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

Qiang Wang, Robert Shoeman,* and Peter Traub*

Max-Planck-Institut für Zellbiologie, Rosenhof, 68526 Ladenburg, Germany

Received February 1, 2000; Revised Manuscript Received April 3, 2000

ABSTRACT: The amino acid residues responsible for stable binding of nucleic acids by the intermediate filament (IF) subunit protein vimentin were identified by a combination of enyzmatic and chemical ladder sequencing of photo-cross-linked vimentin−oligodeoxyribonucleotide complexes and analysis by MALDI-TOF mass spectrometry. Three tryptic peptides of vimentin (vim_{28-35}, vim_{36-49}, and vim_{50-63}) were found to be cross-linked to oligo(dG·BrdU)$_{12}$·dG·3′-FITC. From a methodological standpoint, it was necessary to remove the bulk of the bound oligonucleotide by digestion with nuclease P1 to get reproducible spectra for most of the peptides studied. Additionally, removal of the phosphate group of the residually bound dUMP or modification of the amino terminus of the peptide−oligonucleotide complexes with dimethylaminoazobenzene isothiocyanate dramatically improved the quality of the MALDI-TOF spectra obtained, particularly for the vim_{28-35} peptide. A single Tyr residue within each of these peptides (Tyr_{29}, Tyr_{37}, and Tyr_{52}) was unequivocally demonstrated to be the unique site of cross-linking in each peptide. These three Tyr residues are contained within the two β-ladder DNA-binding wings proposed for the middle of the vimentin non-α-helical head domain. The experimental approach described should be generally applicable to the study of protein−nucleic acid interactions and is currently being employed to characterize the DNA-binding sites of several other IF subunit proteins.

The intermediate filament (IF)[1] subunit protein vimentin binds nucleic acids via its non-α-helical head domain (*1*, *2*). We have shown that several types of nucleic acids are capable of transporting vimentin into the nucleus in vivo (*3*). This result lends considerable support to our hypothesis that vimentin may be involved in the regulation of gene expression and other DNA-based nuclear events (*4* and references cited therein). While a molecular biology and biochemical approach to defining the binding site for nucleic acids of vimentin was successful in delineating the site to the middle of the head domain (*5*), the individual residue(s) within the two DNA-binding wings (*1*, *5*) responsible for this binding was (were) not identified. We therefore chose to identify those residues of vimentin directly involved in binding of nucleic acids using a protein biochemical approach. Photochemical cross-linking (*6*, *7*) of a BrdU-containing oligodeoxyribonucleotide permitted the production and isolation of stable, covalently linked vimentin−oligonucleotide complexes. Since the non-α-helical head domain of vimentin contains 13 Arg, 2 Lys, and only 4 Pro residues (*1*), a variety of tryptic peptides can be generated from the head domain. These peptides are within a size range that permits accurate mass determination in a matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer (*8*). MALDI-TOF analysis has been employed previously for the characterization of peptide−DNA complexes (*9*, *10*), but no sequencing was performed in these experiments. The contact sites of ribosomal proteins and ribosomal RNA were identified after cross-linking by a combination of automated N-terminal amino acid microsequencing for the peptide moiety and by chemical and enzymatic cleavage of the RNA followed by MALDI-TOF analysis (*11*). Therefore, we thought it likely that MALDI-TOF analysis of these tryptic peptide−oligonucleotide complexes, in combination with enzymatic or chemical sequencing, might yield useful information concerning the site(s) of oligonucleotide binding. In this study, we describe a combination of methods of photochemical cross-linking, chemical and enzymatic peptide sequencing, and peptide sequence analysis via conventional MALDI-TOF mass spectrometry that permits the unequivocal identification of vimentin amino acid residues cross-linked to a bound oligonucleotide.

EXPERIMENTAL PROCEDURES

Materials. Recombinant mouse vimentin was prepared as described (*5*). Oligo(dG·BrdU)$_{12}$·dG·3′-fluorescein isothiocyanate (FITC) and other BrdU-containing oligonucleotides were synthesized as described (*12*) using specialty reagents from Glen Research (Eurogentec, Seraing, Belgium). Oligonucleotides were stored at 4 °C in the dark in 10 mM Tris-HCl, 0.1 mM EDTA, pH 7.4, and concentrations were deter-

* Corresponding authors. Telephone: +49-6203-106-200. Fax: +49-6203-106-122. Email: ptraub@zellbio.mpg.de, rshoeman@zellbio.mpg.de.

[1] Abbreviations: DABITC, dimethylaminoazobenzene isothiocyanate; FITC, fluorescein isothiocyanate; HPLC, high-pressure liquid chromatography; IF, intermediate filament; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; *m*/*z*, mass/charge; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid.

10.1021/bi000199s CCC: $19.00
Published on Web 05/12/2000

mined by using a constant factor of 1 mg/mL = 20 OD_{260} for all nucleic acids (*2*). Enzymes, standard peptides, matrix chemicals, and reagents for Edman degradation were purchased from Sigma−Aldrich (Deisenhofen, Germany), Roche Molecular Biochemicals (Mannheim, Germany), Promega (Mannheim, Germany), Amersham−Pharmacia Biotech (Freiburg, Germany), or New England Biolabs (Frankfurt, Germany). Other reagents were as described in the references or were of reagent grade from Merck (Darmstadt, Germany). UV light (312 nm) was supplied by a VL-215M light (2 × 15 W-312 tubes), and UV fluence was measured with a VLX-3W radiometer equipped with a CX-312 filter (light and radiometer were both from Bioblock Scientific, Illkirch, France). Disposable, reverse-phase mini-columns (Ziptip) for peptide purification were purchased from Millipore (Eschenborn, Germany) and used according to their instructions.

