture were taken after 30 min and 5 h of incubation. In the chromatograms, a dashed line indicates the chromatographic profile of native avidin and straight lines the profiles of the ZQG-conjugated species formed after 30 min and 5 h of reaction. ZEG (carbobenzoxy-L-glutamyl-glycine) indicates the deamidated form of ZQG, that is produced by the TGase-catalyzed hydrolysis of the Gln residue to glutamic acid (measured mass for ZEG is 337.14 Da; theoretical mass 337.13 Da).

dx.doi.org/10.1021/bc400378h | Bioconjugate Chem. 2014, 25, 470−480

Table 2. Molecular Masses Measured for the Tryptic Peptides of Avidin Detected in the ESI Mass Spectra of the Digests of Avidin and Avidin[1,2ZQG]

Tryptic peptide	Peptide sequence	N° of ZQG	molecular mass (Da)		
			found in avidin[a]	found in avidin[1,2ZQG a]	calculated[b]
27–45	GEFTGTYITAVTATSNEI**K**	–	2002.00	2001.91	2001.99
46–59	ESPLHGTQNTIN**K**$_{58}$R	–	1593.76	1593.72	1593.82
		1ZQG	–	1913.82	1913.92
60–71	TQPTFGFTVNW**K**	–	1424.70	1424.69	1424.71
4–9/ 72–87[c]	C**K** / FSESTTVFTGQCFIDR	–	2442.09	2442.02	2442.12
95–100	TMWLLR	–	818.41	818.42	818.45
101–111	SSVNDIGDDW**K**	–	1234.55	1234.52	1234.55
115–122	VGINIFTR	–	918.52	918.51	918.53
123–128	LRTQ**K**$_{127}$E	–	773.43	–	773.44
		1ZQG	–	1093.53	1093.54
125–128	TQ**K**$_{127}$E	–	–[d]	–	504.25
		1ZQG	–	824.29	824.36

[a]Experimental molecular masses determined by ESI-MS analysis of the tryptic digest of avidin or avidin[1,2ZQG]. The reported mass values were measured on the tryptic digest of non-reduced protein samples. [b]Calculated monoisotopic molecular masses. [c]Tryptic peptides 4–9 and 72–87 are linked by the disulfide bridge between Cys4 and Cys83. [d]Peptide 125–128 was not detected in the ESI mass spectrum of the tryptic digest of avidin.

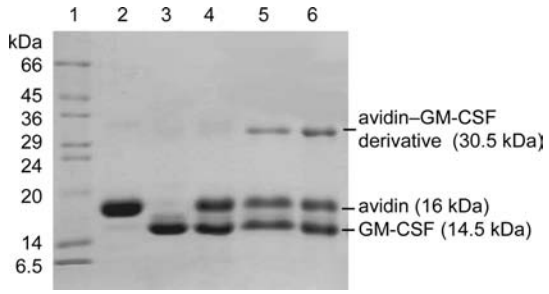


Figure 5. TGase-mediated conjugation of avidin to GM-CSF. SDS-PAGE analysis of the conjugation reaction between avidin and GM-CSF catalyzed by TGase. Protein samples were loaded in the following order: MW standards (lane 1); avidin (lane 2); GM-CSF (lane 3); reaction mixture containing avidin, GM-CSF and TGase after 0, 1, and 5 h of incubation (lane 4, 5, and 6, respectively).

and peptides 112–127 of GM-CSF and 125–128 of avidin (Table 3) were found. Analysis of the MS/MS fragmentation

Table 3. Molecular Masses Measured for Peptide 112–127 of GM-CSF Conjugated to Peptides 46–59 or 125–128 of Avidin (AVI)

species	molecular mass (Da)	
	found[a]	calculated[b]
112–127$^{GM\text{-}CSF}$/46–59^{AVI}	3582.99	3582.74
112–127$^{GM\text{-}CSF}$/125–128^{AVI}	2493.33	2493.17

[a]Experimental molecular masses determined by LC-MS/MS analysis of the tryptic digest of the avidin/GM-CSF conjugation product. [b]Calculated monoisotopic molecular masses.

pattern of these two cross-linked species (Figures 6) confirmed that cross-linking occurs between residues Gln126 of GM-CSF and either Lys58 or Lys127 of avidin, demonstrating that these two residues are both primary sites of TGase derivatization. In the LC-MS/MS analysis the *m/z* signals of peptide 112–127$^{GM\text{-}CSF}$/46–59^{AVI} have lower intensity than those of peptide 112–127$^{GM\text{-}CSF}$/125–128^{AVI}, an indication that derivatization of avidin occurs preferentially at Lys127.

