Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 5879–5883, August 1988
Biochemistry

Protein engineering of antibody binding sites: Recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*

(biosynthetic analogue/anti-digoxin antibodies/variable regions/gene synthesis/*Escherichia coli* expression)

JAMES S. HUSTON*†, DOUGLAS LEVINSON*, MEREDITH MUDGETT-HUNTER‡, MEI-SHENG TAI*, JIŘÍ NOVOTNÝ‡, MICHAEL N. MARGOLIES§, RICHARD J. RIDGE*, ROBERT E. BRUCCOLERI‡, EDGAR HABER‡, ROBERTO CREA*, AND HERMANN OPPERMANN*

*Creative Biomolecules, 35 South Street, Hopkinton, MA 01748; and ‡Molecular and Cellular Research Laboratory, and §Department of Surgery, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114

Communicated by Michael Sela, May 6, 1988 (received for review March 9, 1988)

ABSTRACT **A biosynthetic antibody binding site, which incorporated the variable domains of anti-digoxin monoclonal antibody 26-10 in a single polypeptide chain (M_r = 26,354), was produced in *Escherichia coli* by protein engineering. This variable region fragment (Fv) analogue comprised the 26-10 heavy- and light-chain variable regions (V_H and V_L) connected by a 15-amino acid linker to form a single-chain Fv (sFv). The sFv was designed as a prolyl-V_H-(linker)-V_L sequence of 248 amino acids. A 744-base-pair DNA sequence corresponding to this sFv protein was derived by using an *E. coli* codon preference, and the sFv gene was assembled starting from synthetic oligonucleotides. The sFv polypeptide was expressed as a fusion protein in *E. coli*, using a leader derived from the *trp* LE sequence. The sFv protein was obtained by acid cleavage of the unique Asp-Pro peptide bond engineered at the junction of leader and sFv in the fusion protein [(leader)-Asp-Pro-V_H-(linker)-V_L]. After isolation and renaturation, folded sFv displayed specificity for digoxin and related cardiac glycosides similar to that of natural 26-10 Fab fragments. Binding between affinity-purified sFv and digoxin exhibited an association constant [$K_a = (3.2 \pm 0.9) \times 10^7\ M^{-1}$] that was about a factor of 6 smaller than that found for 26-10 Fab fragments [$K_a = (1.9 \pm 0.2) \times 10^8\ M^{-1}$] under the same buffer conditions, consisting of 0.01 M sodium acetate, pH 5.5/0.25 M urea.**

It is known that antigen binding fragments of antibodies (1, 2) can be refolded from denatured states with recovery of their specific binding activity (3–6). The smallest such fragment that contains a complete binding site is termed Fv, consisting of an M_r 25,000 heterodimer of the V_H and V_L domains (2, 5–11). Givol and coworkers were the first to prepare an Fv by peptic digestion of murine IgA myeloma MOPC 315 (2). However, subsequent development of general cleavage procedures for Fv isolation has met with limited success (7–11). As a result, the M_r 50,000 Fab (1) has remained the only monovalent binding fragment used routinely in biomedical applications.

An Fv analogue was constructed in which both heavy- and light-chain variable domains (V_H and V_L) were part of a single polypeptide chain. Synthetic genes for the 26-10 anti-digoxin V_H and V_L regions were designed to permit their connection through a linker segment, as well as other manipulations (12, 13). The synthetic gene for single-chain Fv (sFv) was expressed in *Escherichia coli* as a fusion protein, from which the sFv protein was isolated.¶ The sFv was renatured with recovery of binding specificity and affinity similar to those of the parent molecule. Thus, variable domains connected artificially to form one polypeptide chain can be renatured into properly folded Fv regions.

MATERIALS AND METHODS

Model Antibody. The digoxin binding site of the IgG2a,κ monoclonal antibody 26-10 has been analyzed by Mudgett-Hunter and colleagues (14–16). The 26-10 V region sequences were determined from both protein sequencing (17) and DNA sequencing of 26-10 H- and L-chain mRNA transcripts (D. Panka, J.N., and M.N.M., unpublished data). The 26-10 antibody exhibits a high digoxin binding affinity ($K_a = 5.4 \times 10^9\ M^{-1}$) (14) and has a well-defined specificity profile (15), providing a baseline for comparison with the biosynthetic sFv.

Protein Design. X-ray coordinates for Fab fragments (18–20) were obtained from the Brookhaven Data Bank (21) and analyzed with the programs CHARMM (22), CONGEN (23), and FRODO (24). X-ray data indicated that the Euclidean distance between the C terminus of the V_H domain and the N terminus of the V_L domain was ≈3.5 nm. A 15-residue linker should bridge this gap, since the peptide unit length is ≈0.38 nm. The linker should not exhibit a propensity for ordered secondary structure or any tendency to interfere with domain folding. Thus, the 15-residue sequence $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$ was selected to connect the V_H carboxyl and V_L amino termini (Fig. 1).

