Intrabody Construction and Expression. I The Critical Role of V_L Domain Stability **Ettore Ohage, Boris Steipe** 000–000

Reprinted from *J. Mol. Biol.* (1999) **00**, 000–000

JMB

Intrabody Construction and Expression. I The Critical Role of V_L Domain Stability

Ettore Ohage and Boris Steipe

Article No. jmbi.1999.3019 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 00, 1–10

JMB

Intrabody Construction and Expression. I The Critical Role of V_L Domain Stability

Ettore Ohage and Boris Steipe*

Genzentrum der Ludwig-Maximilians-Universität Feodor-Lynen Straße 25 81377, München, Germany

We have constructed a panel of hyperstable immunoglobulin V_L domains by a rational approach of consensus sequence engineering and combining stabilizing point mutations. These prototype domains unfold fully reversibly, even when the conserved structural disulfide bridge is reduced. This has allowed us to probe the factors that limit the expression yield of soluble immunoglobulin domains in the reducing environment of the cytoplasm (intrabodies). The most important factor is thermodynamic stability, and there is an excellent quantitative correlation between stability and yield. Surprisingly, an unprocessed N-terminal methionine residue can severely compromise V_L stability, but this problem can be overcome by changing the amino acid following the initiator methionine residue. Transcription from the strong T7 promoter does not increase the amount of soluble material over that obtained from the *tet*A promoter, but large amounts of inclusions bodies can be obtained. Elevated temperature shifts protein from a productive folding pathway to aggregation. The structural disulfide bridge does not form in the cytoplasm, but the two consensus cysteine residues can be quantitatively oxidized *in vitro*. In summary, stability engineering provides a plannable route to the high-yield cytoplasmic expression of functional intrabody domains.

Keywords: intrabodies; recombinant expression; V_L domain; protein stability; protein engineering

*Corresponding author

Introduction

The antibody molecule is nature's paradigm of evolutionary engineering. A process of somatic hypermutation on the level of individual domains, followed by clonal expansion of successful variants, generates a repertoire of complementary protein paratopes. These circulate throughout the extracellular space to respond to virtually any challenge by non-self molecules. The recombinant production of immunoglobulin domains, first established in *Escherichia coli* (Skerra & Plückthun, 1988; Better & Horwitz, 1988), has contributed important new perspectives to exploit this principle for research, biotechnology and medical diagnosis and therapy (Hudson, 1998, and references therein). The evolutionary paradigm itself has been employed to evolve new antibodies *in vitro*, e.g. in phage display systems (Hoogenboom *et al*., 1998). Considering the successes of recombinant antibody technology, it is remarkable that the recent change of focus of the life sciences to the intracellular space has not been accompanied by a widespread use of intrabodies: intracellular antibody domains.

Variable domain structures are highly homologous (Padlan, 1996) consisting of a β-sandwich connected by a conserved, structural disulfide bond between Cys23 and Cys88 ($V_\kappa^\dagger$). This disulfide bond contributes a major part of the domain's folding stability but, since the formation of disulfide bonds requires oxidation, it cannot form in a reducing environment such as the cellular cytoplasm. Disruption of the disulfide bond has been shown to result in aggregation (Glockshuber *et al*., 1992). Therefore, although the engineering of antibody domains as intracellular complementary surfaces would be an attractive goal, the soluble cyto-

Abbreviations used: β-ME, β-mercapto-ethanol; CDR, complementarity-determining region; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); DTT, dithiothreitol; EDTA, ethylenediaminotetraacetic acid; Fv, heterodimer containing V_L and V_H; ΔG_F, free energy of folding; GdmCl, guanidinium chloride; PBD, phosphate-buffered saline; scFv, single chain Fv fragment; V_L and V_H, variable domains of immunoglobulin light and heavy chains, respectively; ASA, accesible surface area. The subscript κ denotes the kappa V_L subtype.

E-mail address of the corresponding author: steipe@lmb.uni-muenchen.de

0022-2836/99/0000001–10 $30.00/0

plasmic expression of variable antibody domains was long considered impossible. Intrabodies would minimally have to compensate for a severe loss of stability in the cytoplasm but other factors, such as promoter strength, temperature, N-terminal processing, slow folding reactions and the requirement for dimerisation may prove to become limiting for expression even when stability would be sufficient in principle.

Serendipitously, Fritz and co-workers have found a stabilizing mutation which lead to soluble cytoplasmic expression of the human V_κ myeloma protein REI as an N-terminal fusion to β-lactamase (Frisch *et al*., 1994). This single mutation, Y32H, was sufficient to compensate for the subsequent replacement of Cys23 without significant alteration of the domain structure (Uson *et al*., 1997). But although a series of further stabilized mutants was reported by the same group, only one could be expressed solubly in the periplasm as a fusion protein in significant amounts (Frisch *et al*., 1996). Recently, two reports utilizing random mutations and elegant selection strategies have described scFv intrabodies (Proba *et al*., 1998; Martineau *et al*., 1998). Both these intrabodies had a number of sequence changes, but thermodynamic data on individual mutations that would aid in generalizing these results have not been reported. Thus, while soluble intrabodies can be obtained by sophisticated methods of evolutionary engineering, the contribution and relative importance of the factors that determine successful cytoplasmic expression remain to be investigated.

As a basis for an in-depth analysis of factors determining the expression of intrabodies, we have attempted to engineer intrabody domains in a rational way. Motivated by a formal analogy to the concept of an ensemble in statistical mechanics, we have compiled the amino acid distributions in aligned immunoglobulin domain sequences. These reflect the selective pressure on features common to all sequences in the ensemble, like stability and heterodimer interactions. The resulting distributions are expected to be Boltzmann-like with respect to the selected function of the individual residues; we approximate the consensus residue to be the best choice for stability. This canonical sequence approximation has reliably predicted stabilizing point mutations (Steipe *et al*., 1994). The approach does not depend on special features of the target immunoglobulin domain and mutations can be predicted by comparison of target sequences with a freely accessible database the canonical sequence approximation; http://www.lmb.uni-muenchen.de/groups/bs/canonical.htm. We have further generalized the approach to structural motifs, such as β-turns (Ohage *et al*., 1997). Additionally, we have prepared a series of stabilizing mutations, based on the structure of a disulfide-free variant of the myeloma protein REI (Uson *et al*., 1997). While these have highly context-dependent effects (E.O. & B.S., unpublished results) some provide additional stability. This panel of well-characterised V_κ domain stability mutants has now been used as a starting point for an analysis of the factors contributing to the cytoplasmic expression of immunoglobulin domains. Here, we focus on the issues of folding stability of V_κ domains, the domain N terminus, promoter strength and expression temperature. The application of this approach to a soluble, functional intrabody Fv is reported in a companion manuscript (Ohage *et al*., 1999).

