Mass spectrometric analysis of a UV-cross-linked protein–DNA complex: Tryptophans 54 and 88 of *E. coli* SSB cross-link to DNA

HANNO STEEN, JØRGEN PETERSEN, MATTHIAS MANN, AND OLE N. JENSEN

Center for Experimental BioInformatics, Department of Biochemistry and Molecular Biology, University of Southern Denmark/Odense University, DK-5230 Odense M, Denmark

(RECEIVED February 20, 2001; FINAL REVISION July 2, 2001; ACCEPTED July 12, 2001)

Abstract

Protein–nucleic acid complexes are commonly studied by photochemical cross-linking. UV-induced cross-linking of protein to nucleic acid may be followed by structural analysis of the conjugated protein to localize the cross-linked amino acids and thereby identify the nucleic acid binding site. Mass spectrometry is becoming increasingly popular for characterization of purified peptide–nucleic acid heteroconjugates derived from UV cross-linked protein–nucleic acid complexes. The efficiency of mass spectrometry-based methods is, however, hampered by the contrasting physico-chemical properties of nucleic acid and peptide entities present in such heteroconjugates. Sample preparation of the peptide–nucleic acid heteroconjugates is, therefore, a crucial step in any mass spectrometry-based analytical procedure. This study demonstrates the performance of four different MS-based strategies to characterize *E. coli* single-stranded DNA binding protein (SSB) that was UV-cross-linked to a 5-iodouracil containing DNA oligomer. Two methods were optimized to circumvent the need for standard liquid chromatography and gel electrophoresis, thereby dramatically increasing the overall sensitivity of the analysis. Enzymatic degradation of protein and oligonucleotide was combined with miniaturized sample preparation methods for enrichment and desalting of cross-linked peptide–nucleic acid heteroconjugates from complex mixtures prior to mass spectrometric analysis. Detailed characterization of the peptidic component of two different peptide–DNA heteroconjugates was accomplished by matrix-assisted laser desorption/ionization mass spectrometry and allowed assignment of tryptophan-54 and tryptophan-88 as candidate cross-linked residues. Sequencing of those peptide–DNA heteroconjugates by nanoelectrospray quadrupole time-of-flight tandem mass spectrometry identified tryptophan-54 and tryptophan-88 as the sites of cross-linking. Although the UV-cross-linking yield of the protein–DNA complex did not exceed 15%, less than 100 pmole of SSB protein was required for detailed structural analysis by mass spectrometry.

Keywords: Nanoelectrospray tandem mass spectrometry; DNA–protein cross-linking; 5-iodouracil; MALDI mass spectrometry

Reprint requests to: Ole Nørregaard Jensen, Department of Biochemistry and Molecular Biology, University of Southern Denmark/Odense University, Campusvej 55, DK-5230 Odense M, Denmark; e-mail: jenseno@bmb.sdu.dk; fax: +45-65-50-24-67.

Abbreviations: CID, low-energy collision-induced fragmentation; Da/kDa, Dalton/(kilo)Dalton = atomic mass units; ESI, electrospray ionization; IMAC, immobilized metal affinity chromatography; 5-IU, 5-iodouracil; LRF, laboratory reference frame; MALDI, matrix-assisted laser desorption/ionization; *m/z*, mass-to-charge ratio; MS, mass spectrometry; MS/MS, tandem mass spectrometry; NTA, nitrilotriacetic acid; PAGE, polyacrylamide gel electrophoresis; SDS, sodium dodecyl sulfate; SSB, single-stranded DNA-binding protein; TOF, time-of-flight.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1101/ps.07601>.

Protein–nucleic acid interaction is involved in many cellular processes including transcription, translation, and DNA duplication, warranting the development of sensitive methods to study those interactions. UV-induced photochemical cross-linking of protein to nucleic acid is a commonly used method to study molecular assemblies of protein and oligonucleotides. UV-irradiation of natural or derivatized nucleobases generates highly reactive intermediates that form zero-length covalent cross-links to protein molecules in the vicinity. It can be assumed that the amino acids that are cross-linked to nucleobases are located at the site of nucleic

acid–protein interaction (for a recent review of this topic, see Meisenheimer and Koch 1997). Thus, the identification of cross-linked peptide and/or amino acid residues leads to the characterization of the nucleic acid binding site of the protein. Irradiation at wavelengths around 254 nm induces photochemical cross-linking of nucleobases, typically thymine, to protein. Light at this wavelength is, however, not only absorbed by the nucleobases but also by other functional groups of nucleic acid and protein, and may lead to photodamage (e.g., strand cleavage, oxidation), which in turn, may affect the integrity and functionality of the components present in the reaction mixture. This problem can be circumvented by using modified nucleotides containing azido-, bromo-, iodo-, or thio-substituted nucleobases, which absorb UV-light at longer wavelengths, i.e., above 300 nm. The risk of photodamage is thereby reduced at the expense of potential structural disturbances of the RNA/DNA induced by exchange of native nucleobases with photoreactive analogs.

Traditionally, covalent nucleic acid–protein heteroconjugates are sequenced by Edman degradation to determine the modified amino acid (Merrill et al. 1984; Allen et al. 1991; Prasad et al. 1993; Musier-Forsyth and Schimmel 1994; Malkov and Camerini-Otero 1995; Stump and Hall 1995; Urlaub et al. 1995; Wang and Adzuma 1996; Holz et al. 1999; Pingoud et al. 1999; Wong and Reich 2000). In Edman degradation, the cross-linked amino acid is deduced by observing a blank sequencing cycle, i.e., the absence of a standard PTH-derivatized amino acid peak, during the termination of the amino acid sequence. Further complications may arise if several modified amino acids are located within one peptide, or when complete purification of the heteroconjugate has not been achieved. Mass spectrometry has been used in a number of cross-linking studies for molecular weight determination and amino acid sequencing (Jensen et al. 1993; Qin and Chait 1997; Connor et al. 1998; Golden et al. 1999; Gafken et al. 2000; Rieger et al. 2000). MS-based approaches have several advantages over traditional Edman sequencing: (1) better sensitivity and specificity, (2) direct identification of the modified amino acid by its molecular mass (difference between two fragment ions in the MS/MS-spectrum), (3) the possibility to sequence oligonucleotide as well as peptide components in a heteroconjugate by tandem mass spectrometry or by enzyme degradation combined with mass spectrometry (Jensen et al. 1996; Urlaub et al. 1997; Wang et al. 2000), and (4) mixtures can be analyzed.

To take full advantage of mass spectrometry and tandem mass spectrometry for analysis of peptide–oligonucleotide heteroconjugates several complications arising from the opposing physico-chemical properties of proteins and oligonucleotides have to be overcome (Jensen et al. 1996).

Matrices used for sample preparation of nucleic acid oligomers for matrix-assisted laser desorption/ionization

mass spectrometry (MALDI) mass spectrometry (Karas and Hillenkamp 1988) are usually not suitable for peptides, and vice versa (Jensen et al. 1997). This raises the questions of what kind of matrix is appropriate for MALDI mass spectrometric analysis of heteroconjugates comprising those two different classes of analytes. Any choice of MALDI matrix tends to compromise sensitivity and performance with respect to unmodified protein/peptide, unmodified nucleic acid, or heteroconjugates. As a result, only the unreacted components were observed when crude cross-linking reaction mixtures were analyzed, whereas the heteroconjugate present in the sample escaped detection (Jensen et al. 1993; Wong and Reich 2000).

Optimal conditions for analysis of oligonucleotides and peptides, respectively, by using electrospray ionization (ESI) mass spectrometry (Fenn et al. 1989) are very different. Oligonucleotides are usually analyzed in the negative ion mode by ESI mass spectrometry using alkaline solutions with ammonium salts as ion pairing reagent for the phosphate diester backbone of the nucleic acid. In contrast, peptides are analyzed and sequenced as cations using salt-free acidic solutions. Alkaline conditions usually prevent the formation of sequence-revealing fragment ions under low-energy collision-induced fragmentation (CID) tandem mass spectrometry, a process that requires protonation to be efficient. Additional problems arising from the contrasting behavior of nucleic acids and peptides lead to ionization suppression when cross-linked heteroconjugates are present with unreacted oligonucleotides or peptides. This suppression of ionization might be due to the higher hydrophilicity of oligonucleotide–peptide/protein heteroconjugates compared to the unmodified peptides/proteins. A similar problem is observed for phosphopeptides (Chen et al. 2001; Wind et al. 2001).

