Electronic Supplementary Material (ESI) for Chemical Science


SUPPLEMENTARY INFORMATION

# Structural insights into the potential of 4-fluoroproline to modulate biophysical properties of proteins

Bastian Holzberger, Samra Obeid, Wolfram Welte, Kay Diederichs, and Andreas Marx*

Departments of Chemistry and Biology, Konstanz Research School Chemical Biology
University of Konstanz, Universitätsstr. 10, 78457 Konstanz, Germany

Andreas.Marx@uni-konstanz.de

## 1. General

MinElute Reaction Cleanup Kit was from Qiagen; restriction enzymes were purchased from Fermentas; LB medium (LB broth (Lennox)), carbenicillin, and lysozyme were from Roth; polyethylenimine (PEI) was from Sigma-Aldrich; Q sepharose, superdex 75, and ÄKTApurifier are from GE Healthcare; Viva Spin columns were from Sartorius Stedim Biotech; CD quartz cuvettes are from Hellma; CD spectrometer J-815, temperature controller MPTC-4090S and Spectra Manager v2 are from Jasco; oligonucleotides were purchased from Metabion or Purimex (HPLC purified); and ddCTP was from Jena Bioscience.

## 2. Mass spectrometry

For LC-ESI-MS, the protein sample was desalted (Viva Spin) to 5 mM Tris·HCl pH 9.2 and 0.25 mM $MgCl_2$. The samples (20 µL, 50-100 µM) were applied to a Vertex-Column (250×8 mm Nucleosil 120 C4) and eluted with a gradient of 0.05% TFA in $H_2O$ and 0.05% TFA in acetonitrile. ESI-MS was performed on an ESI-MS microTOF II (Bruker). Data were analyzed with DataAnalysis from Bruker.

## 3. Protein expression and purification

The *E. coli* codon optimized wild-type KlenTaq gene (purchased from GeneArt; amino acids 293-832 of *Taq* gene cloned into pET21b using the restriction sites NdeI and NotI) was amplified together with the N-terminal RBS of the pET vector by PCR. Products were purified by using MinElute Reaction Cleanup Kit and cloned into pGDR11[1] using the restriction sites EcoRI and HindIII. This modified pGDR11 vector harboring the *E. coli* optimized wild-type gene was used to express wild-type and (4*R*)-FPro-KlenTaq in *E. coli* JM83 without any additional N-or C-terminal amino acids except for the N-terminal Met.

Wild-type KlenTaq: A 1 L culture was grown in LB medium (100 mg/L carbenicillin) to an $OD_{600}$ of 0.6. Protein synthesis was induced (1 mM IPTG) and carried out for 5 h.

(4*R*)-FPro-KlenTaq: An in total 4 L culture was grown in NMM medium (7.5 mM $(NH_4)_2SO_4$, 8.5 mM NaCl, 22 mM $KH_2PO_4$, 50 mM $K_2HPO_4$, 1 mM $MgSO_4$, 1 mg/L $CaCl_2$, 0.55 mg/L $FeSO_4$, 1 µg/L $CuCl_2$, 1 µg/L $MnCl_2$, 1 µg/L $ZnCl_2$, 1 µg/L $Na_2MoO_4$, 20 mM glucose, 10 mg/L thiamine, 10 mg/L biotin, 100 mg/L carbenicillin, and 50 mg/L of each amino acid except Pro) with 35 µM Pro until stationary phase was reached ($OD_{600}$ 0.4-0.6). The cells were additionally incubated for 3 h at 37 °C. Cells were harvested (3,800 rpm, 12 min, 4 °C) and resuspended in 4 L of NMM without Pro. (4*R*)-FPro was added to a final concentration of 1 mM. After incubation for 20 min at 37 °C, protein synthesis was induced (1 mM IPTG) and carried out overnight.

