Biophysical Journal Volume 71 October 1996 2002–2011

Symmetry of Fv Architecture Is Conducive to Grafting a Second Antibody Binding Site in the Fv Region*

Peter C. Keck and James S. Huston
Creative BioMolecules, Hopkinton, Massachusetts 01748 USA

ABSTRACT Molecular modeling studies on antibody Fv regions have been pursued to design a second antigen-binding site (χ-site) in a chimeric single-chain Fv (χsFv) species of about 30 kDa. This analysis has uncovered an architectural basis common to many Fv regions that permits grafting a χ-site onto the Fv surface that diametrically opposes the normal combining site. By using molecular graphics analysis, chimeric complementarity-determining regions (χCDRs) were defined that comprised most of the CDRs from an antibody binding site of interest. The chain directionality of χCDRs was consistent with that of specific bottom loops of the sFv, which allowed for grafting of χCDRs with an overall geometry approximating CDRs in the parent combining site. Analysis of 10 different Fv crystal structures indicates that the positions for inserting χCDRs are very highly conserved, as are the corresponding χCDR boundaries in the parent binding site. The results of this investigation suggest that it should be possible to generally apply this approach to the development of chimeric bispecific antibody binding site (χBABS) proteins.

INTRODUCTION

During the past decade considerable research has focused on the Fv region of antibodies, a 25-kDa heterodimer of V_H and V_L domains that form the antigen combining site. The unique binding properties of a particular antibody derive primarily from the structure of its complementarity-determining regions (CDRs), which comprise loops that connect structurally conserved β-strands forming the framework regions of each variable domain. The inherent stability of the framework region was first suggested by the process of CDR grafting or humanization (Jones et al., 1986), in which a hybrid Fv region may be genetically constructed by taking the CDRs from a murine monoclonal antibody and grafting them onto human framework regions (Jones et al., 1986; Verhoeyen et al., 1988). During this early period in antibody engineering, recombinant Fv fragments were produced as discrete entities by secretion from bacteria (Skerra and Plückthun, 1988; Field et al., 1988, 1990) or mammalian cells (Reichmann et al., 1988), and Fv regions were successfully engineered as the single-chain Fv (sFv), with both variable domains connected by a linker peptide introduced at the DNA level (Huston et al., 1988; Bird et al., 1988).

The present investigation has explored the possibility of constructing a second binding site as an integral part the Fv region (forming a χsFv), rather than making single-chain Fv fusions with distinct protein domains, such as a radiometal chelate (George et al., 1995), an Fc-binding protein (Tai et al., 1990), or another sFv (Huston et al., 1991; Neri et al., 1995). The minimal size of such bispecific Fv regions could provide significant advantages in vivo, and this design offers an opportunity for allosteric linkage between pairs of binding sites in a χsFv. The many examples of successful CDR grafting (Adair, 1992) suggest that greater advantage might be taken of the structural principles underlying the process of humanization. For example, framework regions (FRs) and CDRs may be viewed as separate contributors to the Fv folding reaction, and a primary driving force in Fv refolding could then be interpreted to involve the recovery of FR interactions between β-strands linked by bottom loops. As an alternative, the FRs could be defined to consist of β-strands connected by top loops, which include CDRs, in which case the bottom loops could be considered available for the design of a second combining site. This reasoning led the present investigation to analyze Fv architecture in search of a general strategy for changing the bottom surface into a second combining site.

The secondary structure of immunoglobulin variable domains has long been recognized to comprise a double layer of β-pleated sheets (Poljak et al., 1973), which may also be viewed within the Fv region as two concentric β-barrels with the combining site formed by CDR loops mainly on the inner barrel (Novotny et al., 1983). The extremely high conservation of three-dimensional backbone architecture in the Fv region was confirmed by the first Fab crystal structures (reviewed by Amzel and Poljak, 1979), which also suggested other invariant features of these structures (Novotny et al., 1983; Novotny and Haber, 1985). A large increase in the number of Fab crystal structures reported in the literature has allowed the generalization of conserved Fv structural features to an invariant coordinate system in which the V_H and V_L are symmetrical about a twofold axis (Gelfand et al., 1996). In the present investigation we have discovered symmetry between binding site CDRs and the bottom loops, which diametrically oppose the primary combining site. Chimeric sFv (χsFv) designs were explored using computer-generated models in which bottom loops were altered to form a second binding site (Fig. 1, *A* and *B*).

Received for publication 21 June 1996 and in final form 5 August 1996.

Address reprint requests to Dr. J. S. Huston, Creative BioMolecules, 45 South Street, Hopkinton, MA 01748.

*This paper is dedicated to Professor Serge N. Timasheff on the occasion of his 70th birthday.

0006-3495/96/10/2002/10 $2.00

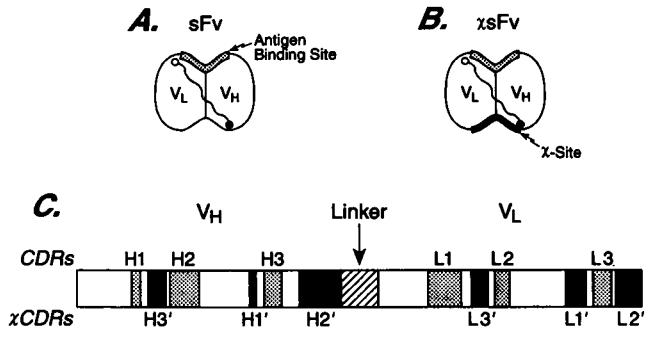

FIGURE 1 Schematic comparison of single-site sFv and dual-site χsFv, with position of χCDRs indicated on linear map of the χsFv. (*A*) The single-chain Fv is diagrammed as a V_H-linker-V_L molecule, with the linker drawn diagonally across the face of the Fv region, connecting the C-terminus of V_H to the N-terminus of V_L. (*B*) The chimeric single-chain Fv (χsFv) analog is drawn to emphasize the twofold symmetry relating the primary site (*dotted surface*) and the bottom χ-site (*solid black*). The sFv protein comprises about 250 amino acids with a molecular weight in the range of 26,000–27,000, whereas the χsFv is slightly larger, with a molecular weight that approaches 30,000. (*C*) The linear placement of χCDRs in the χsFv gene or protein is dictated by bottom loop geometry, which derives from the β-structure of each domain and the twofold symmetry relating the V_H and V_L domains within the Fv region (Searle et al., 1995). The order of the χCDRs (H3′, H1′, H2′, L3′, L1′, L2′) contrasts with that of the normal CDRs (H1, H2, H3, L1, L2, L3).

Within limits, it was found that much of the CDR structure from a typical binding site could be positioned on the bottom of the Fv in a nearly normal configuration, although χCDRs were grafted onto the FRs out of the normal order seen with CDRs (Fig. 1 *C*). Discussions of a second binding site in the Fv have previously been incidental to meeting presentations (Keck et al., 1994; Huston et al., 1994; Huston, 1994) and have appeared in a related patent application (Huston and Keck, 1993).

MATERIALS AND METHODS

The Fv coordinates used for molecular modeling were obtained from the Protein Data Bank (PDB) of the Brookhaven National Laboratory (Bernstein et al., 1977). These studies utilized the following Fab structures (with the corresponding PDB data set identified in parentheses): anti-phosphocholine McPC603 (2MCP); anti-digoxin 26-10 (1IGJ); anti-lysozyme D1.3 (1FDL); KOL (2FB4); anti-fluorescein 4-4-20 (4FAB); anti-lysozyme; HyHEL-5 (2HFL); anti-lysozyme HyHEL-10 (3HFM); anti-galactan J539 (2FBJ); anti-*p*-azobenzenearsonate R19.9 (1F19); anti-neuraminidase NC41 (1NCA).