Results

Proteins of this study

Proteins investigated in this study are derived from the immunoglobulin variable domain (V_κ) of the light chain of the murine antibody McPC603 (Skerra & Plückthun, 1988) (for an overview see Table 1). The oxidized wild-type V_κ domain of McPC603 has a free energy of folding (ΔG_F^P) of only −13.5 kJ/mol (Steipe *et al*., 1994). Combination of five stabilizing point mutations had previously resulted in domain V_L-500 with a ΔG_F^P of −27.2 kJ/mol (Ohage *et al*., 1997). The engineering of mutations suggested by β-turn propensities had resulted in a further stabilised V_L domain with a ΔG_F^P of −30.2 kJ/mol (Ohage *et al*., 1997). An additional opportunity for stabilization was suggested from an analysis of the length distributions for V_κ CDR-I. Introducing the consensus loop-length and sequence into this domain results in the prototype framework V_L-601 (Table 1). This further enhances stability by 4.1 kJ/mol, to a total ΔG_F^P of −34.3 kJ/mol. The mutations are distributed over the domain and the prediction does not depend on the special context of an individual fold. V_L-601 served as the reference domain for further stabilization through a network of charged residues based on variants of the myeloma protein REI (E.O. & B.S., unpublished results). From the latter study, two variants were used here: mutant V_L-705 contains the point mutations Tyr32His and His92Gln, and mutant V_L-703 contains Tyr32His and Gly50Glu (Table 1). The free energy of folding is −37.9 kJ/mol for V_L-705 and −43.6 kJ/mol for V_L-703. V_L-703 is the most stable V_κ domain

Table 1. Sequences of proteins used in this study

Protein	Sequence
V_L-500[a]	McPC603 V_L+A15L, M21I, F32Y, T63S, N90Q
V_L-601	V_L-500+S56P+CDRI: 12→6[b]
V_L-703	V_L-601+Y32H+G50E
V_L-705	V_L-601+Y32H+H92Q

[a] Stability of oxidized domain reported previously (Ohage *et al*., 1997).

[b] The loop length of CDR1 is shortened from 12 to six residues and the consensus sequence for this loop length is introduced; the sequence given in Methods.

investigated here, its stability being more than three times that of the wild-type domain.

Further, we analyzed an immunoglobulin domain derived from V_L-601 which contains the three CDR-loops of the catalytic antibody 17E4 (Guo *et al.*, 1994). This domain, constructed towards engineering of a synthetic catalytic intrabody, was named IcaL-14 (intracellular catalytic antibody). IcaL-14 forms a dimer with a ΔG of −78.1 kJ/mol in the oxidized and −59.1 kJ/mol in the reduced state (Ohage *et al.*, 1999). These values for ΔG were obtained according to a model of unfolded monomer and native dimer ($2U \rightleftharpoons N_2$) and thus they are not directly comparable to the stabilities of the monomeric domains of the V_L series.

Folding stability of reduced domains

Figure 1 shows unfolding transition curves for V_L-703 in the oxidized and reduced state. In the oxidized, folded state the fluorescence of Trp35 is completely quenched by the spatially adjacent disulfide bridge. As a consequence, the signal rises as the domain unfolds (Figure 1). For the reduced domain, however, fluorescence is more intense in the folded than in the unfolded state as is generally observed for tryptophane residues that are transferred to the polar solvent. Both curves can be well fitted to a two-state reaction using a linear free-energy model. A least-squares fit gives a folding free energy of −24.4 kJ/mol for the reduced protein. The disulfide bridge therefore contributes −19.2 kJ/mol to the stability of domain V_L-703 (Table 2). For domains V_L-601 and V_L-705, values for ΔG_F^P were −18.8 and −21.6 kJ/mol, which amounts to a loss of folding energy of 15.5 and 16.3 kJ/mol, respectively. Since this is almost exactly what theory predicts for the change in entropy of the unfolded state†, this value places a limit on the effects the disulfide bridge may have on the native structure. For these three domains, reduction causes an increase in the cooperativity of the folding reaction (*m*-value, $d(\Delta G_F)/d$[GdmCl]) by 3.9(±0.6) kJ/molM. Increased *m*-values can be correlated with an increase in the accessible surface area (ΔASA) upon unfolding (Schellman, 1978). Thus, breaking the covalent bond between two cysteine residues appears to result in a significant increase in denaturant binding sites of the unfolded state. Since the *m*-value change was the same for V_L-601, V_L-703 and V_L-705, the same increase was assumed in the analysis of the transition curve of domain V_L-500. Here, the curve was difficult to fit because the pre-unfolding baseline is very short. We therefore applied a curve fit with a fixed *m*-value of 17.8 kJ/molM, which is that of the oxidised domain plus 3.9 kJ/molM. Experimental values accomodate this biased curve fit very well (not shown). We obtained a ΔG_F^P of −13.3 kJ/mol for V_L-500 in the reduced state, which translates into a stability loss ($\Delta\Delta G_F^P$) of 13.9 kJ/mol upon reduction. This is in good agreement with the values for $\Delta\Delta G_F^P$obtained in the unbiased curve fits for the other domains.

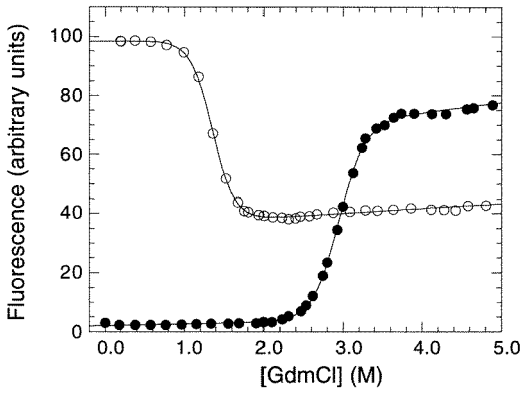

Figure 1. Unfolding transition curves for V_L-703 in the oxidized (filled circles) and reduced (open circles) state at pH 7.4 and 20 °C, protein concentration was 2 μM. λ_{ex} 280(±1.25) nm, λ_{em} 350(±2.5) nm (oxidized) or 323(±2.5) nm (reduced domains). Curves were calculated by non-linear least squares fit following a two-state unfolding model. Note that the disulfide bridge quantitatively quenches tryptophan fluorescence in the native state, while fluorescence increases in the folded state of the reduced domain.

Expression strategies

Since all mutants were at least as stable in the reduced state as the oxidized wild-type protein, the expression in the reducing environment of the cytoplasm appeared feasible, if no further limiting factors were operative. To investigate this, two expression vectors with different promoters were compared. One vector was derived from pASK75 (Skerra, 1994) and contains the moderately strong but very tightly regulated *tet*A-promoter (Ohage *et al.*, 1999). Alternatively, we used the pRSET5d vector, which enables expression under the very strong T7-promoter (Schoepfer, 1993).