After harvesting (4,500×g, 40 min, 4 °C), cells were in both cases resuspended in lysis buffer (50 mM Tris·HCl pH 8.55, 10 mM $MgCl_2$, 16 mM $NH_4SO_4$, 0.1% Thesit, 0.1% Triton X, 1 mM PMSF; 20 mL per 1 L expression volume). Lysozyme was added to a final concentration of 0.7 mg/mL and cells were lysed for 1 h at r.t. followed by heat denaturation of host proteins (75 °C, 45 min) and centrifugation (4,500×g, 60 min, 4 °C). Supernatants were treated stepwise with 5% polyethylenimine (PEI) to remove bacterial DNA. After each PEI addition step the suspensions were shaken (30 min, 4 °C) and centrifuged (4,500×g, 45 min, 4 °C). These steps were repeated until almost no more precipitate was formed. Afterwards, lysates were filtered through syringe sterile filters to remove residual particles. Anion exchange chromatography was carried out using Q sepharose on an ÄKTApurifier. The protein was eluted applying 0-1 M NaCl in 20 mM Tris·HCl (pH 8.55), 1 mM EDTA, and 1 mM $\beta$-mercaptoethanol (4 °C, 1 mL/min). Protein fractions were controlled by SDS PAGE, pooled and concentrated using Viva Spin columns. Size exclusion chromatography was carried out using superdex 75 on an ÄKTApurifier (20 mM Tris·HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM $\beta$-mercaptoethanol, 4 °C; 1 mL/min). Fractions were controlled by SDS PAGE. Pure protein was pooled and stored at 4 °C until further use.

## 4. Figures

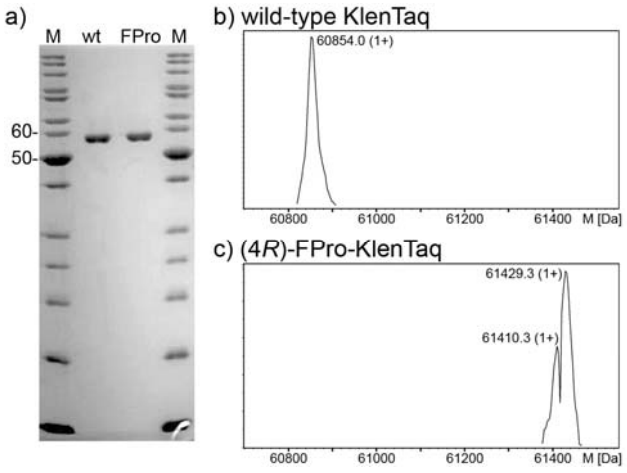


**Figure S1.** a) SDS PAGE gel of purified wild-type KlenTaq (wt) and (4*R*)-FPro-KlenTaq (FPro). M: Marker [kDa]. b, c) Deconvoluted ESI mass spectra profiles of b) wild-type KlenTaq (60854.0 Da; calcd (1+): 60855.1) and c) (4*R*)-FPro-KlenTaq (61429.3 and 61410.3 Da; calcd (1+): 61430.9 (100.0%) and 61412.9 Da (96.9%)). M : molecular mass [Da].