Source χCDR sequences consist of the portion of each CDR that can be grafted into bottom loop positions defined by splice points. Target segments are the bottom loop sequences that would be displaced in the grafting process; the boundaries of the target segments are the target χCDR splice points of a χBABS protein. Flanking the first and last residues of each source or target sequence, the splice points are each defined as the midpoint between the α-carbons of a given end-point residue and its adjacent flanking residue (the residue preceding the χCDR or target segment's amino terminus or following its carboxyl terminus). The first three data sets were used for superposition models of the χsFv, and all 10 structures were used to calculate average χCDR source and target splice points.

Model building

Molecular modeling was conducted on either a Silicon Graphics Indigo II or 4D/70 GT graphics workstation. Aside from the SUPER algorithm for superposition of multiple Fv structures, which was written for this study (Keck, unpublished), all modeling utilized the INSIGHT and HOMOLOGY protein modeling packages (Molecular Simulations, San Diego, CA). The HOMOLOGY program was used in the construction of χsFv models and in the examination of specific molecules to correlate protein sequence with structure. To create a χsFv model, the carboxyl-terminal end of H2′ was fused with the amino-terminal end of the interdomain linker, consisting of a $(Ser_4\text{-}Gly)_3$ type of linker (Adams et al., 1993; McCartney et al., 1995; Tai et al., 1995; Huston et al., 1996b) or $(Gly_4\text{-}Ser)_3$ linker (Huston et al., 1988; Tai et al., 1990), as well as others in the literature (reviewed by Huston et al., 1991, 1996a).

Structure alignment

For superposition models, manual alignment of pairs of Fv structures was accomplished using the INSIGHT program to display the α-carbon backbones of the source and target structures. The CDR loops of the source structure were aligned with the bottom loops of the target structure. The H3 and L3 CDR loops (H TL4 and L TL4, respectively) of the source structure were respectively aligned with bottom loops H BL2 and L BL2 of the target structure (Fig. 2). Small translations and rotations then resulted in the optimal alignment of H1 (H TL1) with H BL4 and L1 (L TL1) with L BL4. Finally, the H IS2 strand from the source structure was optimally aligned with H IS4 from the target structure, and L IS2 was aligned with L IS4.

The MCP/MCP superposition model yielded the splice points for McPC603, and splice points for the other nine Fv fragments were based on homology with the MCP/MCP model. The source or target splice points of all 10 Fv structures were superimposed. The program SUPER, written in Microsoft C to run on personal computers under DOS, optimized the three-dimensional alignment of multiple sets of corresponding points. For each Fv structure, the MCP/MCP homology positions were used to identify two point sets, comprising the 10 splice points for source χCDRs (H3n, H3c, H1n, H1c, H2n, L3n, L3c, L1n, L1c, L2n) and 10 splice points for target segments (H3′n, H3′c, H1′n, H1′c, H2′n, L3′n, L3′c, L1′n, L1′c, L2′n). Point groups were then formed, each consisting of the corresponding point from all 10 structures (e.g., each Fv contributed the coordinates of its H3n splice point to form the H3n point group). A central point (average splice point) was defined for each point group by averaging the coordinates for members of that group. Averaging of the 10 groups of source splice points yielded one set of average source splice point values, and a corresponding set of average target splice points was similarly calculated.

The average source and target splice points were then aligned by SUPER to compare their relative geometries. Alignment of the 10 Fv structures was achieved by linking each central point to each point in a group by a first-order force or "spring." For each point set from a specific Fv structure, the net translational force and torque on that set resulting from the net force of the "springs" was calculated, and the point set was translated and rotated by an amount proportional, respectively, to the net force and torque. Within a point group, the mean square distance between individual points and the average central point was calculated, and, after each iteration of translation and rotation, its sum over all of the groups was calculated as a measure of the overall alignment. The alignment process was stopped when additional iterations changed these sums by less than 10^{-5} Å^2. The average radial distribution of end points around the center point of a given point group was calculated as a measure of the compactness of a point group's spatial distribution. Source and target vectors were defined as pointing from the amino- to the carboxyl-terminal splice points for H1′, H3′, L1′, and L3′, and vectors were presented in three dimensions by using stereo diagrams.

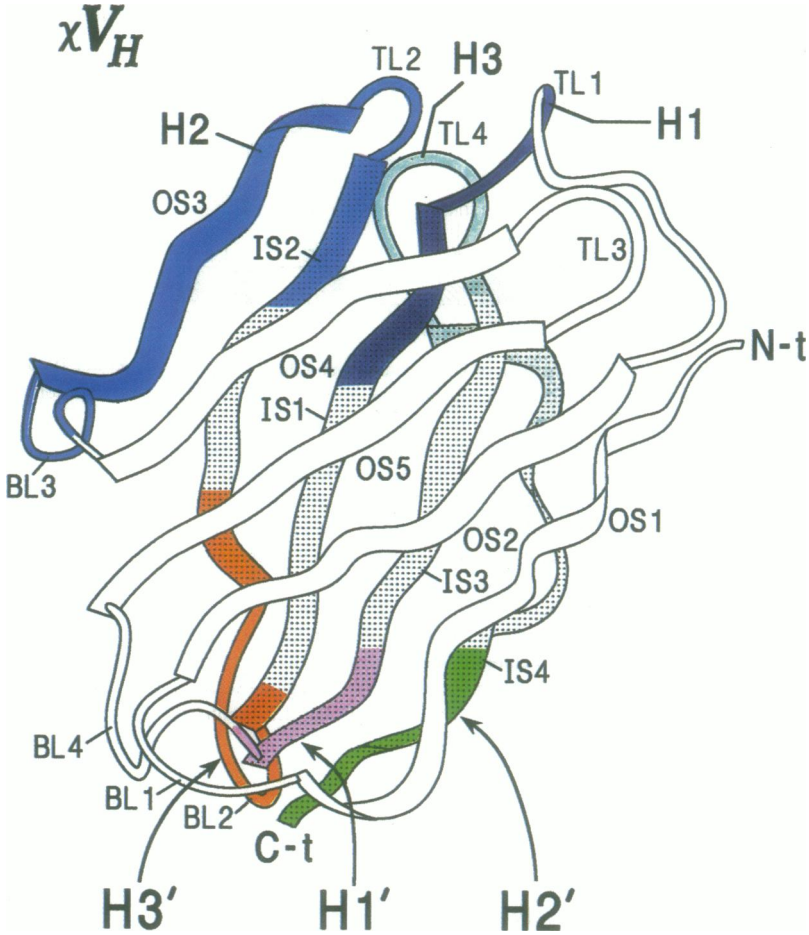

FIGURE 2 A backbone model of the V_H domain showing bottom loop positions suitable for χCDR grafting. The protein architecture that underlies the χsFv design is suggested in this schematic. This diagram depicts the D1.3 V_H α-carbon backbone based upon crystallographic coordinates for the Fv region of D1.3 Fab, from data set 1FDL in the Protein Data Bank (PDB) of Brookhaven National Laboratory (Bernstein et al., 1977). The β-strands are drawn as ribbons in this diagram, which is based on the crystal structure of the anti-lysozyme D1.3 Fab fragment (Fischmann et al., 1991). The primed loop notations indicate target sites where χCDRs would be grafted onto bottom loops of the V_H domain.

The variable region amino terminus (N-t) and carboxyl terminus (C-t) are noted on V_H. Loops in this figure are denoted as top loops (TL) and bottom loops (BL), which correspond to front and back loops, respectively, as defined by Amzel and Poljak (1979); β-strands correspond to the definition of structurally conserved regions (Greer, 1991), being common to all aligned structures. The CDRs (*curved lines* in three shades of *blue*) are part of top loops that connect β-strands (*ribbons*); bottom loops are represented by connections between α-carbons, noted as BL1-4.

