The tricorn-interacting factor F1 of the archaeon *Thermoplasma acidophilum* cleaves small hydrophobic peptide products of the proteasome and tricorn protease. F1 mutants of the active site residues that are involved in substrate recognition and catalysis displayed distinct activity patterns toward fluorogenic test substrates. Crystal structures of the mutant proteins complexed with peptides Phe-Leu, Pro-Pro, or Pro-Leu-Gly-Gly showed interaction of glutamates 213 and 245 with the N termini of the peptides and defined the S1 and S1' sites and the role of the catalytic residues. Evidence was found for processive peptide cleavage in the N-to-C direction, whereby the P1 product is translocated into the S1 site. A functional interaction of F1 with the tricorn protease was observed with the inactive F1 mutant G37A. Moreover, small angle x-ray scattering measurements for tricorn and inhibited F1 have been interpreted as formation of transient and substrate-induced complexes.

Controlled degradation of intracellular proteins is performed in the majority of species initially by proteasomes and subsequently by associated peptidases. This protein degradation machinery of the model organism *Thermoplasma acidophilum* has been intensively studied and characterized by x-ray structures, such as the 20 S proteasome that appears to be a major component in the cleavage of oligopeptides produced by the proteasome. Recent investigations of tricorn structure and inhibited F1 have been inter- preted as formation of transient and substrate-induced complexes.

**EXPERIMENTAL PROCEDURES**

**Preparation of Mutant F1 Protein**—Site-directed mutagenesis was performed with PCR using *Pfu* Turbo (Stratagene) on the expression plasmid pRS616 (16) that contained the F1 gene of *T. acidophilum* (15). After DpnI (New England Biolabs) digest of the original plasmid, plasmid pRset6c (15) that contained the F1 gene of *T. acidophilum* was isolated. After DpnI (New England Biolabs) digest of the original plasmid, plasmid pRS616 (15) that contained the F1 gene of *T. acidophilum* was isolated. After DpnI (New England Biolabs) digest of the original plasmid, the isolated plasmid was transformed into *E. coli* DH5α competent cells (15). MIDI preparations of the plasmids (Qiagen) were used for DNA sequencing and direct transformation of BL21(DE3)RIL cells for expression (Stratagene).

**Preparation of Mutant F1 Protein**—Site-directed mutagenesis was performed with PCR using *Pfu* Turbo (Stratagene) on the expression plasmid pRS616 (15) that contained the F1 gene of *T. acidophilum* (15). After DpnI (New England Biolabs) digest of the original plasmid, the isolated plasmid was transformed into *E. coli* DH5α competent cells (15). MIDI preparations of the plasmids (Qiagen) were used for DNA sequencing and direct transformation of BL21(DE3)RIL cells for expression (Stratagene).

The mutant proteins were purified as described elsewhere for native recombinant F1 in three chromatographic steps on Q-Sepharose, octyl-Sepharose, and Superose 12 (15). The mutants were analyzed by mass spectrometry and found to have molecular masses of ± 2 Da differing from the expected calculated values.

**Activity Assays and Reverse Phase High Performance Liquid Chromatography Analysis of Insulin Degradation Products**—A test for the degradation of fluorogenic substrates by F1 has been described previously.
X-ray Snapshots of Peptide Processing in Mutants of F1

(5, 6) and was measured with 7-amido-4-methyl coumarin (AMC) substrates (Bachem). Oxidized insulin B-chain (Sigma-Aldrich) was incubated with each recombinant tricorn protease and F1 protein samples of wild type and mutants at 50 °C in a buffer composed of 20 mM Tris/HCl, pH 7.5, and 100 mM NaCl. For control reactions, an aliquot was taken immediately after substrates and the proteases were mixed. At different times, further aliquots were taken. Degradation products of insulin B-chain were analyzed by reverse phase high performance liquid chromatography on an X-Terra MS C8 reverse phase column and detected by UV absorption at 210 nm. The column was equilibrated with 5% acetonitrile (AcCN) and 95% of 2% H3PO4 and eluted with a linear gradient of 5 to 90% of AcCN and 95 to 10% of 2% H3PO4 in 20 min at a flow rate of 1.5 ml/min.

**Biacore Measurements**—Binding of tricorn, WT-F1, and F1 mutants was investigated with a Biacore 1000 device. Proteins were coupled in 2.4 mg/ml (20 μM) tricorn solution added at concentrations from 0.12 mg/ml (1 μM) up to 2.4 mg/ml (20 μM) in 50 mM bis-Tris, pH 6.5, and 100 mM NaCl. Samples with 2 μM tricorn and 5 μM Phe-Leu-Gly-Gly (PLGG) were measured under the same conditions. In another test series, tricorn was coupled to the chip matrix at 0.2 mg/ml and F1-G37A at 0.34 and 0.68 mg/l (10 and 20 μM). Moreover, F1-G37A and S105A were added in 50 mM bis-Tris, pH 6.0, and 100 mM NaCl at 10 and 20 μM concentrations; S105A was added at 50, 75, and 100 μM and PLGG at 5 μM PLGG.

