The importance of surface loops for stabilizing an eightfold $\beta\alpha$ barrel protein

ROMAN URFER AND KASPER KIRSCHNER

Abteilung Biophysikalische Chemie, Biozentrum der Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

(RECEIVED July 29, 1991; ACCEPTED September 3, 1991)

Abstract

An important step in understanding how a protein folds is to determine those regions of the sequence that are critical to both its stability and its folding pathway. We chose phosphoribosyl anthranilate isomerase from *Escherichia coli*, which is a monomeric representative of the $(\beta\alpha)_8$ barrel family of proteins, to construct a variant that carries an internal tandem duplication of the fifth $\beta\alpha$ module. This $(\beta\alpha)_9$ variant was enzymically active and therefore must have a wild-type $(\beta\alpha)_8$ core. It had a choice a priori to fold to three different folding frames, which are distinguished by carrying the duplicated segment as an insert into one out of three different loops. Steady-state kinetic constants, the fluorescence properties of a crucial tryptophan residue, and limited proteolysis showed that the stable $(\beta\alpha)_9$ variant carries the insertion between β -strand 5 and α -helix 5. This preference can be explained by the important role of loops between α helices and β strands in stabilizing the structure of the enzyme.

Keywords: eightfold $\beta\alpha$ barrel protein; loops; mutagenesis; protein evolution; protein folding

Close packing of secondary structural elements is necessary for the correct folding and stability of proteins (Finkelstein & Ptitsyn, 1988; Chothia & Finkelstein, 1990), but the loops that connect these elements may also contribute toward stabilizing the folded state (Leszczynsky & Rose, 1986; Thornton et al., 1988). Mutational analysis using recombinant DNA technology is a versatile approach for probing the importance either of individual residues or of the entire sequence of such loops (Riechmann et al., 1988; Fetrow et al., 1989; Hynes et al., 1989). Fragmentation of the sequence in a given loop region (Taniuchi et al., 1986), as well as circular permutation of the amino acid sequence, which also involves cleavage of a surface loop (Goldenberg & Creighton,

1983; Luger et al., 1989; Buchwalder et al., 1991), are alternative approaches.

It is desirable to assess quantitatively the stabilizing effect of a given loop. However, it is difficult to carry out such comparative measurements commensurately, either with fragment complexes or with circularly permuted variants. Therefore, we sought another approach to determine the relative importance of surface loops in stabilizing the structure of a protein. It occurred to us that the tandem duplication of an internal $\beta\alpha$ module of a $(\beta\alpha)_8$ barrel protein would create a $(\beta\alpha)_9$ variant that has the choice to fold into several isomeric folding frames. Because each folding frame has the same native $(\beta\alpha)_8$ core, the thermodynamic stabilities should be comparable except for the perturbation caused by inserting the duplicated segment into one of the native loops. Thus, the protein should reveal automatically its most favorable conformation.

Here we describe the construction and expression of a mutant of the *trpF* gene (coding for phosphoribosyl anthranilate [PRA] isomerase) that carries such an internal tandem duplication. The gene product adopts only one out of three possible folding frames. The particular choice suggests that the loops between α helices and β strands are more important to the stability of PRA isom-

Reprint requests to: Kasper Kirschner, Abteilung Biophysikalische Chemie, Biozentrum der Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland.

Abbreviations: ANT, anthranilic acid; BPTI, bovine pancreatic trypsin inhibitor; CD, circular dichroism; DEAE, diethylaminoethyl; DTE, dithioerythrol; EDTA, (ethylenedinitrilo)-tetraacetic acid; GuCl, guanidine hydrochloride; HPLC, high performance liquid chromatography; IPTG, isopropyl- β -D-thiogalactopyranoside; PRPP, phosphoribosyl pyrophosphate; PRT, anthranilate phosphoribosyl transferase; PVDF, polyvinylidene difluoride; rCdRP, reduced 1-(2-carboxyphenylamino)-1-deoxyribose-5-phosphate; SDS-PAGE, sodium dodecylsulfate polyacrylamide gel electrophoresis.

erase than the loops between β strands and α helices. This rule also explains the relative stabilities of two circularly permuted as well as a terminally duplicated variant of PRA isomerase.

Results

Tactical considerations

Phosphoribosyl anthranilate isomerase possesses an ideal folding pattern for the insertional mutagenesis in mind. It has the eightfold $\beta\alpha$ barrel structure (Priestle et al., 1987) common to about 20 different enzymes (Farber & Petsko, 1990; Wilson et al., 1991). Our previous studies of circular permutation (Luger et al., 1989) and terminal extension by a duplicated region (Luger et al., 1990) were carried out with the naturally occurring monomeric and monofunctional PRA isomerase from yeast. In this work we turned to the artificially separated monomeric and monofunctional PRA isomerase domain of the bifunc-

tional indoleglycerol phosphate synthase:PRA isomerase from *Escherichia coli* because its structure, but not the structure of the yeast enzyme, is known at 2.0 Å resolution (Wilmanns et al., 1991b). The sequence of the PRA isomerase domain begins with the first natural residue of β_1 (E256), preceded by M254' and L255', and ends at the natural carboxyl terminus. Because the steady-state kinetic constants of the separated and naturally fused domains of PRA isomerase are practically indistinguishable (Eberhard, 1990b), we assume that the two structures are essentially the same (Fig. 1A; Kinemage 1). In formal terms the structure of PRA isomerase consists of a sequence of eight unique $\beta\alpha$ modules (Fig. 1B), where several of the loops between β strands and α helices carry conserved residues (Niermann & Kirschner, 1991; Wilmanns et al., 1991b) that are involved in the catalytic mechanism (unpubl.). We reasoned that tandem duplication of any one of the $\beta\alpha$ modules would allow this $(\beta\alpha)_9$ variant to decide which one of the two modules is to be incorporated preferably into the $(\beta\alpha)_8$ protein core (Fig. 2).

Fig. 1. The structure of PRA isomerase from *E. coli*, a limit $(\beta\alpha)_8$ barrel protein. **A:** Schematic ribbon diagram. The β strands are numbered sequentially from the amino-terminus. The carboxyl-terminal ends of the β strands (arrows) and loops at the active site face the reader. **B:** Schematic representation of supersecondary structure. The $(\beta\alpha)$ modules are rolled out in the plane. They are boxed and numbered sequentially from the amino-terminus. Arrows represent β strands and spirals α helices. Each secondary structural element is connected by a semicircle that represents a surface loop. Active site residues are located at the carboxyl-terminal ends of β strands and the following loops.

We selected $(\beta\alpha)^5$ for duplication for the following reasons: (1) Neither the carboxyl-terminus of β_5 nor the loop residues between β_5 and α_5 carry any conserved, i.e., potentially catalytic, residues (Kinemage 1); (2) β_5 carries a tryptophan residue (W356) that is almost completely buried in the core of the central β barrel and is a useful indicator of solvent accessibility (Kinemage 2); and

(3) α_5 carries two glutamate residues (E363 and E369) that are potential cleavage sites for V8-protease (Kinemage 2). The refined structure (illustrated by Kinemage 1) shows that the region between the end of β_5 and the beginning of β_6 has an irregular structure rather than an α helix. Because the region from G362 to F370 is predicted to be α -helical on average for a set of 14 orthologous sequences (Niermann & Kirschner, 1991), we refer to it simply as α_5 .