Methods. Photochemical cross-linking (*6*, *7*) of vimentin and BrdU-containing oligonucleotides was performed in analytical and preparative scales. For analysis, 8 μg of vimentin and 1.5 μg of oligonucleotide in 30 μL of buffer I (10 mM Tris-HCl, 1 mM dithiothreitol, 1 mM EDTA, pH 7.4) were incubated alone, or separately with NaCl added to final concentrations of 50−250 mM, for 30 min at 37 °C. The samples were then placed on ice and irradiated with UV light (312 nm) for various times (5−60 min) at a distance of 16 cm from the UV light source (UV fluence ∼0.15 J cm^{-2} min^{-1}). Reaction products were separated by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE)(*13*) and analyzed after staining with Coomassie Brilliant blue (*14*) to visualize vimentin or illumination with 312 nm light to visualize the FITC-labeled products. Preparative reactions contained ∼1:1 molar ratios of vimentin (5.4 mg) and oligonucleotide (1 mg) in 20 mL of buffer I, with added NaCl as indicated under Results, and were incubated and illuminated for 30 min as described for the analytical scale reactions. Cross-linked vimentin was isolated by preparative SDS−PAGE. The band of interest was cut out from the gel, the gel was crushed, and the vimentin−oligonucleotide complex was eluted with multiple changes of buffer I containing 2 M urea. The eluate was dialyzed against buffer I and concentrated to 1 mL by vacuum rotoevaporation. The quantity of oligo$(dG{\cdot}BrdU)_{12}{\cdot}dG{\cdot}3'$-FITC bound to vimentin was measured with a spectrofluorometer (Aminco SPF-500), using the unreacted oligo$(dG{\cdot}BrdU)_{12}{\cdot}dG{\cdot}3'$-FITC as a calibration standard.

Digestion with Trypsin and Isolation of Oligonucleotide Cross-Linked Vimentin Peptides. Sequencing grade, modified trypsin (Promega) was added directly to the concentrated, cross-linked vimentin solution at a ratio of 1:200 (w/w, trypsin/vimentin), and the mixture was incubated for 18 h at 37 °C. The extent of the digestion was confirmed by analytical denaturing gel electrophoresis. The peptides containing the cross-linked oligo$(dG{\cdot}BrdU)_{12}{\cdot}dG{\cdot}3'$-FITC were isolated by preparative denaturing gel electrophoresis with a buffer containing 89 mM Tris, 89 mM boric acid, 2 mM EDTA, and 7 M urea, excised from the gel, and eluted with buffer I. Each peptide band, designated hereafter upper peptide band and lower peptide band, was separately purified by chromatography on a Mono Q (HR 5/5) column (Amersham−Pharmacia), using a NaCl gradient for elution. The peptide fractions were dialyzed against 10 mM sodium acetate, pH 5.3, concentrated, and exhaustively digested with nuclease P1. Where indicated, cross-linked peptides were also treated with alkaline phosphatase. Nuclease P1-treated peptides were further purified by chromatography using a LKB HPLC system and a Nucleosil 120 C18 column (K. Ziemer, Mannheim, Germany), equilibrated with 0.1% trifluoroacetic acid (TFA) and eluted with an acetonitrile gradient (0−60% in 0.1% TFA over 40 min). Cross-linked peptides were identified by their mass in MALDI-TOF analysis (see below).

Peptide Ladder Sequencing. Enzymatic sequencing was performed with aminopeptidase M or carboxypeptidase Y (Boehringer Mannheim, Roche Molecular Biochemicals), essentially according to standard protocols supplied by the manufacturer. It was necessary to carefully titrate each enzyme for each peptide to achieve the desired extent of digestion (*15*). Manual Edman degradation, using phenyl isothiocyanate in place of dimethylaminoazobenzene isothiocyanate (DABITC), was performed essentially as described (*16*). Some peptides were additionally modified with DABITC to permit their facile detection in MALDI-TOF analysis.

MALDI-TOF Analysis. MALDI-TOF analysis was performed in the linear, high mass, positive ion mode with pulsed (time-delayed) extraction using a Kratos Maldi IV instrument (Shimadzu Deutschland, Duisburg, Germany). Initial experiments were performed in the linear, high mass, positive ion mode using a Kratos Maldi III instrument (Shimadzu Deutschland). Most peptides, particularly following enzymatic peptide ladder sequencing, were purified and concentrated on Ziptips. Samples (usually in 0.1% TFA) either were applied via a sandwich technique (in which 0.7 μL of matrix was dried onto the sample spot, followed by 0.7 μL of sample and then another 0.7 μL of matrix) or, for more concentrated samples, were simply mixed 1:10 with the matrix solution and 0.7 μL of this mixture was dried onto the stainless-steel sample holder. Matrix solutions were usually α-cyano-4-hydroxycinnamic acid or sinapinic acid [either reagent dissolved at 10 mg/mL in 50% acetonitrile, 50% 0.1% TFA (all % v/v)]. Spectra were calibrated using near-external standards consisting of a mixture of substance P, fragment 1−4 (M+H *m*/*z* 497.6), angiotensin II (M+H *m*/*z* 1047.2), angiotensin I (M+H *m*/*z* 1297.5), angiotensinogen, fragment 1−13 (M+H *m*/*z* 1646.9), and oxidized insulin B chain (M+H *m*/*z* 3496.9) for samples with α-cyano-4-hydroxycinnamic acid matrix or a mixture of substance P, fragment 1−4, angiotensinogen, fragment 1−13, oxidized insulin B chain, and angiotensin, fragment 1−7 (M+H *m*/*z* 900.0) for samples with sinapinic acid. All of the data presented in the figures in this paper were obtained using α-cyano-4-hydroxycinnamic acid as matrix, except for those in the insets of Figures 1 and 2, for which sinapinic acid was employed. Absolute *m*/*z* values occasionally varied by up to 1 Da depending on the individual calibration and the distance and time between individual samples or measurements; for these samples, spectra were recalibrated using the known *m*/*z* values of the largest peak(s). Spectra were collected and analyzed using the standard Kratos software (Sun OS, Release 5.4, OpenWindows Ver. 3.4, Kratos Kompact Software Ver. 5.2.0) and were usually the average of 50−100 individual laser shots across the width of the sample spot. Data were smoothed and baseline-corrected, generally with a window width of 30 channels. Reported values of *m*/*z* are for the protonated species (M+H) observed in the spectra. Each individual spectrum was plotted on a

relative scale of 100% intensity (i.e., the highest peak in the visible portion of each spectrum was set to 100% intensity), which is indicated on each figure on the *y* axis as %Int. The spectra were exported to a PC version of the Kratos software (Ver. 1.2.1) for preparation of the final figures.