**Small Angle X-ray Scattering and Data Processing**—The synchrotron radiation x-ray scattering data were collected on the X33 beamline (18, 19) at the European Molecular Biology Laboratory (EMBL) of the Deutsches Elektronen Synchrotron (DESY) using a wavelength of λ = 1.5000 Å and multibeam proportional chambers with delay line readout (20). The sample detector distance was 2.2 m covering the range of momentum transfer 0.015 < s < 0.35 Å−1 (s = 4π sin(θ)/λ with the scattering angle 2θ). Tricorn solutions were measured in 100-μl cuvettes at concentrations of 2.0, 4.0, 6.0 and 10.0 mg/ml and those of F1-G37A and S105A at 1.1:1.0 molar ratio of F1 to tricorn monomer) at several concentrations ranging from 1.9 to 7.1 mg/ml. To check for radiation damage the data were collected in 10 successive 1-min frames that were averaged after normalization to the beam intensity, corrected for the detector response and the background scattering of the buffer with the program package PRIMUS (21). The maximum dimensions of the particles Dmax were estimated with the orthogonal expansion program ORTOG-NOM (22). Parameters for forward scattering I(0) and the radius of gyration Rg were evaluated using both the Guinier approximation (23) and calculations from the entire scattering patterns with the program package GNOM (24), which also provides the distance distribution function p(r) of the particles. The effective molecular masses of the solutes were evaluated by comparison of the forward scattering with that from reference solutions of bovine serum albumin (66 kDa). For a quantitative description the experimental scattering intensity Iexp(s) was represented as shown in Equation 1

\[ I(s) = \sum_{j=1}^{K} v_j I_j(s) \]  

(Eq. 1)

where K is the number of components in the mixture and 0 ≤ vj ≤ 1 and Ij(s) are the volume fraction and the scattering intensity from the j-th component, respectively. The volume fractions were determined by the program OLIGOMER (21), which uses scattering curves Ij(s) of the components and solves linear equations with respect to the unknown values vj to minimize the discrepancy χ as shown in Equation 2

\[ \chi^2 = \frac{1}{N - 1} \sum_{j} \left[ \frac{\mu_j I_j(s) - I_{exp}(s)}{\sigma(s)} \right]^2 \]  

(Eq. 2)

where N is the number of experimental points, σj(s) are experimental errors, μj is a scaling coefficient, and Ij(s) the scattering intensity from the linear combination (1). For the mixtures of tricorn monomers and hexamers (K = 2 in Equation 1), the scattering patterns Ij(s) from the crystallographic models of monomeric and hexameric tricorn (Protein Data Bank entry 1K32) were calculated by the program CRYSTAL (25). The corresponding calculations for tricorn-F1 mixtures were performed with a model of F1 (Protein Data Bank entry 1MTZ). A tentative model of the tricorn-F1 complex was built containing three F1 molecules/tricorn hexamer bound to the three supposed docking sites, namely the β-6 propeller domains. The underlying 1:1 complex model was built manually in MAIN (26) according to the analogy of prolyl oligopeptidase and qualitatively confirmed by the docking program FTDOCK (28) as described previously (15).

**Crystallization of F1 Mutants**—The concentration for F1 mutants was 30 mg/ml mixed with equal volumes of 100 mM bis-Tris/HCl, pH 6.0, with 7–14% polyethylene glycol 6000 and yielded orthorhombic crystals of space group P21212 and P212121. Crystals were grown by the hanging drop vapor diffusion method. Peptides Phe-Leu, Pro-Pro, Pro-Leu-Gly-Gly, and the inhibitor phenyl chloromethyl ketone (Bachem) were dissolved at 10–100 mM final concentration in reservoir buffers. Crystals
processes with MOSFLM (30), SCALA, TRUNCATE, and CAD of the CCP4 program suite. Final model building was performed with the interactive graphics program MAIN (26, 27) and refinement in CNS version 1.1 (31). The quality of the model was cross-validated by using 5% of independent reflections in test data sets. Ligand coordinates for PCK were obtained on the HICU server (x-ray.bmc.uu.se/hicup) and refined together with standard topology and parameter files (32) in CNS. The refinement of the investigated mutants and their complexes resulted in models with R-factors less than 25% and R_free = 30% for data at resolutions better than 2.8 Å (TABLE ONE). The stereochemistry of the models was inspected with PROCHECK (33); density surfaces and figures were created with PyMOL (35). Coordinates and structure factors were deposited in the Protein Data Bank with the

