PROTEINS: Structure, Function, and Genetics 39:89–101 (2000)

Plasticity and Steric Strain in a Parallel β-Helix: Rational Mutations in the P22 Tailspike Protein

Benjamin Schuler,[1] Frank Fürst,[2] Frank Osterroth,[1] Stefan Steinbacher,[3] Robert Huber,[3] and Robert Seckler[2]*
[1]*Institut für Biophysik und Physikalische Biochemie, Universität Regensburg, Regensburg, Germany*
[2]*Physikalische Biochemie, Universität Potsdam, Luckenwalde, Germany*
[3]*Max-Planck-Institut für Biochemie, Abteilung Strukturforschung, Martinsried, Germany*

ABSTRACT **By means of genetic screens, a great number of mutations that affect the folding and stability of the tailspike protein from *Salmonella* phage P22 have been identified. Temperature-sensitive folding (*tsf*) mutations decrease folding yields at high temperature, but hardly affect thermal stability of the native trimeric structure when assembled at low temperature. Global suppressor (*su*) mutations mitigate this phenotype. Virtually all of these mutations are located in the central domain of tailspike, a large parallel β-helix. We modified tailspike by rational single amino acid replacements at three sites in order to investigate the influence of mutations of two types: (1) mutations expected to cause a *tsf* phenotype by increasing the side-chain volume of a core residue, and (2) mutations in a similar structural context as two of the four known *su* mutations, which have been suggested to stabilize folding intermediates and the native structure by the release of backbone strain, an effect well known for residues that are primarily evolved for function and not for stability or folding of the protein. Analysis of folding yields, refolding kinetics and thermal denaturation kinetics in vitro show that the *tsf* phenotype can indeed be produced rationally by increasing the volume of side chains in the β-helix core. The high-resolution crystal structure of mutant T326F proves that structural rearrangements only take place in the remarkably plastic lumen of the β-helix, leaving the arrangement of the hydrogen-bonded backbone and thus the surface of the protein unaffected. This supports the notion that changes in the stability of an intermediate, in which the β-helix domain is largely formed, are the essential mechanism by which *tsf* mutations affect tailspike folding. A rational design of *su* mutants, on the other hand, appears to be more difficult. The exchange of two residues in the active site expected to lead to a drastic release of steric strain neither enhanced the folding properties nor the stability of tailspike. Apparently, side-chain interactions in these cases overcompensate for backbone strain, illustrating the extreme optimization of the tailspike protein for conformational stability. The result exemplifies the view arising from the statistical analysis of the distribution of backbone dihedral angles in known three-dimensional protein structures that the adoption of ϕ/ψ angles other than the most favorable ones is often caused by side-chain interactions. Proteins 2000;39:89–101.**

Key words: protein folding; protein stability; backbone dihedral angles; parallel beta-helix; active site; temperature-sensitive mutants

INTRODUCTION

Parallel β-sheets have long been perceived as rather unstable structures compared to their antiparallel counterparts, due to their more distorted inter-strand hydrogen bonds, and reflected by the low abundance of large parallel sheets in known protein structures. Before the discovery of a parallel β-helix in the pectate lyase PelC,[1] the small number of protein structure types included all-α-helical domains, entities composed of mixed α-helices with parallel β-sheets, all antiparallel β-structures, and a class of small proteins rich in disulfides or containing metal binding sites.[2,3] Parallel β-sheets were thought to only coexist in domains with other, more stable structural elements and had never been observed at all with their outer surface exposed to solvent.[4] Although this has changed with the PelC structure and the identification of parallel β-helices in a number of other proteins since,[1,5–18] little is known about the interactions stabilizing such all-parallel β-structures and about the requirements for their efficient folding.

Apart from the two left-handed structures found so far,[10,11] all parallel β-helices are formed by coiling a β-strand into a right-handed helix.[19] Except for alkaline protease, each turn around the helix axis is completed by typically three β-strand segments, connected by turns or loops of varying length. The β-strands of consecutive turns

Abbreviations: TSP, tailspike protein; wt, wild-type; K_D, dissociation constant; CD circular dichroism; SDS, sodium dodecyl sulfate.

Grant sponsor: Deutsche Forschungsgemeinschaft; Grant sponsor: Fonds der Chemischen Industrie.

B. Schuler's present address is Laboratory of Chemical Physics, National Institutes of Health, Bethesda, Maryland.

*Correspondence to: Robert Seckler Physikalische Biochemie, Universität Potsdam, Im Biotechnologiepark, 14943 Luckenwalde, Germany. E-mail: seckler@rz.uni-potsdam.de

Received 1 July 1999; Accepted 5 November 1999

Fig. 1. Stereo ribbon drawing of one subunit of the trimeric TSPΔN prepared with MOLMOL.[59] The domain formed by the large right-handed β-helix is in dark grey. Thirteen complete turns wind around the helix axis and form a long binding groove for the lipopolysaccharide receptor of phage P22 present on the surface of its host *Salmonella*. Bound octasaccharide, a product of the hydrolysis catalysed by the endorhamnosidase, is indicated by thick lines. The mutation sites addressed here are indicated as light spheres. The N-terminus is located in the lower left.

are aligned to form parallel β-sheets. Two of them, usually called PB1 and PB2 (in pectate lysases), or C and A (in P22 tailspike), respectively, form a parallel β-sandwich. One side of the sandwich is closed either by turns of variable length or very short β-strands (PB1a in rhamnogalacturonase). On the other side lies a third longer β-sheet (PB3 or B), giving the cross-section of each coil an L-shaped appearance. All of the known enzymes with a parallel β-helical fold are involved in the recognition and cleavage of large polysaccharides. Whether this reflects a particular suitability of this fold for polysaccharide binding, and if so, whether these enzymes are homologous or analogous structures, remains to be determined. Sequence alignment studies so far offer only very limited information on this question, as they are hampered by the low accuracy of prediction methods for β-structure, the different lengths of regular structure in each coil, and the high variability of size and composition of the loops protruding from the parallel β-helix. But, although it is not yet possible to identify a distinctive sequence pattern characterizing all β-helices, certain features typical of some of the known structures, like internal Asn ladders or the high content of aromatic and large hydrophobic groups oriented towards the helix core, will possibly aid the identification of further candidates with the new motif.[20]

P22 Tailspike Protein

In search of factors contributing to the stability and folding of β-helices, we chose to approach the problem by rational site-directed mutagenesis of the P22 tailspike endorhamnosidase, a β-helix protein for which a large amount of biochemical and biophysical data are available. Bacteriophage P22 attaches to its host *Salmonella* via six trimeric tailspikes non-covalently linked to the baseplate. The receptor is the lipopolysaccharide structure on the surface of the enterobacterium, containing 4 to 30 O-antigenic tetrasaccharide repeats. It is bound and hydrolyzed in the long groove of the β-helix domain of tailspike (Fig. 1). Both oligosaccharide binding and hydrolysis have been thoroughly characterized,[41] and the crystal structures of several protein-oligosaccharide complexes have been determined.[21,7]

The central domain of tailspike is a 13-turn parallel β-helix (Fig. 1) with a tightly packed hydrophobic core containing very few polar or charged side chains. It does not contain the stacked Asn residues typical of some other β-helix proteins. The terminal domains of the three subunits are intertwined and thereby responsible for the thermostability of the trimer. This thermostability, combined with a resistance towards denaturation by SDS, allowed the identification of SDS- and heat-sensitive assembly intermediates, leading to an extensive characterization of the folding and assembly pathway in vivo[22,23] and in vitro.[24,25] Moreover, King and co-workers identified a large number of amino acid substitutions in the central part of tailspike, which affect the yield of folding and assembly.[26] Tailspike folding yields decrease dramatically with increasing temperature, with very similar dependencies in vivo and in vitro.[27,28] The temperature sensitivity is more drastic for a type of mutants designated *temperature sensitive for folding* or *tsf*.[29–31] These *tsf* mutations pre-

vent the formation of the native trimer at restrictive temperature and promote aggregation, but once assembled at a lower permissive temperature, the mutant tailspike trimers have full biological activity at high temperatures.[32] *Tsf*-mutations have been identified at some 60 sites exclusively in the central β-helix domain.[33] A second class of mutations, isolated as *global suppressors* (*su*) of the *tsf* phenotype, have the opposite effect: they increase the yield of correctly assembled tailspikes at higher temperature, thus compensating for the effect of *tsf* mutations.[34,35] Only four such *su* mutations have been identified so far.[36] An analysis of the refolding pathways of wild-type tailspike and of a number of single- and double-mutant proteins at varied temperature has confirmed that both types of mutations act by altering the global stability of well-structured, native-like folding intermediates.[25,37,38] Accordingly, there is a good correlation between the effect of a mutation on tailspike folding and its effect on the stability of the finally folded structure.[39]

tsf and *su* Analogs

In view of this wealth of data, we set out to investigate the effects of substituting amino acids not previously identified as sites of folding mutations. We selected three positions in the central part of the parallel β-helix, close to the oligosaccharide-binding site. Two of the sites correspond to active-site residues associated with unfavorable backbone dihedrals in the high-resolution crystal structure, such that stabilization by glycine or alanine substitutions might be expected. The side chain of the third residue (Thr326) is directed towards the β-helix core. Residues of varying side-chain volume were introduced at this site as a further test of our model for the mechanism of action of *temperature-sensitive folding* mutations. The results of these mutations on the one hand prove that the *tsf* phenotype can be generated by rational mutation designed to moderately destabilize the native parallel β-helix fold, and on the other hand illustrate the remarkable amount of plasticity in the β-helix lumen.

