



**HAL**  
open science

## A Poly-Proline II Helix in YadA from *Yersinia enterocolitica* serotype O:9 Facilitates Heparin Binding through Electrostatic Interactions

Ina Meuskens, Per Eugen Kristiansen, Benjamin Bardiaux, Vladimir Rosenov Koynarev, Daniel Hatlem, Kristian Prydz, Reidar Lund, Nadia Izadi-pruneyre, Dirk Linke

### ► To cite this version:

Ina Meuskens, Per Eugen Kristiansen, Benjamin Bardiaux, Vladimir Rosenov Koynarev, Daniel Hatlem, et al.. A Poly-Proline II Helix in YadA from *Yersinia enterocolitica* serotype O:9 Facilitates Heparin Binding through Electrostatic Interactions. *FEBS Journal*, In press, 10.1111/febs.17001 . hal-04290267

**HAL Id: hal-04290267**

**<https://hal.science/hal-04290267>**

Submitted on 16 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Bardiaux Benjamin (Orcid ID: 0000-0003-4014-9195)  
Izadi-Pruneyre Nadia (Orcid ID: 0000-0002-6864-2961)  
Linke Dirk (Orcid ID: 0000-0003-3150-6752)

Color : Fig 1-6

# A Poly-Proline II Helix in YadA from *Yersinia enterocolitica* serotype O:9 Facilitates Heparin Binding through Electrostatic Interactions

Ina Meuskens<sup>a#</sup>, Per Eugen Kristiansen<sup>a</sup>, Benjamin Bardiaux<sup>b</sup>, Vladimir Rosenov Koynarev<sup>c</sup>, Daniel Hatlem<sup>a</sup>, Kristian Prydz<sup>a</sup>, Reidar Lund<sup>c</sup>, Nadia Izadi-Pruneyre<sup>d</sup>, and Dirk Linke<sup>a\*</sup>

<sup>a</sup> Department of Biosciences, University of Oslo, PO Box 1066 Blindern, NO-0316 Oslo, Norway <sup>b</sup> Institut Pasteur, Université de Paris-Cité, CNRS UMR3528, Structural Bioinformatics Unit, F-75015 Paris, France

<sup>c</sup> Department of Chemistry, University of Oslo, PO Box 1033 Blindern, NO-0316 Oslo, Norway

<sup>d</sup> Institut Pasteur, Université de Paris-Cité, CNRS UMR3528, Bacterial Transmembrane Systems Unit, F-75015 Paris, France

<sup>#</sup> present address: Institute of Molecular Biology & Biophysics, ETH Zurich, Otto-Stern-Weg 5, CH-8093 Zurich, Switzerland

## Correspondence

\* D. Linke, Department of Biosciences, University of Oslo, PO Box 1033 Blindern, NO-0316, Norway, Tel: +47 22857654, E-mail: [Dirk.Linke@ibv.uio.no](mailto:Dirk.Linke@ibv.uio.no)

## Running title

YadA<sub>O:9</sub> Poly-Proline Helix promotes heparin binding

## Abbreviations

COG (conserved oligomeric Golgi), ECM (extracellular matrix), GAG (glycosaminoglycan), HS (heparan sulphate), PPII helix (poly proline II helix), TAA (trimeric autotransporter adhesin), YadA (*Yersinia* adhesin A), KD (dissociation constant), CSP (chemical shift perturbation), DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid), CD (circular dichroism), SD (standard deviation), RMSD (root mean square deviation), sfGFP (superfolder GFP).

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.1111/febs.17001](https://doi.org/10.1111/febs.17001)

## Keywords

*Yersinia enterocolitica*, YadA, poly proline II helix, glycosaminoglycans, heparin

## Conflict of Interest

The authors declare no conflict of interest.

## Abstract

Poly-proline II helices are secondary structure motifs frequently found in ligand binding sites. They exhibit increased flexibility and solvent exposure compared to the strongly hydrogen-bonded  $\alpha$ -helices or  $\beta$ -strands and can therefore easily be misinterpreted as completely unstructured regions with an extremely high rotational freedom. Here, we show that the adhesin YadA of *Yersinia enterocolitica* serotype O:9 contains a poly-proline II helix interaction motif in the N-terminal region. The motif is involved in the interaction of YadA<sub>O:9</sub> with heparin, a host glycosaminoglycan. We show that the basic residues within the N-terminal motif of YadA are required for electrostatic interactions with the sulphate groups of heparin. Biophysical methods including CD spectroscopy, solution-state NMR, and SAXS all independently support the presence of a poly-proline helix allowing YadA<sub>O:9</sub> binding to the rigid heparin. Lastly, we show that host cells deficient in sulphation of heparin and heparan sulphate are not targeted by YadA<sub>O:9</sub>-mediated adhesion. We speculate that the YadA<sub>O:9</sub>-heparin interaction plays an important and highly strain-specific role in the pathogenicity of *Yersinia enterocolitica* serotype O:9.

## Introduction

*Yersinia* adhesin A (YadA) is a surface-exposed protein found in *Yersinia enterocolitica* (*Y. enterocolitica*) as well as in *Yersinia pseudotuberculosis* (*Y. pseudotuberculosis*). As a member of the family of trimeric autotransporter adhesins (TAAs) [1], [2], it is essential for pathogenesis and host colonization [3], [4]. YadA aids in interaction of the pathogen with host extracellular matrix (ECM) proteins like collagen and vitronectin [5]–[8]. Structurally, YadA can be separated into three regions, the C-terminal,  $\beta$ -barrel anchoring YadA in the bacterial outer membrane, a coiled-coil stalk domain facing towards the bacterial environment and a N-terminal  $\beta$ -roll head domain conferring most of the ligand-binding ability [9]–[11]. While YadA is highly conserved between *Yersinia* species, the N-terminal region of the head domain displays length variations between different *Y. enterocolitica* serotypes (Fig. 1a) [7]. In most *Y. enterocolitica* serotypes, YadA exhibits a short, N-terminal region that I not resolved in the crystal structure of YadA of *Y. enterocolitica* serotype O:3 (YadA<sub>O:3</sub>) due to its flexibility (Fig. 1b) [11]. Compared to that, YadA from *Y. enterocolitica* serotype O:9 (YadA<sub>O:9</sub>) exhibits a 35-residue insertion in the N-terminal region (Fig. 1a, b) [7]. We were recently able to show that this insertion in YadA<sub>O:9</sub> facilitates binding to the N-linked glycans of vitronectin as well as to heparin [8]. Heparin and heparan sulphate (HS) are structurally similar glycosaminoglycans (GAGs) which are heavily sulphated

and thus negatively charged (Fig. 1c) [12]. GAGs are part of the host ECM where they confer hydration [12], [13]. The interaction between YadA<sub>0:9</sub> and heparin is to date one of only two Type V secreted proteins that interact with a glycan [8], [14]. Here we elucidate the chemical and structural nature of the interaction between YadA<sub>0:9</sub> and heparin. We show that the interaction is based on electrostatic interactions. Furthermore, we conclude from various, complementary experiments that the N-terminal insertion within YadA<sub>0:9</sub> adopts a poly-proline II (PPII) helix conformation. We hypothesize that this structural motif helps to orient the basic residues of the YadA<sub>0:9</sub> N-terminal region towards the ligand. These results tie in well with host cell binding experiments showing that cells deficient in HS synthesis are significantly less affected by YadA<sub>0:9</sub>-mediated bacterial adhesion.

## Results

### Glycan binding of YadA<sub>0:9</sub> relies on electrostatic interactions

Recently, we described that YadA<sub>0:9</sub> binds to heparin via an N-terminal motif that is composed of a strikingly high number of proline and basic residues (Fig. 1a). We hypothesized that the interaction between YadA<sub>0:9</sub> and the negatively charged heparin might rely on electrostatic interactions. Heparin and HS consist of repeating units of D-Glucuronic acid  $\beta$ -(1 $\rightarrow$ 4)-linked to D-N-acetylglucosamine (Fig. 1c) [15]. The repeating units exhibit variable amounts of O-sulphates in the C3 and C6 position of the glucosamine unit, the C2 position of the hexuronic acid as well as N-sulphate groups (Fig. 2a) [12]. Heparin exhibits on average 2.7 sulphates per repeating unit, HS typically only exhibits a single sulphate group per repeating unit [16]. To investigate whether the negative charges conferred by sulphate groups are a prerequisite for the interaction, we performed binding experiments between YadA<sub>0:9</sub> head domains and three different glycans: Heparin was used as a heavily negatively charged, linear GAG. Hyaluronic acid was included as a GAG without sulphates [13]. Dextran, an uncharged, branched-chain glycan of bacterial origin was included as a non-GAG control (Fig. 2a) [17].

Binding experiments using biotinylated glycans with surface-coated YadA<sub>0:9</sub> head domains (Fig. 2b) demonstrate that heparin binds YadA<sub>0:9</sub> in a concentration-dependent manner, while no binding was observed for the non-sulphated glycans (Fig. 2c). The results indicate that heparin is bound to YadA<sub>0:9</sub> in a 1:1 molar stoichiometry, suggesting that three heparin molecules can be bound per homotrimeric YadA<sub>0:9</sub> head domain (Fig. 2c). The apparent binding affinity was determined to be in the lower nM range with an approximate dissociation constant ( $K_D$ ) of 30 nM. Our results thus show that indeed, strong negative charges conferred by sulphate groups are a prerequisite for YadA<sub>0:9</sub>-mediated glycan interaction.

To confirm the direct involvement of the basic residues in the YadA<sub>0:9</sub> head domain in heparin binding, we performed binding studies using purified YadA<sub>0:9</sub> head domains harboring point mutations of the basic residues to alanine (R/K $\rightarrow$ A) (Fig. 2d). YadA<sub>0:8</sub> was included as a control as it lacks the heparin-binding motif (Fig. 1a/b). Significant differences in heparin binding could be observed for proteins with alanine substitutions of residue K7, R21 or R24 all showing reduced binding to heparin. Mutation K12A shows a significant increase in heparin binding (Fig. 2d). We hypothesize that not all basic residues within the YadA<sub>0:9</sub> N-terminal region contribute equally to heparin binding. This might be due to steric hindrance in context of the trimeric configuration of the full-length head domain. We thus confirm that some of the basic residues within the YadA<sub>0:9</sub> N-terminal peptide are involved in binding of heparin. To investigate whether point mutations of the basic residues within the YadA<sub>0:9</sub> head domain would impact heparin binding in the context of full-length YadA<sub>0:9</sub> being expressed on the cell surface, we

introduced the aforementioned mutations into the N-terminal motif of full-length YadA<sub>O:9</sub> (Fig. 2e) [8]. Interestingly, and in contrast to the binding experiments with purified YadA<sub>O:9</sub> head domains, we could not observe any significant changes in heparin binding. Notably, the strength of YadA-mediated binding to surfaces might be caused by an avidity effect rather than the individual affinity of YadA to its ligand [18]–[20]. This could explain why we could monitor changes in heparin binding when using individual YadA<sub>O:9</sub> head domains as opposed to full length YadA<sub>O:9</sub> displayed on the surface.

To investigate the involvement of specific residues in the interaction, we decided to employ NMR as an additional method. As we used a synthetic peptide derived from the native heparin binding motif as found in YadA<sub>O:9</sub>, we initially confirmed that the peptide alone binds to heparin outside of its native, trimeric YadA<sub>O:9</sub> head domain environment (Fig. 1b). A construct was made inserting the heparin-binding peptide into one of the loops of sfGFP. Note that sfGFP, in contrast to the obligate trimeric YadA, is a monomeric protein. Using this construct, we could monitor binding between heparin and the sfGFP\_heparin-binding peptide in a dot blot (Fig. 2f). Indeed, like for the head domain of YadA<sub>O:9</sub> we observed binding of heparin to the sfGFP\_YadA<sub>O:9NT</sub>, indicating that no other region within the YadA<sub>O:9</sub> head domain is required for heparin binding and that trimerization was not necessary for binding.

