J. Mol. Biol. (1994) **242**, 655–669

Correctly Folded T-cell Receptor Fragments in the Periplasm of *Escherichia coli*

Influence of Folding Catalysts

Christoph Wülfing and Andreas Plückthun

Max-Planck-Institut für Biochemie
Protein Engineering Group, Am Klopferspitz, 82152 Martinsried, Germany

The T-cell receptor is the central recognition molecule in cellular immunity. Its extracellular domains are homologous with and thought to be structurally similar to an antibody Fab fragment. Despite the biological importance of the TCR and the ease of bacterial expression of antibody fragments, there are only few reports of TCR-fragment expression in *E. coli*. In order to understand the difficulties of expressing correctly folded TCR fragments in *E. coli*, we have characterized the expression behavior of single-chain Fv analogs of three different TCRs (scTCR). All of them can be folded into the correct conformation in the periplasm of *E. coli*, yet the extent of correct folding varies greatly. In order to overcome the folding problems of some of the scTCRs, we have developed a system with enhanced *in vivo* folding capability based on the simultaneous induction of the heat-shock response and over-expression of the *E. coli* disulfide isomerase DsbA at low temperature. We present a model describing the folding of the scTCRs in the periplasm of *E. coli* and possible points of folding assistance. The role of the periplasm as an independent folding compartment is emphasized and the existence of a general periplasmic chaperone is postulated. We have also shown that a bivalent scTCR, dimerized *in vivo* with helix-turn-helix modules, can be expressed in a correctly folded form.

Keywords: T-cell receptor; protein secretion; protein folding; molecular chaperones; heat shock

1. Introduction

The T-cell receptor controls the cellular immune response. Even though it does not recognize peptide antigens directly, but only in the context of the major histocompatibility complex (Zinkernagel & Doherty, 1979), the extracellular domains of the TCR† are thought to be structurally similar to antibody Fab fragments (Chothia *et al.*, 1988). Despite the immense importance of the TCR and the ease of antibody fragment expression in the periplasm of *Escherichia coli* (Plückthun, 1992), so far only one TCR has been reported to be expressible in soluble form in the periplasm of *E. coli* (Ward, 1991, 1992). Other groups have described the production of TCR fragments as inclusion bodies in *E. coli* (Novotny *et al.*, 1991; Ganju *et al.*, 1992; SooHoo *et al.*, 1992; Kurucz *et al.*, 1993). In order to make full use of *E. coli* technologies, especially of the *in vivo* screening methods using filamentous phages (McCafferty *et al.*, 1990), it is desirable to find a general way to express TCRs functionally in the periplasm of *E. coli*.

It is generally assumed that the entire information for the correct fold of a protein is encoded in its amino acid sequence (Anfinsen, 1973; Matthews, 1993). Protein folding pathways have been described for a number of proteins *in vitro*, e.g. for BPTI (Creighton, 1988; Kim & Baldwin, 1990; Weissman & Kim, 1991), hen egg-white lysozyme (Radford *et al.*, 1992), or for barnase (Horowitz & Fersht, 1992) and *in vivo* e.g. for the bacteriophage P22 tailspike protein (Mitraki & King, 1992). From such studies, it emerged that the rate-limiting step in protein folding is late in the pathway, close to the final tertiary (and quaternary) structure (Jaenicke, 1991; Matthews, 1993). In simplified terms, folding can often be described as a first-order process with one rate-limiting step. First-order folding then competes with higher-order diffusion-controlled aggregation (Rudolph, 1990; Goldberg *et al.*, 1991). Therefore, not the final stability of the correctly folded protein,

Correspondence: A. Plückthun, Biochemischer Institut der Universität Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland.

Present addresses: C. Wülfing, Beckman Center, Stanford University, Stanford, CA 94305-5428, U.S.A.; A. Plückthun, Biochemisches Institut der Universität Zürich, Winterthurerstr. 190, 8057 Zürich, Switzerland.

† Abbreviations used: TCR, T-cell receptor; scTCR, single-chain Fv analog of TCR.

0022–2836/94/400655–15 $08.00/0

which is often in the range of 50 kJ/mol (Pace, 1991), but the rate of folding as opposed to the rate of aggregation is the major determinant of the extent of correct folding. Misfolded protein aggregates are often efficiently removed *in vivo*, notably in eukaryotes, thereby creating the misleading impression of folding yields of 100% (Hurtley & Helenius, 1989; Pelham, 1989).

Even though the information for the final fold of a protein is encoded in its amino acid sequence, a large set of proteins is known to be involved in the *in vivo* folding process (Gething & Sambrook, 1992). Some proteins seem to catalyze specific steps: disulfide bond formation in the periplasm of *E. coli* is catalyzed by DsbA (Bardwell *et al.*, 1991; Kamitani *et al.*, 1992). In contrast, molecular chaperones (Craig, 1993; Agard, 1993; Hendrick & Hartl, 1993) seem to help this process by preventing aggregation as the competing reaction of protein folding (Buchner *et al.*, 1991), and they do so by stabilizing an unfolded state of their substrate (Zahn *et al.*, 1994). For many (but not all) substrate proteins, release from the chaperones is an ATP-dependent process (Martin *et al.*, 1991). A model was proposed that in mitochondria or in the cytoplasm of *E. coli* the protein to be folded is handed from one chaperone to the other (Manning-Krieg *et al.*, 1991; Langer *et al.*, 1992). Folding of at least some proteins therefore seems to be an extensively guided, energy consuming process.

The periplasm of *E. coli* is of particular interest for the heterologous expression of recombinant proteins. Because of its non-reducing environment, disulfide bridges can be formed there. This was used for the functional expression of a wide variety of recombinant proteins, such as antibody fragments (Plückthun, 1992). These are of special interest for the present study, being the proteins most similar to T-cell-receptors expressed in the periplasm so far. Despite the considerable importance of the functional expression of proteins in the periplasm, this process has not received much attention to date. Apart from the catalysts of disulfide bond formation (Bardwell *et al.*, 1993; Knappik *et al.*, 1993) and the suspected catalysts of the prolyl-peptidyl *cis–trans* isomerisation (Compton *et al.*, 1992) no periplasmic proteins involved in general protein folding are known. Since the existence of ATP in the periplasm is very questionable (Rosen, 1987), there are doubts as to whether there can be any ATP-dependent periplasmic chaperones. Folding in the physiological context of the periplasm has recently been reviewed more extensively (Wülfing & Plückthun, 1994).

In this paper we report that the expression of scTCR fragments is essentially governed by their folding capability under the specific conditions of the periplasm of *E. coli*. We show that the fraction of the TCRs that is soluble in the periplasm is correctly folded. We present a model that rationalizes the fate of the different TCRs in the periplasm of *E. coli* in the presence or absence of simultaneously over-expressed proteins that are involved in protein folding. We furthermore present indirect evidence for the existence of an extensive, heat-shock controlled folding machinery in the periplasm of *E. coli* and we use this machinery to increase the folding yields of one of our TCRs. Finally we present evidence of correctly folded dimeric TCRs that show enhanced binding to anti-TCR antibody affinity columns.

2. Materials and Methods

(a) *Vector constructions*

DNA manipulations were performed following Sambrook *et al.* (1989). The so-called standard expression system is based on one of our previously described vectors pASK30 (Skerra *et al.*, 1991) using the coding region of the recombinant TCRs. The genes of the variable domains of three TCRs used are: CR15 (Hünig & Bevan, 1982), P14 (Pircher *et al.*, 1987) and 8/10-2 (Iwamoto *et al.*, 1987). The genes were included up to the amino acids of the mature protein with the numbers (consecutive numbering): CR15, α chain 114, β chain 116; P14, α chain 124, β chain 115; 8/10-2, α chain 114, β chain 112. The peptide bond between codons 21 and 22 of the α chain of P14 was assigned as the signal cleavage site based on statistical methods (von Heijne, 1990; Folz & Gordon, 1987). The two chains were linked in the order α before β with a peptide linker given in Figure 1 and tagged with a *myc*-tag and five histidines as described by Lindner *et al.* (1993). The expression strain used in the standard system is JM83 (Yanisch-Perron *et al.*, 1985). For making bivalent scTCR, identical vectors were used except for the substitution of the *myc*-tag and the five histidines by the dimerization motif scdHLX (Pack *et al.*, 1993). In this case, the five amino acids DYKDE were added between the signal sequence and the α chain for immunological detection (Hopp *et al.*, 1988; Knappik & Plückthun, 1994). The plasmids carrying the *lac*UV5 and the *tac* promoter are identical to the standard vector, except for carrying the *lac*UV5 and the *tac* mutations (de Boer *et al.*, 1983).

For the construction of the direct chaperone co-expression system, a cloning site was created in the unique *Hin*dIII site of the standard vector by insertion of the oligonucleotide 5'-AGCTGCGGCCGCAGTCAGAAT-TCA-3' and the corresponding opposite strand oligonucleotide. This creates a *Not*I and an *Eco*RI site. The *rpo*H gene including its own promoter was cloned between these sites using PCR with the oligonucleotides 5' - ACTGAGCGGCCGCGTACAACATTTACGCCACTT - 3' and, 5'-AGTCTGAATTCGGAAATTGATTATTACAGA-GG-3'. The *dsbA* gene including its own promoter was obtained as a *Sna*BI/*Hin*dIII restriction fragment from the plasmid p12-7 (Bardwell *et al.*, 1991) and cloned after blunt-ending into the filled-in unique *Bam*HI site of the standard vector. All genes are transcribed clockwise as given in Figure 1B. The expression strain of the direct co-expression system was JM83.

