Probing the Interaction between a High-Affinity Single-Chain Fv and a Pyrimidine (6-4) Pyrimidone Photodimer by Site-Directed Mutagenesis†

Hiroyuki Kobayashi,‡ Hiroshi Morioka,‡ Kunihiro Tobisawa,‡ Takuya Torizawa,§ Koichi Kato,§ Ichio Shimada,§ Osamu Nikaido,|| Jon D. Stewart,*,⊥ and Eiko Ohtsuka*,‡

Graduate School of Pharmaceutical Sciences, Hokkaido University, Kita-ku, Sapporo 060-0812, Japan, Graduate School of Pharmaceutical Sciences, The University of Tokyo, Tokyo 113-0033, Japan, Faculty of Pharmaceutical Sciences, Kanazawa University, Kanazawa 920-0934, Japan, and Department of Chemistry, University of Florida, Gainesville, Florida 32611

Received April 22, 1998; Revised Manuscript Received August 25, 1998

ABSTRACT: We have used site-directed mutagenesis to uncover the origin of the binding affinity differences exhibited by a series of monoclonal antibodies that recognize pyrimidine (6-4) pyrimidone photoproducts in the context of single- or double-stranded DNA. In this study, we have focused on two antibodies—64M3 and 64M5—that share the same binding specificity but differ in their affinities. We used single-chain Fv (scFv) derivatives for these studies since they can be easily expressed in *Escherichia coli*. To facilitate this, we also developed a simple, on-column refolding procedure for scFvs that is rapid and does not require high dilution. We took several precautions to ensure that the scFvs faithfully reflected the behavior of the parent monoclonal antibodies. Results obtained from chimeric scFvs constructed from 64M3 and 64M5 suggested that the higher affinity of the 64M5 antibody was mainly due to its VL region. Loop-grafting studies in which VH CDR loops of 64M3 were individually transplanted into 64M5 were consistent with this hypothesis. Since the VL sequences of 64M3 and 64M5 differ at only three positions (L30, L50, and L90), alanine-scanning mutagenesis was used to assess the importance of these three residues in DNA binding by 64M5. These studies highlighted the importance of all three VL CDR loops; furthermore, they suggested that photoproduct binding involved conformational changes within the VL region.

DNA damage caused by ultraviolet (UV)[1] light is known to be a principal cause of toxicity, mutation, and cell death in living organisms (*1*). In vivo, UV-induced photoproducts are corrected by the base excision repair (BER) and nucleotide excision repair (NER) pathways (*2*). In the spectrum of UV photoproducts, (6-4) pyrimidine-pyrimidone photoproducts are of special concern since they most commonly produce mutations (Figure 1). We are therefore particularly interested in elucidating mechanisms of molecular recognition between repair proteins and this type of damaged DNA. While several proteins that interact with photodamaged DNA have been characterized, e.g., UV damage endonuclease (UVDE) (*3*−*5*), DNA photolyase (*6*, *7*), and damaged DNA-binding protein (*8*−*10*), details of the protein−DNA associations remain unknown.

To detect and quantitate photolesions in DNA, several monoclonal antibodies have been established (*11*−*14*). Previous studies from our laboratories have focused on a series of monoclonal antibodies isolated from a single mouse immunized with UV-irradiated calf thymus DNA (*12*). Two of the monoclonal antibodies specifically bound *cis*,*syn*-cyclobutane thymine dimers and four recognized pyrimidine (6-4) pyrimidone photodimers (designated 64M2, 64M3, 64M4, and 64M5) (*12*). Variable regions from the four antibodies recognizing (6-4) photoproducts have been cloned and sequenced, and computer models for each have been constructed (*15*). Amino acid sequences of three of the four are highly related (64M2, 64M3, and 64M5; ≈92% identical in the variable regions), and they are therefore likely to possess very similar three-dimensional structures. However, despite their similarities, these three antibodies display a range of binding affinities for photodamaged DNA (*12*). Specifically, 64M5 binds DNA with the highest affinity whereas 64M2 and 64M3 bind with at least an order of

† This work was supported by Grants-in-Aid by the Monbusho International Scientific Research Program and by Specially Promoted Research funding from the Ministry of Education, Science, Culture and Sports of Japan.

* To whom correspondence should be addressed. (E.O.) Phone: +81-11-706-3975. Fax: +81-11-706-4979. E-mail: ohtsuka@pharm.hokudai.ac.jp. (J.D.S.) Phone: 352-846-0743. Fax: 352-846-2095. E-mail: jds2@chem.ufl.edu.

‡ Hokkaido University.

§ The University of Tokyo.

|| Kanazawa University.

⊥ University of Florida.

[1] Abbreviations: Fv, antibody variable region fragment containing the VH and VL portions of a monoclonal antibody; scFv, single-chain Fv, consisting of antibody VH and VL domains linked by a flexible linker; VL, light-chain variable domain; VH, heavy-chain variable domain; CDR, complementarity-determining region(s); UV, ultraviolet; HPLC, high-performance liquid chromatography; PCR, polymerase chain reaction; cDNA, DNA complement derived from messenger RNA; *E. coli*, *Escherichia coli*; IPTG, isopropyl-1-thio-β-D-galactopyranoside; EDTA, ethylenediaminetetraacetic acid; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; NTA, nitrilotriacetic acid; GuHCl, guanidinium hydrochloride; SPR, surface plasmon resonance; RU, resonance unit; k_{ass}, association rate constant; k_{diss}, dissociation rate constant; T[6-4]T, (6-4) photoproduct of TpT; Fab, antigen-binding fragment(s); ELISA, enzyme-linked immunosorbent assay(s).