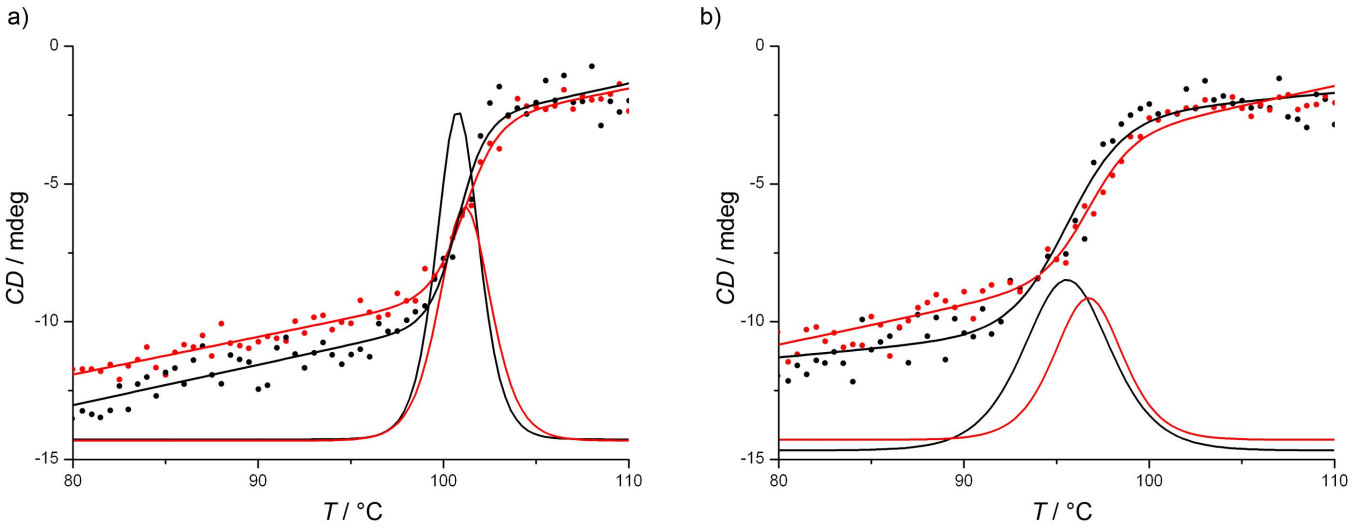


**Figure S2.** CD spectroscopy. Thermal denaturation of a) wild-type KlenTaq and b) (4*R*)-FPro-KlenTaq following the ellipticity at 208 nm (black) and 222 nm (red) in PBS buffer (pH 7.4). Melting temperatures were determined by sigmoidal fitting and calculating the derivatives of two separate experiments.

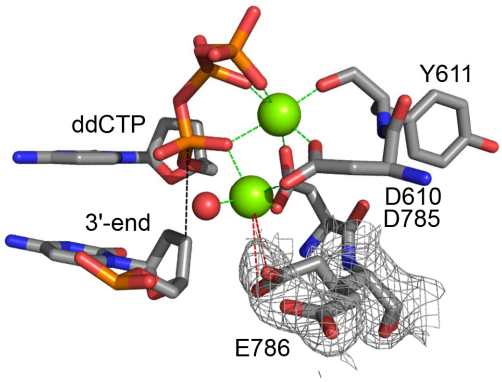


**Figure S3.** Active site of wild-type KlenTaq. Inner coordination spheres showing $Mg^{2+}$ ions (green). The 3'-end of the primer, the incoming triphosphate (ddCTP) and the respective amino acids (Y611, D610, D785, and E786) are shown in grey. A water molecule is shown as red sphere. The lower metal ion is probably coordinated by the second conformation of E786 (red dashes). 2mFo-DFc map at 1$\sigma$ is shown around E786.

## 5. Tables

**Table S1.** Data collection and refinement statistics.

| | **wild-type KlenTaq** | **(4*R*)-FPro-KlenTaq** |
|---|---|---|
| **Data collection** | | |
| Space group | $P3_121$ | $P3_121$ |
| Cell dim. | | |
| *a*, *b*, *c* [Å] | 107.8, 107.8, 89.6 | 109.6, 109.6, 91.0 |
| $\alpha, \beta, \gamma$ [°] | 90, 90, 120 | 90, 90, 120 |
| Resolution [Å] | 46.67 – 1.89 (2.01 – 1.89) | 47.47 – 2.44 (2.59 – 2.44) |
| $R_{meas}$ [%] | 14.0 (189.9) | 23.6 (172.7) |
| $I / \sigma_I$ | 14.52 (1.13) | 12.34 (1.17) |
| Completeness [%] | 99.6 (97.7) | 99.7 (98.5) |
| Redundancy | 10.6 (8.6) | 10.5 (8.7) |
| **Refinement** | | |
| Resolution [Å] | 46.67 – 1.89 (1.91 – 1.89) | 42.09 – 2.44 (2.49 – 2.44) |
| No. reflections | 92174 | 45237 |
| $R_{work}$ / $R_{free}$ | 21.4 / 25.2 (40.6 / 39.1) | 23.3 / 28.2 (37.9 / 35.7) |
| No. residues / average B-factors | | |
| Protein | 540 / 36.3 | 539 / 44.6 |
| Primer,Template | 12, 14 / 34.0, 39.2 | 12, 16 / 30.7, 43.4 |
| Triphosphate | ddCTP / 36.4 | ddCTP / 30.7 |
| Ion (active site) | 2 $Mg^{2+}$ / 32.8 | $Mg^{2+}$ + $Mn^{2+}$ / 36.8 |
| Water | 445 / 37.9 | 156 / 28.8 |
| R.m.s. deviations | | |
| Bond lengths [Å] | 0.005 | 0.008 |
| Bond angles [°] | 0.628 | 0.587 |
| Ramachandran statistics [%] | | |
| Most favored regions | 92.6 | 94.2 |
| Additionally allowed regions | 7.2 | 5.5 |
| Generously allowed regions | 0.0 | 0.0 |
| Disallowed regions | 0.2 | 0.2 |
| PDB ID | 4DLG | 4DLE |