MATERIALS AND METHODS

Chemicals

Ultrapure urea was obtained from ICN Biochemicals. Concentrations of urea solutions were determined by refractive index measurements.[40] Dithioerythritol was from Roth, Karlsruhe; Tween 20, SDS, and electrophoresis chemicals were from Serva (Heidelberg). Oligonucleotides were obtained from MWG-Biotech (Munich), and enzymes were from Boehringer-Mannheim Biochemicals (Indianapolis, IN), Pharmacia (Gaithersburg, MD) or New England Biolabs (Beverly, MA). Other chemicals were analysis grade, and quartz-bidistilled water was used throughout. Lipopolysaccharide fragments from *Salmonella enteritidis* and dodecasaccharide fragments labelled with 7-amino-4-metylcoumarin[41] were generously provided by A. Weintraub and U. Baxa.

Site-Directed Mutagenesis and Protein Purification

The plasmid pTSF1 used here is a derivative of pASK30 with an insert coding for an amino-terminal methionine followed by residues 109 to 666 of P22 tailspike protein, i.e., for the tailspike protein lacking its amino-terminal head-binding domain.[39] Site-directed mutagenesis was carried out essentially as described by Kunkel et al.[42] with slight variations.[43] The oligonucleotides used for mutagenesis were:

T326V: 5′-GGGCGCTACTTACGGATCCATAGCTGACTCGTCCGCCAATG-3′

T326S: 5′-GGGCGCTACTTACGGATCCATAGCTGCTTCGTCCGCCAATG-3′

T326F: 5′-GGGCGCTACTTACGGATCCATAGCTGAATCGTCCGCCAATG-3′

E359A: 5′-CAAGTTTTAACGCCCGAGGCCCCAGCGCGATATG-3′

E359G: 5′-TTGCCAAGTTTTAACGCCGGATCCCCCAGCGCGATATGAAGT-3′

W391A: 5′-AAATCGAATCCGTCAGCGACTGGGTAAATAACGAC-3′

W391G: 5′-CCTAAATCGAATCCGTCTCCTACGGGGTAAATAACGACC-3′

The double mutant E359G/G244R was constructed by restriction fragment exchange between the plasmid coding for the G244R mutant protein (pTSF6; Ref. 39) and the plasmid containing a mutation leading to the exchange E359G created here using the unique restriction sites NheI and HindIII. The mutations introduced were confirmed by DNA chain termination sequencing. Amino-terminally shortened tailspike protein was expressed and purified as described.[6,39] Protein concentrations were determined using a specific absorbance at 280 nm of $A_{1\ mg/ml} = 1.20$ for wt.[39] Extinction coefficients for the mutants involving exchange of aromatic residues (T326F, W391A, and W391G) were determined according to Pace et al.[44] by measurement of the absorbance of solutions of equal protein concentration in neutral buffer and 6 M guanidinium chloride. Purified protein was stored as a suspension in 40% saturated ammonium sulfate, 50 mM Tris/HCl, pH 8.0 at 4°C.

Octasaccharide Binding Isotherms and Determination of Endorhamnosidase Activity

Dissociation constants for the specific binding of tailspike protein mutants to octasaccharide fragments from *Salmonella enteritidis* lipopolysaccharide were determined by fluorescence titration[41] exploiting the static fluorescence quench of two Trp residues located in the saccharide binding groove. The protein concentration used was 10 μg/ml (170 nM binding sites) in 50 mM sodium phosphate buffer pH 7.0. Data were fit to the equation

$$F = F_0 + \Delta F(E_0 + L_0 + K_D - ((E_0 + L_0 + K_D)^2 - 4L_0E_0)^{1/2})/2E_0$$

by nonlinear regression to determine the dissociation constants, where F is the measured fluorescence, F_0 is the fluorescence of the free protein, ΔF is the change in fluorescence at saturation relative to F_0, E_0 and L_0 are the total concentrations of protein and oligosaccharide, respectively, and K_D is the dissociation constant.

Endorhamnosidase activity was measured using labelled dodecasaccharide derived from *Salmonella* lipopolysaccharide.[41] At varied times, samples were removed from the reaction vessel containing 0.17 μM tailspike and 10 μM substrate, the reaction was stopped by acidification of the sample and analyzed by reversed phase HPLC using fluorescence detection. Relative amounts of labelled tetrasaccharide and unhydrolyzed labelled dodecasaccharide were determined from the elution profiles by integration using the program PeakFit (Jandel Scientific, Corte Madera, CA).

Protein Denaturation and Reconstitution

Tailspike protein was dissociated and unfolded in 5 M urea, pH 3.0 (50 to 65 mM H_3PO_4) for ≥ 30 min at room temperature.[45] For measurements of folding kinetics, a solution of denatured protein was diluted to a final urea concentration of 50 mM and a final protein concentration of 10 μg/ml. In order to guarantee rapid mixing, the refolding buffer (50 mM sodium phosphate, 1 mM EDTA, 1 mM dithioerythritol (pH 7.0)) was stirred vigorously using the thermostatted cell holder of a Spex Fluoromax spectrofluorimeter. Fluorescence emission was observed at 342 nm using an excitation wavelength of 280 nm. For the reconstitution yields at different temperatures, refolding was performed in Eppendorf tubes pre-treated by a rinse with 5% (v/v) Tween 20 (polyoxyethylenesorbitan monolaurate) followed by drying to avoid loss of protein to the vessel walls. Refolding was initiated by rapid dilution of denatured protein solution with refolding buffer to a final urea concentration of 100 mM and a final protein concentration of 25 μg/ml. The samples were then incubated at constant temperature for 16 hr (above 20 °C) or 70 hr (20°C and below) to ensure completion of the refolding reaction. After SDS polyacrylamide gel electrophoresis and Coomassie Blue staining, the ratios of the intensities of monomer to native trimer bands were determined densitometrically.[46,39]

Thermal Unfolding in the Presence of SDS

For each time point, 40 μl of protein solution containing 50 mM Tris/HCl, 150 mM 2-mercaptoethanol, 20 g/l SDS at pH 7.0 at room temperature[28] were incubated submerged in a thermostatted water bath using thin-walled reaction vessels (for PCR, Roth, Karlsruhe) to maximize the heating rate of the solution.[39] The unfolding reactions were stopped by rapidly cooling the samples corresponding to the respective times in a water/ice mixture. Electrophoresis and densitometry were used as described above to monitor the decrease in intensity of the detergent-resistant trimer band relative to the increasing intensity of the monomer band. Although the specific staining intensities of monomer and native trimer bands are different, the time courses analyzed under these rigorously identical experimental conditions can well be compared between the mutants.

Crystallization and Structure Determination

The hanging droplet method was used for crystallization with ammonium sulfate as a precipitant.[6] After one week, cube-shaped crystals of approximately 0.2 to 0.4 mm were harvested and X-ray intensities measured using a MAResearch image plate (MAResearch, Hamburg) with CuKα-radiation from a Rigaku RU 200 X-ray generator operated at 5.4 kW at 16°C. Data were processed with MOSFLM,[47] and the mutant was analyzed by difference Fourier using phases of the wild-type structure[7] after energy-restrained crystallographic refinement with X-PLOR and parameters derived by Engh and Huber.[48] The coordinates have been submitted to the Protein Data Bank (entry codes 1QQ1 (E359G), 1QRB (T326F), 1QRC (W391A)).