We proceeded to analyze the interaction between the heparin-binding peptide and heparin by NMR. For initial peak assignments the pep3 peptide was used (Fig. 3a). Employing a combination of TOCSY, NOESY, DQCOSY, <sup>1</sup>H-<sup>13</sup>C HSQC, <sup>1</sup>H-<sup>15</sup>N HSQC and HMBC spectra, we were able to assign 100 % of the backbone resonances. During the assignment, we observed a noticeable downfield shift for H $\alpha$  chemical shifts of residues preceding Pro residues in the TOCSY spectra. This simplified the assignment of the affected residues. Similar downfield shifts in the context of Pro-rich peptides have been described earlier and might indicate a distinct behavior of these residues due to added rigidity in the amide bond between Pro and the preceding residue [21]. It is worth noticing that in NOESY spectra of pep3 as well as pep2 (Fig. 3a), both in the presence and absence of heparin, we failed to detect significant inter-residue NOEs even with mixing times between 150 and 600 ms. This suggests that the peptide is overall flexible and adopts a relaxed conformation as *e.g.* found in poly-proline helices (PPII) or random coils [22]. After full assignment of the peptide, <sup>1</sup>H-<sup>15</sup>N HSQC and titration experiments were used to determine chemical shift perturbations (CSPs) upon binding of <sup>15</sup>N-labeled pep2 to heparin (Fig. 3). Though pep2 differs slightly in sequence as compared to pep3, we were able to transfer the assignment done for pep3 to pep2 (Fig. 3a). We verified assignment transferability by 3D TOCSY-<sup>15</sup>N HSQC and 3D NOESY-<sup>15</sup>N HSQC experiments. The results of the <sup>1</sup>H-<sup>15</sup>N HSQC titration experiment using pep2 is presented in Fig. 3. Here, the <sup>1</sup>H-<sup>15</sup>N HSQC spectra recorded in the presence of different concentrations of heparin are shown as an overlay. We observe most of the residues to exhibit CSPs upon addition of heparin, indicating that the entire peptide contributes to binding to heparin, either by direct interactions or conformational rearrangement. The CSPs are concentration-dependent and saturable, with maximal chemical shift changes at a heparin/pep2 molar ratio of approx. 1/6 (Fig. 3a). While CSPs were observed for most of the residues, they differ in overall chemical shift change. Chemical shift changes after addition of 1.6 mg of heparin are shown in Fig. 3b. Chemical shift perturbations upon titration of heparin are observed throughout the peptide with the most severe changes being observed for the first 24 residues of the peptide (Fig. 3b). Interestingly, the highest number of Pro residues can be found within this region of the peptide, which causes an increased rigidity of the peptide backbone due to their fixed  $\Phi$ -angle. We hypothesize that despite the added rigidity of the backbone, the remaining residues undergo conformational changes due to re-orientation of the side-chains of certain residues towards or away from the ligand during binding. The C-terminal region of the peptide (residue 25-37) mostly exhibits minor chemical shift perturbations indicating that these residues are not directly involved in the interaction and that conformational changes in this region might be of a lesser importance to the interaction (Fig. 3a). This coincides with a low number of Pro residues which might impose less structural constraint. The most striking chemical shift change was observed for residue E23 which shows a maximal chemical shift change of approx.

0.8 ppm (Fig. 3a, b). This suggests a severe change in the chemical environment of this particular residue upon heparin binding. Interestingly, also the neighboring residues, L22 and R24, show relatively large changes in chemical shifts. The result is in concordance with the mutation studies, where we observed a significant decrease in heparin binding for R24A which supports our hypothesis that these residues are of particular importance to the interaction between YadA<sub>0:9</sub> to heparin (Fig. 3d). Residue R21 on the other hand, while showing significantly reduced binding of heparin upon mutation to alanine, does not show much CSP in our titration experiments (Fig. 3a, b). The  $K_D$  between Yad<sub>0:9</sub> N-terminal peptide (pep2) and heparin was calculated to approx. 30  $\mu$ M based on the chemical shift perturbations obtained in the titration experiment (Fig. 3c) [23]. This  $K_D$  is approx. 1.000-fold higher than the  $K_D$  determined prior for the YadA head domains (Fig. 2b). One explanation might be that within the framework of the YadA<sub>0:9</sub> head domain, the proper orientation of the peptide is facilitated, or its structure somewhat more stabilized. The NMR experiments point towards an approximate molar binding ratio of 6:1, meaning that six N-terminal peptides bind to one molecule of heparin. The different binding stoichiometry as determined by NMR compared to the plate-based binding experiments (Fig. 2c) could be explained by steric hindrance of the peptide binding to heparin in the context of the YadA head domain. Overall, our experiments provided a full assignment of the YadA<sub>0:9</sub> N-terminal peptide as well as information on potential regions of interest for binding to heparin.

To further substantiate our NMR-based binding data, we used pep3 (Fig. 3a) and heparin in SAXS experiments. Fig. 3d depicts the small-angle X-ray scattering data of various mixtures of heparin and the peptide. As a reference for the hypothetical case of no binding or interaction, the calculated average intensity of the two components with the same concentration as in the mixture is plotted. A significant change in the scattering curves can be observed for all mixtures. At low Q, which is mostly sensitive to the large scales of the complex, an increase in the overall intensity is observed indicating an increase in the molecular weight upon peptide binding to heparin. The structure factor peak furthermore is observed resulting from initial strong repulsions, progressively disappear with the amount of added peptide indicating charge neutralization and electrostatic screening. This is well in line with the fact that heparin is strongly negatively charged and individual heparin chains might repulse each other. Peptide binding leads to neutralization of the charged groups exposed on heparin. In order to obtain more quantitative information, the data were analyzed using a detailed scattering model. In short, assuming the peptide binds to heparin, we modelled the complex as a “superchain” described as a semi-flexible polymer chain with  $N_p$  peptides attached. In this way we additionally obtain the overall chain size, radius of gyration,  $R_g$ , and the molecular weight of the complexes. First, fitting the individual peptides and heparin chains and keeping the individual molecular weight fixed, we analyzed the complexes to obtain  $N_p$  to be about 11, 14 and 10 for the 2+ 2 mg/ml, 3.3 + 3.3 mg/ml and 4+ 2 mg/ml peptide/heparin solutions respectively.  $R_g$  increased from 25, 32 and 37 Å in the same series. Curiously, the amount of bound peptide seems to apparently go down for the highest peptide/heparin ratio. At the same time the overall size of the complex increases from about 25, 32 and 37 Å. However, this can be explained by the presence of free peptides at large ratios which is not resolved by SAXS without the use of fractionation techniques. The SAXS experiments thus further confirm the interaction between the YadA<sub>0:9</sub> N-terminal peptide. They are well in line with the NMR results, suggesting 10-14 bound peptide molecules per heparin chain which fits the binding ratios determined via NMR.

The N-terminal region of the YadA<sub>0:9</sub> head domain exhibits PPII helix characteristics.

We wondered whether the YadA<sub>0:9</sub> heparin-binding peptide might be structured, or potentially adopt a more rigid structure upon heparin binding. In the crystal structure of the head domain of YadA<sub>0:3</sub> the N-terminal region (Y51-G62) was not resolved properly due to the flexibility and/or disordered nature of this region [11]. As structural methods as *e.g.* X-ray crystallography tend to fail identifying PPII



Accepted Article

helices due their increased flexibility, we decided to employ solution-state NMR for this. The fact that few NOEs were observable in the NMR spectra is congruent with both, a flexible, random coil conformation as well as a PPII helix [22]. The H $\alpha$  downfield shifts of residues preceding Pro residues have also been observed for the PPII helix of titin indicating that the peptides have similar structures [21]. We thus aimed to investigate whether the YadA<sub>0-9</sub> N-terminal peptide could, at least partially, adopt a PPII helix conformation. As PPII helices are difficult to identify unambiguously, we used a combination of CD (circular dichroism) spectroscopy, NMR, and *in silico* modelling to investigate our hypothesis. PPII helices, in CD spectra, have a minimum at around 200 nm and a weak positive signal between 220-228 nm [24]. A disordered peptide would have a minimum at approx. 200 nm but no other features in the CD spectrum, and an  $\alpha$ -helical peptide should display a maximum at 190-195 nm and two minima at about 208 nm and 222 nm. As can be seen from the CD spectrum of the pure peptide (Fig. 4a), a minimum at about 200 nm is observed, while no other distinguishing features are visible. However, a weak positive peak may be unobservable in a peptide that does not purely form PPII helices, and it cannot be excluded that there is a tendency for the N-terminal part to form an PPII helix. When the sample is mixed with heparin, a clear red-shift can be observed, that is significantly different from the sum of the individual spectra of the peptide and heparin alone, indicating a change in structure upon ligand binding (Fig. 4a). Overall, the CD data strongly suggests that a distinct secondary structure is adopted upon addition of heparin to the sample.

To further investigate the hypothesis that the peptide might form a secondary structure different from a random coil, we analyzed NMR-based secondary structure propensities calculated from backbone chemical shifts. Backbone  $\phi/\psi$  torsion angles were predicted based on the chemical shifts from the NMR experiments using TALOS (Fig. 4b) [25]. The backbone torsion angles predicted for the N-terminal region of the peptide (residue 1-18) as well as the C-terminal region (residue 32-40) correspond to the dihedral angles reported for PPII helices ( $\phi = -75^\circ$ ,  $\psi = 150^\circ$ ) [26]. Next,  $\delta 2D$  population analysis was employed to predict populations based on the H $\alpha$ , HN, C', C $\alpha$ , C $\beta$  and N chemical shifts [27]. The total prediction population of PPII helix conformation is 24.2%, and some regions are close to 30.0 % as determined by  $\delta 2D$  (Fig. 4c). PPII helices appear to be the most frequent class of secondary structure in the peptide. Based on dihedral angles derived from Talos-N predictions, an ensemble of 1.000 conformers of pep3 was then generated using CNS [28]. Frequencies of observed secondary structures in the ensemble were calculated with PROSS and reported for each residue position in the peptide (Fig. 4d) [29]. Here, a high frequency for a PPII helix for the N-terminal region of the peptide (residue 1-18) was found. Residues 21-31 either involve a turn or a helical conformation. The high frequency for PPII helices further supports that the Pro-rich portion of the N-terminal peptide might form a PPII helix, that is potentially adopted upon heparin binding. In a next step, we attempted to model the N-terminal peptide secondary structure. While the overall structural similarity of the conformer ensemble was low with an RMSD of  $7.6 \pm 2.1$  Å among the 20 lowest-energy conformers, four well-ordered regions could be identified using CYRANGE [30]. Each well-ordered region was visually analyzed after superimposition of the 20 lowest-energy conformers on the backbone atoms of the corresponding regions (Fig. 5a).

The first well-converged region encompasses residues 0 to 11, exhibiting a canonical PPII helix conformation (Fig. 5a). Interestingly, in this model, the two basic residues, R3 and R7, face in opposite directions suggesting that only one residue would be able to contact the ligand. This observation is well in line with the mutational studies where we observed a significant decrease in heparin binding for YadA<sub>0-9</sub> R7A but not for R3A (Fig. 2d). The second well-converging region encompasses residue 13-18 also showing a PPII helix structure (Fig. 5a). Here, two Pro residues follow each other directly, potentially causing increased rigidity. The conformation for the region between residues 19 and 28

could explain why in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC binding study (Fig. 3a, b), we observe the largest chemical shift change for residue E23. The model shows a turn placing residue E23 and R26 in proximity to each other. Facing the other site R21 and R24 are also placed in proximity (Fig. 5a). Such a structural arrangement of the residues is well in line with the results of the of the mutational studies (Fig. 2d) where mutating R21 and R24 to alanine led to a significant decrease in heparin binding while mutating residue R26 had no significant impact on binding. Finally, the region from residue 29 to 38 resemble a PPII helix despite having only one Pro residue (Fig. 5a). Maybe counterintuitively, Pro residues are not integral to the formation of PPII helices [31]. While a model for the full peptide did not converge, our results provide evidence for a more ordered conformation of the peptide. To further consolidate our findings, we furthermore attempted to predict the structure of the entire  $\text{YadA}_{\text{O:9}}$  head domain including the heparin-binding motif using AlphaFold2 (Fig. 5b) [32]. Even though AlphaFold 2 did predict a PPII helix for the region between residue L12 and K17, the pLDDT scores for the region are low to very low (Fig. 5b). Interestingly, also in this model, all basic residues, except for residue R26, are projected outwards, supporting the data of our mutational study (Fig. 2d). We can at this point not exclude that within the context of the trimeric  $\text{YadA}_{\text{O:9}}$  head domain, a secondary structure of the N-terminal peptide might be stabilized through interactions with the short  $\beta$ -strands or the hydrophobic interior of the  $\beta$ -roll head domain of  $\text{YadA}_{\text{O:9}}$ . Furthermore, the individual N-terminal regions might stabilize each other within the context of the trimeric head domain. While our structural work as well as the AlphaFold2 model aid in interpreting the  $^1\text{H}$ - $^{15}\text{N}$  HSQC peptide-heparin binding data and our mutational binding studies, note that they show possible, local secondary structures that might exist only transiently in the cellular context.

