Protein Science (1995), *4*:2073–2081. Cambridge University Press. Printed in the USA.

Destabilizing loop swaps in the CDRs of an immunoglobulin V_L domain

LARRY R. HELMS AND RONALD WETZEL

Macromolecular Sciences Department, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406

(RECEIVED May 15, 1995; ACCEPTED August 3, 1995)

Abstract

It is generally believed that loop regions in globular proteins, and particularly hypervariable loops in immunoglobulins, can accommodate a wide variety of sequence changes without jeopardizing protein structure or stability. We show here, however, that novel sequences introduced within complementarity determining regions (CDRs) 1 and 3 of the immunoglobulin variable domain REI V_L can significantly diminish the stability of the native state of this protein. Besides their implications for the general role of loops in the stability of globular proteins, these results suggest previously unrecognized stability constraints on the variability of CDRs that may impact efforts to engineer new and improved activities into antibodies.

Keywords: complementarity determining regions; hypervariable loops; immunoglobulin stability; protein stability; Stern–Volmer plots; unfolding

In the structures of antibody heavy and light chain variable domains, complementarity determining region (CDR) loops of diverse sequence and conformation project from β-sandwich frameworks to determine the affinity and specificity of antigen binding. These diverse sequences arise in the development of an antibody response via a complex combination of gene rearrangements and somatic mutations (Branden & Tooze, 1991). Protein engineers have transferred antigen binding activity from one framework to another by swapping hypervariable CDR loops between frameworks (Jones et al., 1986; Riechmann et al., 1988). In addition, totally artificial loop sequences have been introduced into variable domain frameworks to generate novel binding functions (Barbas et al., 1993; Fisch et al., 1994). This apparent tolerance of the immunoglobulin fold for both sequence diversity and loop swaps is consistent with the view that the major determinants of structural stability of globular proteins lie in the sequence-specific formation and packing of regular secondary structure elements (Shortle, 1992; Matthews, 1993), whereas surface loops are more forgiving of sequence changes (El Hawrani et al., 1994).

Although the CDR loops differ greatly in sequence, they are not infinitely variable in either sequence or conformation. Positions with highly conserved amino acids are found within many CDR sequences (Kabat et al., 1991). Further, X-ray crystallographic studies have demonstrated that only a limited number of "canonical structures" appear to be available for each CDR. For example, in an examination of the structural database available in 1989, only four different loop conformations were identified in the light chain crystallographic database for CDR1, and only three for CDR3 (Chothia et al., 1989). If these patterns do suggest constraints on the sequence variability of CDRs, however, it remains difficult from examination of evolved antibodies to either gauge the severity of these constraints or to understand their structural or functional basis.

In this report we describe the preparation of a number of mutants of the immunoglobulin light chain variable domain REI, in which wild-type CDR sequences were replaced by loops of differing sequence and length. These replacements cause dramatic reductions—and in some cases apparent elimination—of the net folding stability of the protein. The results have implications for the role of loops in determining the stabilities of immunoglobulins and other globular proteins and also suggest a folding stability constraint on the evolution of CDR structures.

Results

Figure 1 shows a ribbon diagram of the REI V_L domain highlighting its three CDRs. Figure 2 lists the mutants examined in this work and shows the CDR sequence changes carried out to produce the proteins. The RGD-containing sequences inserted into the REI CDRs are either derived from loop regions of the snake venom disintegrins kistrin (Dennis et al., 1989), barbourin (Scarborough et al., 1991), and echistatin (Gan et al., 1988) or (in REI-RGD18 and REI-RGD26) were designed, unnatural

Reprint requests to: Ronald Wetzel, Macromolecular Sciences Department, SmithKline Beecham Pharmaceuticals, 709 Swedeland Road, King of Prussia, Pennsylvania 19406; e-mail: wetzelrb%phvax.dnet@sb.com.

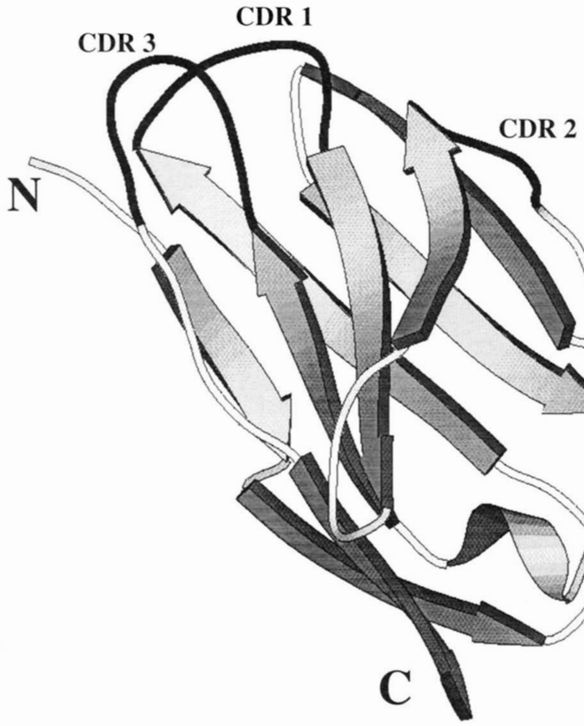

Fig. 1. Ribbon diagram of the REI V_L domain from the X-ray structural coordinates (structure 1REI from the Brookhaven protein structure database [Abola et al., 1987]) with the CDRs highlighted. Drawing produced using MOLSCRIPT (Kraulis, 1991).

sequences. Loop replacements were constructed in both CDR1 and CDR3, yielding domains that by several criteria appeared to be stably folded proteins. Thus, each mutant was stably expressed in *Escherichia coli* by secretion into the periplasm. In addition, the intradomain disulfide bonds of these mutants—like that of the wild type—form readily in the *E. coli* periplasm. Despite these indications of stability, however, and despite the solubility of these proteins in native buffer, further examination revealed significant destabilization in most of the mutants.

The folding integrity of an immunoglobulin domain can be assessed (Tsunenaga et al., 1987) by the intrinsic fluorescence of a conserved tryptophan, which is dramatically quenched in the folded state by the close proximity of the intradomain disulfide bond (Epp et al., 1974). Thus, the fluorescence of the wild-type domain in native buffer at RT is quite low compared to the fluorescence in 4 M guanidine hydrochloride (Gdn-HCl), where the protein is unfolded. Table 1 shows the fluorescence yield for WT REI V_L in native buffer is only 11% of the fluorescence in denaturant, and that some of the mutants (RGD12, RGD1, RGD34, RGD23, RGD18, RGD35, RGD32) also exhibit low fluorescence and are therefore also essentially completely folded in native buffer. In contrast, fluorescence ratios in Table 1 for the other mutants suggest they are substantially unfolded in native buffer under these conditions.