Lasters et al. (1988) have argued persuasively that only eightfold $\beta\alpha$ barrels are stable. Detailed studies of the packing of side chains in the cores of several $(\beta\alpha)_8$ proteins by Lesk et al. (1989) support this view. On one hand, even if the $(\beta\alpha)_9$ variant were to adopt a labile ninefold $(\beta\alpha)$ barrel (Fig. 2A), the active site would be disrupted. On the other hand it is unlikely that the $(\beta\alpha)_9$ variant would form a $(\beta\alpha)_8$ barrel core from eight contiguous $\beta\alpha$ modules. Because each of the eight $\beta\alpha$ modules of PRA isomerase is unique, the packing of internal side chains would not be optimal. Moreover, one of these folding frames would have $(\beta\alpha)^1$ and the other $(\beta\alpha)^8$ as a terminal appendage to the central core. Each of these structures would be enzymically inactive because both $(\beta\alpha)^1$ (unpubl.) and $(\beta\alpha)^8$ (Wilmanns et al., 1991a) carry catalytically important residues.

It is more likely that the $(\beta\alpha)_9$ variant would fold to a core comprising the sequence of the eight natural $\beta\alpha$ modules and an insertion of the duplicated region into one of three surface loops of the original $(\beta\alpha)_8$ molecule (Kinemage 2). Two isomeric folding frames would carry the insert in a loop between an α helix and a β strand (Fig. 2B, C). The third (Fig. 2D) would carry a transposed version of the duplicated module α_5 -linker- β_5 as an insert in the loop between β_5 and α_5 . This insert would be at the carboxyl-terminal end of the β barrel that harbors the active site.

In order to distinguish between the three possible folding frames of $(\beta\alpha)_9$, we decided to exploit the dependence of the fluorescence of tryptophan side chains on their environment. Because $(\beta\alpha)_8$ wild type (wt) contains two Trp residues (W356 in β_5 and W391 in α_6) and the $(\beta\alpha)_9$ WW' contains three (W356, W356', and W391, cf. Fig. 4), it was necessary to determine the contribution of each tryptophan residue to the total fluorescence of these proteins. For this purpose we replaced individual tryptophan residues with phenylalanine, which does not fluoresce significantly and has a hydrophobic side chain, albeit of somewhat smaller volume than that of tryptophan.

Constructions

The coding sequence of the PRA isomerase domain ($[\beta\alpha]_8$ wt) was first subcloned into the mutagenesis vector pMc5-10 (Stanssens et al., 1989; Figs. 3A, 5A). Then a new BglIII site was introduced into one copy of the cod-

Fig. 2. The $(\beta\alpha)_9$ variant has a choice to adopt three different folding frames with the same $(\beta\alpha)_8$ barrel core. Secondary structural elements of $(\beta\alpha)^5$ (red) and $(\beta\alpha)^{5'}$ (green). Artificial linker connecting α_5 with β_5 (black, dashed). **A:** Hypothetical ninefold β -barrel core. Hypothetical eightfold barrel cores with duplicated sequence inserted into: **B:** α_4 loop β_5 ; **C:** α_5 loop β_6 ; and **D:** β_5 loop α_5 .

ing sequence, namely in the region coding between α_4 and β_5 (Fig. 3C, giving the mutant I_1). Next a new BamHI site was introduced into another copy of the sequence in the region coding between α_5 and β_6 (Fig. 3C, giving the mutant I_2). By ligating the appropriate (1.1-kb) EcoRI/BamHI fragment of I_2 and the appropriate (0.4-kb) BglII/HindIII fragment of I_1 into pMc5-10, we obtained the gene of the $(\beta\alpha)_9$ wild-type variant (Fig. 3E, $[\beta\alpha]_9$ WW') in an expressible form. The amino acid sequence of $(\beta\alpha)_9$ WW' encompassing the region between α_4 and β_6 is given in Figure 4. To emphasize the internally duplicated sequence we have used the same numbering

for both regions and arbitrarily assigned primes to the carboxyl-terminal one. Note that neither one nor the other region can be defined a priori as the insert. The sequence between Q371 and V353' is the newly introduced unnatural linker between α_5 and β_5 .

Then we replaced W365 of the $(\beta\alpha)_8$ wt by phenylalanine (Fig. 3B, $[\beta\alpha]_8$ W356F) and introduced the BglII and BamHI sites as before (Fig. 3D, giving the analogous mutants I_3 and I_4). Finally, we combined the appropriate fragments as described above to obtain point mutants of $(\beta\alpha)_9$ WW' (Fig. 3E) in which the Trp residue either in strand β_5 ($[\beta\alpha]_9$ FW') or in strand β_5' ($[\beta\alpha]_9$ WF')

Fig. 3. Construction of the $(\beta\alpha)_8$ and $(\beta\alpha)_9$ variants. Schematic flow-sheet. See text for details. □, Coding sequence. ▤, Modules $(\beta\alpha)^5$ or $(\beta\alpha)^{5'}$. W, Trp residue, or F, Phe residue at position 356 or 356'. Partial amino acid and nucleotide sequences. Replaced amino acids are in bold type; new restriction sites are underlined. **A:** $(\beta\alpha)_8$ wt. **B:** $(\beta\alpha)_8$ W356F. **C:** I_1 and I_2 are mutants of $(\beta\alpha)_8$ wt with a new BglII or BamHI site, respectively. **D:** I_3 and I_4 are mutants of $(\beta\alpha)_8$ W356F with a new BglII or BamHI site, respectively. **E:** Assembly and ligation of the appropriate EcoRI/BamHI fragments of I_2 or I_4 and the BglII/HindIII fragments of I_1 or I_3 into the multiple cloning site of the vector pMc5-10 leads to three variants $(\beta\alpha)_9$ WW', $(\beta\alpha)_9$ WF', and $(\beta\alpha)_9$ FW' in expressible form. **F:** $(\beta\alpha)_8$ wt; **G:** $(\beta\alpha)_8$ W391F.

Fig. 4. Partial amino acid sequence of $(\beta\alpha)_9$ WW'. Assignments of secondary structure according to Niermann and Kirschner (1991) and Wilmanns et al. (1991b). The residues used as probes for establishing the preferred folding frame are boxed (Trp residues 356 and 356') or encircled (Glu residues 363, 369, 363', and 369'). L352' is the only unnatural residue (bold). Q371 to L352' is the new linker between α_5 and β_5 ' (cf. Figs. 2, 10).

was replaced by phenylalanine. We further constructed a variant of $(\beta\alpha)_8$ wt (Fig. 3F), where W391 was exchanged for phenylalanine (Fig. 3G, $[\beta\alpha]_8$ W391F).

To express the $(\beta\alpha)_8$ variants at high levels we used the expression vector pDS56/RBSII/SphI (Certa et al., 1986). We first introduced a new SphI restriction site into the $(\beta\alpha)_8$ wt gene (Fig. 5A) at the position of the start codon, without altering any other codons (Fig. 5B). The 0.7-kb

Table 1. $(\beta\alpha)_8$ and $(\beta\alpha)_9$ variants used in this study

Number of $(\beta\alpha)$ modules	Residue at position			Designation
	356	356'	391	
8	W	—	W	$(\beta\alpha)_8$ wt
8	F	—	W	$(\beta\alpha)_8$ W356F
8	W	—	F	$(\beta\alpha)_8$ W391F
9	W	W	W	$(\beta\alpha)_9$ WW'
9	W	F	W	$(\beta\alpha)_9$ WF'
9	F	W	W	$(\beta\alpha)_9$ FW'

SphI/HindIII fragment was then ligated into the multiple cloning site of the vector pDS56/RBSII/SphI (Fig. 5C) to give the new expression vector $(\beta\alpha)_8$ wt/pDS (Fig. 5D). From this plasmid both $(\beta\alpha)_8$ point mutations were constructed in a highly expressible form by exchanging the 0.4-kb ClaI/HindIII fragment of $(\beta\alpha)_8$ wt/pDS for the corresponding mutated fragment of $(\beta\alpha)_8$ /pMc5-10. The designations of all variants are given in Table 1.