For cytoplasmic expression, the periplasmic signal sequence has to be removed. This causes an unexpected problem: translation of any gene must begin with a (formyl-) methionine, but in most native V_κ sequences, as well as in the sequences investigated here, the first residue is an aspartic acid. Large residues in the second position, however, can prevent processing of the N-terminal methionine residue (Hirel *et al.*, 1989). Therefore,

† The contribution of a structural disulfide bridge to protein stability predominantly results from entropic destabilisation of the unfolded state (Goto & Hamaguchi, 1982). For a disulfide bond bridging 65 residues it can be estimated as −8.8 J/K mol −(1.5 *R* ln(65)), which amounts to −17.8 kJ/mol at 20 °C (Pace *et al.*, 1988). This can be more than the entire folding stability of typical, monomeric immunoglobulin variable domains.

Table 2. Free energies of folding and folding cooperativities of oxidized and reduced proteins used in this study

Protein	Reduced		Oxidised		Contribution of SSBr	
	ΔG_F^p (kJ/mol)	m (kJ/mol M)	ΔG_F^p (kJ/mol)	m (kJ/mol M)	$\Delta\Delta G_F^p$ (kJ/mol)	Δm (kJ/mol M)
Wild-type	-	-	−13.5[a]			
V_L-500	−13.3 (0.3)	17.8[b]	−27.2 (0.1)[a]	13.9 (0.1)	−13.9 (0.3)	−3.9[b]
V_L-601	−18.8 (1.1)	17.7 (0.9)	−34.3 (0.6)	13.6 (0.3)	−15.5 (1.3)	−4.1 (0.9)
V_L-705	−21.6 (1.1)	18.9 (0.9)	−37.9 (1.1)	14.6 (0.4)	−16.3 (1.6)	−4.3 (1.0)
V_L-703	−24.4 (0.7)	18.0 (0.5)	−43.6 (1.1)	14.8 (0.4)	−19.2 (1.3)	−3.2 (0.6)
IcaL-14[c]	−59.1 (2.0)	20.0 (1.2)	−78.1 (1.2)	20.5 (0.5)	−19.0 (2.3)	0.5 (1.3)

At pH 7.4 and 20 °C, the protein concentration was 2 μM; the experimental error is given in parentheses (1σ).
[a] Data published previously (Steipe *et al*., 1994; Ohage *et al*., 1997).
[b] For the reduced domain V_L-500, the *m*-value was fixed to 17.8 kJ/mol M.
[c] Stability values obtained from dimer transition curves (Ohage *et al*., 1999).

effects on protein stability caused by an unprocessed N-terminal residue had to be investigated. Three different N termini were constructed. In a ''naïve'' approach, an initiator methionine codon was simply fused to Asp1, thus giving rise to the sequence MetAspIle. Constructs of this kind were called μ-mutants. To provide better substrates for the Met-aminopeptidase, the so-called α and η-mutants were designed. In α-mutants, an alanine is inserted before Asp1, which results in translation starting MetAlaAspIle. In η-mutants, Asp1 was replaced by alanine (MetAlaIle).

Expression into inclusion bodies and refolding

In a first set of experiments, N-terminally modified domains were expressed in *E. coli* BL21(DE3) under the control of the phage T7-promoter. After cell lysis, insoluble inclusion bodies were solubilized under non-reducing, denaturing conditions and subjected to IMAC on Ni^{2+}-columns. After chromatography, protein concentration and oxidation state were determined. Protein yields were in the range of 20 mg per g of bacterial pellet, approximately independent of stability. Only about 50 % of the cysteine residues had free thiol groups. On non-reducing gels, the amount of protein with intermolecular disulfide bonds, forming oligomers, could be estimated to comprise less than 10 % (not shown). After refolding and oxidation of inclusion body material, yields of soluble protein were only 50 to 70 % of the eluted material. Apparently, a significant fraction of domains possess non-native disulfide bonds with small-molecular mass compounds and will not refold unless these are first removed. After *in vitro* reduction of the insoluble material, refolding yields under reducing conditions approached 100 %.

Analysis of N-terminally extended proteins

In order to investigate the fate of the initiator methionine, μ, α and η-variants of IcaL-14 were analyzed by mass spectrometry. The results show that processing of the initiator methionine residue was quantitatively inhibited when methionine was followed by the native N-terminal residue, aspartic acid (Table 3). Processing appears to be tightly coupled to translation, as it takes place before the protein segregates into inclusion bodies.

The effect of the N-terminal modification on folding stability was investigated by measuring transition curves for a number of variants under oxidizing and reducing conditions. Surprisingly, an unprocessed N-terminal methionine destabilizes the domains by approximately 3 kJ/mol (Table 4). This effect is independent of the oxidation state. Stability effects are even more pronounced for domain IcaL-14: here, the unprocessed methionine reduces stability by 10 kJ/mol, both in the reduced and oxidised state (mutant IcaL-μ14). To a large extent, this is caused by a decrease of the folding cooperativity. It should be noted, that such a major loss in stability alone would prevent expression of many immunoglobulin domains.

Table 3. Processing of N-terminal methionine residue

Protein	N terminus	M_r (Da)		Processed?
		Theoretical	Experimental	
IcaL-μ14	(M)DIEL...	+Met: 12,396.0	12,397.0	No
		−Met: 12,264.8		
IcaL-α14	(M)ADIEL...	+Met: 12,467.1		
		−Met: 12,335.9	12,335.0	Yes
IcaL-η14	(M)AIEL...	+Met: 12,352.0		
		−Met: 12,220.8	12,221.0	Yes

Proteins were cytoplasmically expressed into inclusion bodies, refolded and analyzed on a mass spectrometer. The N-terminal methionine residue is displayed in parentheses.

