

Structure and Stability of FlgD from the Pathogenic 26695 Strain of *Helicobacter pylori*

Ivana Kekez,^{1,2,*} Laura Cendron,³ Mario Stojanović,¹ Giuseppe Zanotti,^{2,#} Dubravka Matković-Čalogović¹

¹ University of Zagreb, Faculty of Science, Department of Chemistry, Division of General and Inorganic Chemistry, Horvatovac 102a, Zagreb, HR-10000, Croatia

² Department of Biomedical Sciences, University of Padua, Via Ugo Bassi 58/B, Padua, IT-35131, Italy

³ Department of Biology, University of Padua, Via Ugo Bassi 58/B, Padua, IT-35131, Italy

* Corresponding author's e-mail address: ikekez@chem.pmf.hr

Corresponding author's e-mail address: giuseppe.zanotti@unipd.it

RECEIVED: June 24, 2016 * REVISED: July 22, 2016 *ACCEPTED: July 22, 2016

THIS PAPER IS DEDICATED TO THE LOVING MEMORY OF IVANA WEYGAND-ĐURAŠEVIĆ (1952–2014)

Abstract: Flagellin component D (FlgD) is a hook capping protein involved in the proper structural assembly of bacterial flagella but is not present in the mature flagella. We report here the crystal structure of the truncated form of FlgD from the strain 26695 of *Helicobacter pylori* (HpFlgD_26695) in the *P2* space group at 2.8 Å resolution, along with data on protein stability. The structure includes only the central globular part of the protein composed of conserved Fn-III and tudor domains, while the N- and C-terminal regions are absent. This structure, together with the tetragonal crystal structure of HpFlgD_26695 and the monoclinic crystal structure of HpFlgD_G27 reported previously, revealed a different mutual orientation of the two domains in the HpFlgD structures in respect to FlgD from *Xanthomonas campestris* and *Pseudomonas aeruginosa*. In addition, investigation of protein stability in solution showed higher protein stability than in other bacterial orthologues reported so far.

Keywords: *Helicobacter pylori*, strain 26695, FlgD, stability, monoclinic crystal structure.

INTRODUCTION

HELICOBACTER PYLORI (*H. pylori*) is a small (3.5 μm × 0.5 μm), slow growing, microaerophilic, non-spore forming and spiral-shaped Gram-negative rod bacterium. The bacterium became an important and interesting target of research since it is known that it can colonize human gastric cells and lead to development of severe gastrointestinal diseases such as chronic gastritis, peptic ulcer disease, gastric mucosa associated lymphoid tissue (MALT) lymphoma, and gastric cancer.^[1,2] *H. pylori* flagellum is a complex structure composed of about 50 different proteins needed for its proper regulation and assembly. It provides the bacterial cell a chance to move to a more favorable environment and to prevent stressful conditions.^[3] Major sections that define the flagellum are: the filament, the hook and the basal

body.^[4] FlgD plays an important role in *H. pylori* virulence, owing to its role in forming the hook cap component.^[5] Moreover, FlgD interacts with FlhK and FlhB and regulates appropriate assembly of FlgE monomers in the hook structure.^[6] Until now crystal structures of two other FlgD homologues, from *Xanthomonas campestris* (XcFlgD, PDB code: 3C12)^[5] and *Pseudomonas aeruginosa* (PaFlgD, PDB code: 3OSV),^[7] have been determined, as well as two crystal structures of *H. pylori* FlgD from two different strains and in two different space groups (HpFlgD_26695_t, tetragonal space group *I422*, and HpFlgD_G27_m, monoclinic space group *P2*, with PDB codes 4ZZF and 4ZZK, respectively).^[8]

In this paper we describe the crystal structure of truncated FlgD from the *H. pylori* strain 26695 in a new space group, and discuss stability of the full length protein.