For the mutants that do exhibit folding stability in native buffer, stability was quantified by Gdn-HCl denaturation experiments to determine $\Delta\Delta G_{stab}$'s, the difference in free energy of stabilization between the wild-type sequence and each mutant. The wild-type domain at pH 7 exhibits an unfolding transition (Fig. 3), which gives a midpoint for Gdn-HCl unfolding (C_m) of 1.55 M and a corresponding free energy of stabilization of −6.8 kcal/mol (Table 1) (Hurle et al., 1994). These values are typical of globular proteins in this size range (Pfeil, 1981) and are also similar to values obtained for other V_L domains (Ahmad & Bigelow, 1986; Tsunenaga et al., 1987).

All of the loop insertion mutants examined prove to be substantially destabilized compared to the wild type (Fig. 3; Table 1). Some mutants, such as RGD1, RGD12, RGD23, and RGD34, are stably folded at pH 7 and RT in native buffer, but at the same time require significantly lower concentrations of Gdn-HCl, compared to wild type, to induce unfolding. Other mutants, such as RGD14, RGD17, and RGD22, are even more unstable, exhibiting significant fluorescence at pH 7 in native buffer at RT. Although it is difficult to rigorously determine the standard free energies of unfolding from the incomplete unfolding curves obtained for many of these proteins, because the na-

```
                     90         100
A REI-VL WT   ..YYCQQY------QSLP------YTFGQ...

  REI-RGD17   ..YYCQQYGTVSRVAKGDWNDDTSYTFGQ...
  REI-RGD4    ..YYCQ--GKISRIPRGDMPDDRS--FGQ...
  REI-RGD18   ..YYCQQ-----KGGRGDSGGK--YTFGQ...
  REI-RGD12   ..YYCQQY----RIPRGDMP----YTFGQ...
  REI-RGD26   ..YYCQ---KGGGGGRGDSK------FGQ...
  REI-RGD23   ..YYCQQ-----RIPRGDMP-----TFGQ...
  REI-RGD21   ..YYCQQ-----RIPRGDMP------FGQ...
  REI-RGD22   ..YYCQ------RIPRGDMP-----TFGQ...
  REI-RGD14   ..YYCQ------RIPRGDMP------FGQ...
  REI-RGD15   ..YYCQQY------PRGD------YTFGQ...

                20                 40
B REI-VL WT   ..TITCQA---SQD---IIKYLNWYQQTP..

  REI-RGD35   ..TITCQASRIPRGDMPIIKYLNWYQQTP..
  REI-RGD32   ..TITCQAKRARGDDM-IIKYLNWYQQTP..
  REI-RGD34   ..TITCQARIPRGDMP-IIKYLNWYQQTP..
  REI-RGD37   ..TITCQAKIGRGDLV-IIKHLNWYQQKP..
  REI-RGD1    ..TITCQA---RGD---IIKYLNWYQQTP..
```

Fig. 2. Sequences of the loop swap mutants in this study. Residues underlined in the wild-type sequence indicate the CDR as defined by sequence homology within the immunoglobulin family (Kabat et al., 1991). Most of the RGD sequences introduced are derived from the active site loops of the snake venom disintegrins kistrin (Dennis et al., 1989) (RGD4, 12, 14, 15, 21, 22, 23, 34, and 35), echistatin (Gan et al., 1988) (RGD32), and barbourin (Scarborough et al., 1991) (RGD17). REI-RGD1 and RGD4 were described previously (Lee et al., 1993). Hyphens are included to accommodate the longest sequence in the lineup. **A:** Sequences of mutants derived from loop swaps in CDR3. **B:** Sequences of mutants derived from loop swaps in CDR1.

Table 1. *Properties of REI V_L-RGD hybrid proteins*

Name	Structure[a]	Δ Residues[b]	Fluorescence[c]	T_m (°C)[d]	C_m (M)[e]	$\Delta\Delta G_{unf}$[f]
RGD17	REI(Δ92-95/+GTVSRVAKGDWNDDTS)	+12	0.78		<0	>6.8
RGD4	REI(Δ90-97/+GKISRIPRGDMPDDRS)	+8	0.75	25	0	>6.8
RGD18	REI(Δ91-95/+KGGRGDSGGK)	+5	0.25		0.13	6.2
RGD12	REI(Δ92-95/+RIPRGDMP)	+4	0.12	37	0.40	5.1
RGD26	REI(Δ90-97/+KGGGGGRGDSK)	+3	0.85			>6.8
RGD23	REI(Δ91-96/+RIPRGDMP)	+2	0.25	33	0.28	5.6
RGD21	REI(Δ91-97/+RIPRGDMP)	+1	0.66			>6.8
RGD22	REI(Δ90-96/+RIPRGDMP)	+1	0.82			>6.8
RGD14	REI(Δ90-97/+RIPRGDMP)	0	0.84			>6.8
RGD15	REI(Δ92-95/+PRGD)	0	0.54		0	6.8
RGD35	REI(Δ27-28/+RIPRGDMP)	+6	0.21		0.20	5.9
RGD32	REI(Δ26-28/+KRARGDDM)	+5	0.22		0.32	5.4
RGD34	REI(Δ26-28/+RIPRGDMP)	+5	0.13	34	0.42	5.0
RGD1	REI(S26R/Q27G)	0	0.12		0.5	4.6
WT REI		–	0.11	55	1.55[e]	–

[a] Structural nomenclature for loop insertion mutants is as described by Wetzel (1988). Deleted wild-type residues are indicated on the left side of the slash, and the new sequence introduced into the gap is indicated to the right of the slash.
[b] Difference in REI length introduced by the loop swap.
[c] Intrinsic fluorescence at 350 nm upon excitation at 295 nm by a 4 μM protein solution at 25 °C relative to the signal in 4 M Gdn-HCl.
[d] Temperature for 50% unfolding as described in Figure 4.
[e] C_m for Gdn-HCl unfolding in 10 mM sodium phosphate, pH 7. The C_m value for REI WT is from Hurle et al. (1994).
[f] ΔG_{unf} for the mutant REI V_L minus ΔG_{unf} for the WT, calculated at the C_m of the WT; see the Materials and methods.

tive baseline is short or nonexistent, reasonable estimates for the amount of destabilization by the loop swaps range from 5 to greater than 7 kcal/mol (Table 1).