Gene Synthesis. Design of the 744-base sequence for the synthetic sFv gene was derived from the sFv protein sequence by choosing codons preferred by *E. coli* (25). Synthetic genes encoding the *trp* promoter-operator, the modified *trp* LE leader peptide (MLE), and V_H were prepared largely as described (26). The gene encoding V_H was assembled from 46 overlapping synthetic 15-base oligonucleotides. The V_L gene was derived from 12 synthetic polynucleotides ranging in size from 33 to 88 base pairs, prepared in automated DNA synthesizers (model 6500, Biosearch, San Rafael, CA; model 380A, Applied Biosystems, Foster City, CA). They spanned major restriction sites (*Aat* II, *BstE*II, *Kpn* I, *Hin*dIII, *Bgl* I, and *Pst* I), and several fragments were flanked by *Eco*RI and *Bam*HI cloning ends. All segments were cloned and assembled in pUC vectors. The linker between V_H and V_L, encoding $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$, was cloned from two polynucleotides spanning *Sac* I and *Aat* II sites. The complete sFv gene was assembled from the V_H, V_L, and linker genes to yield a single sFv gene, corresponding

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*"

Abbreviations: Fv, variable region fragment (V_HV_L dimer); MLE, modified *trp* LE leader sequence; sFv, single-chain Fv; V_H, heavy-chain variable region; V_L, light-chain variable region.

†To whom reprint requests should be addressed.

¶The sequence reported in this paper is being deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA,

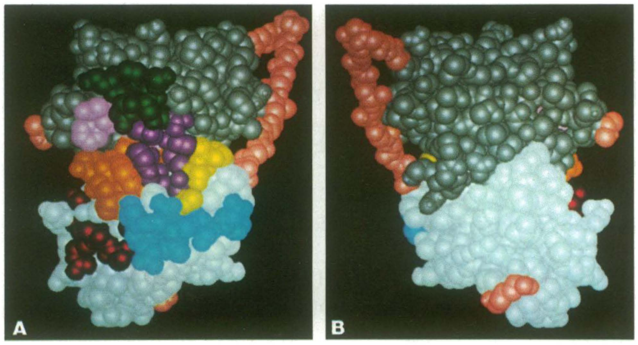


FIG. 1. Computer-generated views of an sFv model. These pictures display the sFv based on the x-ray structure of the Fab from murine IgA myeloma McPC 603 (20). One possible conformation of the linker (Gly-Gly-Gly-Gly-Ser)$_3$ is shown connecting the C terminus of the V_H domain and the N terminus of the V_L domain. (*A*) View showing the linker and binding site of McPC 603. (*B*) View of opposite side showing linker and free terminal residues (Asp-1 of V_H and Lys-113 of V_L); this orientation was obtained by rotating the molecule in (*A*) 180° about a vertical axis. Color coding: gray, V_H domain; white, V_L domain; salmon, linker and free terminal residues; other colors, complementarity determining region (CDR) segments [lavender (H1), green (H2), orange (H3), blue (L1), red (L2), yellow (L3)]; purple, side chains from CDRs that directly contact the phosphorylcholine hapten in McPC 603.

to aspartyl-prolyl-V_H-(linker)-V_L, flanked by *Eco*RI and *Pst* I restriction sites (Fig. 2). The *trp* promoter-operator, starting from the *Ssp* I site, and the MLE leader gene, ending in the *Eco*RI site, were assembled from 36 overlapping 15-base oligomers. The final expression plasmid, based on the pBR322 vector, was constructed by a three-part ligation using the sites *Ssp* I, *Eco*RI, and *Pst* I (Fig. 2*D*). Intermediate DNA fragments and assembled genes were sequenced by the dideoxy chain-termination method (28).

Fusion Protein Expression. Single-chain Fv was expressed as a fusion protein (Fig. 2) with the MLE leader gene (29) and under the control of a synthetic *trp* promoter-operator (29). *E. coli* strain JM83 was transformed with the expression plasmid and protein expression was induced in M9 minimal medium by addition of indoleacrylic acid (10 μg/ml) at a cell density with A_{600} = 1. The high expression levels of the fusion protein resulted in its accumulation as insoluble protein granules, which were harvested from cell paste (Fig. 3, lane 1).

Fusion Protein Cleavage. The MLE leader was removed from sFv by acid cleavage of the Asp-Pro peptide bond (32–34) engineered at the junction of the MLE and sFv sequences. The washed protein granules containing the fusion protein were cleaved in 6 M guanidine hydrochloride/10% acetic acid, pH 2.5, incubated at 37°C for 96 hr. The reaction was stopped by ethanol precipitation, and the precipitate was stored at −20°C (Fig. 3, lane 2).