RESULTS

Cross-Linking of Oligo(dG·BrdU)$_{12}$·dG·3′-FITC to Vimentin and Isolation of Peptide−Oligonucleotide Complexes. Analytical scale cross-linking reactions were employed to determine the optimum time and salt concentrations for formation of vimentin−oligo(dG·BrdU)$_{12}$·dG·3′-FITC complexes. Reactions illuminated with 312 nm UV light for 30 min were found to represent the best tradeoff between coupling efficiency and suppression of vimentin−oligo(dG·BrdU)$_{12}$·dG·3′-FITC−vimentin multimers (data not shown). The extent of cross-linking was diminished in direct proportion to the NaCl concentration over the range of 0−250 mM, with about 40% efficiency (relative to no added NaCl) at 75 mM NaCl and about 5% efficiency at 250 mM (data not shown). Vimentin−oligo(dG·BrdU)$_{12}$·dG·3′-FITC complexes were isolated from large-scale reactions (incubated without added NaCl): an average total of 11% (n = 3, range 10.2−12%) of the input oligo(dG·BrdU)$_{12}$·dG·3′-FITC was obtained covalently bound to vimentin. Other preparations (see below) were isolated from reaction mixtures containing 150 mM NaCl.

The vimentin−oligo(dG·BrdU)$_{12}$·dG·3′-FITC complexes were digested to completion with trypsin, and the peptides obtained were analyzed by gel electrophoresis. This approach was designed to isolate all cross-linked peptides regardless of their site(s) of origin and without bias based on previous experimental results (*5*). Two fluorescent peptide bands were identified and the materials subsequently isolated by denaturing gel electrophoresis. These upper and lower peptide−oligo(dG·BrdU)$_{12}$·dG·3′-FITC complexes were further purified by FPLC on a Mono Q column; each preparation yielded a single peak of peptide−oligonucleotide complex (data not shown). These peptide−oligonucleotide complexes were initially analyzed directly via MALDI-TOF in a variety of matrixes in both positive and negative ion modes. No interpretable spectra were obtained (data not shown), presumably due to the complexities inherent in analyzing an oligo(dG·BrdU)$_{12}$·dG·3′-FITC-containing molecule. Therefore, the complexes were treated with nuclease P1, which should remove most of the oligonucleotide, leaving (ideally) a dUMP residue (*m*/*z* 307) attached to the vimentin peptide. The upper band yielded two major mass peaks, *m*/*z* 1752 and 1803 (Figure 1, lower spectrum). Subtracting *m*/*z* 307 (the mass of dUMP) yields peptide masses of 1445 and 1496, respectively, which correspond to those predicted for vimentin tryptic peptides vim$_{50-63}$ and vim$_{36-49}$. In some preparations of the upper peptide band, additional peaks were found at *m*/*z* 2081 and 2132 (i.e., *m*/*z* 1752 + 329 and 1803 + 329) when sinapinic acid was used as matrix (Figure 1, inset). Treatment of these preparations with nuclease P1 resulted in a shift of the *m*/*z* 2081 peak to *m*/*z* 1752 and the *m*/*z* 2132 peak to *m*/*z* 1803 (data not shown). Since the *m*/*z* of dGMP is 329, these additional peaks represent the peptides containing a bound dUMP·dGMP dinucleotide.

MALDI-TOF analysis of the upper band peptide complex obtained after cross-linking in the presence of 150 mM NaCl yielded much cleaner spectra (data not shown). Extensive digestion of preparations containing the molecular species of *m*/*z* 1752 and 1803 with nuclease P1 had no effect (data not shown). Treatment with alkaline phosphatase resulted in a mass shift of *m*/*z* 80 (Figure 1, upper spectrum), exactly as expected for removal of the phosphate group from the bound dUMP.

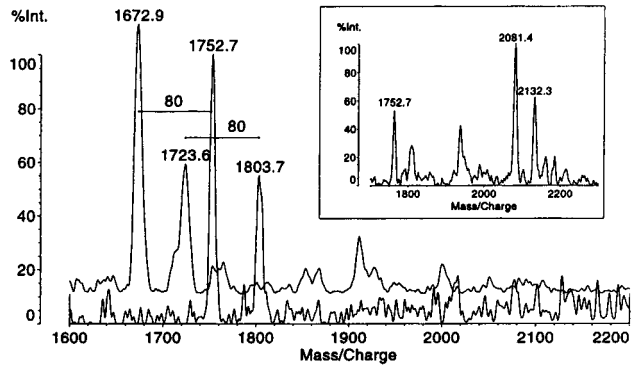

FIGURE 1: MALDI-TOF analysis of the peptide−oligonucleotide complexes present in one of the gel-purified complexes (the upper band) of tryptic peptides from vimentin cross-linked to oligo(dG·BrdU)$_{12}$·dG·3′-FITC. After exhaustive digestion with nuclease P1 (lower spectrum), two peaks of *m*/*z* 1752 and 1803 are observed. These probably correspond to peptides vim$_{50-63}$ (*m*/*z* 1445) and vim$_{36-49}$ (*m*/*z* 1495), each containing a bound dUMP (*m*/*z* 307), respectively. Treatment of such preparations with alkaline phosphatase (upper spectrum) results in a mass shift of *m*/*z* 80, consistent with the loss of a $-PO_4$ group from each complex. Inset: In some preparations, particularly when sinapinic acid was employed as matrix, additional complexes with *m*/*z* values of 2081 and 2132 were observed. These *m*/*z* values are those predicted for the vim$_{50-63}$ and vim$_{36-49}$ peptides bound to a dUMP−dGMP dinucleotide.