Expression and purification

All $(\beta\alpha)_9$ and $(\beta\alpha)_8$ variants were expressed in the strain W3110 [*trpR*, Δ *trpEA2*] and isolated from inclusion bodies. The $(\beta\alpha)_9$ variants were expressed constitutively from the vector pMc5-10, and the $(\beta\alpha)_8$ variants were expressed from the vector pDS56/RBSII/SphI after induction with 1 mM IPTG.

Fig. 5. The plasmids used for mutagenesis and expression of the $(\beta\alpha)_8$ variants. □, Coding sequences; cat, gene for chloramphenicol acetyltransferase; bla, gene for β -lactamase; am, amber stop codon used for selection of the mutated DNA strand; ■, promoter/operator region; ▨, ribosome binding site; |, restriction site. **A:** Mutagenesis vector containing the sequence of $(\beta\alpha)_8$ wt. **B:** Silent mutation of A with a new SphI site. **C:** Expression vector with multiple cloning site. **D:** B subcloned in C (cf. Fig. 3F). **E:** $(\beta\alpha)_8$ W391F subcloned into C (cf. Fig. 3G).

Up to 100% of the $(\beta\alpha)_9$ variants and up to 90% of the $(\beta\alpha)_8$ variants were found in the insoluble fraction of the cell homogenate (Table 2; Fig. 6A, lane 2; Fig. 6B, lane 2). The pellets were dissolved in GuCl, refolded by dialysis, and purified as described in Materials and methods. The results of the purification of the $(\beta\alpha)_8$ wt and the $(\beta\alpha)_9$ WW', which are representative for all variants, are summarized in Table 2. The various preparations were at least 90% pure as judged by SDS-PAGE (Fig. 6C). All native proteins were soluble to about 3 mg protein/mL without aggregation.

The $(\beta\alpha)_9$ variants were typically obtained with yields of 1.5–6 mg of pure protein from 1 g of wet cells. For the $(\beta\alpha)_8$ variants, the yields were 9–15 mg of pure protein

from 1 g of wet cells, due to the high expression efficiency of pDS56/RBSII/SphI.

Protein-folding studies

Native $(\beta\alpha)_9$ WW' is a monomer with the expected molecular weight (M_r , ~23,000) as assured by gel filtration (data not shown). The far-UV CD spectra of $(\beta\alpha)_9$ WW' and $(\beta\alpha)_8$ wt indicate that $(\beta\alpha)_9$ WW' has about the same content of apparent secondary structure on a per residue basis as $(\beta\alpha)_8$ wt (Fig. 7A). Because the segment between β_5 and β_6 is not an α helix, the difference between these spectra cannot give unequivocal information

Fig. 6. Purification of the PRA isomerase variants $(\beta\alpha)_8$ wt and $(\beta\alpha)_9$ WW'. Polyacrylamide gel electrophoresis in the presence of SDS. See Table 2 for quantitative data. M, marker proteins with indicated $10^{-3} M_r$ values. **A:** $(\beta\alpha)_8$ wt. **B:** $(\beta\alpha)_9$ WW'. Lanes: 1, homogenate supernatant; 2, homogenate pellet; 3, fraction refolded from pellet; 4, DEAE-Sepharose eluate. **C:** Purified proteins $(\beta\alpha)_8$ wt and $(\beta\alpha)_9$ WW'. Lanes: 1, $(\beta\alpha)_8$ wt; 2, $(\beta\alpha)_9$ WW'; 3, mixture of $(\beta\alpha)_8$ wt and $(\beta\alpha)_9$ WW'.

Table 2. Purification of $(\beta\alpha)_8$ wt and $(\beta\alpha)_9$ WW'^a

Protein	Purification step	Total activity (units)	Total protein (mg)	Specific activity (units/mg)	Yield (%)
$(\beta\alpha)_8$ wt	Homogenate supernatant	1,650	860	6	10
	Refolded pellet	15,100	810	19	90
	DEAE-Sepharose eluate	17,100	198	86	102
$(\beta\alpha)_9$ WW'	Homogenate supernatant	0.6	2,454	0.0003	1
	Refolded pellet	83	400	0.2	99
	DEAE-Sepharose eluate	59	19	3.1	70

^a See Materials and methods for details. Starting material was 15 g of wet cell paste in both cases.

Fig. 7. Reversible unfolding transitions of the $(\beta\alpha)_8$ and $(\beta\alpha)_9$ variants. Buffer: 50 mM potassium phosphate buffer at pH 7.5 supplemented with 0.1 mM EDTA and 0.1 mM DTE at 22 °C. Unfolding transitions monitored at 222 nm. The linear change of θ_{222} above 4.0 M was extrapolated to 0 M GuCl. f_N is the apparent fraction of native protein and represents the deviation of θ_{222} from this baseline normalized by the value at 0 M GuCl concentration. See Materials and methods for details. **A:** Normalized far-UV CD spectra of $(\beta\alpha)_8$ wt (---) and $(\beta\alpha)_9$ WW' (—). **B:** Unfolding (○) and refolding (●) of $(\beta\alpha)_8$ wt and unfolding transitions of the point mutants $(\beta\alpha)_8$ W356F (----) and $(\beta\alpha)_8$ W391F (.....). **C:** Unfolding (□) and refolding (■) of $(\beta\alpha)_9$ WW' and unfolding transitions of the point mutants $(\beta\alpha)_9$ WF' (----) and $(\beta\alpha)_9$ FW' (.....).

of whether the insert possesses secondary structure or not.

Both proteins unfold reversibly in GuCl as shown by the superposition of the equilibrium unfolding and the refolding transitions monitored by far-UV CD. At equilibrium the unfolding of $(\beta\alpha)_9$ WW' (Fig. 7C) shows a small shift of the transition midpoint (about 0.1 M) toward lower GuCl concentrations compared to the $(\beta\alpha)_8$ wt (Fig. 7B). The transition midpoints of all $(\beta\alpha)_8$ and $(\beta\alpha)_9$ variants lie between 1.3 M and 1.5 M.

Steady-state kinetics

Table 3 shows that the W356F replacement reduces the k_{cat} value of the $(\beta\alpha)_8$ wt about 5-fold and increases the K_M^{PRA} value about 13-fold, so that the efficiency constant $k_{\text{cat}}/K_M^{\text{PRA}}$ decreases 61-fold. In $(\beta\alpha)_8$ wt the indolyl side chain of W356 lies parallel to the barrel axis, deeply buried in the interior of the barrel at its amino-terminal end, i.e., away from the active site (Wilmanns et al., 1991b). The exchange of the indolyl for the phenyl side chain ap-

parently alters the packing of the hydrophobic core such that the enzymic activity is affected indirectly (Lim & Sauer, 1991).

The $(\beta\alpha)_9$ variant carrying the wild-type residue W356 in both $(\beta\alpha)^5$ modules (i.e., $[\beta\alpha]_9$ WW') has a 25-fold lower value of k_{cat} than $(\beta\alpha)_8$ wt, but the same K_M^{PRA} value. Spectrophotometric titration of both $(\beta\alpha)_8$ wt and $(\beta\alpha)_9$ WW' with the chromophoric substrate analogue rCdRP (Bisswanger et al., 1979) gave almost identical difference spectra and $K_d^{\text{rCdRP}} = 14 \mu\text{M}$ and $24 \mu\text{M}$, respectively (data not shown). By inference, the $(\beta\alpha)_9$ variant has the $(\beta\alpha)_8$ core of the PRA isomerase $(\beta\alpha)_8$ wt. Therefore the ninefold $(\beta\alpha)$ barrel topology of Figure 2A is excluded.