Although MS has been used in a number of studies of cross-linked peptide–oligonucleotide species, there are only few reports on the use of tandem mass spectrometry for the structural analysis of such heteroconjugates. The first example was provided by Jensen et al. (1996), who used a synthetic DNA–peptide heteroconjugate, consisting of a deoxythymidine hexamer and an undecapeptide, for ESI MS/MS studies in positive ion mode. This approach was also used in three recent studies (Golden et al. 1999; Gafken et al. 2000; Rieger et al. 2000). Qin and Chait (1997) employed MALDI in combination with ion trap MS/MS for analysis of a purified peptide–DNA heteroconjugate to obtain peptide sequence information that revealed the site of cross-linking.

In all published studies enrichment of the peptide–oligonucleotide heteroconjugates was a prerequisite for successful analysis of heteroconjugates by MS because it eliminates the problem of ionization suppression. Most published protocols dealing with the localization of protein–nucleic acid interaction sites, describe extensive one- or multidimen-

sional liquid chromatography for the purification of cross-linked species from complex peptide mixtures. Anion exchange chromatography (Merrill et al. 1984; Bennett et al. 1994; Pingoud et al. 1999), reverse-phase liquid chromatography (Urlaub et al. 1995; Golden et al. 1999; Rieger et al. 2000), and size-exclusion chromatography (Blatter et al. 1992; Urlaub et al. 1995) have been applied. Another widely applied approach uses polyacrylamide gel electrophoresis for separation followed by a passive elution protocol to obtain purified heteroconjugate (Hicke et al. 1994). This approach is simple, but it requires radioactive isotope labeling of the oligonucleotide to enable visualization of the DNA-containing species by autoradiography after enzymatic or chemical cleavage of the protein–nucleic acid complex (Malkov and Camerini-Otero 1995; Green et al. 1996; Golden et al. 1999). However, extensive chromatographic separation and/or gel separation compromises sensitivity because sample recovery is low for peptide–nucleic acid heteroconjugates. Protein amounts in the nanomole range are typically used as starting material to be able to localize the nucleic acid binding site by UV-cross-linking and standard protein chemistry techniques. High microgram to milligram amounts of purified protein are often difficult to produce, highlighting the need for more sensitive methods for systematic studies of protein–nucleic acid complexes and the corresponding UV-cross-linked heteroconjugates.

The aim of this study was to improve the overall sensitivity of analysis of UV-cross-linked protein–nucleic acid conjugates by using mass spectrometry while eliminating the need for gel electrophoresis or liquid chromatography. The model protein used was single-stranded DNA binding protein (SSB), which is involved in DNA repair, replication, and recombination in *E. coli* (Meyer and Laine 1990). A DNA 19-mer containing one 5-iodouracil was the model ligand. The SSB protein has previously been used for UV-cross-linking studies (Merrill et al. 1984; Prasad et al. 1993). A crystal structure based on X-ray diffraction of an SSB–DNA complex was recently published (Raghuathan et al. 2000), and was used to validate the data obtained in the present UV-cross-linking study.

Results and Discussion

Single-stranded DNA binding protein (SSB) from *E. coli* and a DNA 19-mer were used as the model system. The custom-made oligonucleotides used for this study had the sequence TGTAGCTGTTGATCTAAGT. A Stratalinker UV Crosslinker equipped with 312 nm bulbs was used to initiate the photochemical cross-linking reaction. Three oligonucleotides with a single 5-iodouracil incorporated at positions 9, 10, and 13, respectively, were incubated with SSB and tested for cross-linking efficiency. SDS-PAGE analysis showed that oligonucleotides with the halogen-substituted nucleobase located in the 9- and 10-positions gave the highest yields (data not shown). The yield of protein–DNA het-

eroconjugate was estimated to be around 15% by phosphorimaging of ^{32}P -labeled DNA (relative to the amount of DNA used; data not shown). TGTAGCTG(5-IU)TGATCTAAGT (My9IU) was used as the DNA probe in all subsequent experiments.

Characterization of protein–DNA cross-link by differential peptide mass mapping

UV-cross-linked SSB–DNA heteroconjugate was separated from unreacted SSB and DNA by SDS-PAGE. Two clear bands were observed after colloidal Coomassie staining of the gel, corresponding to the unmodified SSB and the cross-linked species, respectively (Fig. 1A). Both bands were cut out of the gel. Two sets of samples were analyzed by in-gel digestion and MS: one was enzymatically degraded using trypsin, the other one was treated with endoprotease Glu-C.

Small fractions of the digestion supernatants were used for peptide mass mapping by MALDI-TOF MS. In the tryptic digest of the control band (unreacted SSB) ion signals corresponding to the tryptic peptide T_{88-96} (m/z 1092.45) and T_{87-96} (m/z 1220.57) were clearly present (Fig. 1C), but these species were missing in the digest of the UV-cross-linked SSB–DNA heteroconjugate (Fig. 1B). The absence of these two ion signals in the digest of the cross-linked protein indicated that one of the amino acids in the region 87–96 was involved in the UV-cross-linking reaction. This interpretation was confirmed by peptide mass mapping subsequent to endoprotease GluC digest of the SSB–DNA conjugate (Fig. 1D) and the unreacted SSB (Fig. 1E). The control digest (unreacted SSB) showed an ion signal corresponding to the GluC–peptide G_{81-100} at m/z 2426.2, but this signal was absent in the GluC digest of the UV-cross-linked SSB–DNA heteroconjugate.

These data suggested that one major cross-link had been formed upon UV-irradiation of the SSB–My9IU mixture because peptides containing the amino acid involved in cross-linking were missing in the peptide mass map of the protein–DNA heteroconjugate, but were present in the control peptide mass map. This is only the case when a nearly homogeneous cross-link was generated, and therefore, cannot be applied when multiple, nearly equimolar cross-links are formed at different amino acid residues in the protein. As MALDI peptide mass mapping rarely provides complete sequence coverage of proteins, some peptides may be missing even in the control experiments. Therefore, it is recommended to use at least two different proteases for this approach.

Isolation of the DNA–peptide heteroconjugate by urea gel electrophoresis

More direct evidence about the site of cross-linking was obtained by gel electrophoretic isolation of peptide–DNA

Fig. 1. SDS-PAGE and differential peptide mass mapping of SSB-DNA heteroconjugate. (A) SDS-PAGE separation of protein-DNA conjugate from unreacted protein (visualized by colloidal Coomassie staining). (B) Peptide mass map obtained after in-gel trypsinolysis of SSB-DNA cross-link. (C) Control: peptide mass map of unreacted SSB, in-gel digested with trypsin. (D) Peptide mass map obtained after in-gel digestion of SSB-DNA cross-link with endoproteinase GluC. (E) Control: peptide mass map of unreacted SSB, in-gel digested with endoproteinase GluC. Peptides, which are present in the control but not in the digest of the heteroconjugate, are marked with arrows.

heteroconjugates as reported previously (Malkov and Camerini-Otero 1995; Green et al. 1996; Golden et al. 1999). The entire SSB-DNA cross-linking reaction mixture was digested with trypsin, and peptide-DNA heteroconjugates were isolated by urea-PAGE. In urea-PAGE, unreacted oligonucleotide migrates faster than the DNA-peptide heteroconjugate, whereas unmodified peptides enter the urea-polyacrylamide gel with low efficiency due to the relatively low number of negative charges at the pH value of the running buffer (TBE: pH 8.4).