RESULTS

Selection of Mutation Sites

The mutations were introduced into the N-terminally shortened variant of tailspike that retains all functional properties of the endorhamnosidase, as well as thermostability, SDS-resistance, and the folding pathway of the complete protein, but allows the effects of *su-* and *tsf*-mutations on the stability of the native structure to be more readily detected in thermal denaturation kinetics.[28,39]

Two of the mutation sites, 359 and 391, are partially solvent-exposed and located at equivalent positions in two successive turns of the β-helix in the sections containing β-sheets B and C and forming one rim of the saccharide binding groove. Five residues located in the corresponding position of consecutive β-helix coils—Asp 303, Val 331, Glu 359, Trp 391, and Gly 438—form a stack with very similar backbone dihedral angles (Fig. 2). In a Ramachandran diagram, their ϕ/ψ coordinates correspond to the border of an "additionally allowed region" (in ProCheck terminology[49]) for Val 331 (ϕ = −121°, ψ = 145°) and a "generously allowed region" for Glu 359 (ϕ = −122°, ψ = −104°) and Trp 391 (ϕ = −107°, ψ = −109°), thus indicating substantial steric strain in these residues. Indeed, if Val 331 is mutated to Ala or Gly, an increase in stability of the native protein is observed, along with higher renaturation yields at elevated temperature. Accordingly, V331A and V331G are the most potent mutations suppressing the folding defect of *tsf* mutants in tailspike. The reason why these mutants have not evolved as the wild-type is probably their diminished receptor binding affinity.[50] As the backbone strain is even more prominent for Glu 359 and Trp 391, the question was whether exchanging these residues to Ala or Gly would have similar effects, although these mutations had never shown up in the extensive genetic screens for *su* sites. A possible improvement of folding and stability caused by these mutations might have been compromised by their effects on the function of the endorhamnosidase, as structural and biophysical data indicate that Trp 391 is important for saccharide binding and suggest a possible role of Glu 359 in substrate hydrolysis (see below).

The other site chosen for directed mutagenesis is Thr 326, stacked between Ser 354 and Pro 298 in β-sheet B, and pointing to the interior of the β-helix (cf. Fig. 7). In the crystal structure, neither hydrogen bonds to other residues nor any neighboring cavities are detectable, thus

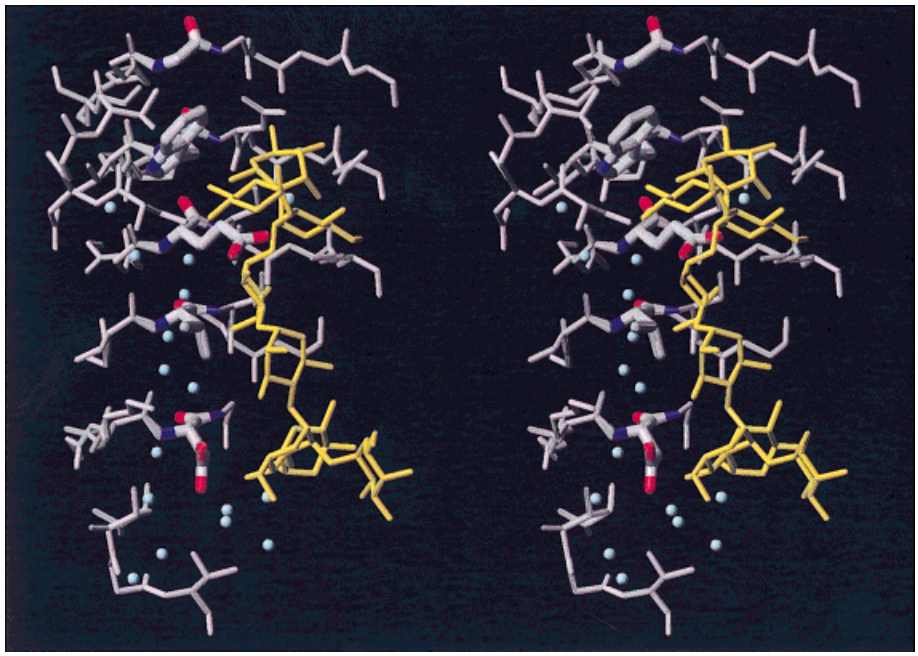

Fig. 2. Stereo plot of the vicinity of the stack consisting of Asp 303, Val 331, Glu 359, Trp 391, and Gly 438 (from bottom to top), which are drawn with increased bond radius for clarity; N atoms are colored blue and O atoms red. The bound octasaccharide is yellow, water molecules found in the crystal structure are indicated by light blue balls. To the right of Trp 391, Asp 392, and Asp 395 can be recognized, which are probably involved in oligosaccharide hydrolysis. The water molecule assumed to be responsible for the attack on the C1-atom of rhamnose I (upper end of the oligosaccharide as depicted here) is hydrogen bonded to Glu 359, Ser 360, and Asp 395 below rhamnose I. The group of residues drawn to the left of Trp 391 belong to the dorsal fin subdomain of tailspike and strongly contribute to the hydrophobic side-chain interactions of Trp 391. Among other interactions, Glu 359 is involved in a salt bridge to Lys 363, here to the right of Glu 359. Plot drawn using MOLMOL.[59]

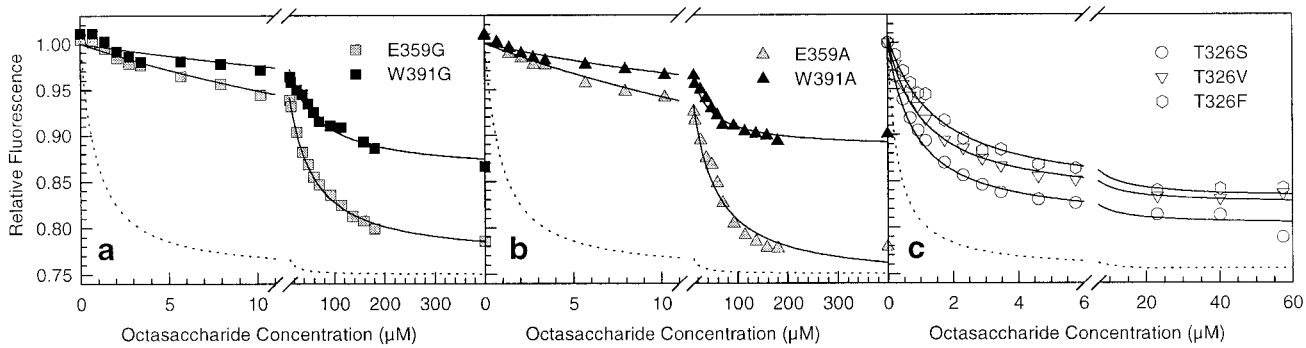

Fig. 3. Fluorescence titrations with octasaccharide for the determination of the dissociation constants at 10°C. The dotted line indicates the binding isotherm for wt tailspike, for the purpose of comparison. All data were normalized by the fluorescence intensity of unliganded protein as obtained from the fit. The variation in the amplitude of the fluorescence quench is due to the slightly different content of impurities in different oligosaccharide preparations, and the missing tryptophan residue in the case of W391G and W391A.

minimizing the possible factors contributing to changes in stability brought about by the introduction of other side chains. Under these conditions, mostly changes in hydrophobicity and size of the amino acid should influence stability and folding. A threonine residue was chosen, because valine could be introduced as an isosteric and serine as another uncharged polar substitution, giving us the possibility to observe the effect of very slight changes on the structure. Additionally, Thr 326 was exchanged to Phe in order to test the tolerance of the β-helix domain towards a major increase in side-chain volume and the resulting steric hindrance. In the case of a possible destabilization of the protein, a *tsf* phenotype was expected, in correspondence with the mechanism of folding mutations in tailspike presented above, which suggests a crucial role of native-like intermediates at the switch between folding and aggregation on the assembly pathway of the protein.

Effects on Tailspike Function

Both binding of an octasaccharide and hydrolysis of a dodecasaccharide fragment derived from *Salmonella enteritidis* lipopolysaccharide have been characterized in detail for tailspike protein.[41] We used the methods described to investigate a possible influence of the mutations on the functional properties of tailspike (Fig. 3). As expected for conservative exchanges like T326S and T326V of a residue that is not involved in function according to the crystal structure, no significant reduction in substrate affinity was observed. Only for the bulky Phe mutation, a slight but significant increase of the dissociation constant by about a factor of two was observed compared to wt (Fig. 3c, Table I). More prominent effects are caused by mutating residues Glu 359 and Trp 391. As indicated by the crystal structure, both residues interact with the substrate, and for Glu 359 an involvement in the catalytic mechanism is

TABLE I. Binding, Stability and Folding of Tailspike Mutants

TSPΔN mutant	K_D (μM)[a]	k_{denat}[b]	k_{fold} (10^{-2} s^{-1})[c]	Refolding yields[d]
wt	0.8 ± 0.1	1.0	1.5 ± 0.2	0
T326S	0.8 ± 0.2	1.7	1.8 ± 0.2	0
T326V	1.1 ± 0.2	2.5	1.6 ± 0.1	0
T326F	1.5 ± 0.2	5.0	0.66 ± 0.04	−
E359A	36 ± 7	16	1.6 ± 0.1	0
E359G	36 ± 5	2.0	1.4 ± 0.1	0
W391A	25 ± 3	1.8	1.8 ± 0.2	0
W391G	48 ± 11	1.9	1.5 ± 0.1	0
su V331A	10 ± 3[f]	0.20	2.6 ± 0.2[e]	+[e]
su V331G	(1.7 ± 0.3) · 10^2 [f]	0.20	2.0 ± 0.1[e]	+[e]
tsf G244R	n.d.	10	1.1 ± 0.1[e]	−[e]