10.1021/bi9809167 CCC: $18.00
Published on Web 12/17/1998

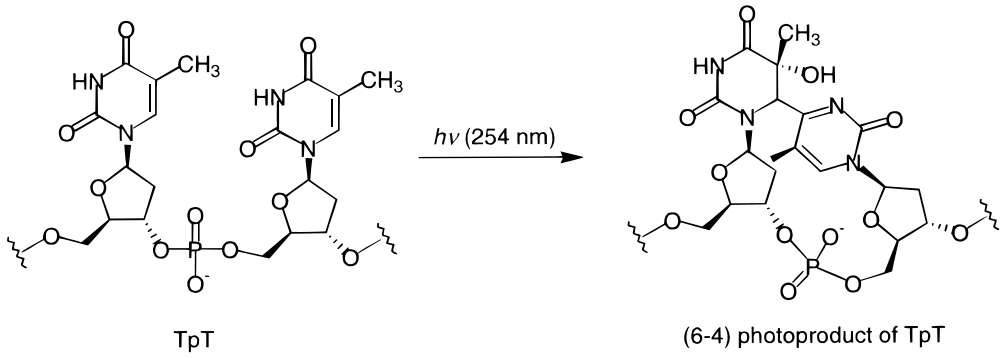

FIGURE 1: Photoconversion of two adjacent thymines to the (6-4) photoproduct.

magnitude lower affinities (*12*). This paper describes our efforts to determine the origin of these affinity differences. Chimeric scFvs in which one chain from 64M5 was replaced with the corresponding portion of 64M3 revealed that the VL portion of 64M5 conferred high-affinity DNA binding. A series of site-directed mutants was also created in order to systematically explore the residues that differed between 64M2, 64M3, and 64M5. On the basis of our computer modeling of 64M5 (*15*), these nonconserved residues in the VL CDRs (L30, L50, and L90) and the VH CDRs (H31−32, H58 and 60, H96−100) are located near the center of the putative photodimer binding channel on the protein surface. Our results indicate that the light chains play key roles in determining binding affinity and that induced fit is likely to play an important role in protein−DNA recognition. As part of these studies, we also developed a simple and rapid method for on-column renaturation and purification of antibody single-chain Fv derivatives directly from inclusion bodies.

MATERIALS AND METHODS

Enzymes and Other Reagents. Restriction and DNA-modifying enzymes were purchased from Takara Shuzo, New England Biolabs, Bethesda Research Labs, Stratagene, Toyobo, or Boehringer Mannheim. Chemicals were obtained from Wako Pure Chemical Industries, Nacalai Tesque, or Sigma Chemical. DNA phosphoramidite reagents were obtained from Perkin-Elmer Applied Biosystems.

Preparation of Oligonucleotides. Solid-phase oligonucleotide synthesis was carried out on an Applied Biosystems model 394 DNA/RNA synthesizer using standard β-cyanoethyl chemistry according to the manufacturer's protocol. The oligomers were cleaved from the support and deprotected by treatment with 4 mL of 28% ammonium hydroxide for 6 h at 55 °C and then purified by reversed-phase HPLC (μBondapak C18, Waters). 3′-Biotinylated oligonucleotides containing a (6-4) photoproduct were prepared as described previously (*16*).

Construction of Expression Plasmids for the 64M3 and 64M5 scFvs. The preparation of cDNA libraries from 64M3 and 64M5 hybridomas was performed as described previously (*15*). To determine the N-terminal amino acid sequences of the 64M5 VH and VL genes, the collection of double-stranded cDNAs was inserted into pBluescript II KS(+) (Stratagene) vector using *Eco*RI adaptors. PCR amplification was then performed on the pooled plasmids using a primer that annealed in the vector region one that annealed to the VH or VL regions: T7 [5′−d(GTA ATT CGA CTT ACT ATA GGG CGA)−3′]; VHB [5′−d-(CCG GAA TTC CCA GTG CAT CCA GTA)−3′]; VLB [5′−d-(CCG GAA TTC GTT CTG ACT AGA TCT)−3′]. The PCR products were directly sequenced using an automated DNA sequencer (Applied Biosystems model 373A). The C-terminal regions of VH and VL genes were obtained by similar means. The primers used for PCR amplification were VHF [5′−d(CCG GAA TTC AGT TAT GCT ATG GAC)−3′], located at the VH CDR3 region; MOCG12FOR [5′−d(CTC AAT TTT CTT GTC CAC CTT GGT GC)−3′], which annealed at the 3′ end of the heavy-chain constant region (CH1); VLF [5′−d(CCG GAA TTC ACA TGT TCC CAC GTT)−3′], which bound at the VL CDR3 region; and CKFOR [5′−d(CTC ATT CCT GTT GAA GCT CTT GAC)−3′], which annealed at the 3′ end of light-chain constant region (CL) (*15*, *17*). The VH and VL N-terminal sequences were also determined by direct amino acid sequencing of both chains of the 64M5 monoclonal antibody. This provided the first eight amino acids of the VL chain and the first six of the VH chain (*data not shown*). Using these data, primers were designed for PCR amplification of genes encoding the VL and VH regions of 64M5 without introducing undesirable mutations: VL−*Nde*I, 5′−d(G GAA TTC CAT ATG GAT GTT TTG ATG ACC)−3′; VL−*Nhe*I, 5′−d(CTA GCT AGC CCG TTT TAT TTC CAG CTT GGT CCC CCC CCC GA)−3′; VH−*Eco*RI, 5′−d(CCG GAA TTC GAG GTT CAG CTC CAG CAG TCT GGG ACT GTG CT)−3′; VH−*Xho*I, 5′−d(CCG CCG CTC GAG TGA GGA GAC GGT GAC TGA)−3′. Following amplification from a 64M5 cDNA library as described previously (*15*), the DNA fragments were digested with the appropriate restriction enzymes and purified. The complete 64M5 scFv gene was then assembled in a four-component ligation reaction that included DNAs encoding the VL and VH regions, a synthetic oligonucleotide encoding $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$ that possessed *Nhe*I and *Eco*RI overhangs at its 5′ and 3′ ends, respectively, and pET22b(+) that had been digested with *Nde*I and *Xho*I. After transformation of competent *E. coli* cells, plasmids were isolated from randomly picked colonies, mapped by restriction digestions, and the scFv gene was completely sequenced. The final expression plasmid was designated pETC64M5LH15.