$\text{YadA}_{\text{O:9}}$ -mediated adhesion is crucial for host-cell interaction

To elucidate whether binding of negatively charged heparin and HS via  $\text{YadA}_{\text{O:9}}$  is directly relevant for binding *in vivo*, we investigated whether defects in host cell glycosylation negatively impact on  $\text{YadA}_{\text{O:9}}$ -mediated binding to host cells. To this end,  $\text{YadA}_{\text{O:9}}$  adhesion assays using HEK293T wt cells and different COG (Conserved Oligomeric Golgi) mutants were performed. The HEK293T mutant cell lines each lack one subunit of the COG (COG1-8/-) leading to alterations in protein sialylation and fucosylation as well as aberrations in polymerization and turnover of GAGs [33] [34]. We used HEK293T wt cells as well as HEK293T COG1/- and COG3/- cells. While COG1/- cells exhibit reduced HS synthesis, the GAG chain length is increased resulting in comparable numbers of binding sites at the cell surface compared to wt cells [35]. HEK293T COG3/- cells are completely deficient in cell surface-exposed HS.

$\text{YadA}_{\text{O:9}}$ -mediated adhesion of *E. coli* AS75 was tested for all three HEK293T cell types. Adhesion was unchanged between HEK293T wt and HEK293T COG1/- cells, while adhesion to HEK293T COG3/- cells was significantly reduced (Fig. 6). This indicates that  $\text{YadA}_{\text{O:9}}$ -mediated adhesion is reduced in cells severely defective in the synthesis of GAGs, displaying little HS at the cell surface. Defects in glycosylation of glycoproteins and wild-type-like levels of HS at the cell surface as in HEK293T COG1/- cells on the other hand did not affect adhesion, suggesting that HS is crucial for *Y. enterocolitica* serotype O:9 binding and potentially, virulence.

Discussion



Accepted Article

PPII helices are a common structural motif in proteins. Due to their solvent-exposed nature, they are important in ligand interactions [31]. Despite being well defined through their backbone dihedral angles ( $\phi = -75^\circ$ ,  $\psi = 150^\circ$ ), discriminating PPII helices from random coils experimentally is still difficult [22]. Structural biology techniques as for example X-ray crystallography or cryo-EM often fail to identify PPII helices due to their flexible nature. Furthermore, despite the recent progress in the development of neural network approaches for structure prediction like AlphaFold2 [32], [36], prediction of PPII helices remains challenging with oftentimes low confidence scores for regions with potential PPII helical content [37], [38]. Thus, NMR remains one of the few methods suitable to determine the presence of PPII helices even though there are few inter-residue NOEs, making structure determination difficult [26]. We applied a series of complementary techniques and determined that the N-terminal motif within YadA<sub>O:9</sub> exhibits physical characteristics in agreement with a PPII-helix. Our data show that the PPII helix interacts with heparin through electrostatic interactions between basic residue within the YadA<sub>O:9</sub> peptide and the negatively charged sulphate groups of heparin. Finally, we were able to show that the short motif within YadA<sub>O:9</sub> is crucial for binding to host cells, which might have important consequences for *Y. enterocolitica* serotype O:9 virulence. We speculate that this might also cause serotype-specific differences in the course of an infection but were unable to find robust patient data in the literature to support this hypothesis. A key question in this context is whether heparin actually is the primary binding target of YadA<sub>O:9</sub> *in vivo*. Heparin is stored in secretory granules of Mast cells and thus not directly accessible for extracellular pathogens [39]. It is probable that other sulphated GAGs in the ECM like HS are the *in vivo* target of YadA<sub>O:9</sub>. As heparan sulphate and heparin are chemically almost identical, heparin constitutes a sufficient model for investigation of these types of interactions between YadA<sub>O:9</sub> and sulphated GAGs [12]. In terms of physiological relevance of the interaction described here, we speculate that it might serve various functions. YadA<sub>O:9</sub> interacting with surface-exposed sulphated glycans might aid in host interaction and invasion [5], [40], [41]. Furthermore, binding of heparin to the bacterial cell surface might be beneficial in pathogen immune evasion in the bloodstream [42]–[44]. Heparin is an inhibitor of the complement cascade [7], [45], [46]. This leads to local increase of complement-inhibitory molecules and could aid in complement evasion [4], [45], [47]; alternatively, it might simply decorate the bacteria in host molecules, making them more difficult to detect for immune cells. Thus, binding to sulphated glycans might serve a double-purpose in tissue adhesion and immune evasion. Further research will show whether this is indeed the case in the course of an infection with *Y. enterocolitica*.

In summary, we showed how a short, structured insertion in the YadA<sub>O:9</sub> N-terminal region aids in adhesion to heparin and thus completely alters adhesion properties compared to YadA from other *Y. enterocolitica* serotypes. We hence exemplify how one simple genetic event can alter virulence-specific traits of a bacterial pathogen – a process that is typically more associated with viruses [48].

## Materials and Methods

### Cloning

Mutagenesis for introduction of the point mutations into the N-terminal region of the YadA<sub>O:9</sub> head domain as well as for full-length constructs was performed using overlapping primers harboring the point mutations (Tab. 1). In short, a PCR reaction using overlapping mutagenesis primers was performed using Q5 polymerase to insert the mutations in both the YadA<sub>O:9</sub> head domain construct pASK-Iba2-YadA<sub>O:9</sub> as well as into the YadA<sub>O:9</sub> full length construct pASK-Iba4c-YadA<sub>O:9</sub>-FL [7], [8]. After the PCR reaction, methylated DNA was digested with DpnI. 50 ng of circular PCR product were transformed into *E. coli* Top10 cells. A Sumo-TEV-YadA<sub>O:9</sub> heparin binding peptide construct for

Accepted Article

purification of labelled YadA<sub>0:9</sub> N-terminal peptide (Fig. 3a, pep2) was cloned in multiple steps. First, an existing pASK-Iba2-Sumo construct (data unpublished) was linearized by PCR using “linearization primers” (Sumo linear, Tab.1). The sequence coding for the YadA<sub>0:9</sub> N-terminal peptide was amplified *via* PCR from the YadA<sub>0:9</sub>-head domain construct (“N-terminal peptide insertion primers”, Tab. 1) [8]. Gibson cloning was used for insertion of the YadA<sub>0:9</sub> N-terminal peptide into the Sumo construct [49]. Insertion primers were designed to directly insert a TEV cleavage site into the pASK-Iba2-Sumo N-terminal region construct (“TEV insertion primers”, Tab. 1).

Insertion of sequence coding for the YadA<sub>0:9</sub> N-terminal peptide into sfGFP was done using Gibson assembly. The construct was designed based on [50]. For cloning a pASK-Iba2-sfGFP construct was linearized by PCR (“sfGFP open” Tab. 1). The sequence encoding the N-terminal peptide was amplified from (“N-terminal region insertion sfGFP”, Tab. 1). Assembly of the N-terminal peptide into the linearized sfGFP construct was done using Gibson assembly [49].

### Proteins and peptides used in this study

In this study several different proteins and peptides derived from the YadA<sub>0:9</sub> head domain were used (Fig. 1b, 7). Well-plate based binding assays were performed with purified YadA<sub>0:9</sub> head domains entailing the native N-terminal peptide (pep1) (residue 29-65 in YadA<sub>0:9</sub>) [8]. Initial peak assignments and NMR experiments were performed using an unlabeled, synthesized peptide (pep3). <sup>1</sup>H-<sup>15</sup>N HSQC-based binding studies were done with a <sup>15</sup>N labelled peptide that was produced as a fusion peptide with Sumo and cleaved off using TEV protease (pep2). To unify the residue numbering for all peptides, the numbering is based on the native peptide sequence, pep1 as found in YadA<sub>0:9</sub> as shown in Fig. 3a.

### Synthesized, unlabeled peptide for initial NMR assignment

For NMR experiments and resonance assignments pep3 was used (Fig. 2). This peptide was purchased from Genscript with a purity of ≥ 98.9% as determined by HPLC. For NMR experiments the peptide was resuspended in NMR buffer (20 mM Phosphate pH 6.0, 50 mM NaCl, 5% (v/v) D<sub>2</sub>O and 0.2 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS, Larodan).

### Production and purification of labelled peptide for NMR experiments

For production of <sup>15</sup>N-labeled pep2 an over-night culture was prepared inoculating a single colony of *E. coli* BL21 (DE3) Gold harboring the Sumo-pep2 construct into 50 mL <sup>15</sup>N-labeled M9 minimal medium supplemented with 0.5 g/L <sup>15</sup>NH<sub>4</sub>Cl and Ampicillin at a final concentration of 100 µg/mL. The next day a subculture was prepared in 4 L <sup>15</sup>N-labeled M9 medium and grown at 37°C to an OD<sub>600</sub> of approx. 0.5. Protein production was induced by addition of AHTC (anhydrotetracycline hydrochloride) to a final concentration of 200 ng/mL. Expression was performed at 30 °C for 4 h. Cells were harvested by centrifugation at 4.500 x g for 45 min at 12°C. The pellet was resuspended in 25 mL TBS buffer (20 mM Tris-HCl pH 7.2, 300 mM NaCl) supplemented with DNase, Lysozyme and MgCl<sub>2</sub> at final concentrations of 10 µg/mL, 8 µg/mL and 1 mM, respectively. 100 x HALT protease inhibitor cocktail (Thermo Scientific) was added at a 1:100 dilution. The suspension was passed through the French press 3 times at 1000 psi. The sample was centrifuged at 24.000 g for 1h at 4°C and the supernatant was filtered through a 0.2 µm syringe filter. The sample was then subjected to Ni-NTA affinity chromatography. Fractions containing Sumo-pep2 were collected and concentrated to a final volume of 1 mL using a 10 kDa cutoff Vivaspin concentrator (Merck Millipore). Subsequently, the protein was subjected to size exclusion chromatography using a HiPrep™ 26/60 Sephacryl® S200 HR size exclusion column (GE) equilibrated with TBS buffer (20 mM Tris pH 7.5, 100 mM NaCl). Fractions containing Sumo-pep2 were combined and TEV protease was added to a final concentration of 20 µg/mL. Cleavage was performed for 16 h at 30 °C. Then, another 20 µg/mL TEV protease were added, and the reaction was continued for another

24h. After cleavage the salt concentration was increased to 300 mM NaCl, and the sample was subjected to Ni-NTA affinity chromatography. Fractions containing pep2 were collected, combined, and lyophilized. C<sub>18</sub> spin columns (ThermoFisher Scientific; Catalogue number: 89870) for desalting. Acetonitrile was removed by heating the sample to 65 °C for 2h. Finally, the sample was lyophilized again to remove TFA. The peptide purity was confirmed by SDS Polyacrylamide gel electrophoresis and Mass spectrometry. For NMR experiments, the lyophilized sample was resuspended in NMR buffer.

#### Protein production and purification

Production and purification of the entire YadA<sub>0-9</sub> head domain (residue 16-260 of YadA<sub>0-9</sub> wt) was performed as described elsewhere [8], [11]. Production and purification of the sfGFP-N-terminal peptide fusion protein was performed as described in [50]. The anion exchange purification step was omitted.

#### Glycan binding assays with purified protein

YadA-glycan binding assays were performed as following: 100 µL of a 10 µg/mL solution of purified YadA wt or mutant head domains were immobilized in a clear, polystyrene 96-well plate (Sarstedt, 82.1581) by incubation at RT for 1h. The wells were washed three times with 200 µL TBS buffer (20mM Tris-HCl pH 7.5, 150 mM NaCl) and blocked using 200 µL of a 3% (w/v) BSA solution in TBS for 1h at RT. The wells were washed as described above and a dilution series of the biotinylated glycan of interest was added; Heparin-Biotin (Merck, B9896), Hyaluronan-Biotin (Merck, B1157), Dextran-Biotin (Merck, B9139), Biotin (Merck, 14400). After incubation for 1 h at RT the wells were washed as described above. For detection of YadA<sub>0-9</sub>-bound glycans, 100 µL of a 1:10.000 Streptactin-HRP (IBA Lifesciences) solution in 3% (w/v) BSA in TBS were incubated for 30 min at RT. The wells were washed as described above. Detection was performed using ABTS (2,2'-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) diammonium salt). 150 µL of a 1 mg/mL ABTS solution were added and after a development time of approx. 30 min, the color change was stopped by addition of 100 µL 1% (w/v) SDS solution. Absorbance at 405 nm was measured for quantification of bound glycans.