[a]Dissociation constants for octasaccharide binding as determined by fluorescence titration at 10°C.
[b]Rate of denaturation at 71°C (E359A and W391A), 73°C (E359G and W391G) or 69°C (T326S, T326V, and T326F) relative to wt ($3.6 \cdot 10^{-3}$ min^{-1} at 71°C, $1.9 \cdot 10^{-2}$ min^{-1} at 73°C, and $1.4 \cdot 10^{-2}$ min^{-1} at 69°C). The values shown result from global fits of at least five independent measurements to single exponentials. Different temperatures had to be used to optimize the differentiation of the unfolding rate constants. Assuming the applicability of Arrhenius' law for these reactions, the ratios of reaction rates do not vary significantly with temperature if the differences in temperature are small compared to the absolute temperature used (see text).
[c]Rate of subunit folding at 25°C (as determined by nonlinear regression using single exponentials) ± standard deviation (at least 3 independent measurements).
[d]Influence on the temperature dependent refolding yields relative to wt. 0: no significant difference; +: higher yields than wt at temperatures above 25°C; −: lower yields than wt at temperatures above 25°C.
[e]Taken from Miller et al.[39] for the purpose of comparison.
[f]Taken from Baxa et al.[50] for the purpose of comparison.

suggested (Fig. 2, see below). A clear increase in K_D for octasaccharide binding by a factor of about 30 to 60 in the four mutants compared to wt confirms the essential role of these two residues in receptor binding in the wt structure (Fig. 3a and b, Table I). It should also be noted that removal of the Trp residue at position 391 coincides with a drastic decrease in the total amplitude of the fluorescence quench upon titration with octasaccharide by almost a factor of two, in agreement with the assumption that static quench of both Trp 365 and Trp 391 is responsible for the quench amplitude found in wt protein.

Similarly, exchanging Gly or Ala for Glu 359 or Trp 391 obliterates the activity of the protein as measured by hydrolysis of fluorophore-labelled dodecasaccharide. Prolonged incubation (up to 88 h) of the mutant proteins with this substrate at 37°C yields specific but minute amounts of product, with a corresponding decrease in the catalytic rate by at least a factor of 10^4 compared to wt. This impairment of tailspike activity would certainly be sufficient to exclude a detection of the corresponding mutations in an in vivo screen relying on intact function of the protein during host infection by the virus. Consequently, it corroborates the assumption that possible changes in folding and stability brought about by mutating Glu 359 or Trp 391 can only be investigated by directed mutagenesis and an in vitro analysis independent of protein function.

Effects on Tailspike Stability

Equilibrium folding transitions, the method usually employed to measure the conformational stability of proteins, can neither be applied for the complete nor for the N-terminally shortened tailspike (TSPΔN) used here, as its unfolding at high concentrations of denaturant or elevated temperature is not reversible on an accessible timescale. Hence, denaturation curves exhibit marked apparent hysteresis and cannot be analyzed thermodynamically, which forces us to resort to unfolding kinetics at high temperature to obtain information about the differences in stability brought about by mutations.[39] Protein unfolding kinetics reflect the difference in stability between the native state and the rate-limiting transition state of unfolding, and may thus be affected by changes in either state. However, because many native interactions are broken in the transition state, unfolding kinetics are very sensitive to changes in the native structure, and there is a good correlation between the unfolding rates and the free energies of folding for a series of mutants of a given protein.[51] Moreover, a clear correlation between the rate of thermal unfolding and the folding phenotype has been found for the tailspike mutants investigated so far.[28,39] In this context, it should also be noted that effects on the stability of the unfolded state of the protein do not affect unfolding kinetics, which is of particular importance for mutations to Gly leading to an increased conformational entropy of the unfolded state. Ignoring the lag phase at the beginning of the denaturation kinetics, thermal unfolding of the N-terminally shortened tailspike can be described by a single exponential.[39] The fits and resulting rate constants for the different mutants are shown in Figure 4 and Table I. In order to obtain optimal differentiation of the unfolding rates, the experiments had to be performed at slightly different temperatures, as the absolute rates are very sensitive to this parameter. But, assuming the approximate applicability of Arrhenius' law for these reactions, the ratios of the reaction rates measured with different mutants at the same temperature should not vary significantly with temperature, if the differences in temperature are small compared to the value of the absolute temperature used. Experiments performed at different temperatures with the same mutants support this approximation (unpublished results). In the following we thus only discuss acceleration or deceleration of unfolding compared to wt (Table I).

For the mutations at position 326, the slightest variation was observed for 326 Ser, which accelerated unfolding by only a factor of 2. Slightly more pronounced is the effect of Val with an acceleration factor of 3. As expected for the large Phe side chain, which can only with difficulties be accommodated in the interior of the tightly packed β-helix, a more drastic destabilization of the native structure is the outcome. An increase by a factor of 5.0 is still relatively small compared to known temperature sensitive mutants (see discussion), but a folding phenotype different from wt might already be expected. A more complex picture arises for the substitutions of Glu 359 and Trp 391. Removing the atoms distal to Cβ from the side chain leads to a slight

TABLE II. Data Collection and Refinement for the Mutants

	T326F	E359G	W391A
Data collection			
Resolution range (Å)	15 to 2.0	8 to 1.8	15 to 2.5
Total observations	167323	232408	54352
Unique reflections	38358	53301	19727
Completeness (%)	96.8 (15 to 2.0 Å)	98.8 (8.0 to 1.8 Å)	96.2 (15 to 2.5 Å)
	79.7 (2.11 to 2.0 Å)	98.3 (1.84 to 1.80 Å)	97.2 (2.63 to 2.5 Å)
R_{merge} (%)[a]	9.4 (15 to 2.0 Å)	7.1 (8.0 to 1.8 Å)	13.2 (15 to 2.5 Å)
	23.3 (2.11 to 2.0 Å)	32.4 (1.84 to 1.80 Å)	33.0 (2.63 to 2.5 Å)
Refinement			
Resolution range (Å)	8.0 to 2.0	8.0 to 1.80	8.0 to 2.5
Unique reflections for $I > 0\sigma(I)$	37769	53301	19246
R.m.s.-deviation			
bonds (Å)	0.009	0.009	0.010
angles (degree)	1.64	1.61	1.67
dihedrals	26.91	27.04	27.02
improper	1.29	1.302	1.36
R-factor (%)[b] for $I > 0\sigma(I)$	15.6	18.4	14.9

[a] $R_{merge} = \Sigma_h \Sigma_i \|I(h,i)| - |\langle I(h)\rangle\|/\Sigma_h \Sigma_i I(h,i)$; where $I(h,i)$ is the intensity of the ith measurement of the reflection h, $\langle I(h)\rangle$ is the mean of h for all i measurements of h. The sum is over all reflections.
[b] $R = (\Sigma|F_o - F_c|)/\Sigma F_o$; where F_o is the observed and F_c is the structure factor amplitude calculated from the model. The sum is over all reflections.

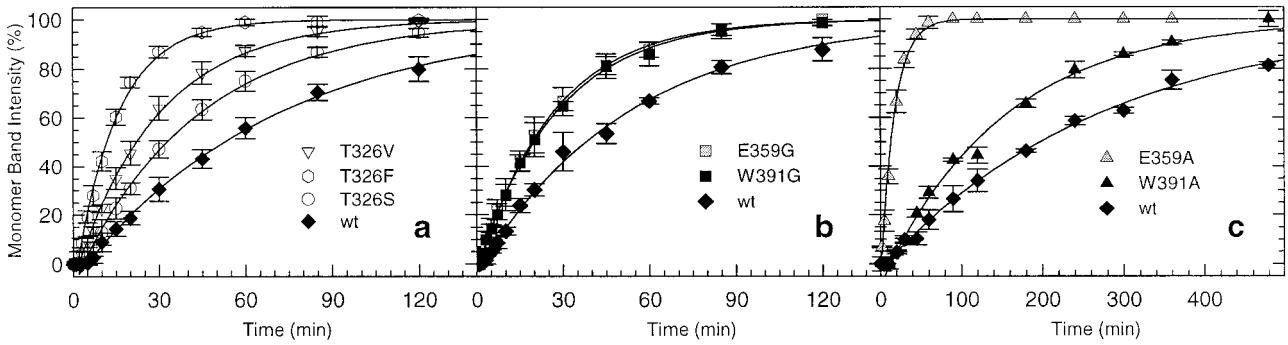

Fig. 4. Thermal unfolding kinetics in the presence of SDS at 69°C **(a)**, 73°C **(b)**, and 71°C **(c)**, respectively. At the times indicated, the reactions were stopped by rapid cooling and analyzed by SDS gel electrophoresis followed by densitometry. The fraction of monomer band intensity—corresponding to denatured protein—is plotted against time. TSPΔN wt was included in each experiment for comparison.

destabilization at position 391, reflected by an acceleration of unfolding by a factor of 2, but has a huge effect in the case of Glu 359: here, thermal unfolding is accelerated by a factor of 20, a value higher than for several genetically isolated *tsf* mutants. The mutation to Gly leaves the stability unaffected at site 391 (acceleration factor 2), but the exchange of Ala 359 to Gly leads to a drastic decrease in the rate of unfolding—only a factor of 2 compared to wt as opposed to 20 in E359A. Summing up, all mutations lead to an increase in the rate of thermal unfolding, probably corresponding to a destabilization of the native state.