The DNA used for cross-linking was spiked with ^{32}P -labeled DNA to enable visualization of the DNA-containing bands by autoradiography of the urea-PAGE gel. The gel displayed one intense band, attributable to the unreacted DNA, and two bands of lower intensity corresponding to a major (see Fig. 2A, upper arrow) and a minor (lower arrow) peptide-DNA heteroconjugate, respectively. A very weak band was visible in the upper part of the gel and corresponded to the intact SSB-DNA heteroconjugate. This observation indicated that the cross-linked SSB-DNA complex was quite resistant to tryptic digestion, as also observed by others (Hicke et al. 1994). This was rather surprising, because a trypsin to protein ratio of 1:2 was applied, and the digestion reaction was allowed to proceed for 48 h. It showed that the DNA protects the protein from digestion.

The gel band containing the cross-linked peptide-DNA heteroconjugates was excised from the gel and eluted into

water. The aqueous sample was then desalted and concentrated using a POROS R2 microcolumn (Wilm and Mann 1996; Kussmann et al. 1997), and eluted with matrix solution (3-hydroxy picolinic acid) directly onto the MALDI probe. The resulting MALDI mass spectrum displayed an intense ion signal corresponding to a singly charged species at 7308.1 Da, accompanied by a smaller one at 6816.0 Da (Fig. 2B). For comparison, a MALDI mass spectrum of the unreacted DNA oligomer with a molecular mass of 5961.7 Da is shown in Figure 2C.

The mass differences of 1346.4 and 854.3 Da between the heteroconjugates and the unreacted DNA allowed to calculate the mass of the peptides attached to the DNA probe (1475.3 Da and 983.2 Da; considering the loss of HI during the reaction), which in turn, allowed the identification of the tryptic SSB-peptide moieties that were covalently attached to the DNA. The calculations here and in the following sections were based on the photochemical reaction sequence of 5-iodouracil published by Norris et al. (1996) (Fig. 3): (1) upon UV-irradiation at wavelengths <330 nm the carbon-iodine bond of the 5-iodouracil is homolytically cleaved, generating a highly reactive vinyl radical in the 5-position; (2) the vinyl radical adds preferentially to the π -system of an aromatic amino acid side chain that is close in space, forming a C—C bond that is stable under normal MS and MS/MS conditions; (3) a hydrogen atom is abstracted, regenerating the aromatic system of the amino acid residue. In

Fig. 2. Isolation of the DNA-peptide conjugate by urea PAGE and analysis by MALDI-MS. The reaction was spiked with My9IU-DNA labeled at its 5'-terminus with [32 P]-phosphate. (A) Autoradiography of an urea PAGE after trypsinolysis of the SSB-DNA cross-link reaction. Bands attributable to unreacted DNA, peptide-DNA conjugate, and undigested SSB-DNA are indicated. (B) MALDI mass spectrum of the peptide-DNA heteroconjugate after elution from the gel band (indicated in [A]) showing ion signals attributable to DNA-T₈₅₋₉₆ (TRKWTDQSGQDR) at m/z 7308.1 and to DNA -T₅₀₋₅₆ (EQTEWHR) at m/z 6816.0. (C) Control: MALDI mass spectrum of unreacted My9IU, showing singly, doubly, and triply charged ion species. The ion signals marked with asterisks in (B) and (C) correspond to a DNA synthesis artifact.

this reaction mechanism hydroiodic acid (HI) is lost upon formation of the covalent bond.

The SSB amino acid sequence was inspected for potential peptides with the masses of 1475.3 and 985.45 Da fulfilling the trypsin cleavage conditions (cleavage C-terminal to Arg or Lys). A mass deviation of 3 Da was allowed for the mass search, corresponding to a mass accuracy of 500 ppm for the measured mass of the intact peptide-oligonucleotide heteroconjugates. This low accuracy was used because (1) the

spectrum was only externally calibrated, and (2) the external calibrant and the analyte belonged to two different classes of compounds. Different classes of compounds behave differently during the MALDI process; the flight time, and therefore, the mass calibration, can be affected by this (Glückmann and Karas 1999).

This investigation identified two tryptic peptides that matched the experimentally determined masses. One tryptic peptide 85-TRKWTDQSGQDR-96 (theoretical average

Fig. 3. Suggested reaction mechanism for the cross-linking of 5-iodouracil to tryptophan, adapted from Norris et al. (1996). The assumption, that the 2-position of the indol group of the tryptophan side chain is involved in the formation of the cross-link, is based on studies by Ito et al. (1980) and Dietz and Koch (1987).

mass: 1477.6 Da) was already assigned as potential site of cross-linking in the earlier differential peptide mass mapping (see Fig. 1). The second tryptic peptide that was retrieved by this search (T_{50–56} [EQTEWHR]) confirmed the presence of a second cross-linking site in SSB.

Although this method provided very useful information, it had several disadvantages: (1) gel electrophoresis and passive elution provided relatively poor recovery of the heteroconjugate, (2) it required ³²P-labeling with all precautions associated with this approach, and (3) no answer with respect to the exact site of cross-linking was obtained. These limitations prompted us to explore alternative approaches that would eliminate the need for gel electrophoresis, liquid chromatography, and radioactive labeling.

Proteinase K digest of DNA–protein cross-link

The new approach should preferentially (1) be fast and simple (one-step reaction), (2) minimize the sample handling to improve the overall sensitivity of analysis, and (3) replace any kind of liquid chromatography or gel electrophoresis by miniaturized efficient sample handling methods such as microcolumns (Wilm and Mann 1996; Kussmann et al. 1997).

Proteinase K, a highly active unspecific protease, was added to the complete cross-linking reaction solution comprising SSB, My9IU–DNA, and protein–DNA cross-link. The resulting mixture of DNA, small peptides, and DNA–peptide heteroconjugates was analyzed by MALDI TOF mass spectrometry (see Fig. 4A). One major ion signal at m/z 5961.7 attributable to unreacted My9IU–DNA was observed. This was not unexpected, due to the low cross-linking yield of 15% and, therefore, the presence of a large excess of unreacted DNA. In addition, six low abundance ion signals were identified in the m/z -range just above the unreacted My9IU–DNA (Fig. 4B). These six ion signals corresponded to heteroconjugates of DNA and short peptides of varying lengths generated by incomplete proteolysis. This was most likely the result of the covalently bound DNA protecting the peptide bonds near the site of cross-linking.

The masses of the small peptide moieties in these heteroconjugates were subsequently derived from the six m/z values (considering the elimination of hydroiodic acid; see above). Five of these peptide masses corresponded to stretches of SSB amino acids covering the residues 84 to 92 with the sequence RTR**KWTDQS** (see Table 1). Underlined in bold face is a “core” pentapeptide (87–91: KWTDQ) that was common to these species. The ion signal at m/z 6382.1 could correspond to either KWTD or WTDQ covalently linked to the DNA probe (Q and K are isobaric), so both amino acids had to be considered as potential sites of cross-linking. Additionally, the core sequence is extended successively on the N- and C-terminal by up to four

Fig. 4. (A) MALDI mass spectrum of proteinase K digest of SSB–DNA cross-link; the major ion signal corresponds to unreacted DNA. (B) Magnification of the m/z range above the unreacted My9IU–DNA showing diagnostics peaks attributable to DNA–peptide conjugates that identify the DNA-binding domain of SSB (see Table 1).

amino acid residues due to incomplete proteinase K digestion of the SSB–protein. In this way, the number of amino acids potentially involved in cross-link formation was reduced to five by this simple method. This agreed very well with the data obtained by the first two methods described above.

The sixth annotated peak used for the mass search (m/z 6561.1) corresponded to the tryptic peptide 52–TEWHR–56 (see Table 1), consistent with the presence of a second heteroconjugate species as also observed in the previous experiment.