EXPERIMENTAL

Cloning, expression and purification of *HpFlgD_26695*

Cloning of the HP0907 gene from *Helicobacter pylori* (strain 26695, Figure 1) was performed as described in Pulić *et al.* (2016).^[8] pETite C-His vector (*Lucigen*), bearing gene coding for the *HpFlgD_26695* protein was transformed into *Escherichia coli* BL21(DE3) cells by electroporation on a Micropulser (*BioRad*). Transformed cells were grown on a solid Luria Bertani (LB) agar medium supplemented with 30 µg mL⁻¹ kanamycin. Grown colonies were used for protein overexpression in a 2 L LB liquid medium supplemented with 30 µg mL⁻¹ kanamycin at 301 K. Protein overexpression was induced with isopropyl-β-D-1thiogalactopyranoside (IPTG, *c* = 0.5 mmol L⁻¹) when OD₆₀₀ reached the value between 0.6–0.8. After 4 hours of induction the cells were harvested by centrifugation (*Sigma 6K15*) at 4000 rpm and 15 min at 277 K. Harvested cells were resuspended in buffer (20 mmol L⁻¹ Tris-base pH 7.5, 150 mmol L⁻¹ NaCl, 3 mmol L⁻¹ phenylmethylsulfonyl fluoride (PMSF), 0.1 µg mL⁻¹ lysozyme, 0.05 µg mL⁻¹ DNase I, 1 mmol L⁻¹ MgCl₂) and broken down by sonication (*Vibra Cell, Bioblock Scientific*). The soluble portion of the lysate was separated by centrifugation (4500 rpm, 20 min, 277 K) and further purified on a 1 mL HisTrap column (*GE Healthcare*) equilibrated with buffer A (20 mmol L⁻¹ Tris-base pH 7.5, 150 mmol L⁻¹ NaCl). Washing step of 50 mL of buffer A supplemented with 20 mmol L⁻¹ imidazole was performed to eliminate unspecifically bound proteins. Finally, elution of the protein was performed in buffer A with a stepwise gradient of imidazole from 20 to 500 mmol L⁻¹. Size-exclusion chromatography was conducted on an ÄKTA Purifier (*GE Healthcare*) using Superdex 200 10/300 column (*GE Healthcare*). Purity of the obtained fractions was monitored by electrophoresis under denaturing conditions (SDS-PAGE), and fractions with the highest concentration of protein were further concentrated by ultrafiltration in the Centricon system.

Protein Crystallization and X-ray Data Collection

Purified *HpFlgD_26695* at a concentration of 17.5 mg mL⁻¹ was used in the crystallization experiment. Crystallization was performed by the sitting drop vapor diffusion method with the Oryx 8 crystallization robot (*Douglas Instruments*) using 1 µL of protein mixed with 1 µL precipitant solution and equilibrated against 75 µL of the mother liquor in the reservoir. Monoclinic crystals of *HpFlgD_26695* were grown from the screening solution 6 of the PACT suite (*Qiagen*) (0.1 mol L⁻¹ SPG pH 9.0, 25 % PEG 1500) at 293 K.

A crystal of approximate size of 0.15 × 0.05 × 0.02 mm³ was frozen directly in liquid nitrogen with no cryoprotectant. Diffraction data were collected at 100 K at the Elettra synchrotron, beamline XRD1, on the 2M Pilatus detector. 360 frames at 0.5° rotation range were collected. Diffraction data were indexed and integrated with iMOSFLM,^[9] and scaled with Scala^[10] contained in the CCP4i crystallographic package.^[11]

Structure Solution and Refinement

The structure of *HpFlgD_26695* was solved by molecular replacement using AMoRe,^[12] implemented in the CCP4i suite, with the structure of the tetragonal crystal form of *HpFlgD_26695* (PDB entry 4ZZF) as the search model. Iterative model building (using Coot),^[13] and refinement cycles (Refmac5,^[14] implemented in CCP4i) were used to complete the structure. MolProbity^[15] was used for the evaluation of the refinement quality. Figures of the protein structure were prepared with ccp4mg.^[16]