The gray ribbons indicate that β-strands are normally buried at the interface between V_H and V_L domains in the sFv, and white ribbons indicate surface-exposed β-structure. OS, outer strand; IS, inner strand; TL, top loop; BL, bottom loop; N-t, amino terminus; C-t, carboxyl terminus; H1, H2, and H3, normal CDRs; H1′ (*lavender*), H2′ (*green*), or H3′ (*copper-brown*), target segments, defined as the residues on the bottom of the Fv that will be displaced by χCDRs.

RESULTS

Symmetry and design of a χsFv

The general relationship between an sFv and a χsFv is suggested by the schematic diagrams in Fig. 1, *A* and *B*. Examination of the bottom loop arrangement in McPC603 revealed that certain combinations of loops displayed a geometric relationship that was very similar to that of the original CDRs in the McPC603 combining site. In an analysis described at the end of this section, these observations were found to hold true for 10 different crystal structures of Fv regions available in the PDB. Bottom loops that might be used to graft modified CDRs onto the Fv were identified by molecular modeling. As a manifestation of the complex Greek key fold of the β-strands in the variable regions (Searle et al., 1995), these chimeric CDRs (χCDRs) mapped onto the linear χsFv sequence in the order H3′, H1′, H2′, L3′, L1′, and L2′ (Fig. 1 *C*).

Spatial positions of bottom V_H loops that would be modified in the process of grafting χCDRs may be visualized in Fig. 2. Using a nomenclature based on the structural homology of Ig variable domains (Huston et al., 1993a,b), the diagram indicates where the H3′ and H1′ regions would be positioned, respectively, in the BL2 and BL4 bottom loops and adjacent inner β-strands. The position of the C-terminal H2′ is indicated by the green coloration of the final stretch of IS4, but H2′ would actually continue beyond the end of V_H, with extra residues fused to the linker of a χsFv. If V_L were drawn beside Fig. 2, its framework and bottom loop structures would be essentially the same as those of the V_H schematic. The design of a corresponding χV_L would thus involve the bottom loops positioned as in χV_H, but they would be labeled L3′, L1′, and L2′. In a χsFv configured as χV_H-linker-χV_L, the final residues of L2′ would lack a fixed C-terminus unless they were attached artificially to the framework, for example by disulfide bonding through a pair of cysteinyl residues engineered at appropriate positions.

Structural considerations in χCDR grafting

Our initial χsFv modeling studies were conducted using the 2MCP coordinates for the Fv region of McPC603 Fab, the anti-phosphorylcholine Fab fragment first solved by Segal et al. (1974) and brought to higher resolution by Satow et al. (1986). To determine the positions where χCDR loops may be appropriately grafted onto the bottom framework (defined as splice points), we initially embarked on building a model of McPC603 χsFv with McPC603 χCDRs grafted onto the bottom of the McPC603 sFv (defined as MCP/MCP χsFv). Two copies

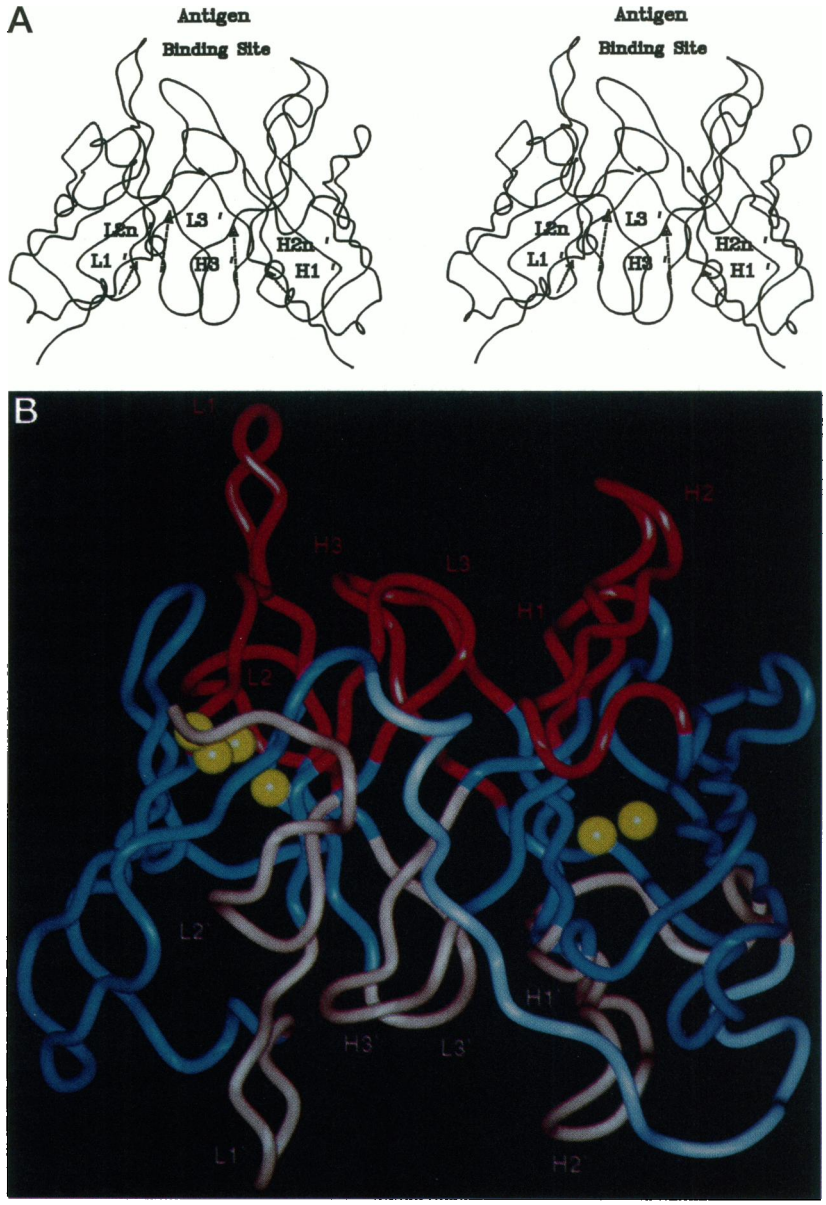

FIGURE 3 Molecular models showing splice point geometry in McPC603 Fv and MCP/MCP χsFv, emphasizing the twofold symmetry of binding sites in the χsFv. (*A*) Visual definition of the target splice sites identified in the McPC603 Fv. Stereo backbone diagrams of McPC603 Fv drawn to show the positions of its χCDR splice sites. The bottom loop geometry of each variable domain provides potential sites for grafting χCDRs H3′, L3′, H1′, and L1′ in a reasonable spatial approximation of the CDR loops that are most central to the original binding site. The McPC603 χCDR loops were grafted onto bottom framework positions, as indicated by vectors drawn approximately between the fixed endpoints (pointing from N- to C-terminal positions flanking χCDR residues) that are compatible with the normal chain directionality of χCDRs 1 and 3 in the context of the binding site. (*B*) The resulting χsFv constructed computationally by grafting the McPC603 χCDRs into the defined target splice points. This computer-generated molecular model of a χsFv represents a bivalent analog of the McPC603 sFv, in which the χCDRs from the natural phosphocholine-binding site on top have been transplanted to bottom loop positions.

In addition to the loops altered as noted in *A*, the remaining two χCDRs based on H2′ and L2′ were positioned with only their N-terminal ends fused genetically to rigid framework positions, at splice points noted in *A* of this figure; the C-terminus of H2′ was fused to the linker, and the C-terminus of L2′ was affixed to the V_L surface with a disulfide bond between a carboxyl-terminal cysteinyl residue and surface Cys mutation at L5.