Since the N- and C-terminal amino acid sequences of 64M3 were found to be identical to those of 64M5 (using procedures analogous to those described above), the same primers were used to amplify DNA fragments encoding the VL and VH portions of 64M3. These amplifications used the previously reported 64M3 clones as templates (*15*).

Table 1: Yields of ScFv Proteins

protein	altered CDR sequence[a]	mutations	yield (mg/L)[b]
wild-type 64M3			0.3
chimera 1	VL_{M3}-VH_{M5}		0.8
chimera 2	VL_{M5}-VH_{M3}		0.3
wild-type 64M5			2.6
LCDR1 mutant	RSSQNIVHSNGATYLE	Y-L30-A	1.6
LCDR2 mutant	AVSNRFS	T-L50-A	3.2
LCDR3 Ala mutant	FAGSHVPT	R-L90-A	0.7
LCDR3 Gln mutant	FQGSHVPT	R-L90-Q	0.6
HCDR1 mutant	SFWMH	N-H31-S, Y-H32-F	2.2
HCDR2 mutant	TTYPGNSDTSYNQKFKG	T-H58-S, S-H60-N	0.1
HCDR3 mutant	RSGYKYYAMDY	N-H96-S, Y-H97-G, G-H98-Y, S-H99-K, S-H100-Y	2.0

[a] Mutated amino acids are underlined. [b] After metal ion affinity and gel filtration chromatography.

Following amplification, an expression plasmid for the 64M3 scFv was constructed by steps analogous to those described above.

Mutant Constructions. Chimeric scFVs incoporating portions of the 64M3 and 64M5 variable regions were constructed by standard methods. For example, the chimeric scFv containing the VL region derived from 64M3 and the VH portion from 64M5 was created by digesting the 64M5 scFv expression plasmid with *Eco*RI and *Xho*I, then this segment was replaced by the VH region of 64M3. The latter was obtained by PCR amplificiation using the primers described above and the original 64M3 VH plasmid (*15*) as the template. The chimeric scFv containing the VL region from 64M5 and the VH portion from 64M3 was constructed in the same manner.

Site-directed mutations were introduced into the 64M5 scFv gene by PCR methods (*18*) and the scFv variants prepared for this study are summarized in Table 1. Two primers were used to construct each mutant, and specific base substitutions were introduced as a mismatch between a PCR primer and the target sequence. The first round PCR amplifications contained 20 pmol of the each appropriate primer, 10 ng of template, 5 μL of 10× reaction buffer, 4 μL of 2.5 mM dNTPs, and 2.5 units of *Pfu* DNA polymerase in a total volume of 50 μL. PCR reactions were carried out for 30 cycles of 94 °C (0.5 min), 45 °C (0.5 min), and 72 °C (2 min) after initially heating to 94 °C for 5 min. Second round PCR amplifications were carried out for 30 cycles using the first-round PCR products in 100 μL reaction mixture similar in composition to those described above. The final PCR products were digested with the appropriate enzymes and purified by gel electrophoresis. They were then used to replace the appropriate regions of the wild-type scFv in the expression vector. All mutant sequences were confirmed using an automated DNA sequencer (Applied Biosystems model 373A).

Expression and Purification of Wild-Type and Mutant 64M5scFvs. A 10 mL portion of L-broth containing 100 μg/mL ampicillin was inoculated with a single bacterial colony and incubated overnight at 37 °C. This preculture was then used to inoculate a 1 L portion of L-broth containing 100 μg/mL ampicillin. This culture was incubated at 37 °C until the absorbance at 660 nm was between 0.3 and 0.5. Single-chain Fv expression was then induced by adding IPTG to a final concentration of 0.1 mM, and the culture was incubated at 30 °C for 16 h. The cells were harvested by centrifuging at 5000 rpm for 10 min at 4 °C, then they were suspended in 30 mL of lysis buffer [50 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 0.5 mM EDTA, and 0.1 M NaCl] and sonicated with a microtip horn at 0 °C (12 × 30 s, with 1 min intervals between pulses). The sample was centrifuged at 12 000 rpm for 30 min at 4 °C. According to SDS−PAGE and western blotting with Omni-probe (M-21) (Santa Cruz Biotechnology), the scFv was present only in the pellet due to the formation of inclusion bodies. Proteins in the pellet were therefore dissolved overnight at 4 °C in 20 mL of solubilization buffer [6 M guanidine hydrochloride, 0.1 M Tris-HCl (pH 8.0), and 0.5 M NaCl]. After this time, insoluble debris was removed by centrifuging at 15 000 rpm for 30 min, then the solution containing the denatured scFv was loaded onto an Ni-NTA column (1 × 6 cm) that had been equilibrated with solubilization buffer. The scFv was renatured on the column eluting with a 200 mL linear gradient from 6 to 1 M GuHCl in 0.1 M Tris-HCl (pH 8.0) and 0.5 M NaCl over a period of 2 h. This was followed by a second 100 mL linear gradient from 1 to 0 M GuHCl in 0.1 M Tris-HCl (pH 8.0) and 0.5 M NaCl over 2 h. After renaturation, the Ni-NTA column was washed with 20 mL of wash buffer [0.1 M Tris-HCl (pH 8.0) and 20 mM imidazole], then the scFv protein was eluted with 0.1 M Tris-HCl (pH 8.0) and 500 mM imidazole. The partially purified scFv was loaded onto a Superose 12 HR 10/30 column that had been previously equilibrated with HBS (10 mM HEPES, 150 mM NaCl, 3.4 mM EDTA, and 0.005% Tween-20, pH 7.4). This column was eluted isocratically using a SMART-system (Pharmacia Biotech).

BIAcore Measurement of 64M5scFv Binding to the (6-4) Photoproduct. The binding of scFvs to oligonucleotides containing a (6-4) photoproduct was determined by surface plasmon resonance (SPR) measurements using a Pharmacia BIAcore instrument as described previously (*16*). All experiments were performed at 25 °C. Sensor chip SA surfaces with streptavidin preimmobilized to the dextran substrate were used (Pharmacia Biosensor). Oligonucleotides were captured onto the sensor chips by injecting 5−10 μL of biotinylated oligonucleotides diluted to 0.01 pmol/μL in HBS at a flow rate of 20 μL/min. Injections were repeated until the SPR signal was increased by 10−30 resonance units (RU) above the original baseline.