Protein Purification. The acid-cleaved sFv was separated from remaining intact MLE–sFv species by chromatography on DEAE-cellulose. The precipitated cleavage mixture was redissolved in 6 M guanidine hydrochloride/0.2 M Tris·HCl, pH 8.2/0.1 M 2-mercaptoethanol, and dialyzed exhaustively against column buffer (6 M urea/2.5 mM Tris·HCl, pH 7.5/1 mM EDTA), made 0.1 M in 2-mercaptoethanol, and chromatographed on a column (2.5 × 45 cm) of Whatman DE 52. Elution of the intact fusion protein was retarded relative to sFv during the column buffer wash, with leading fractions

A

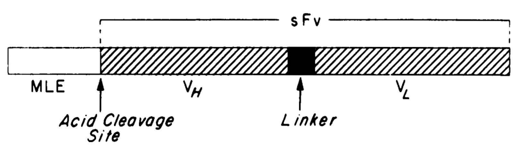
```
                                    10                                      20                                      30                                      40
Met Lys Ala Ile Phe Val Leu Lys Gly Ser Leu Asp Arg Asp Leu Asp Ser Arg Leu Asp Leu Asp Val Arg Thr Asp His Lys Asp Leu Ser Asp His Leu Val Leu Val Asp Leu Ala
ATG AAA GCA ATT TTC GTA CTG AAA GGT TCA CTG GAC AGA GAT CTG GAC TCT CGT CTG GAT CTG GAC GTT CGT ACC GAC CAC AAA GAC CTG TCT GAT CAC CTG GTT CTG GTC GAC CTG GCT
                                                    BglII                                                                                           SalI
                                    50                                      60
Arg Asn Asp Leu Ala Arg Ile Val Thr Pro Gly Ser Arg Tyr Val Ala Asp Leu Glu Phe Asp
CGT AAC GAC CTG GCT CGT ATC GTT ACT CCC GGG TCT CGT TAC GTT GCG GAT CTG GAA TTC GAT
                                    SmaI                                    EcoRI
```

B

```
                                    10                                      20                                      30                                      40
Pro Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys Pro Gly Ala Ser Val Arg Met Ser Cys Lys Ser Ser Gly Tyr Ile Phe Thr Asp Phe Tyr Met Asn Trp Val Arg Gln
CCC GAA GTT CAA CTG CAA CAG TCT GGT CCT GAA TTG GTT AAA CCT GGC GCC TCT GTG CGC ATG TCC TGC AAA TCC TCT GGG TAC ATT TTC ACC GAC TTC TAC ATG AAT TGG GTT CGC CAG
                                                            NarI            FspI                                                                            BstXI
                                    50                                      60                                      70                                      80
Ser His Gly Lys Ser Leu Asp Tyr Ile Gly Tyr Ile Ser Pro Tyr Ser Gly Val Thr Gly Tyr Asn Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala
TCT CAT GGT AAG TCT CTA GAC TAC ATC GGG TAC ATT TCC CCA TAC TCT GGG GTT ACC GGC TAC AAC CAG AAG TTT AAA GGT AAG GCG ACC CTT ACT GTC GAC AAA TCT TCC TCA ACT GCT
                XbaI                                PflMI       BstEII                              DraI                    SalI
                                    90                                     100                                     110                                     120
Tyr Met Glu Leu Arg Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Gly Ser Ser Gly Asn Lys Trp Ala Met Asp Tyr Trp Gly His Gly Ala Ser Val Thr Val Ser Ser
TAC ATG GAG CTG CGT TCT TTG ACC TCT GAG GAC TCC GCG GTA TAC TAT TGC GCG GGC TCC TCT GGT AAC AAA TGG GCC ATG GAT TAT TGG GGT CAT GGT GCT AGC GTT ACT GTG AGC TCT
                                        SacII                                                   NcoI                            NheI            SacI
                                   130                                     140                                     150                                     160
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Val Val Met Thr Gln Thr Pro Leu Ser Leu Pro Val Ser Leu Gly Asp Gln Ala Ser Ile Ser Cys Arg Ser
GGT GGC GGT GGC TCG GGC GGT GGT GGG TCG GGT GGC GGC GGA TCT GAC GTC GTA ATG ACC CAG ACT CCG CTG TCT CTG CCG GTT TCT CTG GGT GAC CAG GCT TCT ATT TCT TGC CGC TCT
                                                            AatII                                   BstEII
                                   170                                     180                                     190                                     200
Ser Gln Ser Leu Val His Ser Asn Gly Asn Thr Tyr Leu Asn Trp Tyr Leu Gln Lys Ala Gly Gln Ser Pro Lys Leu Leu Ile Tyr Lys Val Ser Asn Arg Phe Ser Gly Val Pro Asp
TCC CAG TCT CTG GTC CAT TCT AAT GGT AAC ACT TAC CTG AAC TGG TAC CTG CAA AAG GCT GGT CAG TCT CCG AAG CTT CTG ATC TAC AAA GTC TCT AAC CGC TTC TCT GGT GTC CCG GAT
  PflMI       BstXI                                           KpnI                                  HindIII
                                   210                                     220                                     230                                     240
Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile Ser Arg Val Glu Ala Glu Asp Leu Gly Ile Tyr Phe Cys Ser Gln Thr Thr His Val Pro Pro Thr Phe Gly Gly
CGT TTC TCT GGT TCT GGT TCT GGT ACT GAC TTC ACC CTG AAG ATC TCT CGT GTC GAG GCC GAA GAC CTG GGT ATC TAC TTC TGC TCT CAG ACT ACT CAT GTA CCG CCG ACT TTT GGT GGT
                                                      BglII

Gly Thr Lys Leu Glu Ile Lys Arg *oc
GGC ACC AAG CTC GAG ATT AAA CGT TAA CTGCAG
            XhoI            HpaI    PstI
```