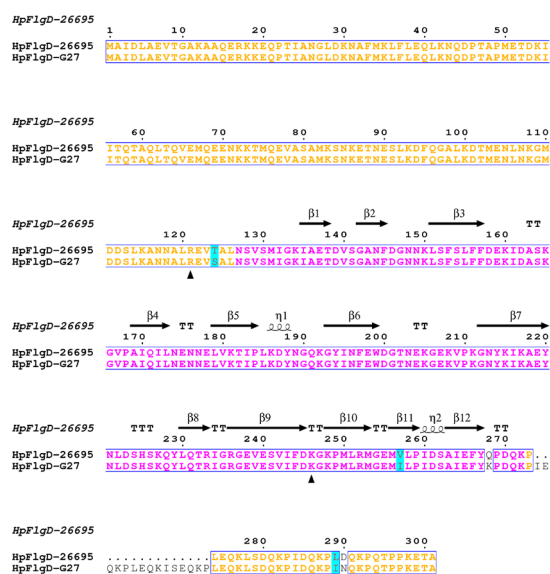


Figure 1. Sequence alignment of *HpFlgD_26695* (HP, O25565_HELPHY) and *HpFlgD_G27* (HP, B5Z7R3_HELPG) using CLUSTAL OMEGA.^[20] The alignment figure was generated using ESPrpt 3.0.^[21] The numbering system is based on the *HpFlgD_26695* sequence. Residues visible in the monoclinic crystal structures of *HpFlgD_26695* and *HpFlgD_G27* are labelled in pink while the rest of the sequences residues are labelled in orange. Similar residues in sequences are highlighted in blue, while residues of low or no similarity are labelled in black. Labels of the secondary structure elements are shown as well. Possible cleavage sites accessible to trypsin in both sequences are signed with black triangles.

Stability Studies

To investigate stability of *HpFlgD_26695* a limited proteolysis experiment was carried out. Each of the four samples of *HpFlgD_26695* ($\gamma = 3 \text{ mg mL}^{-1}$) in buffer A was digested with a selected protease (trypsin, α chymotrypsin, thermolysin or thrombin) in a 1 : 1000 and 1 : 500 ratio at 298 K, and incubated for an hour. The yield of digestion was monitored by SDS-PAGE. The product obtained by trypsin digestion was further purified by size-exclusion chromatography (SEC) using the Superdex 200 10/300 column (*GE Healthcare*).

The influence of the components of the crystallization solution on the protein stability was also examined. *HpFlgD_26695* was mixed with PEG 1500 and SPG buffer to reach the final values of (w/v) 25 % and 0.1 mol L^{-1} , respectively, at different pH values (pH 4, pH 6–9) and left at room temperature (298 K). In addition, time-course stability experiment of protein in buffer A without any additives was performed to monitor the cleavage kinetics. Stability of the protein over time was followed by SDS-PAGE.

RESULTS AND DISCUSSION

Crystal Structure of the Monoclinic Form of *HpFlgD_26695* and Its Comparison with Other FlgD Structures

Data collection and processing statistics of the monoclinic form of the *HpFlgD_26695* protein (*HpFlgD_26695_m*) are given in Table 1. *HpFlgD_26695_m* crystallizes in the monoclinic form, space group *P2*, which is a new crystal

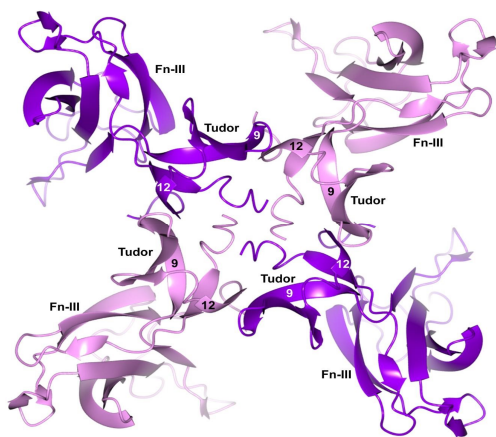


Figure 2. *HpFlgD_26695_m* tetramer with monomer A shown in purple and monomer B in pink color. Beta sheets (β_9 and β_{12}) of the two tudor domains from neighbouring monomers responsible for tetramerization are labelled in white (in monomer A) and black (in monomer B).