Effect of the Amino Acid Sequence on the Reactivity of Gln126. In avidin, Gln126 is located N-terminally to the TGase-reactive Lys127. However, TGase derivatizes this Gln residue with a very low yield even if the C-terminal region of avidin is highly flexible and not resolved in the X-ray structure.[5] In order to study the reactivity of this Gln residue without the interference of the protein quaternary structure, a set of peptides corresponding to the region 120–128 was synthesized. First, we substituted Phe120 with Tyr, in order to better quantify the concentration of the peptides by UV absorbance spectroscopy. The sequence of the peptide was: H-Tyr-Thr-Arg-Leu-Arg-Thr-Gln-Lys-Glu-OH (120–128K). Since Gln residues flanked at the C-terminus by positively charged Lys residues were reported not to be substrate of tissue TGase,[36,37] we also synthesized the peptide carrying an Ala residue in position 127 instead of a Lys (120–128A) to verify the effect of Lys proximity on Gln126 reactivity. In addition, acidic amino acid residues at the C-terminus of Gln residues were reported to slow down the kinetic of the modification because of unfavorable interactions with the negatively charged active site of TGase.[17,38] Thus, we investigated the effect of the negatively charged Glu128 and the C-terminus of the protein on the reactivity of Gln126 by esterification with a methyl group of the −COOH groups of the previous peptide sequences (120–128K^{2Me} and 120–128A^{2Me}) and by synthesis of the peptide 120–128K amidated at the C-terminus (120–128K-NH_2).

Peptides 120–128K and 120–128A as well as their esterified forms and peptide 120–128K-NH_2 were subjected to derivatization at the level of Gln126 using HTA as model substrate. DC could not be used, because in the RP-HPLC analysis, it elutes at the same retention time of esterified peptides, thus preventing the quantification of the reaction yield. The reaction mixtures were quenched after incubation for 4 h. The reaction kinetic was followed by RP-HPLC analysis. The reaction yield was estimated on the basis of the percentage of residual unmodified peptide calculated from the area of the RP-HPLC peak after 4 h of incubation with respect to 0 h. Both

dx.doi.org/10.1021/bc400378h | Bioconjugate Chem. 2014, 25, 470–480

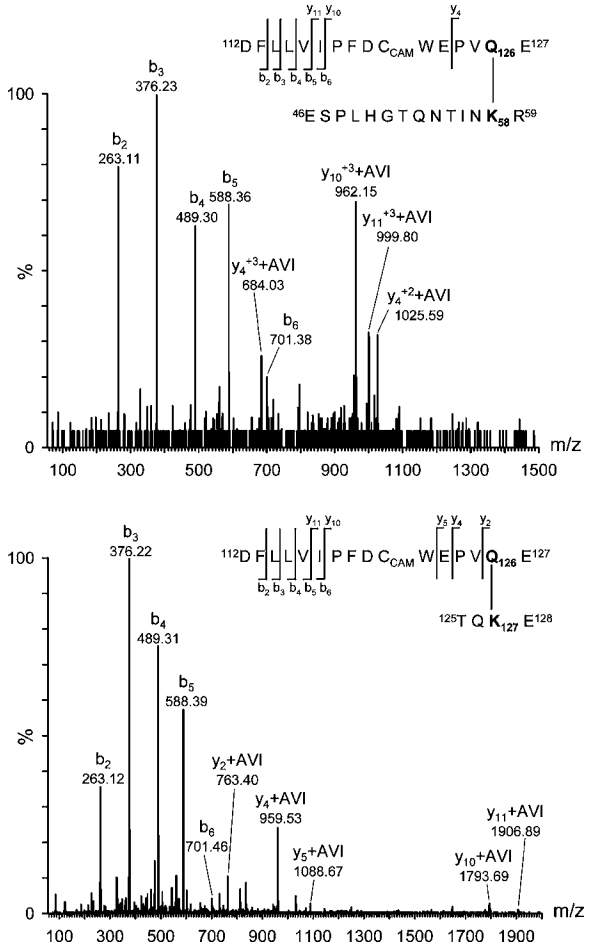


Figure 6. Electrospray MS/MS mass spectra of the fourth charged ion at 896.69 *m*/*z* of the avidin tryptic peptide 46−59 conjugated to the tryptic peptide 112−127 of GM-CSF (top) and of the doubly charged ion at 1247.59 *m*/*z* of the avidin tryptic peptide 125−128 conjugated to the tryptic peptide 112−127 of GM-CSF (bottom). In the MS/MS spectra, ions assigned to fragments of the series b and y of peptide 112−127 of GM-CSF are indicated and fragment ions of the y series that show a mass increment corresponding to the conjugation to avidin peptides are indicated as "+AVI". An inset in each MS/MS spectrum shows the sequences of the cross-linked peptides. In each inset, a line connecting Lys58 or Lys127 of avidin and Gln126 of GM-CSF represents the TGase-catalyzed formation of an isopeptide bond between the two residues. The carbamidomethylation of the cysteine residue of peptide 112−127 of GM-CSF is indicated as C_{CAM}. Fragments of the series b and y of peptide 112−127 of GM-CSF that were identified in the MS/MS spectra are indicated on its sequence.