#### Heparin binding studies with bacteria expressing full-length YadA

Binding of biotinylated heparin to bacteria expressing full length YadA<sub>0-9</sub> harboring the point mutations was monitored using *E. coli* AS75 strain genomically carrying sfGFP under an arabinose inducible promoter [51]. Bacteria were transformed with pASK-lba4c\_YadA\_FL. A colony was inoculated in 5 mL LB medium supplemented with chloramphenicol to a final concentration of 25 µg/mL and 0.01% (w/v) arabinose. The following day, a 25 mL subculture was prepared in the same medium and incubated at 37°C. At an OD<sub>600</sub> of 0.5, production of YadA was induced by adding AHTC to a final concentration of 200 ng/mL. Protein production was carried out for 3 h. Clumping (autoagglutination) of the bacterial cells indicated expression of functional YadA [52], [53]. The cultures were harvested by centrifugation at 4.500 x g and resuspended in PBS to an OD<sub>600</sub> of 0.2. 100 µL of the bacterial dilutions were added to the wells of a clear, polystyrene 96-well plate. The plate was centrifuged at 3.500 x g for 10 min. After removal of the supernatant, 100 µL of a 10 µg/mL heparin-biotin solution in 3% BSA were incubated for 30 min at RT. The plate was centrifuged at 3.500 x g for 10 min and the supernatant was discarded. The wells were washed three times with PBS. Between every wash the plate was centrifuged for 5 min at 3.500 g. Finally, a 1:10.000 Streptactin-HRP dilution in 3% (w/v) BSA in PBS solution was added and incubated for 30 min. The plate was centrifuged as prior. The plate was washed with PBS as described before. To normalize the number of bacterial cells expressing YadA<sub>0-9</sub>, fluorescence of the bacterial cells was measured at 488 nm absorption and 533 nm emission wavelength. After that, 150 µL of a 1 mg/mL ABTS solution was added. The plate was incubated for 30 min and the color development of the ABTS was stopped by addition of 100 µL of a 1% w/v SDS solution. Absorption at 405 nm was measured in a plate reader to assess the amount of heparin bound to the YadA<sub>0-9</sub>-expressing bacterial cells.

NMR experiments for assignment of the peptide

Initial peak assignments of the YadA<sub>0-9</sub> N-terminal peptide was performed using pep3. The peptide was resuspended in NMR buffer at a concentration of 1mM. The concentration was determined spectrophotometrically at 280 nm. A full set of NMR spectra was acquired on an 800 MHz Bruker Avance III spectrometer equipped with a 5 mm TCI cryo probe (<sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N) at a temperature of 298.15 K: homonuclear <sup>1</sup>H-<sup>1</sup>H 2D TOCSY (80 ms mixing), NOESY (100, 200, 300, 400, 600 ms mixing), DQCOSY, <sup>1</sup>H-<sup>15</sup>N HSQC, <sup>1</sup>H-<sup>13</sup>C HSQC, and <sup>13</sup>C-<sup>1</sup>H HMBC. In homonuclear spectra, water suppression was obtained using excitation sculpting [54]. The data was processed using Topspin 3.5 pl6 (Bruker Biospin). DSS was used as a chemical shift standard, and <sup>15</sup>N and <sup>13</sup>C data were referenced using frequency ratios [55]. Peak assignment was performed using the program SPARKY [56] and standard methods [57].

NMR experiments for <sup>1</sup>H-<sup>15</sup>N HSQC titration experiments

For titration experiments <sup>15</sup>N-labeled pep2 was used. Spectra were recorded as described above. The peptide was resuspended in NMR buffer at a final concentration of 0.059 mM as determined by BCA assays and comparison of a <sup>1</sup>H proton spectrum with a delay of 60 s between the scans of the 18 methyl groups of branched chain amino acids in the peptide at 0.8 ppm to 1.0 ppm to the methyl group peaks of DSS of known concentration by integrating the area under the peaks. For binding experiments <sup>1</sup>H-<sup>15</sup>N HSQC titration experiments were performed. A 25 mg/mL heparin solution (Merck, B9896-10MG) was prepared in NMR buffer and titrated stepwise into the NMR tube containing the peptide sample. <sup>1</sup>H-<sup>15</sup>N HSQC spectra were obtained at each titration step: 0 µg, 50 µg, 100 µg, 200 µg, 400 µg, 800 µg and 1600 µg. To determine molar binding ratios the average molecular weight of heparin (10 kDa) was presumed.

At the final titration step 3D <sup>15</sup>N TOCSY- HSQC and 3D <sup>15</sup>N NOESY- HSQC spectra were recorded. The <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of the final titration step was assigned using the 3D spectrum, programs and methods described for pep3. Chemical shift perturbations (CSP) were calculated for chemical shifts occurring between the peptide alone and chemical shifts acquired at 1.6 mg of total heparin added. The changes of amide <sup>1</sup>H and <sup>15</sup>N chemical shifts were averaged using the following equation:

$$\Delta\delta_{av} = \left[ \frac{1}{2} [(\Delta\delta_H)^2 + (0.2\Delta\delta_N)^2] \right]^{\frac{1}{2}} \quad \text{e.q. 1}$$

The dissociation constants ( $K_D$ ) were determined from the changes in chemical shifts upon addition of heparin to the YadA<sub>0-9</sub> N-terminal peptide [23]:

$$\Delta\delta_{av} = \Delta\delta_{av,max} [K_D + [L]_0 + [P]_0 - \sqrt{(K_D + [L]_0 + [P]_0)^2 - 4[L]_0[P]_0}] / 2[P]_0 \quad \text{e.q. 2}$$

Here  $[P]_0$  and  $[L]_0$  are the concentrations of the peptide (P) and heparin as a ligand (L).  $\Delta\delta_{av,max}$  is the maximal CSP that can be obtained upon addition of a ligand to the peptide.  $K_D$  and  $\Delta\delta_{av,max}$  were free parameters during the fit of the experimental data.

Dot blots

Dot blots were performed as described previously [8]. In short, 1.4 µg of purified YadA head domain was immobilized on a nitrocellulose membrane. The membrane was blocked with 5% BSA in TBS-T (20 mM Tris-HCl pH 7.4, 100 mM NaCl, 0.05% (v/v) Tween 20) for 1h at RT. The blot was then incubated with 500 µL of a 100 µM solution of biotinylated glycans of interest in TBS-T buffer for 1h at RT. The blot was then washed three times with TBS-T. Lastly the blot was incubated with a 1:10.000 dilution of Streptactin-HRP conjugate in 5% BSA in TBS-T for 30 min at RT. After washing the blot three times with TBS-T buffer and once with TBS buffer, signals were recorded using ECL reagent. Chemiluminescence was observed using a Kodak Image Station 4000R.

## Circular Dichroism spectroscopy

Samples of pep3 at a final concentration of 20  $\mu\text{M}$  in MilliQ water and a mix of 1:1 pep3/heparin have been obtained. CD spectra were recorded using a Jasco J-810 spectropolarimeter (Jasco International Co) calibrated with D-camphor-10-sulfonate (Aldrich). All measurements were done using a quartz cuvette (Starna) with 0.1 cm path length. Samples were scanned five times with a scanning rate of 50nm/min with a bandwidth of 1nm, a response time of 2 s, and sampling every 0.5 nm over the wavelength range 180–260 nm. For each of the obtained spectra the background of water was subtracted.

## Chemical shift analysis

Predictions for secondary structure populations ( $\beta$ -strand,  $\alpha$ -helix, PPII and coils) were obtained from  $^{15}\text{N}$ ,  $^{13}\text{C}$  and  $^1\text{H}$  backbone chemical shifts using  $\delta 2\text{D}$  software [27]. Backbone dihedral angles were predicted with TALOS-N [25] from HN, H $\alpha$ , C $\alpha$ , C $\beta$ , C' and N chemical shifts. Predictions classified as “Bad” were not considered for further modelling analysis.

## Modelling of local peptide secondary structure

Predicted backbone dihedral angles from TALOS-N were converted into  $\varphi/\psi$  dihedral angle restraints, using errors corresponding to twice the TALOS standard deviation (sd) for “Strong” predictions and 3\*sd for “Dynamical” or “Generous” predictions. Using simulated annealing protocols implemented in ARIA, 1000 conformers for pep3 were generated with CNS [28], [58]. Using CYRANGE well converging regions were found with backbone root mean square deviations (RMSDs) of around 0.83-1.04  $\text{\AA}$  [30]. All 1000 conformers were analyzed with PROSS for statistical analysis of the frequency of secondary structure elements at each amino acid position [29].

## Alphafold2 modelling of trimeric YadAO:9 head domain

The YadA<sub>0:9</sub> head domain (residue 1-487, excluding the signal peptide) structure was predicted using Alphafold2 [32] Multimer via the Colabfold Multimer pipeline [36] with default parameters, 20 recycles and final amber relaxation. The best ranked model was used for analysis and figure preparation. pLDDT scores were plotted against the residue number using the generated .json file extracted using Alphapickle (<https://github.com/mattarnoldbio/alphapickle/tree/v1.4.0>).

## SAXS data modelling

To describe the SAXS data from single peptides in solution, the following expression was used:

$$I(Q) = \varphi \cdot M_w/d_p \cdot (\rho_p - \rho_0)^2 \cdot P(Q)_{peptide} \quad (1)$$

$\varphi$  is the volume fraction of the peptide,  $M_w$  is the molecular weight,  $d_p$  is the solution density of the peptide,  $\rho_p$  and  $\rho_0$  are the scattering length densities of the peptide and buffer, respectively. For the form factor,  $P(Q)_{peptide}$ , either a random chain form factor,  $P(Q)_{chain}$  or cylindrical form factor,  $P(Q)_{cylinder}$ , describing alpha-helical structures, were assumed. These are given by:

$$P(Q)_{chain} = \frac{2 \cdot \exp[-(QR_g)^2] - 1 + (QR_g)^2}{(QR_g)^4} \quad (2)$$

where  $R_g$  is the radius gyration of the peptide chain. The cylindrical form factor,  $P(Q)_{cylinder}$ , is given by:



$$P(Q)_{cyl} = \int_0^{\pi/2} |A(Q, \alpha)_{cyl}|^2 \sin \alpha d\alpha, \quad (3)$$

$$\text{where } A(Q, \alpha)_{cyl} = \frac{2J_1(QR \sin \alpha) \sin(QL \cos \alpha)}{QR \sin \alpha \quad QL \cos \alpha}. \quad (4)$$

$R$  and  $L$  are the radius and length of the cylinder,  $\alpha$  is its angle to the scattering vector  $Q$ , and  $J_1(x)$  is the first order Bessel function. The integral over alpha averages the form factor over all possible orientations of the cylinder with respect to  $Q$ .

To describe the scattering of the heparin solution, the electrostatic interactions as well as the rigid nature of the polysaccharides must be taken into account. To take repulsions into account, the well-known Percus-Yevick structure factor,  $S(Q, \eta, R_{HS})$ , for hard spheres is included. This potential is determined by the effective hard sphere volume fraction,  $\eta$  and radius,  $R_{HS}$ . Moreover, we use a rather crude approximation where we assume a decoupling approximation where the scattering amplitude,  $A(Q)_{chain}$  can be approximated with that of a Gaussian chain [59]:

$$I(Q) = \varphi \cdot M_w/d_h \cdot (\rho_h - \rho_0)^2 \cdot P(Q, L_c, l_K, R)_{WLC} \cdot \left(1 + \frac{A(Q)_{chain}^2}{P(Q)_{WLC}} (S(Q, \eta, R_{HS}) - 1)\right) \quad (5)$$

Where  $d_p$  is the solution density of heparin,  $\rho_h$  the corresponding scattering length density.

$$A(Q)_{chain} = \frac{1 - \exp[-(QR_g)^2]}{(QR_g)^2} \quad (6)$$

$P(Q, L_c, l_K, R)_{WLC}$  is the form factor for a worm-like chain with excluded volume effects described by the Kuhn length,  $l_K$ , contour length,  $L_c$  and the cross section of the chain [60]. For the mixtures, we assumed peptide-heparin complex into "super-chain" which is modelled in the same fashion as a generalized polymer chain which consist of heparin with  $N_p$  peptide chains attached.

$$I(Q)_{complexes} = \varphi \cdot \left( N_p \cdot M_w(peptide)/d_p + M_w(heparin)/d_h \right) \cdot (\rho_{complex} - \rho_0)^2 \cdot P(Q, L_c, l_K, R)_{WLC} \cdot \left(1 + \frac{A(Q)_{chain}^2}{P(Q)_{WLC}} (S(Q, \eta, R_{HS}) - 1)\right) \quad (7)$$

where the scattering length density of the complex is given by:

$$\rho_{complex} = \frac{\frac{N_p \cdot M_w(peptide)}{d_p} \cdot \rho_p + \frac{M_w(heparin)}{d_h} \cdot \rho_h}{\frac{N_p \cdot M_w(peptide)}{d_p} + \frac{M_w(heparin)}{d_h}} \quad (8)$$



## Bacterial binding assay to HEK cells and HEK $\Delta$ COG1 and $\Delta$ COG3 cells

HEK293T wt and mutant cells were grown to 80% confluency in DMEM/F12 medium (Gibco, Life Technologies) supplemented with 10% fetal calf serum (HyClone, GE Healthcare Life Sciences), and 1% penicillin/streptomycin (PS; Lonza) in 24-well polystyrene plates at 37°C with 5% CO<sub>2</sub>. *E. coli* AS75 cells carrying full-length YadA<sub>0:9</sub> encoded on pASK-lba4c were grown in 25 mL LB from an o/n culture in the presence of a final concentration of 50 µg/mL chloramphenicol and 0.02 % w/v Arabinose to an OD<sub>600</sub> of 0.5. Production of full-length YadA<sub>0:9</sub> was induced by adding AHTC to a final concentration of 200 ng/mL. YadA<sub>0:9</sub> production was allowed for 3 h at 37°C. The cells were then harvested, and the bacterial cell density was normalized to an OD<sub>600</sub> of 1.0 in PBS. The bacterial cells were then diluted 1:20 in 1 mL DMEM medium. HEK293T wt and mutant cells were carefully washed three times with 1 mL PBS and 1 mL DMEM without antibiotics was added. The cells were incubated at 37°C and 5 % CO<sub>2</sub> for 30 min to recover from washing. After recovery, 50 µL of the bacterial solution in DMEM were added to each well and incubated for 30 min at 37°C and 5 % CO<sub>2</sub>. The wells were washed carefully three times with 1 mL PBS and fluorescence of adhering bacteria was measured at an excitation wavelength of 488 nm and an absorption wavelength of 533 nm in a BioTek Synergy H plate reader. The background adhesion of *E. coli* AS75 not expressing YadA<sub>0:9</sub> was subtracted from the fluorescence values of *E. coli* AS75 expressing YadA<sub>0:9</sub>.