Table 4. Folding stabilities and transition cooperativities for N-terminally modified domains

Protein	N Terminus	Reduced			Oxidized		
		ΔG_F^P (kJ/mol)	m (kJ/mol M)	$\Delta\Delta G_F^P$ (kJ/mol)	ΔG_F^P (kJ/mol)	m (kJ/mol M)	$\Delta\Delta G_F^P$ (kJ/mol)
IcaL-μ14[b]	MDIEL…	−49.0 (2.7)	13.2 (1.5)	10.1 (3.4)	−69.8 (1.4)	17.6 (0.6)	8.3 (1.8)
IcaL-α14[b]	ADIEL…	−62.0 (2.8)	22.1 (1.8)	−2.9 (3.4)	−70.7 (0.9)	18.3 (0.4)	7.4 (1.5)
IcaL-η14[b]	AIEL…	−63.3 (3.7)	23.0 (2.0)	−4.2 (4.2)	−74.2 (1.1)	19.1 (0.5)	3.9 (1.6)
V_L-μ500	MDIVM…	−12.4 (0.3)	15.7[d]	0.9 (0.4)	−25.2 (0.7)	13.3 (0.4)	2.0 (0.7)
V_L-μ601	MDIVM…	−16.4 (1.0)	14.9 (0.4)	2.4 (1.5)			
V_L-μ703	MDIVM…	−20.3 (0.4)	16.4 (0.3)	4.1 (0.8)			
V_L-η703	AIVM…	−22.6 (0.5)	16.9 (0.4)	1.8 (0.9)			
V_L-μ705	MDIVM…	−18.5 (0.9)	17.0 (0.7)	3.1 (1.4)			

At pH 7.4 and 20 °C, the protein concentration was 2 μM; the experimental error (1σ) is given in parentheses.

[a] Stability effects resulting from the respective modifications at the N terminus ($\Delta\Delta G = \Delta G_{modified} - \Delta G_{unmodified}$). Values for $\Delta G_{unmodified}$ are from Table 2.

[b] Stability values obtained from dimer transition curves (Ohage *et al.*, 1999).

[c] The *m*-value for the transition of the reduced mutant V_L-μ500 was fixed to 15.7 (=17.8 − 2.1) kJ/mol M: 17.8 is the value for domain V_L-500, −2.1 is the average change after modification of the N terminus for mutants V_L-μ601, V_L-μ703 and V_L-μ705.

If Asp1 is mutated to an alanine residue, the N-terminal methionine is correctly processed. This results in a significantly reduced destabilization (V_L-η703 *versus* V_L-703) or even stabilization (IcaL-η14 *versus* IcaL-14) under reducing conditions (Table 4). Oxidized IcaL-η14 was modestly destabilized by this mutation. If Asp1 is preceded by an alanine residue (IcaL-α14), a similar behavior can be observed but the effect of the mutation on the oxidised state is more pronounced.

We conclude that an unprocessed N-terminal methionine can cause unpredictable stability problems. For the proteins investigated in this study this problem can be solved by modifying the N terminus to include a residue that is a good substrate for the methionyl-aminopeptidase: the mutation of Asp1 to alanine was consistently neutral or even slightly stabilizing under reducing conditions.

Intracellular expression of soluble immunoglobulin domains

Unless stated otherwise, experiments to quantify cytoplasmic expression were performed with domains of the μ-series under control of the *tet*A-promoter at 20 °C. After harvest, cells were lysed mechanically under reducing conditions and the soluble fraction was subjected to IMAC. Protein yields were determined by the A_{280} of the eluate. The expression yields of soluble immunoglobulin domains are strongly correlated to their folding stability: While the V_κ wild-type domain of McPC603 cannot be detected in the soluble fraction of the cell extract, yields increased up to 3.5 mg/g bacterial pellet for the most stable variant V_L-μ703 (Figure 2(a)) which corresponds to about 15 mg/l in a shaker flask, after three hours of induction at 20 °C ($A_{600} = 1.7$). The minimal folding stability for the expression of soluble immunoglobulin domains is about −12 kJ/mol under reducing conditions (Figure 2(b)). Considering the loss of the disulfide bond and the modification of the N terminus, this requires a ΔG_F^P of at least −27 kJ/mol in oxidized domains, which is approximately a factor two above the stability for the McPC603 V_κ domain.

The stability effects of the modified N termini are also reflected in the expression yields. For domain IcaL-14, the unstable μ-variant can hardly be detected, while the more stable domain IcaL-η14 is obtained in 2.6(±0.4) mg/l (Figure 3(a)). In contrast, expression yields for domains V_L-μ703 and V_L-η703 are similar, in good correlation with their comparable thermodynamic stabilities (Table 4).

The temperature during expression has a marked effect on the yields of soluble material. The optimum temperature appears to be 20 °C (Figure 3(b)). Western blots of the insoluble fractions suggest that aggregation becomes a limiting factor at higher temperatures (not shown).

The effect of promoter strength on the yield of soluble material is small: under control of the stronger phage T7-promoter the amount of soluble material was similar to that detected in the *tet*-system (not shown). However, aggregated material accumulates to high levels (20 mg/g, see above) relative to expression under the *tet*A-promoter (see below).

Fate of unstable proteins

The question arises whether unstable proteins are lost by degradation or aggregation of soluble proteins, or by the abortion of productive folding due to aggregation of newly formed chains. To investigate this, insoluble material was resuspended in 10 M urea after cell lysis and aliquots were compared with soluble fractions on a Western blot. For detection, an anti-His-tag scFv-fragment was used (Lindner *et al.*, 1997), rendering the signal independent of mutated residues within domains. Although this method is only semi-quantitative, a clear increase of the amount of soluble protein with stability could be detected in the soluble fraction (Figure 2(c)). In contrast, the amount of protein in the insoluble fraction appeared to be fairly independent of stability. Aggregates accumu-