The scFvs were diluted in HBS buffer prior to SPR analysis. The scFv solution was injected in the presence of immobilized oligonucleotides at a flow rate of 100 μL/min over a concentration range from 10 to 150 nM. Sensorgrams were recorded and normalized to a baseline of 0 RU. Equivalent volumes of the diluted antibodies were also injected over an underivatized surface to allow subtraction

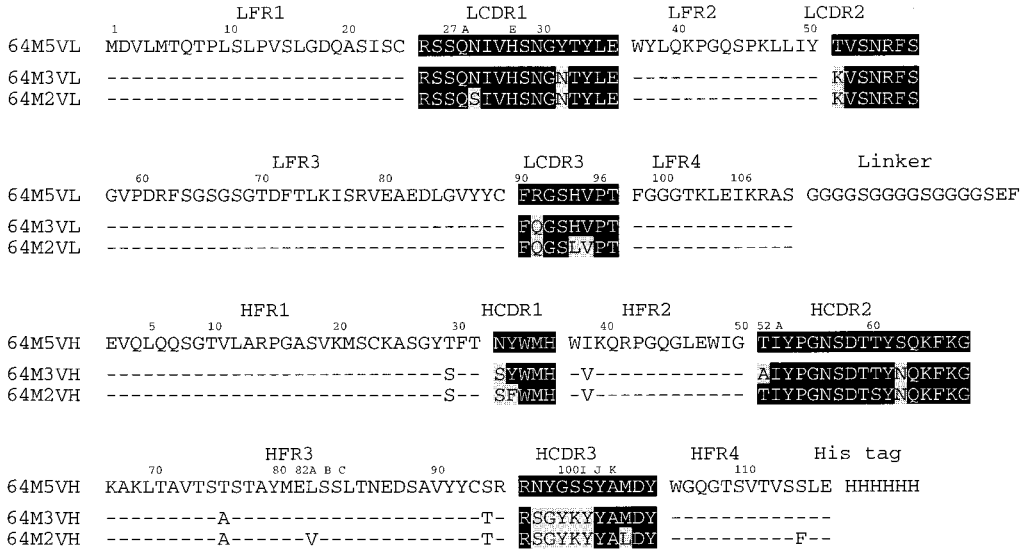

FIGURE 2: Variable region amino acid sequences of anti-(6-4) photoproduct antibodies. These data incorporate the results from directly sequencing the N- and C-terminal regions of the monoclonal antibodies. The sequences and positions of the flexible linker and hexahistidine tag used to construct scFv derivatives are also shown.

of the bulk refractive index background. The association was monitored by measuring the rate of binding to antigen at different protein concentrations. The dissociation of the scFvs was monitored after the end of the association phase. Any antibodies that remained after the conclusion of a measurement were removed by injecting 50 μL of 100 mM HCl.

Kinetic rate constants were calculated from the SPR data using BIAevaluation 2.1 software (Pharmacia). A single-site binding model (A + B = AB) was used for the analysis. The ratio of the apparent rate constants allowed the apparent equilibrium constant to be calculated, $K_D = k_{diss}/k_{ass}$.

Fluorescence Titrations. The thermodynamic K_D value for binding of the wild-type 64M5 scFv to d(T[6-4]T) was determined by fluorescence quenching. For fluorescence measurements, 64M3 and 64M5 Fab fragments were dissolved at concentrations of 1.0 μM and 30 nM, respectively, in 5 mM sodium phosphate buffer (pH 7.4), containing 150 mM NaCl and 3.4 mM EDTA. Emission spectra of protein intrinsic fluorescence were recorded in the region from 310 to 350 nm upon excitation at 260 nm using a Shimadzu RF-5300PC spectrofluorometer. Excitation and emission bandpasses were 5 and 10 nm, respectively. All measurements were carried out at 30 °C. Fluorescence quenching at 334 nm upon addition of d(T[6-4]T) was monitored and binding constants of Fab for d(T[6-4]T) were determined from the slope of the Scatchard plot. The linearity of these plots suggested that isomerization to the Dewar photoproduct did not occur to a significant extent during these measurements.

Computer Modeling. The 64M5 Fv model constructed previously (*15*) was used as the basis of the scFv models. To model the scFv derivatives, the linker sequences were added at the appropriate chain termini using the Biopolymer module of Sybyl v. 6.3. To generate a reasonable conformation for the linker, a loop search was performed with the ends of the loop fixed. The *c*-myc and hexahistidine tags were added in the same manner in random conformations. Since all of these elements are likely to be highly conformationally mobile, further energy minimizations were not performed.

RESULTS

Single-Chain Fv Design and Construction. The N-terminus of the VL region lies in close proximity to the antigen-binding site, and this region was therefore of special concern when designing scFv derivatives. In our previous studies of the four monoclonal antibodies that recognize pyrimidine (6-4) pyrimidone photodimers (*15*), the N-terminal sequences of both the VL and VH portions were derived from the standard PCR primers used to amplify the variable region genes. In addition, the original scFv construct for the antibody with the highest binding affinity (64M5) had the flexible linker joining the C-terminus of the VH region with the N-terminus of the VL portion. While the specificity of this scFv was identical to that of the corresponding Fab fragment prepared by proteolysis of the monoclonal antibody, the association rate for binding of this scFv to photoproduct-containing DNA was ca. 10-fold slower than the value obtained for the Fab fragment (*16*). For this study, we have used scFv constructs in which the flexible linker instead joins the C-terminus of the VL portion with the N-terminus of the VH region. Moreover, we have determined N-terminal amino acid sequences directly from the 64M5 monoclonal antibody so that the corresponding scFv genes encoded the native residues at these positions and thereby better reflected the behavior of the natural protein. A hexahistidine tag was appended to the C-terminus of the VH region to facilitate purification. The sequence of the final 64M5 scFv construct is shown in Figure 2, and computer models of the original and current scFv are shown in Figure 3.