Most importantly, replacement of W356' in $(\beta\alpha)_9$ WW' by F has no effect, whereas replacement of W356 by F leads to a 17-fold decrease of k_{cat} and a 4-fold increase of K_M^{PRA} . Because alteration of the tryptophan side chain of β -strand 5 decreases the $k_{\text{cat}}/K_M^{\text{PRA}}$ of $(\beta\alpha)_9$ WW' by the same factor (about 60-fold) as the same replacement does in the case of $(\beta\alpha)_8$ wt, we conclude that

Table 3. Steady-state kinetic constants of all $(\beta\alpha)_8$ and $(\beta\alpha)_9$ variants^a

PRA isomerase wt and mutants		Reference in Figure 3	k_{cat} (s^{-1})	$K_{\text{M}}^{\text{PRA}}$ (μM)	$k_{\text{cat}}/K_{\text{M}}^{\text{PRA}}$ ($\mu\text{M}^{-1}\text{s}^{-1}$)	$(k_{\text{cat}}/K_{\text{M}}^{\text{PRA}})_{\text{wt}}/(\text{mutant})$
$(\beta\alpha)_8$	wt	A, F	32.00	4.7	6.8	1
	W356F	B	6.91	62.0	0.11	0.016
	W391F	G	4.81	5.1	0.94	0.138
$(\beta\alpha)_9$	WW'	E	1.20	4.6	0.26	1
	WF'	E	1.14	4.3	0.26	1
	FW'	E	0.07	15.3	0.0044	0.017

^a k_{cat} and K_{M} values were measured as described in Materials and methods. Buffer: 50 mM potassium phosphate at pH 7.5, containing 4 mM K_2MgEDTA and 1 mM DTE; temperature: 25 °C.

the β -strand 5 is integrated into the $(\beta\alpha)_8$ core of $(\beta\alpha)_9$ WW'. Conversely, the tolerance of enzymic activity toward alteration of the tryptophan side chain of β -strand 5' indicates that it is not part of the $(\beta\alpha)_8$ core. These data also show that the W356F mutations in $(\beta\alpha)_9$ WW' do not change the preference of the $(\beta\alpha)_8$ barrel core for a particular conformer.

Fluorescence emission spectra

The wild-type PRA isomerase domain ($(\beta\alpha)_8$ wt) contains two Trp residues (W356 in β_5 and W391 in α_6) and the wild-type $(\beta\alpha)_9$ WW' contains three Trp residues W356 in β_5 , W356' in β_5' (shown in Kinemage 2), and W391 in α_6 (cf. Fig. 4 and Fig. 8A, inset).

Figure 8A shows that the extra Trp residue in $(\beta\alpha)_9$ WW' leads to the expected increase in the relative fluorescence (~50% increase of area) compared to $(\beta\alpha)_8$ wt. The λ_{max} value is nearly indistinguishable from that of $(\beta\alpha)_8$ wt. Both point mutations of $(\beta\alpha)_9$ WW' had the expected decrease in fluorescence. However, the replacement of W356' by F in the variant $(\beta\alpha)_9$ WF' led to a blue-shift and the replacement by W356 by F in the variant $(\beta\alpha)_9$ FW' to a red-shift of the respective emission maximum. Therefore W356' must have a red-shifted emission spectrum due to its exposure to solvent. By the same reasoning W356 must have a blue-shifted spectrum due to its shielding from solvent. These observations further support the hypothesis that β_5 but not β_5' is part of the $(\beta\alpha)_8$ core of $(\beta\alpha)_9$ WW'. Thus, folding frame B is excluded (cf. Fig. 2).

Quenching of Trp fluorescence by acrylamide

The quenching experiments were performed with all $(\beta\alpha)_8$ and $(\beta\alpha)_9$ variants to assess the accessibility of their Trp residues. The Stern–Volmer plot (Eftink & Ghiron, 1976) of Figure 8B shows the quenching effect of acrylamide on the Trp fluorescence of the $(\beta\alpha)_8$ wt and its two variants $(\beta\alpha)_8$ W356F and $(\beta\alpha)_8$ W391F. The

quenching constant (i.e., the slope) of $(\beta\alpha)_8$ wt is significantly smaller than that of *N*-acetyl-Trp-methylester, due to partial shielding of the two Trp residues. Because replacement W356F had no effect on the quenching of the fluorescence, quenching of $(\beta\alpha)_8$ wt must be due mainly to the single Trp W391. The relative insensitivity of the fluorescence of $(\beta\alpha)_8$ W391F toward acrylamide supports this conclusion.

Figure 8C shows that the quenching constant of the $(\beta\alpha)_9$ WW' is increased compared to that of the $(\beta\alpha)_8$ wt. This observation suggests that one of the duplicated Trp residues W or W' is exposed. The mutation of W356 in $(\beta\alpha)_5$ to F in $(\beta\alpha)_9$ FW' does not affect the sensitivity of its fluorescence toward acrylamide, indicating that the removed Trp residue is inaccessible. Therefore W356' must be the exposed residue. Conversely, the mutation of W356' in $(\beta\alpha)_5'$ to F in $(\beta\alpha)_9$ WF' causes a decrease in the slope to a value even smaller than that of the $(\beta\alpha)_8$ wt. This result confirms that the W356' is responsible for increasing the quenching constant of $(\beta\alpha)_9$ WW' over that of $(\beta\alpha)_8$ wt. In summary, these results support the conclusion that β -strand 5 in $(\beta\alpha)_9$ WW' is buried and β -strand 5' is exposed. Because β -strand 5 is buried in both folding frames C and D (cf. Fig. 2), quenching of the Trp fluorescence cannot distinguish between them.

Limited proteolysis with V8-protease

The $(\beta\alpha)_9$ variant carries two glutamate residues in each of the tandem α_5 regions (cf. Fig. 4 and Kinemage 2). In the two remaining alternative folding frames C and D of Figure 2 either α_5' (with E363' and E369') or α_5 (with E363 and E369) is part of the extruded insert. Therefore we used digestion of $(\beta\alpha)_9$ WW' by V8-protease, which cleaves Glu–Xaa bonds, to distinguish between the protected α_5 region of the $(\beta\alpha)_8$ core and the presumably exposed α_5 region of the insert.

Addition of V8-protease to native PRA isomerase ($(\beta\alpha)_8$ wt) caused slow inactivation. As monitored by SDS-PAGE (Fig. 9A), the chain ($M_r = 21,100$) was

Fig. 8. Fluorescence properties of the $(\beta\alpha)_8$ and the $(\beta\alpha)_9$ variants. Fluorescence was excited at 280 nm. The protein concentration was 10 μ M for all samples in 50 mM potassium phosphate at pH 7.5 containing 1 mM EDTA and 1 mM DTE. **A:** Fluorescence spectra. $(\beta\alpha)_8$ wt (---), $(\beta\alpha)_9$ WW' (—), $(\beta\alpha)_9$ WF' (---), and $(\beta\alpha)_9$ FW' (· · · · ·). Inset: Approximate positions of the Trp residues in the $(\beta\alpha)$ modules 5, 5', and 6 of $(\beta\alpha)_9$ WW'. Vertical line indicates emission maximum of $(\beta\alpha)_8$ wt. **B:** Quenching of Trp fluorescence of the $(\beta\alpha)_8$ variants by acrylamide. F_0/F is the quotient of the initial fluorescence of the sample and the fluorescence in presence of acrylamide (Eftink & Ghiron, 1976). Conditions as in A. Fluorescence was excited at 295 nm and emission was observed at 330 nm. $\circ$, $(\beta\alpha)_8$ wt; $\bullet$, $(\beta\alpha)_8$ W356F; Δ , $(\beta\alpha)_8$ W391F; —, *N*-acetyl-Trp-methylester. **C:** Quenching of Trp fluorescence of the $(\beta\alpha)_9$ variants by acrylamide. Conditions as in B. $\square$, $(\beta\alpha)_9$ WW'; $\blacksquare$, $(\beta\alpha)_9$ FW'; $+$, $(\beta\alpha)_9$ WF'; - - - - -, $(\beta\alpha)_8$ wt, cf. Figure 8B.

degraded to a transient fragment of $M_r \sim 16,000$. Its amino-terminal sequence corresponded to that of the wild-type protein (MLENKVCGLTRGQ). We assume that the E415–A416 peptide bond, which is located in α_7 , is the most probable cleavage site, as cleavage at the next nearest neighbors (E369 and E431) would lead to a smaller (M_r , $\sim 12,000$) or a larger (M_r , $\sim 18,000$) fragment.