## Statistical analysis

Binding data are shown as means  $\pm$  SD. The data were analyzed using one-way ANOVA including Tukey's test. For data plotting and statistical analysis, OriginPro was used. Significance levels are indicated in the graphs with p-values: p < 0.05 (\*), p < 0.01 (\*\*), or p < 0.001 (\*\*\*).

## Author Contributions

DL and IM conceived the project. IM, PEK, and VRK performed the experiments. IM, PEK, BB, VRK, RL and DL analyzed the data. IM, PEK, BB, VRK, DH, KP, RL, NIP and DL wrote and revised the manuscript.

## Acknowledgements

The authors thank Annika Teresa Lehmann for support in protein production during her Erasmus+ internship. We furthermore thank Régine Dazzoni, Theis Jacobsen, Maylis Lejeune, and Sébastien Cardon at the Institute Pasteur and Athanasios Saragliadis at the University of Oslo for useful discussions and technical support. Finally, we thank the NMR facility as well as the Mass Spectrometry facility at the University of Oslo for access to excellent equipment, and Bernd Thiede for mass spectrometry analysis of the peptides. We are thankful for the funding received in form of a Horizon 2020 Innovative Training Network "ViBrANT" (grant ID: 765042) to D.L. and N.I.P and a FEMS mobility grant (grant-ID: FEMS-GO-2019- 608) to I.M.

## Data Availability statement

The data that support the findings of this study are available from the corresponding author (Dirk.Linke@ibv.uio.no) upon reasonable request.

## References

- [1] D. Linke, T. Riess, I. B. Autenrieth, A. Lupas, and V. A. J. Kempf, "Trimeric autotransporter adhesins: variable structure, common function," *Trends Microbiol*, vol. 14, no. 6, pp. 264–270, 2006, doi: 10.1016/j.tim.2006.04.005.

- [2] J. C. Leo, I. Grin, and D. Linke, "Type V secretion: Mechanism(S) of autotransport through the bacterial outer membrane," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, pp. 1088–1101, 2012, doi: 10.1098/rstb.2011.0208.
- [3] M. Mühlenkamp, P. Oberhettinger, J. C. Leo, D. Linke, and M. S. Schütz, "Yersinia adhesin A (YadA)--beauty & beast.," *Int J Med Microbiol*, vol. 305, no. 2, pp. 252–8, Feb. 2015, doi: 10.1016/j.ijmm.2014.12.008.
- [4] I. Meuskens, A. Saragliadis, J. C. Leo, and D. Linke, "Type V Secretion Systems: An Overview of Passenger Domain Functions," *Front Microbiol*, vol. 10, no. May, pp. 1–19, 2019, doi: 10.3389/fmicb.2019.01163.
- [5] J. C. Leo *et al.*, "First analysis of a bacterial collagen-binding protein with collagen toolkits: Promiscuous binding of YadA to collagens may explain how YadA interferes with host processes," *Infect Immun*, vol. 78, no. 7, pp. 3226–3236, 2010, doi: 10.1128/IAI.01057-09.
- [6] B. Westerlund and T. K. Korhonen, "Bacterial proteins binding to the mammalian extracellular matrix," *Mol Microbiol*, vol. 9, no. 4, pp. 687–694, 1993, doi: 10.1111/j.1365-2958.1993.tb01729.x.
- [7] M. C. Mühlenkamp *et al.*, "Vitronectin binds to a specific stretch within the head region of yersinia adhesin a and thereby modulates yersinia enterocolitica host interaction," *J Innate Immun*, vol. 9, no. 1, pp. 33–51, 2017, doi: 10.1159/000449200.
- [8] I. Meuskens, J. Leva-bueno, P. Millner, M. Schütz, S. A. Peyman, and D. Linke, "The Trimeric Autotransporter Adhesin YadA of Yersinia enterocolitica Serotype O : 9 Binds Glycan Moieties," *Front Microbiol*, vol. 12, no. February, pp. 1–12, 2022, doi: 10.3389/fmicb.2021.738818.
- [9] S. A. Shahid *et al.*, "Solid-state NMR study of the YadA membrane-anchor domain in the bacterial outer membrane," *Angewandte Chemie - International Edition*, vol. 54, no. 43, pp. 12602–12606, 2015, doi: 10.1002/anie.201505506.
- [10] B. H. Alvarez, M. Gruber, A. Ursinus, S. Dunin-Horkawicz, A. N. Lupas, and K. Zeth, "A transition from strong right-handed to canonical left-handed supercoiling in a conserved coiled-coil segment of trimeric autotransporter adhesins," *J Struct Biol*, vol. 170, no. 2, pp. 236–245, 2010, doi: 10.1016/j.jsb.2010.02.009.
- [11] H. Nummelin, M. C. Merckel, J. C. Leo, H. Lankinen, M. Skurnik, and A. Goldman, "The Yersinia adhesin YadA collagen-binding domain structure is a novel left-handed parallel  $\beta$ -roll," *EMBO Journal*, vol. 23, no. 4, pp. 701–711, 2004, doi: 10.1038/sj.emboj.7600100.
- [12] D. L. Rabenstein, "Heparin and heparan sulfate: Structure and function," *Nat Prod Rep*, vol. 19, no. 3, pp. 312–331, 2002, doi: 10.1039/b100916h.
- [13] K. Prydz and K. T. Dalen, "Synthesis and sorting of proteoglycans," *J Cell Sci*, vol. 113, no. 2, pp. 193–205, 2000, doi: 10.1242/jcs.113.2.193.
- [14] J. J. Paxman *et al.*, "Unique structural features of a bacterial autotransporter adhesin suggest mechanisms for interaction with host macromolecules," *Nat Commun*, vol. 10, no. 1, pp. 1–12, 2019, doi: 10.1038/s41467-019-09814-6.
- [15] Z. Shriver, I. Capila, G. Vencataraman, and R. Sasisekharan, "Heparin and Heparan Sulfate: Analyzing Structure and Microheterogeneity," *Handb Exp Pharmacol*, vol. 207, no. 207, pp. 159–176, 2012, doi: 10.1007/978-3-642-23056-1\_3.

- [16] T. Toida *et al.*, "Structural differences and the presence of unsubstituted amino groups in heparan sulphates from different tissues and species," *Biochemical Journal*, vol. 322, no. 2, pp. 499–506, 1997.
- [17] E. Diaz-Montes, "Dextran : Sources , Structures , and Properties," *MDPI polysaccharides*, pp. 554–565, 2021.
- [18] E. Hoiczky, "Structure and sequence analysis of Yersinia YadA and Moraxella UspAs reveal a novel class of adhesins," *EMBO J*, vol. 19, no. 22, pp. 5989–5999, 2000, doi: 10.1093/emboj/19.22.5989.
- [19] N. Chauhan *et al.*, "Insights into the autotransport process of a trimeric autotransporter, Yersinia Adhesin A (YadA)," *Mol Microbiol*, pp. 0–3, 2019, doi: 10.1111/mmi.14195.
- [20] T. Heise and P. Dersch, "Identification of a domain in Yersinia virulence factor YadA that is crucial for extracellular matrix-specific cell adhesion and uptake.," *Proc Natl Acad Sci U S A*, vol. 103, no. 9, pp. 3375–80, 2006, doi: 10.1073/pnas.0507749103.
- [21] K. Ma, L. S. Kan, and K. Wang, "Polyproline II helix is a key structural motif of the elastic PEVK segment of titin," *Biochemistry*, vol. 40, no. 12, pp. 3427–3438, 2001, doi: 10.1021/bi0022792.
- [22] A. A. Adzhubei, M. J. E. Sternberg, and A. A. Makarov, "Polyproline-II helix in proteins: Structure and function," *J Mol Biol*, vol. 425, no. 12, pp. 2100–2132, 2013, doi: 10.1016/j.jmb.2013.03.018.
- [23] L. Fielding, "NMR methods for the determination of protein-ligand dissociation constants," *Prog Nucl Magn Reson Spectrosc*, vol. 51, no. 4, pp. 219–242, 2007, doi: 10.1016/j.pnmrs.2007.04.001.
- [24] J. L. S. Lopes, A. J. Miles, L. Whitmore, and B. A. Wallace, "Distinct circular dichroism spectroscopic signatures of polyproline II and unordered secondary structures: Applications in secondary structure analyses," *Protein Science*, vol. 23, no. 12, pp. 1765–1772, 2014, doi: 10.1002/pro.2558.
- [25] Y. Shen and A. Bax, "Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks," *J Biomol NMR*, vol. 56, no. 3, pp. 227–241, 2013, doi: 10.1007/s10858-013-9741-y.
- [26] T. J. Narwani *et al.*, "Recent advances on polyproline II," *Amino Acids*, vol. 49, no. 4, pp. 705–713, 2017, doi: 10.1007/s00726-017-2385-6.
- [27] C. Camilloni, A. De Simone, W. F. Vranken, and M. Vendruscolo, "Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts," *Biochemistry*, vol. 51, no. 11, pp. 2224–2231, 2012, doi: 10.1021/bi3001825.
- [28] A. T. Brünger *et al.*, "Crystallography & NMR system: A new software suite for macromolecular structure determination," *Acta Crystallogr D Biol Crystallogr*, vol. 54, no. 5, pp. 905–921, 1998, doi: 10.1107/S0907444498003254.
- [29] R. Srinivasan and G. D. Rose, "A physical basis for protein secondary structure," *Proc Natl Acad Sci U S A*, vol. 96, no. 25, pp. 14258–14263, 1999, doi: 10.1073/pnas.96.25.14258.
- [30] D. K. Kirchner and P. Güntert, "Objective identification of residue ranges for the superposition of protein structures," *BMC Bioinformatics*, vol. 12, 2011, doi: 10.1186/1471-2105-12-170.