On-Column Refolding of scFvs. Preliminary attempts to directly express soluble scFv in the periplasmic space of *Escherichia coli* cells were unsuccessful and only inclusion bodies containing the scFv were obtained. Moreover, SDS−PAGE analysis revealed the presence of both the desired scFv and a protein with slightly higher molecular weight. The latter likely corresponds to scFv in which the *pelB* leader sequence remained intact as a result of incomplete processing. We therefore expressed the scFv without a leader sequence and employed an in vitro refolding system. After dissolving the

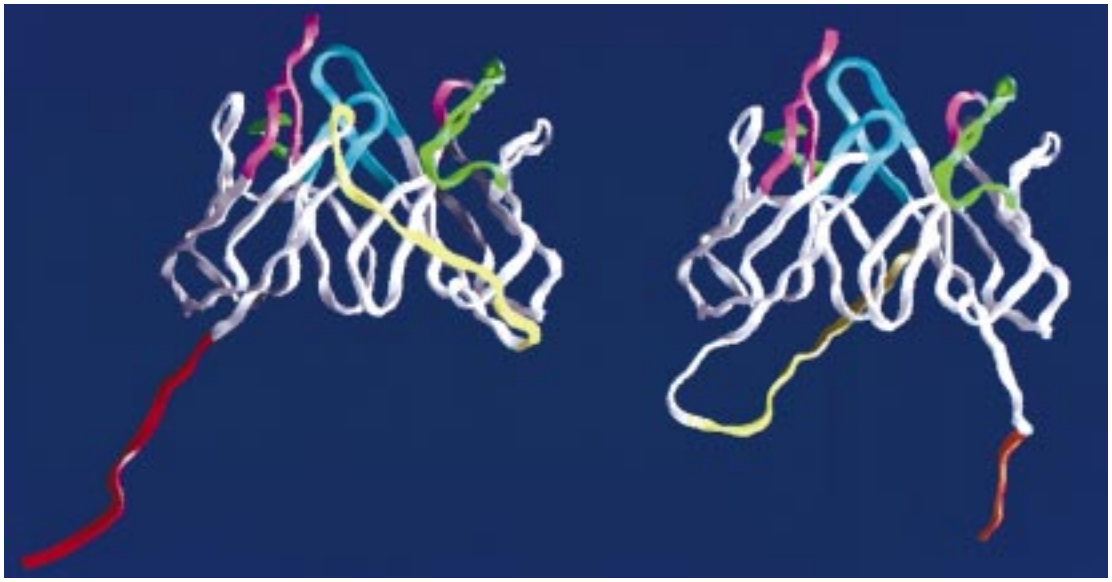

FIGURE 3: Computer models of the 64M5 scFvs. Framework residues are shown in white, and CDR loops are colored pink (LCDR1 and HCDR1), green (LCDR2 and HCDR2), and blue (LCDR3 and HCDR3) and the flexible linkers are shown in yellow. (A) Original scFv design. The flexible linker connects the C-terminus of the VH region to the N-terminus of the VL and a twelve amino acid *c*-myc tag (red) is appended to the C-terminus of the VL portion to facilitate detection in ELISA experiments. Note the close approach of the linker to the antigen binding site in this design. (B) Single-chain Fv design used in the present study. Here, the flexible linker connects the C-terminus of the VL region with the N-terminus of the VH. A hexahistidine tag (orange) is appended to the VH region to facilitate purification. Note that the linker is well-removed from the antigen binding site in this design. Both figures were produced using GRASP (*30*).

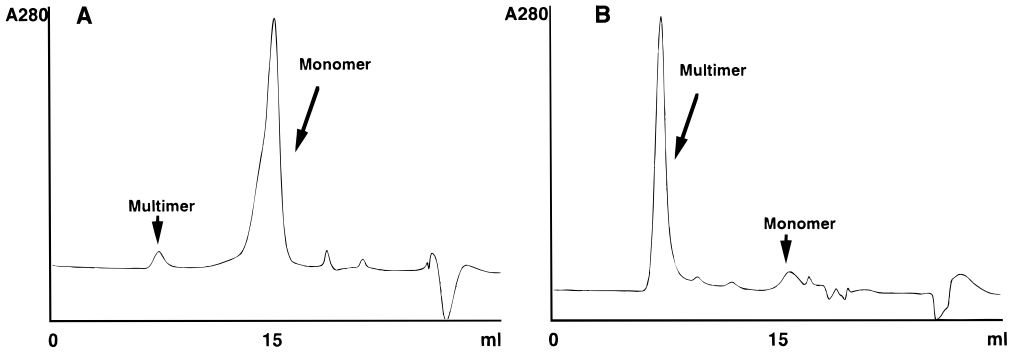

FIGURE 4: Gel filtration chromatography of refolded scFvs. (A) Analysis of the wild-type 64M5 scFv after on-column refolding showed = 90% monomer. (B) Analysis of the HCDR2 mutant after on-column refolding revealed only a small amount of monomer, suggesting that this mutant was unable to maintain its proper three-dimensional structure.

inclusion bodies in 6 M guanidine hydrochloride, the mixture was applied to a Ni-NTA column. After washing with buffer containing 6 M guanidine hydrochloride to remove nonspecifically bound proteins, the scFv was refolded in situ by washing the column with buffer containing a linear gradient from 6 to 0 M guanidine hydrochloride, then the refolded scFv was eluted from the column. Analysis of the eluted scFv by analytical gel filtration chromatography revealed that only small amounts of aggregate were present in the sample, except in the case of the HCDR2 mutant (vide infra). The aggregates were removed by preparative gel filtration chromatography (Figure 4) and the final yields of scFvs prepared by this method are listed in Table 1.

Fluorescence quencing titrations were used to determine the thermodynamic K_D values for binding of the 64M5 Fab and the refolded and purified 64M5 scFv to a dinucleotide containing a (6-4) photoproduct (2.0×10^{-8} and 1.4×10^{-8} M, respectively). The similar affinities suggested that antigen binding by this type of scFv closely resembles that of the native monoclonal antibody. Unfortunately, it was not possible to determine K_D values for an octanucleotide containing a central (6-4) photoproduct by this method since infinitely tight binding was observed even at the lowest practical antibody concentrations.