Effects on Tailspike Folding

Two parameters have proved particularly useful to measure the effect of genetically isolated tailspike folding mutations on the folding process: subunit folding kinetics and the temperature dependence of refolding yields.[25] We thus use the same experiments here to analyze the generated mutants.

Two phases are observed during refolding of tailspike monitored by fluorescence or CD, the first of which is complete in the dead time of manual mixing[24] and low-resolution stopped flow experiments (~100 ms, unpublished results). The second phase, on the other hand, with a half-time of about 0.6 min for wt at 25°C, corresponds to the folding at the subunit level[24] and can easily be observed using fluorescence. Neither T326S nor T326V show any significant differences to the folding kinetics of wt (Table I), again confirming the conservative character of these mutations. T326F, in contrast, exhibits folding kinetics very much resembling those of known *tsf* mutants: both folding rate and amplitude of the observable phase are clearly decreased (Fig. 5a). All mutations investigated at positions 359 and 391 are not significantly different from wt protein in their folding kinetics (Table I).

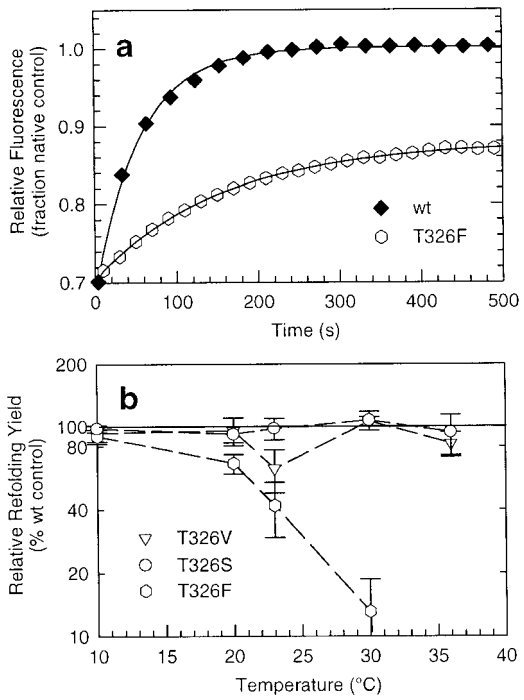

Fig. 5. Folding of T326F. **a:** Kinetics of refolding upon rapid dilution of denatured protein in buffer without denaturant at 25°C, as measured by the fluorescence increase at 342 nm. A representative wt trace is included for the purpose of comparison. After a rapid increase of fluorescence in the dead time of manual mixing, a slower phase corresponding to subunit folding is observed. For T326F, both rate and amplitude of this phase are clearly decreased (see Table I), resembling the characteristics of known *tsf* mutants. Single exponential fits are represented by solid lines. **b:** Temperature sensitivity of T326F refolding. Yields of renaturation are plotted relative to the yields observed for wt at the same temperature. Reconstitution yields of TSPΔN wt trimers decrease from about 80% at 10°C to 0 above 40°C. No refolding is observed for T326F above 30°C.

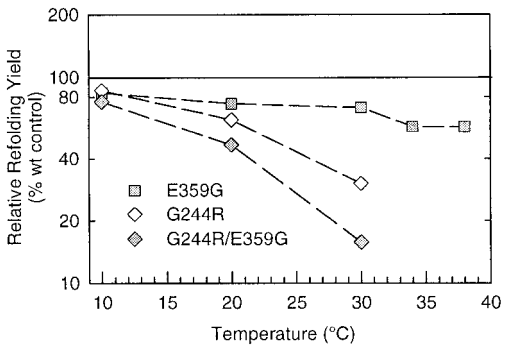

Fig. 6. Effect of E359G on the temperature sensitivity of refolding of *tsf* mutant G244R. Yields of renaturation are plotted relative to the yields observed for wt at the same temperature. Reconstitution yields of TSPΔN wt trimers decrease from about 80% at 10°C to 0 above 40°C. No significant differences could be detected for the double mutant in comparison to the G244R single mutant, excluding an effect of E359G analogous to known *su* mutations.

The more sensitive criterion for the influence of mutations on tailspike folding is the temperature dependence of the folding efficiency as depicted in Figure 5b. Although the relative error is high with this method, clear differences can be observed between wt, mutants stabilizing, and mutants destabilizing crucial folding intermediates.[25,28] No significant deviations from wt behavior are detectable for T326S and V, E359A and G, or W391A and G, particularly in the upper temperature range, which is very sensitive for differences in the thermostability of folding intermediates.[39] Again, T326F deviates markedly from the wt reference. Whereas the folding yields for wt converge to zero at about 40°C, no formation of native protein is observed for T326F at 36°C already; at 30°C, the yield is decreased by a factor of 8, but at low temperatures, folding yields of T326F approach those of the wild-type. This corresponds well to its decreased stability as measured by thermal denaturation kinetics, and resembles the behavior known from *tsf* mutants. For all other mutants, even those that show faster denaturation than wt, no effect on folding could be observed.

Especially in one of the *su* mutations known for tailspike, A334I, the characteristic phenotype suppressing the decrease in folding yields at elevated temperature brought about by *tsf* mutations is not apparent for the *su* mutation in wt background, but can only be observed in the *tsf* mutant background. This can be explained by the mutations having different or even reciprocal effects on the stability of early and late folding intermediates or the native structure.[37] Therefore, we tested this possibility for E359G, one of the mutants with presumptive release of backbone strain isolated here, as a representative example. The well-characterized *tsf* mutant G244R was chosen as a background, and the results of the refolding yields of wt, G244R and the double mutant G244R/E359G are illustrated in Figure 6. Clearly, the double mutant does not show any increase in yields compared to the *tsf* mutant alone. The convergence to zero renaturation is about 34°C in both cases; at lower temperature even a slight decline in folding yields becomes apparent, whose significance is doubtful in view of the relatively large error inherent in this experimental technique. Consequently, no suppression of the *tsf* phenotype of G244R results upon combination with E359G, which excludes the possibility that this mutation preferentially affects early folding intermediates in a way similar to A334I.

Effects on Tailspike Structure

According to spectroscopic data, including UV absorption, fluorescence, and circular dichroism, all tailspike mutations introduced in this study did not affect the integrity of the native three-dimensional structure (unpublished observations). Moreover, the octasaccharide binding experiments described above confirm the presence of native structure, although for some of the mutants with decreased receptor affinity. But to ensure complete validity of our conclusions, and to be able to assess the local rearrangements resulting from mutating the protein, we crystallized the mutants T326F, E359G, and W391A, and solved the crystal structures by molecular replacement (Table II). For E359G, there are no significant differences detectable compared to the wt structure, apart from the obviously missing side chain. In W391A, the side chains in