- [31] A. A. Adzhubei and M. J. E. Sternberg, "Left-handed Polyproline II Helices Commonly Occur in Globular Proteins," *J Mol Biol*, vol. 229, no. 2, pp. 472–493, 2002.
- [32] J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, Aug. 2021, doi: 10.1038/s41586-021-03819-2.
- [33] J. B. Blackburn, I. Pokrovskaya, P. Fisher, D. Ungar, and V. V. Lupashin, "COG complex complexities: Detailed characterization of a complete set of HEK293T cells lacking individual COG subunits," *Front Cell Dev Biol*, vol. 4, no. March, pp. 1–16, 2016, doi: 10.3389/fcell.2016.00023.
- [34] J. B. Blackburn and V. V. Lupashin, "Creating Knock-outs of Conserved Oligomeric Golgi complex subunits using CRISPR-mediated gene editing paired with a selection strategy based on glycosylation defects associated with impaired COG complex function," *Methods in Molecular Biology*, pp. 1–16, 2017, doi: 10.1007/978-1-4939-6463-5.
- [35] R. Adusumalli, H. C. Åsheim, V. Lupashin, J. B. Blackburn, and K. Prydz, "Proteoglycan synthesis in conserved oligomeric Golgi subunit deficient HEK293T cells is affected differently, depending on the lacking subunit," *Traffic*, vol. 22, no. 7, pp. 230–239, 2021, doi: 10.1111/tra.12804.
- [36] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger, "ColabFold: making protein folding accessible to all," *Nat Methods*, vol. 19, no. 6, pp. 679–682, Jun. 2022, doi: 10.1038/s41592-022-01488-1.
- [37] D. V. Laurents, "AlphaFold 2 and NMR Spectroscopy: Partners to Understand Protein Structure, Dynamics and Function," *Frontiers in Molecular Biosciences*, vol. 9. Frontiers Media S.A., May 17, 2022. doi: 10.3389/fmolb.2022.906437.
- [38] A. G. de Brevern, "An agnostic analysis of the human AlphaFold2 proteome using local protein conformations," *Biochimie*, vol. 207, pp. 11–19, Apr. 2023, doi: 10.1016/j.biochi.2022.11.009.
- [39] E. Forsberg and L. Kjellén, "Heparan sulfate : lessons from knockout mice Find the latest version : Heparan sulfate : lessons from knockout mice," *Journal of Clinical Investigation*, vol. 108, no. 2, pp. 175–180, 2001, doi: 10.1172/JCI200113561.Kidney.
- [40] N. F. Müller *et al.*, "Trimeric autotransporter adhesin-dependent adherence of *Bartonella henselae*, *Bartonella quintana*, and *Yersinia enterocolitica* to matrix components and endothelial cells under static and dynamic flow conditions," *Infect Immun*, vol. 79, no. 7, pp. 2544–2553, 2011, doi: 10.1128/IAI.01309-10.
- [41] Y. El Tahir and M. Skurnik, "YadA, the multifaceted *Yersinia* adhesin.," *Int J Med Microbiol*, vol. 291, pp. 209–218, 2001, doi: 10.1078/1438-4221-00119.
- [42] H. Yu, E. M. Muñoz, R. E. Edens, and R. J. Linhardt, "Kinetic studies on the interactions of heparin and complement proteins using surface plasmon resonance," *Biochim Biophys Acta Gen Subj*, vol. 1726, no. 2, pp. 168–176, 2005, doi: 10.1016/j.bbagen.2005.08.003.
- [43] G. L. Logue, "Effect of heparin on complement activation and lysis of paroxysmal nocturnal hemoglobinuria (PNH) red cells," *Blood*, vol. 50, no. 2, pp. 239–247, 1977, doi: 10.1182/blood.v50.2.239.239.
- [44] A. Sahu and M. K. Panoburn, "Identification of multiple sites of interaction between heparin and the complement system," *Mol Immunol*, vol. 30, no. 7, pp. 679–684, 1993, doi: 10.1016/0161-5890(93)90079-Q.

- [45] M. K. H. Schindler *et al.*, "Yersinia enterocolitica YadA Mediates Complement Evasion by Recruitment and Inactivation of C3 Products," *The Journal of Immunology*, vol. 189, no. 10, pp. 4900–4908, 2012, doi: 10.4049/jimmunol.1201383.
- [46] B. Keller *et al.*, "Yersinia enterocolitica exploits different pathways to accomplish adhesion and toxin injection into host cells," *Cell Microbiol*, vol. 17, no. 8, pp. 1179–1204, Aug. 2015, doi: 10.1111/cmi.12429.
- [47] M. Schütz *et al.*, "Trimer stability of YadA is critical for virulence of Yersinia enterocolitica," *Infect Immun*, vol. 78, no. 6, pp. 2677–2690, 2010, doi: 10.1128/IAI.01350-09.
- [48] M. N. Sorin, J. Kuhn, A. C. Stasiak, and T. Stehle, "Structural insight into non-enveloped virus binding to glycosaminoglycan receptors: A review," *Viruses*, vol. 13, no. 5, pp. 1–11, 2021, doi: 10.3390/v13050800.
- [49] D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith, "Enzymatic assembly of DNA molecules up to several hundred kilobases," *Nat Methods*, vol. 6, no. 5, pp. 343–345, 2009, doi: 10.1038/nmeth.1318.
- [50] Y. Chen, X. Huang, R. Wang, S. Wang, and N. Shi, "The structure of a GFP-based antibody (fluorobody) to TLH, a toxin from *Vibrio parahaemolyticus*," *Acta Crystallographica Section:F Structural Biology Communications*, vol. 71, pp. 913–918, 2015, doi: 10.1107/S2053230X15008845.
- [51] A. Saragliadis and D. Linke, "Assay development for the discovery of small-molecule inhibitors of YadA adhesion to collagen," *The Cell Surface*, vol. 5, no. May, p. 100025, 2019, doi: 10.1016/j.tcs.2019.100025.
- [52] U. Grosskinsky *et al.*, "A conserved glycine residue of trimeric autotransporter domains plays a key role in Yersinia adhesin A autotransport," *J Bacteriol*, vol. 189, no. 24, pp. 9011–9019, 2007, doi: 10.1128/JB.00985-07.
- [53] T. Trunk, H. S. Khalil, and J. C. Leo, "Bacterial autoaggregation," *AIMS Microbiol*, vol. 4, no. 1, pp. 140–164, 2018, doi: 10.3934/microbiol.2018.1.140.
- [54] K. Stott, J. Keeler, T. L. Hwang, A. J. Shaka, and J. Stonehouse, "Excitation Sculpting in High-Resolution Nuclear Magnetic Resonance Spectroscopy: Application to Selective NOE Experiments," *J Am Chem Soc*, vol. 117, no. 14, pp. 4199–4200, 1995, doi: 10.1021/ja00119a048.
- [55] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, and B. D. Sykes, "1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects," *J Biomol NMR*, vol. 5, no. 1, pp. 67–81, 1995, doi: 10.1007/BF00227471.
- [56] W. Lee, M. Tonelli, and J. L. Markley, "NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy," *Bioinformatics*, vol. 31, no. 8, pp. 1325–1327, 2015, doi: 10.1093/bioinformatics/btu830.
- [57] K. Wüthrich, "NMR with Proteins and Nucleic ACids," 1986.
- [58] B. Bardiaux, T. Malliavin, and M. Nilges, *ARIA for Solution and Solid-State NMR.*, vol. 831. 2012.
- [59] S. H. Chen, "Interaction in Micellar and Microemulsion," *Annu Rev Phys Chem*, vol. 37, pp. 351–399, 1986.

[60] J. S. Pedersen and P. Schurtenberger, "Scattering functions of semiflexible polymers with and without excluded volume effects," *Macromolecules*, vol. 29, no. 23, pp. 7602–7612, 1996, doi: 10.1021/ma9607630.

## Tables

**Table 1.:** Primers used in this study.

Name	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')
R3A	ATATCCAGCAGCCCAATATTACG	TGCTGGATATAATCCCACAC
R7A	ATATTAGCCCCAGAAAACCCAAAAT TAC	CTGGGGCTAATATTGGGGCTG
K12A	CAGAAAACCCAGCCTTACCTCCAG	GGGTTTTCTGGACGTAATATTGG
K17A	CCTCCAGAAGCCCCAGGATCAC	TTCTGGAGGTAATTTGGGTTTTTC
R21A	CAGGATCAGCGCTAGAAAGATCTAG	TTCTAGCGCTGATCCTGGTTTTTC
R24A	GTCTAGAAGCATCTAGATTACATCTA GC	CTAGATGCTTCTAGACGTGATCC
R26A	GATCTGCATTACATCTAGCAGAATC	GATGTAATGCAGATCTTTCTAGACG
R36A	ACTACCAGCAGTACCAGGC	GTAAGTCTGGTAGTATTGATTCTG
pASK- Iba2_Sumo linearization	TGTGCGACATTTTTTTTGTGTC	ACCACCAATCTGTTCTCTG
N-terminal region insertion Sumo	ATTGAGGCTCACAGAGAACAGATTG GTGGTCCAGCAAGACCAATATTACG	GTAACGGCAGACAAAAAAAATGTCGCACATTATA CTCGTGGTAGTATTGATTCTG
TEV insertion	CTTGTATTTTCAGGGCCCAGCAAGAC CAAT	GAAAATACAAGTTTTCCACCAATCTGTT
sfGFP open	ATTATGACGGAATTTTAAAGTTTGC	CAATTAGCAGATCATTATCAACAA
N-terminal region insertion sfGFP	CTTTAAAATTCGTCATAATGTAGAAG ATGGTTCACCAGCAAGACCAATATT CCGTC	ATTTGTTGATAATGATCTCGCTAATTGTACTCGTGGT AGTATTGATTCTGCTAG



## Figure Legends

**Figure 1.** Overview of YadA and its heparin binding peptide and the minimal units of heparin and heparan sulphate (HS). **(a)** Sequence alignment of YadA head domain sequences from YadA of different *Y. enterocolitica* serotypes (YadA<sub>O:3</sub> GI:48607, YadA<sub>O:8</sub> GI:122815846, YadA<sub>O:9</sub> GI:972903261). The 35-residue N-terminal insertion in YadA from *Y. enterocolitica* serotype O:9 is colored in red **(b)** Cartoon representation of the structure of the YadA head domain monomer from *Y. enterocolitica* serotype O:3 (green) (PDB: 1P9H). The figures were generated using PyMol 2.5.3. In the left panel the non-resolved loop is indicated with a red dotted line. In the right panel, the longer N-terminal peptide as found in YadA<sub>O:9</sub> is indicated with a dotted red line for comparison. **(c)** Repeating disaccharide units of heparin and HS with potential sites for sulphation indicated with green ellipses. 3-O-sulphation can only be found in heparin.

**Figure 2.** YadA binding to heparin depends on the interaction between the basic residues within the YadA<sub>O:9</sub> N-terminal motif and negative charges of the sulphate groups within heparin. **(a)** Minimal repeating units of the glycans used in the binding assays shown in b and c. **(b)** Bar diagram showing binding of different glycans to purified, surface immobilized YadA<sub>O:9</sub> head domains. Detection relies on biotinylation of the glycans. Biotin was therefore included as a control. The experiment was performed in triplicate. The error bars show standard deviation from the depicted averaged three individual experiments. Significance levels are indicated with  $p < 0.05$  (\*),  $p < 0.01$  (\*\*), or  $p < 0.001$  (\*\*\*) based on a one way ANOVA and Tukey's t-test. **(c)** Binding curve showing the concentration dependency of glycan binding. The curve was fitted using a Hill fit. **(d)** Bar diagram showing binding of heparin to YadA<sub>O:8</sub> and YadA<sub>O:9</sub> head domains, as well as YadA<sub>O:9</sub> entailing point mutations within the N-terminal heparin-binding motif. The residue numbering refers to the residues within the motif as shown in Fig. 3a. The experiment was performed in triplicates. The error bars show standard deviations from the depicted averaged three individual experiments. Significance levels are indicated with  $p < 0.05$  (\*),  $p < 0.01$  (\*\*), or  $p < 0.001$  (\*\*\*) based on a one way ANOVA and Tukey's t-test. **(e)** Bar diagram showing heparin binding to YadA<sub>O:9</sub> point mutants within the N-terminal peptide in the context of YadA<sub>O:9</sub> wild-type (wt) exhibited on the surface of *E. coli*. The residue numbering refers to the residues within the motif as shown in Fig. 3a. The experiment was performed in triplicates. The error bars show standard deviations from the depicted averaged three individual experiments. No statistical significance was found according to one way ANOVA testing and Tukey's t-test. **(f)** Dot blot showing binding of heparin to sfGFP (superfolder GFP), YadA<sub>O:9</sub> and the sfGFP\_YadA<sub>Nt</sub> fusion protein immobilized on a nitrocellulose membrane. Each protein was immobilized in triplicates and bound heparin was detected via its biotinylation.

**Figure 3.** The entire heparin-binding motif of YadA<sub>O:9</sub> is involved in the YadA<sub>O:9</sub>-heparin interaction. **(a)** Overview of peptides used in this study. Peptide 1 (pep1) is the peptide as found in the YadA<sub>O:9</sub> wild-type (wt) protein (residue 29-65). Peptide 2 (pep2) was recombinantly expressed as a Sumo fusion peptide. The peptide in its <sup>15</sup>N-labeled form was used for heparin titration experiments as well as other functional NMR studies. Peptide 3 (pep3) was synthesized and was used for initial NMR peak assignments as well as SAXS measurements. <sup>15</sup>N HSQC titration experiment. <sup>15</sup>N labeled pep2 was used for recording of the spectra. Chemical shift changes were recorded upon addition of 50 µg (orange), 100 µg (yellow), 200 µg (green), 400 µg (blue) and 800 µg (purple) of heparin, equaling molar binding ratios of 2:1, 1:1, 1:2, 1:3, 1:6 and 1:10, respectively. Maximal shifts of the peaks were observed upon addition of 400 µg of heparin. **(b)** Bar diagram depicting the maximal pep2 combined <sup>1</sup>H-<sup>15</sup>N  $\delta\Delta$  chemical shift changes binding to heparin. The dotted line indicates the threshold for differentiation between CSP considered strong vs. weak. **(c)** Representative curves showing chemical shift perturbations upon addition of different concentrations of heparin for residue L6, E23 and R24. The curves indicate binding of heparin by the peptide in the lower µM range. **(d)** SAXS measurements showing the scattering intensity plotted against the modulus of the scattering vector Q of varying

peptide/heparin ratios obtained at 37°C. The SAXS curves result from 10 frames of the same sample, each with an exposure time of 1s.