**Table S2.** Torsion angles [°], peptide bond conformations, and prolyl ring puckering conformations of wild-type KlenTaq. Alternative conformations are depicted as **a** and **b**, respectively.

| Pro | $\omega$ (Xaa-Pro) | peptide [a] | $\Phi$ | $\Psi$ | $\chi^1$ | $\chi^2$ | $\chi^3$ | $\chi^4$ | $\chi^{Pro}$ [b] | pucker [c] |
|---|---|---|---|---|---|---|---|---|---|---|
| **298-a** | -178.7 | trans | -66.7 | 153.4 | 24.7 | -35.3 | 31.3 | -15.9 | 107.2 | endo |
| **298-b** | -179.1 | trans | -65.3 | 152.6 | -21.6 | 34.2 | -32.5 | 19.3 | -107.5 | exo |
| **300** | 1.1 | cis | -92.1 | 162.8 | 34.9 | -34.5 | 20.2 | 2.6 | 87.0 | endo |
| **301** | 175.2 | trans | -83.4 | 162.5 | 32.4 | -37.3 | 27.0 | -6.2 | 102.9 | endo |
| **302** | 174.6 | trans | -70.0 | 167.2 | 24.5 | -34.7 | 30.5 | -15.2 | 104.8 | endo |
| **316** | -177.2 | trans | -62.1 | -27.7 | 22.8 | -34.5 | 31.9 | -17.9 | 107.2 | endo |
| **336-a** | 177.3 | trans | -64.5 | -42.4 | 21.7 | -33.0 | 30.5 | -17.1 | 102.2 | endo |
| **336-b** | 176.8 | trans | -63.0 | -44.1 | -21.9 | 34.3 | -32.3 | 18.9 | -107.3 | exo |
| **338-a** | -174.1 | trans | -55.7 | -46.7 | 19.4 | -33.2 | 33.2 | -21.6 | 107.5 | endo |
| **338-b** | -174.5 | trans | -54.5 | -47.9 | -26.1 | 35.5 | -30.3 | 13.9 | -105.7 | exo |
| **368** | -178.4 | trans | -73.5 | 142.6 | 27.9 | -34.9 | 27.5 | -9.7 | 100.1 | endo |
| **369** | 177.1 | trans | -78.6 | 143.4 | 29.2 | -33.7 | 24.3 | -5.4 | 92.7 | endo |
| **373** | 179.3 | trans | -56.8 | -28.2 | -26.2 | 35.5 | -30.2 | 13.7 | -105.7 | exo |
| **382-a** | -178.9 | trans | -62.0 | -11.2 | -23.6 | 35.0 | -31.9 | 17.1 | -107.7 | exo |
| **382-b** | -178.4 | trans | -63.3 | -9.8 | 22.1 | -33.7 | 31.4 | -17.7 | 104.9 | endo |
| **387-a** | 177.8 | trans | -62.5 | -29.3 | 21.8 | -34.1 | 32.3 | -19.0 | 107.2 | endo |
| **387-b** | 177.4 | trans | -61.2 | -30.8 | -24.1 | 35.0 | -31.4 | 16.4 | -107.0 | exo |
| **436-a** | 177.6 | trans | -67.8 | -29.0 | -20.6 | 34.3 | -33.7 | 21.2 | -109.8 | exo |
| **436-b** | 177.8 | trans | -69.0 | -27.6 | 24.2 | -35.2 | 31.6 | -16.5 | 107.5 | endo |