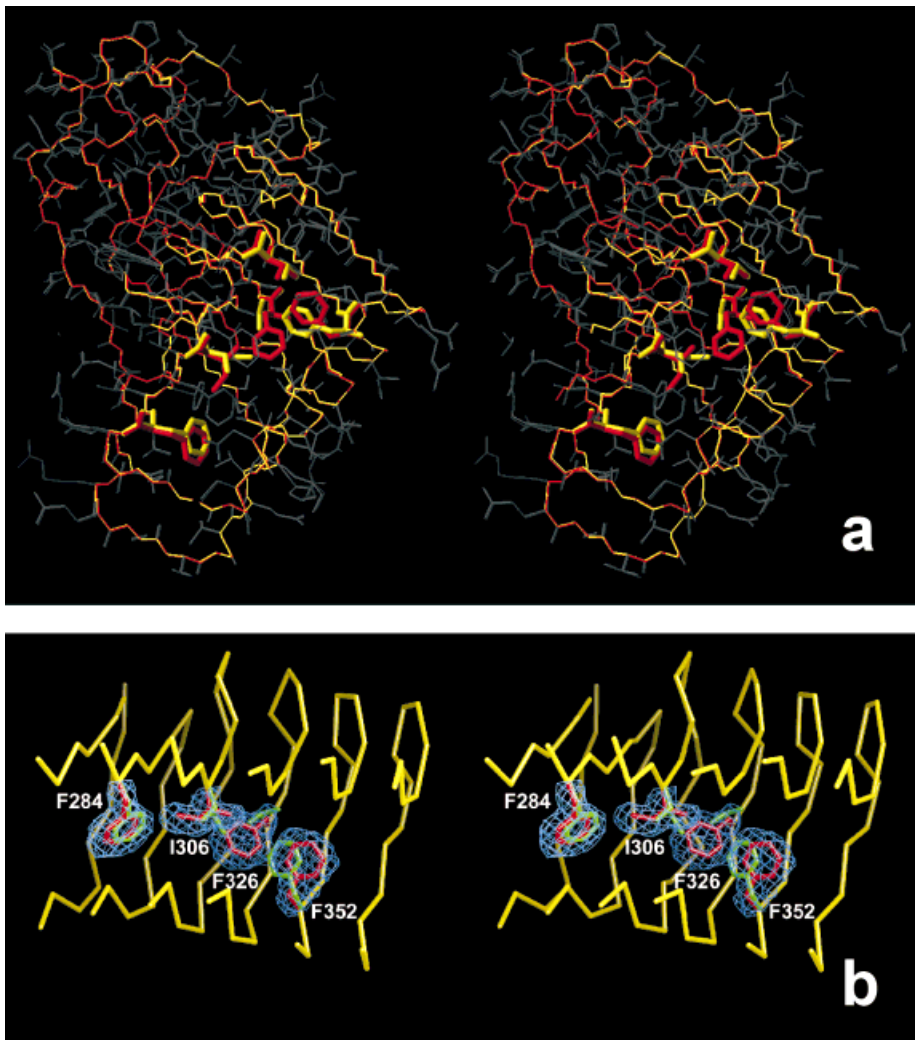

Fig. 7. **a:** Detectable changes in the crystal structure of TSPΔN caused by the mutation T326F. The wt structure is drawn in yellow, the mutant structure in red. The side chains of residues F284, I306 T/F326, F352, and V362 are colored, all other side chains are in dark grey. For details, see text. **b:** Stereoview of the $2F_o - F_c$ electron density map of T326 and F326 in the superimposed wt and mutant structures contoured at 1σ.

the vicinity of the mutation site slightly rearrange, indicating relaxation due to the missing interactions of neighboring residues with Trp 391. In T326F, on the other hand, a major reorganization of side chains in the interior of the β-helix was observed (Fig. 7). The introduction of Phe 326 forces Phe 352 into a very unfavorable side-chain conformation ($\chi_1 = 177°$, $\chi_2 = 69°$ for wt, $\chi_1 = 112°$, $\chi_2 = 138°$ for T326F[52]), associated with a displacement of the Phe-C_ζ by 2.9 Å, and leading to a slight propagation of the steric clashes to the side chain of Val 362, located in β-sheet C on the opposite side of the helix. Furthermore, Phe 326 induces a rotation of the Ile 306 side chain by 120° about χ_1, causing C_ζ of Phe 284 to move by about 1 Å. Remarkably, no significant changes in the position of backbone atoms and bond angles could be identified.

DISCUSSION

In search of a more detailed understanding of the factors that stabilize the β-helix domain of P22 tailspike protein, we introduced several mutations at three sites by directed mutagenesis, and related their stability, folding, and structural context to folding mutants of known phenotype.

Mutating Thr 326, which is situated in one of the β-sheets, and whose side chain is buried in the interior of the β-helix, to Ser, Val, and Phe leads to two distinct folding phenotypes. The Ser and Val mutants on the one hand differ from wt neither in their refolding kinetics, refolding yields at varied temperature, nor oligosaccharide binding. Only in the kinetics of thermal denaturation, some destabilization relative to wt can be observed. This indicates that neither removal of a methyl group (T326S) nor substitution of a methyl- for a hydroxyl group (T326V) cause any detectable impairment of folding, in spite of a slight destabilization of the protein. The loss of stability by decreased side-chain packing or introduction of a hydrophobic cavity is quite straightforward for T326S, but the destabilization of Thr 326 by the nearly isosteric Val is less obvious, as no hydrogen bond acceptor is apparent in the vicinity of the side chain, although Ser 333 and Ser 354 might provide a hydrophilic environment. Here, the slight rearrangements brought about by the changed character of the side chain seem to be dominant over the dehydration free energy that has to be summoned up for burying a hydroxyl group in the core.

The second phenotype found at position 326, represented by T326F, very closely resembles the behavior of known, genetically isolated *tsf* mutants: its folding kinetics are retarded, relative refolding yields show a pronounced decrease at temperatures above 20 °C, and thermal denaturation is accelerated by a factor of five with respect to wt. Incorporation of an additional phenyl moiety into the tightly packed core of the β-helix causes significant deformation of neighboring side-chain packing and interactions. Particularly striking is the altered conformation of Phe 352, which corresponds to a rotamer very rarely found in known protein structures[52] and is considered to be energetically very unfavorable. The steric strain enforced by this mutation substantially destabilizes the native state and at the same time, a *tsf* phenotype is clearly observed. This is in agreement with the hypothesis that an aggregation-prone, thermolabile folding intermediate with native-like structure on the assembly pathway of tailspike is similarly destabilized by these mutations as the native structure, and consequently shows increased aggregation compared to wt at restrictive temperature.[25] Yet, the core of the tailspike β-helix appears to be remarkably plastic, and tolerant of major distortions that can be compensated for by structural rearrangements propagated through large parts of the helix lumen, whereas the backbone remains largely invariant, indicating enormous rigidity due to optimized hydrogen bonding. This is in agreement with the extremely low crystallographic temperature factors found for backbone atoms in this part of the structure.[6,7] Accordingly, as the scaffolding of the β-structure and thus the exterior remains invariant upon mutation, the creation of specific aggregation sites on the surface of the protein can be excluded as a mechanism for its increased aggregation tendency.

Our findings with this *tsf* mutant—the first one generated by rational design—relate well to results from site-saturation mutagenesis experiments at positions 334,[36] 235, 238, and 244.[53] Ala 334 is located in the substrate binding groove of tailspike with its side chain directed towards the interior of the β-helix. Lee et al. found that introduction of large polar (Gln, Asn) or hydrophobic (Met) residues resulted in a *tsf* phenotype, whereas charged or even larger side chains prevented folding completely, even at low temperature. At the *tsf* sites 235, 238, and 244 in the dorsal fin domain of tailspike, only the wt and one or two very similar residues are tolerated, indicating that specific native-like interactions have to be maintained in order to sufficiently stabilize intermediates and thus prevent their aggregation.

Residues Glu 359 and Trp 391 show aberrant backbone dihedrals in the wt structure of tailspike, similar to those of neighboring residues in a stack located in the turn region between β-sheets B and C, next to the active site (Fig. 2). Such Ramachandran outliers are rare in high-resolution crystal structures of proteins and thus in many cases represent residues with very specific requirements of structure, often connected with function.[54,55] Replacing Glu 359 or Trp 391 by Gly is expected to release the backbone strain completely, whereas replacement by Ala almost exclusively obliterates side-chain interactions with only little release of steric strain. The determination of the three-dimensional structures of E359G and W391A confirmed our assumption that the removal of a mostly solvent-exposed side chain does not strongly disturb the structure.

In agreement with their role in substrate binding and possibly catalysis suggested by the crystal structure, the mutations at both positions decrease the binding free energy for octasaccharide by about 10 kJ/mol, and essentially eliminate activity. This would certainly be expected for Glu 359, which is intimately involved in interactions with the sugar and has been suggested to act as a general base together with Asp 395 in the presumed catalytic mechanism of the endorhamnosidase.[7,21] The crucial role of Trp 391 for function, on the other hand, is less understandable from the crystal structure. Its only obvious interaction with the bound oligosaccharide is a hydrophobic contact to the 6-methyl group of the rhamnose. That it does not seem to be dispensable may also be due to an effect on the hydrophobic environment of Asp 392,[7] which probably serves as a general acid during hydrolysis. The data show that there is a clear necessity for the protein to conserve these amino acids in the course of evolution in order to retain its specific function essential for replication of the phage.