Although fluorescence can be sensitive to even small structural changes, the observed tryptophan fluorescence dequenching appears to be associated with a radical disorganization of the folded structure of REI. Unfolding curves such as those shown in Figure 3 are normally associated with global folding transitions. In addition, the unfolded states of REI V_L defined by fluorescence dequenching are nonnative when analyzed by other spectroscopic probes. CD (Fig. 4) of REI-RGD4, a mutant that is about 50% unfolded in native buffer at pH 7 and RT by fluorescence measurements, shows that the molecule's conforma-

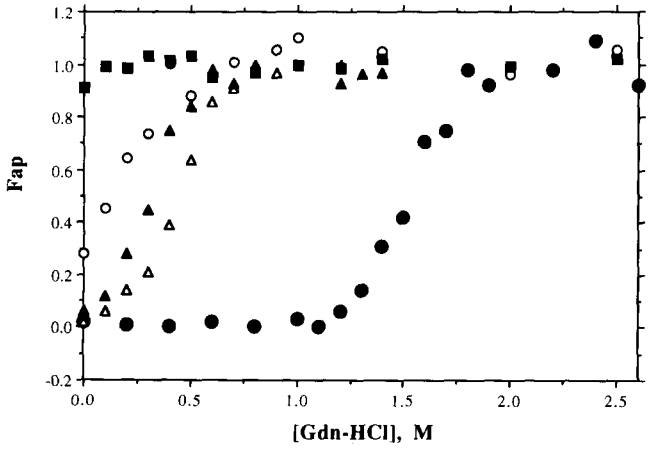

Fig. 3. Mole fraction of unfolded protein (Fap) versus Gdn-HCl concentration for denaturant unfolding of selected REI V_L sequence variants. ●, wild type; ■, RGD17; ○, RGD18; ▲, RGD32; △, RGD34. The data set for wild-type REI V_L is reproduced from Hurle et al. (1994).

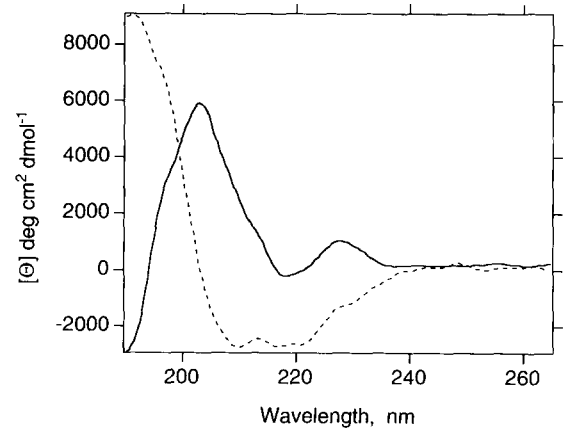

Fig. 4. CD spectra of WT (———) and RGD4 (– – – –) REI V_L. Measurements were made at 25 °C in 10 mM sodium phosphate, pH 7.4, on a Jasco model J-500-C spectropolarimeter equipped with a Macintosh computer-based data acquisition system. The wild-type data were collected on a sample of 0.7 mg/mL in a 0.02-cm-pathlength cell. REI-RGD4 mutant data were collected on a sample of 0.13 mg/mL in a 0.02-cm-pathlength cell. Each spectrum is presented as an average of 20 scans over the 265–190-nm range.

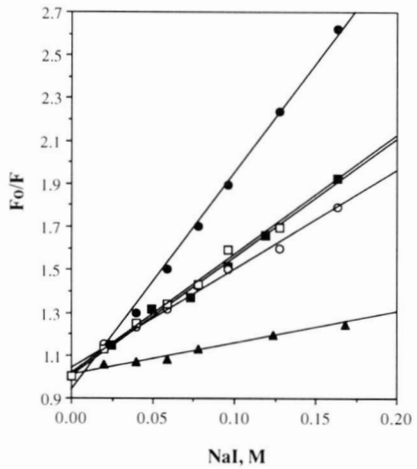

Fig. 5. Stern–Volmer plots of iodide quenching of the tryptophan in various REI V_L sequence variants. All samples were at 10 μg/mL in 10 mM sodium phosphate, 50 mM NaCl, pH 7.4. The wild-type sample also contained Gdn-HCl at 4 M. Fluorescence was read in a Perkin-Elmer MPF-66 spectrofluorometer in a thermostatted cell equilibrated at 22 °C. The $K_{S\text{-}V}$ values obtained for the REI V_L variants plotted are: ●, WT, 10.1; ▲, RGD4, 1.5; ○, RGD14, 4.6; □, RGD21, 5.5; ■, RGD22, 5.5.

tion under the same conditions, although not in statistical coil, is significantly different from the β-sheet conformation exhibited by the wild type (the spectrum of the *E. coli*-produced wild type matches well with the published spectrum of the REI V_L derived from Bence-Jones protein [Brahms & Brahms, 1980]).

It is possible that the observed fluorescence dequenching is due to misfolding into an alternatively packed structure in which Trp^{35} is no longer proximal to the disulfide bond. Stern–Volmer analysis was conducted on several of the REI V_L mutants to determine the accessibility of the Trp to a small molecule quencher. Figure 5 shows Stern–Volmer plots for WT REI V_L unfolded in 4 M Gdn-HCl and a number of apparently unfolded mutants in native buffer. (Because the Trp in the native folded structure is already quenched by the disulfide, it is not possible to derive a $K_{S\text{-}V}$ for Trp in the native state of REI.) This analysis shows that the tryptophan of the unfolded WT is fully accessible (Eftink & Ghiron, 1981) to iodide quenching, giving a $K_{S\text{-}V}$ of 10.1. The mutants RGD14, RGD21, and RGD22 are also sensitive to quenching, with $K_{S\text{-}V}$ values around 5. RGD4 is more resistant to iodide quenching, with a $K_{S\text{-}V}$ value of 1.5. A fully buried tryptophan would be expected to yield a $K_{S\text{-}V}$ value of 0.2 or less (Eftink & Ghiron, 1981). The accessibility of Trp^{35} to the large, charged iodide quencher supports the interpretation that these REI V_L variants are significantly unfolded or misfolded in native buffer.

In principle, inability to achieve a highly quenched fluorescent (native-like) state might be attributable to the absence of the intradomain disulfide bridge. Figure 6 shows an SDS-PAGE gel of a wild-type REI V_L and series of loop-swap mutants. The gel shows that, in general, a nonreduced, alkylated sample migrates more rapidly through the gel than a reduced, alkylated sample of the same protein. This is consistent with the existence of a disulfide in the nonreduced sample (Pollitt & Zalkin, 1983). Such mobility shifts can be difficult to observe in small proteins like REI V_L, and examination of Figure 6 shows the shifts to be small and variable. The absence of a shift in RGD14 probably simply reflects the limit of detection for this mutant. Other means were also used to confirm the existence of the disulfide in some of the mutant V_L domains. An alkylated sample of REI-RGD4 gives a parent ion in mass spectrometry consistent with molecular weight expected for the oxidized mutant and inconsistent with the alkylated molecule (L. Helms & D. McNulty, unpubl. result).

Although these proteins were purified by reverse-phase HPLC, in which they required acetonitrile concentrations for elution similar to that found for wild type, it is formally possible that some of the mutants exist as disulfide-linked dimers or oligomers that would likely exhibit alternate fluorescence characteristics. However, Figure 6 shows there are no disulfide-linked dimers or trimers in these preparations. Neither was there any Coomassie blue-stained material at higher molecular weights in the gel (not shown), consistent with the absence of larger covalent aggregates.