Addition of V8-protease to native $(\beta\alpha)_9$ WW' caused more rapid inactivation compared to $(\beta\alpha)_8$ wt. The chain (M_r , $\sim 23,000$) was degraded to two intermediate fragments (Fig. 9B; M_r 's, $\sim 19,000$ and $\sim 11,500$), followed by further digestion to a limit fragment with $M_r \sim 9,000$. This time the amino-terminal sequences of the major fragments were determined after 0, 10, 240, and 480 min. The 23-kDa protein had the expected wild-type amino terminus of $(\beta\alpha)_9$ WW', and the 19-kDa intermediate had the same amino-terminus. By analogy to the cleavage pattern for the $(\beta\alpha)_8$ wt, we propose that the 19-kDa band corresponds to the large amino-terminal fragment of $(\beta\alpha)_9$ WW' with E415 at its carboxyl-terminus. The 11.5-kDa band contained two peptides, one with the sequence of the natural amino-terminus (MLENKVCGLTRGQ), and a second peptide with the

new amino-terminal sequence expected for cleavage after E369 in α_5 (FQHVDLVIAWKALSV, cf. Fig. 4). The 9-kDa fragment had the sequence of the natural amino-terminus and is identical to the 9-kDa fragment that accompanies the 16-kDa fragment when $(\beta\alpha)_8$ wt is digested with V8-protease for 240–480 min (Fig. 9A). None of the fragments contained a peptide with a sequence expected for a cut in α_5 (FQHVDKYVLD...).

These results suggest that the $(\beta\alpha)_9$ WW' is cleaved by V8-protease predominantly at E369. That is, E369 in α_5 is more exposed than E369' in α_5 , and the inserted segment is probably disordered. This cleavage pattern is consistent with folding frame D (Fig. 2) but not with folding frame C. The transient appearance of the 19-kDa fragment is compatible with a parallel, albeit slower, cleavage at E415, followed by a second cleavage at E369.

Discussion

As inferred from its enzymic activity, the PRA isomerase variant with an internal tandemly duplicated $(\beta\alpha)^5$ module folds to the native $(\beta\alpha)_8$ core. A ninefold $\beta\alpha$ barrel (cf. Fig. 2A) would disrupt the array of catalytic

Fig. 9. The $(\beta\alpha)_9$ WW' variant is fragmented by Glu-Xaa-specific V8-protease. The kinetics were measured by gel electrophoresis in the presence of SDS. Lanes: M, M_r of marker proteins with indicated $10^{-3} M_r$ values; 0–480, incubation period in minutes; C, control incubation for 480 min in the absence of V8-protease. Arrows with estimated M_r values of fragments. **A:** $(\beta\alpha)_8$ wt. **B:** $(\beta\alpha)_9$ WW'.

residues that are located on the loops between β_1 and α_1 , β_6 and α_6 , and the phosphate-binding site provided by $\beta_7\alpha_7\beta_8\alpha_8$ (Wilmanns et al., 1991a,b). The $(\beta\alpha)_9$ variant closely resembles the $(\beta\alpha)_8$ parent protein in terms of its K_M^{PRA} value, the similarity of the absorption difference spectra as well as the K_d values of bound rCdRP and the CD spectra. Because the GuCl-induced unfolding transitions are also similar, the $(\beta\alpha)_8$ cores of both $(\beta\alpha)_8$ wt and $(\beta\alpha)_9$ WW' must have comparable stabilities. Most importantly, a number of independent experimental criteria establish that the protein has folded into only one out of three possible folding frames; the preferred one carries the duplicated secondary structural elements in the transposed sequence α_5 -linker- β_5 , as an insert between β_5 and α_5 (Fig. 10). β -strand 5 and α -helix 5' now form the natural module $(\beta\alpha)^5$.

Steric hindrance by the disordered insert is probably the reason for the 25-fold decrease of k_{cat} of the $(\beta\alpha)_9$ variant compared to that of the $(\beta\alpha)_8$ reference molecule. The response of both the steady-state kinetic constants and the fluorescence of $(\beta\alpha)_9$ WW' to the replacement of W356 or W356' by phenylalanine is inconsistent with folding frame B (Fig. 2). The remaining folding frames C and D were distinguished by relative susceptibilities of the E369-F370 peptide bonds within α_5 and α_5' toward proteolysis by V8-protease. The clear preference of the initial cleavage at E369 in α_5 is consistent with the hypothesis that the helix α_5' has the conformation of α_5 in the parent molecule ($[\beta\alpha]_8$ wt), which protects the E369-F370 bond.

The argument above is based on the implicit assumption that D is the predominant folding frame. However, if C were the predominant species (cf. Fig. 2), the observed fragmentation pattern (Fig. 9B) would require (1) that the inserted module of B were protected from V8-protease, and (2) that equilibration of C with protease-sensitive D were rapid. These assumptions are not plausible.

Assuming that thermodynamics and not kinetics determines the folding of a protein (Jaenicke, 1991), the preference for folding frame D shows that insertion of the duplicated natural module β_5 -loop- α_5 into the loops between either α_4 and β_5 or α_5 and β_6 (cf. Fig. 2) is energetically less favorable than insertion of the transposed sequence α_5 -linker- β_5 between β_5 and α_5 .

What is the physical reason for these energetic differences? It is pertinent to note that a much higher percentage (44%) of the residues of PRA isomerase from *E. coli* are located in loop positions (Wilmanns et al., 1991b) than are found in proteins on average (~30%). Thus, in agreement with the theoretical considerations of Chou and Carlacci (1990), the loops may contribute significantly toward stabilizing PRA isomerase.

The simplest interpretation of the preference of the $(\beta\alpha)_9$ variant for the folding frame D is that loops and their neighboring regions at the amino-terminal face of the central β barrel core are generally more important for

Fig. 10. Presumptive structure of $(\beta\alpha)_9$ WW'. Schematic two-dimensional representation as in Figure 1B. Partial amino acid sequence in Figure 4. ○, Amino acid residues of the insert. ●, Glu residues 363, 369, 363', and 369'. ◐, Trp residues 356 and 356'. Arrow indicates the primary cleavage site for V8-protease.

stability than their counterparts at the carboxyl-terminal face; Kinemage 2 highlights the $\alpha\beta$ versus the $\beta\alpha$ loops of PRA isomerase. The $(\beta\alpha)_8$ barrel fold provides for a clear distribution of responsibility between those residues that participate in function and those that stabilize the framework (Brändén, 1991). It is plausible that loops at the active site need to be flexible for catalysis. Consequently, mutational insertions in the β -loop- α segments should not destabilize the fold. Conversely, insertions in the α -loop- β segments, which belong to the bowl-shaped framework, should destabilize the fold significantly.