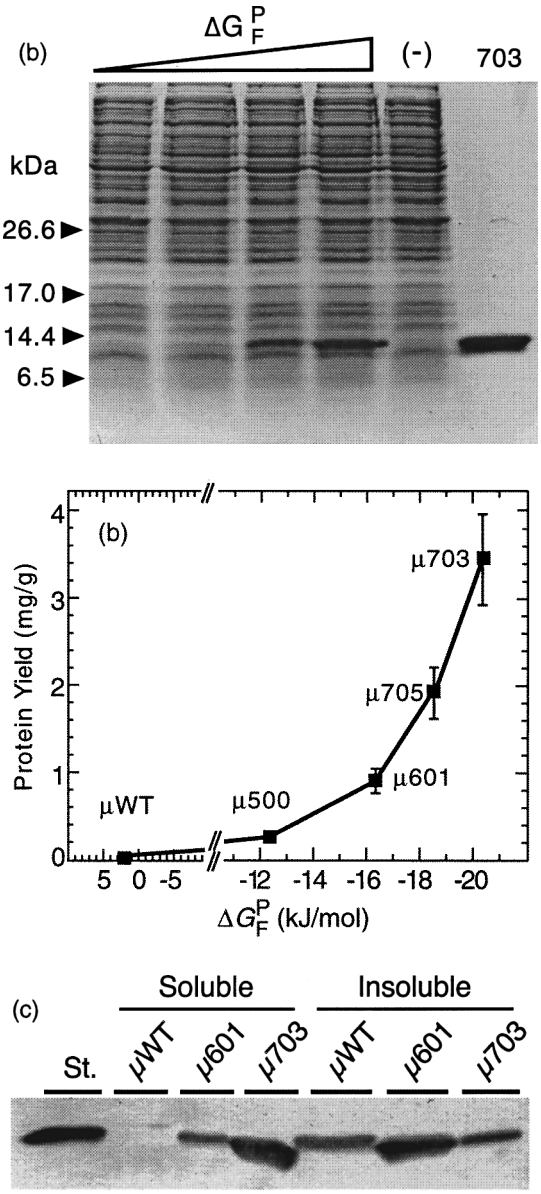

Figure 2. Expression yield of immunoglobulin domains in the cytoplasm. (a) Soluble whole-cell protein of *E. coli* JM109 cells. V_κ domains with increasing folding stabilities were expressed under the *tet*A-promoter at 20 °C. From left to right: domain V_L-μWT, V_L-μ601, V_L-μ705, V_L-μ703, uninduced negative-control, purified domain V_L-μ703. Reducing SDS 17 %-PAGE, Coomassie-stained. (b) The yield of soluble protein determined after IMAC is given in mg per g of bacterial wet mass. Folding stabilities are given for the reduced, N-terminally modified domains (Table 4). The stability for the V_L-μWT domain under these conditions was estimated. (c) Western-blot of the soluble and insoluble cytoplasmic fractions for three domains from (a). One slot contains the equivalent of 100 μl culture volume (0.4 mg cells), the standard (left-most lane) was 2 μg purified domain V_L-μ601. Proteins were detected using an anti-His-tag-scFv fused to alkaline phosphatase (Lindner *et al.*, 1997).

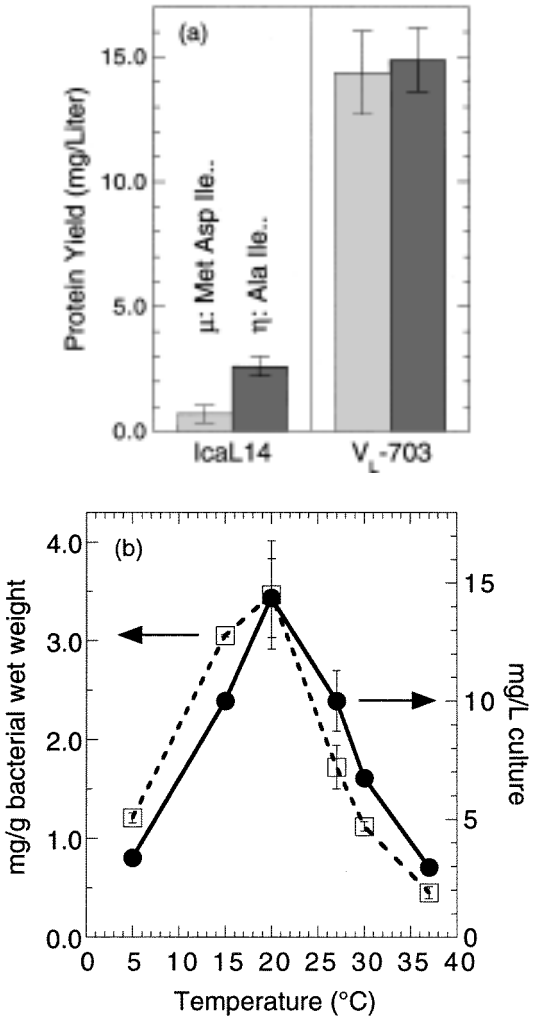

Figure 3. (a) Influence of N-terminal residue on the yield of soluble protein after expression in *E. coli* JM109 under control of the *tet*A-promoter. The μ-variant is shown in light, the η-variant in dark grey bars. (b) Effect of the expression temperature on the yield of soluble domain V_L-μ703 (*tet*A-promoter). Results are normalized to bacterial mass (arrow pointing to left axis, open squares) or culture volume (arrow pointing to right axis, filled circles).

lated to a level of about 1-2 mg/g irrespective of folding stability of the mutant and irrespective of the total amount present in the cell. We conclude that inclusion body formation is a side-reaction of nascent protein chains, it changes little with constant translation rates and is independent of thermodynamic stability (although it may well depend on folding rates), while the variations in the yield of soluble material are mainly due to degradation.

Oxidation state of soluble material

To investigate the oxidation state of the soluble intracellular material, domain V_L-η703 was analyzed in more detail. After harvest, cells were disrupted and chromatography of the soluble fraction was performed without additional reduction. The eluate was then diluted with phosphate-buffered

saline (PBS) and probed with DTNB for free thiols. The amount of thiols that reacted with DTNB was very small and hardly changed with time (Figure 4(a)). However, addition of GdmCl to a final concentration of 3.2 M resulted in a biphasic exponential increase of the signal. Owing to the deadtime of mixing, the fast phase was not resolved. Once equilibrium was reached, the absorption difference between native and denaturing buffer corresponded to 95% free thiols. Thus the cysteine residues remain reduced after this very stable mutant has folded to the native state in the cytoplasm. Subsequently, the cysteine residues, located in the core of the folded domain, are protected from oxidation during chromatography and from reaction with DTNB. This emphasizes the fact that proximity of thiols alone is not sufficient for oxidation, but a reaction with an oxidant is required.

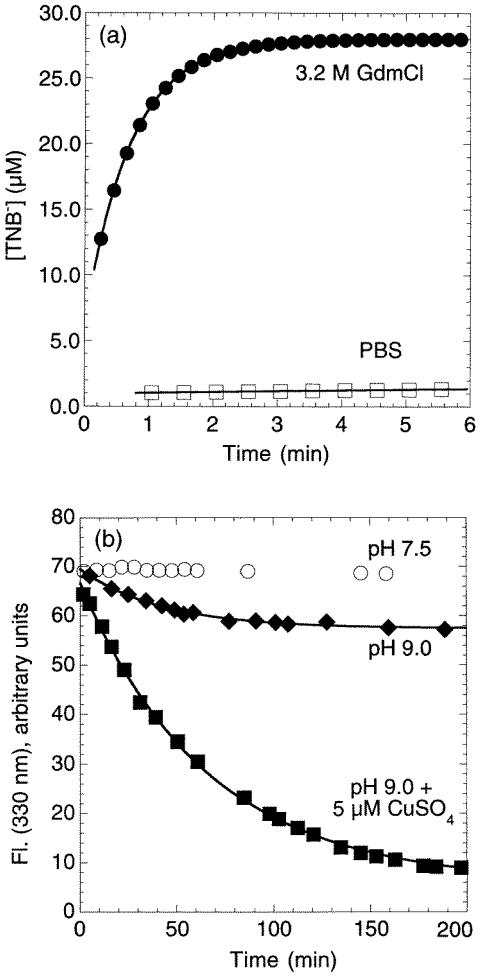

Figure 4. (a) Oxidation state of the two cysteine residues in domain V_L-η703 after cytoplasmic expression in *E. coli* JM109. Cell lysis and protein purification *via* IMAC was performed without additional reducing agents. The purified protein (final concentration 14.6 μM) was mixed with DTNB under native (PBS, open squares) or denaturing conditions (PBS + 3.2 M GdmCl, filled circles) and the absorbance at 412 nm was recorded. From the absorbance signal the concentration of the thio-nitrobenzoate anion (TNB^-) was calculated. The curves are monoexponential fits. (b) Air oxidation of the free thiols in domain V_L-η703. Directly after chromatography, aliquots of the purified protein (final concentration 5 μM) were incubated in 0.1 M Tris buffers at the indicated pH values. In one case, $CuSO_4$ was added. The fluorescence of the samples was recorded at 320 nm after exciation at 280 nm. The curves are monoexponential fits.

