

variation in the V_L and V_H domains (6). Analysis of antibody primary sequence data has established the existence of two classes of variable region sequence: hypervariable sequences and framework sequences (7). The framework sequences are responsible for the correct β -sheet folding of the V_L and V_H domains and for the interchain interactions that bring domains together. Each variable domain contains three hypervariable sequences, which appear as loops at one end of the β -sheets. The six hypervariable sequences of the variable region, three from the V_L and three from the V_H , form the antigen-binding site and are therefore named the complementarity determining regions (CDRs).

The design of the single-chain antigen-binding protein was based on the assumption that the molecular interactions responsible for the conserved structure, determined by the framework sequences, would assure that proper folding of the individual V_L and V_H domains would occur when tethered by a short peptide linker. Linkers of different designs have been used to join the V_L and V_H sequences.

A computer-assisted method (8) was used in designing the first group of linkers. Since the variable domains of antibodies appear to have homologous three-dimensional structures, we based our modeling on the previously published structure of the Fab fragment of MCPC603, a mouse myeloma protein that binds phosphorylcholine (9). Design of the polypeptide linker was initiated by selecting specific amino acids, one near the carboxyl terminus of the V_L sequence and one near the amino terminus of the V_H sequence. A computer program was then

used to search libraries of three-dimensional peptide structures derived from the Brookhaven Protein Data Bank for peptides of the proper molecular dimensions to span the distance in space between the selected amino acids. The number of potential peptides was reduced by specifying that the angle of the peptide bonds at the ends of the linker peptide match the angles of the bonds at the selected amino acids on either side of the linker. Structures of the remaining potential linker peptides were superimposed onto the MCPC603 variable region structure by computer graphics. The peptides that had a conformation that interfered with the Fv structure were discarded. This process was repeated with different amino acids as linker attachment sites to identify a number of peptides that could be used to link the V_L and V_H regions of antibodies to create single-chain antigen-binding proteins. The linker in single-chain antigen-binding proteins to bovine growth hormone (BGH), 3C2/59, and to fluorescein, 18-2-3/59, was designed by this method. A model of the three-dimensional structure of 3C2/59 protein is shown in Fig. 1.

Additional linkers were designed by the incremental addition of single amino acids or short peptides extending from the carboxyl terminus of the V_L to the amino terminus of the V_H . Some linkers were designed to minimize interactions with the Fv and others were designed to fit into a groove on the back of the Fv structure with the use primarily of alternating glycine and serine residues with glutamic acid and lysine residues included to enhance solubility. The linkers in the single-chain antigen-binding proteins to fluorescein 18-2-3/202 and 4-4-

20/202' were designed in this way. Linker 202 has the sequence EGKSSGSGSESKST. In the 4-4-20/202' protein, an additional amino acid, Gln, is present between the linker and the V_H .

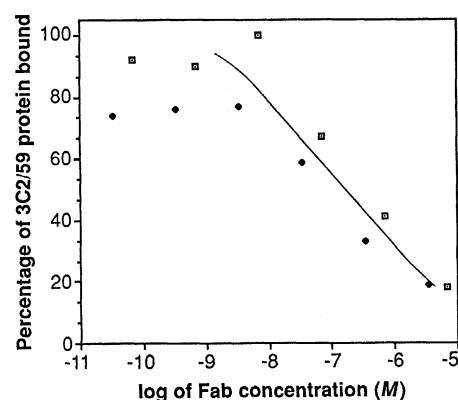
The first single-chain antigen-binding protein genes were constructed from the V_L and V_H sequences of 3C2, an immunoglobulin G1 (IgG1) monoclonal antibody to BGH (anti-BGH). Five versions of an anti-BGH single-chain gene containing different linkers were constructed and expressed in *Escherichia coli*. The resulting proteins were purified, renatured, and tested for their ability to bind to BGH immobilized on nitrocellulose or covalently linked to Sepharose. The anti-BGH single-chain antigen-binding protein containing a linker designated 59 (3C2/59) was chosen for further study based on its affinity for BGH-Sepharose.