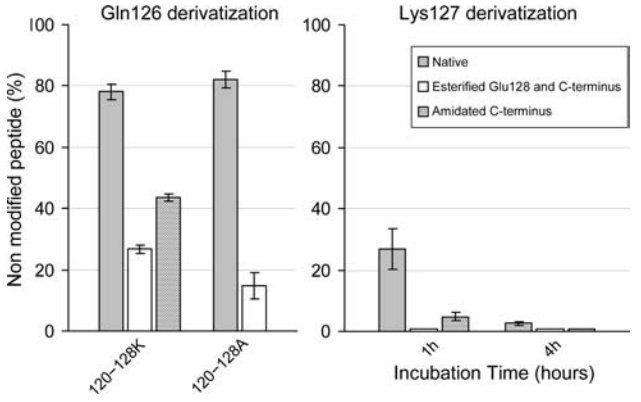


Figure 7. Histograms displaying the different yields of derivatization of the C-terminal peptides of avidin with HTA or ZQG mediated by TGase. The progress of the reaction was monitored as the area percentage of the RP-HPLC peak of the non-modified peptide at the different incubation times with respect to the 0 h of reaction. (left) Derivatization with HTA of peptides 120−128K, 120−128A-(Lys127Ala), of their esterified forms and of peptide 120−128K amidated at the C-terminus. The reaction mixtures were incubated for 4 h at 37 °C. (right) Derivatization of peptide 120−128K, of its esterified form, and of the C-terminal amidated form with ZQG after 1 and 4 h of reaction. Results are reported as mean values of two experiments, while error bars indicate the standard error from the mean.

peptides 120−128K and 120−128A are quite resistant to conjugation to HTA irrespective of the presence of the Lys residues since after 4 h of reaction about 80% of the peptides is not modified (Figure 7; Figure S7 and Table S3). The same results were obtained if the peptides were reacted with DC (data not shown). Esterification of Glu128 and of the C-terminus of the peptides significantly increases the reaction yield, since the unmodified peptide is 27% for 120−128K^{2Me} and 15% for 120−128A^{2Me} after 4 h of incubation (Figure 7; Figure S7 and Table S3), the species with Ala being slightly more reactive. In the case of peptide 120−128K-NH_2, there is still 44% of unmodified peptide after 4 h of incubation which is an intermediate value between the native peptide and the esterified peptide (Figure 7; Figure S8 and Table S3). We can conclude that charge repulsion of both the acidic −COOH groups of Glu128 and of the C-terminus of the protein with the negatively active site of TGase mainly determines the extremely low reactivity of residue Gln126.[17]

We also investigated the effect of the negative charges on the derivatization of peptide 120−128K with ZQG mediated by TGase. While peptides 120−128K^{2Me} and 120−128K-NH_2 are completely derivatized already after 1 h of reaction (Figure 7; Figure S8 and S9, Table S3), there is still 27% of unmodified peptide 120−128K after 1 h of reaction. These results suggest that the negatively charged −COOH group at the C-terminus of the protein slows down the derivatization of Lys127, while the lateral chain of Glu128 has a negligible effect on Lys reactivity.

Biotin-Binding Properties and Aggregation State of the Avidin-ZQG Derivative. The number of functional biotin binding sites in avidin^{1,2ZQG} was determined using the HABA assay.[33] This dye-binding method exploits the spectral changes occurring upon binding of the dye 4-hydroxyazobenzene-2-carboxylic acid (HABA) to avidin. Indeed, binding of HABA to avidin is accompanied by a red shift of the maximum of absorbance of the dye from 350 nm to 500 nm. If biotin is added to the avidin−HABA complex, the dye is stoichiometrically displaced by biotin and its UV−visible spectrum returns to be that of unbound HABA. Measurement of the variation in the absorbance at 500 nm upon biotin addition allows the number of biotin binding sites and, consequently, the concentration of the protein to be estimated in the avidin sample. We determined the protein concentration (mg/mL) of native avidin and the avidin^{1,2ZQG} sample by measuring the absorbance of the protein at 280 nm and by using the HABA colorimetric assay, and then we compared the values obtained for each sample (Table 4). A good agreement was observed on comparing the values measured with the two different methods for both native avidin and the avidin^{1,2ZQG} sample. This result

Table 4. Concentration of Avidin Measured Using the Colorimetric HABA Assay on Samples of Native Avidin and Avidin^{1,2ZQG}

sample	Concentration (mg/mL) based on $A_{280\ nm}$	Concentration (mg/mL) based on HABA assay
Avidin	0.37	0.35
Avidin^{1,2ZQG}	0.38	0.35