In vitro air oxidation of V_K domains

Since formation of the intradomain disulfide bond is accompanied by a marked decrease of the fluorescence signal (Figure 1), the oxidation process can be analyzed *in situ* by measuring the fluorescence. After chromatography under non-reducing conditions, mutant V_L-η703 was diluted into non-degassed neutral or slightly alkaline buffers (Figure 4(b)). While at neutral pH the fluorescence signal remained constant, it slowly decreased at pH 9.0. Catalysis by 5 μM $CuSO_4$ further accelerated oxidation, so that after 20 hours no remaining free thiols could be detected. Thus, cysteine residues within the domain are unreactive, as long as the pH remains neutral. Alkaline conditions that favor the thiolate anion through which the reaction proceeds, and catalytic amounts of metal ions enhance the rate of formation of the disulfide bond.

Discussion

The isolated, monomeric V_κ domain has proven to be an excellent model system to investigate factors influencing the soluble expression of intrabodies. In contrast to V_H domains it folds reversibly (Wirtz & Steipe, 1999); in contrast to Fv or scFv fragments it follows a two-state model of folding. Thus thermodynamic stability can be accurately determined. The advantages of using consensus frameworks for immunoglobulin engineering have been discussed (Steipe *et al.*, 1994; Ohage *et al.*, 1999). Individual consensus mutations can be easily predicted by comparing target sequences with a freely accessible database. Since we have consistently observed the effects of consensus sequence mutations to be independent and additive to a good approximation and the predictions require no special features on the part of the target domain, the approach is expected to be completely general.

With these intrabody V_L domains that differ in their stability in a well-characterized way, we have been able to probe the factors influencing and possibly limiting the expression of intrabodies. A simple model that summarises these results is presented in Figure 5.

Transcription rates that are increased beyond those of a moderately strong promoter do not seem to be advantageous. The observation that the amount of folded material is approximately constant for a given stability raises the possibility that

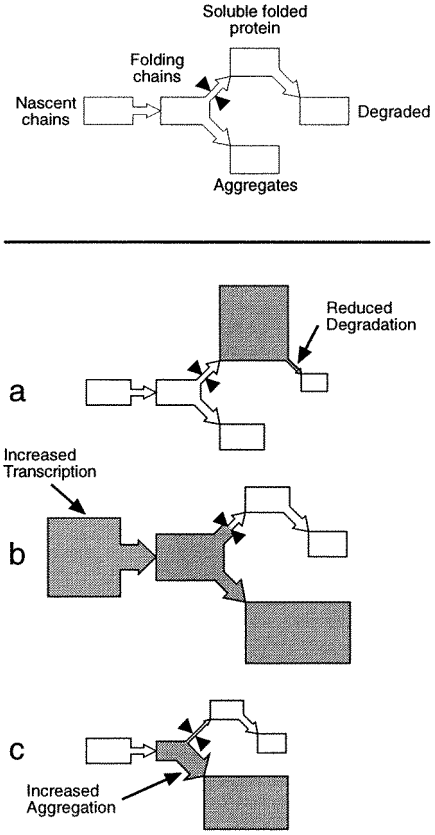

Figure 5. A simple model for the pools of protein material in the expression, folding and cytoplasmic survival of intrabody domains that accomodates our results. The size of the rectangles symbolizes the amount of protein in a given state, the diameter of the arrows symbolises the rates of interconversion between states. The amount of soluble folded protein is determined by an equilibrium of newly folded chains and intracellular degradation. It appears that the rate of folding is limited by an external factor (black triangles). (a) Increased stability reduces intracellular degradation, this raises the amount of soluble folded material. (b) An increased transcription rate leads to increased aggregation as the transition to the folded state becomes rate-limiting. (c) Increased temperature leads to increased aggregation, a reduced amount of chains that go into productive folding and a reduced amount of soluble folded protein.

extrinsic factors, possibly chaperones, may be rate-limiting after translation. For those proteins that do not go into inclusion bodies, there is an excellent correlation of stability and yield of folded material. As the soluble material does not appear to be lost by aggregation, this suggests that degradation of unstable, soluble domains limits expression yields in this case. The observed independence of aggregation and thermodynamic stability is in contrast to results reported previously for the periplasmic space (Chan *et al*., 1996). This suggests that distinct mechanisms may be operative in the formation of inclusion bodies in the cytoplasm and in the formation of insoluble aggregates in the periplasm.

Surprisingly, stability can be significantly compromised by an unprocessed N-terminal initiator methionine coden. A similar situation was observed in α-lactalbumin, where an N-terminal methionine resulted in a stability loss of almost 8 kJ/mol (Ishikawa *et al*., 1998). This is a general problem, since variable immunglobulin domains of any class usually begin with an Asp, Glu or Gln (Kabat *et al*., 1992), residues that are poor substrates for the methionyl-aminopeptidase (Hirel *et al*., 1989). Fortunately, rational engineering of the N terminus can alleviate this problem. Since the N terminus is a potential source of problems for expression and stability, engineering of the N terminus may guide other intrabody design projects.

Stable immunoglobulin domains can comprise a significant fraction of soluble cellular protein. The minimum stability from which expression appears feasible is about −27 kJ/mol in the oxidized state. This is in contrast to a panel of disulfide-free mutants of the human myeloma V_κ domain REI; only two variants could be expressed in reasonable yield, although all domains had a ΔG of ~ -30 kJ/mol and beyond (Frisch *et al*., 1996). The reason for this may be that the REI variants had Cys23 replaced by a Val residue, expression was attempted in the periplasm and an N-terminal fusion to β-lactamase was required (Frisch *et al*., 1994).