Identification of the cross-linked amino acid residues by MS/MS

The previous three MALDI-based experiments demonstrated that the regions 87–91 and 52–56 of SSB cross-link to the DNA probe. We next employed nanoelectrospray tandem mass spectrometry for exact localization of the site

Table 1. Results obtained after proteinase K digest of the SSB-DNA cross-linking reaction mixture

Measured m/z (ave)	Calculated peptide mass [Da] (ave)	Theoretical peptide mass [Da] (ave)	From-To	Sequence
6382.1	549.3	548.6	87–90 or 88–91	KWTD or WTDQ
6597.1	764.3	763.8	87–92	KWTDQS
6752.1	919.3	920.0	86–92	RKWTDQS
6854.7	1021.9	1021.1	85–92	TRKWTDQS
7009.8	1177.0	1177.3	84–92	RTRKWTDQS
6561.1	728.3	727.8	52–56	TEWHR

Masses matching to the SSB sequence are listed in Column 1: Measured m/z -value of the DNA-peptide heteroconjugates. Column 2: Average peptide masses calculated from the experimental data by subtracting the molecular mass of the My9IU-DNA probe and correcting by the mass of hydroiodic acid, eliminated during the cross-linking reaction. Column 3: Theoretical average masses of the peptides shown in column 5 and 6. Column 5: Position of the SSB-peptides matching the calculated mass. Column 6: Sequence of the SSB-peptides matching the calculated mass. The amino acid residues forming the “core” peptide are shown in bold.

of UV-cross-linking, i.e., for identification of the modified amino acid residues.

The heteroconjugate had to be enriched prior to analysis to avoid ionization suppression due to presence of unreacted peptide species (see introduction). Furthermore, the oligonucleotide part of the heteroconjugate had to be as small as possible to ensure that the species under investigation exhibited peptidic properties to facilitate peptide sequencing by ESI MS/MS in the positive ion mode. Our new strategy took advantage of immobilized metal (Fe^{3+}) affinity chromatography (IMAC) resin for the purification of heteroconjugates. Phosphate groups have a high affinity towards Fe(III) -cations (Andersson and Porath 1986), so that IMAC can be used to enrich phosphopeptides from protein digests (Nuwaysir and Stults 1993; Posewitz and Tempst 1999; Stensballe et al. 2001). DNA has a high affinity for the Fe(III) -IMAC resin because of the phosphodiester backbone, and therefore, harsh washing conditions can be applied to remove noncross-linked peptide species from the reaction mixture.

After photochemical cross-linking of the SSB-protein to the DNA probe the entire reaction mixture was incubated with trypsin. The complete tryptic digest mixture was loaded onto a Fe(III) -IMAC microcolumn prepared according to published protocols (Stensballe et al. 2000, 2001). Subsequently, the DNA-containing components were eluted from the column. The DNA was then degraded by using phosphodiesterase I to reduce the size of the oligonucleotide part of the conjugates and eliminate free DNA (Golden et al. 1999). Calf intestine alkaline phosphatase was added to the reaction mixture at the same time to remove any 5'-phosphate groups (Rieger et al. 2000). This method generated heteroconjugates consisting of tryptic peptides cross-linked to di- and trinucleotides.

A tandem purification setup consisting of a POROS R2 and an OLIGO R3 microcolumn (Neubauer et al. 1999) was used to remove salts and nucleosides from the enzyme-

treated sample. This arrangement minimizes sample losses during desalting because small and hydrophilic peptides, which are not retained by the POROS R2-column, will be trapped by the OLIGO R3 material.

Nanoelectrospray MS analysis of the POROS R2-column eluate, i.e., the more hydrophobic analytes, revealed some minor species that were identified by MS/MS sequencing as tryptic peptides derived from SSB (data not shown). The nanoelectrospray mass spectrum of the eluate from the OLIGO R3-column, i.e., the more hydrophilic species, showed one major ion signal and several minor ion signals, the latter corresponding to copurified tryptic SSB peptides. The major signal corresponded to a triply charged ion at m/z 779.62 (Fig. 5A). The fact that a species with a mass of approximately 2.3 kDa was not retained by the POROS R2 material suggested that the compound was much more hydrophilic than regular peptides of this size. Additional evidence for an unusual peptide came from its exact mass of 2335.8 Da. The fractional mass value of 0.8 at a mass of 2.3 kDa indicated a high ratio of mass-deficient atoms such as oxygen, phosphorus, and sulfur in this species (Mann 1995; Lehmann et al. 2000). Because this experiment was performed on a nanoelectrospray quadrupole TOF-type instrument, where the calibration is not affected by the type of analyte (see above), accuracies of less than 25 ppm are expected even with external calibration.

Fragmentation of the major ion species at m/z 779.61 at a CID energy setting normally used for peptide species of this size (approximately 60 eV collision energy in the laboratory reference frame [LRF; Shukla and Futrell 2000]) produced a fragmentation pattern lacking the characteristics of peptide MS/MS spectra (Fig. 5B). This first fragmentation regime was dominated by product ions that could be rationalized as oligonucleotide-related w_1 (3+: 595.26 and 2+: 892.38) and w_2 fragment ions (3+: 704.94 and 2+: 1056.91) using the nomenclature according to McLuckey et al. (1992). These fragments originated from the trinucleotide

Fig. 5.

TpGpX, where X denotes the cross-linked nucleotide (the former 5-iodouracil). This sequence stretch matched to the nucleotides 7–9 of the DNA 19-mer used in this study. In addition, fragment ion signals corresponding to the nucleobase-peptide heteroconjugate were observed, which were formed upon cleavage of an N-glycosidic bond of the nucleotide (marked with a triangle). No fragment ions originating from the peptide part of the heteroconjugate were observed in this spectrum. Localization of the protein cross-linking site was, therefore, not possible when using standard CID conditions. To observe peptide backbone cleavage, the heteroconjugate molecular ions had to undergo more than one fragmentation event in two different fragmentation regimes (see Fig. 5D): (1) cleavage of the trinucleotide, and (2) cleavage of the peptide backbone. Peptide backbone fragmentation was induced by increasing the collision energy to approximately 120 eV(LRF). Figure 5c shows the MS/MS spectrum of the triply charged species at m/z 779.61 acquired under increased collision energy settings. The spectrum was dominated by small molecular weight fragment ions, probably originating from the trinucleotide. In addition, several peptide-derived fragment ions were observed in the higher m/z range that identified the peptide as 85-TRKWTDQSGQDR-96. The two fragments at m/z 906.4 and m/z 1202.5, corresponding to the y_8^- and y_9^- -ions, respectively, revealed the exact site of modification to be Trp-88.

A threefold increase in the amount of starting material allowed the analysis of the second heteroconjugate by nano-electrospray MS/MS. On the basis of the product ion spectrum (Fig. 5E) it was possible to identify Trp-54 as the second amino acid residue that was UV-cross-linked to the DNA probe.

The fact that the CID spectrum was dominated by fragmentation of the trinucleotide was not surprising, considering that DNA is built from carbohydrates and alkylphosphates. This situation is similar to CID product ion spectra obtained from glycopeptides and alkylphosphates containing peptides. In both cases, loss of the modifying moiety is often the predominant fragmentation pathway due to the

lability of the corresponding bonds (Medzihradszky et al. 1990; Covey et al. 1991). Golden et al. (1999) observed similar fragmentation behavior of a heteroconjugate consisting of a nonapeptide with a dinucleotide attached to a tyrosine side chain.

The problem related to the fragmentation behavior and lability of DNA-peptide heteroconjugates can be circumvented in some cases. DNA-binding proteins that do not have any specific DNA sequence requirement for binding, i.e., any DNA can act as a ligand, will bind and cross-link to poly(dT) (Jensen et al. 1996; Gafken et al. 2000). Because thymine is a nonbasic nucleobase, the affinity for protons is very low, and, therefore, polyT or T-rich oligonucleotides are much more stable under positive ion CID conditions than mixed base oligonucleotides (Wang et al. 1997; Vrkic et al. 2000). The MS/MS spectra of polyT-peptide heteroconjugates demonstrate that peptide backbone fragmentations occur with a similar yield as the fragmentation of the oligonucleotide attached.