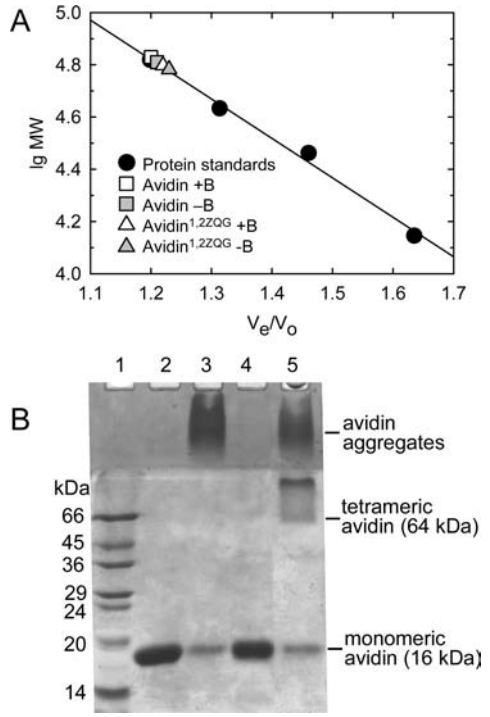


Figure 8. Analysis of the aggregation state of avidin and avidin^{1,2ZQG} and of its thermostability in the presence or in the absence of biotin. (A) Gel filtration profiles of avidin and avidin^{1,2ZQG} (the reaction mixture of avidin with ZQG and TGase after 5 h of incubation) in the absence (−B) and in the presence of biotin (+B). A Superdex 75 column was used for the analysis. Serum albumin (66 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), and lactalbumin (14 kDa) were used as molecular weight standards to calibrate the column. The logarithm of molecular weight (MW) is plotted versus V_e/V_o (V_e is the elution volume and V_o the void volume of the column). (B) Native avidin and avidin^{1,2ZQG} (the reaction mixture) were preincubated for 20 min at 100 °C in the presence of SDS without and with biotin (−B and +B, respectively) and analyzed by SDS-PAGE. Protein samples were loaded in the following order: MW standards (lane 1); avidin −B (lane 2); avidin +B (lane 3); avidin^{1,2ZQG} −B (lane 4); avidin^{1,2ZQG} +B (lane 5).

suggests that the number of functional binding sites in avidin^{1,2ZQG} is maintained.

Modification of Lys58 and Lys 127 of avidin could also affect the quaternary structure of the protein by destabilizing the homotetramer. The aggregation state of avidin^{1,2ZQG} was analyzed by gel filtration chromatography in the presence and absence of free biotin and compared to native avidin (Figure 8A). Avidin and avidin^{1,2ZQG} have similar elution volumes with and without the addition of biotin. Moreover, their migration is consistent with the theoretical molecular mass of the homotetramer of avidin (64 kDa) since the observed masses range from 67.3 kDa (avidin in the absence of biotin) to 60.7 kDa (avidin^{1,2ZQG} in the presence of biotin). The thermal stability of the tetrameric form of avidin^{1,2ZQG} was also analyzed by SDS-PAGE (Figure 8B).[32] Avidin is a positively charged glycoprotein (pI ∼ 10.5) and it aggregates extensively when mixed with anionic detergents as sodium dodecyl sulfate (SDS) leading to aggregated tetramers that remain in the stacking gel of SDS-PAGE analyses. However, when incubated at 100 °C in the presence of SDS, avidin migrates as a monomer (Figure 8B, lane 2), while avidin bound to biotin forms aggregated tetramers that fail to enter the separating gel (Figure 8B, lane 3). Similarly to native avidin, avidin^{1,2ZQG} migrates as a monomer, when incubated at 100 °C in the presence of SDS (Figure 8B, lane 4). However, avidin^{1,2ZQG} bound to biotin still forms a tetramer, but with a reduced tendency to aggregate after incubation at 100 °C in the presence of SDS (Figure 8B, lane 5). Indeed, part of the tetramer formed by avidin^{1,2ZQG} enters the separating gel. This observation can be explained in light of the fact that TGase-mediated derivatization of avidin modifies two Lys residues that are no longer available for SDS binding. Analogously, it has been demonstrated that acetylation of avidin inhibits the aggregation of the biotin-bound tetramer when boiled in the presence of SDS.[32] We can conclude that avidin^{1,2ZQG} forms tetramers that display stability characteristics similar to those of native avidin. Interestingly, derivatization of two of the nine lysine residues of the protein reduces the aggregation of the tetramers in the presence of SDS.

DISCUSSION

Structural and Sequence Determinants of the TGase Site-Specific Modification of Avidin. We previously demonstrated that Gln and Lys residues selectively modified by TGase are those embedded in flexible−unfolded chain segments of the protein.[30] The rationale is that only a disordered region of a polypeptide chain can bind to the active site of TGase in an extended conformation. This conformation is required to maximize the interaction with the enzyme active site and for an efficient catalysis.[30,39] Avidin is a tightly structured protein, since even in the apo-form its melting temperature is of 85 °C.[40] Several 3D structures of avidin in the apo- and holo- forms have been published with similar results.[3−6] They indicate that the most flexible regions in the apo-form of the protein are located at the level of the loop connecting strands 3 and 4 (residues Ala36−Ile44), at the N-terminus (Ala1−Arg2) and C-terminus (Arg124−Glu128) of the protein and in the Asp86−Arg88 segment.[5] The avidin sequence contains nine Lys residues and four Gln residues (Figure 1A). Among these residues, only Gln126 and Lys127 are located in the disordered C-terminal region of the protein. Indeed, Lys127 is modified by TGase. Lys58, which is also a substrate of TGase, is located in the middle of a loop region (residues 54−62) connecting strands 4 and 5, that is expected to display some level of flexibility.