Noncovalent aggregation could in principle also compromise proper interpretation of the fluorescence data, by offering a competing "folding" pathway for the proteins and thus exaggerating their apparent instability as monitored by fluorescence dequenching. In fact, some REI V_L point mutants associated with light chain deposition disease exhibit substantial unfolding-related aggregation by gel-filtration analysis (L.R. Helms & R. Wetzel, unpubl.). However, gel filtration analysis indicates no large aggregates to be present in solutions of RGD 15 and RGD 21 (data not shown).

The destabilization of REI V_L variants assessed by denaturant unfolding was confirmed for some of the mutants by thermal unfolding studies. Figure 7B shows that the wild-type domain at pH 7 undergoes a cooperative thermal unfolding transition, as monitored by fluorescence, with a T_m of 55 °C. In contrast, RGD12, RGD23, and RGD34, three of the more stable of the

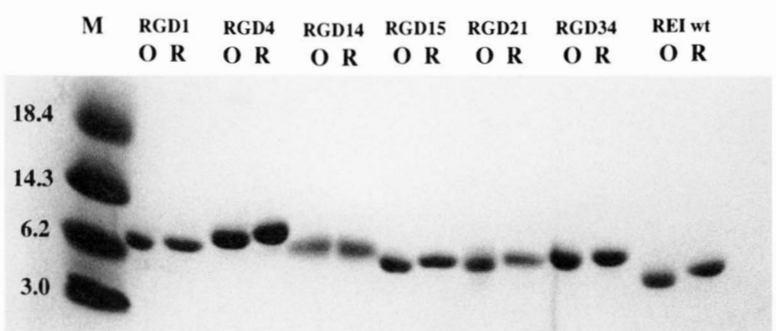

Fig. 6. Coomassie brilliant blue-stained, nonreducing SDS–polyacrylamide gel of various mutants of REI V_L. Reduced (R) samples were treated 30 min at 37 °C with 1 mM dithiothreitol in SDS-PAGE gel loading buffer, after which iodoacetic acid was added to 5 mM and incubated an additional 30 min. Nonreduced (O) samples were treated only with 5 mM iodoacetic acid. Samples were run on a 12% acrylamide gel (Laemmli, 1970).

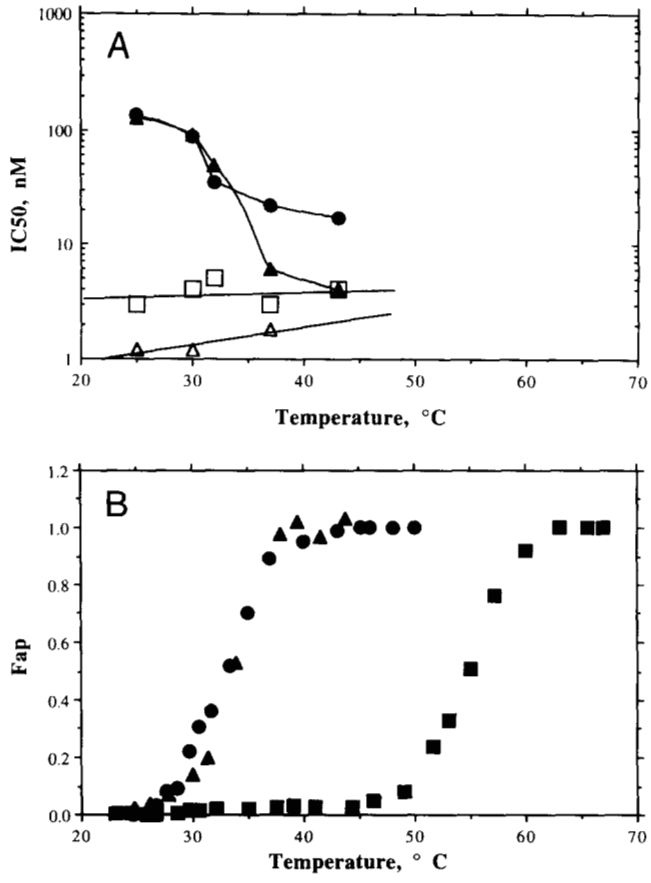

Fig. 7. Temperature dependence of binding and folding. ■, WT REI V_L; ●, REI-RGD23; ▲, REI-RGD34; □, SKF#106760; △, echistatin. **A:** Temperature dependence of IC_{50}'s for inhibition of binding of biotinylated fibrinogen to $\alpha_{IIb}\beta_3$. Assays were conducted as described (Lee et al., 1993), with all incubations done at 30 °C, except for the competition step, which was done at the temperatures indicated. Data points are the mean values of three or four independent determinations. The lack of significant temperature dependence to the inhibitory action of the peptide antagonist SKF#106760 (Samanen et al., 1991) and the snake venom disintegrin echistatin (Gan et al., 1988) show that the temperature dependence for RGD23 and RGD34 derives from a thermal effect on these REI mutants and not on the receptor. The wild-type REI V_L exhibits no binding at any temperature (Lee et al., 1993). Other destabilized RGD-containing mutants also exhibit thermal dependent binding (data not shown). **B:** Thermal unfolding curves for REI V_L domains.

loop replacement mutants, have T_m's that are at or below physiological temperature, about 20 °C lower than the WT (Fig. 7B; Table 1). Interestingly, the thermal unfolding transitions of these mutants are mirrored in the thermal sensitivity of their receptor binding. Figure 7A shows that these mutants are good antagonists of fibrinogen binding to the platelet receptor $\alpha_{IIb}\beta_3$ by virtue of their installed RGD sequences (Lee et al., 1993; Helms & Wetzel, 1994), but that binding affinity depends on assay temperature. The transitions in the temperature dependence of the IC_{50} values exactly follow the independently measured unfolding transitions. Because unrelated receptor antagonists do not exhibit temperature-dependent activity (Fig. 7A), the thermal dependence is not due to effects on the receptor.

The sequence RGD37 listed in Figure 2, containing the replacement of wild-type residues 26–28 with the sequence KIGRGDLV, was generated in an REI V_L sequence that contained two stabilizing mutations, Y32H (1.1 kcal/mol) and T39K (1.3 kcal/mol) (Frisch et al., 1994; H.-J. Fritz, pers. comm.). REI V_L with the two stabilizing mutations exhibited a C_m of 2.25 M Gdn-HCl, indicative of a $\Delta\Delta G$ with respect to WT of −3.1 kcal/mol. RGD37 exhibited a C_m of 0.55 M Gdn-HCl, for a $\Delta\Delta G$ with respect to REI(Y32H/T39K) of 7.0 kcal/mol destabilization. Thus, a sequence lacking prolines inserted into CDR1 is actually more destabilizing than is RIPRGDMP in both REI-RGD34 and REI-RGD35. Table 1 shows that another CDR1 insertion lacking prolines, KRARGDDM in REI-RGD32, is also highly destabilizing.