The involvement of Trp-88 of SSB in DNA-binding was until recently under discussion. Early results from fluorescence and optical detection of triplet state magnetic resonance spectroscopy indicated that Trp-88 was not involved (Khamis et al. 1987). Later, fluorescence and site-directed mutagenesis experiments suggested that it might be involved in DNA binding (Curth et al. 1993). Last year the crystal structure of the DNA-binding domain of SSB bound to a poly(dC) 35-mer was solved by X-ray diffraction (Raghunathan et al. 2000). This structure model clearly shows that Trp-88 contributes to DNA binding. Trp-88 is located in an exposed loop-structure between two β -sheets. This amino acid residue interacts with nucleobases in two out of the four subunits, which form the DNA-binding homotetrameric protein complex. Trp-88 in one of those two subunits is arranged such that a stacking interaction with one of the nucleobases occur; the average distance between the two π -systems was determined to be around 4.5 Å (3.5–5.7 Å) within the range of π -stacking interactions. Trp-88 of the other subunit is arranged perpendicular to two nucleobases, with a ring centroid separation of less than 6.5 Å, indicating

Fig. 5. Identification of the UV cross-linked amino acid residue in SSB by nanoelectrospray tandem mass spectrometry of the peptide–DNA heteroconjugate. (A) Nanoelectrospray mass spectrum obtained from the Oligo R3 eluate after tryptic digest, IMAC purification, and phosphodiesterase I digest. The major ion signal is attributable to a triply charged species at m/z 779.62. (B) First MS/MS fragmentation regime: product ion spectrum of the triply charged species at m/z 779.62 at a Q_0 setting of 40 V (about 60 eV collision energy in LRF). The major fragments corresponding to the fragmentation of a trinucleotide with the sequence TpGpX are labeled. The fragments corresponding to the loss of the former 5-iodouracil are marked with triangles. No peptide backbone cleavage products are observable. (C) Second MS/MS fragmentation regime: product ion spectrum of the same species at a Q_0 setting of 60 V (about 120 eV collision energy in LRF). The region above m/z 300 is enlarged 46-fold showing fragments deriving from the cleavages of the amide bond. The fragment ions y_8 and y_9 unambiguously demonstrate that W88 was cross-linked to the DNA. Sequence specific fragment ions are labeled. (D) MS/MS fragmentation scheme. The upper part depicts the MS/MS fragmentation of the dodecapeptide–trinucleotide heteroconjugate at a collision energy around 60 eV: nucleobase one-letter code are in bold italic face; oligonucleotide fragment ions according to (McLuckey et al. 1992) in bold face; amino acid one-letter code underlined in bold face. The lower part shows the fragmentation of the peptide at a collision energy of about 120 eV. The peptide fragment ions are labeled according to Biemann (1988). (E) Product ion spectrum of the triply charged peptide–trinucleotide heteroconjugate at m/z 615.5, revealing Trp-54 to be the second UV-cross-linked amino acid residue.

that there might be some H- π interaction (Burley and Petsko 1985, 1986). This attractive interaction of two perpendicular π -systems is often observed not only in the crystals of aromatic compounds but also in proteins, where it plays an important role for structural stabilization. In contrast to Trp-88, the involvement of Trp54 in DNA binding has never been disputed because it had been conclusively shown in early studies (Casas-Finet et al. 1987).

Apart from this study there are, to our knowledge, only two other published reports where a tryptophan residue within a protein was found to be involved in photocross-linking to either a 5-bromouracil modified DNA probe (Katouzian-Safadi et al. 1991b) or a 5-iodouracil containing DNA ligand (Kubareva et al. 2000). Considering that Shetlar et al. (1984) demonstrated a high reactivity of tryptophan residue towards DNA upon UV-irradiation, and Zubarev et al. (1999) found that tryptophan residues have a very high reactivity towards radicals when compared to the other natural amino acid residues, one should expect that tryptophan would be a common cross-linking partner for DNA probes. One possible explanation for the rare observation of tryptophan cross-links was provided by Katouzian-Safadi et al. (1991a), who reported photochemical degradation of tryptophan residues upon prolonged UV-irradiation at 300 to 400 nm.

In an earlier SSB-DNA UV-cross-linking study Phe-60 was identified as the site of cross-linking, which was not identified in our study (Merrill et al. 1984). In these experiments 254-nm UV-light was used in combination with octathymidine as a DNA probe, which exhibits eight potential cross-linking sites. Assuming that each thymine is photoactivated to the same degree, cross-linking is controlled by the presence of a reactive amino acid side chain in the direct neighborhood. In contrast to this, in our study only one photo-activatable nucleobase is present, i.e., cross-linking will only occur to amino acid residues that are close in space to this modified nucleobase. The fact that one amino acid residue was identified as the major site of cross-linking indicates that the DNA probe used in this study is bound in such a defined way that the modified, photosensitive nucleotide is close to other residues such as Trp-88 and Trp-54, but not Phe-60.

Conclusions and Perspectives

In this study, four different MS-based methods were used to characterize *E. coli* SSB-protein that was UV-cross-linked to a 5-iodouracil-labeled DNA 19-mer. Three of the methods employed MALDI-MS to demonstrate that the SSB region 87-91 is the major cross-linked site and 52-56 is the minor one. In the fourth approach, using enzyme treatment, microcolumns, and nanoelectrospray MS/MS, we unambiguously showed that Trp-88 and Trp-54 are involved in the formation of the heteroconjugate with DNA. Our data

are in agreement with previous results obtained by other analytical methods (Casas-Finet et al. 1987; Khamis et al. 1987).

Two of the protocols used polyacrylamide gel electrophoresis for the isolation of heteroconjugates from unreacted species in the reaction mixture, whereas the other two protocols included only in-solution sample handling procedures in combination with custom-made microcolumns and mass spectrometry.

Differential peptide mapping with proteolytic digests of protein-DNA heteroconjugates and the protein as control was performed. The clear absence of some peaks in the peptide mass map of the digest mixtures of the heteroconjugate allowed us to conclude (1) that we were working with a near-homogeneous cross-link, and (2) the amino acid residue, which was predominantly involved in SSB-DNA cross-linking, was located within the region 87-96 of the protein.

More direct evidence about the nature of the heteroconjugate was obtained by isolating the peptide-DNA cross-link by urea-polyacrylamide gel electrophoresis after proteolytic digestion. The hybrid DNA-peptide species was eluted from the gel matrix and analyzed by MALDI-TOF MS. Ion signals from two heteroconjugates were observed, indicating a major cross-linking product with the tryptic peptide T₈₅₋₉₆ and a minor one with T₅₀₋₅₆. However, no exact localization of the cross-linked amino acids was obtained.

A fast, simple, and sensitive method to identify a short stretch of amino acids cross-linked to DNA was implemented by treating the crude cross-linking mixture with proteinase K followed by MALDI-TOF MS of the generated sequence ladder. Two pentapeptides were identified to be linked to the DNA-probe: 52-TEWHR-56 and 87-RTRKWTDQST-91. This direct approach worked at the low picomole range (30 pmole protein as starting material), and was completed in a matter of hours.

Nanoelectrospray ionization combined with quadrupole-TOF hybrid tandem mass spectrometry was used for peptide sequencing to pinpoint exactly those amino acid residues involved in the cross-linking reaction with oligonucleotide. Sample preparation was performed by using microcolumns only without any gel electrophoresis or liquid chromatography to purify the heteroconjugates. The data acquired on the quadrupole-TOF tandem mass spectrometer enabled exact assignment of the cross-linking site as Trp-54 and Trp-88 of SSB. Under standard CID conditions fragmentation of a short oligonucleotide, which remained attached to the peptide after nuclease digestion, dominated the MS/MS spectrum of the heteroconjugate. Therefore, higher collision energies than normally used for peptide species were required to induce higher order fragmentation and generating peptide sequence information. Because higher order fragment ions had to be generated, the overall sensitivity for sequencing of

peptide–DNA heteroconjugates was slightly lower than normally obtained for sequencing peptides.

The sensitivity of the presented method is sufficient to allow exact identification of cross-linked amino acid residues with only 100 pmole of protein starting material, even though the cross-linking yield did not exceed 15%. This represents a significant sensitivity improvement over previously published protocols. The information obtained by MS/MS experiments enabled sequencing of both the oligonucleotide and the peptide moiety of the heteroconjugate.

This study and numerous other studies show that the site of cross-linking reflects protein–oligonucleotide interaction sites, and that it is possible to obtain detailed information about pivotal amino acid residues in the oligonucleotide binding site in proteins. Thus, it might be conceivable to use cross-linking strategies with several oligonucleotide probes, where the site of photosensitive nucleobase substitution is successively varied, to model oligonucleotide–protein interaction in cases when structural data of such complexes are not available.