The effects of the Ala and Gly mutants are not only similar in terms of function but also in terms of stability and folding of P22 tailspike protein. The results (Fig. 4, Table I) show that all four mutations destabilize TSPΔN as reflected by increased rates of denaturation at about 70°C. Although this indicates already that backbone strain is not a dominant destabilizing factor for Glu 359 and Trp 391, the differential effects of Ala and Gly mutants are worth noticing. For Glu 359, removal of the side chain down to the methyl group in E359A accelerates thermal unfolding by a factor of about 20, reflecting the loss of most of the side-chain interactions while still retaining substantial steric strain. Further removal of the methyl group in E359G releases this steric strain with the only trade-off being the non-covalent interactions of the methyl group. Accordingly, stability is again increased compared to E359A. But still, the Gly mutant is less stable than wt, the inescapable conclusion of which is that interactions of the side chain overcompensate for the destabilization of the protein due to steric strain. A similar overcompensation is found for Trp 391, although with other differential effects for the two mutants. Here, both the Ala and the Gly variant exhibit a twofold increase in the rate of thermal denaturation. Two possible explanations are that either (1) Trp 391 is less sterically strained than Glu 359, reflected by the slightly closer proximity to the allowed region in the Ramachandran diagram, or (2) mutating Trp 391 leads to a noticeable change of the energetics of the transition state ensemble, thus rendering the conclusion irrelevant that an acceleration of the thermal unfolding kinetics corresponds to a destabilization of the native state. At the present state of knowledge, this detail cannot be answered, but leaves the corollary unaffected that in

both cases the backbone strain in the native structure is balanced, and even outweighed by side-chain interactions.

Regarding the folding phenotype of the protein as assayed by fluorescence measurements of refolding kinetics and the determination of renaturation yields at different temperatures, no significant differences to wt were observed for the 359 and 391 mutants. A related observation has been made for A334I, located in the β-helix core: although thermal denaturation is highly accelerated, this mutant does not show a *tsf* phenotype, it even acts as a global suppressor in a *tsf* background.[36] This was explained by improved hydrophobic stacking mediated by the enlarged hydrophobic side chain in loosely structured intermediates on the one hand, but steric clashes in the tightly packed native structure on the other hand.[37] In addressing the question of whether the mutants investigated here might similarly act as suppressors of the *tsf* phenotype, we additionally created a double mutant containing both E359G and the well-characterized *tsf* mutation G244R. E359G, in combination with the *tsf* mutation, is not able to avoid the excessive aggregation of the destabilized intermediates at elevated temperature, and can thus not be classified as a suppressor mutation. Rather do the mutants at positions 359 and 391 created here, together with T326S and T326V, constitute a new folding phenotype for tailspike, characterized by wt folding, but simultaneously impaired stability of the native state. This may be another example of different requirements upon the side-chain geometry in intermediates and the native structure, respectively.

This leaves the question why the mutations at positions 359 and 391 are so different from mutations at position 331. Val 331 is a typical example of an amino acid residue with unfavorable backbone dihedral angles in so far as it is involved in the function of tailspike. And although both V331G and V331A result in a more stable protein with improved folding properties, Val is invariably found in the wt, because both mutants bind lipopolysaccharide with much lower affinity.[50] Similar observations of this evolutionary compromise between protein stability and function have been made for a large number of amino acid residues, indicating that the structural requirements of function are more stringent compared with those of folding and stability (Shoichet et al.,[55] and references therein). Of particular relevance to this study, a statistical analysis of known protein structures revealed that sterically-strained backbone conformations are very often located in regions concerned with function.[54] Making use of this rule, it is often possible to find mutations of active-site residues that lead to an increased conformational stability of the protein while at the same time reducing activity or ligand binding affinity. Whereas V331A and V331G agree with this rule, the 359 and 391 mutants created here are even slightly less stable than wt—apparently, because the side-chain interactions present in the wt overcompensate for the steric strain in the backbone. From the structure it is evident that the possibilities for side-chain interactions are much less prominent for Val 331 than for Glu 359 or Trp 391 (Fig. 2). The highly solvent-exposed isopropyl group of the former participates in very few non-covalent interactions, whereas the hydrophilic Glu undergoes extensive hydrogen bonding and seems to form a salt bridge to Lys 363; Trp 391 packs against the dorsal fin of tailspike, a long loop protruding from the β-helix domain that folds back on the rim of the binding groove, and this way forms various hydrophobic interactions. It seems as though tailspike has managed to circumvent the destabilization often caused by residues with unfavorable backbone angles in these two cases by developing an extended network of side-chain interactions that make up for the loss in stability. This shows that thermostability is a dominant factor in the evolution of this protein. Moreover, these strong interactions formed by the long side chains Trp and Glu obviously are an extreme example of a more general observation made recently. The distribution of the main-chain dihedral angels in crystallographic structures[56] yields a value for the conformational entropy of Ala relative to Gly very similar to the one found in experimental and computational analyses.[57] But for other residues, i.e., longer side chains, the data base approach grossly overestimates the backbone conformational entropy, corresponding to larger areas of allowed conformations in the Ramachandran diagram. This might very well be caused by the side-chain interactions being able to distort their backbone geometry more than possible for Ala.

CONCLUSION

In our characterization of the effects of seven rationally designed mutations on P22 tailspike protein and a comparison to the properties of known folding mutants, we find that the *tsf* phenotype, which has been observed many times in genetically isolated mutants, can be designed by introducing an exchange in the protein core expected to destabilize the β-helix domain of tailspike, realized here by T326F. This supports the notion that changes in stability of an intermediate in which the β-helix domain is largely formed are the essential mechanism of tailspike folding mutations. The crystal structure of T326F reveals a high degree of plasticity of the side-chain arrangement in the lumen of the β-helix, whereas the backbone involved in β-sheet formation with neighboring helix coils appears to be very rigid. T326F is the first *tsf* mutant of tailspike created by rational design.

Finding locations in the tailspike protein that are not optimized for stability, on the other hand, seems to be very difficult. Even residues with a highly strained backbone, which are frequently found in the active site regions of proteins, are optimized for stability via an overcompensation of steric strain by extensive side-chain interactions, as in the cases of Glu 359 and Trp 391. This finding is supported by the fact that only two *su* mutations have repeatedly been isolated by extensive genetic screening.[58] Additionally, it exemplifies the view arising from the statistical analysis of the distribution of backbone dihedral angles in known three-dimensional protein structures that the adoption of ϕ/ψ angles other than the most favorable ones can be caused by side-chain interactions, and illus-

trates the extreme optimization of the tailspike protein for conformational stability.

ACKNOWLEDGMENTS

We thank Andrej Weintraub for providing lipopolysaccharide fragments, Ulrich Baxa for providing labelled lipopolysaccharide fragments, and Ulrich Baxa and Stefan Miller for helpful discussion.