form in addition to the previously reported tetragonal form found for *HpFlgD_26695* (*HpFlgD_26695_t*).^[8] The structure, as was the case in the previously described structures (PDB codes: 4ZZK, 4ZZF),^[8] is truncated with missing C- and N-termini (Figure 1). From the electron density map 146 amino acids (Asn127–Lys272) can be built. Statistics on the data refinement are given in Table 1. The overall B-factor of the monoclinic form of *HpFlgD_26695* (69.2 \AA^2) is slightly higher than in the previously solved monoclinic structure of *HpFlgD_G27* (62.1 \AA^2). Such high values of B-factors are consequence of low resolution data. The asymmetric unit contains 4 protein molecules and 17 water molecules. *HpFlgD_26695_m* occupies a V_m of $2.64 \text{ \AA}^3 \text{ Da}^{-1}$ and has a solvent content of around 53 %. As in the case of the previously reported *HpFlgD* structures,^[8] the monomer consists of two domains, the fibronectin domain III (Fn-III) and the tudor domain (Figure 2). Overlaid monomer structures of the two published structures, the monoclinic (*HpFlgD_G27_m*) and tetragonal (*HpFlgD_26695_t*) forms, with the present monoclinic form (*HpFlgD_26695_m*) show r.m.s.d.s. of 0.27 \AA and 0.44 \AA , respectively, between 146 equivalent C^α atoms (Figure 3). As shown in Figure 3, the largest structural difference in all three superposed *HpFlgD* structures is present in the loop region positioned in the tudor domain (from Asp245 to Lys248).

As in the previously reported crystal structures of *HpFlgD_26695_t* and *HpFlgD_G27_m* the quaternary structure in *HpFlgD_26695_m* is a tetramer.^[8] Figure 2

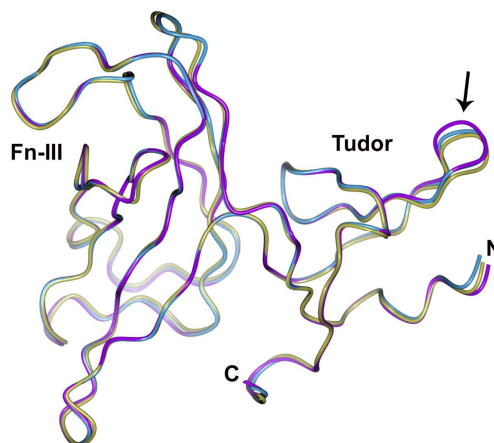


Figure 3. Superposition of the monomers from the crystal structures of *HpFlgD_26695_m_A* (purple), *HpFlgD_26695_t* (gold, PDB code: 4ZZF) and *HpFlgD_G27_m_A* (light blue, PDB code: 4ZZK).^[8] The r.m.s.d.s. for the superposition of 146 aligned C^α atoms of *HpFlgD_26695_t* on *HpFlgD_26695_m_A* and 146 aligned C^α atoms of *HpFlgD_G27_m_A* on *HpFlgD_26695_m_A* are 0.44 \AA and 0.27 \AA , respectively. A black arrow shows the region with major structural differences in all three *HpFlgD* structures.

shows the tetramer assembly in the structure of *HpFlgD_26695_m* formed, as in the structure of *HpFlgD_G27_m*, by monomers A and B, together with their pair related by a two-fold symmetry axis. Monomers C and D form another tetramer in the same way. The two tudor domains of the neighbouring monomers within the tetramer are connected through the network of hydrogen bonds and salt bridges (Table 2). PISA program^[17] was used

Table 1. Statistics of data collection, processing and refinement. Values for the outer shell are given in parentheses.

<i>HpFlgD_26695</i>	
Crystallographic data	
Crystal system	Monoclinic
Space group	<i>P</i> 2
Wavelength / Å	1.000
Unit cell parameters / Å, °	
<i>a</i>	77.51
<i>b</i>	34.00
<i>c</i>	133.34
β	98.90
unit cell	8
<i>Z</i> asymmetric unit	4
Resolution range / Å	76.58–2.80 (2.95–2.80)
<i>R</i> merge	0.113 (0.586)
<i>R</i> pim	0.108 (0.553)
$\langle I / \sigma(I) \rangle$	4.9 (1.3)
<i>CC</i> half	0.988 (0.639)
Completeness / %	94.8 (94.5)
Redundancy	2.7 (2.7)
Refinement	
Resolution range / Å	62.16–2.85
<i>R</i> work / <i>R</i> free	0.227 / 0.256
No. protein / solvent atoms	4624 / 17
Mean <i>B</i> / Å ²	69.2
R.m.s.d. from ideal value	
Bond length / Å	0.011
Bond angles / °	1.51
Geometry	
Ramachandran favored / %	97.0
Ramachandran outliers / %	0
Rotamer outliers / %	2.2
Overall score	1.37