The two binding sites are related by an approximate twofold axis passing through the molecule from left to right. The red loops are the original CDRs, the copper loops are added χCDRs, the blue strands are framework regions, the light blue line is the linker, and the yellow balls are sulfur atoms, two in each domain forming the buried disulfide bond.

of McPC603 Fv were aligned, one with its binding site pointed upward and the other with its bottom surface upward. The heavy- and light-chain CDR loops of one structure were thus superimposed, respectively, on the heavy- and light-chain bottom loops of the other structure to give the greatest overall coincidence of the top and bottom loops. Splice points were chosen for grafting chimeric CDRs (χCDRs) using the following three criteria: 1) retain as much of the CDR sequence as possible; 2) preserve framework residue positions that are judged to be critical to the structural integrity of the V domain; 3) achieve optimal superposition of the α-carbon backbone in regions of source and target splice points.

From this superposition model we identified the CDR residues that were to be inserted (source χCDRs) and the framework residues that would be displaced (target sequences). The positions of target splice points for McPC603 Fv are displayed in a stereo diagram (Fig. 3 *A*), which indicates where the McPC603 χCDRs were grafted to produce the MCP/MCP χsFv, displayed in Fig. 3 *B*. The splice points in Fig. 3 *A* have been connected, defining a vector for each χ-loop with fixed end points, to make their geometric relationship more readily apparent; the vectors point from the N- to C-terminal splice points for each target segment. It is apparent in the MCP/MCP χsFv structure that the normal shape of CDRs on the top is closely mimicked by χCDRs configured on the bottom. The two binding sites are related by an approximate twofold axis that runs horizontally through the molecule, as drawn. In this model, the L2′ loop has been given a fixed end point at its C-terminus by adding a C-terminal Cys residue that can disulfide bond with an opposing Cys added at residue L5 in the OS1 β-strand.

Generally, a given χCDR sequence represents a subset of its parent CDR sequence as an outgrowth of several design

criteria: 1) the relative shapes of the top and bottom loops, 2) the actual positions of splice overlap, and 3) the structural limits on the alteration of bottom loop sequences. The result of these considerations can be seen by comparing the sequences of MCP/MCP χsFv and McPC603 sFv, which are presented at the top of Fig. 4. Structurally aligned sequence positions for CDRs and χCDRs are also listed in Table 1, in columns labeled MCP/MCP. Binding site geometry is conserved in part by interchain hydrogen bonding between glutamine residues (Novotny and Haber, 1985) that are amino-terminal to H3′ and L3′ (Gln H39 and Gln L45 by structural homology (Figs. 4 and 5) equivalent to H39 and L38 according to Kabat et al., 1991). Likewise, there are important electrostatic contributions from the aspartate residues amino-terminal to H1′ and L1′ (Asp H92 and Asp L89, or H86 and L82 according to Kabat), which form intrachain salt bridges with Arg or Lys at positions H69 and L68, respectively (H66 and L60 according to Kabat).

Modeling of χBABS proteins by homology with structurally aligned sequences

It should be possible to generalize the χsFv design from the MCP/MCP model to other Fv regions and χ-site combinations, by virtue of the strong structural conservation and sequence homology among V regions. Sequence alignment by structural superposition (discussed previously by Huston et al., 1993a,b) provides a mechanism for using homology in the design of χBABS proteins. In our analysis, sequences were aligned with the HOMOLOGY program by superimposing the three-dimensional α-carbon backbone of the given Fv structure on that for McPC603 Fv. Structurally homologous regions identified for the sequences in Figs. 4 and 5 correspond to the structural designations noted on the V_H schematic of Fig. 2. Using the MCP/MCP χsFv design, the homology-based source and target χCDR sequences of nine other representative Fv fragments were identified, and their structurally aligned V domain sequences are given in Fig. 5. Sequence positions are noted for their χCDR sequences on the Source line and for their target segments (bottom loop sequences to be displaced) on the Target line.

As a test of the homology modeling approach, we compared a 26-10/D1.3 χsFv superposition model with its homology model (its sequence in Fig. 4 is based upon the MGP/MCP model). The anti-digoxin 26-10 Fv sequence was structurally aligned using the PDB coordinates for the 1IGJ structure of 26-10 Fab (Jeffrey et al., 1993), and the D1.3 sequence was aligned using the 1FDL structure of D1.3 (Satow et al., 1986). Sequence homology with the MCP/MCP superposition model was used to design the 26-10/D1.3 χsFv, in which all six D1.3 χCDRs were grafted onto the 26-10 sequence (Fig. 4). The 26-10 and D1.3 structures were also used to determine the χ-loop splice points by superposition of computer-generated models of the two Fv regions (identified in Table 1 in the columns labeled 26-10/D1.3). The resulting source and target boundaries (Table 1, listed in terms of residue positions in the aligned sequences) were found to be nearly the same as those based on sequence homology (Table 1, columns labeled MCP/MCP). Only small differences in splice points were found between the χsFv models derived by homology and superposition, which supports the use of homology-based models as reasonable first approximations for χBABS design.

TABLE 1 Comparison of MCP/MCP and 26-10/D1.3 χsFv superposition splice points

	Target splice points		Source splice points	
χ-Loop	MCP/MCP	26-10/D1.3	MCP/MCP	26-10/D1.3
H3′	H40-H46	H39-H47	H102-H116	H101-H117
H1′	H93-H95	H93-H95	H31-H33	H30-H33
H2′	H122	H122	H51-H72	H51-72
L3′	L46-L52	L46-L52	L101-L106	L101-L106
L1′	L91-L92	L91-L92	L30-L39	L31-L39
L2′	L109	L109	L56-L68	L55-L65

Generality of splice point geometry

The rationale for using homology-based splice points for the boundaries of χCDRs and their target segments was tested further through computational analysis of representative Fv crystal structures. A necessary condition for the use of generalized splice points is that their geometry should be highly conserved across many Fv structures, for both the source χCDR and target splice points. Average source and target splice points should likewise be in close proximity when they are superimposed. We have compared the actual coordinates for the standard MCP/MCP splice positions among 10 Fv regions of known structure (see Materials and Methods). Averaging each group of homologous splice points gave 10 average splice points for the source and 10 for the target. For each χCDR with two fixed end points (H1′, H3′, L1′, L3′), a vector has been drawn from the amino-terminal to the carboxyl-terminal splice point, and pairs of related source and target χCDR vectors were compared (e.g., H1 from source and H1′ from target positions). These data were presented in a stereo vector diagram (Fig. 6), showing that the source and target vectors are very close, and hence that the geometries of source and target splice sites are indeed very similar.