C



D

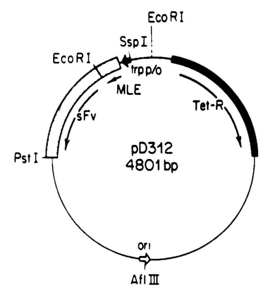


FIG. 2. Sequence of MLE–sFv fusion protein and sFv protein, with corresponding DNA sequence and some major restriction sites, and design of expression plasmid. (*A*) Leader peptide. (*B*) sFv protein. (*C*) Schematic drawing of the fusion protein. (*D*) Schematic drawing of expression plasmid. The leader and sFv are shown as they appear after acid cleavage of the fusion protein. During construction of the gene, fusion partners were joined at the *Eco*RI site that is shown as part of the leader sequence. The complementarity determining regions of V_H and V_L are boldface and the linker peptide is underlined. The pBR322 plasmid, opened at the unique *Ssp* I and *Pst* I sites, was combined in a three-part ligation with an *Ssp* I/*Eco*RI fragment bearing the *trp* promoter-operator and MLE leader and with an *Eco*RI/*Pst* I fragment carrying the sFv gene. The resulting expression vector of 4801 base pairs (bp) (*D*) confers tetracycline resistance on positive transformants.

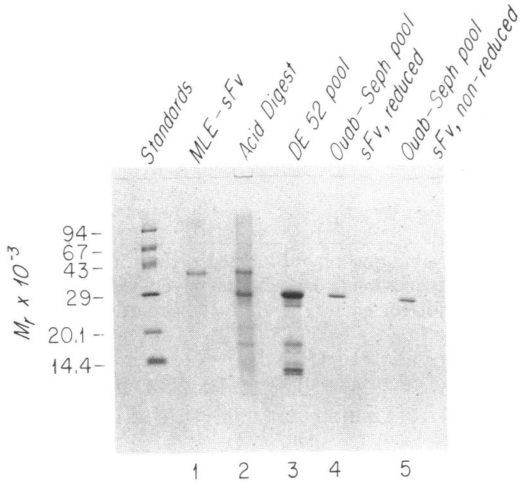


FIG. 3. Analysis of protein at progressive stages of purification by $NaDodSO_4$/PAGE on a 15% polyacrylamide gel (30, 31). The same ouabain-Sepharose pool of sFv was run in lanes 4 and 5, but the gel sample for lane 5 was not reduced, while all others were reduced before electrophoresis. The M_r values calculated for sFv (26,354) and MLE-sFv (33,203) polypeptides were less than gel migration indicated, as a result of the low mean residue weight for sFv (106) and other sources of error in $NaDodSO_4$/PAGE experiments (30).

being devoid of MLE–sFv. Rechromatography of impure fractions yielded more purified material that was added to the DE 52 pool (Fig. 3, lane 3).

Refolding. The DE 52 pool of sFv in 6 M urea/2.5 mM Tris·HCl/1 mM EDTA was adjusted to pH 8 and reduced with 0.1 M 2-mercaptoethanol at 37°C for 90 min. This was diluted at least 1:100 with 0.01 M sodium acetate (pH 5.5) to a concentration below 10 μg/ml and dialyzed at 4°C for 2 days against acetate buffer. Under these conditions, sFv remained soluble throughout refolding, whereas substitution of 0.15 M NaCl/0.05 M potassium phosphate, pH 7.0/0.03% NaN_3 (PBSA) caused precipitation of sFv.

Affinity Chromatography. Active sFv was purified by affinity chromatography at 4°C on a ouabain-amine-Sepharose column, as described (14), except that all eluants were dissolved in 0.01 M sodium acetate (pH 5.5). Bound sFv was eluted from the resin with 20 mM ouabain (Fig. 3, lanes 4 and 5) and dialyzed against 0.01 M acetate buffer. Protein concentrations were quantitated by amino acid analysis (35) (Table 1).

Sequence Analysis of Gene and Protein. The complete sFv gene was sequenced in both directions by the dideoxy method of Sanger *et al.* (28). Automated Edman degradation was conducted on intact sFv protein, as well as on two major CNBr fragments (residues 108–129 and 140–159) with a model 470A gas-phase sequencer equipped with a model 120A on-line analyzer (Applied Biosystems) (37). The CNBr fragments of gel-purified sFv were separated by $NaDodSO_4$/PAGE and transferred electrophoretically onto an Immobilon membrane (Millipore) from which stained bands were cut out and sequenced (38).