The invertible promoter system is based on plasmid pNH16aKan (Wülfing & Plückthun, 1993). Briefly, the *lac* promoter is arranged between two sites recognized by the λ-integrase, initially in the orientation reading away from the gene. The integrase gene is provided on a separate plasmid (pCW 107) under very tight regulation (Wülfing & Plückthun, 1993). The expression plasmid that was used to obtain the results reported here uses the invertible promoter system, controlling expression of the CR15 scTCR gene. This is followed by the *rpo*H and *dsb*A genes arranged exactly as in the direct chaperone co-expression system, followed by the *lac*I^q gene (*Sal*I fragment of pDM1.1, H. Bujard, personal communication), all cloned into the multiple cloning site. Translation of the TCR is

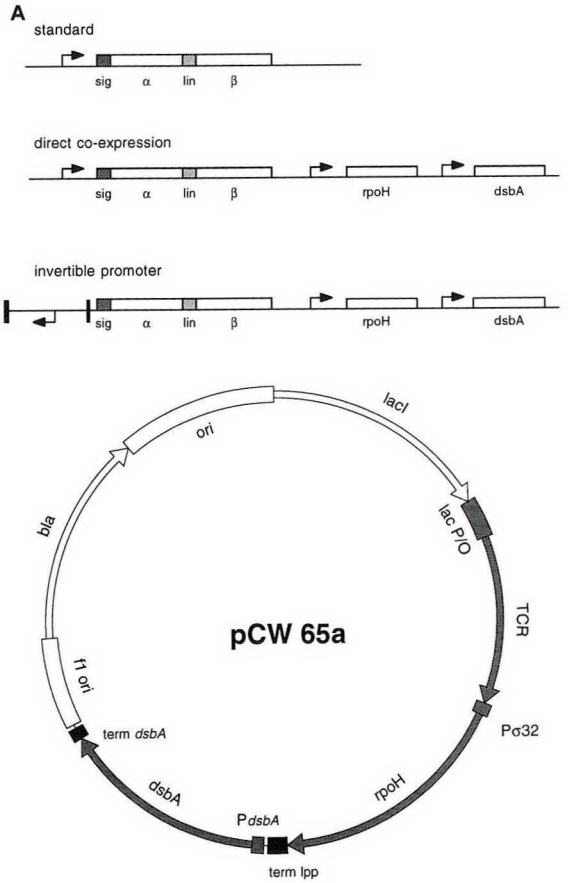

Figure 1. Expression systems. A, Schematic overview. The 3 basic expression systems are denoted by their names used in this paper. α and β represent the variable domains of the α-, respectively β-chain of the TCRs. *sig* is the *omp*A-signal sequence, *lin* the peptide linker between the variable domains. *rpo*H and *dsb*A denote the respective genes. Promoters are indicated by arrows. In the invertible promoter system, the promoter is shown in the off-position, the cassette between the 2 vertical bars gets inverted upon heat-shock allowing the transcription of the scTCR gene (Wülfing & Plückthun, 1993). B, Expression vector pCW65a for the direct co-expression system. Open lines indicate functional elements of the vector retained from pASK30 (f1 origin, which allows for the production of single-stranded DNA, the *bla* gene, encoding the β-lactamase, the colE1 origin and the *lac*I gene, encoding the *lac* repressor). The filled lines define the complete expression cassette, which includes clockwise the CR15 scTCR gene under the control of the *lac* promoter, the *rpo*H gene under its own promoter and the *dsb*A gene under its own promoter. Terminators within the expression cassette are indicated by filled bars.

driven by a T7g10 translation initiation region. The sequence inserted in front of the start codon is 5'-AATAATTTTGTTTAACTTTAAGAAGGAGATATACAT-3'. The expression strain was JM83 (pCW107: Wülfing & Plückthun, 1993).

(b) *Protein expression*

E. coli cells were grown at room temperature for all systems except for the invertible promoter system. For cells containing plasmids that use this system the procedure of Hasan & Szybalski (1987) is followed, with 1 mM IPTG added for induction 30 minutes after heat-shock.

(c) *Cell fractionation*

In all cases cells were harvested three hours after induction. The separation of the subcellular fractions was carried out as described elsewhere (Schmidt, 1991; Glockshuber *et al.*, 1992). Briefly, cells of a one liter culture were resuspended in 2·8 ml of 0·2 M Tris-HCl (pH 8·0), 0·5 mM EDTA, 0·5 M sucrose (buffer A). They were diluted with 7·2 ml of a 1 to 1 mixture of buffer A with water. After 30 minutes on ice the spheroplasts were centrifuged leaving the periplasmic extract as the supernatant and the spheroplasts plus the insoluble periplasmic material as the pellet. For small scale cultures, the extract was made fivefold more dilute, and lysozyme was added at a final concentration of 0·6 mg/ml. Furthermore, for analytical purposes the resuspended spheroplast pellet was separated into a soluble and an insoluble fraction by centrifugation after French Pressure Cell treatment. These fractions were termed "cytoplasmic". As shown for the antibody McPC603 (Schmidt, 1991; Glockshuber *et al.*, 1992), the subcellular fractions termed cytoplasmic essentially contain periplasmic material of the antibody that is either loosely membrane-associated or aggregated and thus can be pelleted together with the spheroplasts. Since the signal peptide of all scTCRs is cleaved almost quantitatively in the corresponding cytoplasmic fractions, a similar situation is assumed for the scTCRs as for the antibody McPC603. scTCR was detected by Western blotting using either the anti *myc*-tag antibody 9E10 (Oncogene Science, Uniondale, NY) or the anti FLAG antibody M1 (International Biotechnologies, New Haven, CT) following the protocols suggested by the suppliers.

(d) *Other methods*

Antibody affinity columns were prepared by covalently crosslinking the antibody bound to Protein G Sepharose fast flow (Pharmacia, Uppsala) as described by Harlow & Lane (1988). The antibody density was 1 mg per ml resin for the FAZ antibody (Prowald, 1990), and 5 mg per ml resin for the F23.1 antibody (Staerz *et al.*, 1985). For binding studies, one liter of *E. coli* expressing the respective scTCR genes in the respective expression system was grown as described above. The periplasmic soluble fraction (see above) was precipitated with saturated ammonium sulfate overnight, the pellet was resuspended in 1 ml 50 mM Tris (pH 8·0), 150 mM NaCl and incubated overnight at 4°C with 0·25 ml of immuno-affinity column slurry with slow shaking. After packing the slurry into a column, it was washed and eluted following the method described by Harlow & Lane (1988). The composition of the elution buffers used is given in the respective Figure captions. Based on the antibody-density of the respective column and the volume of the immuno-affinity column used, the maximal binding capacity of the affinity columns, as used in the binding experiments, was 200 μg (F23.1), or 40 μg (FAZ) of scTCR, respectively. It seems reasonable to assume that during loading of the column part of the antibody is blocked by unspecific binding. These capacities thus represent upper limits.

3. Results

Single-chain TCR fragments (scTCRs) are predicted to be structurally homologous to scFv

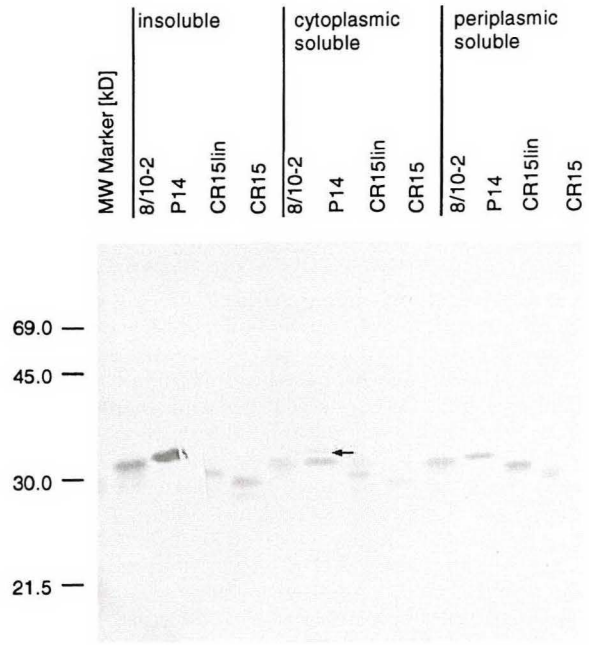

Figure 2. Expression of the different scTCRs. CR15lin, P14 and 8/10-2 have as the peptide linker connecting the α and the β chain the peptide GSGSSSSGSSSSGSG, CR15 has as a linker the peptide GGGGSGGGGSGGGGS. In most experiments, unless noted otherwise, CR15 was used, not CR15lin. A Western blot is shown with the anti-*myc*-tag antibody 9E10, comparing different subcellular fractions of *E. coli* cells expressing the respective scTCRs in the standard system (Figure 1). The precursor of the P14 scTCR can be seen in the corresponding cytoplasmic fraction, indicated with an arrow.

antibody fragments. Since the expression of scFv antibody fragments in *E. coli* is well established (Plückthun, 1992 and references therein) we initially followed an expression strategy for the scTCRs completely analogous to the system we used previously for the secretion of antibody fragments to the periplasm of *E. coli* (Skerra *et al.*, 1991). This system, including JM83 as the expression strain, will be referred to as the "standard system" throughout this paper (Figure 1A). Three different TCR clones were used: CR15 (Hünig & Bevan, 1982), P14 (Pircher *et al.*, 1987) and 8/10-2 (Iwamoto *et al.*, 1987). All of them are secreted to the periplasm of *E. coli* with efficiencies >95%, as judged from the complete removal of the signal-sequence. That the amount of non-processed precursor in the cytoplasm is small can be seen in several Western blots (e.g. Figure 2). Disulfide bridges are formed in all three scTCRs in the periplasm as judged from SDS/PAGE gels, run under reducing and non-reducing conditions in direct comparison (data not shown). The expression levels in the standard system, however, differ considerably. Whereas the P14 scTCR is expressed with a yield of at least 0·5 mg/l culture (legend to Figure 5B), the other two scTCRs are expressed with yields that are more than an order of magnitude smaller (legend to Figure 3A; data not shown). Furthermore, using Western blots of *E. coli* cell fractions, proteolytic degradation of CR15 and 8/10-2 was observed (data not shown).

(a) *Expression of the CR15 scTCR using enhanced* in vivo *folding*

In order to improve the expression level of the CR15 scTCR, several modifications were introduced into the basic host/vector system (summarized in Table 1). Some variation in the expression levels was seen, but the basic picture as outlined above did not change. Two results, however, will be discussed below. We increased the expression of the CR15 scTCR by increasing the promoter strength by changing the *lac* promoter to the *lac*UV5 or the *tac* promoter. The amount of CR15 scTCR produced increased by about a factor of two in each step, in all cell fractions (as estimated by Western blot, data not shown). Using

Table 1
Modifications in the standard expression system tested

Parameter	Variable tested
Variation of the promoter strength	*lac*, *lac*UV5, *tac*
mRNA secondary structure	Removal of TIR secondary structure, introduction of the T7g10 TIR, including the 5'-stem-loop
Chain connection	Single-chain *versus* F_v-analog, different linkers
Fusion protein	Fusion with *bla*
Different expression system	T7 system, pET-based
	T7 system, own constructs
Expression strain with relevant marker	KS476 (*degP*)
	AB1899, M6 (*lon*)
	UT4400 (*omp*T)
	MRE600 (RNAse III)
	CAG629 (*lon*, *rpo*H)

Summary of a whole set of modifications of the standard system introduced in order to increase the expression level of the CR15 scTCR. Transcription, translation, chain connectivity or accidental sensitivity to a specific protease do not seem to be the reason for the low yields of expression of the CR15 scTCR in the standard system, since changing these parameters does not have a major influence on the expression level (C.W. & A.P., unpublished data).