A numerical comparison of splice site positions is given in Table 2, which indicates that the distances between average source and target positions are in most cases no greater than the error typical of atomic coordinates reported for protein crystal structures. The Source and Target columns labeled Group Radius list, for each point group, the average radial distance between the 10 actual points and single average point for that group. The low values of these radii further emphasize the compactness of these splice point distributions. It is apparent that target end points from the bottom loops of the framework architecture show a smaller variation from the average than do source splice points. The striking compactness of each point group's

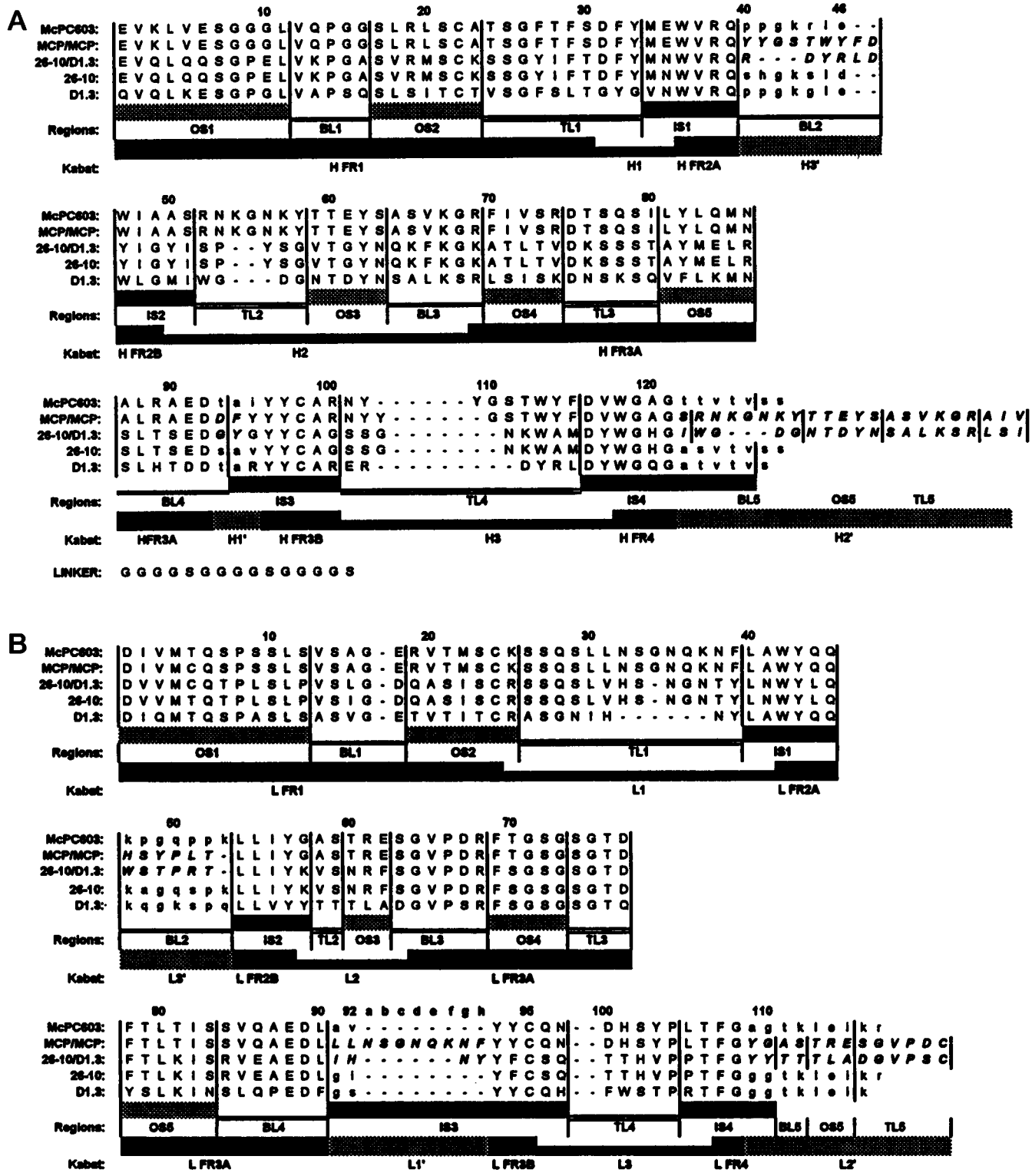

FIGURE 4 Alignment of χsFv and parent sFv sequences for bivalent MCP/MCP and bispecific 26-10/D1.3 χsFv proteins. The 26-10/D1.3 χsFv was designed by homology with MCP/MCP χsFv, which was determined by superposition of two McPC603 model structures; for reference purposes the original variable region sequences are given for anti-phosphorylcholine McPC603, anti-digoxin 26-10, and anti-lysozyme D1.3 Fv. Conserved elements of V regions are designated by nomenclature that is related to structural elements in Fig. 2. The β-sheet contacts between the V_H and V_L domains are formed by inner β-strands (IS), which were indicated by shaded ribbons in Fig. 2, and the solvent-exposed outer β-strands (OS) are indicated as white ribbons. Framework regions (FR1-4) and CDRs (H1-3, L1-3) were defined according to the method of Kabat et al. (1991) on the line denoted Kabat; structurally defined regions were designated on the line denoted Regions. The χCDR residues (H1′–H3′, L1′–L3′) are printed in italicized and boldfaced type.