Table 1. Estimated yields during purification, normalized to a 1-liter fermentation

Step	Wet pellet weight, g	Protein weight, mg	Mol % of prior step
Cell paste	12	1440*	
MLE–sFv granules	2.3	480*†	100
DE 52 pool		144†‡	38
Active sFv		18.1‡	12.6§

*Determined by Lowry analysis (36).
†Determined by absorbance measurements.
‡Determined by amino acid analysis.
§Calculated from the concentration of sFv protein specifically eluted from ouabain-Sepharose, compared to DE 52 pool, both measured after dialysis against 0.01 M acetate buffer by amino acid analysis. Yield is 4.7% relative to MLE–sFv.

Specificity and Affinity Determinations. Specificities of anti-digoxin 26-10 Fab and sFv were assessed by radioimmunoassay (16) (Fig. 4 and Table 2). Equilibrium binding measurements utilized immunoprecipitation techniques to separate bound and free [^{3}H]digoxin, with association constants calculated from Sips plots (39) and binding isotherms (40) (Fig. 5) as well as Scatchard plots (16).

RESULTS

DNA sequencing of the complete sFv gene and its major oligonucleotide fragments indicated that the sFv gene possessed the intended sequence (Fig. 2), incorporating the V_H and V_L sequences of monoclonal antibody 26-10. The protein sequence expected for this sFv gene product was confirmed at the amino terminus of intact sFv (residues 1–15 for affinity-purified sFv; residues 1–40 for DE 52 pool) and for CNBr fragments over internal regions extending from residue 107 of V_H into the linker (sFv residues 108–129) and over V_L residues 5–24 (sFv residues 140–159). These results suggest that purified sFv protein was a faithful expression product of the synthetic sFv gene.

The yields of protein at various stages of isolation are given in Table 1, with purity assessed by $NaDodSO_4$/PAGE (Fig. 3). Affinity chromatography of the renatured DE 52 pool provided the final purification step (Fig. 3, lanes 4 and 5) and yielded 12.6% of the renatured protein as active sFv (Table 1). Successful affinity purification of very dilute renatured sFv suggests that monomeric protein was adsorbed to resin. Thus, all of the sFv bound to ouabain-Sepharose may be considered to have been active, with stoichiometric binding capacity for digoxin.

The sFv specificity profile was reproduced for samples in various buffers at different stages of purification, appearing

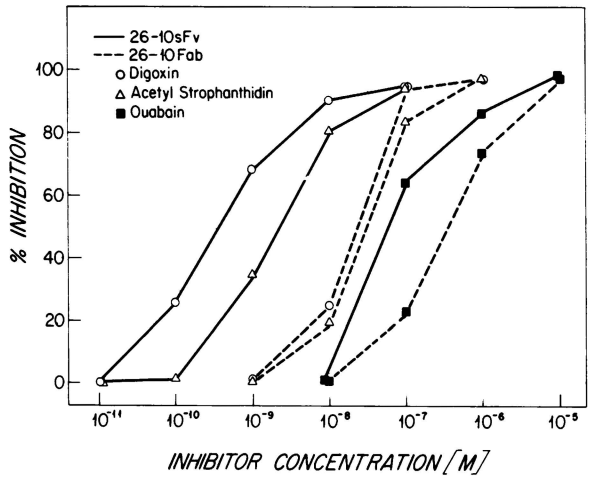


FIG. 4. Specificity profiles for sFv and 26-10 Fab species. Microtiter plates were coated first with affinity-purified goat anti-mouse Fab antibody, followed by sFv or 26-10 Fab in 1% horse serum/0.01 M sodium acetate, pH 5.5. Then ^{125}I-labeled digoxin (50,000 cpm) having specific activity of 1800 μCi/μg (Cambridge Diagnostics; 1 Ci = 37 GBq) was added in the presence of a series of glycoside concentrations. The inhibition of radioligand binding by each of seven cardiac glycosides was plotted and relative affinities for each digoxin analogue were calculated (Table 2). The sFv inhibition curves have been displaced to lower glycoside concentrations than corresponding 26-10 Fab curves, because the concentration of active binding sites on the plate was less for sFv than for 26-10 Fab. When 0.25 M urea was added to the sFv in 0.01 M sodium acetate (pH 5.5), more active sFv bound to the goat anti-mouse Fab on the plate. Hence, the sFv specificity profile shifted toward higher glycoside concentrations, closer to the position of that for 26-10 Fab

Table 2. Specificity analysis

26-10 antibody species	Normalizing glycoside	Digoxin	Digoxigenin	Digitoxin	Digitoxigenin	Acetyl strophanthidin	Gitoxin	Ouabain
Fab	Digoxin	1.0	1.2	0.9	1.0	1.3	9.6	15
	Digoxigenin	0.9	1.0	0.8	0.9	1.1	8.1	13
sFv	Digoxin	1.0	7.3	2.0	2.6	5.9	62	150
	Digoxigenin	0.1	1.0	0.3	0.4	0.8	8.5	21

Results are expressed as normalized concentration of inhibitor giving 50% inhibition of ^{125}I-labeled digoxin binding. Relative affinities for each digoxin analogue were calculated by dividing the concentration of each cardiac glycoside at 50% inhibition by the concentration of digoxin or digoxigenin that gave 50% inhibition for each type of 26-10 species.

to be independent of refolding conditions (data not shown). The comparison of sFv with 26-10 Fab revealed some differences between their specificity profiles. Relative to acetyl strophanthidin, digoxin bound more tightly to the sFv than to 26-10 Fab, ouabain bound less tightly (Fig. 4), and other cardiac glycosides exhibited slight shifts in specificity (Table 2). When relative affinities are expressed in relation to digoxigenin, good agreement between sFv and 26-10 Fab values can be found for all analogues except digoxin. These data indicate that the sFv is slightly more specific for digoxin than the parent 26-10 Fab, but that major features of the 26-10 combining site have been reproduced in the sFv.