Properties of Chimeric scFvs. We created a series of chimeric scFvs in which one chain from the highest-affinity antibody (64M5) was paired with a partner from one of the two antibodies with lower affinities (64M3). After refolding and purification, the affinities of the chimeric scFvs for an octamer containing a central (6-4) photoproduct were determined by surface plasmon resonance using a procedure similar to that in our previous report (*16*) (Table 2). The amount of immobilized DNA was kept as low as possible (10−30 RU) in order to minimize scFv rebinding during the dissociation phase (*19*). No binding was detected for sensor chips either derivatized with streptavidin alone or coupled with the undamaged oligonucleotide (data not shown). The binding affinity (k_{diss}/k_{ass}) of 64M5 to the octanucleotide containing T[6-4]T was approximately 26-fold higher than that of 64M3. The binding affinity of Chimera 1 (consisting of VL portion of 64M3 and VH portion of 64M5) was almost same as that of 64M3; however, Chimera 2 (which contained

Table 2: Kinetic Constants for the Binding of ScFvs to Oligonucleotides Containing a (6-4) Photoproduct

	d2-mer-(6-4)[a]			d8-mer(6-4)[b]		
scFv	k_{ass} (M^{-1} s^{-1})	k_{diss} (s^{-1})	k_{diss}/k_{ass} (M)	k_{ass} (M^{-1} s^{-1})	k_{diss} (s^{-1})	k_{diss}/k_{ass} (M)
wild-type 64M3	$(1.4 \pm 0.2) \times 10^5$	$(2.1 \pm 0.6) \times 10^{-3}$	$(1.5 \pm 0.3) \times 10^{-8}$	$(2.4 \pm 0.1) \times 10^5$	$(6.2 \pm 0.2) \times 10^{-4}$	$(2.6 \pm 0.1) \times 10^{-9}$
VL_{M3}-VH_{M5}	$(5.0 \pm 0.2) \times 10^4$	$(2.0 \pm 0.1) \times 10^{-3}$	$(3.9 \pm 0.1) \times 10^{-8}$	$(2.4 \pm 0.2) \times 10^5$	$(7.9 \pm 0.7) \times 10^{-4}$	$(3.3 \pm 0.4) \times 10^{-9}$
VL_{M5}-VH_{M3}	$(1.4 \pm 0.3) \times 10^5$	$(2.2 \pm 0.1) \times 10^{-3}$	$(1.8 \pm 0.2) \times 10^{-8}$	$(8.4 \pm 0.3) \times 10^5$	$(1.4 \pm 0.1) \times 10^{-4}$	$(1.7 \pm 0.1) \times 10^{-10}$
wild-type 64M5	$(9.5 \pm 0.7) \times 10^5$	$(1.5 \pm 0.2) \times 10^{-3}$	$(1.6 \pm 0.2) \times 10^{-9}$	$(1.1 \pm 0.1) \times 10^6$	$(1.1 \pm 0.1) \times 10^{-4}$	$(1.0 \pm 0.1) \times 10^{-10}$
LCDR1 mutant	$(1.4 \pm 0.3) \times 10^5$	$(1.9 \pm 0.3) \times 10^{-3}$	$(1.4 \pm 0.4) \times 10^{-8}$	$(4.1 \pm 0.2) \times 10^5$	$(2.0 \pm 0.2) \times 10^{-3}$	$(4.9 \pm 0.5) \times 10^{-9}$
LCDR2 mutant	$(1.7 \pm 0.6) \times 10^5$	$(2.5 \pm 0.5) \times 10^{-3}$	$(1.5 \pm 0.6) \times 10^{-8}$	$(3.2 \pm 0.3) \times 10^5$	$(1.0 \pm 0.4) \times 10^{-3}$	$(3.1 \pm 1.3) \times 10^{-9}$
LCDR3 Ala mutant	undetectable	undetectable		undetectable	undetectable	
LCDR3 Gln mutant	$(1.5 \pm 0.3) \times 10^5$	$(4.9 \pm 0.6) \times 10^{-3}$	$(3.2 \pm 0.4) \times 10^{-8}$	$(1.6 \pm 0.3) \times 10^5$	$(5.2 \pm 1.4) \times 10^{-4}$	$(3.3 \pm 1.1) \times 10^{-9}$
HCDR1 mutant	$(5.1 \pm 1.2) \times 10^5$	$(2.8 \pm 0.7) \times 10^{-3}$	$(5.5 \pm 1.9) \times 10^{-9}$	$(1.1 \pm 0.1) \times 10^6$	$(1.9 \pm 0.2) \times 10^{-4}$	$(1.7 \pm 0.2) \times 10^{-10}$
HCDR2 mutant	undetectable	undetectable		undetectable	undetectable	
HCDR3 mutant	$(9.3 \pm 0.5) \times 10^5$	$(3.9 \pm 0.6) \times 10^{-3}$	$(4.2 \pm 0.7) \times 10^{-9}$	$(1.1 \pm 0.1) \times 10^6$	$(6.3 \pm 0.6) \times 10^{-4}$	$(5.7 \pm 0.8) \times 10^{-10}$

[a] d(T⟨6-4⟩T). [b] d(CAAT⟨6-4⟩TAAG).

the VL portion of 64M5 and VH portion of 64M3) possessed almost the same affinity as 64M5. These data indicated that VL portion of 64M5 is mainly responsible for its high affinity.