Table 2
Co-expression of parts of the E. coli *folding apparatus*

Co-expressed gene(s)	Way of co-expression	Effect
*gro*E	Under its own promoter	Cell lysis stability
rot	In an operon with the TCR	None
*dsb*A	In an operon with the TCR	None
	under its own promoter	None
*rpo*H	In an operon with the TCR	Cell lysis stability
	under its own promoter	Cell lysis stability
*rpo*H + *rot*	Both under the *rpo*H promoter	Cell lysis stability
*rpo*H + *dsb*A	Both under the *rpo*H promoter	Cell lysis stability
*rpo*H + *dsb*A	Both under their own promoter	High expression level, no degradation

The purpose of this Table is to show that the *in vivo* folding enhancement effect is not simply a question of accelerating one rate-determining step in the folding of the CR15 scTCR. At least 2 components are involved. DsbA acts as a specific component and σ^{32} as a general component inducing one (or more) yet unidentified further components. The heat-shock specific sigma-factor σ^{32} is encoded by the *rpo*H gene (Grossman *et al.*, 1984). Its over-expression leads to a modest over-expression of all heat-shock proteins of *E. coli* (Grossman *et al.*, 1987). As heat-shock leads to increased denaturation of proteins, we can expect among the heat-shock proteins components that are competent in dealing with misfolded protein. Over-expression of σ^{32} thus might lead to especially "folding-competent" *E. coli*. Over-expression of DsbA (Bardwell *et al.*, 1991), the periplasmic rotamase (Liu & Walsh, 1990) and GroE (Hemmingsen *et al.*, 1988) was verified by SDS/PAGE of subcellular fractions of *E. coli* with and without over-expression (Figure 3A). Since massive (direct) over-expression of GroE leads to the same lysis stability effect of *E. coli* as over-expression of the *rpo*H gene on a plasmid, we conclude that plasmid encoded *rpo*H must be functional.

the protease deficient strain CAG629 (relevant markers: *lon*$^-$, *rpo*H$^-$) as the expression host, large amounts of CR15 scTCR were obtained using either the *lac* or *tac* promoter (>2 mg/l culture, as judged from IMAC purification), but the scTCR produced was completely insoluble and heavily degraded. Using a *lon*$^-$ strain (Table 1) without heat-shock deficiency, no effect on expression level and amount of soluble protein was seen (data not shown).

The sensitivity of the CR15 scTCR to proteolysis and its tendency to aggregate in these studies were interpreted as signs of misfolding. In order to assist *in vivo* folding, we constructed a series of new vectors for coordinately over-expressing a number of proteins likely to be involved in the *E. coli* folding process (Table 2). In these experiments, the gene in question was placed downstream of the scTCR gene, and under the control of the promoter given in Table 2. The expression strain in all experiments outlined in Table 2 was JM83. Whereas the over-expression of single proteins of the folding apparatus did not have a significant effect on CR15 scTCR expression, over-expressing the heat-shock sigma factor σ^{32} *together* with the *E. coli* periplasmic disulfide isomerase *dsb*A (Figure 1), both from their own promoters, resulted in an expression behavior of the CR15 scTCR comparable to the P14 scTCR (Figure 3). Thus, the expression level was now in the range of 1 mg/l culture, with no signs of proteolytic degradation. Since the same transcriptional and translational signals and the same expression strain are used as in the standard system, the over-expression of parts of the folding apparatus being the only difference, we attribute this drastic effect to an enhanced *in vivo* folding. This system is called the "direct co-expression system" in this publication (Figure 1). The possible role of proteolysis in changing the amount of CR15 scTCR is discussed further below.

The expression plasmid for the direct co-expression system, pCW65a, was found to be genetically unstable (data not shown). Since the other scTCRs can be maintained in the same direct chaperone co-expression system, the instability of pCW65a cannot be due to the plasmid-based over-expression of σ^{32} and DsbA, but must be caused by specific properties of the CR15 scTCR. Therefore, a genetically stable version of this system (including the presence σ^{32} (*rpo*H) and *dsb*A genes in an identical fashion) had to be constructed (see Materials and Methods). We used an improved version of the invertible promoter system of Szybalski and co-workers (Podhajska *et al.*, 1985; Hasan & Szybalski, 1987; Wülfing & Plückthun, 1993). In this invertible promoter system (also expressing σ^{32} (*rpo*H) and *dsb*A) the amount of soluble periplasmic material of the CR15 scTCR is intermediate between the standard and the direct chaperone co-expression system, and proteolytic degradation is diminished in comparison to the standard system (Wülfing & Plückthun, 1993; and unpublished data). During the following experiments this system (called the invertible promoter system) (Figure 1A) was used as a stable substitute for the direct chaperone co-expression system for the expression of the CR15 scTCR.

(b) *Characterization of the CR15 scTCR*

In order to characterize the material produced with the different CR15 scTCR expression systems, different types of chromatography were used. French Press extracts were used as the starting material, as they would contain all soluble scTCR protein. With the standard expression system for CR15 scTCR, most of the scTCR protein did not bind to different standard chromatographic columns, such as DEAE-Sepharose, Ni^{2+}-NTA Agarose or Phenyl-

Sepharose (data not shown). The large majority of the scTCR protein was found in the run-through fraction. This suggests that the soluble scTCR of the French Press extract is largely present as aggregated misfolded material.

The folding of the CR15 scTCR in the periplasmic soluble fractions was assayed, using a monoclonal antibody (called FAZ; Prowald, 1990), which recognizes a clonotypic, conformational epitope, covalently linked to a Protein G Sepharose column. Because of the conformational specificity of the FAZ antibody (it does not bind to the protein on Western blots), all scTCR able to bind the column can be regarded as at least containing the correctly folded epitope of the CR15 TCR, and most likely the same structure as the TCR expressed on the surface of the original T-cell line. Soluble periplasmic material obtained with the standard system and with the invertible promoter system was compared. Both systems yield CR15 scTCR in the periplasmic soluble fraction that behaves the same with respect to binding and elution conditions from the FAZ affinity column. In both cases only a very small fraction of the material bound to the column (Figure 4A, and data not shown). Reloading the run-through fraction of the affinity column onto another column gave the same binding and elution behavior (data not shown). Since all the material bound to the column elutes already under mildly acidic conditions (Figure 4A), the binding energy between CR15 scTCR and FAZ must be low. These data are compatible with a low affinity interaction between FAZ and the CR15 scTCR. Thus, we conclude that correctly folded material is present, the relative amount of which cannot be quantified. Quantification was, however, possible using bivalent CR15 scTCRs.

(c) *Bivalent CR15 scTCR*

A bivalent scTCR was constructed by fusing a helix-turn-helix module, connected by a hinge peptide, to the end of the scTCR. The helices, being amphipathic, act as a dimerization domain *in vivo*, as previously demonstrated for single-chain antibody fragments (Pack & Plückthun, 1992; Pack *et al.*, 1993). A correctly folded dimer can be expected to show a polyvalency effect of the apparent avidity in binding to the FAZ antibody affinity-column. However, if only a minor fraction of the periplasmic soluble material is correctly folded, and assuming independent folding of the two domains within the dimer, which are separated by long hinge regions, hardly any functional dimer would be obtained for statistical reasons. Observing an avidity effect would therefore imply that a substantial amount of the periplasmic soluble material is correctly folded. Such

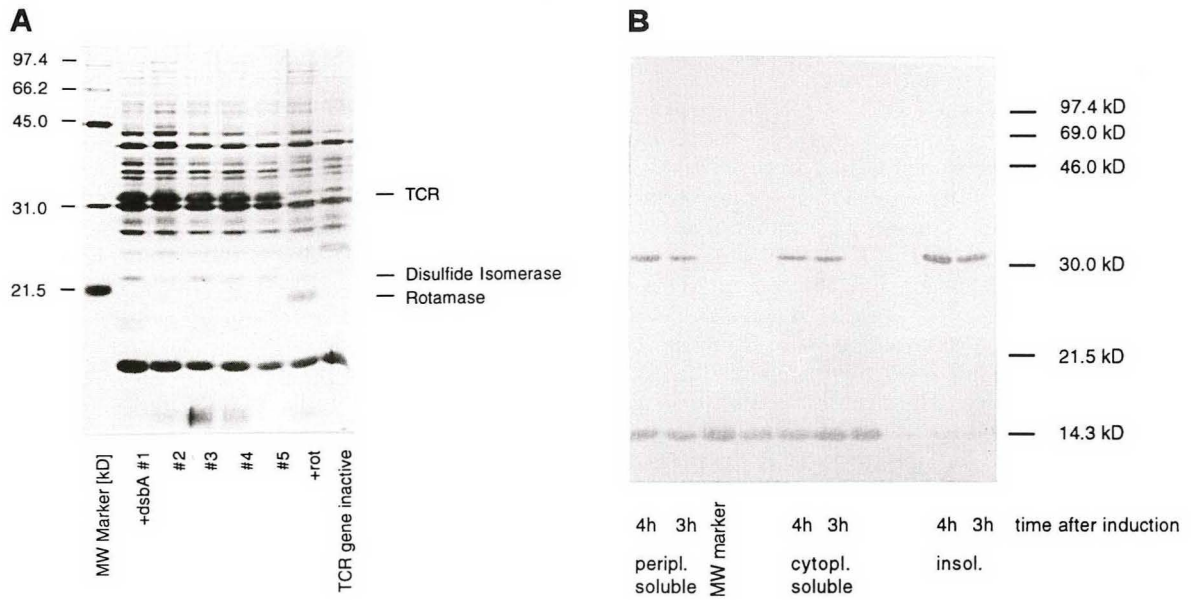

Figure 3. Expression of CR15 scTCR with the direct σ^{32}/DsbA co-expression system. A, A silver-stained SDS/PAGE gel of periplasmic fractions of different constructs. Lanes *+ DsbA #1* to *+ DsbA #5* show 5 independent transformants of the direct co-expression system. *+ rot* indicates a similar construct over-expressing σ^{32} and (instead of *dsb*A) the *rot* gene, *TCR gene inactive* denotes a clone of the direct over-expression system, carrying a mutation in the *lac*-promoter region disabling CR15 scTCR expression. The last 2 lanes are included as controls. CR15 scTCR expression can be seen as an additional band in the lanes of the 5 transformants as indicated. Comparing this gel with similar gels for the expression of an McPC603 Fv antibody fragment (using the bands of *E. coli* periplasmic proteins as internal standards of comparison) the expression level of the CR15 scTCR as judged by band intensities is higher than the antibody fragment expression level. The CR15 scTCR expression level in the direct co-expression system must therefore be in the range of 1 mg/l culture. Since with the standard system (i.e. without σ^{32}/DsbA over-expression), in contrast to these results, no TCR band is seen in a silver-stained SDS/PAGE gel, the expression level there must be lower by orders of magnitude. Increased amounts of rotamase and DsbA due to plasmid-based over-expression are indicated. B, Western blot of the different subcellular fractions of *E. coli* expressing the CR15 scTCR gene with the direct co-expression system. The scTCR is detected with the anti-*myc*-tag antibody 9E10.