| F1-Mutant Peptide/Inhibitor | G37A | S105A FL | S105A PCK | S105A PLGG | Y205F PCK | E213Q AF |
|-----------------------------|------|---------|----------|-----------|-----------|---------|
| Resolution (in Å)           |      |         |          |           |           |         |
| Space group P2_1 2_1 2_1, cell constants (Å) | 55.28 | 57.35 | 52.27 | 54.17 | 57.33 | 56.97 |
| Space group P2_1 2_1 2_1, cell constants (Å) | 57.73 | 60.88 | 60.52 | 60.83 | 61.85 | 61.92 |
| Space group P2_1 2_1 2_1, cell constants (Å) | 82.70 | 80.50 | 79.79 | 80.24 | 80.77 | 80.42 |
| Unique reflections (highest resolution shell) | 11843 (1868) | 18335 (2876) | 13669 (1730) | 4316 (695) | 23723 (2733) | 7637 (800) |
| Completeness (%)             |      |         |          |           |           |         |
| I/σ (°)                      |      |         |          |           |           |         |
| R_free α                     |      |         |          |           |           |         |
| Root mean square deviation bonds (Å) | 0.008 | 0.009 | 0.007 | 0.011 | 0.008 | 0.008 |
| Root mean square deviation angles (°) | 1.045 | 1.113 | 1.080 | 1.298 | 1.112 | 1.025 |
| Protein atoms                |      |         |          |           |           |         |
| Water molecules              |      |         |          |           |           |         |
| B-average (Å²)               |      |         |          |           |           |         |
| Protein                      | 19.4 | 36.0    | 34.4     | 50.0      | 27.2      | 40.9    |
| Ligands                      | 42.1 | 48.8    | 38.3     | 21.6      | 63.3      |
| Water                        | 37.1 | 37.1    | 27.2     | 32.4      |
| Protein Data Bank entry      | 1XQV | 1XQW    | 1XQX     | 1XQY      | 1XRL      | 1XRM    |
| F1-Mutant Peptide/Inhibitor | E213Q FA | E213Q FL | E213Q PLGG | E245Q FL | E245Q PP |
| Resolution (in Å)            |      |         |          |           |           |         |
| Space group P2_1 2_1 2_1, cell constants (Å) | 57.41 | 51.35 | 50.42 | 57.56 | 52.80 |
| Space group P2_1 2_1 2_1, cell constants (Å) | 61.54 | 60.42 | 60.52 | 61.91 |
| Space group P2_1 2_1 2_1, cell constants (Å) | 80.69 | 79.93 | 79.91 | 80.03 |
| Unique reflections (highest resolution shell) | 7239 (1184) | 21603 (3395) | 9462 (1523) | 6889 (720) | 9533 (1445) |
| Completeness (%)             |      |         |          |           |           |         |
| I/σ (°)                      |      |         |          |           |           |         |
| R_free α                     |      |         |          |           |           |         |
| Root mean square deviation bonds (Å) | 0.008 | 0.006 | 0.009 | 0.008 | 0.008 |
| Root mean square deviation angles (°) | 1.004 | 1.036 | 1.175 | 1.094 |
| Protein atoms                |      |         |          |           |           |         |
| Water molecules              |      |         |          |           |           |         |
| B-average (Å²)               |      |         |          |           |           |         |
| Protein                      | 47.4 | 28.2    | 28.5     | 43.6      | 33.1      |
| Ligands                      | 54.1 | 31.0    | 19.9     | 52.6      |
| Water                        | 34.7 | 35.1    | 34.9     | 33.8      |

were soaked prior to measurements for 1–10 min with 15% glycerol as cryoprotectant and flash frozen in a stream of nitrogen gas at 100 K.

**Data Collection, Model Refinement, and Graphic Representation**—Data sets were measured with synchrotron radiation for wavelengths 0.9795 and 1.0500 Å at the BW6 beamline at Deutsches Elektronen Synchrotron in Hamburg, using a MAR CCD detector. The images were processed with DENOVO and SCALEPACK (29); data reduction and further optimization were performed with TRUNCATE, CAD, and SCALCAT of the CCP4i software. Preliminary refinements were attempted with REFMAC5 (see www.ccp4.ac.uk/) against the F1 model (Protein Data Bank code 1MTZ) without water molecules, and a first inspection of the calculated electron density maps was done with the program O. Additionally, data were collected with Cu-Kα radiation generated by a rotating anode generator on a MAR 345 image plate system and processed with MOSFLM (30), SCALA, TRUNCATE, and CAD of the CCP4 program suite. Final model building was performed with the interactive graphics program MAIN (26, 27) and refinement in CNS version 1.1 (31). The quality of the model was cross-validated by using 5% of independent reflections in test data sets. Ligand coordinates for PCK were obtained on the HICU server (x-ray.bmc.uu.se/hicup) and refined together with standard topology and parameter files (32) in CNS. The refinement of the investigated mutants and their complexes resulted in models with R-factors less than 25% and R_free = 30% for data at resolutions better than 2.8 Å (TABLE ONE). The stereochemistry of the models was inspected with PROCHECK (33); density indices and correlations were analyzed with SFHECK (34). Molecular surfaces and figures were created with PyMOL (35). Coordinates and structure factors were deposited in the Protein Data Bank with the
accession codes 1XQV, 1XQW, 1XQX, 1XQY, 1XRL, 1XRM, 1XRO, 1XRP, 1XRQ, and 1XRR.

RESULTS

Functional Characteristics of F1 Mutants—For activity assays mutants and wild type (WT) F1 were incubated with the fluorogenic AMC substrates Pro-AMC, Phe-AMC, and Ala-Ala-Phe-AMC (AAF-AMC) at 60 °C. Remarkably, the exchange of the P1 binding residue Gly-37 to alanine generated a mutant that was completely inactive against these substrates (Fig. 2, A and B). In contrast, the removal of the O-/H9253/nucleophile of the catalytic serine in the S105A mutant resulted in a residual activity of 9% against AAF-AMC. S105A cleaved Phe-AMC with 6% and Pro-AMC with 3% of WT-F1 activity (Fig. 2, A and B).