The lower band peptide complex was more recalcitrant to analysis, but finally yielded a weak peptide−dUMP peak at *m*/*z* 1222, which shifted by *m*/*z* 80 after digestion with alkaline phosphatase (data not shown). The peptide mass of this complex (*m*/*z* 1222 − 307 = 915) corresponds to that of vimentin peptide vim$_{28-35}$.

Sequencing of the Vimentin Upper Band Peptides and Identification of the Residues Cross-Linked to dUMP. The vimentin peptide−dUMP complexes thought to contain vim$_{36-49}$ and vim$_{50-63}$ were separated by HPLC on a reverse-phase column (Figure 2A) and yielded highly pure fractions (Figure 2B, lower and upper spectra in large graph). The cross-linked vim$_{50-63}$−dUMP complex was treated with aminopeptidase M, and a ladder of dUMP-substituted peptide derivatives with *m*/*z* 1753, 1666, and 1553 (Figure 2B, inset, lower spectrum) was obtained, yielding an N-terminal sequence of Ser-Ile/Leu, which is in agreement with the Ser-Leu expected for this peptide. Exhaustive digestion with aminopeptidase M failed to remove the next residue (Tyr)- (data not shown). As a control, native vim$_{50-63}$ (i.e., not cross-linked to oligonucleotide) was isolated and treated with aminopeptidase M: a ladder of peptides was obtained (Figure 2B, inset, upper spectrum) with *m*/*z* 1445, 1358, 1245, 1082, 995, and 908, which corresponds to an N-terminal sequence of Ser$_{50}$-Ile/Leu$_{51}$-Tyr$_{52}$-Ser$_{53}$-Ser$_{54}$ (the next bond, Ser−Pro, cannot be cleaved by this enzyme). These results suggest that the Tyr$_{52}$ in the *m*/*z* 1752 peak is modified, i.e., contains the bound dUMP.

The cross-linked vim$_{36-49}$−dUMP complex was analyzed in a similar manner. Aminopeptidase M could only remove

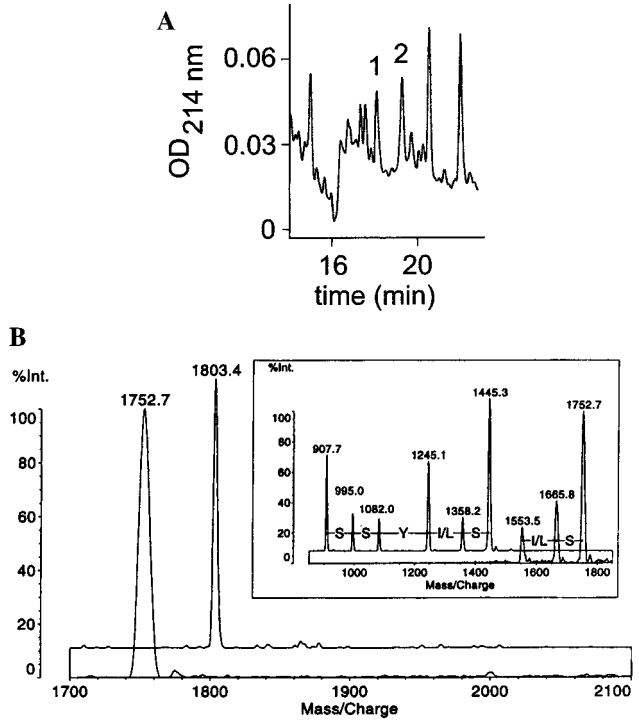

FIGURE 2: (A) Purification of the peptide−oligonucleotide complexes present in the upper band by reverse-phase HPLC (see Experimental Procedures for details). The chromatogram shows the relevant part of the separation. The column eluates containing the two indicated peaks were collected separately and were subjected to MALDI-TOF analysis (see Figure 2B). The other peaks did not contain peptide−oligonucleotide complexes (data not shown). (B) MALDI-TOF analysis of the peptide−oligonucleotide complexes present in the upper band after their purification by HPLC. The individual complexes are quite pure and yield exceptionally clean MALDI-TOF spectra (lower spectrum, complex from HPLC chromatogram peak #1, with *m*/*z* 1752; upper spectrum, complex from HPLC chromatogram peak #2, with *m*/*z* 1803). Inset: Peptide ladder sequencing with aminopeptidase M of the *m*/*z* 1752 peptide−oligonucleotide complex (lower spectrum) and of the vim_{50-63} (*m*/*z* 1445) peptide (upper spectrum). This analysis yielded an N-terminal sequence for the *m*/*z* 1752 complex of Ser_{50}-Leu_{51} and for the vim_{50-63} (*m*/*z* 1445) peptide of Ser_{50}-Leu_{51}-Tyr_{52}-Ser_{53}-Ser_{54}. Both of these partial sequences match that of the vim_{50-63} peptide.

the amino-terminal Thr_{36}, but not Tyr_{37} from this complex, yet residues Thr_{36} through Arg_{44} were rapidly removed from the native peptide (data not shown). This suggests that Tyr_{37} was cross-linked to dUMP. Attempts to sequence vim_{36-49} cross-linked to dUMP after digestion with carboxypeptidase Y were largely unsuccessful. Only a single peak of *m*/*z* 1647 (which represents the removal of Arg_{49}) was observed (data not shown). However, peptides produced by carboxypeptidase Y digestion of this complex were readily observed in MALDI-TOF spectra if the phosphate group of the bound dUMP was removed with alkaline phosphatase (Figure 3A). The C-terminal sequence of this peptide was Ser_{46}-Thr_{47}-Ser_{48}-Arg_{49}, thereby confirming the identity of vim_{36-49} in this complex.