Materials and methods

Materials

Chemicals were obtained from Aldrich or Sigma. High-purity solvents used for nanoelectrospray experiments were purchased from Labscan. Recombinant *E. coli* single-strand DNA-binding protein (SSB) was purchased from Stratagene. The 5-iodouracil (5-IU) containing oligonucleotides TGTAGCTG(5IU)TGATCTAAGT (My9IU), TGTAGCTGT(5IU)GATCTAAGT (My10IU), and TGTAGCTGTTGA(5IU)CTAAGT (My13IU) were custom made by DNA technology. Aliquots of the oligonucleotide were ^{32}P -labeled at their 5'-termini using [γ - ^{32}P]ATP (Amersham Pharmacia) and polynucleotide kinase (Roche Diagnostics) according to the procedure described in the supplier's manual. To remove unincorporated ATP, the reaction mixture was desalted using Micro Bio-Spin 6 Tris Spin Columns (Bio Rad).

Cross-linking

SSB (final concentration: 1.7 μM) and the 5-iodouracil containing oligonucleotide (final concentration: 2 μM) were incubated in 50 mM NH_4HCO_3 and 18 mM NaCl at room temperature for 15 min. Aliquots (30 μL) were irradiated at 0°C for 40 min in the lids of Eppendorf tubes in a UV Stratalinker 1800 (Stratagene) equipped with 312-nm bulbs.

SDS-PAGE, quantification, and in-gel digest

Cross-linking mixture (40 μL) was concentrated in a speedvac and loaded onto 4–12% NuPage gels (Novex) and visualized by colloidal Coomassie Blue staining (Colloidal Blue Staining Kit, Novex). Quantification of the cross-linking yield was done on a PhosphorImager (Storm 840, Molecular Dynamics) using ImageQuant software, version 5.0. In-gel reduction, alkylation, and proteolytic digestion with either trypsin or endoprotease GluC

(both sequencing grade, Roche Diagnostics) were performed as described previously (Shevchenko et al. 1996).

Urea polyacrylamide gel electrophoresis

Trypsin (0.4 μg) was added to 60 μL cross-linking mixture (DNA partially ^{32}P -labeled) and incubated for 24 h at 37°C. An additional 0.4 μg of trypsin were added and the digestion was allowed to proceed for another day. The tryptic digest was reduced to dryness in a speedvac and redissolved in DNA sample buffer (1 $\times$ TBE, 30% formamide, 0.2% bromophenol blue) and loaded onto a 13% urea-PAGE (7 M urea). The positions of the radioactive bands were visualized by autoradiography. The bands containing the oligonucleotide–peptide cross-links were excised and shaken for 24 h in 200 μL H_2O to elute passively the heteroconjugate from the gel plugs. The resulting solution was loaded onto POROS R2 (Perseptive Biosystems) microcolumns, packed in GELoader tips (Eppendorf) as described previously (Wilm 1996; Gobom 1999) and washed with 5% aqueous formic acid. Peptide mass searches against the SSB amino acid sequence were performed using GP-MAW software (Lighthouse Data).

Proteinase K digest

Cross-linking mixture (15 μL) was incubated with 1.4 μL of proteinase K (2 mg/mL; Sigma), for 2 h at 50°C and concentrated and desalted using POROS R2 microcolumns prior to mass spectrometric analysis (see below). A solution of 15 mM triethylammonium acetate solution containing 3% acetonitrile was used for washing.

Isolation of the peptide–DNA heteroconjugate

Cross-linking mixture (50 μL) (cross-linking yield about 15%) containing 85 pmole SSB was incubated with trypsin (0.5 μg) for 48 h at 37°C. An additional 0.5 μg protease was added after 24 h. The entire mixture was loaded onto a micro-Fe(III)-IMAC column prepared in a GELoader tip as described previously (Stensballe et al. 2000). To ensure that only the DNA-containing components of the digestion mixture (unreacted oligonucleotides and peptide–DNA heteroconjugate) were retained on the IMAC column, the column was washed, in turn, with 0.1 M acetic acid, 0.1 M acetic acid and acetonitrile (3:1 v/v), 0.1 M acetic acid and acetonitrile (1:1 v/v), and 0.1 M acetic acid and acetonitrile (1:3 v/v). The retained components were eluted with 50 μL H_2O , pH 10.5 (adjusted with NH_3). After addition of 35 μL 0.1 M Tris-HCl (pH 9.3), 20 μL 1 M acetic acid, 70 μL H_2O , 2 μL 0.1 M MgCl_2 , 2 μL alkaline phosphatase (1 U/ μL , Roche Diagnostics), and 2 μL phosphodiesterase I (snake venom nuclease from *Crotalus adamanteus*, an exonuclease that attacks the free 3'-hydroxyl group yielding 5'-nucleotides; 1 mg/mL according to the supplier's manual, Amersham Pharmacia) the solution was incubated overnight at 37°C. Peptide components in the phosphodiesterase I stock solution were removed by gel-filtration using Micro Bio-Spin 6 Tris Spin Columns (Bio Rad) after changing the buffer in the spin columns to 110 mM NaCl, 15 mM MgCl_2 , 110 mM Tris-HCl (pH 8.9), and 50% glycerol. Due to the high viscosity of this solution centrifugation times had to be doubled compared to those stated in the manual. Salts and nucleosides formed in the course of the digestion with phosphodiesterase and phosphatase were removed by loading the reaction mixture onto tandem microcolumns filled with POROS R2/OLIGO R3 (Perseptive Bio-

systems) as described earlier (Neubauer and Mann 1999) and washed with 5% aqueous formic acid.

Mass spectrometry

All MALDI-TOF mass spectra were acquired in positive ion reflector mode on a Reflex III mass spectrometer (Bruker Daltonik GmbH) equipped with delayed extraction. For differential peptide mapping experiments a minor fraction of the acidified digestion supernatant was prepared by the "thin layer" method with α -Cyano-4-hydroxycinnamic acid (4-HCCA) as matrix according to published procedures (Vorm et al. 1994). Protease autolysis peaks were used for internal calibration. For MALDI-TOF mass spectra of the DNA or the DNA-peptide heteroconjugates the samples were eluted from the microcolumns directly onto the target with 1.5 μ L of 15 mM triethylammonium acetate/30% acetonitrile saturated with 3-hydroxypicolinic acid. The sample was allowed to dry at ambient temperature after addition of a few NH_4^+ -loaded cation exchange beads (Bio-Rad, 50 W $\times$ 8, mesh size 100–200 μ m) as described earlier (Nordhoff et al. 1993). The MALDI mass spectra obtained by using the DNA conditions were externally calibrated using My9IU–DNA as calibrant.

The nano-electrospray mass spectra were acquired in positive ion mode on a QSTAR Pulsar quadrupole-TOF hybrid tandem mass spectrometer (AB/MDS-Sciex) equipped with a nano-electrospray ion source (MDS Proteomics). The samples were eluted from the microcolumns directly into palladium gold-coated nano-electrospray needles (MDS Proteomics) by using 5% formic acid in 60% methanol. For MS/MS-spectra the Q_0 setting (the collision energy setting on QSTAR instruments) was increased to 60 V to ensure the generation of higher order fragment ions that reveal peptide sequence information.

Acknowledgments

The authors thank Dr. Finn Kirpekar (Department of Biochemistry and Molecular Biology, University of Southern Denmark/Odense University, Denmark) for sharing his extensive knowledge about MALDI-TOF MS analysis of DNA and RNA samples, and Morten M. Nedertoft (Department of Biochemistry and Molecular Biology, University of Southern Denmark/Odense University, Denmark) for the help to interpret the X-ray crystal structure of the SSB–DNA complex. We also acknowledge Søren Andersen (Department of Biochemistry and Molecular Biology, University of Southern Denmark/Odense University, Denmark) for technical assistance, and Dr. Philip Gafken (Fred Hutchinson Cancer Research Center, Seattle, WA) and Prof. Douglas Barofsky (Department of Chemistry, Oregon State University, OR) for helpful discussions and for communication of experimental details prior to publication. Parts of this research were supported by a grant from the Carlsberg Foundation (Denmark). The work was part of the activities of the Center for Experimental BioInformatics (CEBI) sponsored by the Danish National Research Foundation.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Allen, T.D., Wick, K.L., and Matthews, K.S. 1991. Identification of amino acids in lac repressor protein cross-linked to operator DNA specifically substituted with bromodeoxyuridine. *J. Biol. Chem.* **266**: 6113–6119.