When the 3C2/59 gene was expressed in *E. coli*, the single-chain protein accumulated in insoluble inclusion bodies. The 3C2/59 protein had an apparent molecular size of 26 kD as determined by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). This compares favorably with the molecular size of 26,652 daltons calculated from the amino acid sequence. The 3C2/59 protein cross-reacted on immunoblots with antiserum prepared against purified 3C2 light chain.

The anti-BGH 3C2/59 protein was renatured and purified by affinity chromatography on BGH-Sepharose. The affinity-purified protein migrated as a single band of 26 kD when analyzed by SDS-PAGE under both reduced and nonreduced conditions. The amount of protein that was successfully folded and therefore able to bind to a BGH-Sepharose column varied between 5% and 30% in different experiments. To demonstrate that the affinity-purified protein retained binding activity after thiocyanate elution, the protein was analyzed again by chromatography on a second BGH-Sepharose column. More than 90% of this protein bound to BGH-Sepharose and was eluted with thiocyanate. Renatured single-chain protein produced from a modified 3C2/59 gene in which the sequence of five of the six hypervariable regions had been changed did not bind to BGH-Sepharose, demonstrating that binding of 3C2/59 protein to BGH occurs at the antigen-binding site.

The relative affinity of the purified 3C2/59 protein for BGH was determined by competition with Fab fragments isolated from the 3C2 monoclonal antibody. Increasing amounts of unlabeled Fab were mixed with [35 S]methionine-labeled 3C2/59 protein, and the mixture was incubated with BGH-Sepharose. After incubation, the amount of bound labeled protein was determined. The competition curves for two ex-

Fig. 2. The 3C2/59 gene was constructed for easy insertion into an *E. coli* expression vector such that an ATG codon is placed directly in front of the first codon of the light chain sequence. Expression is controlled by the hybrid λ phage promoter (18). The resulting strain containing the CI857 temperature-labile repressor gene and the expression plasmid was induced by raising the culture temperature to 42°C. Overnight growth of the expression strain GX6539 at 42°C resulted in the production of 3C2/59 protein at greater than 10% of the total cell protein. Cells that had been induced in the presence of [35 S]methionine were disrupted by two passes through a French Pressure Cell at 1600 psi. The crude inclusion bodies were recovered by differential centrifugation, dissolved in 50 mM glycine, pH 10.8, 9M urea, 1 mM EDTA, and 20 mM β -mercaptoethanol (19), and diluted to a final protein concentration of ~ 100 μ g/ml in the same buffer. The diluted protein was renatured as described by Boss *et al.* (19). The renatured protein in phosphate-buffered saline (PBS) was bound to a BGH-Sepharose column and eluted with 3M sodium thiocyanate. After dialysis against PBS, 90% of the 3C2/59 protein bound to a second BGH-Sepharose column. We performed competition assays using 8×10^{-7} M 3C2/59 protein, BGH-Sepharose beads, and varying concentrations of Fab derived from the monoclonal antibody in a total volume of 175 μ l. After washing, we counted individual wells containing the Sepharose beads in a Beckman liquid scintillation counter to quantitate binding. The results from two experiments are shown by the different symbols.



periments are shown in Fig. 2. The concentration of Fab that inhibited binding of 3C2/59 protein to BGH by 50% was $2 \times 10^{-7}M$, one-fourth the concentration of the labeled protein ($8 \times 10^{-7}M$), indicating that the equilibrium constant (K_a) of the 3C2/59 single-chain protein was within a factor of 4 of the K_a of the Fab.

Monoclonal antibodies to fluorescein were chosen for continued development of the single-chain antigen-binding protein technology because this antibody-antigen system has been well characterized (10).