Discussion

We have been developing the immunoglobulin V_L domain REI for use as a "presentation scaffold" (Lee et al., 1993; Helms & Wetzel, 1994)—a small domain expected to hold an installed peptide sequence into a defined conformation and thus help elucidate the receptor-binding conformations of peptides. We chose the integrin receptor-binding motif Arg-Gly-Asp (RGD) for installation and study because of its small size and because of speculations on the importance of conformation in its receptor binding selectivity (Ruoslahti & Pierschbacher, 1987). We chose an immunoglobulin V_L domain as a scaffold, in part, based on the expectation that its CDRs should be particularly adept at accommodating sequence changes without incurring significant losses in folding stability.

The data presented here, however, suggest that sequence replacements in the CDR loops of immunoglobulin domains can be significantly destabilizing to the native structure of the domain. The destabilizing effects are observed for insertions in two separate CDRs and are seen for such changes as a single point mutation, a short replacement that leaves the loop length unchanged, and insertions that increase loop size to lengths equal to or greater than those known to be accommodated at these positions in immunoglobulins. The results have important implications for antibody engineering experiments as well as for the general perception that loops are particularly resilient parts of proteins that are capable of accommodating many sequence changes.

The nature of the nonnative state generated in destabilized REI V_L mutants

A number of experiments were conducted to investigate the degree of unfolding in the fluorescence dequenched state monitored in these studies. These proteins contain no large covalent or noncovalent aggregates, as assessed by nonreducing SDS-PAGE and native gel-permeation chromatography. The proteins contain the expected intramolecular disulfide bond, so that the absence of quenching of Trp in apparently unstable mutants cannot be ascribed to the absence of the disulfide. CD of the destabilized RGD4 in native buffer indicates secondary structure substantially different from the folded wild-type REI V_L, but not statistical coil. Whatever the structure of this nonnative state, it must be substantially reorganized from the native because the normally buried tryptophan is accessible to dynamic quenching by iodide. Further studies on the solution structures of the highly destabilized REI V_L mutants are warranted, especially because these molecules may hold important clues to the

partially unfolded states implicated in light chain amyloidosis and deposition disease (Hurle et al., 1994).

Loop swaps and destabilization

A number of mutagenesis studies have been conducted suggesting that loop sequences have relatively little effect on the thermodynamic stability of globular proteins (reviewed in El Hawrani et al., 1994). Our observation of significant destabilization by loop replacements is not unprecedented, however. Two *Staphylococcus* nuclease mutants constructed by replacing a five-residue β-turn with a β-turn sequence from concanavalin A were found to be destabilized by about 4.5 kcal/mol with a corresponding reduction in T_m of about 20 °C (Eftink et al., 1991). A mutant constructed by replacing the loop region of chymotrypsin inhibitor 2 with a nonapeptide sequence found in an α-helix in subtilisin Carlsberg is also significantly destabilized by about 8 kcal/mol (Osmark et al., 1993). In both of these cases, crystallographic analysis suggests that destabilization may be attributable to the loss of favorable long-range interactions in the WT between loop residues and residues elsewhere in the "host" structure (Hynes et al., 1989; Osmark et al., 1993).

The mechanism(s) by which the loop replacements described here destabilize REI V_L are not obvious. Most of the installed sequences are derived from the long, relatively disordered loops of the platelet antagonist disintegrins (Gould et al., 1990) and thus might be expected to be both relatively flexible and compatible with extended, solvated loop conformations. The RGD (or KGD) sequence itself, which is common to all of the inserted loops, would also be expected to be compatible with a solvated loop environment; in fact, the RGD sequence has been observed in a topologically similar loop to CDR3 in the fibronectin domain of several matrix proteins (Leahy et al., 1992; Main et al., 1992; Dickinson et al., 1994) as well as in the loops of disintegrins. The inserted loops vary considerably in length (leading to net changes in domain length of between 0 and +12 residues) and in the presence or absence of residues that increase (Gly) or decrease (Pro) peptide backbone configurational flexibility. Although many of the destabilizing replacements contain proline residues, two CDR1 mutants lacking prolines, REI-RGD32 (Table 1) and REI-RGD37 (Results), suffer destabilization equal to or greater than that of a mutant containing an insert of equal length and including prolines (REI-RGD34).

Two different CDRs—one between sheets (CDR1) and one within a sheet (CDR3)—were replaced with very similar effects on stability. In CDR3, different insertion points were explored, including points in the loop removed from any possible interaction with the body of the protein. In all cases, however, the destabilization was significant. For example, although the entire 90–97 sequence is within the CDR3 as defined by sequence homology (Fig. 2), it is possible that some replacements of residues 90, 91, and 97 may make altered packing interactions with framework residues, resulting in destabilization. However, replacements that leave residues 90, 91, 96, and 97 unchanged are also observed to destabilize the domain by 5 (RGD12) to 7 (RGD15) kcal/mol (Table 1).

As in the other examples of destabilizing loop replacements described previously (see above), interactions lost from the wild type in the process of replacement may account for some of the observed destabilization. Table 2 lists the hydrogen bonds found in the crystal structure of wild-type REI involving residues 26–29 and 90–97. There are only a few long-range H-bonds in each case. In CDR1, the serine side-chain OH is involved in several H-bonds thqat would be lost if Ser were replaced by another residue; however, REI-RGD35 retains this serine and is just as destabilized as the related REI-RGD34, which lacks this residue. The other long-range H-bonds, at residues 3–26 and 29–68, may be lost if the inserted loop occupies a conformation that brings atoms 26:N and 29:N out of position with respect to the wild-type structure.

Table 2 also lists a number of long-range H-bonds that may be lost in loop-swap mutants involving CDR3. The side chain of Gln 90 plays a very important role in orienting the wild-type loop, by lying in the loop plane and making H-bonds with residues 93, 95, and 97 (Table 2). This is a standard feature of CDR3 loops (Tramontano et al., 1989). This may be important in the destabilization of some of the mutants reported here, because even if Gln 90 is retained (as it is in many of the mutants), loops of different sizes compared with WT may not be able to adopt these H-bonds.

Although it is thus possible that lost H-bond contributions may account for some stability decrease in the mutants, further experiments and structural determination will be required to elucidate the mechanism of destabilization. X-ray crystal structure analysis of REI-RGD34, in progress, shows that the basic immunoglobulin fold is maintained in the folded state of this destabilized mutant (B. Zhao, L.R. Helms, E. Winborne, S. Abdel-Meguid, & R. Wetzel, unpubl. results).