Chemical Cleavage of Vimentin Peptide−Oligonucleotide Complexes. Since it was not possible to identify the residues containing the bound dUMP via enzymatic cleavage, the peptide−dUMP complexes were subjected to one or more cycles of manual Edman degradation. Cross-linked vim_{50-63}−dUMP contained two peptide complexes after aminopeptidase M treatment: one, *m*/*z* 1752, is undigested vim_{50-63}−dUMP, and the second, *m*/*z* 1552, is vim_{50-63}−dUMP with Ser-Leu removed (i.e., is a vim_{52-63}−dUMP complex)(Figure 3B, lower spectrum). After one cycle of Edman degradation, two new peaks of *m*/*z* 1666 (vim_{50-63}−dUMP minus Ser_{50}) and *m*/*z* 1082 were detected (Figure 3B, upper spectrum). This latter peak is that expected if a Tyr-dUMP moiety rather than just a Tyr is removed from the vim_{52-63}−dUMP complex. This conclusively demonstrates that Tyr_{52} is the single site of bound dUMP within the vim_{50-63} peptide.

An identical analysis was performed for the vim_{36-49}−dUMP complex (*m*/*z* 1803). Digestion of this complex with aminopeptidase M yielded a single product of *m*/*z* 1702, in addition to undigested starting material (results not shown). After one cycle of Edman degradation (Figure 3C), the undigested complex exhibited a mass shift of *m*/*z* 101 (i.e., Thr_{36} was removed), and the vim_{37-49}−dUMP complex exhibited a mass shift of *m*/*z* 471 (i.e., a Tyr-dUMP moiety was removed). The mass of the remaining material (*m*/*z* 1231) matches that predicted for residues vim_{38-49}. Thus, Tyr_{37} is the single site of bound dUMP within the vim_{36-49} peptide.

MALDI-TOF Analysis of the Cross-Linked Lower Band Peptide. MALDI-TOF spectra of both native and oligonucleotide cross-linked vimentin peptide vim_{28-35} were very noisy and difficult to interpret regardless of the matrix employed (data not shown). It was noticed that many peptides, including the cross-linked vim_{28-35}−dUMP, generated much cleaner spectra after modification with DABITC and using sinapinic acid as matrix. DABITC-modified vim_{28-35}−dUMP (Figure 3D, lower spectrum) was subjected to two complete rounds of manual Edman degradation, and the DABITC-modified reaction products were analyzed by MALDI-TOF. The sequencing ladder composite (Figure 3D) displays peaks of *m*/*z* 1505 for the DABITC-modified vim_{28-35}−dUMP (Figure 3D, lower spectrum), of *m*/*z* 1418 after one cycle (Figure 3D, middle spectrum), and of *m*/*z* 948 (Figure 3D, upper spectrum) after two cycles of manual Edman degradation. These data demonstrate that the N-terminal sequence of the cross-linked vim_{28-35} is Ser_{38}-Tyr_{29}-dUMP; the remaining peak of *m*/*z* 948 corresponds to that predicted for uncross-linked vim_{30-35}, modified with DABITC. Thus, Tyr_{29} is the single site of cross-linking to dUMP within this peptide.

DISCUSSION

Our initial attempts to identify the peptides to which an intact oligo(dG·BrdU)$_{12}$·dG·3′-FITC molecule was cross-linked based on the mass of the entire complex were unsuccessful. This was perhaps not surprising since the MALDI-TOF spectrum of the oligo(dG·BrdU)$_{12}$·dG·3′-FITC molecule alone was very noisy and basically unusable (data not shown). While our original approach also used oligo(dT·BrdU)$_{12}$·dT·3′-FITC, which yielded more tractable MALDI-TOF spectra, the precision of the mass determination of the peptide−oligonucleotide complexes was also too low to identify the peptides. These results are in concordance with those of Jensen et al. (*17*), who found that the quality of MALDI-TOF spectra from model peptide−oligothymidylic acid complexes was largely a function of the oligonucleotide. We therefore chose (a) to employ oligo(dG·BrdU)$_{12}$·dG·3′-FITC for our experiments, since dG-rich oligonucleotides are bound by vimentin better than T-rich (or other) oligonucleotides (*1*, *18*), and (b) to remove the bulk of the