The observed expression yields of several milligrams per liter of shake-flask cultures, without optimization, are already high. Up-scale into high-density fermentation protocols should be routine. For those cases where overall expression yields are a concern, we have shown that our stabilized domains can be obtained at very high yields as insoluble material after transcription from a T7-promoter, subsequently refolded quantitatively under reducing conditions and oxidized *in vitro*. This approach may also be advantageous in situations where fusion proteins, such as immunotoxins (Pai *et al*., 1996), are prone to form non-native disulfides during *in vitro* folding. To further improve expression yields, one option is the introduction of specific residues that have been reported to improve folding properties (Knappik & Plückthun, 1995). In this respect, consensus frameworks can at least provide baseline stability to allow optimisation with methods of evolutionary biotechnology.

In summary, the successful expression of a whole panel of soluble, isolated V_L intrabody domains suggests a plannable route to designed intrabodies.

Methods

Sequences of proteins in this study

The sequence of domain V_L-601 is ^{1}DIVMTQSPSS ^{11}LSVSLGERVT ^{21}ISCRASQSVS ^{31}NYLAWYQQKP ^{41}GQ PPKLLIYG ^{51}ASTREPGVPD ^{61}RFSGSGSGTD ^{71}FTLTISS VQA ^{81}EDLAVYYCQQ ^{91}DHSYPLTFGA ^{101}GKLELKHH

H ^{111}HH (numbering according to Kabat *et al.*, 1992). See Ohage *et al.*, 1999 for the sequence of IcaL-14. Oligonucleotide directed mutagenesis was performed by standard methods using single-strand DNA prepared from an f1 origin of replication in the vector. N-terminal modifications were introduced into the domains by PCR using standard protocols. All modified genes were sequenced.

Expression and purification of disulfide-intact domains

For expression of oxidized immunoglobulin domains in *E. coli*, genes of interest were fused to the signal peptide sequence of the outer membrane protein A (ompA) and secreted into the bacterial periplasm. Due to the oxidizing properties of this compartment, the disulfide bridge can form and the domains fold spontaneously into their native conformation (Skerra & Plückthun, 1988). Domains were expressed in *E. coli* JM83 and purified *via* immobilized metal ion affinity chromatography as described (Ohage *et al.*, 1997).

Cytoplasmic expression

For cytoplasmic expression, genes were N-terminally modified by PCR to yield the amino acid sequences MetAspIle (μ), MetAlaAspIle (α) or MetAlaIle (η-domains). The respective genes were cloned into expression vectors derived from pASK75 (transcription driven by the *tet*A-promoter) or into pRSET5d (transcription driven by the T7-promoter) (Schoepfer, 1993). For expression into inclusion bodies, cultures of *E. coli* BL21(DE3) harboring the respective plasmid were grown to an A_{600} of 1 and induced with 0.5 mM IPTG. The cultures were incubated for three hours at 37 °C, harvested and lysed in a French press. After centrifugation, the insoluble material was resuspended in sonication buffer (300 mM NaCl, 50 mM Tris-HCl (pH 8.0)) and washed with IB-buffer I (50 mM EDTA, 10 mM Tris-HCl (pH 8.0), 1% (v/v) Triton-X-100), IB-buffer II (500 mM GdmCl, 100 mM Tris-HCl (pH 7.5), 20 mM EDTA) and twice with water. The pellet was then solubilized in DC-buffer (6 M GdmCl, 50 mM $NaPO_4$ (pH 8.0)) and subjected to an equilibrated nickel-NTA-column (Qiagen, Hilden). After washing with DC buffer containing 30 mM imidazole, proteins were eluted with DC buffer containing 200 mM imidazole.

To quantify the amount of soluble V_L domains after cytoplasmic expression, *E. coli* JM109 or *E. coli* BL21(DE3) containing the respective plasmid were grown to an A_{600} of 1 and induced with 430 nM anhydro-tetracycline or 0.5 mM IPTG, respectively. Cells were incubated at the desired temperature for three hours, harvested, and mechanically lysed with glass beads under reducing conditions (1 M NaCl, 50 mM $NaPO_4$ (pH 8.0), 10 mM imidazole, 2 mM β-mercaptoethanol, (β-ME)) to prevent formation of intermolecular disulfide bonds. After centrifugation, the soluble fraction was isolated. For the following steps, β-ME had to be removed as it jeopardizes precise absorbance measurements. However, since the absence of reducing agents could result in aggregation of formerly soluble material due to formation of intermolecular disulfide bonds during workup, urea was added to the soluble fraction to a final concentration of 8 M. The proteins were then purified on a nickel-NTA-column as described above, except that buffers contained 8 M urea and 1 M NaCl and no GdmCl or β-ME. To quantify the protein yield, the A_{280} of the eluate was measured and normalised to the A_{600} before harvest. For *E. coli* JM83 the bacterial mass was assumed to be 2.5 g/l A_{600}, and for *E. coli* BL21(DE3) 1.7 g/l A_{600}. The molar extinction coefficients ε_{280} for oxidized proteins were 13,490 $M^{-1}cm^{-1}$ for domains V_L-500 and V_L-601, 12,210 $M^{-1}cm^{-1}$ for domains V_L-703 and V_L-705, and 14,770 $M^{-1}cm^{-1}$ for domain IcaL-14 (Gill & von Hippel, 1989). Molar extinction coefficients for N-terminally modified proteins are identical, those for reduced domains can be obtained by subtracting 120 $M^{-1}cm^{-1}$.

To probe the oxidation state of the cytoplasmically expressed material, harvested cells were lysed in a French press and IMAC was performed under non-reducing, non-denaturing buffer conditions (1 M NaCl, 50 mM $NaPO_4$ (pH 7.4)).

Refolding and oxidation

Material purified from inclusion bodies was refolded by dilution or dialysis at room temperature to final concentrations of 1 M GdmCl, 50 Tris-HCl (pH 9) and 5 μM protein. After incubation overnight, 5 μM $CuSO_4$ was added and incubated for another six hours. The solution was concentrated, centrifuged and dialyzed against PBS (4 mM KH_2PO_4, 16 mM Na_2HPO_4, 115 mM NaCl, (pH 7.4)) containing 20 mM EDTA. The yield of soluble protein was determined by its A_{280}.

Fluorescence spectroscopy

Fluorescence spectroscopy of oxidised domains was performed as described previously (Ohage *et al.*, 1997). For reduced domains, samples were unfolded in 4 M GdmCl in the presence of 50 mM DTT to ensure complete reduction. Refolding was performed by dilution into PBS which contained 3 mM DTT. In the native, disulfide-intact state, tryptophane emission is completely quenched by the disulfide bond, and only a tyrosine signal at 305 nm can be detected. Reduction of the domain results in an intense tryptophane emission with a maximum at 320 nm. Spectra of the unfolded oxidized and reduced state are identical within experimental error when the intensity of the exciting beam, I, is corrected for light absorption by DTT ($I = I_0\ 10^{-A}$), (data not shown). The folding stability was determined by reversible chemical denaturation with GdmCl as described (Ohage *et al.*, 1997, 1999).