Table 2. *Hydrogen bonds*[a] *in wild-type REI-*V_L

Donor	Acceptor	Molecule A	Molecule B
3:N	26:OG	3.24	–
26:N	3:O	3.27	3.23
26:OG	26:O	3.23	
26:OG	3:O		3.34
29:N	27:O	3.16	
29:N	28:OD2	3.02	2.98
29:N	68:O	3.29	3.14
30:N	28:OD2	2.75	2.48
90:N	97:O	3.31	3.36
90:NE2	1:OD2	3.16	–
90:NE2	93:O	3.14	–
90:NE2	95:O	3.02	–
90:NE2	97:OG1	3.21	–
91:N	32:O	3.29	3.22
91:N	90:OE1	3.32	–
92:N	90:OE1	2.88	–
93:OG	93:O	3.18	3.02
95:N	1:OD2	–	3.16
97:N	95:O	3.10	–
97:OG1	1:OD1	2.94	–
97:OG1	2:O	2.85	2.71
97:OG1	97:O	3.16	3.12
97:OG1	90:OE1	–	3.28

[a] Hydrogen bonds of acceptable bond lengths and geometries were identified in the coordinate set 1REI in the Brookhaven database (Bernstein et al., 1977; Abola et al., 1987) using Insight II™ (Biosym Technologies, San Diego, California). The independent data for Molecule A and Molecule B reflect the two independent monomers in the single dimer of the unit cell (Epp et al., 1974).

An important implication of these results is that there may be significant limitations on the allowed sequences and lengths of protein loops, including antibody hypervariable loops.

Thermal dependence of binding activity

The temperature dependence of receptor binding by these mutants confirms the ability of appropriately chosen scaffolds to control the conformation of an installed bioactive ligand (Wetzel, 1991). At the same time, the results underscore the importance of considering domain stability in the choice of a domain for a presentation scaffold. If the destabilizations of 5–7 kcal/mol found here and elsewhere for loop swaps prove to be somewhat general, it is clearly advantageous to choose a domain with a free energy of folding of significantly more than 7 kcal/mol for phage display and other scaffold experiments.

Stability of the immunoglobulin fold

Although immunoglobulin molecules are known to be highly stable proteins, relatively little is known about the properties of the fundamental immunoglobulin fold. Studies on wild-type V_L domains released by limited proteolysis of Bence-Jones proteins or expressed by recombinant DNA methods show that these immunoglobulin domains are about as stable as expected for proteins of about 100 amino acids, with ΔG^0_{stab} values around 3–9 kcal/mol (Ahmad & Bigelow, 1986; Tsunenaga et al., 1987; Hurle et al., 1994; Steipe et al., 1994; Stevens et al., 1995). Wild-type REI V_L falls within this range, with a ΔG^0_{stab} of about 6.5 kcal/mol (Hurle et al., 1994; Kolmar et al., 1994). As with other globular proteins, the folding stability of the V_L domain is sensitive to point mutations. Previously, we showed that single point mutations in framework residues of REI V_L can destabilize the domain by as much as 3 kcal/mol when transferred into REI V_L (Hurle et al., 1994). Similarly, Steipe et al. (1994) identified nonconserved residues in the framework sequences of the V_L domain of the antibody McPC603 and showed that conversion of these residues to the most common residues at these positions provided enhanced stability by up to 1.5 kcal/mol. Stabilizing point mutations for REI V_L have also been described (Frisch et al., 1994). The role of loop residues on immunoglobulin domain stability has not been investigated previously.

The work described here suggests that many loop sequences are not compatible with the variable domain framework, and thus that evolution of the natural repertoire of CDR sequences has occurred in the context of significant selective pressure against destabilizing sequences in the CDR positions. This is consistent with the observation of conserved residues (Kabat et al., 1991) and canonical structures (Chothia et al., 1989) in antibody CDR loops. Although canonical CDR structures are usually discussed in terms of the role of CDRs in antigen binding, our results suggest an additional important role of CDR sequence and structure: their contribution to domain stability. Given this possibility, it will be interesting to see the frequency with which new "canonical" CDR structures arise in antibodies generated through the emerging use of CDR sequence randomization methods (Barbas et al., 1993; Fisch et al., 1994).

V_L domains that possess poor stability might still be capable of forming functional antibodies, because additional stabilization energy is expected from the interaction between the folded V_H and V_L domains (Pantoliano et al., 1991) and would thus be expected to help rescue a V_L of impaired stability.

It remains to be seen whether viable antibodies containing poorly stable V_L domains can be produced; experiments are in progress to address this question. Even if it turns out that functional antibodies containing destabilized variable domains can be produced, they may nonetheless possess reduced net stabilities, which in turn might influence their temperature dependence of binding—as shown here for the $\alpha_{IIb}\beta_3$ interaction with REI V_L—as well as other properties such as immunogenicity and clearance. Furthermore, domain stability within the light chain itself has been suggested to play an important role in determining the metabolic fate of circulating light chains in amyloidosis (Hurle et al., 1994). Domain stability may also play an important role in immunoglobulin biosynthesis.

The unstable proteins described here were isolated from inclusion bodies (IBs) formed in the *E. coli* periplasm. Recently, Knappik and Plueckthun (1995) described amino acid replacements in the V_H domain of McPC603 that influence IB formation in *E. coli* expression of antibody fragments containing both V_L and V_H domains. Consistent with previous studies on the role of mutations in IB formation (Mitraki et al., 1993; Wetzel & Chrunyk, 1994), the authors described results suggesting that their mutations direct material into IBs by influencing folding rates rather than thermodynamic stability. Results of expression of REI V_L containing a variety of stabilizing and destabilizing mutations, both in framework regions (Hurle et al., 1994) and in CDRs (described here), suggest that, in contrast to these other systems, thermodynamic stability considerations may be sufficient to account for IB formation in the expression of the V_L domain. A formal analysis of REI V_L expression data from this point of view will be reported elsewhere (W. Chan, R. Wetzel, L.R. Helms, S. Ngola, & B. Maleeff, unpubl. results).

Materials and methods

Materials

Fibrinogen receptor $\alpha_{IIb}\beta_3$ isolated from human platelets was provided by Kyung Johanson (SmithKline). Cyclic RGD antagonist SKF#106760 was provided by James Samanen (SmithKline). *E. coli* MM294 was used for transformations and expression of REI V_L.

Production of REI V_L variants

Mutants were constructed from the WT coding sequence (Palm & Hilschmann, 1973) using cassette mutagenesis (Wells et al., 1985) and sequences confirmed by dideoxy nucleotide sequencing using Sequenase (USB) and established procedures (Sanger et al., 1977). All other recombinant DNA procedures were as described by Sambrook et al. (1989).