Mutations of the hydrogen-bonded network that binds the N terminus of substrates affected the activity in a differentiated pattern. The loss of the tyrosine OH group in the mutant Y205F reduced the activity against Pro-AMC to 40%, whereas the cleavage of Phe-AMC and AAF-AMC had the same rate as in WT-F1 (Fig. 2, A and B). The importance of the negatively charged glutamate Glu-213 in cleavage of proline in P1 position is emphasized by the strong reduction of Pro-AMC cleavage to 2% and of Phe-AMC cleavage to 8% in the glutamine mutant E213Q (Fig. 2B). In E245Q, the activity against test substrates followed the same order, but the effect was smaller than in E213Q. Interestingly, the cleavage rate of AAF-AMC in both mutants remained relatively high with 43 and 58% of WT-F1 activity (Fig. 2B). The least active F1 mutants were chosen to compare their influence on the functional interaction of F1 and the tricorn protease.

Proteolytic Assay of Tricorn Combined with WT-F1 and Mutants S105A and G37A—Based on previous studies of tricorn-interacting factors (13), a protein degradation assay with oxidized insulin-B as test substrate was performed with tricorn and the above mentioned mutants of F1. Subsequently, cleavage products were analyzed with reverse phase high performance liquid chromatography. The product patterns of tricorn protease alone and combined with WT-F1 are very similar when compared 20 and 40 min after the reaction was started. The peak intensity of the tricorn-F1 degradation products appears to be slightly smaller than for tricorn alone, most likely because F1 cleaved a certain amount of N-terminal residues from the tricorn products (Fig. 2C, a and b). In contrast to the wild type proteins, which degrade insulin-B completely within 40 min, neither of the mutant mixtures with tricorn is capable of significantly breaking down the substrate. Furthermore, the combined activity of tricorn and S105A is obviously retarded, and even after 40 min the product peaks reach only ~30% of the tricorn/WT-F1 reaction (Fig. 2C, c). In the case of the G37A-F1 mutant together with
X-ray Snapshots of Peptide Processing in Mutants of F1

tricorn only traces of products could be detected (Fig. 2C, d). As estimated by the area under the curve of the cleavage product peaks, the tricorn/S105A-F1 activity is reduced to ~10%, whereas the tricorn/G37A-F1 activity is ~1% of the tricorn/WT-F1 reaction (Fig. 2C). Apparently, the total reaction depends mostly on the activity of F1 and its mutants. These results hint of a physical interaction between tricorn and F1, which is examined below.

Solution Studies on F1 Tricorn Interaction—We chose small angle x-ray scattering to investigate the size and shape of solute tricorn, F1, and possible complex particles. Tricorn itself appeared to be a multifaceted system in solution, which complicated analysis of the mixtures with F1. The processed scattering intensity from tricorn (Fig. 3A) yielded the radius of gyration \( R_g \) 59 ± 2 Å and a maximum size of \( D_{max} \) 190 ± 10 Å, which was in good agreement with values from the crystal structure of hexameric tricorn \( (R_g = 59.5 \text{ Å} \text{ and } D_{max} = 190 \text{ Å}) \). However, the scattering pattern computed from the tricorn hexamer failed to fit the experimental data with a discrepancy \( \chi = 7.15 \) (see Fig. 3A, curve 2). Moreover, the estimated average molecular mass of the particles \( (380 \pm 40 \text{ kDa}) \) was significantly lower than the theoretical value of 720 kDa, suggesting partial dissociation of the tricorn hexamers in solution. To evaluate this hypothesis, the experimental data were fitted by a linear combination of the scattering from tricorn monomers and hexamers. A good fit to the experimental data with a discrepancy of \( \chi = 1.25 \) (Fig. 3A, curve 3) was obtained from a mixture that was composed of 35% monomers and 65% hexamers. The scattering patterns from tricorn displayed a pronounced dependence on the solute concentration. In contrast to samples at lower concentrations, such as \( c = 2.0 \text{ and } 4.0 \text{ mg/ml} \), the higher concentration samples exhibit distinct oscillations as depicted for the low angle portions of the scattering patterns (Fig. 3A, insert, curves 1 and 2). These wiggling curves may be explained by the presence of large isometric particles as illustrated by curve 3 computed for a hollow spherical particle with outer and inner diameters of 500 and 320 Å, respectively. Interestingly, the size of this particle corresponds well to the overall size of the icosahedral super molecules observed earlier by cryoelectron microscopy (36). Thus, our results indicate that monomeric, hexameric, and icosasymmetric tricorn are in a concentration-dependent equilibrium in solution.

Mixtures of tricorn protease with wild type F1, G37A, and S105A mutants were measured with and without addition of peptides and inhibitors and at different solute concentrations. In most cases, the radius of gyration of the solute increased to 62 or even 65 Å. As the computed \( R_g \) value of the tentative tricorn–F1 complex \( (67.8 \text{ Å}) \) is larger than that of tricorn alone \( (59 \text{ Å}) \), such an increase pointed to the complex formation between tricorn and F1. To quantitatively assess this process, experimental data were fitted assuming mixtures of monomeric and hexameric tricorn, monomeric F1, and, optionally, hexamer complexes with three F1 monomers. In solutions with low tricorn concentrations \( (2.0 \text{ mg/ml}) \), accounting for the complex formation with wild type F1 allowed noticeable improvement in the fit to the experimental data (Fig. 3B, curve 1, \( \chi \) reduction from 1.69 to 1.49) and with the G37A mutant \( (curve 2, \chi \) reduction from 1.89 to 1.67). In both cases, the improvement of the fit was achieved by transforming ~70% of tricorn hexamers into complexes with F1. Surprisingly, an even stronger effect was observed for the mixture with the G37A mutant after addition of the