- Andersson, L. and Porath, J. 1986. Isolation of phosphoproteins by metal (Fe^{3+}) affinity chromatography. *Anal. Biochem.* **154**: 250–254.
- Bennett, S.E., Jensen, O.N., Barofsky, D.F., and Mosbaugh, D.W. 1994. UV-catalyzed cross-linking of *Escherichia coli* uracil–DNA glycosylase to DNA. Identification of amino acid residues in the single-stranded DNA binding site. *J. Biol. Chem.* **269**: 21870–21879.
- Biemann, K. 1988. Contributions of mass spectrometry to peptide and protein structure. *Biomed. Environ. Mass Spectrom.* **16**: 99–111.
- Blatter, E., Ebright, Y., and Ebright, R. 1992. Identification of an amino acid-base contact in the GCN4–DNA complex by bromouracil-mediated photocrosslinking. *Nature* **359**: 650–652.
- Burley, S.K. and Petsko, G.A. 1985. Aromatic–aromatic interaction: A mechanism of protein structure stabilization. *Science* **229**: 23–28.
- . 1986. Dimerization energetics of benzene and aromatic amino acid side chains. *J. Am. Chem. Soc.* **108**: 7995–8001.
- Casas-Finet, J.R., Khamis, M.I., Maki, A.H., and Chase, J.W. 1987. Tryptophan 54 and phenylalanine 60 are involved synergistically in the binding of *E. coli* SSB protein to single-stranded polynucleotides. *FEBS Lett.* **220**: 347–352.
- Chen, J., Qi, Y., Zhao, R., Zhou, G.W., and Zhao, Z.J. 2001. Assay of protein tyrosine phosphatases by using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Biochem.* **292**: 51–58.
- Connor, D.A., Falick, A.M., Young, M.C., and Shetlar, M.D. 1998. Probing the binding region of the single-stranded DNA-binding domain of rat DNA polymerase beta using nanosecond-pulse laser-induced cross-linking and mass spectrometry. *Photochem. Photobiol.* **68**: 299–308.
- Covey, T., Shushan, B., Bonner, R., Schröder, W., and Hucho, F. 1991. LC/MS and LC/MS/MS screening for the aites of post-translational modifications in proteins. In *Methods in protein sequence analysis* (eds. H. Jönvall, J.O. Höög, and A.M. Gustavsson), pp. 249–256. Birkhäuser Verlag, Basel.
- Curth, U., Greipel, J., Urbanke, C., and Maass, G. 1993. Multiple binding modes of the single-stranded DNA binding protein from *Escherichia coli* as detected by tryptophan fluorescence and site-directed mutagenesis. *Biochemistry* **32**: 2585–2591.
- Dietz, T.M. and Koch, T.H. 1987. Photochemical coupling of 5-bromouracil to tryptophan, tyrosine, and histidine, peptidic derivatives in aqueous fluid solution. *Photochem. Photobiol.* **46**: 971–978.
- Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**: 64–71.
- Gafken, P.R., Mosbaugh, D.W., and Barofsky, D.F. (2000). Characterization of the UV-crosslinked *E. coli* uracil–DNA glycosylase–dT20 nucleoprotein interface by MALDI MS and ESI MS/MS. Proc. 48th ASMS Conf. Mass Spectrom. Allied Top. Long Beach, CA, pp. 54–55.
- Glückmann, M. and Karas, M. 1999. The initial ion velocity and its dependence on matrix, analyte and preparation method in ultraviolet matrix-assisted laser desorption/ionization. *J. Mass Spectrom.* **34**: 467–477.
- Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., and Roepstorff, P. 1999. Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* **34**: 105–116.
- Golden, M., Resing, K., Collins, B., Willis, M., and Koch, T. 1999. Mass spectral characterization of a protein–nucleic acid photocrosslink. *Protein Sci.* **8**: 2806–2812.
- Green, L.S., Jellinek, D., Jenison, R., Ostman, A., Heldin, Ch., and Janjic, N. 1996. Inhibitory DNA ligands to platelet-derived growth factor b-chain. *Biochemistry* **35**: 14413–14424.
- Hicke, B., Willis, M., Koch, T., and Cech, T. 1994. Telomeric protein–DNA point contacts identified by photo-cross-linking using 5-bromodeoxyuridine. *Biochemistry* **33**: 3364–3373.
- Holz, B., Dank, N., Eickhoff, J., Lipps, G., Krauss, G., and Weinhold, E. 1999. Identification of the binding site for the extrahelical target base in 6-adenine DNA methyltransferases by photo-cross-linking with duplex oligodeoxyribonucleotides containing 5-iodouracil at the target position. *J. Biol. Chem.* **274**: 15066–15072.
- Ito, S., Saito, I., and Matsuura, T. 1980. Acetone-sensitized photocoupling of 5-bromouridine to tryptophan derivatives via electron-transfer process. *J. Am. Chem. Soc.* **102**: 7535–7541.
- Jensen, O.N., Barofsky, D.F., Young, M.C., Von Hippel, P.H., Swenson, S., and Seifried, S.E. 1993. Direct observation of UV-crosslinked protein–nucleic acid complexes by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **7**: 496–501.
- Jensen, O.N., Kulkarni, S., Aldrich, J.V., and Barofsky, D.F. 1996. Characterization of peptide–oligonucleotide heteroconjugates by mass spectrometry. *Nucleic Acids Res.* **24**: 3866–3872.