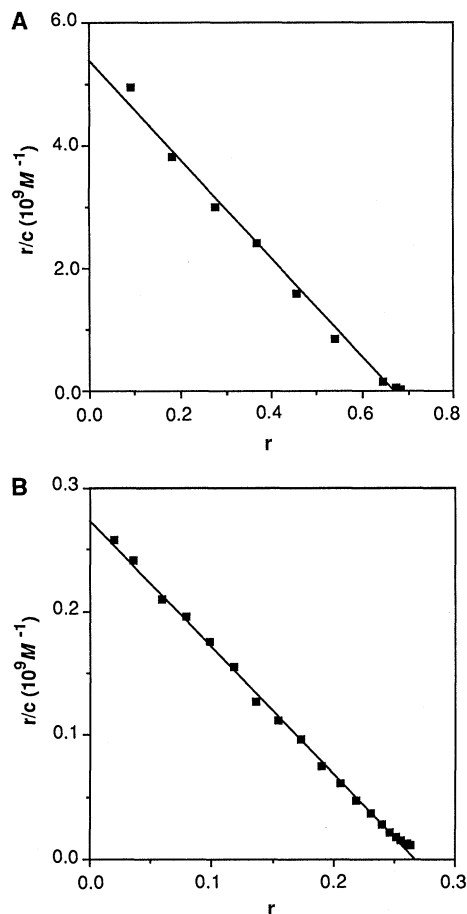


Fig. 3. Analysis of fluorescein binding affinities of the 4-4-20 Fab (A) and 4-4-20/202' protein (B) by the use of a fluorescence quenching assay (11). The r represents the fraction of Fab or 4-4-20/202' protein bound, and c represents the concentration of unbound fluorescein. A sample (590 μ l) was placed in a 0.5-cm path-length cell, and 10- μ l portions of $3.0 \times 10^{-7}M$ fluorescein were added to each sample and to buffer specimens (controls); the fluorescence was measured with a Perkin-Elmer LS-5 fluorescence spectrophotometer. Quenching maxima (previously determined) of 0.965 and 0.970 were used for the 4-4-20 Fab and 4-4-20/202' protein analysis, respectively. We calculated protein concentrations from absorbance at 280 nm using calculated molar extinction coefficients at 280 nm (76,182 and $51,267M^{-1}cm^{-1}$ for the 4-4-20 Fab and 4-4-20/202' protein, respectively, based on the Trp and Tyr content of each protein sequence).

Several monoclonal antibodies with high affinity for fluorescein have been isolated, and quantitative assays for equilibrium constant determinations based on the quenching of fluorescein fluorescence have been described (11). Single-chain antigen-binding protein genes were constructed from the sequences of the variable domains of two different monoclonal antibodies to fluorescein: 18-2-3, an IgM (12), and 4-4-20, in IgG2a (13).

Binding measurements for each sample of anti-fluorescein Ig, Fab, or single-chain antigen-binding protein were performed by fluorescence quenching assays (11). An initial titration of the protein with fluorescein was used to determine the fluorescence quenching maximum (Q_{max}). An estimate of the K_a was then calculated from this titration to select protein and fluorescein concentrations for accurate determination of the K_a . For the monoclonal antibody 4-4-20, the reported values for Q_{max} and K_a are 96.4% and $1.7 \times 10^{10}M^{-1}$, respectively (13).

The 4-4-20/202' protein, renatured from inclusion bodies, was concentrated, dialyzed against 20 mM Tris, pH 8.5, loaded onto a DEAE-SPW high-performance liquid chromatography anion exchange column (LKB), and eluted with a linear gradient of 0 to 2M sodium acetate in 20 mM Tris, pH 8.5. Binding data for samples of 4-4-20 Fabs and 4-4-20/202' protein are shown in Scatchard plots (14) in Fig. 3, A and B, respectively. The K_a s for the 4-4-20 Fab and the 4-4-20/202' protein, calculated from the slopes in Fig. 3, are $8.0 \times 10^9M^{-1}$ and $1.1 \times 10^9M^{-1}$, respectively. The number of binding sites per molecule was determined from the x -intercept. Values of 0.68 and 0.27 indicate the fractions of active Fab and 4-4-20/202' protein in the preparations under these conditions. Thus, the refolding and nonaffinity purification used for the 4-4-20/202' protein produced 27% active protein with high affinity (15).