to validate the biological assembly from the crystal structure packing and the analysis suggests that the tetramer observed in the crystal corresponds to the molecule in solution, as described in Pulić *et al.* (2016).^[8] The largest differences in the hydrogen bond length in the monoclinic structures of *HpFlgD_26695_m* and *HpFlgD_G27_m* are found for hydrogen bonds between the [O ϵ 2] atom of Glu265 and [O γ] atom of Ser241, connecting monomers C[$-x, y, -z$] and D, B and A, D[$x-1, y, z-1$] and C, with values of 0.18 Å, 0.20 Å and 0.35 Å, respectively, however these differences are not significant. Moreover, when comparing the tetragonal and monoclinic crystal structures of *HpFlgD_26695*, the hydrogen and salt bridges bond length differences are even greater, particularly in case of Glu265 [O ϵ 2]...Arg252 [NH2] and Glu265 [O ϵ 2]...Ser241 [O γ] contacts where the range of differences vary between 0.69–0.88 Å and 0.40–0.59 Å, respectively. Dimerization of monomers D and C[$-x, y, -z$] through the β 12 and β 9 sheets, respectively, creates the largest interface area that corresponds to 517 Å² and involves 10 hydrogen bonds and 1 salt bridge (Table 3). In the case of *HpFlgD_G27_m* the symmetrically equivalent monomers occupy a smaller interface area of 492.5 Å² which is formed by 6 hydrogen bonds and 2 salt bridges.^[18] The interface involved in tetramerization in the tetragonal structure of *HpFlgD_26695* occupies the smallest interface area of 478.8 Å² but involves the highest number of hydrogen bonds and salt bridges (12 H-bonds and 2 salt bridges).^[18] In addition to the main interface (labelled t), which is involved in tetramerization, the other type of interfaces (labelled a–c) are also present in the structure of *HpFlgD_26695_m*, but form smaller interface areas (Table 3).

Other bacterial orthologues of the FlgD protein, whose crystal structures have been solved (*XcFlgD*, *PaFlgD*),^[5,7] show the same tendency to lose the N-terminal part during crystallization. The *HpFlgD* sequence contains an extra 44 and 29 amino acids at the C-terminus in the G27 and 26695 strains, respectively in comparison to *XcFlgD* and *PaFlgD*. These amino acids are not present in the crystal structures of *HpFlgD* (Figure 1).

In all FlgD structures, the monomer is composed of conserved Fn-III and tudor domains. The most important difference is the different mutual orientation of the Fn-III and tudor domains in *HpFlgD* with respect to the same domains in *XcFlgD* and *PaFlgD*. When the Fn-III domain of *PaFlgD* is superposed on the same domain in *HpFlgD_26695_m* it can be seen that the tudor domains are not superposed but are rotated in respect to each other by approx. 115° (Figure 4a). The tudor domains by itself superimpose well in these two structures (Figure 4b). For more details on the description of the differences between the *HpFlgD* and its orthologues see Refs. Pulić *et al.* (2016).^[8,18]

Table 2. Hydrogen bonds (Å) between monomers that are relevant for their assembly into a tetramer in the monoclinic crystal structure of *HpFlgD_26695*. * Denotes the salt bridge.

Monomer1...Monomer2 [Symmetry code]	B...A[-x, y, -z]	B...A	C[-x, y, -z]...D	D[x-1, y, z-1]...C
Ile264 [O] Phe244 [N]	2.90	2.78	2.90	2.86
Phe266 [N] Val242 [O]	2.80	2.87	2.83	2.91
Phe266 [O] Val242 [N]	2.89	3.00	2.89	3.07
Glu265 [O] Ser241 [O]	3.51	3.40	3.33	3.52
*Glu265 [O] Arg252 [NH2]	2.68	2.56	2.55	2.74

Table 3. Different types of interfaces between the monomers in the monoclinic crystal structure of *HpFlgD_26695*. N_{HB} and N_{SB} correspond to the number of hydrogen bonds and salt bridges, respectively. The interface type signed as t is involved in tetramerization. The interface areas were calculated using the PISA software.^[17]