Although the sFv bound to ouabain-Sepharose was fully active, upon elution a substantial fraction of protein appeared to form inactive aggregates. The extent of sFv self-association was aggravated further in PBSA at pH 7 but could be reduced by keeping the sFv in dilute acetate buffer at pH 5.5, and further minimized by adding urea to a concentration of 0.25 M. This low concentration of urea enhanced activity by an order of magnitude without any apparent change in the K_a observed in acetate buffer alone. Under these conditions, active binding species represented 22% of sFv and 43% of 26-10 Fab in solution, based on active site concentration (R_1) from Scatchard analysis (R_1 [sFv] = $(2.9 \pm 0.4) \times 10^{-8}$ M; R_1 [26-10 Fab] = $(9.4 \pm 0.5) \times 10^{-9}$ M) and total protein concentration from amino acid analysis ([sFv] = 1.3×10^{-7} M; [26-10 Fab] = 2.2×10^{-8} M).

The association constant for digoxin binding (K_a) was determined from binding data (Fig. 5) by binding isotherm analysis (40) [K_a (sFv) = 5.2×10^7 M^{-1}; K_a (Fab) = 3.3×10^8 M^{-1}], by Sips analysis using linear regression analysis to calculate K_a (39) [K_a (sFv) = 2.6×10^7 M^{-1}; K_a (Fab) = 1.8×10^8 M^{-1}], and by Scatchard analysis (16) [K_a (sFv) = $(3.2 \pm 0.9) \times 10^7$ M^{-1}; K_a (Fab) = $(1.9 \pm 0.2) \times 10^8$ M^{-1}]. In summary, K_a for sFv was 3–5 $\times 10^7$ M^{-1}, while K_a for 26-10 Fab ranged from 2 to 3 $\times 10^8$ M^{-1}, under the conditions of 0.25 M urea in 0.01 M sodium acetate (pH 5.5). Since the 26-10 Fab had a lower K_a in this buffer than in PBSA at pH 7 (K_a = 3.3×10^9 M^{-1}) (12), the sFv binding constant may have been similarly reduced.

DISCUSSION

A single-chain biosynthetic antibody binding site was shown to closely mimic the antigen binding affinity and specificity of the parent antibody. Connection of V_H and V_L by a 15-residue linker may substitute for constant region contacts in the Fab