| **481** | 179.2 | trans | -79.1 | 153.0 | 29.4 | -35.3 | 26.6 | -7.8 | 99.1 | endo |
| **501-a** | -179.9 | trans | -66.8 | 135.1 | 24.4 | -34.9 | 31.0 | -15.7 | 106.0 | endo |
| **501-b** | 179.9 | trans | -65.8 | 134.4 | -21.5 | 34.3 | -32.7 | 19.6 | -108.1 | exo |
| **527** | -178.7 | trans | -63.8 | -38.2 | -22.4 | 34.8 | -32.8 | 19.1 | -109.1 | exo |
| **548-a** | -175.9 | trans | -71.6 | -38.0 | 26.1 | -35.0 | 29.4 | -12.8 | 103.3 | endo |
| **548-b** | -176.3 | trans | -69.8 | -40.1 | -18.4 | 33.4 | -34.6 | 23.6 | -110.0 | exo |
| **550** | -179.8 | trans | -64.8 | -22.0 | -23.9 | 35.7 | -32.8 | 18.0 | -110.4 | exo |
| **555** | -173.9 | trans | -68.9 | -13.1 | 26.7 | -35.8 | 30.1 | -13.3 | 105.9 | endo |
| **579** | -4.8 | cis | -94.8 | 169.4 | 34.6 | -34.1 | 19.7 | 3.0 | 85.4 | endo |
| **585** | -176.0 | trans | -49.6 | 148.1 | -28.6 | 35.4 | -27.6 | 9.3 | -100.9 | exo |
| **589-a** | 178.6 | trans | -53.4 | -37.7 | -28.9 | 35.8 | -28.0 | 9.6 | -102.3 | exo |
| **589-b** | 179.0 | trans | -54.6 | -36.5 | 16.2 | -32.0 | 34.5 | -25.2 | 108.0 | endo |
| **650** | 177.9 | trans | -51.7 | 154.5 | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* |
| **656** | 177.8 | trans | -50.7 | -33.9 | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* |
| **685** | 177.6 | trans | -56.1 | 155.8 | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* |
| **701** | -178.4 | trans | -69.8 | -27.8 | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* |
| **731** | -176.6 | trans | -72.5 | -28.3 | 27.5 | -35.8 | 29.3 | -11.9 | 104.5 | endo |
| **752** | 178.4 | trans | -62.9 | -29.4 | -25.3 | 36.4 | -32.5 | 16.7 | -110.9 | exo |
| **770** | 179.1 | trans | -67.6 | -22.5 | 24.2 | -35.5 | 32.2 | -17.2 | 109.0 | endo |
| **792** | 174.2 | trans | -58.1 | 146.5 | -26.3 | 35.6 | -30.1 | 13.6 | -105.7 | exo |
| **812** | -174.0 | trans | -53.0 | 138.8 | -27.7 | 36.4 | -30.1 | 12.7 | -106.8 | exo |
| **816-a** | 179.4 | trans | -66.5 | 143.9 | 24.4 | -35.2 | 31.3 | -16.1 | 107.0 | endo |
| **816-b** | 179.1 | trans | -65.3 | 143.2 | -22.2 | 34.5 | -32.5 | 19.0 | -108.2 | exo |