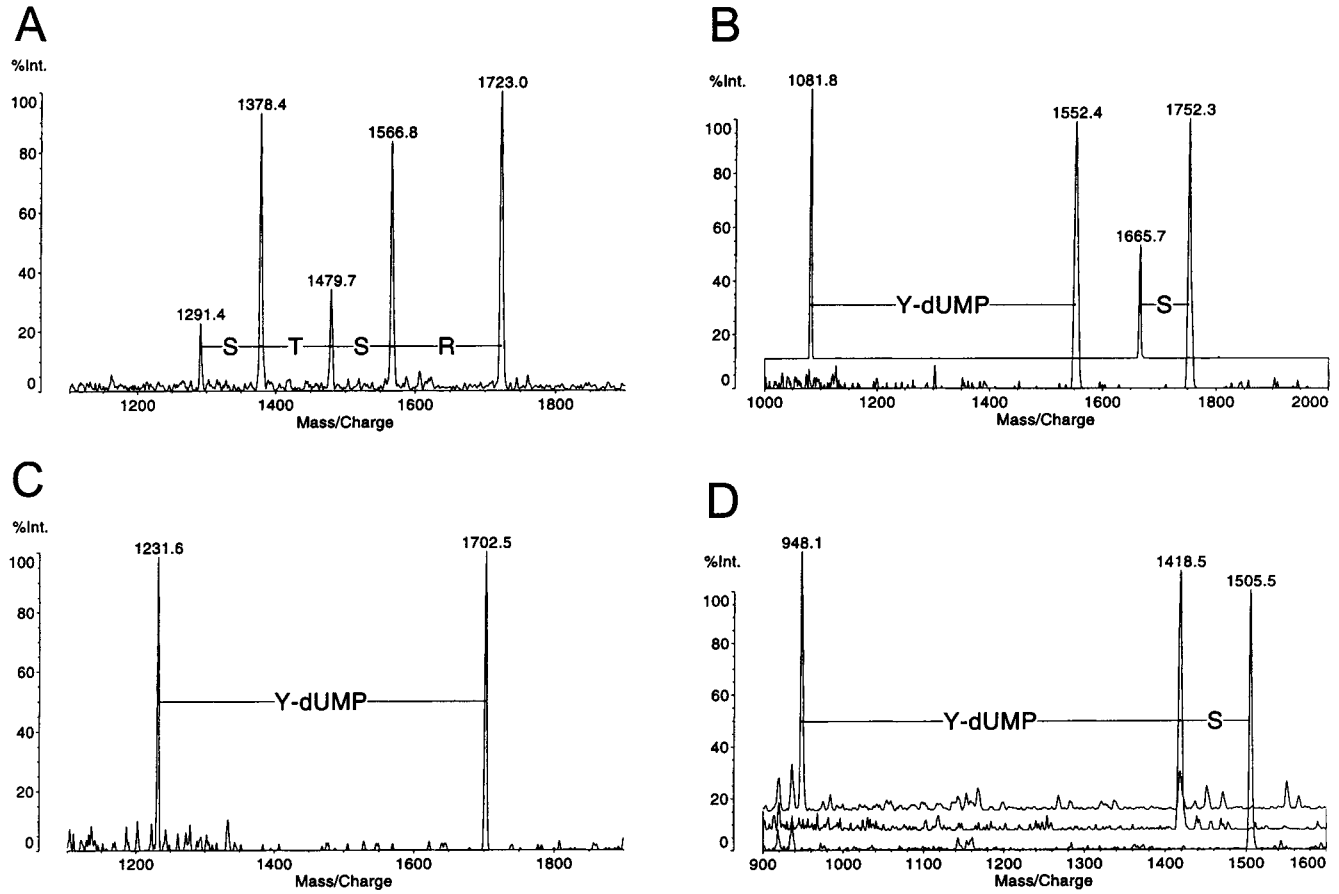

FIGURE 3: (A) MALDI-TOF analysis of the peptide ladder of the *m*/*z* 1803 peptide−dUMP complex treated with alkaline phosphatase and then digested with carboxypeptidase Y. The difference in *m*/*z* values corresponds to a sequence of Ser_{46}-Thr_{47}-Ser_{48}-Arg_{49}, thereby confirming the presence of vim_{36-49} peptide in this complex. (B) MALDI-TOF analysis of the peptide ladder of the *m*/*z* 1752 peptide−dUMP complex treated with aminopeptidase M (lower spectrum). Two peaks, with *m*/*z* 1752 and 1552, are observed, which correspond to undigested starting material and to a vim_{50-63}−dUMP complex lacking Ser_{50} and Leu_{51}. MALDI-TOF analysis of this material after one round of manual Edman degradation resulted in the production of molecules with *m*/*z* 1667 (i.e., the *m*/*z* 1752 complex minus Ser_{50}) and with *m*/*z* 1082 (upper spectrum). This mass shift (1552−1082) corresponds to that predicted if a Tyr−dUMP complex, rather than a free Tyr residue, is removed during this reaction. The mass of the remaining molecule agrees with that predicted for the free, unmodified vim_{53-63} peptide (*m*/*z* 1082.2). (C) MALDI-TOF analysis of the *m*/*z* 1803 vim_{36-49} peptide−dUMP complex treated with aminopeptidase M (which yielded the peak at *m*/*z* 1702 and corresponds to the removal of the N-terminal Thr_{36}) and subjected to one round of manual Edman degradation. A single additional peak is observed, corresponding to the removal of Tyr_{37}-dUMP. The mass of the remaining molecule agrees well with that predicted for the free, unmodified vim_{38-49} peptide (*m*/*z* 1232.4). (D) MALDI-TOF analysis of DABITC-modified vim_{28-35}−dUMP complexes. Analysis of untreated DABITC-modified vim_{28-35}−dUMP complex (lower spectrum) and the same complex subjected to one (middle spectrum) or two (upper spectrum) rounds of manual Edman degradation demonstrates that the N-terminal sequence of this complex is Ser_{28}-Tyr_{29}-dUMP. The mass of the remaining molecule agrees well with that predicted for the DABITC-modified vim_{30-35} peptide (*m*/*z* 948).

oligonucleotide by nuclease P1 digestion before MALDI-TOF analysis. This approach leaves only a dUMP residue attached to the peptide. Most of the vimentin tryptic peptides from the head domain contain a C-terminal Arg and generated reasonable spectra in MALDI-TOF analysis, except for the vim_{28-35} peptide. In general, the quality of the MALDI-TOF spectra was improved by removal of the phosphate from the bound dUMP. It was necessary to modify the amino terminus of the vim_{28-35} peptide with DABITC, after which clear and reproducible spectra were obtained (Figure 3D).