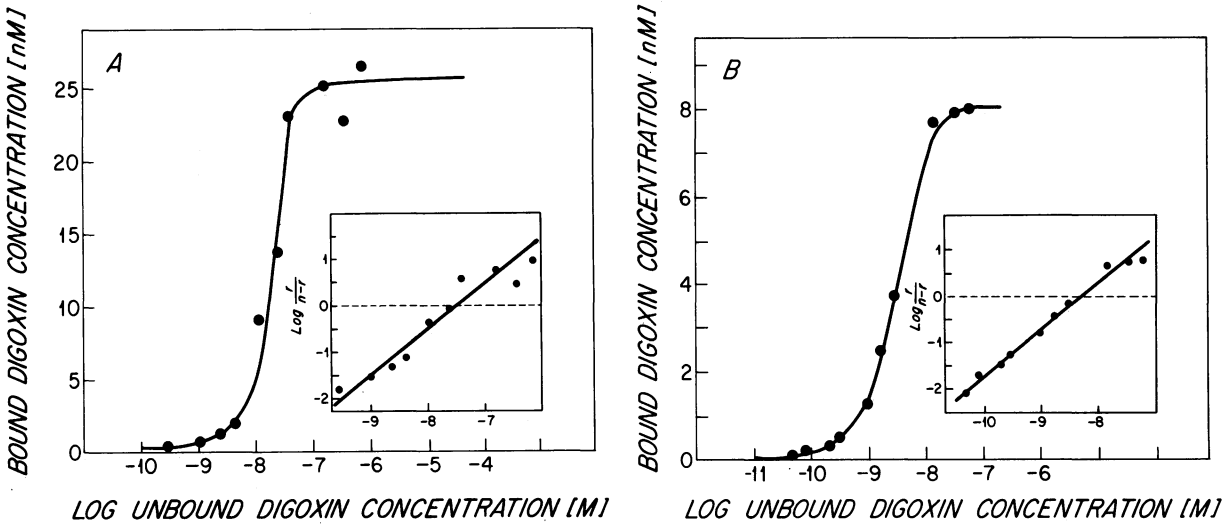


FIG. 5. Analysis of digoxin binding affinity. (*A*) sFv binding isotherm and Sips plot (*Inset*). (*B*) 26-10 Fab binding isotherm and Sips plot (*Inset*). Binding isotherms display data plotted as the concentration of digoxin bound versus the log of the unbound digoxin concentration, and the dissociation constant corresponds to the ligand concentration at 50% saturation (40). Sips plots (*Inset*) present the data in linear form, with the same abscissa as the binding isotherm but with the ordinate representing $\log[r/(n - r)]$ (defined below). The average intrinsic association constant (K_a) was calculated from the modified Sips equation (39), $\log[r/(n - r)] = a \log C - a \log K_a$, where r equals mol of digoxin bound per mol of antibody at an unbound digoxin concentration equal to C; n is mol of digoxin bound at saturation of the antibody binding site, and a is an index of heterogeneity, which describes the distribution of association constants about the average intrinsic association constant, K_a. Least-squares linear regression analysis of the data indicated correlation coefficients for the lines obtained were 0.96 for sFv and 0.99 for 26-10 Fab. Equilibrium binding was conducted in solution, as follows. Aliquots (100 μl) of [^{3}H]digoxin at a series of concentrations (10^{-6}–10^{-10} M) in 0.01 M sodium acetate (pH 5.5) with 1% bovine serum albumin were added to 26-10 Fab or sFv (100 μl) at a fixed concentration in 0.01 M sodium acetate, pH 5.5/0.5 M urea/1% bovine serum albumin. After 2–3 hr of incubation at room temperature, the protein was precipitated by the successive addition of goat anti-mouse Fab serum, the IgG fraction of rabbit anti-goat IgG, and protein A-Sepharose. After 2 hr on ice, bound and free [^{3}H]digoxin were separated by vacuum filtration of samples, and radioligand bound to the protein entrapped on glass fiber filters was measured by scintillation counting.

and thereby aid recovery of native binding properties in the sFv. For example, the Fv(GAR) exhibited a reduction by a factor of 1000 in the riboflavin binding affinity of Fab(GAR) (11), while the MOPC 315 Fv bound dinitrophenol almost as well as the parent antibody (2). Construction of a single-chain Fv may have minimized the refolding problems of two-chain species, such as incorrect domain pairing or aggregation during the renaturation process (41). In fact, reconstitution of 26-10 Fv from separately cloned V_H and V_L domains has thus far proven unsuccessful (J.S.H. and M.M.-H., unpublished data), while *in vitro* recombination of 26-10 H and L chains produced a low yield of antibody with significantly reduced affinity for digoxin (16). Furthermore, past efforts to produce antibodies from cloned H and L chains gave very low recoveries (42–44). The present 12.6% yield of active sFv is 9 times greater than that reported for antibody activity regained from H and L chains expressed in *E. coli* (42).

The sFv and 26-10 Fab both contain identical V_H and V_L polypeptide sequences, but other features of sFv covalent structure might perturb 26-10 combining site properties in the single-chain Fv. Proline has been added to the V_H N terminus and connection of V regions via the linker has eliminated the charge on the V_L α-amino group. The V region N termini may be in sufficient proximity of the combining site to influence binding properties. In a mutant of another anti-digoxin antibody, deletion of the first two residues of V_H dramatically changed its binding affinity (45). Introduction of the linker has also eliminated the V_H terminal carboxyl and may introduce constraints on folded V domains.

Given the feasibility of making the 26-10 single-chain Fv biosynthetically, protein engineering can be used to advantage in further studies. Variation in protein association has been related to restricted changes in primary sequence (46, 47), and one may therefore expect that aggregation of 26-10 sFv can be moderated by alteration of surface residues linked to self-association. Binding site variants of 26-10 sFv may likewise be constructed (13), or its entire framework replaced, while keeping complementarity determining region sequences unchanged (12). The immunopharmacology of biosynthetic antibody binding sites could prove particularly interesting, insofar as their small size may accelerate the pharmacokinetics and reduce the immunogenicity observed for Fab fragments administered intravenously (48). Further research on the single-chain Fv and related immunoconjugates may lead to biomedical applications that have been heretofore impossible with conventional antibody fragments.

We are grateful for the expertise of Sarah Hardy, Abbie White, Denise Maratea, Clare Corbett, Rou-Fun Kwong, Larry Haith, Gay-May Wu, and Robert Juffras, and for the encouragement of Charles Cohen and Prof. Serge Timasheff. This project was supported in part by the National Institutes of Health through SBIR Grant CA 39870 and by HL 19259.