Kranz *et al.* (13) reported that the absorbance spectrum of fluorescein was shifted from a maximum at 493 nm to a maximum at 505 nm when the fluorescein was bound by the 4-4-20 monoclonal antibody. This shift ranged from about 500 to 525 nm for different antibodies to fluorescein. To test whether this shift occurred with fluorescein bound to a single-chain antigen-binding protein, we measured the fluorescence excitation spectra of the 4-4-20 monoclonal antibody, the Fab fragment prepared from this antibody, and the 4-4-20/202' protein (Fig. 4). The 4-4-20/202' protein caused a similar shift in excitation maximum from 493 to 505 nm as the monoclonal antibody and Fab, demonstrating that fluorescein was bound by the 4-4-20/202' protein in a

manner analogous to the monoclonal antibody.

The 18-2-3/59 and 18-2-3/202' proteins also quenched the fluorescence of fluorescein. The Q_{max} of these proteins was equivalent to that of the 18-2-3 monoclonal antibody. Absolute K_a 's were not determined for these proteins.

In summary, single-chain antigen-binding proteins were constructed from the variable region sequences of three different monoclonal antibodies and various linker peptides designed by computational methods and computer graphics. These proteins retain both the affinity and specificity of the starting monoclonal antibodies. The linker peptide that worked in one single-chain antigen-binding protein was also successful in a different protein, as in the anti-BGH 3C2/59 and anti-fluorescein 18-2-3/59 proteins. We are confident that we can produce active single-chain antigen-binding proteins with the sequence of any monoclonal antibody.

Winter and colleagues (16) showed that the specificity of a monoclonal antibody can be changed by substituting the hypervariable sequences of a mouse antibody into a human framework. It should also be possible to change the specificity of a single-chain antigen-binding protein by replacing the hypervariable sequences with those from a

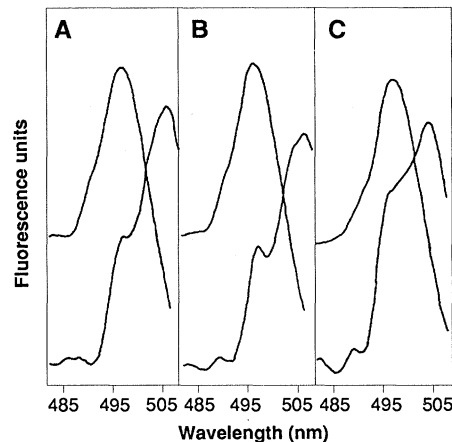


Fig. 4. The fluorescence excitation spectra were obtained on a Perkin-Elmer LS-5 fluorescence spectrometer connected to an R100 recorder and were determined at 2°C with 5-mm path-length microcuvettes in a thermostatted holder. Readings were obtained for samples (600 μ l) of $5 \times 10^{-9}M$ free fluorescein (left peak in all panels) and the same concentration of fluorescein in the presence of (A) 10 μ g/ml 4-4-20 monoclonal antibody, (B) 10 μ g/ml 4-4-20 Fab fragment, and (C) 0.1 μ g/ml 4-4-20/202' protein. The fluorescence emission was monitored at 530 nm while the excitation wavelength was varied from 470 to 515 nm. The nominal absorption maximum of unbound, unquenched fluorescein is about 493 nm [left peak in (A) to (C)] and was monitored on a scale five times that of the protein-bound, quenched fluorescein samples (right peak).

monoclonal antibody having a different specificity. A logical step in the development of a clinically useful antitumor agent would be to insert the hypervariable sequences from antitumor antibodies generated in mice into a single-chain antigen-binding protein derived from human framework regions.

Single-chain antigen-binding proteins are expected to have advantages in clinical applications because of their small size. These proteins should be cleared from serum faster than monoclonal antibodies or Fab fragments. Because they lack the Fc portion of an antibody, which is recognized by cell receptors, they should have a lower background in imaging applications and they should be less immunogenic. They may penetrate the microcirculation surrounding solid tumors better than monoclonal antibodies. We foresee the use of single-chain antigen-binding proteins in applications for which monoclonal antibodies and antibody fragments are currently used, such as (i) imaging and therapy of cancers and cardiovascular or other diseases, (ii) separations, and (iii) biosensors.