Surprisingly, TGase modifies residue Gln126 at a negligible level. Our study on synthetic peptides covering the C-terminus of avidin (120−128K) and carrying the Lys to Ala mutation (120−128A) showed that Lys127 has a minor effect on Gln126 derivatization (Figure 7). On the other end, derivatization with HTA of the esterified peptides (120−128K^{2Me} and 120−128A^{2Me}) and of peptide 120−128K amidated at the C-terminus indicated that the reaction yield increases by 3−5-fold when neutralizing the two nearby negative charges (i.e., the lateral chain of Glu128 and the C-terminal −COOH). The two −COOH groups both contribute to inhibit the derivatization of Gln126 since amidation of the C-terminus increases the yield of the reaction, but not to the same value as with the

dx.doi.org/10.1021/bc400378h | Bioconjugate Chem. 2014, 25, 470−480

neutralization of both the negative charges. Interestingly, in GM-CSF the reactive Gln residue (i) is also flanked at the C-terminus by a Glu residue (i+1; sequence -Val-Gln$_{126}$-Glu-COOH) and by the C-terminal −COOH group of the protein. If we draw the disordered C-terminal sequences of avidin (residues -Thr$_{125}$-Gln-Lys-Glu$_{128}$-COOH) and of GM-CSF in an extended conformation with peptide bonds in the energetically favored *trans* conformation, we observe that Gln126 (i) and Glu128 (i+2) of avidin are located on the same face of the strand while the lateral chains of Gln126 and Glu127 residues of GM-CSF are located on the opposite faces (Figure S10). The orientation of the negatively charged side chain of Glu128 in avidin likely maximizes the repulsion with the negatively charged active site of TGase, thus preventing the derivatization of the nearby Gln residue.[17]

A negative charge in the vicinity of the reactive primary amine also decreases the yield of the TGase catalyzed reaction.[19,41] The derivatization of Lys127 of avidin is affected to a lower extent by the C-terminal −COOH group, while the nearby Glu residue has almost no effect. Indeed, the negatively charged Glu128 is in position i+1 with respect to the Lys residue (i) and its lateral chain is located on the opposite face of the strand resulting in a lower repulsion with the TGase active site. Overall, these results demonstrate that the reactivity of Gln and Lys residues to TGase is determined by the local conformation of the polypeptide chain, but it can be strongly affected by the presence of nearby negatively charged groups. In the case of avidin, we observe a low level of derivatization at the level of Gln126, while this residue is derivatized with a 20% yield in peptide 120−128K (carrying the two −COOH groups). We thus do not exclude that structural features of the avidin tetramer could also affect Gln modification.

Biotechnological Applications of the TGase-Mediated Derivatization of Avidin. Avidin modified with ZQG at the level Lys127 or Lys58 still displays biotin binding properties and maintains its tetrameric quaternary structure. This is expected since residues that participate in the binding of biotin are located on the opposite end of the avidin barrel with respect to the two reactive Lys residues (Figure 1B).[3] Moreover, the interactions that stabilize avidin quaternary structure are not affected by the modification since Lys58 and Lys127 are not among the residues involved in tetramer formation.[3,42] This is also demonstrated by the fact that a circularly permuted form of avidin, where the new N- and C-termini are located at Arg59 and Lys58, respectively, maintains the biotin-binding properties of the native protein and forms a quaternary structure similar to wild-type avidin.[43]

Currently, chemical derivatization at the level of amino groups is mostly used to produce avidin derivatives.[14,15] Avidin contains ten amino groups located on nine lysine residues and at the N-terminus of the protein that can be chemically modified. Among Lys residues, it has been demonstrated that derivatization of any of the three residues Lys45, Lys111, and Lys94 is detrimental for the biotin binding properties of the protein.[44] Chemical derivatization methods have the disadvantage that the protein can be conjugated at the level of heterogeneous sites, leading to mixtures of isomers even when three−four molecules are conjugated per avidin chain.[16] In this respect, the TGase-mediated modification of avidin is a method that allows the site-specific derivatization of this protein, leading to a maximum of two molecules conjugated per avidin chain. Considering the high yield of derivatization obtained with the ZQG substrate, most of the modified avidin tetramers will contain an average of four conjugated molecules and a low percentage of the modified tetramers will be conjugated with five to eight molecules due to the formation of biderivative species at the level of the avidin polypeptide chain. TGase can be used to modify avidin with organic molecules, but also for the production of heterodimeric protein conjugates of avidin with another protein possessing a reactive Gln residue, as demonstrated in this paper with GM-CSF. Alternatively, the target protein can be expressed with a Gln containing tag at the C- or at the N-termini that can be conjugated to avidin using TGase.[45]