We took advantage of the unique chromatographic properties of the vimentin peptide−oligonucleotide complexes to first purify them from the other peptides that might have contaminated the gel-purified complexes. Thereafter, the oligonucleotide was digested with nuclease P1 to leave a dUMP residue at the site of cross-linking. Enzymatic peptide ladder sequencing (Figures 2B and 3A) was only partially successful and marginally useful in this analysis since (a) each peptide required unique combinations of enzyme concentration and digestion time to generate useful ladders and (b) the enzymes were not capable of removing amino acid residues containing bound dUMP. In some cases, enzymatic peptide ladder sequencing was quite useful for unequivocal identification of the peptide present in the cross-linked complexes (i.e., Figure 3A). Manual Edman degradation, on the other hand, removed amino acid residues with or without bound dUMP, but was rather time-consuming when multiple cycles were performed. We therefore employed both approaches in combination: enzymatic degradation was employed to remove unmodified residues, and chemical cleavage was employed to remove the residue containing the bound dUMP.

Only three peptides, vim_{28-35}, vim_{36-49}, and vim_{50-63}, all of which are located within the previously defined nucleic acid-binding region of the non-α-helical head domain (*5*)

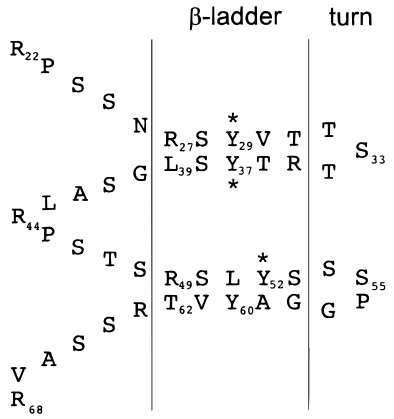

FIGURE 4: Model of the potential β-ladder DNA-binding wing structure of the nucleic acid binding domain of vimentin. Vimentin residues 27−39 can be folded into an antiparallel, β-ladder structure that is highly homologous to that found in the bacteriophage fd G5P single-stranded DNA-binding protein (see ref *5*). In this model, Tyr_{29} and Tyr_{37} are both predicted to be on the same face of the β-ladder to be in contact with the bound oligonucleotide. Vimentin residues 49−62 represent an imperfect match to these β-ladder sequences (see ref *5*). On the assumption that the proposed structure for vim_{49-62} is correct, Tyr_{52} and Tyr_{60} are predicted to be on opposite sides of the β-ladder DNA-binding wing (i.e., only one of these Tyr can contact the bound nucleotide). As demonstrated in this paper, Tyr_{29}, Tyr_{37}, and Tyr_{52} were found to be cross-linked to the oligonucleotide (indicated by an asterisk, *); Tyr_{60} was not. These data support the basic tenets of the model presented in this figure, but do not provide any information about the actual structure of the putative DNA-binding wings nor their orientation relative to each other.

could be cross-linked to the oligo(dG·BrdU)$_{12}$·dG·3′-FITC in these experiments; no other peptides from other domains were cross-linked to the oligonucleotide. Based on the yields of each peptide, the cross-linking of the oligo(dG·BrdU)$_{12}$·dG·3′-FITC to the three peptides was roughly equal. The sites of cross-linking were determined to be Tyr_{29}, Tyr_{37}, and Tyr_{52}. We have previously shown that nitration of the Tyr residues in the vimentin amino-terminal peptide NT1 (residues 1−96) abolishes salt-stable (intercalation) interactions, but not electrostatic binding of oligonucleotides (*5*). In this same study, a deletion mutant vimentin protein lacking residues 43−68 (i.e., lacking Tyr_{52} and Tyr_{60}) still bound oligo dG_{25} in a salt-stable fashion, whereas another deletion mutant vimentin protein lacking residues 25−43 (i.e., lacking Tyr_{29} and Tyr_{37}) did not. Due to the large number of Arg residues scattered throughout the head domain, both these and other deletion mutant proteins still bound oligonucleotides via electrostatic or hydrogen bonding interactions (*5*). This behavior was also exhibited by two synthetic peptides, R23R and R25R, which encompass vim_{22-44} and vim_{44-68}, respectively, and represent two imperfect direct repeats in the central region of the vimentin head domain (*5*): R23R bound oligo(dG_{25}) in a salt-stable fashion while R25R did not. The peptide R23R contains a sequence element, vim_{27-39}, which is highly homologous to the β-ladder DNA-binding wing of prokaryotic ssDNA-binding proteins, particularly to that of the gene 5 protein of bacteriophage fd (*19*, *20*) (Figure 4). It is very likely that the repeat peptide R25R containing Tyr_{52} as an additional vimentin−DNA cross-linking site is also capable of forming an antiparallel β-sheet structure and thus constitutes a second β-ladder DNA-binding wing. In any event, if our model of these DNA-binding wings is correct (*1*, *5*)(Figure 4), both Tyr_{29} and Tyr_{37} would be on the same face of the ladder and, by analogy to the results obtained with the bacteriophage protein, should contact the oligonucleotide. Furthermore, Tyr_{52} and Tyr_{60} are predicted to be on opposite faces of the potential second ladder, so that only one would be expected to contact the bound oligonucleotide. The results presented here demonstrate that both Tyr_{29} and Tyr_{37} contact the bound oligonucleotide, as does Tyr_{52}. In support of our model, Tyr_{60} is clearly not within cross-linking distance of the bound oligonucleotide. Of course, it will be necessary to employ more sophisticated analytical techniques, such as nuclear magnetic resonance spectroscopy, to prove this concept and also to demonstrate whether the two postulated DNA-binding wings in the amino-terminal head domain of vimentin fold into a higher order structure similar to the five-stranded antiparallel β-sheets of the ssDNA-binding proteins of various bacteriophages (*21*, *22*). In this context, it has to be noted that computer predictions (see *23*) suggest that the middle of the head domain of vimentin contains several stretches of amino acids with a propensity for β-sheet formation and that only the carboxyl end of the head domain can adopt an α-helical structure. The solution structure of the bacteriophage Pf3 ssDNA-binding protein (*22*) is a homodimer with two distinct DNA-binding wings. By analogy, the model of the vimentin nucleic acid-binding domain presented in Figure 4 is conceptually dimeric, with two functional DNA-binding wings.