**Figure 4.** The N-terminal heparin-binding peptide from YadA is extremely flexible but shows a high propensity to adopt a PPII helix conformation. **(a)** CD spectra of heparin (black, solid line), peptide (blue, dashed line), equimolar peptide/heparin mix (red, dashed line) and the sum of adding the peptide and heparin spectra (gray, dotted line). Measurements were performed in triplicates with each replicate scanned five times. **(b)** Talos prediction of  $\varphi/\psi$  dihedral angles derived from  $H\alpha$ , HN, C,  $C\alpha$ ,  $C\beta$  and N chemical shifts. Empty circles depict generous predictions, filled circles depict strong prediction and circles with low opacity depict bad predictions based on the error bars shown in the graph. **(c)** Bar graph of predicted populations of secondary structure derived from chemical shifts done using  $\delta 2D$  analysis. **(d)** Bar graph showing secondary structure frequencies of each residue based in 1000 conformers generated. Secondary structures were computed using PROSS [29]. PPII helices are shown as a green bar and the pink bar indicates a turn.

**Figure 5.** Structural determination and modeling of the  $YadA_{0-9}$  N-terminal heparin-binding peptide. **(a)** Superimposition of well-ordered regions within the  $YadA_{0-9}$  N-terminal peptide among the conformer ensemble generated from chemical shift-derived dihedral angles. Pro residues are shown in orange, basic residues are indicated in blue and acidic residues are shown in red. The remaining residues are shown in green. The model of the region spanning residues 0-11 within the YadA heparin binding motif forming a poly proline II (PPII) type helix. Similarly, the region spanning residues 13-18 within the YadA heparin binding motif exhibits the propensity to form a PPII type helix. For the region spanning residues 19-28 within the YadA heparin binding motif a turn was predicted, placing residue R21 and R24 in close proximity. Lastly, the model of the region spanning residues 29-38 within the YadA heparin binding motif forms a PPII type helix. **(b)** Cartoon representation of the best AlphaFold2 model for the  $YadA_{0-9}$  head domain colored in gray with the N-terminal heparin-binding motif colored in magenta as a side and a top view (upper panel). Predicted pep1 structure shown as sticks. Pro residues are shown in orange, basic residues are indicated in blue and acidic residues are shown in red. The remaining residues are shown in gray. All basic side-chains are exposed except for R26 which is buried, in concordance with the mutation studies. All structural models were generated using AlphaFold 2 and representations of the generated pdb were prepared in PyMol 2.5.3. pLDDT scores plotted against the residue number of the  $YadA_{0-9}$  head domain. Overall, the pLDDT scores for the N-terminal heparin binding region are very low in confidence as indicated in the magenta box in the blot.

**Figure 6.** E. coli expressing wild-type (wt)  $YadA_{0-9}$  on the surface show decreased adhesion to HEK cells with decreased cell surface Heparan Sulphate (HS) levels. Surface adherent wt HEK293T cells, COG1-/- HEK293T cells (reduced GAG synthesis but same overall amount of HS binding sites on the surface) and HEK COG3-/- (reduced levels of HS on the surface) were challenged with fluorescent E. coli exposing  $YadA_{0-9}$  wt on the surface. To correct for background adhesion, HEK293T wt and mutants were challenged with fluorescent bacteria not producing  $YadA_{0-9}$ . The fluorescent values were subtracted from the fluorescence values determined for adherent E. coli carrying  $YadA_{0-9}$ .  $YadA_{0-9}$ -mediated bacterial adhesion to HEK cells was significantly reduced in HEK293T COG3-/- cells but not to HEK293T wt and HEK293T COG1-/- cells. Significance levels are indicated with  $p < 0.05$  (\*),  $p < 0.01$  (\*\*), or  $p < 0.001$  (\*\*\*) based on a one way ANOVA and Tukey's t-test. The error bars show standard deviation from the depicted averaged three individual experiments.

**a**

```

0:8  -HFTAVQISPNSDDPSHVMIFQ-----PEVRAPGGTNALAKGTHSIAVGASAEAAERAHAVGAGSIATGVNSVAIGPLSKALGDSAVTYGAG
0:3  PNLTA VQISP NADPALG- LEYPVR-----PPVPGAGGLNASAKGIHSIAGATAEAAKGAAVAVGAGSIATGVNSVAIGPLSKALGDSAVTYGAG
0:9  PRLSAVQISP NADPELGVGLYPARILRPENPKLPPEKPGSRLERSRLHLAESILPRVPGAGGLNASAKGIHSIAGATAEAAKGAAVAVGAGTIATGVNSVAIGPLSKALGDSAVTYGAA
    ..:*****:** ..

```

```

0:8  STAQKDGVAIGARASTDTGVAVGFNSKVDAKNSVAIGHSSHVVDHDYSIAIGDRSKTDRKNSVSIHESLNRQLTHLAAGTKD TDAVNVAQLKKEIEKTQENANKK
0:3  STAQKDGVAIGARASTDTGVAVGFNSKADAKNSVAIGHSSHVAANHGSYIAIGDRSKTDRENSVSIHESLNRQLTHLAAGTKD TDAVNVAQLKKEIEKTQENTNKR
0:9  STAQKDGVAIGARASTDTGVAVGFNSKADAKNSVAIGHSSHVAVDHDYSIAIGDRSKTDRENSVSIHESLNRQLTHLAAGTKD TDAVNVAQLKKEIEKTQENTNKK
    *****:** ..:*****:**

```

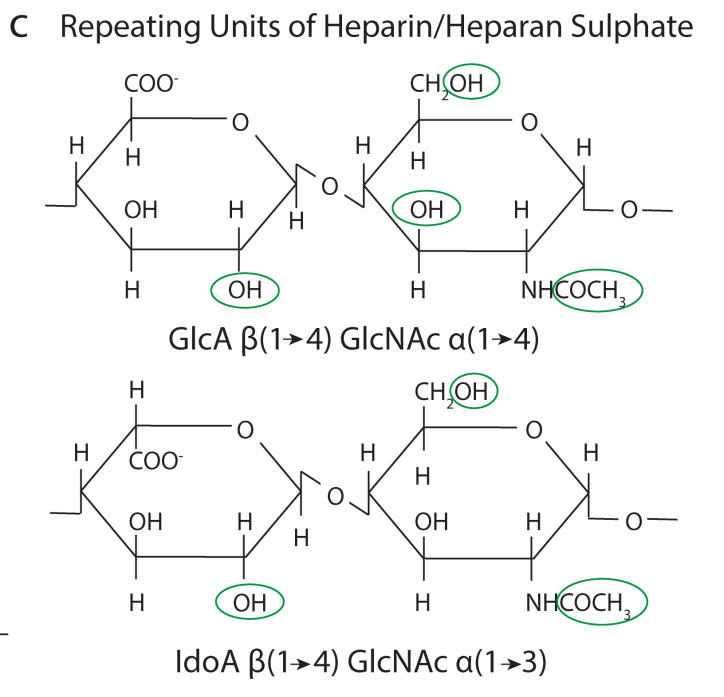
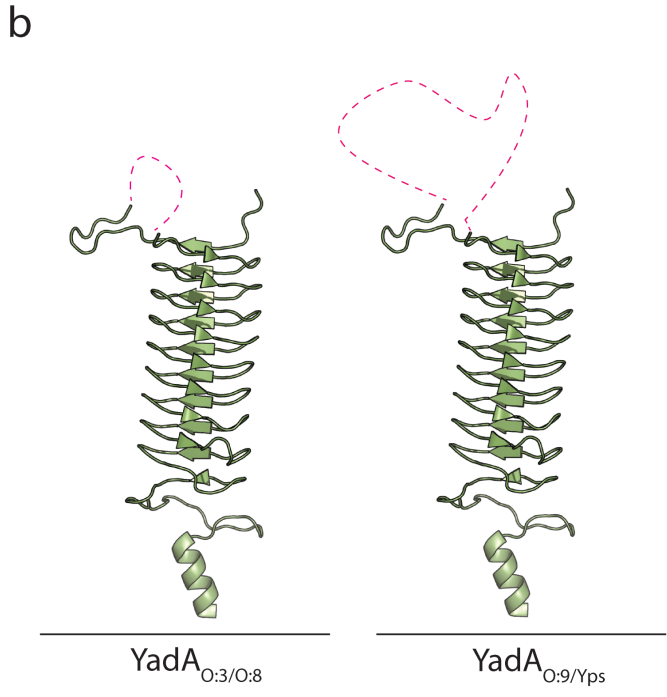