Interface type	Monomer1...Monomer2 [Symmetry code]	Interface area / Å ²	N_{HB}	N_{SB}
t	D C[-x, y, -z]	517.0	10	1
t	C D[x-1, y, z-1]	491.9	8	1
t	B A[-x, y, -z]	481.8	8	1
t	B A[x, y, z]	474.1	7	1
c	D D[x, y-1, z]	307.7	1	3
c	A A[x, y-1, z]	304.6	3	5
c	C C[x, y-1, z]	253.3	4	5
c	B B[x, y-1, z]	251.2	3	5
b	D B[-x+1, y, -z]	267.0	3	-
a	C B[x, y, z]	205.8	-	-

Limited Proteolysis Studies

Since the previously reported crystal structures of *HpFlgD* were truncated and the growth of crystals took 2 months,^[8] we wanted to test if limited proteolysis could eliminate the flexible regions present at the surface of the protein and thus speed up the crystallization process. As previously described, *HpFlgD_26695* was cleaved with four proteases. Three of the four enzymes gave similar results enriching the pool of species at about 14 kDa, while thrombin did not cut the protein (Figure 5a). The most homogenous and discrete digestion pattern was obtained with trypsin and this sample was used for further experiments. The protein digested with trypsin was eluted from the SEC column at a volume of 17.11 mL (Figure 6), showing a reduction of molecular weight in comparison with the full length *HpFlgD_26695*.^[8] A solution containing a monomer of about 15 kDa was obtained. The size of protein estimated by MALS measurement was in agreement with the one estimated by size exclusion calibration with protein standards (MW digested *HpFlgD_26695* = 15 kDa). Analyses of *HpFlgD* sequences

by Peptide cutter (http://web.expasy.org/peptide_cutter/) showed 2 potential cleavage sites for the trypsin protease, positioned at the sequence number 121 (LR|EV) and 246 (DK|GK), Figure 1. The 3D model of the full length *HpFlgD_26695* was built by the ITASSER software.^[19] Investigation of the modelled and truncated structures of *HpFlgD_26695* imply that the potential cleavage sites (Arg121, Lys246) are positioned in loop regions at the surface of the protein and could be accessible for trypsin cleavage (Figure 7). In those cases, cleavage would generate a truncated version of the protein whose size is compatible with the crystallized fragment.

Such *in vitro* conditions somehow reproduce a result similar to what occurs in the crystallization conditions, but they give rise to a protein fragment 2.4 kDa shorter than the one visible in previously described *HpFlgD* crystal structures (16.4 kDa, PDB codes: 4ZZK, 4ZZF).^[8] The distinct processing levels encountered in the different experimental conditions most likely produced multiple oligomerization states observed in our study. Indeed, differently from the full length protein (that forms dimers)

and the crystallized species (that associate as tetramers),^[8] such 14 kDa proteolysis product behaves as a monomer in solution. These results can be explained if we look at the tetramers observed in the crystal structures: a 14 kDa product can be easily obtained thanks to further degradation of the crystallized fragment. Indeed, if a 26 amino acids loss occurs at the C-terminus, implying the removal of the last three beta strands (β_{10} – β_{12}) of the crystallized species. This includes the loss of β -strand 12 that is directly involved in the main interaction (β_{12} and β_9 strands) that enables protein tetramerization (Figure 2).

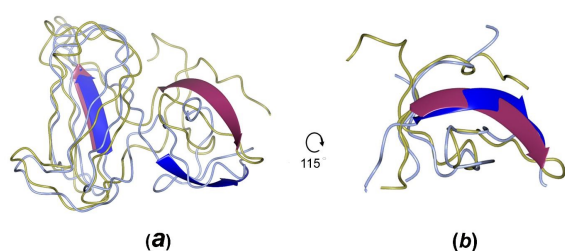


Figure 4. (a) Different spatial orientation of the two domains in *HpFlgD_26695_m* (gold colour) and *PaFlgD* (light blue color, PDB code: 3OSV_A).^[7] Amino acid residues used for the torsion angle calculation are highlighted in dark purple and dark blue in *HpFlgD_26695_m* and *PaFlgD*, respectively. (b) Overlaid tudor domains of *PaFlgD* and *HpFlgD_26695_m*. The r.m.s.d. for the superposition of 43 aligned C α atoms of *PaFlgD* on *HpFlgD_26695_m* is 2.06 Å.