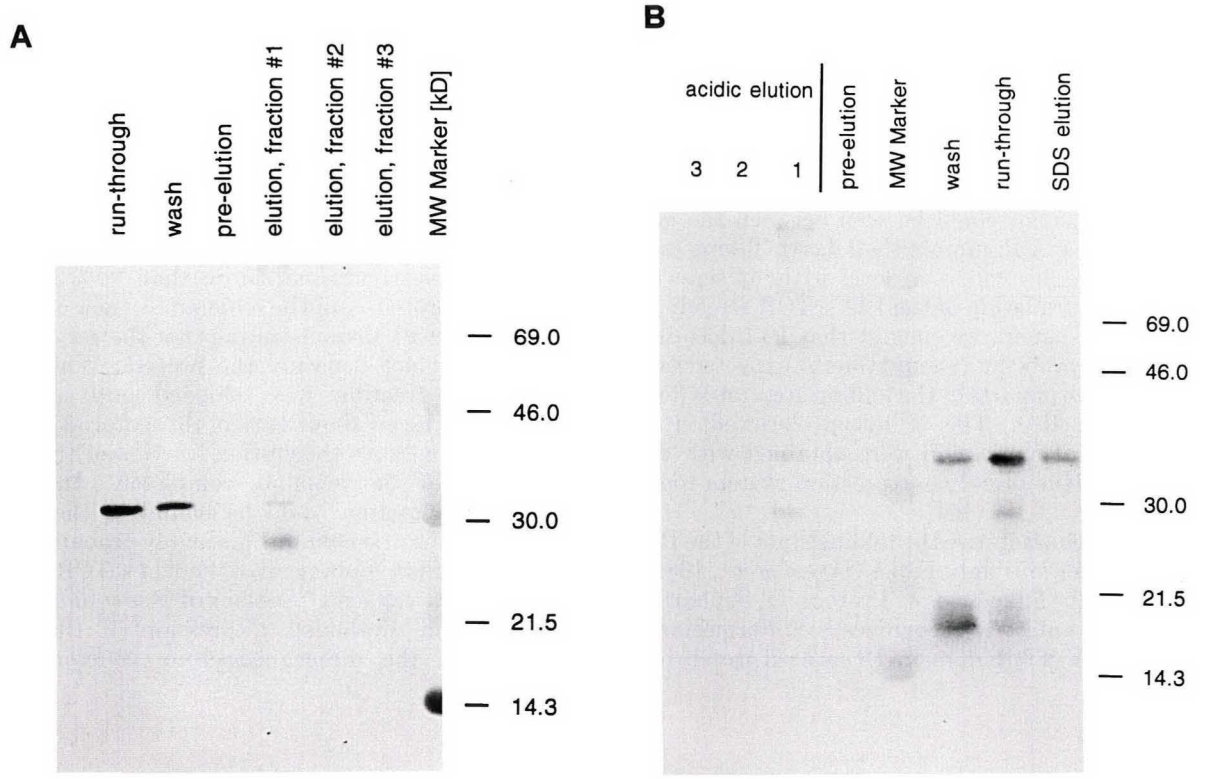

Figure 4. Binding of CR15 scTCR to an FAZ antibody affinity column. A, CR15 scTCR is expressed with the invertible promoter system. A Western blot with the anti-*myc*-tag antibody 9E10 of the affinity column fractions is shown. For elution, 100 mM glycine (pH 3·0) as used. All fractions contain comparable amounts of scTCR, except for the run-through fraction which is 7 times less concentrated. Some FAZ antibody leaking out of the column can be seen (at a molecular weight of 28 kDa). A silver-stained SDS/PAGE gel of the same fractions showed the binding to be specific. Stronger elution conditions did not elute any further scTCR. B, Bivalent CR15 scTCR was produced using the direct co-expression system. A Western blot with anti-FLAG antibody M1 is shown. Elution steps were carried out in the following order: (1) 100 mM glycine, pH 3·0, (2) 200 mM borate, pH 9·0 (neutralization step, not loaded on the gel), (3) 1% SDS. All fractions contain comparable amounts of scTCR, except for the run-through fraction, which is 7 times less concentrated.

an improved binding of the dimer is indeed found (Figure 4B). About 10% of the CR15 scTCR material now binds to the column and can be eluted only under harsh conditions. This increase in binding energy, together with the increased proportion of material bound, can most easily be explained with the predicted avidity effect. Yet, reloading the run-through fraction onto another column once again results in binding of 10% of the CR15 scTCR material. Thus, not all of the correctly folded material has bound to the first column. We attribute this to a geometrical problem: since the density of FAZ-antibody on the column is comparatively low, only a few sites seem to be available with the correct spacing for bivalent binding. Using a similar column with five times the antibody density for similar experiments with the 8/10-2 scTCR, the problem of partial binding of a scTCR is not seen anymore, supporting our hypothesis. Furthermore, some monovalent binding can still be seen with the CR15 scTCR dimer on the FAZ column, as a part of the TCR dimer eluted under mildly acidic conditions (Figure 4B). As this does not occur in similar experiments with the 8/10-2 scTCR dimer, there seems to be a fraction of sites on the FAZ column that are only accessible for monovalent binding according to our hypothesis.

From the FAZ column experiments we conclude that a substantial fraction, and perhaps the majority of the soluble CR15 scTCR in the periplasmic fraction is correctly folded, independently of the expression system used. In contrast, the primary effect of over-expressing σ^{32} and DsbA appears to be a substantial increase in the yield of periplasmic soluble scTCR, of which at least a large part is correctly folded. This increased yield may reflect more efficient folding together with an inherent increased stability of the correctly folded material. One might at first attribute these results simply to increased export levels of the scTCR to the periplasm due to over-expression of cytoplasmic chaperones. Since, however, in the presence of over-produced DsbA the expression level (and therefore the export level as well) of the CR15 scTCR was systematically varied using the invertible promoter system, but enhanced yields of periplasmic soluble material were not observed (data not shown), we attribute the enhanced levels of periplasmic, soluble CR15 scTCR using the direct chaperone co-expression system to an enhanced periplasmic folding.

(d) *Expression and characterization of the P14 scTCR*

The P14 scTCR was also expressed in both the standard and the direct chaperone co-expression system (Figure 1A). The direct co-expression system was genetically stable with P14, and no differences in expression behavior could be seen between the two systems (C.W. & A.P., unpublished data). Taking into account the high expression level without signs of proteolytic degradation of the P14 scTCR already in the standard system, we suggest that P14 does not have a folding problem. It could thus be expected that over-expression of parts of the folding apparatus does not have any effect. This is indeed observed, since similar levels of expression were obtained with the standard and the direct co-expression system (data not shown).

In order to characterize the folding state of the P14 scTCR, another antibody, F23.1 (Staerz *et al.*, 1985), was covalently linked to a Protein G Sepharose column. This antibody recognizes a conformational epitope of the $V_\beta 8$ TCR domain. Denatured protein in a Western blot is again not recognized (data not shown).

In contrast to the situation with the CR15 scTCR where almost no binding was obtained because of low antibody affinity, about 10% of the normal monovalent P14 scTCR from a periplasmic soluble fraction of *E. coli* bound to the column (Figure 5). The binding is specific, since no other proteins were bound to the column, except in trace amounts (Figure 5B). The bound material could only be eluted under strongly acidic conditions. More than 90% of the eluted TCR protein could be rebound to a new column (data not shown), demonstrating that the very short acid step did not denature the protein. When the run-through fraction was reloaded onto a new column, once again about 10% of the material bound to the column. Since the elution fraction of the first column could be rebound completely, and the run-through fraction could be bound to the same extent as on the first column, it is highly probable that the reason for not binding all of the P14 scTCR at the first time is a capacity problem of the column (see Materials and Methods). Expression of the P14 scTCR with the direct chaperone co-expression