Fig1.tif

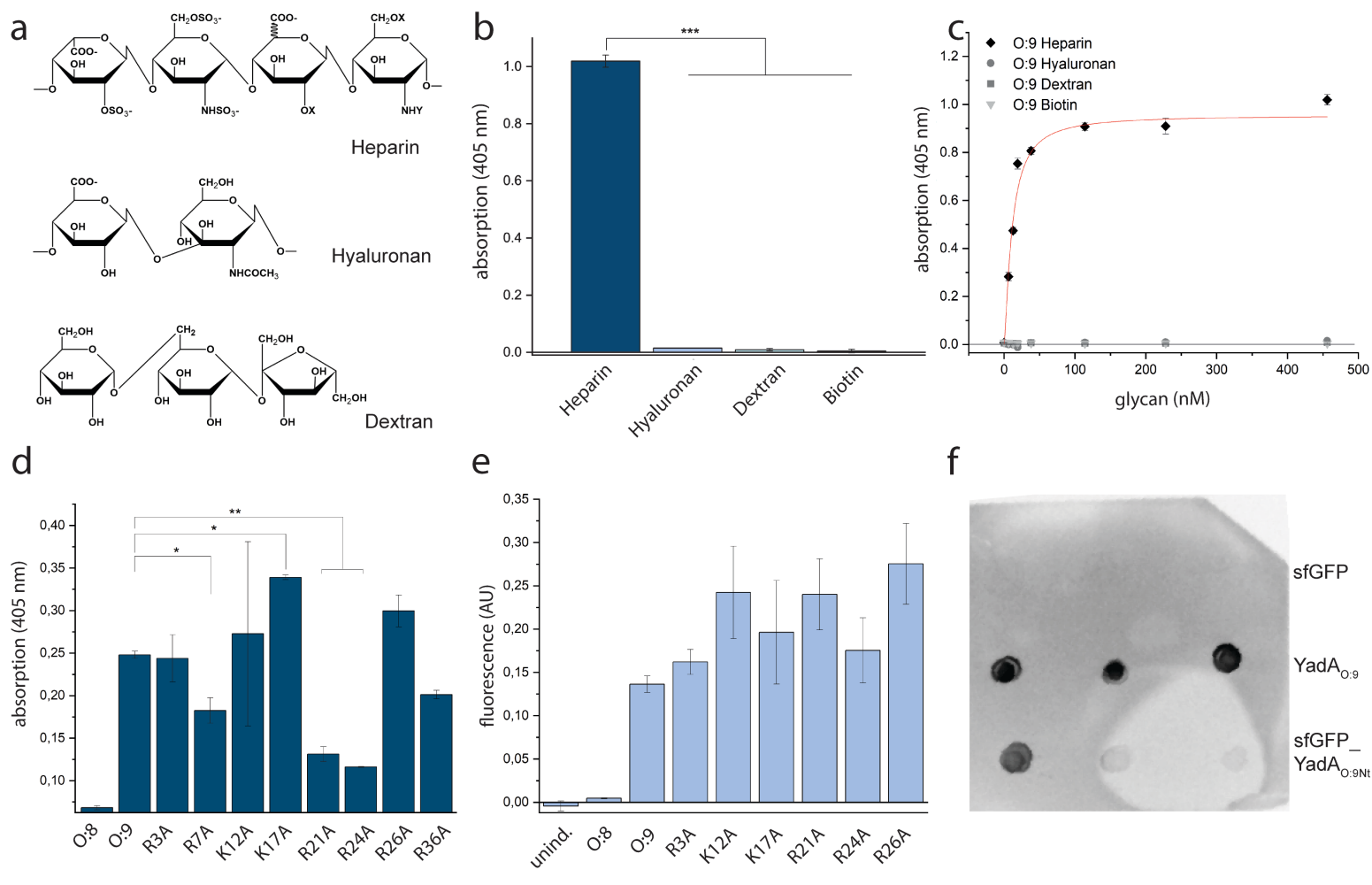


Fig2.tif

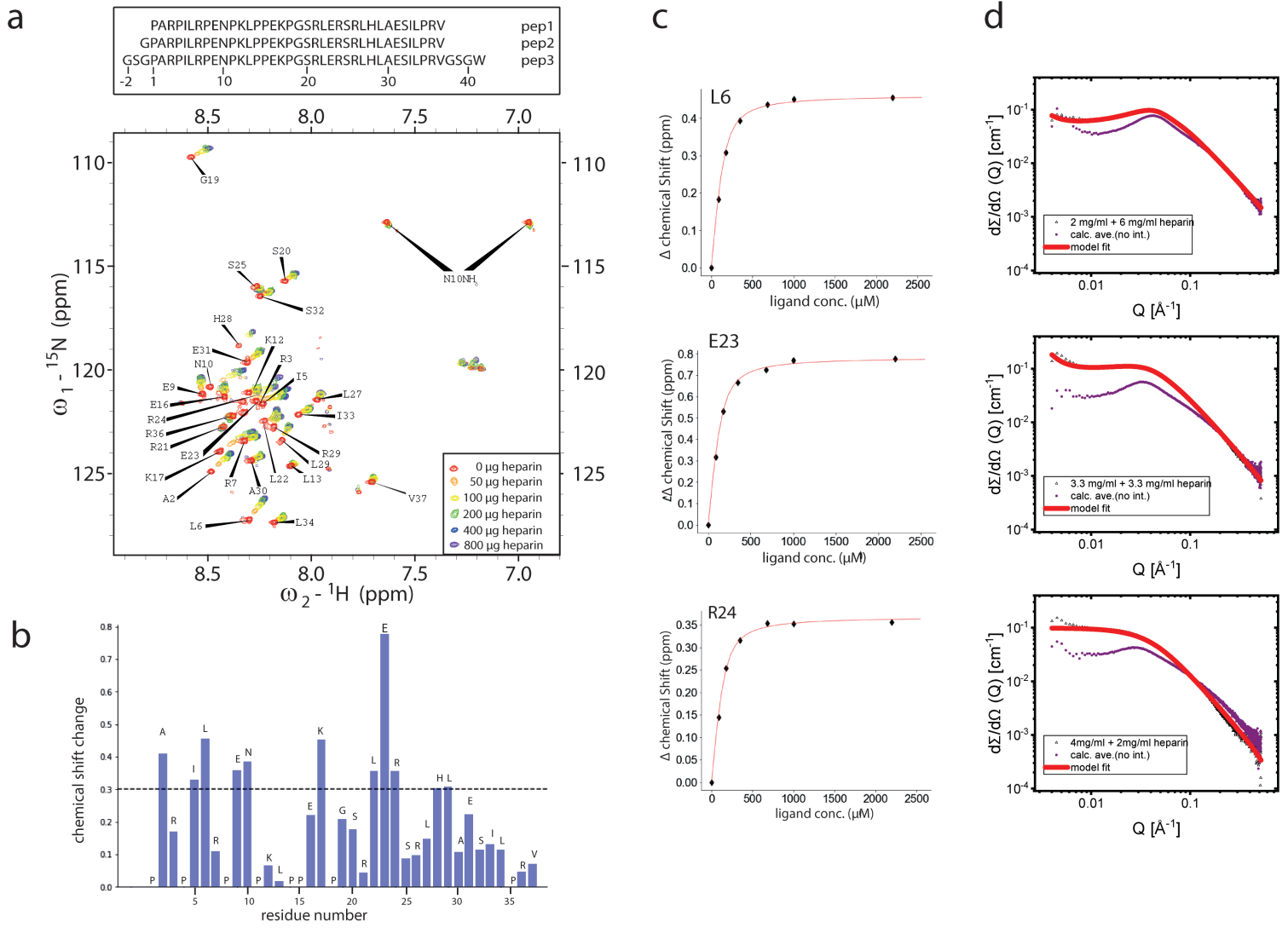


Fig3.tif

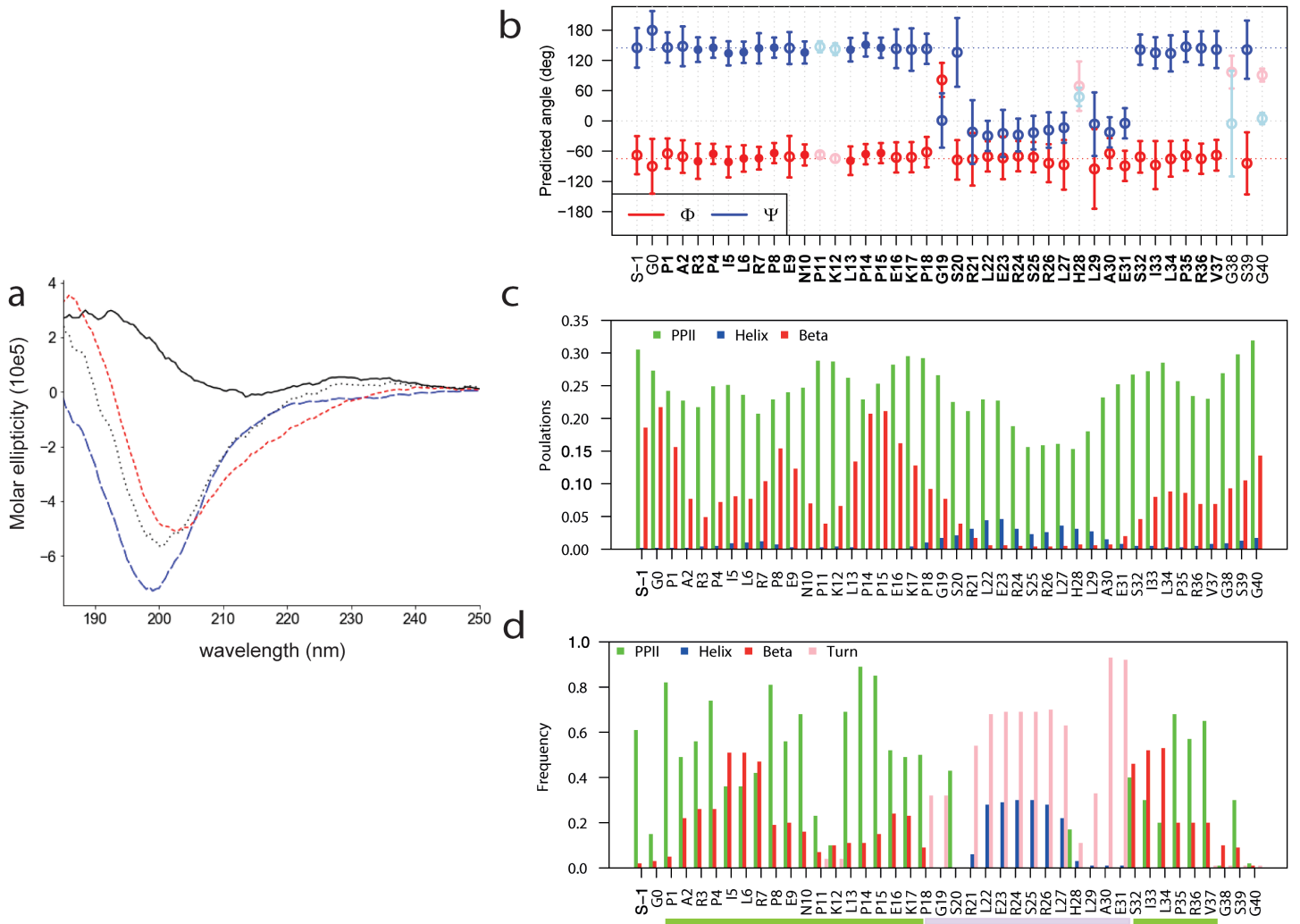


Fig4.tif



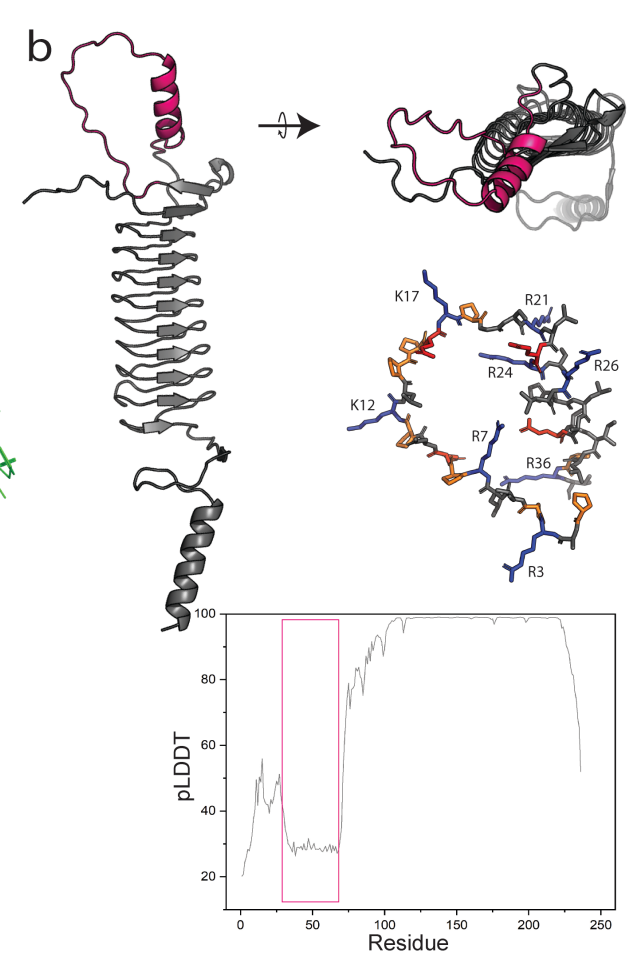
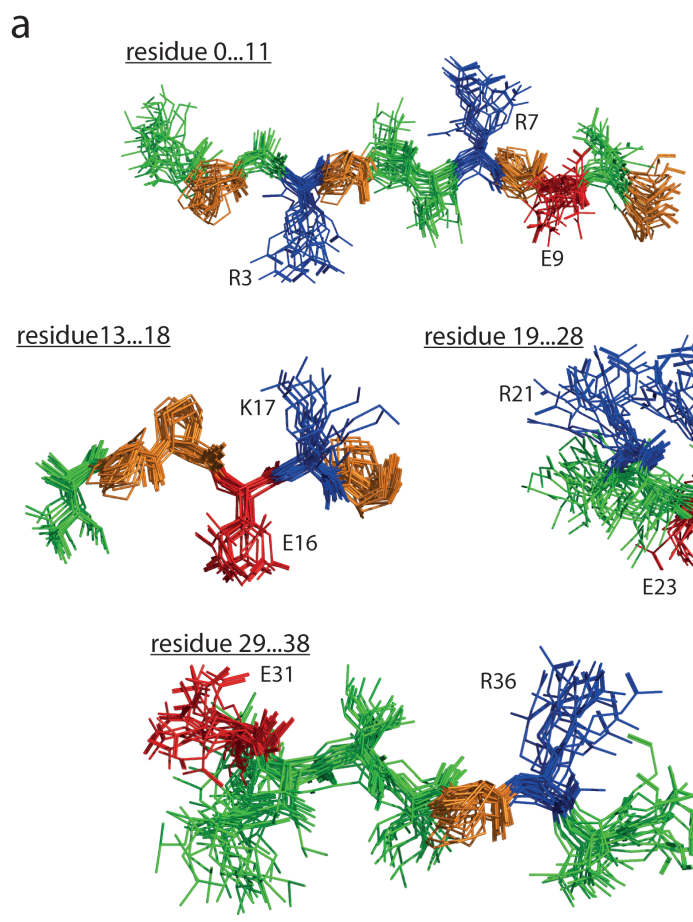


Fig5.tif

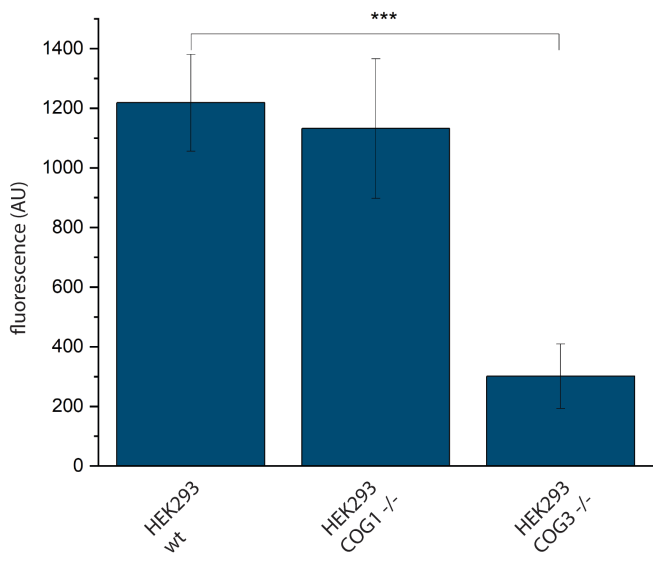


Fig6.tif