This hypothesis is supported by the common motifs comprising α -loop- β but not β -loop- α segments found in several $(\beta\alpha)_8$ barrel proteins (Thornton et al., 1988; Lindqvist, 1989; Rice et al., 1990; Scheerlinck et al., 1991). Briefly, these motifs comprise the interaction between two hydrophobic residues at the carboxyl-terminus of an α helix with a hydrophobic residue at the amino-terminus of the following β strand. The loops connecting the α helices with the β strands frequently contain up to four residues (tight loops). In the case of PRA isomerase from *E. coli*, several loops have more than four residues (extended loops). For these loops the major structural determinant must be their interaction with the rest of the protein (Tramontano et al., 1989). Based on this hypothesis, insertion of the duplicated module perturbs the stabilizing interactions of the loops between α_4 and β_5 as well as between α_5 and β_6 more strongly than those of the loop between β_5 and α_5 .

This hypothesis also provides for a heuristic rationale for two hitherto unexplained facts that are related to the work reported here. First, Luger et al. (1989) found that the circularly permuted variant of yeast PRA isomerase cp₂ was as stable as the wild type, whereas cp₁ was less stable. According to the hypothesis, the preserved inter-

actions of the α_6 -loop- β_6 segment in cp_2 and its loss in cp_1 would explain the relative stability of cp_2 . Second, a $(\beta\alpha)_{10}$ variant of yeast PRA isomerase, which has a choice to fold into five different folding frames, adopts only one of these (Luger et al., 1990). In this case the hypothesis would attribute the relative lability of the four excluded folding frames to the presence of the artificial loop between α_8 and β_1 in their respective $(\beta\alpha)_8$ cores. The remaining stable folding frame retains all of the wild-type loops between α helices and β strands.

The ability of a protein with an internally duplicated module to fold to a stable and active structure is important with regard to protein evolution. Internal tandem duplication at the genetic level can occur spontaneously (Crawford et al., 1987; Mollet & Delley, 1991). If such duplicated segments were preferentially accommodated in loops at the active site, these variants could be selected for altered specificities or new catalytic activities without necessarily abolishing the original function of the protein. In this regard it is remarkable that 14 out of 17 $(\beta\alpha)_8$ barrel enzymes carry inserted secondary structural elements or even domains between β strands and α helices (Farber & Petsko, 1990; Brändén, 1991).

Materials and methods

Materials

Restriction enzymes and other DNA-modifying enzymes were purchased from Boehringer, the DNA-sequencing kit was purchased from United States Biochemical Corp., and both acrylamide and urea for sequencing gels were from BioRad. Oligonucleotides for sequencing and mutagenesis were synthesized on an Applied Biosystems ABI 380B oligonucleotide synthesizer and purified on a Sephadex G-25 column; GuCl grade I, which was used for protein purification, was purchased from Sigma, and ultrapure GuCl for spectroscopic studies was from Schwarz-Mann. Chromatmedia were bought from Pharmacia, V8-protease from Boehringer, Immobilon (PVDF)-membranes from Millipore, and tryptone and yeast extract for bacterial growth media from Oxoid. All other biochemicals were obtained from Merck, Sigma, or Fluka and were of the highest purity available.

Plasmids and strains

The expression/mutagenesis vectors pMac5-10 (Stanssens et al., 1989) were used for both site-directed mutagenesis and expression of the genes of the $(\beta\alpha)_9$ variants. The corresponding *E. coli* strains were, for mutagenesis: WK6mutS ($\Delta(lac-proAB)$, *galE*, *strA*, *mutS*: *Tn10*/F'*lacI*^q, $\Delta\Delta M15$, *proA*⁺*B*⁺) and WK6 ($\Delta(lac-proAB)$, *galE*, *strA*/F'*lacI*^q, $\Delta\Delta M15$, *proA*⁺*B*⁺), and, for expression: W3110 (*trpR*, Δ *trpEA2*) (Schneider et al.,

1981). The $(\beta\alpha)_8$ variants were expressed from the inducible expression vector pDS56/RBSII/SphI (Certa et al., 1986) in the same recipient.

DNA manipulations

Double-stranded plasmid DNA for large-scale preparations was isolated with the Quiagen purification kit according to the manufacturer's instructions. Single-stranded DNA was purified after infection with the helper phage M13KO7 from Pharmacia according to the manufacturer's instructions. All other recombinant DNA methods followed standard protocols (Sambrook et al., 1990).

Site-directed mutagenesis

The gene coding for $(\beta\alpha)_8$ wt was subcloned via an EcoRI/HindIII fragment into the expression/mutagenesis vector pMc5-10 (Figs. 3A,F, 5A) The gapped duplex method in combination with the pMac mutagenesis system (Stanssens et al., 1989) was used to introduce the desired mutations into the gene of the PRA isomerase domain ($(\beta\alpha)_8$ wt). The mutagenic oligonucleotides were 25-42 nucleotides long, depending on the number of mismatched bases, with at least seven matching nucleotides upstream and downstream of the mismatches. The mutants were screened for the presence of new restriction sites where new BglII, BamHI, and SphI sites had been introduced. The W356F and W391F point mutations were screened by sequencing the target region. Clones carrying the desired mutation were checked by sequencing the entire coding sequence by the chain-termination method (Sanger et al., 1977) to detect possible adventitious spontaneous mutations. The DNA of $(\beta\alpha)_8$ wt contained the replacement V329A (GTG $\rightarrow$ GCG), which is therefore also contained in all further variants. This replacement was also detected crystallographically (Wilmanns et al., 1991b).

Enzyme and protein assays

The enzyme assay was performed at 25 °C in 50 mM potassium phosphate buffer at pH 7.5, containing 4 mM K₂MgEDTA and 1 mM DTE. The labile substrate PRA was generated from ANT and PRPP in the cuvette by the PRT-catalyzed reaction (Hommel et al., 1989). The disappearance of PRA was monitored fluorimetrically. The K_M^{PRA} and k_{cat} values were determined by analysis of entire progress curves as well as by initial velocity measurements and direct linear plot analysis according to Eberhard (1990a). Protein concentrations were determined according to Bradford (1976) using the BioRad protein assay with bovine serum albumin as a standard.

Expression and purification of recombinant proteins

The plasmids bearing the coding sequences for the various $(\beta\alpha)_8$ and $(\beta\alpha)_9$ variants were transformed into the recipient strain W3110 (*trpR*, $\Delta trp EA2$). The cells harboring the plasmids coding for the $(\beta\alpha)_9$ variants (Fig. 3E) were grown in LB medium under selection with chloramphenicol. The proteins were constitutively expressed at 37 °C for 14 h. The cells harboring the expression plasmids for the $(\beta\alpha)_8$ variants (Fig. 5D) were induced in LB medium at 37 °C at $A_{600} \sim 0.45$ with 1 mM IPTG and cultivated for another 4.5 h. The cells were disrupted by sonication at a concentration of 0.4 g wet cells/mL at 0 °C in 50 mM potassium phosphate at pH 7.5, containing 1 mM EDTA, 1 mM DTE, and 0.1 mM phenylmethylsulfonylfluoride (buffer A). After centrifugation, the pellet was washed three times with the same buffer and dissolved in 200 mL of 3.5 M GuCl in buffer A, stirred at room temperature for 1 h, and then centrifuged.

The clear supernatant was diluted 2.5-fold with 2 M GuCl in buffer A and dialyzed at 4 °C against buffer A. The precipitate was removed by centrifugation and the supernatant containing the PRA isomerase activity was applied to a column of DEAE-Sepharose fast-flow, equilibrated with buffer A, and eluted with a linear gradient from 10 to 250 mM potassium phosphate buffer at pH 7.5, containing 1 mM EDTA and 1 mM DTE. The pooled active fractions were supplemented with 10% glycerol, then dripped into liquid nitrogen, and stored at -70 °C.