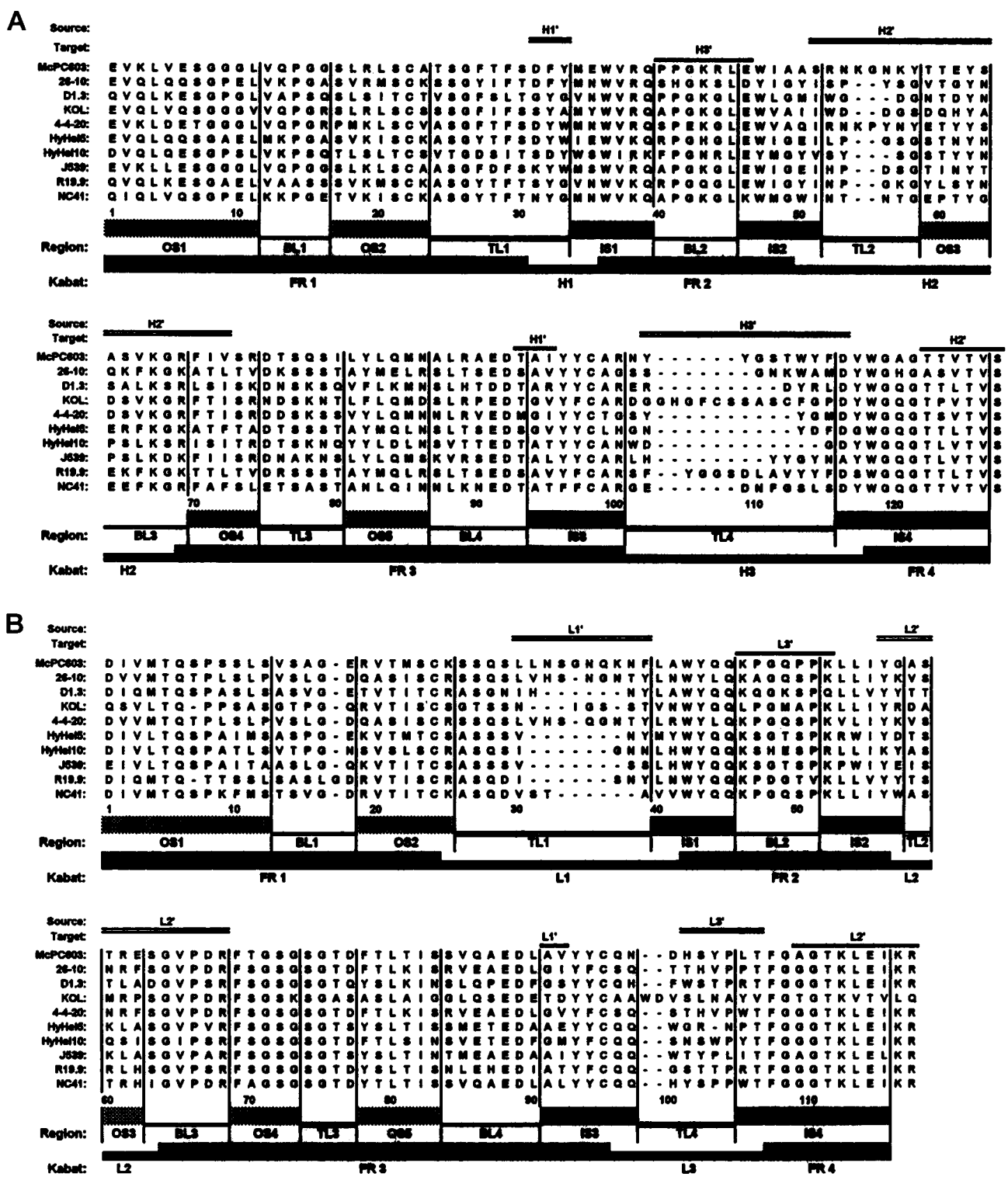

FIGURE 5 Structural alignment of V_H and V_L sequences of Fv structures from the PDB, showing the χCDRs defined by homology with the MCP/MCP model. The labels for structural elements shown in Fig. 2 and defined in the legend to Fig. 4 are used here to relate the structurally aligned sequences of these 10 Fv regions with known crystal structures (PDB data sets utilized are listed in Materials and Methods). The HOMOLOGY program was used to align the α-carbon backbone of each variable domain structure with the McPC603 V_H and V_L reference structures.

spatial distribution was indicated by calculations of the distances between corresponding average source and target points, listed in the right column of Table 2. For some point groups, the sums of source and target group radii were greater than the distance between the average source and target points for those groups, which further underscores the similar-

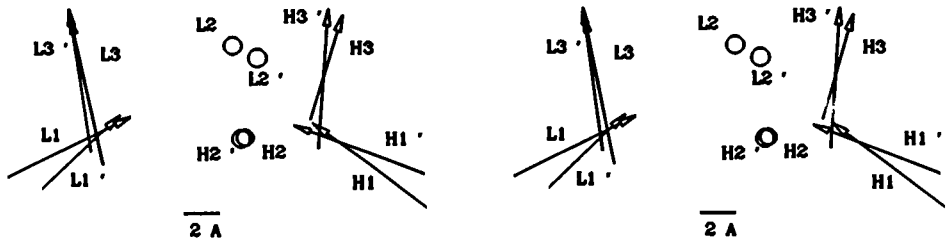

FIGURE 6 Vector diagram in 3-space that compares the average geometry of splice sites for source and target segments, based on 10 Fv crystal structures. The source and target χCDR vectors connect the average splice points (from N- to C-terminal sequence positions), as defined in Materials and Methods. The SUPER algorithm was used to align the source and target splice points for all 10 Fv structures. The averages of the resulting point sets defined central points that were plotted to make this stereo diagram, using the SUPER algorithm to superimpose the averaged source and target sets of splice points. Labels indicate whether source (unprimed) or target (primed) end points were used to define a given vector. The vector diagram is displayed as a stereo pair for a χsFv in the χV_H-linker-χV_L configuration, positioned approximately as in Fig. 3 with the χ-site facing downward. The average source vectors were superimposed by effectively rotating the primary binding site (on top in Fig. 3) to face downward in alignment with the χ-site. The approximate ends of the vectors are shown in the MCP/MCP χsFv ribbon diagram (Fig. 3) as the junctions between blue framework and copper-brown χCDRs, connected so that vectors always point from N- to C-terminal points. These general vectors contrast with those drawn in Fig. 3 *A*, which are the target vectors of the MCP/MCP model, displayed in the context of the McPC603 Fv α-carbon backbone structure.

ity of average target and source splice point geometry. All of these χ-site comparisons involved the mapping of V_H source χCDRs to V_H target segments and, similarly, the mapping of V_L source to V_L target sequences. However, the twofold symmetry of the V domains permits an alternative mapping scheme, of χV_H source to χV_L target and χV_L source to χV_H target. Vector pairs for this alternative χBABS model were calculated using the same set of Fv structures (data not shown), and the average of the distances between corresponding target and source points was 2.25 Å, which was only slightly greater than the 1.33 Å calculated for the homologous mapping case (based on data in Table 2).

DISCUSSION

This investigation has uncovered previously unrecognized symmetry properties that apply to binding-site CDRs and bottom loops of the Fv region. Analysis of the corresponding homology-based splice points in 10 Fv structures has revealed that their source and target positions are conserved to a remarkable degree. These architectural considerations suggested a rationale for the construction of a second binding site that relies on χCDR grafting to preform a binding site surface on the bottom of the Fv region, which would thereby predispose the χ-site to recognition of a given antigenic epitope. Phage antibody display libraries represent an analogous situation, in that they depend upon the preformed binding site surfaces that result from combinatorial pairing of V_H and V_L genes (Winter et al., 1994).

The utility of the molecular designs developed in this investigation will ultimately depend on the ability of χBABS proteins to refold. This appears to be practical, based on promising results in our preliminary experiments (Keck et al., 1994; Tai et al., unpublished results). Another facet of χBABS stability involves the interaction between χV_H and χV_L domains. The CDRs of a normal Fv region account for about 25% of the contact surface between the V_H and V_L domains (Huston et al., 1991), and thus changes in CDRs could significantly modulate the associative interactions between a given pair of variable domains (Huston et al., 1996a). Although contributing a smaller fraction of interdomain contacts, the χ-loops of a χBABS protein could have a similar impact on heterodimer stability. For example, insertion of χCDRs could enhance the association of heterodimers, making a stable χFv practical. Alternatively, there might be a negligible impact, or the χCDRs could lower the association constant of the original V domains, which might be offset by the construction of a χsFv.

TABLE 2 Differences between average source and target positions with group radii of splice points

Source		Target		Distance (Å) between average source and target point group positions
Point group	Group radius (Å)	Group radius (Å)	Point group	
H3n	1.04	0.48	H3′n	1.85
H3c	0.63	0.49	H3′c	1.11
H1n	0.98	0.48	H1′n	2.76
H1c	0.71	0.30	H1′c	1.08
H2n	0.66	0.39	H2′n	0.23
L3n	1.45	0.35	L3′n	1.10
L3c	0.53	0.37	L3′c	0.57
L1n	1.01	0.42	L1′n	2.15
L1c	0.87	0.22	L1′c	0.73
L2n	0.72	0.40	L2′n	1.67

The development of χBABS proteins also relates to more general areas of de novo protein design, particulary those that have sought to explain how the composition of turns (Brunet et al., 1993; Predki et al., 1996) and periodicity of sequence (Kamtekar et al., 1993; Kamtekar and Hecht, 1995; Xiong et al., 1995) affect the recovery and stability of specific secondary structures such as β-sheet (Hecht, 1994). Betadoublet is a recombinant protein that folds spontaneously to dimerize as a β-sandwich (Quinn et al., 1994), and betabellin is a synthetic, cysteine-containing protein analog that folds into a β-sandwich only after the formation of interchain disulfide bonds (Yan and Erickson, 1994). Proteins based on periodic sequence have also been described that rely on designing the antisense or complementary proteins that can be defined by a reversed reading of a given mRNA (Clarke and Blalock, 1990; Blalock, 1994).

Framework engineering of the Fv region has in the past focused primarily on humanization and the idealization of framework properties, or improvement of antigen binding. Thus the improvement of folding and production yields through framework mutations has thus far involved residues located in FR1 (McCartney et al., 1995) or in FR2 (Knappik and Plückthun, 1995). The V_H framework of the B1-8 antibody tolerated the insertion of four residues in its non-CDR TL3 top loop, adjacent to the antibody combining site (Simon and Rajewsky, 1992). Normal hapten binding was recovered in this analog, and the authors proposed that this top loop might be used as a fourth CDR to increase contact with antigen beyond that normally possible. The Fv region of antibodies has also been used as a recombinant carrier of loop peptides. These so-called antigenized antibodies have been made by substituting an antigenic loop sequence for a CDR (reviewed by Billetta and Zanetti, 1993); related work has been reported on sequence substitutions in the top loops of the C_H1 constant region (Sandlie et al., 1996).