Proteins were expressed in *E. coli* under control of the *tac* promoter with a *pel B* leader sequence directing export to the periplasm of the cell. Cells were grown at 37 °C. Wild-type REI V_L was expressed solubly in the periplasm but most of the mutants accumulated in periplasmic IBs, presumably because of their decreased stabilities (W. Chan, L.R. Helms, & R. Wetzel, unpubl.). Despite their low stabilities, these proteins accumulated stably in the cell, presumably because IB formation protects them from proteolysis (Shortle & Meeker, 1989). These IBs contain protein

that is correctly processed (W. Chan, L.R. Helms, & R. Wetzel, unpubl.). These were solubilized in 8 M Gdn-HCl, then dialyzed into 1 M urea, 0.1% trifluoroacetic acid, and centrifuged to remove aggregated protein. Wild-type and solubilized mutants were all subjected to a final purification step of reverse-phase HPLC and analyzed for purity by HPLC and SDS-PAGE.

Denaturant and thermal unfolding studies

The integrity of the REI V_L fold was monitored by the fluorescence of the conserved immunoglobulin domain tryptophan, which is highly quenched in the native state due to its close proximity to the disulfide bond. Fluorescence was determined in a Perkin-Elmer MPF-66 spectrofluorometer with excitation at 295 nm and emission at 350 nm. The extent to which a mutant REI V_L is unfolded in native buffer at room temperature was initially estimated by determining the fluorescence of a 1 μM (monomer) sample in 10 mM sodium phosphate, pH 7.4 in the presence and absence of 4 M Gdn-HCl. The ratio of fluorescence in native buffer versus denaturant is low for a variant that is folded in native buffer (like the wild type, ratio = 0.11; see Table 1) and high for a variant that is substantially unfolded in native buffer (like RGD14, ratio = 0.84).

The degree of stability was confirmed and quantified by using fluorescence to determine the full Gdn-HCl denaturation curves. Samples of approximately 10 μg/mL (~1 μM, monomer) were incubated overnight at room temperature at various Gdn-HCl concentrations in 10 mM sodium phosphate, pH 7.4, and fluorescence was determined. Data were analyzed by nonlinear least-squares fitting to a two-state model using either NLIN (SAS Institute, Cary, North Carolina) or IGOR Pro (WaveMetrics) using the simultaneous baseline fitting method of Santoro and Bolen (1988) to generate fitted curves and values for transition midpoints (C_m's), transition slopes (A values), and free energies of unfolding (ΔG values). $\Delta\Delta G$ values were calculated from the C_m and A values with the WT curve as reference point, using the formula $\Delta\Delta G = \Delta G^{mut} - \Delta G^{WT} = A^{WT} * C_m^{mut} - A^{WT} * C_m^{WT}$, where $A^{WT} = -4.4$ kcal/mol/M (Gdn-HCl), determined experimentally.

Thermal stability was determined with samples of about 2 μM in REI V_L monomer in 10 mM sodium phosphate, pH 7.4, by monitoring fluorescence as a function of temperature. Temperature was controlled using a jacketed cell holder and a circulating water bath, with a temperature ramp rate of about 0.25 °C/min. Cell temperature was monitored using a thermocouple. As expected, fluorescence changed gradually with temperature until the thermal transition was reached, at which point fluorescence sharply increased. The slopes of the fluorescence variation with temperature outside the cooperative unfolding transition were determined and used to generate thermal Fap plots and T_m values.

Integrin receptor antagonist activity

The inhibition of the binding of biotinylated fibrinogen to the integrin receptor $\alpha_{IIb}\beta_3$ was determined by a microtiter assay as described previously (Lee et al., 1993), except the competition step was performed at the temperature indicated. The IC_{50} values were determined from a four-parameter fit using the software Softmax (Molecular Devices, Inc.), and the concentration of the inhibitor at 50% binding was recorded.

Acknowledgments

We thank Mark R. Hurle for helpful suggestions and for critical reading of the manuscript.

References

Abola E, Bernstein FC, Bryant SH, Koetzle TF, Weng J. 1987. Protein Data Bank. In: Allen FH, Bergerhoff G, Sievers R, eds. *Crystallographic databases – Information content, software systems, scientific applications*. Bonn/Cambridge/Chester: Data Commission of the International Union of Crystallography. pp 107–132.

Ahmad F, Bigelow CC. 1986. Estimation of the stability of globular proteins. *Biopolymers 25*:1623–1634.

Barbas CF III, Languino LR, Smith JW. 1993. High-affinity self-reactive human antibodies by design and selection: Targeting the integrin ligand binding site. *Proc Natl Acad Sci USA 90*:10003–10007.

Bernstein FC, Koetzle TF, Williams GJ, Meyer EJ, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Brahms S, Brahms J. 1980. Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J Mol Biol 138*: 149–178.

Branden C, Tooze J. 1991. *Introduction to protein structure*. New York: Garland.

Chothia C, Lesk AM, Tramontano A, Levitt M, Smith-Gill SJ, Air G, Sheriff S, Padlan EA, Davies D, Tulip WR, Colman PM, Spinelli S, Alzari PM, Poljak RJ. 1989. Conformations of immunoglobulin hypervariable regions. *Nature 342*:877–883.

Dennis MS, Henzel WJ, Pitti RM, Lipari MT, Napier MA, Deisher TA, Bunting S, Lazarus RA. 1989. Platelet glycoprotein iib-iiia protein antagonists from snake venoms: Evidence for a family of platelet-aggregation inhibitors. *Proc Natl Acad Sci USA 87*:2471–2475.

Dickinson CD, Veerapandian B, Dai XP, Hamlin RC, Xuong NH, Ruoslahti E, Ely KR. 1994. Crystal structure of the tenth type III cell adhesion module of human fibronectin. *J Mol Biol 230*:1079–1092.

Eftink MR, Ghiron CA. 1981. Fluorescence quenching studies with proteins [Review]. *Anal Biochem 114*:199–227.

Eftink MR, Ghiron CA, Kautz RA, Fox RO. 1991. Fluorescence and conformational stability studies of *Staphylococcus* nuclease and its mutants, including the less stable nuclease-concanavalin A hybrids. *Biochemistry 30*:1193–1199.

El Hawrani AS, Moreton KM, Sessions RB, Clarke AR, Holbrook JJ. 1994. Engineering surface loops of proteins – A preferred strategy for obtaining new enzyme function. *Trends Biotechnol 12*:207–211.

Epp O, Colman P, Fehlhammer H, Bode W, Schiffer M, Huber R, Palm W. 1974. Crystal and molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI. *Eur J Biochem 45*:513–524.

Fisch I, Griffiths A, Winter G. 1994. Generation of large phage antibody repertoires using in vivo recombination. *J Cell Biochem 18D*:T201.

Frisch C, Kolmar H, Fritz HJ. 1994. A soluble immunoglobulin variable domain without a disulfide bridge: Construction, accumulation in the cytoplasm of *E. coli*, purification and physicochemical characterization. *Biol Chem Hoppe-Seyler 375*:353–356.