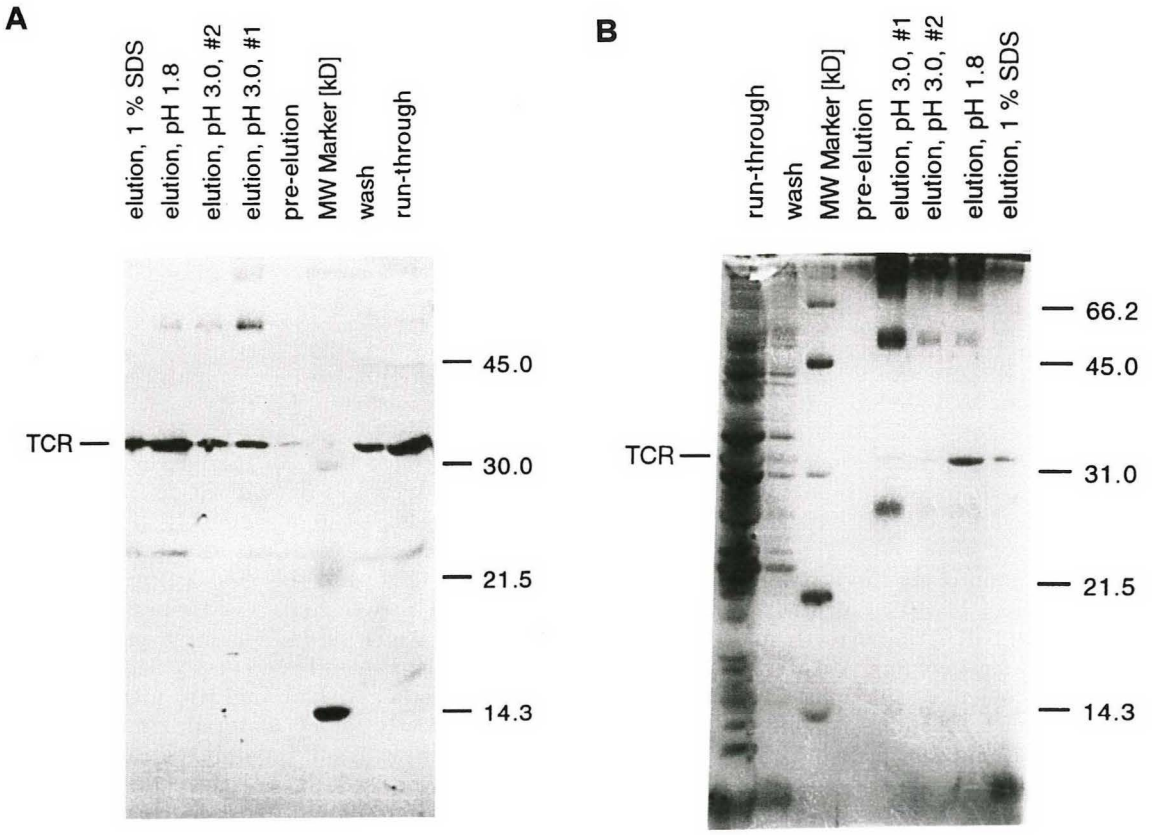

Figure 5. Binding of P14 scTCR to an F23.1 affinity column. P14 scTCR was produced using the standard expression system. A, Western blot with the anti *myc*-tag antibody 9E10 of the affinity column fractions. The column was eluted in the order: (1) 100 mM glycine, pH 3·0, (2) 100 mM glycine, pH 1·8, (3) 200 mM borate, pH 9·0 (neutralization step, not loaded on the gel), (4) 1% SDS. All fractions contain comparable amounts of scTCR, except for the run-through fraction, which is 7 times less concentrated. Some F23.1 antibody leaking out of the column can be seen (at a molecular weight of about 28 and 50 kDa). B, Silver-stained SDS/PAGE gel of the same fractions as A. The amount of P14 scTCR bound to the column was used to estimate the expression yield. Therefore the band intensity was compared to the intensities of the marker bands (containing each 400 ng of protein). Taking into account that only 15% of the material has bound to the column, the expression yield was estimated to be in the range of 0·5 mg/l culture.

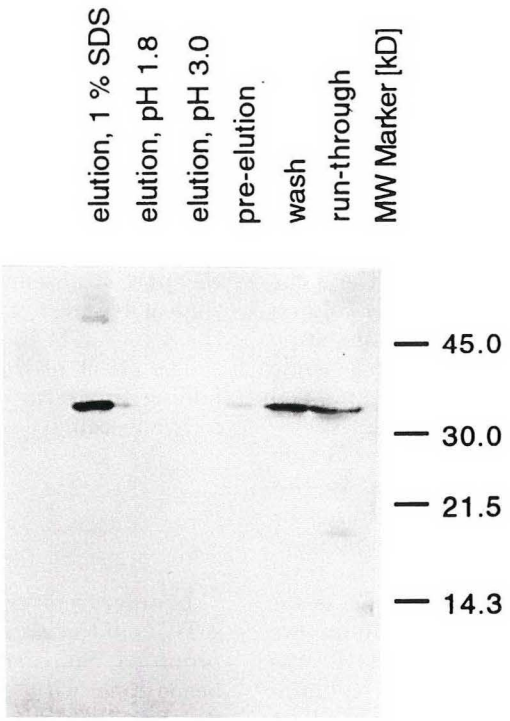

Figure 6. Binding of dimeric P14 scTCR to an F23.1 affinity column. P14 scTCR, dimerized *in vivo*, was produced using the standard expression system. A Western blot with the anti-FLAG antibody M1 of the affinity column fractions is shown. The column was eluted in the order: (1) 100 mM glycine, pH 3·0, (2) 100 mM glycine, pH 1·8, (3) 200 mM borate (pH 9·0: neutralization step, not loaded on the gel), (4) 1% SDS. All fractions contain comparable amounts of scTCR, except for the run-through fraction, which is 7 times less concentrated. Some F23.1 antibody leaking out of the column can be seen (at a molecular weight of about 28 and 50 kDa). A silver-stained SDS/PAGE gel of the same fractions showed the binding to be specific.

system instead of the standard system resulted in the same binding and elution behavior of the periplasmic soluble material (data not shown).

(e) *Bivalent P14 scTCR*

Once again the scTCR was dimerized *in vivo* using the helix module strategy (Pack & Plückthun, 1992; Pack *et al.*, 1993) to assay for an avidity effect. Using the F23.1 affinity column, such an effect was again seen (Figure 6), since all of the bound material could be eluted only under harsh conditions in contrast to the monomeric scTCR described above. Reloading the run-through fraction of the first column onto a second column resulted again in binding of bivalent P14 scTCR to the column. However, the amount of material bound to the second column was smaller than the amount bound to the first column (data not shown). Thus, by the time of binding to the second column, part of the bivalent scTCR was not correctly folded anymore, since otherwise binding of the same amount of bivalent P14 scTCR as to the first column would have been observed. We therefore suggest that the dimerization motif either disturbs folding of the dimeric P14 scTCR in the periplasm or impairs its stability in the periplasmic extract. The specificity of the F23.1 column was verified by loading dimeric CR15 scTCR onto the F23.1 column. CR15 uses a different V_β segment, and is therefore expected not to bind to the column. Indeed, no binding was found (data not shown) demonstrating that binding to the F23.1 column occurs only *via* the variable domains of the bivalent scTCR and not *via* the helices. Therefore, we can attribute the observed tight binding to the bivalent binding of the epitope in the variable domains to the affinity column, and not to an interaction with the dimerization helices.

From the affinity column experiments, we conclude that a substantial part of the P14 scTCR periplasmic soluble material is correctly folded, and it is possible that all of the material has the correct conformation. In this case however, some slow unfolding of the dimeric P14 scTCR has to be postulated to account for the reduced binding of the dimeric P14 scTCR when loaded onto the second affinity column. In contrast to the CR15 scTCR, the expression levels are independent of the expression system used. Having a scTCR that is able to fold without additional assistance to a large extent, the over-expression of parts of the *E. coli* folding apparatus is not required and has no effect.

(f) *Expression and characterization of the 8/10-2 scTCR*

The scTCR 8/10-2 was also expressed in the standard and the direct chaperone co-expression system (Figure 1A). In this case, the scTCR was

highly sensitive to proteolysis (C.W. & A.P., unpublished results). In addition, 8/10-2 scTCR was covalently modified after prolonged induction of expression in the standard system leading to a band of lower mobility in SDS/PAGE gels. This modification was observed without any lag phase in the direct co-expression system (C.W. & A.P., unpublished results). There is no explanation for the modification so far.

Since the TCRs P14 and 8/10-2 share the same β-chain variable segment, the F23.1 affinity column could be used for the evaluation of the folding state of the 8/10-2 scTCR in the soluble periplasmic fraction as well. About 10% of the material, expressed in the standard system, bound to the column. Exactly as with the P14 scTCR, the 8/10-2 scTCR could only be eluted under strongly acidic conditions. In this case however, the run-through fraction could not be rebound to a second column due to the extensive degradation of the 8/10-2 scTCR. Since the total amount of 8/10-2 scTCR, due to proteolysis, is far below the capacity limit of the F23.1 column, we deduce that only 10% of the 8/10-2 scTCR was correctly folded at the time of binding to the column. Using bivalent 8/10-2 scTCR, an avidity effect analogous to the other two TCRs could be seen. Yet, once again, only about 10% of the material bound to the column, suggesting that the majority of the dimeric 8/10-2 scTCR is misfolded at the time of binding to the column. Reloading the run-through fraction onto a second column resulted in no binding (C.W. & A.P., all unpublished data). As with the monomeric material, the majority of the dimeric material must be therefore misfolded. The strength with which the 8/10-2 scTCR binds to the column (as judged by elution conditions) is identical to the P14 scTCR. Sensitivity to proteolytic degradation and the partial absence of correct conformation at the time of loading onto the column are characteristics of the 8/10-2 scTCR.

The effect of the over-expression of parts of the folding apparatus of *E. coli* cannot be evaluated due to the instability of the 8/10-2 scTCR.

4. Discussion

(a) *A general model*

In order to rationalize the behavior of the different scTCRs in the various expression systems, a model is proposed. Since the majority of experiments have been done with the CR15 scTCR, the model is developed based on those experiments (Figure 7A). But, since all the experiments carried out with the other two scTCRs fit into the model easily, the model is extended to all three scTCRs (Figure 7B).