FIGURE 3. A, experimental small angle x-ray scattering patterns from tricorn in plots of relative scattering intensity \( \lg I \) (common logarithm) against the momentum transfer \( s (\text{Å}^{-1}) \); 1) experimental data extrapolated to zero solute concentration, 2) scattering computed for hexameric tricorn, and 3) best fit for a mixture of tricorn monomers and hexamers (ratio 35:65) computed by OLGOMER. The inset displays the low angle scattering from 2.0 and 4.0 mg/ml solutions of tricorn (curves 1 and 2, respectively) and the computed scattering \( (curve 3) \) from a spherical capsid with diameter \( D = 500 \text{ Å} \) and thickness of 70 Å displaying oscillations at the same positions as those observed in curve 2 B, experimental scattering curves \( (dots with error bars) \) of tricorn mixtures with F1 at 2.0 mg/ml and best fits from mixtures containing tricorn–F1 complexes (ratio 2:1, solid lines). Curve 1 corresponds to the wild type protein (70% of hexamers complexed), curve 2 to the G37A mutant, and curve 3 to G37A with phenyl chloromethyl ketone added (100% of hexamers in complex for both G37A curves).
F1 inhibitor PCK, where all hexamers were transformed into complexes (Fig. 3B, curve 3) to yield \( \chi \) reduction from 3.83 to 3.19, whereas the volume fraction of tricorn monomers remained between 30 and 40%.

A second approach to confirm F1-tricorn interaction in solution with the Biacore method largely failed. According to these measurements the binding constants of tricorn and F1 appear to be higher than 100 \( \mu \text{M} \). Thus, only transient complexes between tricorn and F1 may account for the functional interaction behavior of the two peptidases that apparently requires the presence of tricorn products, which are suitable substrates of F1. In the following sections details of the F1 peptidase activity are revealed by a series of x-ray structures that were solved for mutants mostly in complex with ligands.
Crystallographic Snapshots of Peptides Entering the Cavities of F1—The structural basis of the F1 peptidase mechanism, in particular, was investigated by soaking peptide substrates or inhibitor with hydrophobic side chains into crystals of F1 mutants. S105A crystals soaked with the dipeptide Phe-Leu (FL) diffracted to 2.0 Å resolution and displayed strong additional electron densities in the active site region and at the supposed entrance tunnel E1 to the cavity system. In both positions the uncleaved dipeptide FL was modeled. The FL molecule that is located in a groove near E1 points with its N terminus directly toward the active site (Fig. 4A). Apparently, the hydrogen-bonded network of Tyr-205, Glu-213, and Glu-245 attracts the positively charged amino group of the dipeptide. Remarkably, the C terminus of the dipeptide is oriented toward the N terminus of a symmetry-related F1 molecule, whose first three residues are too disordered to be seen in the electron density.

Also located in the E1 region was the tetrapeptide Phe-Leu-Gly-Gly (PLGG) as observed in the complex structure of the mutant E213Q. The N-terminal proline appears to be close to the hydrogen-bonded network, which consists of Tyr-205, Gln-213, and Glu-245. As in S105A with FL, PLGG is bound in the groove at E1 nearly connecting with its C terminus to the first defined N-terminal residue of the symmetry-related F1 molecule (Fig. 4B).

Cleave of the Scissile Bond—An intermediate of the catalytic cleavage step is seen in S105A, where we observed continuous density of the intact dipeptide FL occupying the S1 site. Apparently because of the missing O-γ nucleophile in Ala-105, the peptide bond of FL was not attacked. The aromatic side chain of phenylalanine is bound in the hydrophobic S1 site, whereas its carbonyl-O is fixed by the oxygen hole, which consists of the backbone amides of Gly-37 and Tyr-106 (Fig. 4C). The leucine side chain of FL is located in a second predominately hydrophobic cavity of F1 that had been proposed as the S1’ specificity site (15). The carboxylate group of the leucine residue is hydrogen bonded by the OH group of Tyr-44 and N-ε of His-271. The amino group of the substrate is hydrogen bonded to the carbonyl-O of Gly-37, but not by the carboxylate of Gly-213. Most interestingly, the latter shows a broken density in the side chain, indicating high flexibility. It should be mentioned that the side chain of Glu-213 displays a similarly broken density in the F1 complex with MES but is well defined in WT-F1 and the PCK complexes (15) (Fig. 4D). Notably, the density of the dipeptide around the scissile peptide bond is blurred and the bond of the model is elongated relative to standard values (32). The flexible Glu-213 side chain in the S105A mutant presumably plays the role of the acidic residue in acid protease catalysis and could explain the residual activity of the S105A mutant (Fig. 2A). Moreover, the continuous electron density from the carboxylate of Glu-213 to the scissile bond may be interpreted as water molecule, which probably is the required nucleophile for hydrolysis (Fig. 4C). A better defined water molecule is bound by the amino group of the nearby P1 residue and may represent the location of this catalytic water just before the hydrolytic step.