Gan ZR, Gould RJ, Jacobs JW, Friedman PA, Polokoff MA. 1988. Echistatin. A potent platelet aggregation inhibitor from the venom of the viper, *Echis carinatus*. *J Biol Chem 263*:19827–19832.

Gould RJ, Polokoff MA, Friedman PA, Huang TF, Holt JC, Cook JJ, Niewiarowski S. 1990. Disintegrins: A family of integrin inhibitory proteins from viper venoms. *Proc Soc Exp Biol Med 195*:168–171.

Helms LR, Wetzel R. 1994. Proteolytic excision and in situ cyclization of a bioactive loop from an REI-VL presentation scaffold. *Protein Sci 3*:1108–1113.

Hurle MR, Helms LR, Li L, Chan W, Wetzel R. 1994. A role for destabilizing amino acid replacements in light chain amyloidosis. *Proc Natl Acad Sci USA 91*:5446–5450.

Hynes TR, Kautz RA, Goodman MA, Gill JF, Fox RO. 1989. Transfer of a beta-turn structure to a new protein context. *Nature 339*:73–76.

Jones PT, Dear PH, Foote J, Neuberger MS, Winter G. 1986. Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature 321*:522–525.

Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. 1991. *Sequences of proteins of immunological interest*. Washington, D.C.: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health.

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

Kolmar H, Frisch C, Kleeman G, Goetze K, Stevens FJ, Fritz HJ. 1994. Dimerization of Bence–Jones proteins: Linking the rate of transcription from an *Escherichia coli* promoter to the association constant of REIv. *Biol Chem Hoppe-Seyler 375*:61–70.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature 227*:680–685.

Leahy DJ, Hendrickson WA, Aukhil I, Erickson HP. 1992. Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science 258*:987–991.

Lee G, Chan W, Hurle MR, DesJarlais RL, Watson F, Sathe GM, Wetzel R. 1993. Strong inhibition of fibrinogen binding to platelet receptor αIIb-βIII by RGD sequences installed into a presentation scaffold. *Protein Eng 6*:745–754.

Main AL, Harvey TS, Baron M, Boyd J, Campbell ID. 1992. The three-dimensional structure of the tenth type III module of fibronectin: An insight into RGD-mediated interactions. *Cell 71*:671–678.

Matthews BW. 1993. Structural and genetic analysis of protein stability. *Annu Rev Biochem 62*:139–160.

Mitraki A, Danner M, King J, Seckler R. 1993. Temperature-sensitive mutations and second-site suppressor substitutions affect folding of the P22 tailspike protein in vitro. *J Biol Chem 268*:20071–20075.

Osmark P, Sorenson P, Poulson FM. 1993. Context dependence of protein secondary structure formation: The three-dimensional structure and stability of a hybrid between chymotrypsin inhibitor 2 and helix E from subtilisin Carlsberg. *Biochemistry 32*:11007–11014.

Palm W, Hilschmann N. 1973. Die Primaerstruktur einer kristallinen monoklonalen Immunglobulin-L-Kette vom k-Typ, Subgruppe I (Bence–Jones-Protein Rei.): Ein Beitrag zur Aufklaerung der dreidimensionalen Struktur der Immunglobuline. *Hoppe-Seyler's Z Physiol Chem 354*: 1651–1654.

Pantoliano MW, Bird RE, Johnson S, Asel ED, Dodd SW, Wood JF, Hardman KD. 1991. Conformational stability, folding, and ligand-binding affinity of single-chain Fv immunoglobulin fragments expressed in *Escherichia coli*. *Biochemistry 30*:10117–10125.

Pfeil W. 1981. The problem of the stability of globular proteins. *Mol Cell Biochem 40*:3–28.

Pollitt S, Zalkin H. 1983. Role of primary structure and disulfide bond formation in β-lactamase secretion. *J Bacteriol 153*:27–32.

Riechmann L, Clark M, Waldmann H, Winter G. 1988. Reshaping human antibodies for therapy. *Nature 332*:323–327.

Ruoslahti E, Pierschbacher MD. 1987. New perspectives in cell adhesion; RGD and integrins. *Science 238*:491–497.

Samanen J, Ali FE, Romoff T, Calvo R, Sorenson E, Vasko J, Storer B, Berry D, Bennett D, Strohasacker M, Powers D, Stadel J, Nichols A. 1991. Development of a small RGD peptide fibrinogen receptor antagonist with potent antiaggregatory activity *in vitro*. *J Med Chem 34*:3114.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA 74*:5463–5468.

Santoro MM, Bolen DW. 1988. Unfolding free energy changes determined by the linear extrapolation method. I. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry 27*: 8063–8074.

Scarborough RM, Rose JW, Hsu MA, Phillips DR, Fried VA, Campbell AM, Nannizzi L, Charo IF. 1991. Barbourin: A gpIIb-IIIa-specific integrin antagonist from the venom of *Sistrurus m. barbouri*. *J Biol Chem 266*: 9359–9362.

Shortle D. 1992. Mutational studies of protein structures and their stabilities. *Q Rev Biophys 25*:205–250.

Shortle D, Meeker AK. 1989. Residual structure in large fragments of staphylococcal nuclease: Effects of amino acid substitutions. *Biochemistry 28*:936–944.

Steipe B, Schiller B, Pluckthun A, Steinbacher S. 1994. Sequence statistics reliably predict stabilizing mutations in a protein domain. *J Mol Biol 240*:188–192.

Stevens PW, Raffen R, Hanson DK, Deng YL, Berrios-Hammond M, Westholm FA, Murphy C, Eulitz M, Wetzel R, Solomon A, Schiffer M, Stevens FJ. 1995. Recombinant immunoglobulin variable domains generated from synthetic genes provide a system for in vitro characterization of light chain amyloid proteins. *Protein Sci 4*:421–432.

Tramontano A, Chothia C, Lesk AM. 1989. Structural determinants of the conformations of medium-sized loops in proteins. *Proteins Struct Funct Genet 6*:382–394.

Tsunenaga M, Goto Y, Kawata Y, Hamaguchi K. 1987. Unfolding and refolding of a type kappa immunoglobulin light chain and its variable and constant fragments. *Biochemistry 26*:6044–6051.

Wells JA, Vasser M, Powers DB. 1985. Cassette mutagenesis: An efficient method for generation of multiple mutations at defined sites. *Gene 34*: 315–323.

Wetzel R. 1988. Commentary: Structure, function, nomenclature. *Protein Eng 2*:1–3.

Wetzel R. 1991. Commentary: Learning from the immune system; laboratory methods for creating and refining molecular diversity in polypeptides. *Protein Eng 4*:371–374.

Wetzel R, Chrunyk BA. 1994. Inclusion body formation by interleukin-1β depends on the thermal sensitivity of a folding intermediate. *FEBS Lett 350*:245–248.