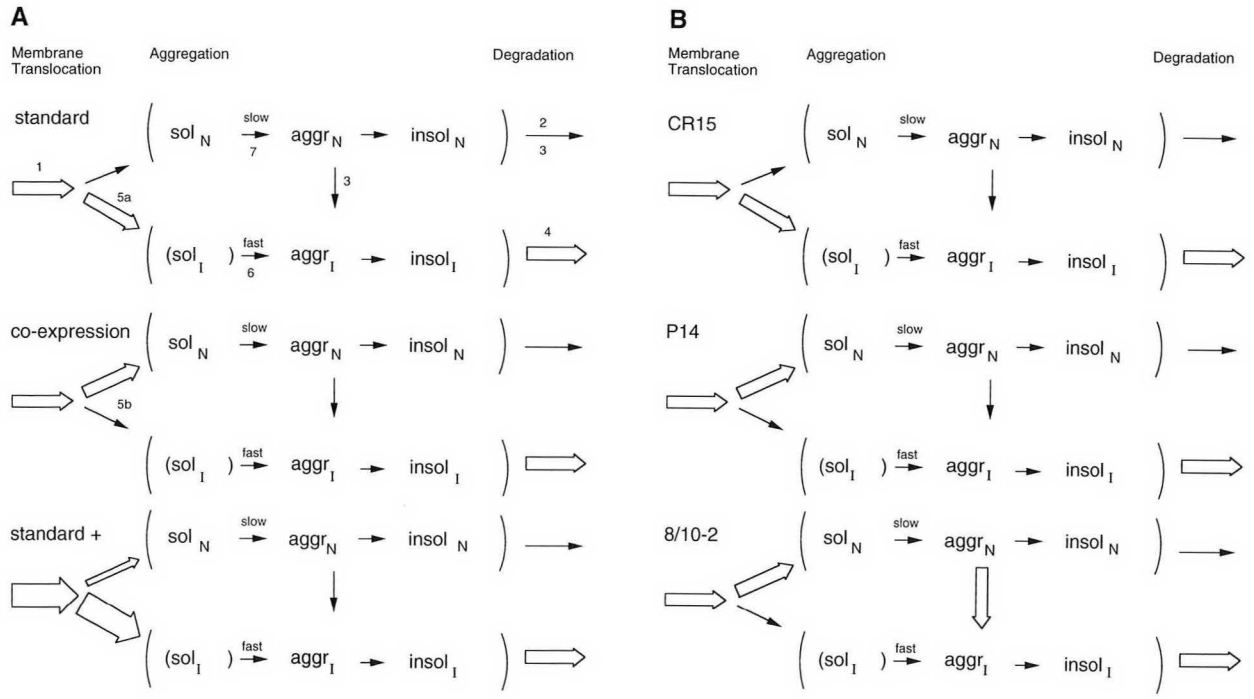

Figure 7. Model rationalizing the expression behavior of the scTCR. The thickness of the arrows outside the brackets represents the amount of material taking this pathway. The subscripts *N* and *I* refer to native and non-native conformations, respectively. *Sol* refers to monomerically soluble, *aggr* to soluble but aggregated protein and *insol* to insoluble protein. The numbers refer to processes discussed in the text. Since details of the precipitation process of the scTCRs are not available, arrows within the brackets only indicate that there is some precipitation, *slow* and *fast* indicate that correctly folded and misfolded scTCR are prone to aggregation to a different extent, as argued in the text. The existence of misfolded soluble material is hypothetical, sol_I is therefore put in parentheses. A, Comparison of 3 different CR15 expression systems. *standard* is the standard expression system, *co-expression* the direct co-expression system and *standard* + is the standard system using a stronger promoter (*tac* or *lac*UV5 *versus lac*). Numbers refer to the explanation of the model in the text. B, Comparison of the 3 different scTCRs, as expressed in the standard expression system.

We suggest that transcription, translation and transport to the periplasm do not constitute a problem for the scTCR molecules. Once in the periplasm, the scTCR may fold or may aggregate. Once folded, it is essentially stable, and only some aggregation at high concentrations is observed. If not folded, however, the scTCR aggregates immediately and is quickly degraded. What kind of evidence do we have for the various steps suggested in the model?

Comparisons of the expression levels of the CR15 scTCR in the standard system, in the direct chaperone co-expression system and with the standard vector in the *rpo*H$^-$ *lon*$^-$ strain CAG629, reveals that they differ drastically. The transcriptional and translational signals used are the same, however. We would like to conclude that the amount of protein leaving the ribosome is essentially the same. In all systems, the signal sequence of the great majority of the scTCR protein is cleaved off. Since the signal peptidase is located on the periplasmic face of the inner membrane, there is only a very minor part of the scTCR material remaining in the cytoplasm. The scTCR molecules are therefore either readily transported out of or quickly degraded in the cytoplasm. Comparing the expression level of the CR15 scTCR in the presence of over-expressed σ^{32} alone (Table 2) *versus* σ^{32} plus DsbA, the major effect on the increased yield in the direct over-expression system (σ^{32} plus DsbA) is on periplasmic protein. If cytoplasmic degradation were responsible for the low expression yield in the standard system, a remedy would have to be cytoplasmic, not periplasmic. Thus, a major role for cytoplasmic degradation is unlikely. We suggest that a considerable amount of scTCR (in the range of 1 mg/l culture) reaches the periplasm in all cases studied (Figure 7A, 1).

Signs of degradation can be easily seen, when *E. coli* cell extracts of cells expressing CR15 scTCR are directly loaded onto an SDS/PAGE gel (data not shown). This contrasts with the situation of the periplasmic soluble fraction used for the FAZ affinity columns. When this material is loaded on a SDS/PAGE gel the bands are still sharp after the two days it takes to perform the column experiment (while the protein remains in contact with periplasmic *E. coli* proteases; Figure 4A). Therefore while part of the CR15 scTCR is prone to fast proteolysis (seen in the immediate extracts), part of it is essentially stable. From the FAZ affinity column experiments, we know that the stable part is essentially correctly folded. It seems reasonable, therefore, to attribute the protease lability of the unstable part to misfolding. Using a French Press extract, part of the misfolded material can be isolated. Misfolding is deduced from the very high tendency of this material to aggregate (in contrast to the soluble periplasmic fraction), which has been observed in different chromatography experiments.

Can the effect of the increased yield in the direct co-expression system (over-expressing σ^{32} and DsbA) be explained by reduced proteolysis of the completely folded material? Heat-shock conditions are known to induce proteases, however not to repress them (Morimoto *et al.*, 1990). Furthermore, using the protease and heat-shock deficient strain CAG629, proteolysis is slowed down, but folding is not improved; large amounts of partially degraded material can still be seen. Over-expressing parts of the *E. coli* folding apparatus in the direct co-expression system, in contrast, leads to large amounts of undegraded material (Figure 3). Thus, heat-shock deficiency (together with a deficiency in the cytoplasmic protease Lon, which does not have an effect on its own, Table 1) and increased heat-shock proficiency (together with DsbA over-expression, which also does not have an effect on its own, Table 1) both give increased yields of total protein. This makes reduced proteolysis due to σ^{32} over-expression highly improbable. A minor instability of the periplasmic soluble fractions used for FAZ affinity chromatography is seen, however (Figure 4b). Whether this is due to slow unfolding, followed by proteolysis, or direct slow proteolysis of the correctly folded material cannot be distinguished (Figure 7A, 3).

Can the effect of increased periplasmic yields be explained by enhanced transport to the periplasm due to over-expression of cytoplasmic chaperones? It might be argued that chaperones in the cytoplasm merely increase the transport competence of the scTCRs and thus overcome a transport limitation. Since we are able to transport considerably larger amounts of scTCR into the periplasm than seen with the chaperone co-expression system, we can exclude this possibility. Increasing the promoter strength (and as such the level of export) well above the level of the *lac*-promoter in the presence of over-expressed DsbA (but not σ^{32}) does not enhance the yield of soluble, periplasmic protein, but merely increases the level of processed insoluble protein. Thus, transport does not seem to limit the production of soluble TCR.

We suggest instead that the effect of increased yields in the direct over-expression system is due to enhanced *in vivo* folding (Figure 7A, 5b) working synergistically with substantially higher intrinsic proteolytic stability of the correctly folded material (Figure 7A, 2). Under *in vivo* conditions, periplasmic folding might well be coupled to late translocation steps. Furthermore, the "periplasmic" chaperone postulated to be induced in response to σ^{32} may of course be a periplasmic domain of a membrane-anchored protein. Nevertheless, the data presented here can only be consistently interpreted if over-expression of σ^{32} and DsbA gives an effect on periplasmic folding, and not the efficiency of transport, be the folding coupled to late translocation steps or not.

Our conclusion from the experiments described is that correctly folded CR15 scTCR is essentially stable (Figure 7A, 2), whereas misfolded scTCR is rapidly degraded (Figure 7A, 4). Since in the standard expression system only small amounts of CR15 scTCR are found, we suggest that the majority of the material reaching the periplasm misfolds (Figure 7A, 5a) and is degraded. The increased yields in the direct

co-expression system are therefore a consequence of enhanced *in vivo* folding in the periplasm.

Despite this beneficial effect on periplasmic folding, some insoluble material is seen in the direct chaperone co-expression system as well. Does, therefore, correctly folded CR15 scTCR aggregate and precipitate as well? Misfolded material quickly aggregates and is partially digested even in the aggregated form (Figure 7A, 6). Correctly folded material (i.e. TCR able to bind to the affinity column) is stable towards this proteolysis. Comparing the expression behavior of the standard vector in the *lon*$^-$ *htpR*$^-$ strain CAG629 to the direct chaperone co-expression system, both give high yields of total protein, the former, however, is partially degraded in all cellular fractions, and the latter not at all (Figure 3B). The material produced with these systems, despite being produced in comparable amounts, thus has a very different sensitivity to proteolysis, suggesting a different degree of folding prior to precipitation.

A prediction of the model would be that increasing the promoter strength from the *lac* promoter of the standard system to e.g. the *lac*UV5 and the *tac* promoter, yet without offering folding assistance, would increase the amount of soluble periplasmic protein only moderately (Figure 7A, standard +). Most of the scTCR produced would still be misfolded and be degraded, perhaps increasing the steady-state concentrations of the aggregated material. A few-fold increase in the amount of material of the different subcellular fractions after increasing the promoter strength is indeed found.

The P14 scTCR behaves as CR15 in the direct co-expression system. However, the 8/10-2 scTCR is rapidly degraded. Using the model, this can be rationalized as a limited thermodynamic stability of the soluble, correctly folded 8/10-2 scTCR. This would result in an increased amount of unfolded protein under equilibrium conditions. In the presence of periplasmic proteases, however, this would result in fast degradation. Alternatively, one would have to postulate that there are large amounts of misfolded soluble periplasmic material *in vivo*. A third possibility, would be an extreme protease sensitivity of the folded protein. Even though we cannot yet distinguish these alternatives, reduced thermodynamic stability of the correctly folded form with concomitant proteolysis of the unfolded form seems to be the explanation requiring the most reasonable assumptions. In conclusion, we believe that the experimental behavior of the P14 and 8/10-2 scTCRs is consistent with this model.

We propose, therefore, that the expression of scTCR fragments in the periplasm is controlled by their folding. If they do not obtain and maintain a correctly folded conformation, they are rapidly degraded. We have presented an expression system that uses over-expression of parts of the *E. coli* folding apparatus to increase the folding yield of one of our scTCRs. We have furthermore shown that it is possible to dimerize scTCRs *in vivo* to make them bivalent without seriously impairing folding or stability of the scTCR.