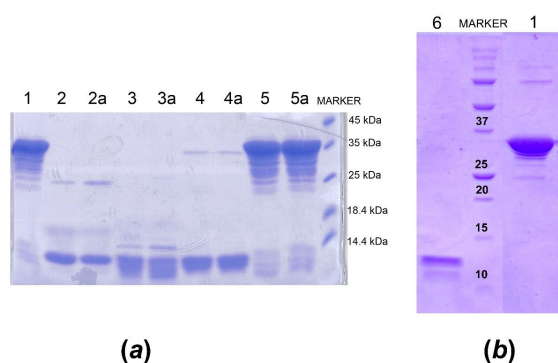


Figure 5. Results of the stability studies of *HpFlgD_26695* monitored by SDS-PAGE. (a) 1 – full length protein (top band) with initial degradation products (lower bands); 2, 3, 4, 5 – 1:1000 ratio of digested *HpFlgD_26695* with trypsin, α -chymotrypsin, thermolysin and thrombin, respectively, and 2a, 3a, 4a, 5a – 1:500 ratio digested *HpFlgD_26695* with the same proteases. (b) 1 – full length protein left at room temperature for 3 weeks; 6 – full length protein left in SPG buffer pH 4.

Stability of *HpFlgD_26695* in Solution

All flagellin components D whose crystal structures are known (PDB codes: 3C12, 3OSV, 4ZZK, 4ZZF) were isolated as full length proteins; however, in the crystal structures they are lacking the C- and N-termini.^[5,7,8] *H. pylori* survives in the stomach environment at pH 4–6.5. The FlgD protein is a hook-capping protein and, similarly to other *H. pylori* flagellar proteins, is predicted to have a pI value of around 5. SDS-PAGE confirmed that full length protein left in buffer A for 3 weeks at room temperature was intact and no lower molecular weight degradation

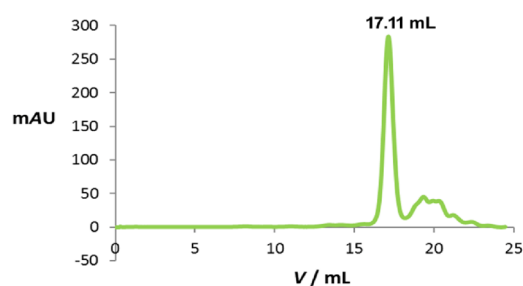


Figure 6. Analytical gel filtration chromatogram of the trypsin digested *HpFlgD_26695* (green curve; MW of the monomer ~14 kDa) which elutes as a single peak at the volume of 17.11 mL corresponding to the molecular weight of ~15 kDa.

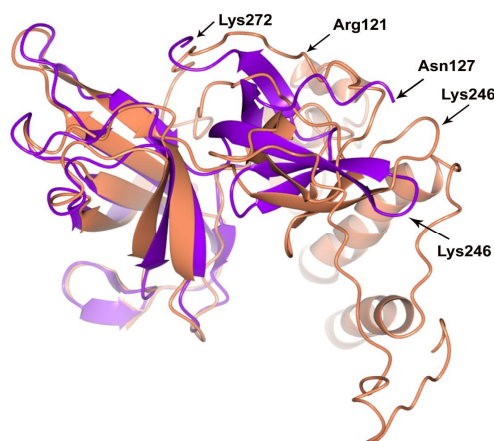


Figure 7. Overlaid structures of the monomers of the *HpFlgD_26695* (purple) and modeled full length *HpFlgD_26695* (orange). Black arrows indicate possible positions for trypsin cleavage (Arg121, Lys246) in both structures and the first and last residues visible in the crystal structure of the monoclinic form of *HpFlgD_26695* (Asn127, Lys272). The r.m.s.d. for the superposition of 112 C α atoms of modeled full length *HpFlgD_26695* on *HpFlgD_26695_m_A* is 2.13 Å

products were obtained (Figure 5b). Addition of PEG 1500 (*w/v* 8.5 % and 25 %) to the solution of the full length protein in buffer A or in the SPG buffer at pH values from 6 to 9 did not provoke any protein degradation after 3 weeks, as well. A lower molecular weight product at around 14 kDa was only observed in the sample left in the SPG buffer at pH 4 (Figure 5b). For XcFlgD a time-course auto-cleavage experiment was performed at 298K.^[5] It was found that XcFlgD showed a tendency to degrade to 20 kDa and 15 kDa products after 5 days, while after 10 days the 20 kDa band disappeared and only a band at approximately 15 kDa remained. All these findings suggest that components of the mother liquor do not affect protein stability in the solution and that HpFlgD_26695 shows greater stability during the 3-week period in comparison to its ortholog from *X. campestris*. It was hard to mimic in solution the cleavage events that happened during crystallization without any protease treatment.