Release of the P1 Residue—An analogue of the next catalytic step, after cleavage of the scissile bond, can be seen in an inhibitor complex of the mutant Y205F. PCK is covalently linked to the O-γ of Ser-105 and exhibits more structural details than the corresponding WT-F1 complex, especially for the inhibitor side chain (15). The chlorine atom is still present in the hydrophobic S1’ site, and the phenyl ring of PCK is well defined in the S1 site (Fig. 4D). The amino group of PCK is coordinated by the Gly-213 side chain, the Gly-37 carbonyl, and a water molecule at a 2.9 Å distance, which also forms a hydrogen bond to the carboxamide group of Asn-209 at 2.9 Å.

The hydrolyzed acyl intermediate of the P1 residue is represented by a 2.0 Å resolution structure of S105A soaked with PCK. Electron density of the non-covalently bound inhibitor was observed in the S1 site with a low occupancy. The binding mode for the amino group of PCK by Glu-213 and Gly-37 is similar to the previously described complexes, as well as the binding of the carbonyl O of the inhibitor into the oxyanion hole. The major difference is the missing covalent bond of Ser-105 O-γ to the tetrahedral sp3-carbon of PCK. There is also no covalent linkage of PCK to O-γ of Ser-104, which rules out the possibility that Ser-104 could play a role as a secondary nucleophile in catalysis.

Comparative frozen intermediates of peptide cleavage are represented by crystal structures of the mutants S105A, E213Q, and E245Q, which were soaked with peptides that contained proline in the P1 position, namely Pro-Leu-Gly-Gly and Pro-Pro. In each case, the peptides were cleaved and proline was located in the S1 cavity with its carboxylate bound by the oxyanion hole and, additionally by the O-γ of Ser-105 in the mutants E213Q and E245Q. Moreover, no P1 cleavage product was found in the S1’, and PLGG was located at the E1 entrance of S105A and E213Q. Intriguingly, the crystalline mutant E245Q was able to cleave the critical bond of the soaked dipeptide Pro-Pro, which requires at least two enzymes in eukaryotes (37). The crystal structure analysis revealed that one proline occupies the S1 site of E245Q, whereas the second proline is located at the E1 entrance with its imino group pointing toward the catalytic residues. Seemingly, the cleaved, zwitterionic P1’ residue cannot remain in the hydrophobic S1’ site and leaves the active site cavities through the E1 opening when S1 is occupied (Fig. 5A).

P1’ Residue Translocation—To elucidate the role of Glu-213 and Glu-245 in catalysis, their glutamine mutants were investigated. Crystals of E213Q and E245Q were soaked with FL, but electron density was only observed in S1, whereas S1’ was largely free. Obviously, the dipeptide was cleaved in both crystals and a single amino acid occupied the hydrophobic S1 sites. The amino group is hydrogen bonded by Glu-213 or Glu-213, and its carboxylate is bound by the oxyanion hole and the O-γ of Ser-105. Surprisingly, the electron density was better fitted with leucine than with phenylalanine (Fig. 5B), indicating that the P1 product had left the S1 site and the P1’ residue moved into the S1 site. There seems to be no preference for leucine over phenylalanine in F1, because the single amino acids do not inhibit the catalytic activity significantly. However, E213Q crystals soaked with Ala-Phe and Phe-Ala demonstrate that the S1 site is predominantly occupied by the former P1’ residue of each dipeptide. In comparison to phenylalanine the alanine is located more centrally in the hydrophobic S1 cave (Fig. 5C). The strongest polar interaction is seen for the amino group of alanine that is hydrogen bonded to the carbonyl-O of Gly-37 and the Asn-209 side chain, whereas the carbonyl group of phenylalanine is bound to the oxyanion hole. Analysis of density indices and correlation coefficients for individual amino acid residues according to the program SFCHECK (34) confirmed in each of the aforementioned examples that a larger fraction of the P1’ amino acid had displaced the P1 product from the S1 site.

G37A and an Alternative Product Exit—The crystal structure of the inactive mutant G37A displays intriguing features that may be involved in product egress, the last step of catalysis in F1. However, the exchange of Gly-37 to alanine caused a major rearrangement of the turn from residues 36 to 40, which moved into the S1’ site and obstruct the access channel E1 (Fig. 5D). The reason for this switch is the energetically unfavorable backbone conformation of Ala-37 in the loop conformation of WT-F1, which is only favorable for glycine. The conformation of Ala-37 prevents peptide binding in the active site in a drastic way. Its amide group that constitutes one half of the oxyanion pocket is shielded by the methyl group of Ala-37 and the carbonyl-O is hydrogen bonded to the O-γ of Ser-105, thereby replacing a water molecule that was observed in WT-F1 and may be the primary catalytic water (Fig. 5D). Concomitant to the loop switch, the aromatic side chain of Tyr-106 moves ~120° into the S1 specificity site and, therefore, binding of P1- side chains is inhibited.
FIGURE 5. Cleavage products of peptides in F1 mutants and inactive G37A in stereo. Electron density is displayed only for relevant residues. A, single proline molecules cleaved from Pro-Pro by mutant E245Q occupy the S1 site and the region of the E1 entrance. The $2F_o - F_c$ electron density map is contoured at 1 σ. B, leucine, the former P1 residue from the soaked peptide FL, is found in the S1 site of E213Q. The amino acid is depicted in green. The corresponding $2F_o - F_c$ map is contoured at 0.8σ. C, the crystal soak with the dipeptide FA in E213Q resulted in a single alanine that is bound in the S1 site mainly by Asn-209 and Gly-37 in contrast to leucine (panel B) or proline (panel A). The $2F_o - F_c$ map is contoured at 1σ. D, stereoview of the inactive mutant G37A rotated anticlockwise by 60° around the y axis with respect to the previous images. The carbonyl of Ala-37 forms a hydrogen bond to the O- of Ser-105, and the loop from Gly-36 to Met-40 has moved into the S1 cave blocking access for peptides through the E1 tunnel. Ala-37 and Tyr-106 are displayed in magenta. The $2F_o - F_c$ map (contour level 1σ) covers Tyr-106, which partially occupies the S1 site, and the loop from Gly-36 to Met-40. The backbone of WT-F1 is superimposed as white sticks.
Additionally, a new small tunnel from the S1 pocket to the protein surface opens by the rotation of the Trp-189 side chain. The diameter of this putative exit channel (E3) would allow single amino acids to leave the S1 cavity.