Identification of Individual Residues Responsible for High-Affinity Binding. Nearly all of the sequence differences between the 64M2, 64M3, and 64M5 antibodies occur in the hypervariable loops (*15*). On the basis of the previous results, we constructed a series of mutants based on 64M5 in which individual VH CDR loops were altered to match those of 64M2 or 64M3 in order to detect loops that were responsible for the higher binding affinity of 64M5. In essence, these studies refined the chimeric scFv strategy to the level of single CDR loops. In the case of the HCDR1 and HCDR2 loops, two amino acids were altered; however, five contiguous amino acids were altered in the case of the HCDR3 region. Since studies of the chimeric scFvs (Chimera 1 and Chimera 2) had underscored the importance of the 64M5 VL region in high-affinity DNA binding, amino acids at LCDR positions that differed between 64M5 and 64M2 and 64M3 were individually converted to alanines. In addition, the R-L90-Q mutant of 64M5 was also created since this substitution was found in both 64M2 and 64M3. All of the mutants created in this study are summarized in Table 1. After on-column refolding and purification, the affinities of the site-directed mutants for both a dimer and octamer containing a central (6-4) photoproduct were determined by surface plasmon resonance (Table 2). The binding affinities of LCDR mutants (Y-L30-A, T-L50-A, and R-L90-Q) were approximately 10-fold lower than that of wild-type 64M5, while those of HCDR mutants were comparable to that of wild-type 64M5. Interestingly, the R-L90-A mutant showed no binding affinity for the (6-4) photoproduct. The HCDR2 mutant was highly unstable and prone to aggregation.

DISCUSSION

For meaningful results to be obtained from site-directed mutagenesis studies, it was critical that the scFvs faithfully reproduced the antigen-binding behavior of the native monoclonal antibodies. The VL N-terminal region was of particular concern since NMR studies of the 64M5 Fab fragment have suggested involvement of these residues in (6-4) photoproduct binding (Shimada et al., manuscript in preparation). To avoid the possibility of steric interference in antigen binding by the flexible linker, the flexible linker joined the VL C-terminus with the VH N-terminus (VL-linker-VH format). A hexahistidine sequence was appended to the C-terminus of the VH region to facilitate purification (Figure 3). The identical thermodynamic K_D values for binding of a dinucleotide containing the (6-4) photoproduct by the 64M5 scFv incorporating the current design and the Fab fragment argue for identical active-site structures. In addition, the binding kinetics of the current 64M5 scFv are very similar to those previously reported for the Fab fragment (*16*).

Antibodies and their derivatives are frequently produced in *E. coli* as insoluble aggregates (*20*). While active proteins can be obtained from such inclusion bodies by solubilization and refolding, this process can be time-consuming and usually requires large volumes of highly dilute solutions to avoid aggregration (*21*, *22*). Furthermore, formation of incorrect intra- and intermolecular disulfide bridges is a major problem with scFv refolding (*23*, *24*). Since proteins bound to a solid support are effectively isolated from one another, on-column refolding has been suggested as an alternative to high dilution (*25*). We therefore developed a simple on-column refolding method that also allows scFv purification in a single operation. A key advantage of this procedure is that immobilization removes intermolecular interferences in protein refolding. The use of metal-chelate affinity chromatography is especially convenient for this process since binding to the solid support occurs even in the presence of 6 M guanidine hydrochloride (*26*, *27*). We expect that this procedure will be generally useful for scFvs and other multidomain proteins prone to aggregation during refolding.

Despite their close sequence relationships, the 64M2, 64M3, and 64M5 antibodies display a range of binding affinities for DNA containing (6-4) photoproducts (*12*). To determine whether the VL or VH portion of 64M5 was mainly responsible for its higher affinity relative to the other two, chimeric scFvs were constructed from the 64M3 and 64M5 antibodies. By measuring the kinetics for binding of an octanuclecotide containing a central (6-4) photoproduct, it was clear that the light chain of 64M5 was the major contributor to high-affinity DNA binding (Table 2). To further demonstrate the key role of the 64M5 VL in high-affinity DNA binding, loop-grafting studies were performed in which one VH CDR sequence derived from a low-affinity antibody (64M2 or 64M3) replaced the corresponding segment of the 64M5 scFv. Unfortunately, the HCDR2 mutant was highly unstable and prone to aggregation, and it was therefore not investigated further. Kinetic analysis of

DNA binding by the remaining two loop graft scFvs revealed only small changes in binding rate constants. These results were surprising since the amino acid sequences of the 64M3 and 64M5 VL regions differ at only three positions (L30, L50, and L90) whereas the VH regions differ at 14 positions.

Each of the three positions that differ between the 64M5 and 64M3 VL sequences was examined individually by alanine-scanning mutagenesis. While replacement of the L30 and L50 residues with alanine was tolerated by the 64M5 scFv, the R-L90-A mutant did not detectably bind DNA containing a (6-4) photoproduct. This result was not completely unexpected since the L90 side chain faces the protein interior and tightly packs with other, predominantly hydrophobic side chains. Moreover, no VL sequence in the current Kabat database contains an alanine at this position (*28*). We therefore created the R-L90-Q mutant since glutamine is found most often at this position and it also occurs in both the 64M2 and 64M3 sequences. Comparison of the binding kinetics obtained from these mutants with those from the wild-type scFvs revealed that residues at the L30, L50, and L90 positions contributed almost equally to the observed affinity difference between the 64M3 and 64M5 VL chains (Table 2).

The most striking finding from the VL mutants were that changes in affinity were mainly due to changes in the association rates. Association rate constants for the dinucleotide containing a (6-4) photoproduct are especially useful since the interactions are due solely to antibody−photoproduct interactions and do not include other, probably auxiliary contacts with flanking nucleotides (*16*). Mutation of either the L30 or L50 residue to alanine significantly decreased the association rate constant while having little impact on the dissociation phase. Such mutations are unlikely to alter the intrinsic collision rate between the scFv and DNA, and the differences in association rate constants are instead more likely to reflect conformational changes in the binding step. One hallmark of multistep binding is the nonequivalence of the thermodynamic dissociation constant and the apparent value calculated from the ratio of k_{diss} and k_{ass}. In the case of the wild-type 64M5 scFv, the thermodynamic K_D value for binding of the (T[6-4]T) dinucleotide was 1.4×10^{-8} M; however, the ratio of k_{diss}/k_{ass} for the same ligand was an order of magnitude smaller ($[1.6 \pm 0.2) \times 10^{-9}$]. While part of this discrepancy could be the result of experimental error in both the fluorescence and SPR values, we propose that conformational equilibria of the unbound scFv also contribute. Such conformational mobility might also explain the occurrence of Arg at position L90, whose charged side chain is surrounded by hydrophobic groups in the 64M5 structure. To our knowledge, an Arg residue at this position is unprecedented in κ light-chain sequences (*28*). Given its position, the side chain of Arg L90 cannot interact directly with the DNA, but instead must affect the affinity indirectly, perhaps by helping to manipulate the conformation of the LCDR3 loop. By contrast, the equivalent position in both 64M2 and 64M3 is occupied by Gln, which is the most commonly observed residue.