REFERENCES

1. Yoder MD, Keen NT, Jurnak F. New domain motif: the structure of pectate lyase C, a secreted plant virulence factor. Science 1994;260:1503–1507.
2. Levitt M, Chothia C. Structural patterns in globular proteins. Nature 1976;261:552–558.
3. Richardson JS. The anatomy and taxonomy of protein structure. Adv Protein Chem 1981;34:167–339.
4. Sprang SR. On a (β-) roll. Trends Biochem Sci 1993;18:313–314.
5. Baumann U, Wu S, Flathery KM, McKay DB. Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif. EMBO J 1993;12:3357–3364.
6. Steinbacher S, Seckler R, Miller S, Steipe B, Huber R, Reinemer P. Crystal structure of P22 tailspike-protein, interdigitated subunits in a thermostable trimer. Science 1994;265:383–386.
7. Steinbacher S, Miller S, Baxa U, Budisa N, Weintraub A, Seckler R, Huber R. Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 Å, fully refined structure of the endorhamnosidase at 1.56 Å resolution, and the molecular basis of O-antigen recognition and cleavage. J Mol Biol 1997;267:865–880.
8. Pickersgill R, Jenkins J, Harris G, Nasser W, Robert-Baudouy J. The structure of *Bacillus subtilis* pectate lyase in complex with calcium. Nature Struct Biol 1994;1:717–723.
9. Lietzke SE, Keen NT, Yoder MD, Jurnak F. The three-dimensional structure of pectate lyase E, a plant virulence factor from *Erwinia chrysanthemi*. Plant Physiol 1994;106:849–862.
10. Raetz CRH, Roderick SL. A left-handed parallel β helix in the structure of UDP-*N*-acetylglucosamine acyltransferase. Science 1995;270:997–1000.
11. Kisker C, Schindelin H, Alber BE, Ferry JG, Rees DC. A left-handed β-helix revealed by the crystal structure of a carbonic anhydrase from the archaeon *Methanosarcina thermophila*. EMBO J 1996;15:2323–2330.
12. Emsley P, Charles IG, Fairweather NF, Isaacs NW. Structure of *Bordetella pertussis* virulence factor P.69 pertactin. Nature 1996; 381:90–92.
13. Mayans O, Scott M, Connerton I, Gravesen T, Benen J, Visser J, Pickersgill R, Jenkins J. Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate-binding clefts of pectin and pectate lyases. Structure 1997;5:677–689.
14. Petersen TN, Kauppinen S, Larsen S. The crystal structure of rhamnogalacturonase A from *Aspergillus aculeatus*: a right-handed parallel β helix. Structure 1997;5:533–544.
15. Beaman TW, Binder DA, Blanchard JS, Roderick SL. Three-dimensional structure of tetrahydrodipicolinate N-succinyltransferase. Biochemistry 1997;36:489–494.
16. Vitali J, Schick B, Kester HCM, Visser J, Jurnak F. The three-dimensional structure of *Aspergillus niger* pectin lyase b at 1.7-Å resolution. Plant Physiol 1998;116:69–80.
17. Pickersgill R, Smith D, Worboys K, Jenkins J. Crystal structure of polygalacturonase from *Erwinia carotovora* ssp. *Carotovora*. J Biol Chem 1998;38:24660–24664.
18. Beaman TW, Sugantino M, Roderick SL. Structure of the hexapeptide xenobiotic acetyltransferase from *Pseudomonas aeruginosa*. Biochemistry 1998;12:6689–6696.
19. Yoder MD, Jurnak F. The parallel β helix and other coiled folds. FASEB J 1995;9:335–342.
20. Heffron S, Moe GR, Siebert V, Mengaud J, Cossart P, Vitali J, Jurnak F. Sequence profile of the parallel β helix in the pectate lyase superfamily. J Struct Biol 1998;122:223–235.
21. Steinbacher S, Baxa U, Miller S, Weintraub A, Seckler R, Huber R. Crystal structure of phage P22 tailspike protein complexed with *Salmonella sp.* O-antigen receptors. Proc Natl Acad Sci USA 1996;93:10584–10588.
22. Goldenberg DP, Berget PB, King J. Maturation of the tail spike endorhamnosidase of bacteriophage P22. J Biol Chem 1982;257: 7864–7871.
23. Goldenberg D, King J. Trimeric intermediate in the *in vivo* folding and assembly of the tail spike endorhamnosidase of bacteriophage P22. Proc Natl Acad Sci USA 1982;79:3403–3407.
24. Fuchs A, Seiderer C, Seckler R. *In vitro* folding pathway of the P22 tailspike protein. Biochemistry 1991;30:6598–6604.
25. Danner M, Seckler R. Mechanism of phage P22 tailspike protein folding mutations. Protein Sci 1993;2:1869–1881.
26. Betts SD, Haase-Pettingell C, King J. Mutational effects on inclusion body formation. Adv Prot Chem 1997;50:243–264.
27. Haase-Pettingell CA, King J. Formation of aggregates from a thermolabile in vivo folding intermediate in P22 tailspike maturation. A model for inclusion body formation. J Biol Chem 1988;263: 4977–4983.
28. Danner M, Fuchs A, Miller S, Seckler R. Folding and assembly of phage P22 tailspike endorhamnosidase lacking the N-terminal, head-binding domain. Eur J Biochem 1993;215:653–661.
29. Yu M-H, King J. Single amino acid substitutions influence the folding pathway of phage P22 tail spike endorhamnosidase. Proc Natl Acad Sci USA 1984;81:6584–6588.
30. Yu M-H, King J. Surface amino acids as sites of temperature-sensitive folding mutations in the P22 tailspike protein. J Biol Chem 1988;263:1424–1431.
31. Villafane R, King J. Nature and distribution of temperature-sensitive folding mutations in the gene for the P22 tailspike polypeptide chain. J Mol Biol 1988;204:607–619.
32. Goldenberg D, King J. Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tailspike protein. II. Active mutant proteins matured at 30°C. J Mol Biol 1981;145:633–651.
33. Haase-Pettingell C, King J. Prevalence of temperature sensitive folding mutations in the parallel beta coil domain of the phage P22 tailspike endorhamnosidase. J Mol Biol 1997;267:88–102.
34. Fane B, Villafane R, Mitraki A, King J. Identification of global suppressors for temperature-sensitive folding mutations in the P22 tailspike protein. J Biol Chem 1991;266:11640–11648.
35. Mitraki A, Fane B, Haase-Pettingell C, Sturtevant J, King J. Global suppression of protein folding defects and inclusion body formation. Science 1991;253:54–58.
36. Lee SC, Koh H, Yu MH. Molecular properties of global suppressors of temperature-sensitive folding mutations in P22 tailspike endorhamnosidase. J Biol Chem 1991;266:23191–23196.
37. Beißinger M, Lee SC, Steinbacher S, Reinemer P, Huber R, Yu M-H, Seckler R. Mutations that stabilize folding intermediates of phage P22 tailspike protein: folding *in vivo* and *in vitro*, stability, and structural context. J Mol Biol 1995;249:185–194.
38. Schuler B, Seckler R. P22 tailspike folding mutants revisited: effects on the thermodynamic stability of the isolated β-helix domain. J Mol Biol 1998;281:227–234.
39. Miller S, Schuler B, Seckler R. Phage P22 tailspike protein: Removal of head-binding domain unmasks effects of folding mutations on native-state thermal stability. Protein Sci 1998;7: 2223–2232.
40. Pace CN. Determination and analysis of urea and guanidine hydrochloride denaturation curves. Methods Enzymol 1986;131: 266–280.
41. Baxa U, Steinbacher S, Miller S, Weintraub A, Huber R, Seckler R. Interactions of phage P22 tails with their cellular receptor, *Salmonella* O-antigen polysaccharide. Biophys J 1996;71:2040–2048.
42. Kunkel TA, Roberts JD, Zakour RA. Rapid and efficient site-directed mutagenesis without phenotypic selection. Methods Enzymol 1987;154:367–382.
43. Yuckenberg PD, Witney F, Geisselsoder J, McClary J. Site-directed in vitro mutagenesis using uracil-containing DNA and phagemid vectors. In: MacPherson MJ, editor. Directed mutagenesis, a practical approach. Oxford: IRL Press, 1991:27–48.
44. Pace CN, Vajdos F, Fee L, Grimsley G, Gray T. How to measure and predict the molar absorption coefficient of a protein. Protein Sci 1995;4:2411–2423.
45. Seckler R, Fuchs A, King J, Jaenicke R. Reconstitution of the thermostable trimeric phage P22 tailspike protein from denatured chains *in vitro*. J Biol Chem 1989;264:11750–11753.

46. Brunschier R, Danner M, Seckler R. Interactions of phage P22 Tailspike protein with GroE molecular chaperones during refolding *in vitro*. J Biol Chem 1993;268:2767–2772,
47. Leslie AGW. Recent changes to the MOSFLM package for processing film and image plate data. In: CCP4 & ESF-EACMB Newsletter on Protein Crystallography, SERC Laboratory, Daresbury, Warrington WA4 4AD, England, 1991.
48. Engh RA, Huber R. Accurate bond and angle parameters for X-ray protein structure refinement. Acta Crystallog Sect A 1991;47:392–400.
49. Laskowski RA, MacArthur MW, Moss DS, Thornton J. M. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 1993;26:283–291.
50. Baxa U, Steinbacher S, Weintraub A, Huber R, Seckler R. Mutations improving the folding of phage P22 tailspike protein affect its receptor binding activity. J Mol Biol 1999;293:693–701.
51. Klemm JD, Wozniak JA, Alber T, Goldenberg DP. Correlation between mutational destabilization of phage T4 lysozyme and increased unfolding rates. Biochemistry 1991;30:589–594.
52. Dunbrack RL Jr, Cohen FE. Bayesian statistical analysis of protein side-chain rotamer preferences. Protein Sci 1997;6:1661–1681.
53. Lee SC, Yu MH. Side-chain specificity at three temperature-sensitive folding mutation sites of P22 tailspike protein. Biochem Biophys Res Commun 1997;233:857–862.
54. Herzberg O, Moult J. Analysis of the steric strain in the polypeptide backbone of protein molecules. Proteins 1991;11:223–229.
55. Shoichet BK, Baase WA, Kuroki R, Matthews BW. A relationship between protein stability and protein function. Proc Natl Acad Sci USA 1995;92:452–456.
56. Stites WE, Pranata J. Empirical evaluation of the influence of side chains on the conformational entropy of the polypeptide backbone. Proteins 1995;22:132–140.
57. D'Aquino JA, Gómez J, Hilser VJ, Lee KH, Amzel LM, Freire E. The magnitude of the backbone conformational entropy change in protein folding. Proteins 1996;25:143–156.
58. Fane B, King J. Intragenic suppressors of folding defects in the P22 tailspike protein. Genetics 1991;127:263–277.
59. Koradi R, Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 1996;14:51–55.