**DISCUSSION**

Active F1 in combination with tricorn allows complete breakdown of B-insulin to small peptides. Less active F1 mutants reduce the B-insulin degradation approximately proportional to their activity, ranging from \(-10\%\) for S105A to nearly \(0\%\) for G37A (Fig. 2C). Gel filtration and Biacore experiments showed only weak interaction of the wild type proteins (15). Also, the Biacore studies with either immobilized F1 mutants or tricorn and in the presence of substrate speak for low affinity interactions. Binding constants are \(100 \mu \text{M}\) or higher; therefore, one has to assume that the complex formation is transient. This phenomenon is well known from prokaryotic protein degradation, for example in the so-called caseinolytic protease complex (ClpAP), which is only formed when substrate is bound. The unfoldase ClpA transfers polypeptide substrates to the internal reaction chamber of the ClpP protease for cleavage, and when proteolysis is completed, the complex falls apart (38). However, the block of tricorn activity by inactive G37A-F1 argues for a permanent complex that is indeed substrate induced. The results of the described small angle scattering experiments confirm the existence of permanent F1-tricorn complexes in solution, although they may be stabilized by mutations and inhibitor. Moreover, the strongest evidence comes from samples that contained the inactive F1 mutant G37A in complex with the inhibitor PCK. Systematic analysis of the data did not reveal significant dependence of the complex formation on mutations of F1 and on the addition of peptides.

The phenomenon of transient complexes of tricorn and F1 appears to be associated with the directionality of the substrate/product flow in both peptidases (39). Mutational and inhibitor studies on tricorn revealed how substrate access through the \(\beta_7\)-propeller and product egress through the \(\beta_6\)-propeller are governed by size and shape of the cavities (40, 41). A dominant feature of tricorn is the directionality of oligopeptide flow to and from the catalytic Ser-965, whereby the proximal arginines Arg-131 and Arg-132 that bind the C terminus of peptide substrates gate the product through the \(\beta_6\)-propeller. Such a unidirectional substrate product flow fulfills the definition of processivity.

Because of the complementary shape and charge of the \(\beta_6\)-propeller and the cap domain of F1, the product outlet of tricorn is ideally located for a continuous substrate flow into the E1 entrance (15) (Fig. 6), where Glu-213 and Glu-245 attract the positively charged N terminus of peptides in order to orient them properly in a groove close to the mouth of the tunnel (Fig. 4, A and B). Even the negative charge of a single glutamate is sufficient to attract and bind peptides in the correct orientation, seen for E213Q and E245Q. Because structure and function of Y205F are very similar to WT-F1 one has to conclude that the role of Tyr-205 is simply to stabilize the hydrogen-bonded network and optimize the arrangement of the two glutamates.

The less active S105A mutant that was soaked with the dipeptide Phe-Leu elucidates the importance of Ser-105 from the catalytic triad as the crucial residue for peptide cleavage. In S105A the C terminus of the P1’ leucine from FL was bound by N-\(\epsilon\) of His-271 and the OH group of Tyr-44. Thus, the S1’ specificity site that was predicted from the F1-PCK structure (15) could be confirmed in detail. The residual activity in S105A, which was below \(10\%\) of WT-F1, cannot be explained by Ser-104 as secondary nucleophile because it did not bind the PCK inhibitor. Each of the WT-F1 and Y205F-PCK complexes displayed a well defined and rigid side chain in Glu-213, whereas this residue was more flexible in the other investigated structures. These findings support the idea that a dynamic Glu-213 is responsible for the residual peptidase activity (Fig. 4, C and D). Most likely, Glu-245 and Glu-213 activate a water molecule as nucleophile for hydrolysis of the scissile bond, analogous to acid proteases, which utilize two aspartates (42). Nevertheless, this water molecule may have only a secondary function, whereas another water molecule that was observed in WT-F1 bound to the oxyanion hole is more suitable for hydrolysis of the peptide substrate in the fully active peptidase (15).