The peptide linkers present in single-chain antigen-binding proteins can be designed with specialized function such as sites for the chelation of metals or the attachment of drugs or toxins for applications in imaging and therapy. In addition, it will be possible to design sequences into the linkers or at the carboxyl terminus for the attachment of the protein to solid supports for use in clinical assays, separations, and sensing devices.

Note added in proof: Since this report was submitted another paper dealing with similar technology has appeared (17).

REFERENCES AND NOTES

1. R. R. Porter, *Science* **180**, 713 (1973).
2. P. M. Schiffer *et al.*, *Biochemistry* **12**, 4620 (1973); R. S. Poljak *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3305 (1973).
3. D. R. Davies and H. Metzger, *Annu. Rev. Immunol.* **1**, 87 (1983).
4. R. L. Wahl, C. Racker, G. W. Philpott, *J. Nucl. Med.* **24**, 316 (1983); P. J. Harwood *et al.*, *Eur. J. Cancer Clin. Oncol.* **21**, 1515 (1985).
5. J. Hochman, D. Inbar, D. Givol, *Biochemistry* **12**, 1130 (1973); J. Sen and S. Beychok, *Proteins* **1**, 256 (1986).
6. C. de Pr val and M. Foug r au, *J. Mol. Biol.* **102**, 657 (1976); E. A. Padlan, *Q. Rev. Biophys.* **10**, 35 (1977).
7. E. A. Kabat *et al.*, in *Sequences of Proteins of Immunological Interest* (National Institutes of Health, Bethesda, MD, 1987), pp. 1–261.
8. R. C. Ladner, U.S. Patent 4 704 692 (1987).
9. D. M. Segal *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4298 (1974); Y. Satow, G. H. Cohen, E. A. Padlan, D. R. Davies, *J. Mol. Biol.* **190**, 593 (1986).
10. E. W. Voss, Jr., in *Fluorescein Hapten: An Immunological Probe*, E. W. Voss, Jr., Ed. (CRC Press, Boca Raton, FL, 1984), pp. 3–20; J. N. Herron, D. M. Kranz, E. W. Voss, Jr., *Biochemistry* **25**, 4602 (1986); J. N. Herron, A. L. Gibson, E. W. Voss, Jr., A. B. Edmundson, *Fed. Proc.* **44**, 1328 (1985).
11. J. N. Herron, in *Fluorescein Hapten: An Immunological Probe*, E. W. Voss, Jr., Ed. (CRC Press, Boca Raton, FL, 1984), pp. 49–76.
12. D. W. Ballard, D. M. Kranz, E. W. Voss, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5071 (1983). The V_L and V_H cDNA sequences of the monoclonal antibodies to fluorescein were synthesized by priming on RNA isolated from hybridoma cells with oligonucleotides complementary to the first constant region of each chain. To verify that the isolated cDNA clones encoded the V_L and V_H chains, the amino acid sequences translated from the nucleotide sequences were compared with the amino-terminal amino acid sequences of the parent antibodies.
13. D. M. Kranz, J. N. Herron, E. W. Voss, Jr., *J. Biol. Chem.* **257**, 6987 (1982).
14. G. Scatchard, *Ann. N.Y. Acad. Sci.* **51**, 660 (1949).
15. The majority of experiments have produced K_a 's within a factor of 2 of these values; therefore, log K_a 's for the 4-4-20 IgG, Fab, and 4-4-20/202' single-chain protein are 10.2, 9.9, and 9.0 \pm 3, respectively. This corresponds to a difference in free energy between the Fab and 4-4-20/202' protein of 1.1 \pm 0.4 kcal mol⁻¹.
16. P. T. Jones, P. H. Dear, J. Fotte, M. S. Newberger, G. Winter, *Nature* **321**, 522 (1986); M. Verhoeyen, C. Milstein, G. Winter, *Science* **239**, 1534 (1988); L. Riechmann, M. Clark, H. Waldmann, G. Winter, *Nature* **332**, 323 (1988).
17. D. Scandella, P. Arthur, M. Mattingly, L. Neuhold, *J. Cell Biochem.* **9B**, 203 (1985).
18. J. S. Huston *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5879 (1988).
19. M. A. Boss, J. H. Kenten, C. R. Wood, J. S. Emtage, *Nucleic Acids Res.* **12**, 3791 (1984).
20. This work was supported by NIH grants 1-R43-GM39662-01 and 1-R43-GM39646-01. We wish to thank R. C. Ladner for initiating the thought processes for this project and for designing the linker in 3C2/59. We thank K. C. Huynh-Pham, N. Essig, K. Knistely, and S. Small for technical assistance, D. Filpula and J. Nagle for DNA sequencing, P. Bryan and M. Pantoliano for discussions, and E. Voss for consultations on the antibody to fluorescein system. We also thank our colleagues for reading the manuscript.