The present study has the limitation that not all amino acid residues can efficiently react with BrdU and be covalently cross-linked to oligonucleotides. Preferred residues for this reaction include the aromatic amino acids Tyr and Trp, but the amino acids His, Cys, Lys, and Ala have also been reported to be involved in BrdU-based cross-linking (*7*). Since the other aromatic residues in vimentin were not cross-linked to the oligo(dG·BrdU)$_{12}$·dG·3′-FITC, it is clear that this approach was successful in specifically identifying those reactive residues localized within the DNA-binding domain. Previous studies (*24*, *25*) have shown that the many Arg residues in the head domain are responsible for the electrostatic interactions and hydrogen bonding and, furthermore, are likely responsible for the preference of vimentin to bind G-rich nucleic acids (*26*, *27*). Our previous (*5*) and present results show that Tyr_{29}, Tyr_{37}, and Tyr_{52} are intimately involved in the salt-stable intercalation reactions that occur after the initial electrostatic binding of nucleic acids by vimentin. Together, these data are very supportive of the proposal that vimentin's nucleic acid-binding site consists of two repeats of a β-ladder DNA-binding wing (*5*).

Many other IF subunit proteins, including the type III proteins desmin, glial fibrillary acidic protein, and peripherin, neurofilament triplet proteins, keratins, etc., also bind nucleic acids. While it remains to be seen whether this property is of physiological relevance (see *4* and references cited therein for a detailed discussion), it is clear that the methods described in this paper can be immediately applied to these and other proteins. Comparison of the results of such experiments should prove interesting and useful in identifying common motifs or gross differences in the respective nucleic acid-binding sites.

ACKNOWLEDGMENT

We thank Mrs. Monika Berthel for excellent technical assistance.

REFERENCES

1. Traub, P., Mothes, E., Shoeman, R., Kühn, S., and Scherbarth, A. (1992) *J. Mol. Biol. 228*, 41.
2. Shoeman, R. L., Wadle, S., Scherbarth, A., and Traub, P. (1988) *J. Biol. Chem. 263*, 18744.
3. Hartig, R., Shoeman, R. L., Janetzko, A., Tolstonog, G., and Traub, P. (1998) *J. Cell Sci. 111*, 3573.
4. Traub, P., and Shoeman, R. L. (1994) *Int. Rev. Cytol. 154*, 1.
5. Shoeman, R. L., Hartig, R., and Traub, P. (1999) *Biochemistry 38*, 16802.
6. Williams, K. R., and Konigsberg, W. H. (1991) *Methods Enzymol. 208*, 516.
7. Meisenheimer, K. M., and Koch, T. H. (1997) *Crit. Rev. Biochem. Mol. Biol. 32*, 101.
8. Jensen, O. N., Larsen, M. R., and Roepstorff, P. (1998) *Proteins*: *Struct., Funct., Genet. Suppl. 2*, 74.
9. Jensen, O. N., Barofsky, D. F., Young, M. C., von Hippel, P. H., Swenson, S., and Seifried, S. E. (1993) *Rapid Commun. Mass Spectrom. 7*, 496.
10. Bennett, S. E., Jensen, O. N., Barofsky, D. F., and Mosbaugh, D. W. (1994) *J. Biol. Chem. 269*, 21870.
11. Urlaub, H., Thiede, B., Müller, E.-C., and Wittmann-Liebold, B. (1997) *J. Protein Chem. 16*, 375.
12. Shoeman, R. L., Hartig, R., Huang, Y., Grüb, S., and Traub, P. (1997) *Antisense Nucleic Acid Drug Dev. 7*, 291.
13. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1994) *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc., New York.
14. Traub, P., and Boeckmann, G. (1978) *Hoppe-Seyler's Z. Physiol. Chem. 359*, 571.
15. Bonnetto, V., Bergman, A.-C., Jörnvall, H., and Sillard, R. (1997) *J. Protein Chem. 16*, 371.
16. Chang, J.-Y. (1983) *Methods Enzymol. 91*, 455.
17. Jensen, O. N., Kulkarni, S., Aldrich, J. V., and Barofsky, D. F. (1996) *Nucleic Acids Res. 24*, 3866.
18. Shoeman, R. L., and Traub, P. (1990) *J. Biol. Chem. 265*, 9055.
19. De Jong, E. A. M., van Duynhoven, J. P. M., Harmsen, B. J. M., Tesser, G. I., Konings, R. N. H., and Hilbers, C. W. (1989) *J. Mol. Biol. 206*, 119.
20. De Jong, E. A. M., van Duynhoven, J. P. M., Harmsen, B. J. M., Tesser, G. I., Konings, R. N. H., and Hilbers, C. W. (1989) *J. Mol. Biol. 206*, 133.
21. Folmer, R. H., Folkers, P. J., Kaan, A., Jonker, A. J., Aelen, J. M., Konings, R. N., and Hilbers, C. W. (1994) *Eur. J. Biochem. 224*, 663.
22. Folmer, R. H. A., Nilges, M., Papavoine, C. H. M., Harmsen, B. J. M., Konings, R. N. H., and Hilbers, C. W. (1997) *Biochemistry 36*, 9120.
23. Shoeman, R. L., Mothes, E., Kesselmeier, C., and Traub, P. (1990) *Cell Biol. Int. Rep. 14*, 583.
24. Traub, P., and Nelson, W. J. (1983) *Hoppe-Seyler's Z. Physiol. Chem. 364*, 575.
25. Traub, P., Nelson, W. J., Kühn, S., and Vorgias, C. E. (1983) *J. Biol. Chem. 258*, 1456.
26. Wang, X., Tolstonog, G., Shoeman, R. L., and Traub, P. (1996) *DNA Cell Biol. 15*, 209.
27. Harada, K., and Frankel, A. D. (1995) *EMBO J. 14*, 5798.

BI000199S