It seems likely that Glu-213 also plays a key role in the translocation of the cleaved P1’ residue into the S1 pocket. In the mutants E213Q and E245Q the observed single amino acid did not leave the S1 site, probably because the electrostatic repulsion of its carboxylate by one negative charge was insufficient. Glu-213 has the capacity to push the charged P1 product out of the S1 site and then pull the amino group of the P1’ product into the S1 site. Concomitantly, Asn-209 could compensate the charges of zwitterionic amino acids as has been observed in E213Q with a single alanine in S1. Seemingly, proline as P1 residue does not leave the
X-ray Snapshots of Peptide Processing in Mutants of F1

S1 site, albeit in the case of Pro-Pro residues P1 and P1’ are indistinguishable and could occupy randomly the S1 site and the position at the entrance E1. In the examples of PLGG in the E213Q and S105A mutants the subsequent entering and cleavage of a tripeptide may be unfavorable because of their strongly reduced activity. However, the possibility that small oligopeptides are pulled into the active site of F1 and processively cleaved without release of intermediates cannot be ruled out. The rather high turnover of the tetrapeptide analogue AAF-AMC by all investigated mutants except G37A, in contrast to the dipeptide analogue turnover, indicates processive catalysis.

The important role of the substrate binding Gly-37 is emphasized by its high conservation in related prolyl peptidases where it is a part of the oxyanion pocket and contributes to the N-terminal binding of the substrate. The proposed exit E3 may be an alternative to E2, which would enhance the turnover rate of the enzyme. The analogous occupation of the S1 site by a tyrosine side chain, like Tyr-106, has been reported for an aspartic protease (43). It is conceivable that the combined movement of Tyr-106 and the loop switch from residues 36 to 40 serve as a kick-out mechanism for cleavage products, to prepare the enzyme for a new catalytic round.

Acknowledgments—We thank Gleb Bourenkov and Hans Bartunik for support at the Deutsches Elektronen Synchrotron (DESY) in Hamburg and Prajna Paramita Pal for critical reading of the manuscript.

REFERENCES

1. Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995) Science 268, 533–539
2. Bochtler, M., Ditzel, L., Groll, M., Hartmann, C., and Huber, R. (1999) Annu. Rev. Biophys. Biomol. Struct. 28, 295–317
3. Groll, M., Bochtler, M., Brandstetter, H., Clausen, T., and Huber, R. (2005) Chembiochem 6, 222–256
4. Brandstetter, H., Kim, J. S., Groll, M., and Huber, R. (2001) Nature 414, 466–470
5. Tamura, T., Tamura, N., Cséka, Z., Hegerl, R., Lottspeich, F., and Baumeister, W. (1996) Science 274, 1385–1389
6. Tamura, T., Tamura, N., Lottspeich, F., and Baumeister, W. (1996) FEBS Lett. 398, 101–105
7. Medrano, F. J., Alonso, J., Garcia, J. L., Romero, A., Bode, W., and Gomis-Ruth, F. X. (1999) EMBO J. 18, 1–9
8. Yoshimoto, T., Kabashima, T., Uchikawa, K., Inoue, T., Tanaka, N., Nakamura, K. T., Tsuru, M., and Ito, K. (1999) J. Biochem. 126, 559–565
9. Fulop, V., Bocskai, Z., and Polgar, L. (1998) Cell 94, 161–170
10. Engel, M., Hoffmann, T., Wagner, L., Wermann, M., Heiser, U., Kiefersauer, R., Huber, B., Bode, W., Demuth, H. U., and Brandstetter, H. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 5063–5068
11. Stoltze, L., Schirle, M., Schwarz, G., Schrotter, C., Thompson, M. W., Hersh, L. B., Kallbacher, H., Stevanovic, S., Ramnensee, H. G., and Schild, H. (2000) Nat. Immunol. 1, 413–418
12. Tamura, N., Lottspeich, F., Baumeister, W., and Tamura, T. (1998) Cell 95, 637–648
13. Yoshimoto, T., Kabashima, T., Uchikawa, K., Inoue, T., Tanaka, N., Nakamura, K. T., Tezuka, Y., and Ito, K. (1999) J. Biochem. 126, 559–565
14. Ko, J., Jang, S.-J., Kusniewski, E., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1996) Acta Crystallogr. Sect. D Biol. Crystallogr. 52, 905–921
15. Engh, R. A., and Huber, R. (1991) Acta Crystallogr. Sect. A 47, 392–400
16. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
17. Vagin, A. A., Richelle, J., and Wodak, S. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 191–205
18. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA
19. Walz, J., Tamura, T., Tamura, N., Grimm, R., Baumeister, W., and Koster, A. J. (1997) Mol. Cell. Biol. 17, 59–65
20. Vanhool, G., Goosens, F., De Meester, L., Hendriks, D., and Scharpe, S. (1995) FASEB J. 9, 736–744
21. Ishikawa, T., Beuron, F., Kessel, M., Winkler, S., Maurizi, M. R., and Steven, A. C. (1994) EMBO J. 13, 4431–4458
22. Brandstetter, H., Kim, J. S., Groll, M., Goettig, P., and Huber, R. (2002) Biochem. Biophys. Res. Commun. 295, 1020–1026
X-ray Snapshots of Peptide Processing in Mutants of Tricorn-interacting Factor F1 from *Thermoplasma acidophilum*

Peter Goettig, Hans Brandstetter, Michael Groll, Walter Göhring, Peter V. Konarev, Dmitri I. Svergun, Robert Huber and Jeong-Sun Kim

*J. Biol. Chem.* 2005, 280:33387-33396.
doi: 10.1074/jbc.M505030200 originally published online July 1, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M505030200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 38 references, 7 of which can be accessed free at http://www.jbc.org/content/280/39/33387.full.html#ref-list-1