In general, we believe that, given the many applications of the avidin−biotin technology, the TGase-mediated site-specific modification of avidin will be a novel tool to produce useful derivatives of this protein. In particular, avidin has potential clinical applications in which the use of homogeneous site-specific derivatives would be beneficial. Indeed, this protein specifically binds to the surface of D-galactose receptor positive tumors mainly due to its glycosilation.[11,46] Because of its specificity for tumor cells, avidin has been studied as a carrier of drugs or radionuclides to be specifically delivered to cancer cells or when conjugated to fluorophores, as an optical imaging agent of tumors.[11,12,46] Another interesting clinical application of the avidin−biotin system is for the immunotargeting of radionuclides to solid tumors. In particular, in pretargeting applications the biotinylated therapeutic agent is delivered to the tumor previously targeted by an avidin conjugated antibody or with a biotinylated antibody that is complexed with avidin.[10] Since when administered in pretargeting applications avidin shows immunological properties and unfavorable pharmacokinetic, PEGylated avidin derivatives were prepared.[16,47] Chemical PEGylation of avidin at the level of Lys residues led to avidin conjugates with reduced immunogenicity and improved pharmacological properties. However, these conjugates displayed a reduced binding to biotin and in particular to biotinylated antibodies likely due to derivatization of Lys residues close to the biotin binding.[16] In this respect, the TGase-mediated PEGlation of avidin could be an alternative approach in an attempt to produce homogeneous and site-specific derivatives to be used in pretargeting applications.[23]

■ ASSOCIATED CONTENT

Supporting Information

Chemical structure of the substrates used in the TGase-mediated modification of avidin (Figure S1); RP-HPLC and ESI-MS analyses of avidin and avidin incubated with TGase and HTA or DC (Figure S2); extracted ion chromatograms for the LC-MS/MS analysis of the digest of avidin incubated with HTA for 24 h in the presence of TGase (Figure S3); ESI-MS/MS mass spectra of modified tryptic peptides of avidin^{1,2ZQG} (Figure S4 and S5); amino acid sequence of GM-CSF (Figure S6); RP-HPLC analyses of peptides 120−128K, 120−128A, of their esterified forms and of peptide 120−128K amidated at the C-terminus modified with HTA and ZQG (Figure S7−S9); extended chain conformations of residues 125−127 of GM-CSF and 125−128 of avidin (Figure S10); MASCOT search results of the LC-MS/MS data acquired on the avidin samples incubated with TGase and with HTA or DC and digested with endoproteinase Lys-C (Table S1); MASCOT search results of the LC-MS/MS analysis of the tryptic digest of avidin conjugated to GM-CSF (Table S2); molecular masses measured for peptides 120−128K, 120−128A, for their methylated forms and for peptide 120−128K amidated at the

C-terminus conjugated to ZQG and/or HTA (Table S3). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone +39(0)49-8276155; Fax +39(0)49-8276159; e-mail: barbara.spolaore@unipd.it.

Present Address

Samanta Raboni, Department of Pharmacy, University of Parma, Parma, Italy.

Funding

This work was supported by the Research Project 2009 of the University of Padua. N. D. was supported by a grant from the CARIPARO Foundation (Padua).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge Dr. Daniele Dalzoppo for technical assistance in the production of methyl ester derivatives, Silvia Gelio for performing some of the experiments here reported, and the peptide synthesis service of CRIBI for the synthetic peptides. We thank also Dr. Margherita Morpurgo for useful suggestions on methods of avidin purification and the Proteomic Center of the University of Padua for the use of the 6520 Q-TOF mass spectrometer.

ABBREVIATIONS

apo-avidin, avidin depleted of biotin; holo-avidin, avidin bound to biotin; AcCN, acetonitrile; DC, dansylcadaverine; ESI-MS, electrospray-ionization mass spectrometry; E/S, enzyme to substrate ratio; GM-CSF, human granulocyte-macrophage colony-stimulating factor; HABA, 4-hydroxyazobenzene-2-carboxylic acid; HPLC, high-performance liquid chromatography; HTA, 5-hydroxytryptamine; IAA, iodoacetamide; MS/MS, tandem mass spectrometry; PAGE, polyacrylamide gel electrophoresis; RP, reverse-phase; RT, retention time; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; TGase, transglutaminase; ZEG, carbobenzoxy-L-glutamyl-glycine; ZQG, carbobenzoxy-L-glutaminyl-glycine

REFERENCES

(1) Green, N. M. (1975) Avidin. *Adv. Prot. Chem. 295*, 85−133.

(2) Green, N. M. (1990) Avidin and streptavidin. *Methods Enzymol. 184*, 51−67.

(3) Livnah, O., Bayer, E. A., Wilchek, M., and Sussman, J. L. (1993) Three-dimensional structures of avidin and the avidin-biotin complex. *Proc. Natl. Acad. Sci. U.S.A. 90*, 5076−5080.

(4) Pugliese, L., Malcovati, M., Coda, A., and Bolognesi, M. (1994) Crystal structure of apo-avidin from hen egg-white. *J. Mol. Biol. 235*, 42−46.

(5) Nardone, E., Rosano, C., Santambrogio, P., Curnis, F., Corti, A., Magni, F., Siccardi, A. G., Paganelli, G., Losso, R., Apreda, B., Bolognesi, M., Sidoli, A., and Arosio, P. (1998) Biochemical characterization and crystal structure of a recombinant hen avidin and its acidic mutant expressed in Escherichia coli. *Eur. J. Biochem. 256*, 453−460.