The results of the present investigation suggest how χBABS proteins might be optimized by phage display techniques. For instance, the refinement process of a χBABS protein could involve site-directed or cassette mutagenesis of given loops in a χsFv, followed by phage display and antigen-driven selection. More extensive application of these techniques to χBABS construction could take full advantage of the homology-based design process delineated in this study. The insertion of randomly synthesized χCDRs that span the predicted pairs of splice points for target segments in an sFv gene would lead to a χsFv mutant library wherein all species would have the same primary combining site but a large population of different χ-site configurations. Appropriate selection by conventional phage display library methods could yield the desired combination of binding-site activities. Selection with and without antigen bound to the primary site could allow selection of variants that exhibit a range of allosteric properties. Future investigations are expected to pursue the construction of χBABS proteins with specific Fv regions by using the highly sophisticated methods of antibody engineering (Borrebaeck, 1995; Huston et al., 1996a). Successful implementation of designs suggested by this paper would have important clinical ramifications, for instance, in the development of minimal bispecific antibody species for cancer therapy. Furthermore, χBABS designs that exhibit significant allosteric linkage between sites would offer entirely novel opportunities for the development of unique therapeutic agents, homogeneous immunoassays, and allosteric catalytic antibodies.

We (P. C. Keck, M.-S. Tai, C. Cohen, and J. S. Huston) thank Professor Timasheff for over a decade of his advice and enthusiasm, which have added immeasurably to our research in protein and antibody engineering. As protein physical chemists, we have been inspired throughout our careers by the beauty and inexhaustible creativity of his thermodynamic studies on protein interactions and cosolvent effects, as well as the rigor, clarity, and eloquence of his research papers.

This research project was stimulated by an SBIR grant (R43 CA52323) from the National Cancer Institute. In addition, we are grateful to Charles Cohen for his encouragement and support.

REFERENCES

Adair, J. R. 1992. Engineering antibodies for therapy. *Immunol. Rev.* 130:5–40.

Adams, G. P., J. E. McCartney, M.-S. Tai, H. Oppermann, J. S. Huston, W. F. Stafford III, M. A. Bookman, I. Fand, L. L. Houston, and L. M. Weiner. 1993. Highly specific in vivo tumor targeting by monovalent and divalent forms of 741F8 anti-c-*erb*B-2 single-chain Fv. *Cancer Res.* 53:4026–4034.

Amzel, L. M., and R. J. Poljak. 1979. Three-dimensional structure of immunoglobulins. *Annu. Rev. Biochem.* 48:961–997.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. J. Tasumi. 1977. The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Billetta, R., and M. Zanetti. 1993. Antigenicity and immunogenicity of antigenized antibodies. Studies on B and T cells. "Antibody Engineering. Medical Applications," ed. by M. Zanetti. *Int. Rev. Immunol.* 10:251–263.

Bird, R. E., K. D. Hardman, J. W. Jacobson, S. Johnson, B. M. Kaufman, S.-M. Lee, T. Lee, S. H. Pope, G. S. Riordan, and M. Whitlow. 1988. Single-chain antigen-binding proteins. *Science.* 242:423–426.

Blalock, J. E. 1994. Protein configurations. *Science.* 265:1511.

Borrebaeck, C., editor. 1995. Antibody Engineering, 2nd ed. Oxford University Press, New York and Oxford.

Brunet, A. P., E. S. Huang, M. E. Huffine, J. E. Loeb, R. J. Weltman, and M. H. Hecht. 1993. The role of turns in the structure of an α-helical protein. *Nature.* 364:355–358.

Clarke, B. L., and J. E. Blalock. 1990. Steroidogenic activity of a peptide specified by the reversed sequence of corticotropin mRNA. *Proc. Natl. Acad. Sci. USA.* 87:9708–9711.

Field, H., G. T. Yarranton, and A. R. Rees. 1988. *Vaccines (Cold Spring Harbor).* 88:29–34.

Field, H., G. T. Yarranton, and A. R. Rees. 1990. Expression of mouse light and heavy chain variable regions in *Escherichia coli* and reconstitution of antigen-binding activity. *Protein Eng.* 3:641–647.

Fischmann, T. O., G. A. Bentley, T. N. Bhat, G. Boulot, R. A. Mariuzza, S. E. V. Phillips, D. Tello, and R. J. Poljak. 1991. Crystallographic refinement of the three-dimensional structure of the FabD1.3-lysozyme complex at 2.5 Å resolution. *J. Biol. Chem.* 266:12915–12920.

Gelfand, I. M., A. E. Kister, and D. Leshchiner. 1996. The invariant system of coordinates of antibody molecules: Prediction of the "standard" C^α framework of V_L and V_H domains. *Proc. Natl. Acad. Sci. USA.* 93: 3675–3678.

George, A. J. T., F. Jamar, M.-S. Tai, B. T. Heelan, G. P. Adams, J. E. McCartney, L. L. Houston, L. M. Weiner, H. Oppermann, A. M. Peters,

and J. S. Huston. 1995. Radiometal labeling of recombinant proteins by a genetically engineered, minimal chelation site: technetium-99m coordination by single-chain Fv antibody fusion proteins through a C-terminal cysteinyl peptide. *Proc. Natl. Acad. Sci. USA.* 92:8358–8362.

George, A. J. T., J. A. Titus, C. R. Jost, I. Kurucz, P. Perez, S. M. Andrew, P. J. Nicholls, J. S. Huston, and D. M. Segal. 1994. Redirection of T cell-mediated cytotoxicity by a recombinant single-chain Fv molecule. *J. Immunology.* 152:1802–1811.

Greer, J. 1991. Comparative modeling of homologous proteins. *Methods Enzymol.* 202:239–252.

Hecht, M. H. 1994. De novo design of β-sheet proteins. *Proc. Natl. Acad. Sci. USA.* 91:8729–8730.

Huston, J. S. 1994. Engineering single-chain Fv species with dual combining sites: covalent $(sFv')_2$ dimers and chimeric bispecific antibody binding sites (χBABS). Fifth IBC International Conference on Antibody Engineering: New Technologies, Applications and Commercialization, La Jolla, CA, December 7–9.

Huston, J. S., and P. Keck. 1993. Chimeric multivalent protein analogues and methods of use thereof. Patent Cooperation Treaty WO 93/23537.

Huston, J. S., P. Keck, M.-S. Tai, D. Jin, J. McCartney, W. F. Stafford III, M. Mudgett-Hunter, H. Oppermann, and E. Haber. 1993a. Single-chain immunotechnology of Fv analogues and fusion proteins. *In* Immunotechnology, J. Gosling and D. Reen, editors. Portland Press, London. 47–60.

Huston, J. S., D. Levinson, M. Mudgett-Hunter, M.-S. Tai, J. Novotny, M. N. Margolies, R. J. Ridge, R. Bruccoleri, E. Haber, R. Crea, and H. Oppermann. 1988. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *E. coli. Proc. Natl. Acad. Sci. USA.* 85:5879–5883.

Huston, J. S., M. N. Margolies, and E. Haber. 1996a. Antibody binding sites. *Adv. Protein Chem.* 49:327–447.

Huston, J. S., M. Mudgett-Hunter, M.-S. Tai, J. E. McCartney, F. D. Warren, E. Haber, and H. Oppermann. 1991. Protein engineering of single-chain Fv analogues and fusion proteins. *Methods Enzymol.* 203: 46–88.