Multinuclear NMR studies of the 64M5 Fab fragment revealed conformational mobility within the LCDR1 and LCDR3 loops in the absence of photoproduct-containing DNA (Shimada et al., unpublished observations) and a similar conformational rearrangement of the BV04−01 LCDR1 loop upon DNA binding has been observed by Herron et al. (*29*). This observation is particularly significant given the very close sequence relationship between the VL regions of BV04-01 and the anti-photoproduct antibodies described in this paper.

ACKNOWLEDGMENT

We thank Drs. Thoru Natsume (JST, ERATO), Hitoshi Nishimura and Yasuo Komatsu (Hokkaido University) for valuable discussions and advice.

REFERENCES

1. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) in *DNA repair and mutagenesis*, ASM Press, Washington, DC.
2. Sancar, A. (1996) *Annu. Rev. Biochem. 65*, 43−81.
3. Bowman, K. K., Sidik, K., Smith, C. A., Taylor, J. S., Doetsch, P. W., and Freyer, G. A. (1994) *Nucleic Acids Res. 22*, 3026−3032.
4. Yajima, H., Takao, M., Sasuhira, S., Zhao, J. H., Ishii, C., Inoue, H., and Yasui, A. (1995) *EMBO J. 14*, 2393−2399.
5. Takao, M., Yonemasu, R., Yamamoto, K., and Yasui, A. (1996) *Nucleic Acids Res. 24*, 1267−1271.
6. Todo, T., Takemori, H., Ryo, H., Ihara, M., Matsunaga, T., Nikaido, O., Kenji, S., and Nomura, T. (1993) *Nature 361*, 371−374.
7. Todo, T., Kim, S. T., Hitomi, K., Otoshi, E., Inui, T., Morioka, H., Kobayashi, H., Ohtsuka, E., Toh, H., and Ikenaga, M. (1997) *Nucleic Acids Res. 25*, 764−768.
8. Treiber, D. K., Chaen, Z., and Esigmann, J. M. (1992) *Nucleic Acids Res. 20*, 5805−5810.
9. Reardon, J. T., Nichols, A. F., Keeney, S., Smith, C. A., Taylor, J. S., Linn, S., and Sancar, A. (1993) *J. Biol. Chem. 268*, 21301−21308.
10. Kai, M., Takahashi, T., Todo, T., and Sakaguchi, K. (1995) *Nucleic Acids Res. 23*, 2600−2607.
11. Roza, L., Wulp, K. J. M. V. D., Macfarlane, S. J., Lohman, P. H. M., and Baan, R. A. (1988) *Photochem. Photobiol. 48*, 627−633.
12. Mori, T., Nakane, M., Hattori, T., Matsunaga, M., Ihara, M., and Nikaido, O. (1991) *Photochem. Photobiol. 54*, 225−232.
13. Mizuno, T., Matsunaga, T., Ihara, M., and Nikaido, O. (1991) *Mutat. Res. 254*, 175−184.
14. Matsunaga, T., Hatakeyama, Y., Ohta, M., Mori, T., and Nikaido, O. (1993) *Photochem. Photobiol. 57*, 934−940.
15. Morioka, H., Miura, M., Kobayashi, H., Koizumi, T., Fujii, K., Asano, K., Matsunaga, T., Nikaido, O., Stewart, J. D., and Ohtsuka, E. (1998) *Biochim. Biophys. Acta 1385*, 17−32.
16. Kobayashi, H., Morioka, H., Torizawa, T., Kato, K., Shimada, I., Nikaido, O., and Ohtsuka, E. (1998) *J. Biochem. 123*, 182−188.
17. McCafferty, J.; Johnson K. S. (1996) in *Phage Display of Peptides and Proteins* (Kay, B. K., Winter, J., and McCafferty, J., Eds.) pp 79−111, Academic Press, San Diego, CA.
18. Higuchi, R. (1989) in *PCR Technology* (Erlich, H. A., Ed.) pp 61−70, Stockton Press, New York.
19. Schuck, P., and Minton, A. P. (1996) *Trends Biochem. Sci. 21*, 458−460.
20. Plückthun, A. (1991) *Methods: A Companion to Methods in Enzymol. 2*, 88−96.
21. Cho, B. K., Schodin, B. A., and Kranz, D. M. (1995) *J. Biol. Chem. 270*, 25819−25826.

22. Proba, K., Honegger, A., and Plückthun, A. (1997) *J. Mol. Biol. 265*, 161−172.
23. Buchner, J., and Rudolph, R. (1991) *BioTechnology 9*, 157−162.
24. Buchner, J., Pastan, I., and Brinkmann, U. (1992) *Anal. Biochem. 205*, 263−270.
25. Suttnar, J., Dyr, J. E., Hamsiková, J., Novák, J., and Vonka, V. (1994) *J. Chromatrogr. B 656*, 123−126.
26. Li, Y., and Inouye, M. (1996) *J. Mol. Biol. 262*, 591−594.
27. Volkov, A., and Jordan F. (1996) *J. Mol. Biol. 262*, 595−599.
28. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S., and Foeller, C. (1992) *Sequences of Proteins of Immunological Interest, Release 5.0*, National Institutes of Health, Bethesda, MD.
29. Herron, J. N., He, X. M., Ballard, D. W., Blier, P. R., Pace, P. E., Bothwell, A. L. M., Voss, E. W., and Edmundson, A. B. (1991) *Proteins: Struct., Funct., Genet. 11*, 159−175.
30. Nicholls, A., Sharp, K., Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.

BI9809167