(6) Pugliese, L., Coda, A., Malcovati, M., and Bolognesi, M. (1993) Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7 Å resolution. *J. Mol. Biol. 231*, 698−710.

(7) Wilchek, M., and Bayer, E. A. (1990) Introduction to avidin-biotin technology. *Methods Enzymol. 184*, 5−13.

(8) Bayer, E. A., and Wilchek, M. (1990) Application of avidin-biotin technology to affinity-based separations. *J. Chromatogr. 510*, 3−11.

(9) Wilchek, M., and Bayer, E. A. (1990) Avidin-biotin mediated immunoassays: Overview. *Methods Enzymol. 184*, 467−469.

(10) Lesch, H. P., Kaikkonen, M. U., Pikkarainen, J. T., and Ylä-Herttuala, S. (2010) Avidin-biotin technology in targeted therapy. *Expert Opin. Drug Deliv. 7*, 551−564.

(11) Yao, Z., Zhang, M., Sakahara, H., Saga, T., Arano, Y., and Konishi, J. (1998) Avidin targeting of intraperitoneal tumor xenografts. *J. Natl. Cancer Inst. 90*, 25−29.

(12) Ogawa, M., Kosaka, N., Longmire, M. R., Urano, Y., Choyke, P. L., and Kobayashi, H. (2009) Fluorophore-quencher based activatable targeted optical probes for detecting in vivo cancer metastases. *Mol. Pharmaceutics 6*, 386−395.

(13) Pardridge, W. M. (2008) Re-engineering biopharmaceuticals for delivery to brain with molecular Trojan horses. *Bioconjugate Chem. 19*, 1327−1338.

(14) Bayer, E. A., and Wilchek, M. (1990) Avidin- and streptavidin-containing probes. *Methods Enzymol. 184*, 174−187.

(15) Haugland, R. P., and Bhalgat, M. K. (2008) Preparation of avidin conjugates. *Methods Mol. Biol. 418*, 1−12.

(16) Salmaso, S., Semenzato, A., Bersania, S., Chinol, M., Paganelli, G., and Caliceti, P. (2005) Preparation and characterization of active site protected poly(ethylene glycol)-avidin bioconjugates. *Biochim. Biophys. Acta 1726*, 57−66.

(17) Kashiwagi, T., Yokoyama, K., Ishikawa, K., Ono, K., Ejima, D., Matui, H., and Suzuki, E. (2002) Crystal structure of microbial transglutaminase from Streptoverticillium mobaraense. *J. Biol. Chem. 277*, 44252−44260.

(18) Folk, J. E. (1983) Mechanism and basis for specificity of transglutaminase-catalyzed ε-(γ-glutamyl)lysine bond formation. *Adv. Enzymol. Relat. Areas Mol. Biol. 54*, 1−56.

(19) Ohtsuka, T., Sawa, A., Kawabata, R., Nio, N., and Motoki, M. (2000) Substrate specificities of microbial transglutaminase for primary amines. *J. Agric. Food. Chem. 48*, 230−233.

(20) Kulik, C., Heine, E., Weichold, O., and Möller, M. (2009) Synthetic substrates as amine donors and acceptors in microbial transglutaminase-catalysed reactions. *J. Mol. Catal. B: Enzym. 57*, 237−241.

(21) Ohtsuka, T., Ota, M., Nio, N., and Motoki, M. (2000) Comparison of substrate specificities of transglutaminases using synthetic peptides as acyl donors. *Biosci. Biotechnol. Biochem. 64*, 2608−2613.

(22) Sato, H. (2002) Enzymatic procedure for site-specific PEGylation of proteins. *Adv. Drug Delivery Rev. 54*, 487−504.

(23) Fontana, A., Spolaore, B., Mero, A., and Veronese, F. M. (2008) Site-specific modification and PEGylation of pharmaceutical proteins mediated by transglutaminase. *Adv. Drug Delivery Rev. 60*, 13−28.

(24) Mariniello, L., Porta, R., Sorrentino, A., Giosafatto, C. V., Rossi Marquez, G., Esposito, M., and Di Pierro, P. (2013) Transglutaminase-mediated macromolecular assembly: production of conjugates for food and pharmaceutical applications. *Amino Acids*, DOI: 10.1007/s00726-013-1561-6.

(25) Kamiya, N., Abe, H., Goto, M., Tsuji, Y., and Jikuya, H. (2009) Fluorescent substrates for covalent protein labeling catalyzed by microbial transglutaminase. *Org. Biomol. Chem. 7*, 3407−3412.

(26) Jeger, S., Zimmermann, K., Blanc, A., Grünberg, J., Honer, M., Hunziker, P., Struthers, H., and Schibli, R. (2010) Site-specific and stoichiometric modification of antibodies by bacterial transglutaminase. *Angew. Chem., Int. Ed. Engl. 49*, 9995−9997.

(27) Porta, R., Giosafatto, C. V., di Pierro, P., Sorrentino, A., and Mariniello, L. (2013) Transglutaminase-mediated modification of ovomucoid: effects on its trypsin inhibitory activity and antigenic properties. *Amino Acids 44*, 285−292.