- Jensen, O.N., Shevchenko, A., and Mann, M. 1997. Protein analysis by mass spectrometry. In *Protein structure—A practical approach* (ed. T.E. Creighton), pp. 29–57. IRL Press, Oxford, UK.
- Karas, M. and Hillenkamp, F. 1988. Laser desorption/ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60**: 2299–2301.
- Katouzian-Safadi, M., Blazy, B., and Charlier, M. 1991a. Photochemical cross-linking of the cyclic adenosine 3',5' monophosphate receptor protein to *Escherichia coli* 5-bromouracil-substituted DNA. Role of the effectors. *Photochem. Photobiol.* **53**: 611–616.
- Katouzian-Safadi, M., Laine, B., Chartier, F., Cremet, J.Y., Belaiche, D., Sautiere, P., and Charlier, M. 1991b. Determination of the DNA-interacting region of the archaeobacterial chromosomal protein MCI. Photocrosslinks with 5-bromouracil-substituted DNA. *Nucleic Acids Res.* **19**: 4937–4941.
- Khamis, M.I., Casas-Finet, J.R., Maki, A.H., Murphy, J.B., and Chase, J.W. 1987. Investigation of the role of individual tryptophan residues in the binding of *Escherichia coli* single-stranded DNA binding protein to single-stranded polynucleotides. A study by optical detection of magnetic resonance and site-selected mutagenesis. *J. Biol. Chem.* **262**: 10938–10945.
- Kubareva, E.A., Thole, H., Karyagina, A.S., Oretskaya, T.S., Pingoud, A., and Pingoud, V. 2000. Identification of a base-specific contact between the restriction endonuclease SsoII and its recognition sequence by photocross-linking. *Nucleic Acids Res.* **28**: 1085–1091.
- Kusmann, M., Nordhoff, E., Rahbek-Nielsen, H., Haebel, S., Rossel-Larsen, M., Jakobsen, L., Gobom, J., Mirgorodskaya, E., Kroll-Kristensen, A., Palm, L. et al. 1997. Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes. *J. Mass Spectrom.* **36**: 593–601.
- Lehmann, W.D., Bohne, A., and Von der Lieth, C.W. 2000. The information encrypted in accurate peptide masses—improved protein identification and assistance in glycopeptide identification and characterization. *J. Mass Spectrom.* **35**: 1335–1341.
- Malkov, V.A. and Camerini-Otero, R.D. 1995. Photocross-links between single-stranded DNA and *Escherichia coli* RecA protein map to loops L1 (amino acid residues 157–164) and L2 (amino acid residues 195–209). *J. Biol. Chem.* **270**: 30230–30233.
- Mann, M. 1995. Useful tables of possible and probable peptide masses. Proc. 43th ASMS Conf. Mass Spectrom. Allied Top. Atlanta, GA, p. 639.
- McLuckey, S., Van Berkel, G.J., and Glush, G.L. 1992. Tandem mass spectrometry of small multiply charged oligonucleotides. *J. Am. Soc. Mass Spectrom.* **3**: 60–70.
- Medzhradszky, K.F., Gillece-Castro, B.L., Settineri, C.A., Townsend, R.R., Masiarz, F.R., and Burlingame, A.L. 1990. Structure determination of o-linked glycopeptides by tandem mass spectrometry. *Biomed. Environ. Mass Spectrom.* **19**: 777–781.
- Meisenheimer, K. and Koch, T. 1997. Photocross-linking of nucleic acids to associated proteins. *Crit. Rev. Biochem. Mol. Biol.* **32**: 101–140.
- Merrill, B.M., Williams, K.R., Chase, J.W., and Konigsberg, W.H. 1984. Photochemical cross-linking of the *Escherichia coli* single-stranded DNA-binding protein to oligodeoxynucleotides. Identification of phenylalanine 60 as the site of cross-linking. *J. Biol. Chem.* **259**: 10850–10856.
- Meyer, R.R. and Laine, P.S. 1990. The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol. Rev.* **54**: 342–380.
- Musier-Forsyth, K. and Schimmel, P. 1994. Acceptor helix interactions in a class II tRNA synthetase: Photoaffinity cross-linking of an rna miniduplex substrate. *Biochemistry* **33**: 773–779.
- Neubauer, G. and Mann, M. 1999. Mapping of phosphorylation sites of gel-isolated proteins by nanoelectrospray tandem mass spectrometry: Potentials and limitations. *Anal. Chem.* **71**: 235–242.
- Nordhoff, E., Cramer, R., Karas, M., Hillenkamp, F., Kirpekar, F., Kristiansen, K., and Roepstorff, P. 1993. Ion stability of nucleic acids in infrared matrix-assisted laser desorption/ionization mass spectrometry. *Nucleic Acids Res.* **21**: 3347–3357.
- Norris, C., Meisenheimer, P., and Koch, T. 1996. Mechanistic studies of the 5-iodouracil chromophore relevant to its use in nucleoprotein photo-cross-linking. *J. Am. Chem. Soc.* **118**: 5796–5803.
- Nuwaysir, L. and Stults, J.T. 1993. Electrospray ionization mass spectrometry of phosphopeptides isolated by on-line immobilized metal-ion affinity chromatography. *J. Am. Soc. Mass Spectrom.* **4**: 662–669.
- Pingoud, V., Thole, H., Christ, F., Grindl, W., Wende, W., and Pingoud, A. 1999. Photocross-linking of the homing endonuclease PI-SceI to its recognition sequence. *J. Biol. Chem.* **274**: 10235–10243.
- Posewitz, M.C. and Tempst, P. 1999. Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal. Chem.* **71**: 2883–2892.
- Prasad, R., Kumar, A., Widen, S.G., Casas-Finet, J.R., and Wilson, S.H. 1993. Identification of residues in the single-stranded DNA-binding site of the 8-kDa domain of rat DNA polymerase beta by UV cross-linking. *J. Biol. Chem.* **268**: 22746–22755.
- Qin, J. and Chait, B. 1997. Identification and characterization of posttranslational modifications of proteins by MALDI ion trap mass spectrometry. *Anal. Chem.* **69**: 4002–4009.
- Ragunathan, S., Kozlov, A.G., Lohman, T.M., and Waksman, G. 2000. Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nat. Struct. Biol.* **7**: 648–652.
- Rieger, R., McTigue, M., Kycia, J., Gerchman, S., Grollman, A., and Iden, C. 2000. Characterization of a cross-linked DNA–endonuclease VIII repair complex by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **11**: 505–515.
- Shetlar, M.D., Christensen, J., and Hom, K. 1984. Photochemical addition of amino acids and peptides to DNA. *Photochem. Photobiol.* **39**: 125–133.
- Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* **68**: 850–858.
- Shukla, A.K. and Futrell, J.H. 2000. Tandem mass spectrometry: Dissociation of ions by collisional activation. *J. Mass Spectrom.* **35**: 1069–1090.
- Stensballe, A., Andersen, S., and Jensen, O.N. 2001. Characterization of phosphoproteins from electrophoretic gels by nano-scale Fe(III) affinity chromatography with off-line mass spectrometry analysis. *Proteomics* **1**: 207–222.
- Stensballe, A., Jensen, O.N., Olsen, J.V., Haselmann, K.F., and Zubarev, R.A. 2000. Electron capture dissociation of singly and multiply phosphorylated peptides. *Rapid Commun. Mass Spectrom.* **14**: 1793–1800.
- Stump, W.T. and Hall, K.B. 1995. Crosslinking of an iodo-uridine–RNA hairpin to a single site on the human U1A N-terminal RNA binding domain. *RNA* **1**: 55–63.
- Urlaub, H., Kruff, V., Bischof, O., Muller, E.C., and Wittmann-Liebold, B. 1995. Protein–rRNA binding features and their structural and functional implications in ribosomes as determined by cross-linking studies. *EMBO J.* **14**: 4578–4588.
- Urlaub, H., Thiede, B., Muller, E.C., Brimacombe, R., and Wittmann-Liebold, B. 1997. Identification and sequence analysis of contact sites between ribosomal proteins and rRNA in *Escherichia coli* 30S subunits by a new approach using matrix-assisted laser desorption/ionization-mass spectrometry combined with n-terminal microsequencing. *J. Biol. Chem.* **272**: 14547–14555.
- Vorm, O., Roepstorff, P., and Mann, M. 1994. Improved resolution and very high sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. *Anal. Chem.* **66**: 3281–3287.
- Vrkic, A.K., O'Hair, R.A.J., Foote, S., and Reid, G.E. 2000. Fragmentation reactions of all 64 protonated trimer oligodeoxynucleotides and 16 mixed base tetramer oligodeoxynucleotides via tandem mass spectrometry in an ion trap. *Int. J. Mass Spectrom.* **194**: 145–164.
- Wang, P., Batlett, M.G., and Martin, L.B. 1997. Electrospray collision-induced dissociation mass spectra of positively charged oligonucleotides. *Rapid Commun. Mass Spectrom.* **11**: 846–856.
- Wang, Q., Shoeman, R., and Traub, P. 2000. Identification of the amino acid residues of the amino terminus of vimentin responsible for DNA binding by enzymatic and chemical sequencing and analysis by maldi-tof. *Biochemistry* **39**: 6645–6651.
- Wang, Y. and Adzuma, K. 1996. Differential proximity probing of two DNA binding sites in the *Escherichia coli* recA protein using photo-cross-linking methods. *Biochemistry* **35**: 3563–3571.
- Wilm, M. and Mann, M. 1996. Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* **68**: 1–8.
- Wind, M., Edler, M., Jakubowski, N., Linscheid, M., Wesch, H., and Lehmann, W.D. 2001. Analysis of protein phosphorylation by capillary liquid chromatography coupled to element mass spectrometry with 31p detection and to electrospray mass spectrometry. *Anal. Chem.* **73**: 29–35.
- Wong, D.L. and Reich, N.O. 2000. Identification of tyrosine 204 as the photo-cross-linking site in the DNA–ecori DNA methyltransferase complex by electrospray ionization mass spectrometry. *Biochemistry* **39**: 15410–15417.
- Zubarev, R.A., Kruger, N.A., Fridriksson, E.K., Lewis, M.A., Horn, D.M., Carpenter, B.K., and McLafferty, F.W. 1999. Electron capture dissociation of gaseous multiply-charged proteins is favored at disulfide bonds and other sites of high hydrogen atom affinity. *J. Am. Chem. Soc.* **121**: 2857–2862.