Determination of oxidation state with DTNB

The oxidation state of domains was probed by measuring the A_{412} of the sample after addition of DTNB to a final concentration of 15 mM in the respective chromatography buffers. For the thionitrobenzoate anion a molar extinction coefficient (ε_{412}) of 14.000 $M^{-1}cm^{-1}$ was assumed.

Western blot

Samples were subjected to SDS-17% PAGE and transferred onto a PVDF membrane. Proteins containing a His-tag were detected as described (Lindner *et al.*, 1997).

Acknowledgements

Anti-his-tag antibody was a kind gift of Dr Peter Lindner. Monika Walter analyzed the effects of CDRI loop length changes for her diploma thesis in the laboratory. We thank Dr Thomas Ruppert for assistance in mass spectroscopy measurements. Thanks are also due to Jan Barnikow and Peter Wirtz for material and valuable discussions and to Heike Bruhn and Monika Walter for critical reading of the manuscript. Parts of this research were funded by DFG grant Ste 563/3-3.

References

Better, M. & Horwitz, A. (1988). *Escherichia coli* secretion of an active chimeric antibody fragment. *Science,* **240**, 1041-1043.

Chan, W., Helms, L. R., Brooks, I., Lee, G., Ngola, S., McNulty, D., Maleeff, B., Hensley, P. & Wetzel, R. (1996). Mutational effects on inclusion body formation in the periplasmic expression of the immunoglobulin VL domain REI. *Fold. Design,* **1**, 77-89.

Frisch, C., Kolmar, H. & Fritz, H. (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.

Frisch, C., Kolmar, H., Schmidt, A., Kleemann, G., Reinhardt, A., Pohl, E., Usón, I., Schneider, T. R. & Fritz, H.-J. (1996). Contribution of the intramolecular disulfide bridge to the folding stability of REI_V, the variable domain of a human immunoglobulin κ light chain. *Fold. Design*, **1**, 431-440.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.

Glockshuber, R., Schmidt, T. & Plückthun, A. (1992). The disulfide bonds in antibody variable domains: effects on stability, folding *in vitro*, and functional expression in *Escherichia coli*. *Biochemistry*, **31**, 1270-1279.

Goto, Y. & Hamaguchi, K. (1982). Unfolding and refolding of the reduced constant fragment of the immunoglobulin light chain. kinetic role of the intrachain disulfide bond. *J. Mol. Biol.* **156**, 911-926.

Guo, J., Huang, W. & Scanlan, T. S. (1994). Kinetic and mechanistic characterization of an efficient hydrolytic antibody: Evidence for the formation of an acyl intermediate. *J. Am. Chem. Soc.* **116**, 6062-6069.

Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G. & Blanquet, S. (1989). Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl Acad. Sci. USA,* **86**, 8247-8251.

Hoogenboom, H. R., de, B. A., Hufton, S. E., Hoet, R. M., Arends, J. W. & Roovers, R. C. (1998). Antibody phage display technology and its applications. *Immunotechnology,* **4**, 1-20.

Hudson, P. J. (1998). Recombinant antibody fragments. *Curr. Opin. Biotechnol.* **9**, 395-402.

Ishikawa, N., Chiba, T., Chen, L. T., Shimizu, A., Ikeguchi, M. & Sugai, S. (1998). Remarkable destabilization of recombinant alpha-lactalbumin by an extraneous N-terminal methionyl residue. *Protein Eng.* **11**, 333-335.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. (1992). *Distribution Files of the Fifth Edition of Sequences of Proteins of Immunological Interest*.

Knappik, A. & Plückthun, A. (1995). Engineered turns of a recombinant antibody improve its *in vivo* folding. *Protein Eng.* **8**, 81-89.

Lindner, P., Bauer, K., Krebber, A., Nieba, L., Kremmer, E., Krebber, C., Honegger, A., Klinger, B., Mocikat, R. & Plückthun, A. (1997). Specific detection of his-tagged proteins with recombinant anti-His tag scFv-phosphatase or scFv-phage fusions. *Biotechniques,* **22**, 140-149.

Martineau, P., Jones, P. & Winter, G. (1998). Expression of an antibody fragment at high levels in the bacterial cytoplasm. *J. Mol. Biol.* **280**, 117-127.

Ohage, E. C., Graml, W., Walter, M. M., Steinbacher, S. & Steipe, B. (1997). Beta-turn propensities as paradigms for the analysis of structural motifs to engineer protein stability. *Protein Sci.* **6**, 233-241.

Ohage, E. C., Wirtz, P., Barnikow, J. & Steipe, B. (1999). Intrabody construction and expression II: A synthetic catalytic Fv. *J. Mol. Biol.*

Pace, C. N., Grimsley, G. R., Thomson, J. A. & Barnett, B. J. (1988). Conformational stability and activity of ribonuclease T_1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem.* **263**, 11820-11825.

Padlan, E. A. (1996). X-Ray chrystallography of antibodies. *Advan. Protein Chem.* **49**, 57-133.

Pai, L. H., Wittes, R., Setser, A., Willingham, M. C. & Pastan, I. (1996). Treatment of advanced solid tumors with immunotoxin LMB-1: an antibody linked to *Pseudomonas* exotoxin. *Nature Med.* **2**, 350-353.

Proba, K., Worn, A., Honegger, A. & Plückthun, A. (1998). Antibody scFv fragments without disulfide bonds made by molecular evolution. *J. Mol. Biol.* **275**, 245-253.

Schellman, J. A. (1978). Solvent denaturation. *Biopolymers,* **17**, 1305-1322.

Schoepfer, R. (1993). The pRSET family of T7 promoter expression vectors for *Escherichia coli*. *Gene,* **124**, 83-85.

Skerra, A. (1994). Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene*, **151**, 131-135.

Skerra, A. & Plückthun, A. (1988). Assembly of a functional immunoglobulin F_v fragment in *Escherichia coli*. *Science,* **240**, 1038-1041.

Steipe, B., Schiller, B., Plückthun, A. & Steinbacher, S. (1994). Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* **240**, 188-192.

Uson, I., Bes, M., Sheldrick, G., Schneider, T., Hartsch, T. & Fritz, H. (1997). X-ray crystallography reveals stringent conservation of protein fold after removal of the only disulfide bridge from a stabilized immunoglobulin variable domain. *Fold. Design*, **2**, 357-361.

Wirtz, P. & Steipe, B. (1999). Intrabody construction and expression III: Engineering hyperstable VH domains. *Submitted,.*

Edited by R. Huber

(Received 16 April 1999; received in revised form 2 July 1999; accepted 2 July 1999)