(b) *Differences between the scTCRs and antibody fragments and between the different scTCRs*

Comparing the three different scTCRs studied here, two questions have to be answered. Why does the P14 scTCR not need either DsbA or σ^{32} over-expression? Why is the 8/10-2 scTCR especially unstable?

Lattman & Rose (1993) suggest that the folding mechanism of proteins in general is very robust, because it is redundantly encoded by their primary sequence. It seems therefore reasonable to assume that the scTCRs all have the potential to obtain their correct fold. The fact that at least a considerable part of the periplasmic, soluble material is indeed correctly folded supports this hypothesis. Only one other scTCR, 1934.4, has so far been reported which can be expressed in a standard system without folding assistance (Ward, 1992). When the amino acid sequences of P14 and 1934.4, the sequences of the rest of the TCRs used in this study, and those that have been expressed in inclusion bodies by others are compared, the total number of charges in the variable domain of the α-chain is found to be greater for P14 and 1934.4 than for the rest. Whereas the standard TCR variable α-domain usually contains considerably less than 20 charged amino acids, the α domain of both P14 and 1934.4 contain a total of well above 20 charged amino acids. Since we have no evidence that the solubility of the final folded proteins differs dramatically, we suggest that the folding intermediates are prone to aggregation to rather different degrees. The number of charged amino acids might influence this tendency to aggregate.

P14 does not seem to benefit from the overexpression of σ^{32} and DsbA. It appears, therefore, that its folding intermediates may be less prone to aggregation, and this could correlate with the number of charged groups in the protein. This makes the P14 scTCR similar to antibodies, which do require the action of DsbA (Knappik *et al.*, 1993), but are independent of any overexpression. DsbA over-expression alone does not have an effect on CR15 scTCR expression, but only when used together with σ^{32} over-expression. We suggest that only after chaperones have prevented the early folding intermediates with exposed hydrophobic surfaces (Matthews, 1993) from aggregation, the CR15 scTCR can make use of the additional DsbA to catalyze the last folding step. A possible explanation for the need of additional DsbA in the case of CR15 scTCR is that this protein may be an especially poorly accessible substrate titrating out the DsbA protein, due to a transient aggregation or interaction with other periplasmic proteins.

In contrast to P14 and CR15, the 8/10-2 scTCR possesses two instead of one potential *N*-glycosylation sites in the variable domains (Marshall, 1974). It is assumed that these sites are used in TCRs (Allison & Lanier, 1987; Clevers *et al.*, 1988). If both glycosylation sites of 8/10-2 were actually used in the eukaryotic cell, the 8/10-2 scTCR expressed in *E. coli* would miss the stabilizing effect of the sugars (Machamer & Rose, 1988; Dubé *et al.*, 1988; Matzuk

& Boime, 1988) to a larger extent. This might explain its sensitivity to proteolysis. Having a correctly folded state, which, even though cross-linked by disulfides, has only a marginal stability, unfolding and subsequent proteolytic degradation would be much more prominent than for P14 and CR15. It should be kept in mind that the dimeric P14 scTCR also seems to unfold slowly. It appears that scTCRs are not particularly stable in general.

Antibody fragments are relatively easily expressed in a functional way in *E. coli*, while scTCRs are not, judging from our own results and the scarceness of other publications in the field. Comparing antibody scFv fragments and scTCRs there are two possible structural differences. Whereas antibodies are usually secreted molecules, the TCR is always part of a complex surface structure (Hunkapiller & Hood, 1989). Nevertheless, when an antibody scFv fragment is expressed, the variable/constant domain interface is exposed. The same is true for scTCRs. In addition however, it is conceivable that further interfaces are exposed, such as a TCR/CD3 interface, and a TCR/CD4 or TCR/CD8 interface, respectively. The hydrophobicity of the surface of a scTCR can thus be assumed to be greater than the hydrophobicity of a corresponding scFv antibody fragment and this is, indeed, found in molecular models. Whereas antibody scFv fragments are not glycosylated, the scTCRs used by us contain at least one attachment site for *N*-linked glycosylation. The scTCRs expressed in *E. coli* will lack their carbohydrates, which may be needed as a stabilizing or solubilizing element.

These structural differences between scTCRs and the corresponding antibody fragments can be used to rationalize their different expression behavior. Since antibody fragments probably have less exposed hydrophobic surfaces, aggregation during and after folding can reasonably be assumed to be slower. A larger proportion of material therefore reaches the correctly folded state.

(c) *Implications*

The expression of scTCRs in the periplasm of *E. coli* is controlled by their folding. We have shown that the folding yield of one of our scTCRs could be increased by the simultaneous over-expression of σ^{32} and DsbA. We therefore suggest the existence of a periplasmic chaperone that is expressed under the control of σ^{32} and that is able to assist the scTCR folding. It seems improbable that a family of specific chaperones, e.g. the PapD family (Hultgren *et al.*, 1993; Holmgen & Bränden, 1989; Kuehn *et al.*, 1993) is able to perform general functions. We thus suggest that the general chaperone in the periplasm of *E. coli*, being able to assist scTCR folding, still has to be discovered. Folding in the periplasm of *E. coli* (reviewed by Wülfing & Plückthun, 1994) might resemble the highly guided folding process in the endoplasmic reticulum of eukaryotic cells more than one would expect from the postulated absence of ATP in the periplasm.

We thank Dr Thomas Hünig for the TCR clone CR15, Michael Tacke for growing the FAZ hybridoma, Dr Hanspeter Pircher and his co-workers for the TCR clone P14 and for large amounts of the F23.1 antibody and Dr Pamela Ohashi for the TCR clone 8/10-2. Dr John Beckwith we thank for plasmid p12-7, and Drs Peter Pack, Achim Knappik, Arne Skerra and Hanspeter Pircher for helpful discussions. Financial support was obtained from the Bundesministerium für Forschung und Technologie and the Deutsche Forschungsgemeinschaft (Schwerpunkt Hitzestressproteine), and C.W. received a predoctoral fellowship from the Stiftung Stipendienfonds des Fonds der chemischen Industrie.

References

Allison, J. P. & Larnier, L. L. (1987). The structure, function, and serology of the T cell antigen receptor complex. *Annu. Rev. Immunol.* **5**, 503–540.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, **181**, 223–230.

Agard, D. A. (1993). To fold or not to fold. *Science*, **260**, 1903–1904.

Bardwell, J. C. A., McGovern, K. & Beckwith, J. (1991). Identification of a protein required for disulfide bond formation *in vivo*. *Cell*, **67**, 581–589.

Bardwell, J. C. A., Lee, J.-O., Jander, G., Martin, N., Belin, D. & Beckwith, J. (1993). A pathway for disulfide bond formation *in vivo*. *Proc. Nat. Acad. Sci., U.S.A.* **90**, 1038–1042.

Buchner, J., Schmidt, M., Fuchs, M., Jaenicke, R., Rudolph, R., Schmidt, F. X. & Kiefhaber, T. (1991). GroE facilitates the folding of citrate synthetase by suppressing aggregation. *Biochemistry*, **30**, 1586–1591.

Clevers, H., Alarcon, B., Wileman, T. & Terhorst, C. (1988). The T cell receptor/CD3 complex: a dynamic protein ensemble. *Annu. Rev. Immunol.* **6**, 629–662.

Chothia, C., Boswell, D. R. & Lesk, A. M. (1988). The outline structure of the T-cell $\alpha\beta$ receptor. *EMBO J.* **7**, 3745–3755.

Compton, L. A., Davis, J. M., MacDonald, J. R. & Baechinger, H. P. (1992). Structural and functional characterization of *Escherichia coli* peptidyl-prolyl *cis-trans* isomerases. *Eur. J. Biochem.* **206**, 927–934.

Craig, E. A. (1993). Chaperones: helpers along the pathways to protein folding. *Science*, **260**, 1902–1903.

Creighton, T. E. (1988). Toward a better understanding of protein folding pathways. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 5082–5086.

de Boer, H. A., Comstock, L. J. & Vasser, M. (1983). The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. *Proc. Nat. Acad. Sci., U.S.A.* **80**, 21–25.

Dubé, S., Fisher, J. W. & Powell, J. S. (1988). Glycosylation at specific sites of erythropoetin is essential for biosynthesis, secretion, and biological function. *J. Biol. Chem.* **263**, 17516–17521.

Folz, R. J. & Gordon, J. L. (1987). Computer-assisted predictions of signal peptidase processing sites. *Biochem. Biophys. Res. Commun.* **146**, 870–877.

Ganju, R. K., Smiley, S. T., Bajorath, J., Novotny, J. & Reinherz, E. L. (1992). Similarity between fluorescein-specific T-cell receptor and antibody in chemical details of antigen recognition. *Proc. Nat. Acad. Sci., U.S.A.* **89**, 11552–11556.

Gething, M.-J. & Sambrook, J. (1992). Protein folding in the cell. *Nature (London)*, **355**, 33–45.

Glockshuber, R., Schmidt, T. & Plückthun A. (1992). The disulfide bonds in antibody variable domains: effects of

stability, folding *in vitro* and functional expression in *Escherichia coli. Biochemistry,* **31**, 1270–1279.

Goldberg, M. E., Rudolph, R. & Jaenicke, R. (1991). A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg white lysozyme. *Biochemistry,* **30**, 2790–2797.

Grossman, A. D., Erickson, J. W. & Gross, C. A. (1984). The *htp*R gene product of *E. coli* is a sigma factor for heat-shock promoters. *Cell,* **38**, 383–390.

Grossman, A. D., Strauss, D. B., Walter, W. A. & Gross, C. A. (1987). σ^{32} synthesis can regulate the synthesis of heat shock proteins in *Escherichia coli. Genes Develop.* **1**, 179–184.

Harlow, E. & Lane, D. (1988). *Antibodies. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Hasan, N. & Szybalski, W. (1987). Control of cloned gene expression by promoter inversion *in vivo*: construction of improved vectors with a multiple cloning site and the p_{tac} promoter. *Gene,* **56**, 145–151.

Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988). Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature (London),* **333**, 330–334.

Hendrick, J. P. & Hartl, F.-U. (1993). Molecular chaperone functions of heat-shock proteins. *Annu. Rev. Biochem.* **62**, 349–384.

Holmgren, A. & Bränden, C.-I. (1989). Crystal structure of chaperone PapD reveals an immunoglobulin fold. *Nature (London),* **342**, 248–251.

Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Cerretti, D. P., Urdal, D. L. & Conlon, P. J. (1988). A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/Technology,* **6**, 1204–1210.

Horovitz, A. & Fersht, A. R. (1992). Co-operative interactions during protein folding. *J. Mol. Biol.* **224**, 733–740.

Hünig, T. R. & Bevan, M. J. (1982). Antigen recognition by cloned cytotoxic T lymphocytes follows rules predicted by the altered-self hypothesis. *J. Exp. Med.* **155**, 111–125.

Hultgren, S. J., Abraham, S., Caparon, M., Falk, P., St. Geme, J. W. III & Normark, S. (1993). Pilus and nonpilus bacterial adhesins: assembly and function in cell recognition. *Cell,* **73**, 887–901.

Hunkapiller, T. & Hood, L. (1989). Diversity of the immunoglobulin gene superfamily. *Advan. Immunol.* **44**, 1–63.

Hurtley, S. M. & Helenius, A. (1989). Protein oligomerization in the endoplasmatic reticulum. *Annu. Rev. Cell. Biol.* **5**, 277–307.

Iwamoto, A., Ohashi, P. S., Pircher, H., Walker, C. L., Michalopoulos, E. E., Rupp, F., Hengartner, H. & Mak, T. W. (1987). T cell receptor variable gene usage in a specific cytotoxic T cell response. *J. Exp. Med.* **165**, 591–600.

Jaenicke, R. (1991). Protein folding: local structures, domains, subunits, and assemblies. *Biochemistry,* **30**, 3147–3161.

Kamitani, S., Akiyama, Y. & Ito, K. (1992). Identification and characterization of an *Escherichia coli* gene required for the formation of correctly folded alkaline phosphatase, a periplasmic enzyme. *EMBO J.* **11**, 57–62.

Kim, P. S. & Baldwin, R. L. (1990). Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **59**, 631–660.

Knappik, A. & Plückthun A. (1994) An improved tag based on the FLAG peptide for the detection and purification of antibody fragments in *Escherichia coli. Biotechniques*, in the press.

Knappik, A., Krebber, C. & Plückthun, A. (1993). The effect of folding catalysis on the *in vivo* folding process of different antibody fragments expressed in *Escherichia coli. Bio/Technology,* **11**, 77–83.

Kuehn, M. J., Ogg, D. J., Kihlberg, J., Slonim, L. N., Flemmer, K., Bergfors, T. & Hultgren, S. J. (1993). Structural basis of pilus subunit recognition by the PapD chaperone. *Science,* **262**, 1234–1241.

Kurucz, I., Jost, C. R., George, A. J. T., Andrew, S. M. & Segal, D. M. (1993). A bacterially expressed single-chain Fv construct from the 2B4 T-cell receptor. *Proc. Nat. Acad. Sci., U.S.A.* **90**, 3830–3834.

Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M. K. & Hartl, F. U. (1992). Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature (London),* **356**, 683–689.

Lattman, E. E, & Rose, G. D. (1993). Protein folding—what's the question. *Proc. Nat. Acad. Sci., U.S.A.* **90**, 439–441.

Lindner, P., Guth, B., Wülfing, C., Krebber, C., Steipe, B., Müller, F. & Plückthun, A. (1993). Purification of native proteins from the cytoplasm and periplasm of *Escherichia coli* using IMAC and histidine tails: a comparison of proteins and protocols. *Methods: a companion of Methods Enzymol.* **4**, 41–56.

Liu, J. & Walsh, C. T. (1990). Peptidyl-prolyl *cis-trans* isomerase from *Escherichia coli*: a periplasmic-homolog of cyclophilin that is not inhibited by cyclosporin A. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 4028–4032.

Machamer, C. E. & Rose, J. K. (1988). Vesicular Stomatitis Virus G proteins with altered glycosylation sites display temperature-sensitive intracellular transport and are subject to aberrant intermolecular disulfide bonding. *J. Biol. Chem.* **263**, 5955–5960.

Manning-Krieg, U. C., Scherer, P. E. & Schatz, G. (1991). Sequential action of mitochondrial chaperones in protein import into the matrix. *EMBO J.* **10**, 3273–3280.

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A. L. & Hartl, F.-U. (1991). Chaperonin-mediated protein folding at the surface of GroEL through a "molten globule" like intermediate. *Nature (London),* **352**, 36–42.

Marshall, R. D. (1974). The nature and metabolism of the carbohydrate-peptide linkages of glycoproteins. *Biochem. Soc. Symp.* **40**, 17–26.

Matthews, C. R. (1993). Pathways of protein folding. *Annu. Rev. Biochem.* **62**, 653–683.

Matzuk, M. M. & Boime, I. (1988). The role of asparagine-linked oligosaccharides of the α subunit in the secretion and assembly of the human chorionic gonadotropin. *J. Cell Biol.* **106**, 1049–1059.

McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature (London),* **348**, 552–554.

Mitraki, A. & King, J. (1992). Amino acid substitutions influencing intracellular protein folding pathways. *FEBS Letters* **307**, 20–25.

Morimoto, R. I., Tissieres, A. & Georgopoulos, C., eds (1990). *Stress Proteins in Biology and Medicine*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Novotny, J., Ganju, R. K., Smiley, S. T., Hussey, R. E., Luther, M. A., Recny, M. A., Siliciano, R. F. & Reinherz, E. L. (1991). A soluble, single-chain T-cell receptor fragment endowed with antigen-combining properties. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 8646–8650.

Pace, C. N. (1991). Conformational stability of globular proteins. *Trends Biochem. Sci.* **15**, 14–17.

Pack, P. & Plückthun, A. (1992). Miniantibodies: use of amphiphatic helices to produce functional, flexibly linked dimeric Fv fragments with high avidity in *Escherichia coli. Biochemistry,* **31**, 1579–1584.

Pack, P., Kujau, M., Schroeckh, V., Knüpfer, U., Wenderoth, R., Riesenberg, D. & Plückthun, A. (1993). Improved bivalent miniantibodies, with identical avidity as whole antibodies, produced by high cell density fermentation of *Escherichia coli. Bio/Technology,* **11**, 1271–1277.

Pelham, H. R. B. (1989). Control of protein exit from the endoplasmatic reticulum. *Annu. Rev. Cell. Biol.* **5**, 1–23.

Pircher, H., Michalopoulos, E. E., Iwamoto, A., Ohashi, P. S., Baenzinger, J., Hengartner, H., Zinkernagel, R. M. & Mak, T. W. (1987). Molecular analysis of the antigen receptor of virus-specific cytotoxic T cells and identification of a new Vα family. *Eur. J. Immunol.* **17**, 1843–1846.

Plückthun, A. (1992). Mono- and bivalent antibody fragments in *Escherichia coli*: Engineering, folding and antigen binding. *Immunol. Rev.* **130**, 151–188.

Podhajska, A. J., Hasan, N. & Szybalski, W. (1985). Control of cloned gene expression by promoter inversion *in vivo*: construction of the heat-pulse-activated *att-nut*L-*p-att-N* module. *Gene,* **40**, 163–168.

Prowald, K. (1990). Klonierung und charakterisierung der exprimierten TCR gene des cytotoxischen T zellklons CR15. PhD Thesis, Univ. of Munich (LMU).

Radford, S. E., Dobson, C. M. & Evans, P. A. (1992). The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature (London),* **358**, 302–307.

Rosen, B. P. (1987). ATP-coupled solute transport systems. In *Escherichia coli* and *Salmonella typhimurium* (Neidhardt, F. C., ed.), pp. 760–767. American Society for Microbiology, Washington, DC, USA.

Rudolph, R. (1990). Renaturation of recombinant, disulfide-bonded proteins from "inclusion bodies". In *Modern Methods in Protein and Nucleic Acid Research* (Tschesche, H., ed.), pp. 149–171. de Gruyter, Berlin, Germany.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schmidt, T. (1991). Untersuchungen zur Bedeutung der Disulfidbindungen für die Stabilität der variablen Domänen des Antikörpers McPC603. Diploma Thesis, Univ. of Munich (LMU).

Skerra, A., Pfitzinger, I. & Plückthun, A. (1991). The functional expression of antibody Fv fragments in *Escherichia coli*: improved vectors and a generally applicable purification technique. *Bio/Technology,* **9**, 273–278.

Soo Hoo, W. F., Lacy, M. J., Denzin, L. K., Voss, E. W. Jr, Hardman, K. D. & Kranz, D. M. (1992). Characterization of a single-chain T-cell receptor expressed in *Escherichia coli. Proc. Nat. Acad. Sci., U.S.A.* **89**, 4759–4763.

Staerz, U. D., Rammensee, H.-G., Benedetto, J. D. & Bevan, M. J. (1985). Characterization of a murine monoclonal antibody specific for an allotypic determinant on T cell antigen receptor. *J. Immunol.* **134**, 3994–4000.

von Heijne, G. (1990). The signal peptide. *J. Membrane Biol.* **115**, 195–201.

Ward, E. S. (1991). Expression and secretion of T-cell receptor Vα and Vβ domains using *Escherichia coli* as a host. *Scand. J. Immunol.* **34**, 215–220.

Ward, E. S. (1992). Secretion of T cell receptor fragments from recombinant *Escherichia coli* cells. *J. Mol. Biol.* **224**, 885–890.

Weissman, J. S. & Kim, P. S. (1991). Reexamination of the folding of BPTI: predominance of native intermediates. *Science,* **253**, 1386–1393.

Wülfing, C. & Plückthun, A. (1994). Protein folding in the periplasm of *Escherichia coli. Mol. Microbiol.* **12**, 685–692.

Wülfing, C. & Plückthun, A. (1993). A versatile and highly repressible *Escherichia coli* expression system based on invertible promoters: expression of a gene encoding a toxic product. *Gene,* **136**, 199–203.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene,* **33**, 103–119.

Zahn, R., Spitzfaden, C., Ottiger, M., Wüthrich, K. & Plückthun, A. (1994). Destabilization of the complete protein secondary structure on binding to the chaperone GroEL. *Nature (London),* **368**, 261–265.

Zinkernagel, R. M. & Doherty, P. C. (1979). MHC-restricted cytotoxic T cells: studies on the biological role of polymorphic major transplantation antigens determining T-cell restriction-specificity, function, and responsiveness. *Advan. Immunol.* **27**, 51–177.

Edited by I. B. Holland

(Received 23 March 1994; accepted 4 July 1994)