CONCLUSIONS

HpFlgD_26695 was crystallized in a new crystal form in the monoclinic, *P2* space group. As in the previously reported crystal structures of truncated HpFlgD, the protein is missing the N- and C-terminus portions and is composed of a globular region divided in the two conserved domains, Fn-III and tudor. The main structural difference among HpFlgD structures occurs in the loop region (Asp245 to Lys248) of the tudor domain. Comparing the structure with other flagellin D orthologues, the most significant discrepancy is found in the mutual orientation of the two domains. Limited proteolysis experiments with trypsin showed that a protein fragment of a similar size as in the crystal could be obtained, but it probably lacks the last three β -strands of the tudor domain and hence is unable to form the tetramer. Time-course stability experiments produced a lower molecular weight (15 kDa) product only when the SPG buffer at pH 4 was used, suggesting higher protein stability in solution than in other flagellin D orthologues.

Acknowledgment. We thank the staff of beamline XRD1 of Elettra synchrotron (Trieste, Italy) for technical assistance during data collection. The research leading to these results were funded by the European Community's Seventh Framework Program under grant agreement BIOSTRUCTX_7860, by the University of Padua (PRIN 2010-2011 (MIUR) "Unraveling structural and functional determinants behind *Helicobacter pylori* pathogenesis and persistence"), and by funds from the University of Zagreb.

Data Deposition. Atomic coordinates and structure factors have been deposited at The Protein Data Bank for immediate release as 5K5Y.

REFERENCES

- [1] C. S. Goodwin, M. M. Mendall, T. C. Northfield, *Lancet*. **1997**, *349*, 265.
- [2] J. G. Kusters, A. H. van Vliet, E.J. Kuipers, *Clin. Microbiol. Rev.* **2006**, *19*, 449.
- [3] M. R. Amieva, E. M. El-Omar, *Gastroenterology* **2008**, *134*, 306.
- [4] O. A. Soutourina, P. N. Bertin, *FEMS Microbiol. Rev.* **2003**, *27*, 505.
- [5] W. T. Kuo, K. H. Chin, W. T. Lo, A. H. Wang, S.H. Chou, *J. Mol. Biol.* **2008**, *381*, 189.
- [6] R. M. Macnab, *Annu. Rev. Microbiol.* **2003**, *57*, 77.
- [7] H. Zhou, M. Luo, X. Cai, J. Tang, S. Niu, W. Zhang, Y. Hu, Y. Yin, A. Huang, D. Wang, *Proteins* **2011**, *79*, 2346.
- [8] I. Pulić, L. Cendron, M. Salamina, P. Polverino de Laureto, D. Matković-Čalogović, G. Zanotti, *J. Struct. Biol.* **2016**, *194*, 147.
- [9] T. G. G. Batty, L. K. Ontogiannis, O. Johnson, H. R. Powell, A. G. W. Leslie, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 271.
- [10] P. Evans, *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 72.
- [11] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235.
- [12] J. Navaza, *Acta Cryst. A* **1994**, *50*, 157.
- [13] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486.
- [14] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R.A., Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355.
- [15] Chen *et al.*, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 12.
- [16] S. McNicholas, E. Potterton, K. S. Wilson, and M. E. M. Noble, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 386.
- [17] E. Krissinel, K. Henrick, *J. Mol. Biol.* **2007**, *372*, 774.
- [18] I. Pulić, L. Cendron, M. Salamina, P. Polverino de Laureto, D. Matković-Čalogović, G. Zanotti, *Data in Brief* **2016**, *7*, 493.
- [19] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, *Nat. Methods* **2015**, *12*, 7.
- [20] H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley, and R. Lopez, *Nucleic Acids Res.* **2013**, *41*, W597.
- [21] X. Robert, P. Gouet, *Nucleic Acids Res.* **2014**, *42*, W320.