Determination of molecular weight

M_r values of native proteins were determined by gel filtration on a column (1 × 30 cm) of Superose 12, using a Merck HPLC system. The buffer (10 mM Tris at pH 7.5, containing 300 mM NaCl) was pumped at a flow rate of 0.5 mL/min, at 20 °C. Samples of 100 μ L at 1.0 mg protein/mL were injected, and the elution profiles were recorded on-line by monitoring the absorption at 280 nm. The column was calibrated by plotting the elution volumes of various standard native proteins versus the logarithm of the known M_r . The standard proteins were cytochrome C (12,400), horse heart myoglobin (16,900), ePRA isomerase (21,100), hen egg white ovalbumin (45,000), indoleglycerol phosphate synthase:PRA isomerase (49,500), and rabbit muscle phosphorylase b (194,800). The M_r values of the unknown proteins were interpolated from this plot.

Polyacrylamide gel electrophoresis in the presence of SDS was performed according to Lämmli (1970). Gels with 12.5% polyacrylamide were used and stained with Coomassie brilliant blue. The following proteins were used as markers: phosphorylase b (97,400), bovine serum albumin (66,300), β -subunit of tryptophan synthase

(43,000), aspartate aminotransferase (43,000), α -subunit of tryptophan synthase (28,700), the PRA isomerase domain from *E. coli* (21,100), myoglobin (16,800), cytochrome C (12,400), and BPTI (6,500).

Circular dichroism measurements

Circular dichroism measurements were carried out with a Cary model 61 instrument at 22 ± 2 °C. Circular dichroism spectra were measured at protein concentrations of 3.0 μ M in a 0.1-cm cell. The buffer was 50 mM potassium phosphate at pH 7.5, containing 0.1 mM EDTA and 0.1 mM DTE. Mean residue ellipticity values ($[\theta]_{mrw}$) were calculated by using the equation $[\theta]_{mrw} = 100 \cdot \theta_{obs} / lc$ (deg cm² dmol⁻¹), where θ_{obs} is the observed ellipticity in degrees, c is the molar concentration of residues, and l is the light path in centimeters.

For equilibrium unfolding experiments, the CD of protein solutions was measured at 222 nm in a cell having a 0.1-cm light path. The proteins were incubated at 4 °C in 50 mM potassium phosphate buffer at pH 7.5, containing 0.1 mM EDTA, 0.1 mM DTE, and various concentrations of GuCl (0, 0.1, 0.2, 0.3, ..., 1.8, 2.0, 3.0, 4.0, 5.0, and 6.0 M). The CD signal was constant after 3 h. For equilibrium refolding experiments, the proteins were first unfolded in 4 M GuCl for 3 h at room temperature. The CD of the protein solutions was measured upon their appropriate dilution to different final concentrations of GuCl (3, 2, 1.5, 1.0, 0.75, 0.5 M) and refolding for 4 h at room temperature.

Fluorescence measurements

Fluorescence emission spectra were measured with a Perkin-Elmer 650-10S fluorimeter equipped with a thermostatted cuvette holder. Measurements were made at 25 °C in 1-cm path-length cuvettes. Protein concentrations were 10 μ M. Fluorescence was excited at 280 nm, and the excitation and emission slit widths were set to 2 nm. Under these conditions no appreciable photodecomposition of the protein samples was detected. All fluorescence spectra were uncorrected.

Quenching of Trp fluorescence by acrylamide

The measurements were performed at 25 °C in 50 mM potassium phosphate at pH 7.5, containing 1 mM EDTA and 1 mM DTE. Acrylamide was dissolved freshly before each experiment at a concentration of 6 M in the described buffer. Fluorescence measurements were carried out as described above. Tryptophan fluorescence was excited at 295 nm and the emission was recorded at 340 nm. Quenching effects were measured as follows. Aliquots (5 μ L) of the stock solution of acrylamide were added successively to a 2-mL sample of protein solution in the

cuvette. The fluorescence intensity measured 2 min after each addition was corrected for the dilution by the acrylamide solution. Correction of the inner filter effect due to the absorption of acrylamide at 295 nm was applied by multiplying the fluorescence intensity by $10^{A/2}$, where A is the absorbance of the solution at 295 nm measured in a 1-cm cuvette (Chaffotte & Goldberg, 1984).

Limited proteolysis with V8-protease

Protein solutions were dialyzed against 100 mM Tris-acetate buffer at pH 7.5 supplemented with 2 mM EDTA and 1 mM DTE, and then concentrated by ultrafiltration to 1 mg/mL. A 350- μ L aliquot of the protein solution was mixed with 9 μ L V8-protease solution (0.2 mg/mL) and incubated at room temperature. At different time intervals, aliquots were mixed with an equal volume of sample buffer (Lämmli, 1970) and heated for 5 min at 100 °C. The samples were subsequently analyzed by gel electrophoresis in the presence of SDS as described above. The activity of the protease was checked by cleavage of indoleglycerol phosphate synthase:PRA isomerase from *E. coli* in the presence of SDS (Kirschner et al., 1980). The characteristic fragments appeared within 5 min at 4 °C.

The fragments generated by limited proteolysis and separated by gel electrophoresis were blotted on an Immobilon (PVDF) membrane using standard techniques (Towbin et al., 1979). The transfer buffer was 10 mM 3-(cyclohexylamino)-1-propane-sulfonic acid at pH 11, containing 10% methanol. The fragments were transferred for 15 min with 500 mA. After washing the membrane, the blot was stained for 5 min with a solution of 0.1% Coomassie blue R-250 in 50% methanol. The blot was allowed to dry at room temperature for 30 min, and the fragments of interest were cut out and stored under nitrogen at -20 °C. These blotted fragments could be used directly for determination of the amino-terminal sequences.

Acknowledgments

We thank M. Wilmanns and J.N. Jansonius for access to the refined coordinates of PRA isomerase from *E. coli* and H. Szadkowski for excellent technical support. We appreciate constructive criticism from E.W. Miles and M. Wilmanns and access to papers prior to publication by C.-I. Brändén and S. Wodak. We are grateful to P. Jenö for performing protein sequence analyses and oligonucleotide syntheses. We also thank E. Johner for typing the manuscript, and both M. Jäggi and V. Gnedder for artwork. This work was supported by the Swiss National Science Foundation, grant no. 31.25711.88.

References

Bisswanger, H., Kirschner, K., Cohn, W., Hager, V., & Hansson, E. (1979). N-(5'-phosphoribosyl)anthranilate isomerase:indoleglycerol