(28) Sorrentino, A., Giosafatto, C. V., Sirangelo, I., De Simone, C., Di Pierro, P., Porta, R., and Mariniello, L. (2012) Higher susceptibility to amyloid fibril formation of the recombinant ovine prion protein modified by transglutaminase. *Biochim. Biophys. Acta 1822*, 1509−1515.

(29) Mero, A., Spolaore, B., Veronese, F. M., and Fontana, A. (2009) Transglutaminase-mediated PEGylation of proteins: direct identification of the sites of protein modification by mass spectrometry using a novel monodisperse PEG. *Bioconjugate Chem. 20*, 384−389.
(30) Spolaore, B., Raboni, S., Ramos Molina, A., Satwekar, A., Damiano, N., and Fontana, A. (2012) Local unfolding is required for the site-specific protein modification by transglutaminase. *Biochemistry 51*, 8679−8689.
(31) Bozzo, C., Spolaore, B., Toniolo, L., Stevens, L., Bastide, B., Cieniewski-Bernard, C., Fontana, A., Mounier, Y., and Reggiani, C. (2005) Nerve influence on myosin light chain phosphorylation in slow and fast skeletal muscles. *FEBS J. 272*, 5771−5785.
(32) Bayer, E. A., Ehrlich-Rogozinski, S., and Wilchek, M. (1996) Sodium dodecyl sulfate-polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin. *Electrophoresis 17*, 1319−1324.
(33) Green, N. M. (1970). Spectrophotometric determination of avidin and biotin. *Methods in Enzymol. 18*, 418−424.
(34) Bruch, R. C., and White, H. B., 3rd (1982) Compositional and structural heterogeneity of avidin glycopeptides. *Biochemistry 21*, 5334−5341.
(35) Rozwarski, D. A., Diederichs, K., Hecht, R., Boone, T., and Karplus, P. A. (1996) Refined crystal structure and mutagenesis of human granulocyte-macrophage colony-stimulating factor. *Proteins 26*, 304−313.
(36) Coussons, P. J., Price, N. C., Kelly, S. M., Smith, B., and Sawyer, L. (1992) Factors that govern the specificity of transglutaminase-catalysed modification of proteins and peptides. *Biochem. J. 282*, 929−930.
(37) Pastor, M. T., Diez, A., Pérez-Payá, E., and Abad, C. (1999) Addressing substrate glutamine requirements for tissue transglutaminase using substance P analogues. *FEBS Lett. 451*, 231−234.
(38) Sugimura, Y., Yokoyama, K., Nio, N., Maki, M., and Hitomi, K. (2008) Identification of preferred substrate sequences of microbial transglutaminase from Streptomyces mobaraensis using a phage-displayed peptide library. *Arch. Biochem. Biophys. 477*, 379−383.
(39) Tyndall, J. D. A., Nall, T., and Fairlie, D. P. (2005) Proteases universally recognize beta strands in their active site. *Chem. Rev. 105*, 973−999.
(40) Donovan, J. W., and Ross, K. D. (1973) Increase in the stability of avidin produced by binding of biotin. A differential scanning calorimetric study of denaturation by heat. *Biochemistry 12*, 512−517.
(41) Gundersen, M. T., Keillor, J. W., and Pelletier, J. N. (2014) Microbial transglutaminase displays broad acyl-acceptor substrate specificity. *Appl. Microbiol. Biotechnol. 98*, 219−230.
(42) Wilchek, M., Bayer, E. A., and Livnah, O. (2006) Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction. *Immunol. Lett. 103*, 27−32.
(43) Nordlund, H. R., Laitinen, O. H., Hytönen, V. P., Uotila, S. T., Porkka, E., and Kulomaa, M. S. (2004) Construction of a dual chain pseudotetrameric chicken avidin by combining two circularly permuted avidins. *J. Biol. Chem. 279*, 36715−36719.
(44) Gitlin, G., Bayer, E. A., and Wilchek, M. (1987) Studies on the biotin-binding site of avidin. Lysine residues involved in the active site. *Biochem. J. 242*, 923−926.
(45) Takazawa, T., Kamiya, N., Ueda, H., and Nagamune, T. (2004) Enzymatic labeling of a single chain variable fragment of an antibody with alkaline phosphatase by microbial transglutaminase. *Biotechnol. Bioeng. 86*, 399−404.
(46) Hama, Y., Urano, Y., Koyama, Y., Kamiya, M., Bernardo, M., Paik, R. S., Krishna, M. C., Choyke, P. L., and Kobayashi, H. (2006) In vivo spectral fluorescence imaging of submillimeter peritoneal cancer implants using a lectin-targeted optical agent. *Neoplasia 8*, 607−612.
(47) Caliceti, P., Chinol, M., Roldo, M., Veronese, F. M., Semenzato, A., Salmaso, S., and Paganelli, G. (2002) Poly(ethylene glycol)-avidin bioconjugates: suitable candidates for tumor pretargeting. *J. Controlled Release 83*, 97−108.