1. Porter, R. R. (1959) *Biochem. J.* **73,** 119–126.
2. Inbar, D., Hochman, J. & Givol, D. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 2659–2662.
3. Haber, E. (1964) *Proc. Natl. Acad. Sci. USA* **52,** 1099–1106.
4. Whitney, P. L. & Tanford, C. (1965) *Proc. Natl. Acad. Sci. USA* **53,** 524–532.
5. Hochman, J., Inbar, D. & Givol, D. (1973) *Biochemistry* **12,** 1130–1135.
6. Hochman, J., Gavish, M., Inbar, D. & Givol, D. (1976) *Biochemistry* **15,** 2706–2710.
7. Sharon, J. & Givol, D. (1976) *Biochemistry* **15,** 1591–1594.
8. Kakimoto, K. & Onoue, K. (1974) *J. Immunol.* **112,** 1373–1382.
9. Lin, L.-C. & Putnam, F. W. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 2649–2653.
10. Reth, M., Imanishi-Kari, T. & Rajewsky, K. (1979) *Eur. J. Immunol.* **9,** 1004–1013.
11. Sen, J. & Beychok, S. (1986) *Proteins: Struct. Funct. Genet.* **1,** 256–262.
12. Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S. & Winter, G. (1986) *Nature (London)* **321,** 522–525.
13. Sharon, J., Gefter, M. L., Manser, T. & Ptashne, M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2628–2631.
14. Mudgett-Hunter, M., Margolies, M. N., Ju, A. & Haber, E. (1982) *J. Immunol.* **129,** 1165–1172.
15. Mudgett-Hunter, M., Anderson, W., Haber, E. & Margolies, M. N. (1985) *Mol. Immunol.* **22,** 477–488.
16. Hudson, N. W., Mudgett-Hunter, M., Panka, D. J. & Margolies, M. N. (1987) *J. Immunol.* **139,** 2715–2723.
17. Novotny, J. & Margolies, M. N. (1983) *Biochemistry* **77,** 1155–1158.
18. Saul, F., Amzel, L. M. & Poljak, R. J. (1978) *J. Biol. Chem.* **253,** 585–597.
19. Marquart, M., Deisenhofer, J., Huber, R. & Palm, W. (1980) *J. Mol. Biol.* **141,** 369–391.
20. Satow, Y., Cohen, G. H., Padlan, E. A. & Davies, D. R. (1986) *J. Mol. Biol.* **190,** 593–604.
21. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. J. (1977) *J. Mol. Biol.* **112,** 535–542.
22. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comput. Chem.* **4,** 187–217.
23. Bruccoleri, R. E. & Karplus, M. (1987) *Biopolymers* **26,** 137–168.
24. Jones, T. A. (1978) *J. Appl. Crystallogr.* **11,** 268–272.
25. Grantham, R., Gautier, C. & Govy, M. (1980) *Nucleic Acids Res.* **8,** 1893–1912.
26. Roberts, D. M., Crea, R., Malecha, M., Alvarado-Urbina, G., Chiarello, R. H. & Watterson, D. M. (1985) *Biochemistry* **24,** 5090–5098.
27. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **3,** 103–119.
28. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
29. Miozzari, G. & Yanofsky, C. J. (1978) *J. Bacteriol.* **133,** 1457–1466.
30. Weber, K., Pringle, J. R. & Osborn, M. (1972) *Methods Enzymol.* **26,** 3–27.
31. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
32. Piszkiewicz, D., Landon, M. & Smith, E. L. (1970) *Biochem. Biophys. Res. Commun.* **40,** 1173–1178.
33. Fraser, K. J., Poulsen, K. & Haber, E. (1972) *Biochemistry* **11,** 4974–4977.
34. Poulsen, K., Fraser, K. J. & Haber, E. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 2495–2499.
35. Bidlingmeyer, B. A., Cohen, S. A. & Tarvin, T. L. (1984) *J. Chromatogr.* **336,** 93–104.
36. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* **193,** 265–275.
37. Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. (1981) *J. Biol. Chem.* **256,** 7990–7997.
38. Matsudaira, P. (1987) *J. Biol. Chem.* **262,** 10035–10038.
39. Smith, T. W., Butler, V. P., Jr., & Haber, E. (1970) *Biochemistry* **9,** 331–337.
40. Klotz, I. M. (1982) *Science* **217,** 1247–1249.
41. Jaenicke, R. (1984) *Angew. Chem.* **23,** 395–413.
42. Cabilly, S., Riggs, A. D., Pande, H., Shively, J. E., Holmes, W. M., Rey, M., Perry, L. J., Wetzel, R. & Heyneker, H. L. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3273–3277.
43. Boss, M. A., Kenten, J. H., Wood, C. R. & Emtage, J. S. (1984) *Nucleic Acids Res.* **12,** 3791–3806.
44. Wood, C. R., Boss, M. A., Kenten, J. H., Calvert, J. E., Roberts, N. A. & Emtage, J. S. (1985) *Nature (London)* **314,** 446–449.
45. Panka, D. J., Mudgett-Hunter, M., Parks, D. R., Peterson, L. L., Herzenberg, L. A., Haber, E. & Margolies, M. N. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 3080–3084.
46. Kumosinski, T. F. & Timasheff, S. N. (1966) *J. Am. Chem. Soc.* **88,** 5635–5642.
47. Huston, J. S., Björk, I. & Tanford, C. (1972) *Biochemistry* **11,** 4256–4262.
48. Smith, T. W., Lloyd, B. L., Spicer, N. & Haber, E. (1979) *Clin. Exp. Immunol.* **36,** 384–396.