11 May 1988; accepted 9 August 1988

Kaposi's Sarcoma Cells: Long-Term Culture with Growth Factor from Retrovirus-Infected CD4⁺ T Cells

SHUJI NAKAMURA, S. ZAKI SALAHUDDIN, PETER BIBERFELD, BARBARA ENSOLI, PHILLIP D. MARKHAM, FLOSSIE WONG-STAAI, ROBERT C. GALLO*

Studies of the biology and pathogenesis of Kaposi's sarcoma (KS) have been hampered by the inability to maintain long-term cultures of KS cells in vitro. In this study AIDS-KS-derived cells with characteristic spindle-like morphology were cultured with a growth factor (or factors) released by CD4⁺ T lymphocytes infected with human T-lymphotropic virus type I or II (HTLV-I or HTLV-II) or with human immunodeficiency virus type 1 or 2 (HIV-1 or HIV-2). Medium conditioned by HTLV-II-infected, transformed lines of T cells (HTLV-II CM) contained large amounts of this growth activity and also supported the temporary growth of normal vascular endothelial cells, but not fibroblasts. Interleukin-1 and tumor necrosis factor- α stimulated the growth of the KS-derived cells, but the growth was only transient and these factors could be distinguished from that in HTLV-II CM. Other known endothelial cell growth promoting factors, such as acidic and basic fibroblast growth factors and epidermal growth factor, did not support the long-term growth of the AIDS-KS cells. The factor released by CD4⁺ T cells infected with human retroviruses should prove useful in studies of the pathogenesis of KS.

K APOSI'S SARCOMA DEVELOPS IN the form of multifocal lesions consisting of characteristic spindle-shaped cells in a stroma of proliferating abnormal vessels, fibroblasts, and infiltrating leukocytes. An indolent form of KS occurs in elderly males in Mediterranean and African countries (1, 2) and a more aggressive, glandular form of the disease occurs in younger Africans. An aggressive form of KS is also associated with HIV-1 infection, primarily in homosexual men (3, 4), and with immune suppression due to other causes (5). A direct transforming involvement of HIV-1 in the development of AIDS-associated KS (AIDS-KS) is unlikely because genomic sequences of the virus have not been detected in KS tissues (6). Furthermore, no other viruses, environmental factors, or genetic factors have been convinc-

ingly linked to any form of KS (4, 6, 7). To gain insight into the nature of KS cells and to search for possible new etiological agents, we developed procedures for establishing cells from AIDS-KS in culture.

A number of growth factors, including endothelial cell growth supplement (ECGS) and fibroblast growth factors (FGF), that were previously shown to stimulate or sup-

S. Nakamura, S. Z. Salahuddin, B. Ensoli, F. Wong-Staal, R. C. Gallo, Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892.

P. Biberfeld, Department of Pathology, Karolinska Institute, 10401 Stockholm, Sweden; and a Visiting Scientist at the Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892.

P. D. Markham, Department of Cell Biology, Bionetics Research, Inc., Rockville, MD 20850.

*To whom correspondence should be addressed.