Huston, J. S., H. Oppermann, and S. N. Timasheff. 1996b. Serine-rich peptide linkers. U.S. Patent 5,525,491 (June 11, 1996).

Huston, J. S., M.-S. Tai, J. E. McCartney, P. Keck, and H. Oppermann. 1993b. Antigen recognition and targeted delivery by the single-chain Fv. *Cell Biophys.* 22:189–224.

Huston, J. S., M.-S. Tai, W. F. Stafford III, H. Oppermann, and P. Keck. 1994. Chimeric bispecific antibody binding sites (χBABS): 30 kD single-chain Fv proteins with dual binding sites. 11th International Hammersmith Meeting, Advances in the Applications of Monoclonal Antibodies in Clinical Oncology, Molyvos, Lesvos, Greece, May 23–26. 9.

Jeffrey, P. D., R. K. Strong, L. C. Sieker, C. Y. C. Chang, R. L. Campbell, G. A. Petsko, E. Haber, M. N. Margolies, and S. Sheriff. 1993. 26–10 Fab-digoxin complex: affinity and specificity due to surface complementarity. *Proc. Natl. Acad. Sci. USA.* 90:10310–10314.

Jones, P. T., P. H. Dear, J. Foote, M. S. Neuberger, and G. Winter. 1986. Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature.* 321:522–525.

Kabat, E. A., T. T. Wu, H. M. Perry, K. S. Gottesman, and C. Foeller. 1991. Sequences of Proteins of Immunological Interest, 5th ed. Publication No. 91-3242. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, Bethesda, MD.

Kamtekar, S., and M. H. Hecht. 1995. The four-helix bundle: what determines a fold? *FASEB J.* 9:1013–1022.

Kamtekar, S., J. M. Schiffer, H. Xiong, J. M. Babik, and M. H. Hecht. 1993. Protein design by binary patterning of polar and nonpolar amino acids. *Science.* 262:1680–1685.

Keck, P. C., M.-S. Tai, H. Oppermann, and J. S. Huston. 1994. Chimeric bispecific antibody binding sites (χBABS): design of a second combining site within the Fv region and recovery of single-chain Fv binding activity after integration of new binding loops between β-strands of V domains. *Biophys. J.* 66:A431.

Knappik, A., and A. Plückthun. 1995. Engineered turns of a recombinant antibody improve its in vivo folding. *Protein Eng.* 8:81–89.

McCartney, J. E., M.-S. Tai, R. M. Hudziak, G. P. Adams, L. M. Weiner, D. Jin, W. F. Stafford III, S. Liu, M. A. Bookman, A. A. Laminet, I. Fand, L. L. Houston, H. Oppermann, and J. S. Huston. 1995. Engineering disulfide-linked single-chain Fv dimers [(sFv')2] with improved solution and targeting properties. Anti-digoxin 26–10 (sFv')2 and anti-*c-erb*B-2 741F8 (sFv')2 made by protein folding and bonded through carboxyl-terminal cysteinyl peptides. *Protein Eng.* 8:301–314.

Neri, D., M. Momo, T. Prospero, and G. Winter. 1995. Hign-affinity antigen binding by chelating recombinant antibodies (CRAbs). *J. Mol. Biol.* 246:367–373.

Novotny, J., R. Bruccoleri, J. Newell, D. Murphy, E. Haber, and M. Karplus. 1983. Molecular anatomy of the antibody binding site. *J. Biol. Chem.* 258:14433–14437.

Novotny, J., R. E. Bruccoleri, and F. A. Saul. 1989. On the attribution of binding energy in antigen-antibody complexes McPC603, D1.3, and HyHEL-5 *Biochemistry.* 28:4735–4749.

Novotny, J., and E. Haber. 1985. Structural invariants of antigen binding: comparison of immunoglobulin V_L-V_H and V_L-V_L domain dimers. *Proc. Natl. Acad. Sci. USA.* 82:4592–4596.

Poljak, R. J., L. M. Amzel, H. P. Avey, B. L. Chen, R. P. Phizackerley, and F. Saul. 1973. Structure of Fab' NEW at 2.8 Å resolution. *Proc. Natl. Acad. Sci. USA.* 70:3305–3310.

Predki, P. F., V. Agrawal, A. T. Brünger, and L. Regan. 1996. Amino-acid substitutions in a surface turn modulate protein stability. *Nature Struct. Biol.* 3:54–58.

Quinn, T. P., N. B. Tweedy, R. W. Williams, J. S. Richardson, and D. C. Richardson. 1994. Betadoublet: de novo design, synthesis, and characterization of a β-sandwich protein. *Proc. Natl. Acad. Sci. USA.* 91: 8747–8751.

Reichmann, L., J. Foote, and G. Winter. 1988. Expression of an antibody Fv fragment in myeloma cells. *J. Mol. Biol.* 203:825–828.

Sandlie, I., B. Bogen, and E. Lunde. 1996. Antigen presentation of a T cell epitope inserted at different positions in an antibody heavy chain. Keystone Symposium, Taos, NM, February 22–28. 47.

Satow, Y., G. Cohen, E. A. Padlan, and D. R. Davies. 1986. Phosphocholine binding immunoglobulin Fab McPC 603, an x-ray diffraction study at 2.7 Å. *J. Mol. Biol.* 190:593–604.

Searle, S. J., J. T. Pedersen, A. H. Henry, D. M. Webster, and A. R. Rees. 1995. Antibody structure and function. *In* Antibody Engineering, 2nd ed. C. Borrebaeck, editor. Oxford University Press, New York and Oxford. 3–51.

Segal, D. M., E. A. Padlan, G. H. Cohen, S. Rudikoff, M. Potter, and D. R. Davies. 1974. The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. Natl. Acad. Sci. USA.* 71:4298–4302.

Simon, T., and K. Rajewsky. 1992. A functional antibody mutant with an insertion in the framework region 3 loop of the V_H domain: implications for antibody engineering. *Protein Eng.* 5:229–234.

Skerra, A., and A. Plückthun. 1988. Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli. Science.* 240:1038–1041.

Tai, M-S, J. E. McCartney, G. P. Adams, D. Jin, R. M. Hudziak, H. Oppermann, A. A. Laminet, M. A. Bookman, S. Liu, W. F. Stafford III, I. Fand, L. L. Houston, L. M. Weiner, and J. S. Huston. 1995. Targeting c-erbB-2 expressing tumors using single-chain Fv monomers and dimers. *Cancer Res.* 55(Suppl.):5983s–5989s.

Tai, M.-S., M. Mudgett-Hunter, D. Levinson, G.-M. Wu, E. Haber, H. Oppermann, and J. S. Huston. 1990. A bifunctional fusion protein containing Fc-binding fragment B of Staphylococcal protein A amino terminal to antidigoxin single-chain Fv. *Biochemistry.* 29:8024–8030.

Verhoeyen, M., C. Milstein, and G. Winter. 1988. Reshaping human antibodies: grafting an antilysozyme activity. *Science.* 239:1534–1536.

Winter, G., A. D. Griffiths, R. E. Hawkins, and H. R. Hoogenboom. 1994. Making antibodies by phage display technology. *Annu. Rev. Immunol.* 12:433–455.

Xiong, H., B. L. Buckwalter, H.-M. Shieh, and M. H. Hecht. 1995. Periodicity of polar and nonpolar amino acids is the major determinant of secondary structure in self-assembling oligomeric peptides. *Proc. Natl. Acad. Sci. USA.* 92:6349–6353.

Yan, Y., and B. W. Erickson. 1994. Engineering of betabellin 14D: disulfide-induced folding of a β-sheet protein. *Protein Sci.* 3:1069–1073.