*n.a.* not assigned (lack of sufficient electron density)

[a] peptide bond conformation of Xaa-Pro

[b] $\chi^{Pro} = \chi^1 + \chi^3 - \chi^2 - \chi^4$

[c] prolyl ring puckering based on $\chi^1$ and $\chi^{Pro}$

**Table S3.** Torsion angles [°], peptide bond conformations, and prolyl ring puckerings of (4*R*)-FPro-KlenTaq. Alternative conformations are depicted as **a** and **b**, respectively.

| (4*R*)-FPro | $\omega$ (Xaa-(4*R*)-FPro) | peptide [a] | $\Phi$ | $\Psi$ | $\chi^1$ | $\chi^2$ | $\chi^3$ | $\chi^4$ | $\chi^{Pro}$ [b] | pucker [c] |
|---|---|---|---|---|---|---|---|---|---|---|
| **298** | -179.0 | trans | -80.0 | 126.5 | -21.8 | 33.7 | -31.8 | 18.6 | -105.9 | exo |
| **300-a** | 4.3 | cis | -84.4 | 173.2 | -1.7 | -15.3 | 25.8 | -28.2 | 67.6 | $C^{\delta}$ out of plane |
| **300-b** | 5.0 | cis | -83.0 | 174.6 | -21.5 | 34.0 | -32.0 | 18.4 | -105.8 | exo |
| **301** | 173.2 | trans | -88.1 | 168.3 | 15.2 | -33.0 | 37.1 | -28.8 | 114.1 | endo |
| **302** | 178.6 | trans | -72.0 | 167.9 | -24.4 | 35.1 | -31.4 | 16.1 | -106.9 | exo |
| **316** | -177.4 | trans | -63.0 | -23.4 | -23.2 | 35.1 | -32.5 | 18.1 | -108.8 | exo |
| **336** | 178.3 | trans | -73.9 | -61.4 | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* |
| **338** | -173.7 | trans | -60.8 | -45.5 | -27.6 | 35.0 | -27.9 | 10.3 | -100.8 | exo |
| **368** | -174.9 | trans | -65.9 | 157.9 | -23.9 | 34.9 | -31.4 | 16.5 | -106.7 | exo |
| **369** | 176.7 | trans | -83.0 | 116.5 | -19.9 | 29.9 | -27.4 | 14.8 | -92.0 | exo |
| **373** | -179.5 | trans | -59.9 | -29.4 | -27.3 | 34.8 | -28.0 | 10.7 | -100.8 | exo |
| **382** | 179.4 | trans | -66.2 | -10.6 | -28.0 | 34.7 | -27.1 | 9.2 | -99.0 | exo |
| **387** | 177.5 | trans | -61.9 | -28.6 | -28.8 | 33.5 | -24.4 | 5.8 | -92.5 | exo |
| **436** | 172.9 | trans | -75.3 | -31.7 | -26.1 | 36.2 | -31.5 | 14.8 | -108.6 | exo |
| **481-a** | 179.1 | trans | -90.3 | 150.1 | 8.6 | -25.8 | 32.1 | -27.9 | 94.5 | endo |
| **481-b** | 179.0 | trans | -90.5 | 150.3 | -25.0 | 35.9 | -32.0 | 16.4 | -109.2 | exo |
| **501** | -177.5 | trans | -84.3 | 137.6 | -23.5 | 34.3 | -30.8 | 16.2 | -104.7 | exo |
| **527** | -179.8 | trans | -69.4 | -19.8 | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* |
| **548** | -175.1 | trans | -71.2 | -44.8 | -21.8 | 36.2 | -35.8 | 22.4 | -116.2 | exo |
| **550** | 179.6 | trans | -63.9 | -33.6 | -27.5 | 36.5 | -30.7 | 13.1 | -107.8 | exo |
| **555** | -174.1 | trans | -63.0 | -35.3 | -25.6 | 34.4 | -29.0 | 12.8 | -101.8 | exo |
| **579** | -2.2 | cis | -98.9 | 163.4 | 31.2 | -35.8 | 25.5 | -5.4 | 98.0 | endo |
| **585** | -173.9 | trans | -54.9 | 151.8 | -28.2 | 34.3 | -26.1 | 7.7 | -96.3 | exo |
| **589** | 177.2 | trans | -59.9 | -43.3 | -32.0 | 32.9 | -20.5 | -0.5 | -84.9 | exo |
| **650** | 176.3 | trans | -63.9 | 162.9 | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* |
| **656** | 179.8 | trans | -62.3 | -17.7 | -33.0 | 32.8 | -19.4 | -2.2 | -83.0 | exo |
| **685** | 177.2 | trans | -62.4 | 156.0 | -31.0 | 34.2 | -23.4 | 3.3 | -91.9 | exo |
| **701** | -176.5 | trans | -87.7 | 49.8 | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* | *n.a.* |
| **731** | -172.7 | trans | -101.2 | 8.0 | -15.8 | 31.6 | -34.2 | 25.2 | -106.8 | exo |
| **752** | -179.2 | trans | -65.7 | -39.5 | -25.3 | 36.9 | -33.6 | 17.6 | -113.5 | exo |
| **770** | 179.2 | trans | -64.5 | -28.2 | -26.1 | 36.1 | -31.4 | 14.8 | -108.4 | exo |
| **792** | 174.3 | trans | -62.1 | 146.2 | -27.0 | 34.0 | -27.0 | 9.7 | -97.7 | exo |
| **812** | -174.3 | trans | -56.6 | 119.2 | -27.4 | 35.3 | -28.6 | 11.1 | -102.5 | exo |
| **816** | 179.8 | trans | -70.6 | 135.3 | -28.7 | 36.0 | -28.6 | 10.2 | -103.6 | exo |

*n.a.* not assigned (lack of sufficient electron density)

[a] peptide bond conformation of Xaa-(4*R*)-FPro

[b] $\chi^{Pro} = \chi^1 + \chi^3 - \chi^2 - \chi^4$

[c] prolyl ring puckering based on $\chi^1$ and $\chi^{Pro}$

1 Peist, R., Koch, A., Bolek, P., Sewitz, S., Kolbus, T., and Boos, W. (1997) Characterization of the aes gene of Escherichia coli encoding an enzyme with esterase activity, *J. Bacteriol. 179*, 7679–7686.