- phosphate synthase. 1. A substrate analogue binds to two different binding sites on the bifunctional enzyme from *Escherichia coli*. *Biochemistry* 18, 5946-5953.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248-254.
- Brändén, C.-I. (1991). The most frequently occurring folding motif in proteins—The TIM barrel. *Curr. Opin. Struct. Biol.*, in press.
- Buchwalder, A., Szadkowski, H., & Kirschner, K. (1991). A variant of dihydrofolate reductase with a circularly permuted sequence is fully active. *Biochemistry*, in press.
- Certa, U., Bannwarth, W., Stüber, D., Gentz, R., Lanzer, M., LeGrice, S., Guillot, F., Wendler, I., Hunsmann, G., Bujard, H., & Mous, J. (1986). Subregions of a conserved part of the HIV gp41 transmembrane protein are differentially recognized by antibodies of infected individuals. *EMBO J.* 5, 3051-3056.
- Chaffotte, A.F. & Goldberg, M.E. (1984). Fluorescence-quenching studies on a conformational transition within a domain of β_2 subunit of *Escherichia coli*. *Eur. J. Biochem.* 139, 47-50.
- Chothia, C. & Finkelstein, A.V. (1990). The classification and origin of protein folding patterns. *Annu. Rev. Biochem.* 59, 1007-1039.
- Chou, K.-C. & Carlacci, L. (1990). Energetic approach to the folding of ($\alpha\beta$) barrels. *Proteins Struct. Funct. Genet.* 9, 280-295.
- Crawford, I.P., Clarke, M., van Cleemput, M., & Yanofsky, C. (1987). Crucial role of the connecting region joining the two functional domains of yeast tryptophan synthase. *J. Biol. Chem.* 262, 239-244.
- Eberhard, M. (1990a). A set of programs for analysis of kinetic and equilibrium data. *Comput. Appl. Biosci.* 6, 213-221.
- Eberhard, M. (1990b). Structural and functional studies on engineered monofunctional domains of an originally bifunctional enzyme of *Escherichia coli*. Ph.D. Thesis, University of Basel, Switzerland.
- Eftink, M.R. & Ghiron, C.A. (1976). Exposure of tryptophanyl residues in proteins. Quantitative determination by fluorescence quenching studies. *Biochemistry* 15, 672-680.
- Farber, G.K. & Petsko, G.A. (1990). The evolution of α/β barrel enzymes. *Trends Biochem. Sci.* 15, 228-234.
- Fetrow, J.S., Cardillo, T.S., & Sherman, F. (1989). Deletions and replacements of omega loops in yeast iso-1-cytochrome c. *Proteins Struct. Funct. Genet.* 6, 372-381.
- Finkelstein, A.V. & Pitsyn, O.B. (1988). Why do globular proteins fit the limited set of folding patterns. *Prog. Biophys. Mol. Biol.* 50, 171-190.
- Goldenberg, D.P. & Creighton, T.E. (1983). Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* 165, 407-413.
- Hommel, U., Lustig, A., & Kirschner, K. (1989). Purification and characterization of yeast anthranilate phosphoribosyl transferase. *Eur. J. Biochem.* 180, 33-40.
- Hynes, T.R., Kautz, R.A., Goodman, M.A., Gill, J.F., & Fox, R.O. (1989). Transfer of a β -turn structure to a new protein context. *Nature* 339, 73-76.
- Jaenicke, R. (1991). Protein folding: Local structures, domains, subunits and assemblies. *Biochemistry* 30, 3147-3161.
- Kirschner, K., Szadkowski, H., Henschen, A., & Lottspeich, F. (1980). Limited proteolysis of N-(5'-phosphoribosyl)anthranilate isomerase:indoleglycerol phosphate synthase from *Escherichia coli* yields two different enzymically active functional domains. *J. Mol. Biol.* 143, 395-409.
- Lämmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage. *Nature* 227, 680-685.
- Lasters, I., Wodak, S.J., Alard, P., & van Cutsem, E. (1988). Structural principles of parallel β -barrels in proteins. *Proc. Natl. Acad. Sci. USA* 85, 3338-3342.
- Lesk, A.M., Brändén, C.-I., & Chothia, C. (1989). Structural principles of α/β barrel proteins: The packing of the interior of the sheet. *Proteins Struct. Funct. Genet.* 5, 139-148.
- Leszczynsky, J.F. & Rose, G.D. (1986). Loops in globular proteins: A new class of protein secondary structures. *Science* 234, 849-855.
- Lim, W.A. & Sauer, R.T. (1991). The role of internal packing interactions in determining the structure and stability of a protein. *J. Mol. Biol.* 219, 359-376.
- Lindqvist, Y. (1989). Refined structure of spinach glycolate oxidase at 2 Å resolution. *J. Mol. Biol.* 209, 151-166.
- Luger, K., Hommel, U., Herold, M., Hofsteenge, J., & Kirschner, K.

- (1989). Correct folding of circularly permuted variants of a ($\beta\alpha$)-barrel enzyme in vivo. *Science* 243, 206–210.
- Luger, K., Szadkowski, H., & Kirschner, K. (1990). An 8-fold ($\beta\alpha$)-barrel protein with redundant folding possibilities. *Protein Eng.* 3, 249–258.
- Mollet, B. & Delley, M. (1991). A β galactosidase deletion mutant of *Lactobacillus bulgaricus* reverts to generate an active enzyme by internal DNA sequence duplication. *Mol. Gen. Genet.* 227, 17–21.
- Niermann, T. & Kirschner, K. (1991). Improving the prediction of secondary structure of “TIM-barrel” enzymes. *Protein Eng.* 4, 359–370.
- Priestle, J.P., Grütter, M.G., White, J.L., Vincent, M.G., Kania, M., Wilson, E., Jardetzky, T.S., Kirschner, K., & Jansonius, J.N. (1987). Three-dimensional structure of the bifunctional enzyme N-(5'-phosphoribosyl)anthranilate isomerase-indole-3-glycerol phosphate synthase from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 84, 5690–5694.
- Rice, P.A., Goldman, A., & Steitz, T.A. (1990). A helix-turn-strand structural motif common in α - β proteins. *Proteins Struct. Funct. Genet.* 8, 334–340.
- Riechmann, L., Clark, M., Waldmann, H., & Winter, G. (1988). Reshaping human antibodies for therapy. *Nature* 332, 323–327.
- Sambrook, J., Fritsch, E.F., & Maniatis, T. (1990). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sanger, F., Nicklen, S., & Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- Scheerlinck, J.-P.Y., Lasters, I., Claessens, M., De Mayer, M., Pio, F., Delhaise, P., & Wodak, S.J. (1991). Recurrent α /beta loop structures in TIM barrel motifs show a distinct pattern of conserved structural features. *Proteins Struct. Funct. Genet.*, in press.
- Schneider, W.P., Nichols, B.P., & Yanofsky, C. (1981). Procedures for production of hybrid genes and proteins and its use in assessing significance of amino acid differences in homologous tryptophan synthetase α polypeptides. *Proc. Natl. Acad. Sci. USA* 78, 2169–2173.
- Stanssens, P., Opsomer, C., McKeown, Y.M., Kramer, W., Zabeau, M., & Fritz, H.-J. (1989). Efficient oligonucleotide-directed construction of mutations in expression vectors by the gapped duplex DNA method using alternating selectable markers. *Nucleic Acids Res.* 17, 4441–4454.
- Taniuchi, H., Parker, H.S., & Juillerat, M.A. (1986). Complementation in folding and fragment exchange. *Methods Enzymol.* 131, 185–217.
- Thornton, J.M., Sibanda, B.L., Edwards, M.S., & Barlow, D.J. (1988). Analysis, design and modification of loop regions in proteins. *Bio-Essays* 8, 63–69.
- Towbin, H., Staehelin, T., & Gordon, T. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. USA* 76, 4350–4353.
- Tramontano, A., Chothia, C., & Lesk, A.M. (1989). Structural determinants of the conformations of medium-sized loops in proteins. *Proteins Struct. Funct. Genet.* 6, 382–394.
- Wilmanns, M., Hyde, C.C., Davies, D.R., Kirschner, K., & Jansonius, J.N. (1991a). Structural conservation of parallel β / α -barrel enzymes that catalyze three sequential reactions in the pathway of tryptophan biosynthesis. *Biochemistry*, 30, 9161–9169.
- Wilmanns, M., Priestle, J.P., Niermann, T., & Jansonius, J.N. (1991b). Three-dimensional structure of the bifunctional enzyme phosphoribosylanthranilate isomerase:indoleglycerol phosphate synthase from *Escherichia coli* refined at 2.0 Å resolution. *J. Mol. Biol.*, in press.
- Wilson, D.K., Rudolph, F.B., & Quirocho, F.A. (1991). *Science* 252, 1278–1284.