



**HAL**  
open science

# The IbeA protein from adherent invasive *Escherichia coli* is a flavoprotein sharing structural homology with FAD-dependent oxidoreductases

Théo Paris, Agneta Kiss, Luca Signor, Georges Lutfalla, Mickaël Blaise,  
Elisabetta Boeri Erba, Laurent Chaloin, Laure Yatime

## ► To cite this version:

Théo Paris, Agneta Kiss, Luca Signor, Georges Lutfalla, Mickaël Blaise, et al.. The IbeA protein from adherent invasive *Escherichia coli* is a flavoprotein sharing structural homology with FAD-dependent oxidoreductases. *FEBS Journal*, 2024, 291 (1), pp.177-203. 10.1111/febs.16969 . hal-04246071

**HAL Id: hal-04246071**



**<https://hal.science/hal-04246071v1>**

Submitted on 9 Oct 2024

**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.

# The IbeA protein from adherent invasive *Escherichia coli* is a flavoprotein sharing structural homology with FAD-dependent oxidoreductases

Théo Paris<sup>1</sup>, Agneta Kiss<sup>2</sup>, Luca Signor<sup>2</sup>, Georges Lutfalla<sup>1</sup>, Mickaël Blaise<sup>3</sup> ,  
 Elisabetta Boeri Erba<sup>2</sup>, Laurent Chaloin<sup>3</sup> and Laure Yatime<sup>1</sup> 

1 LPHI, Univ. Montpellier, CNRS, INSERM, France

2 Univ. Grenoble Alpes, CEA, CNRS, IBS, Grenoble, France

3 IRIM, Univ. Montpellier, CNRS, France

## Keywords

AIEC; FAD-oxidoreductase; flavoprotein; homodimer; invasion of brain endothelium protein A (IbeA)

## Correspondence

L. Yatime, LPHI UMR5294, University of Montpellier, Place E. Bataillon, Bât. 24 CC107, Montpellier cedex 05 34095, France  
 Tel: + 33 (0)4 6714 9203  
 E-mail: [laure.yatime@umontpellier.fr](mailto:laure.yatime@umontpellier.fr)  
 Website: <https://lphi.umontpellier.fr/research-teams/immunity-inflammation-and-bacterial-virulence/yatime/>

(Received 30 June 2023, revised 22 September 2023, accepted 29 September 2023)

doi:10.1111/febs.16969

Invasion of brain endothelium protein A (IbeA) is a virulence factor specific to pathogenic *Escherichia coli*. Originally identified in the K1 strain causing neonatal meningitis, it was more recently found in avian pathogenic *Escherichia coli* (APEC) and adherent invasive *Escherichia coli* (AIEC). In these bacteria, IbeA facilitates host cell invasion and intracellular survival, in particular, under harsh conditions like oxidative stress. Furthermore, IbeA from AIEC contributes to intramacrophage survival and replication, thus enhancing the inflammatory response within the intestine. Therefore, this factor is a promising drug target for anti-AIEC strategies in the context of Crohn's disease. Despite such an important role, the biological function of IbeA remains largely unknown. In particular, its exact nature and cellular localization, i.e., membrane-bound invasin versus cytosolic factor, are still of debate. Here, we developed an efficient protocol for recombinant expression of IbeA under native conditions and demonstrated that IbeA from AIEC is a soluble, homodimeric flavoprotein. Using mass spectrometry and tryptophan fluorescence measurements, we further showed that IbeA preferentially binds flavin adenine dinucleotide (FAD), with an affinity in the one-hundred nanomolar range and optimal binding under reducing conditions. 3D-modeling with AlphaFold revealed that IbeA shares strong structural homology with FAD-dependent oxidoreductases. Finally, we used ligand docking, mutational analyses, and molecular dynamics simulations to identify the FAD binding pocket within IbeA and characterize possible conformational changes occurring upon ligand binding. Overall, we suggest that the role of IbeA in the survival of AIEC within host cells, notably macrophages, is linked to modulation of redox processes.

## Abbreviations

ACN, acetonitrile; AIEC, adherent invasive *Escherichia coli*; APEC, avian pathogenic *Escherichia coli*; Caspr1, contactin-associated protein 1; CD, Crohn's disease; DIMS, direct infusion mass spectrometry; FAD, flavin adenine dinucleotide; FMN, flavin mononucleotide; GGR, geranylgeranyl reductase; GR, glutathione reductase; HPLC, high-performance liquid chromatography; IbeA, invasion of brain endothelium protein A; IFE, inner filter effects; IPTG, isopropyl  $\beta$ -D-1-thiogalactopyranoside;  $K_{d-app}$ , apparent dissociation constant; LC/ESI-TOF, liquid chromatography coupled to electrospray ionization with time-of-flight; MD, molecular dynamics; MS, mass spectrometry; NADH, nicotinamide adenine dinucleotide; NADPH, nicotinamide adenine dinucleotide phosphate; NATA, N-Acetyl-L-tryptophanamide; NMEC, neonatal meningitis-causing *Escherichia coli*; PDB, Protein Data Bank; PSF, polypyrimidine tract-binding protein-associated splicing factor; RMSD, root-mean-square deviation; SEC, size exclusion chromatography; T4SS, type IV secretion system; T6SS, type VI secretion system; TEC, translocating *Escherichia coli*; TEV, tobacco etch virus; TFA, trifluoroacetic acid; WT, wild-type.

## Introduction

Adherent-Invasive *Escherichia coli* (AIEC) constitute a subclass of *E. coli* pathogens that reside in the human gastrointestinal tract [1]. They are often encountered in healthy individuals but their low prevalence, combined with efficient containment by the immune system, prevents any dysregulation of intestinal homeostasis [2]. In contrast, AIEC are abundantly found in the gut microbiota of persons suffering from Crohn's disease (CD) [1,3,4]. In a genetic and environmental background favorable to the development of inflammatory bowel diseases, AIEC become predominant over commensal bacteria and can freely exert their pathogenic traits, thereby generating an inflammatory burden that strongly contributes to the symptoms of CD [5–9]. AIEC are defined by their phenotypic features [10,11] and a hallmark of these *E. coli* strains is their capacity to invade and persist within intestinal epithelial cells and macrophages [12,13]. Intramacrophage AIEC have been shown to induce massive release of tumor necrosis factor  $\alpha$ , thus sustaining the inflammatory burst [14]. AIEC are able to form biofilms along the intestinal cell wall, a property that allows them to escape immune detection and better resist to antibiotic treatments [15,16]. In addition, they can efficiently cross the mucus layer and get internalized in Peyer's patches, through translocation across M cells [17–19]. This provides a first entry point for AIEC into lymphoid tissues, in a region of the intestinal epithelium where the first inflammatory lesions linked to CD are generally observed. An involvement of AIEC in provoking these early lesions has even been suggested but still awaits confirmation [20].

The role of AIEC in the pathogenesis of CD has garnered huge attention to this pathobiont. Many research efforts are currently dedicated to the identification of AIEC virulence factors that may be therapeutically targeted for anti-CD applications [21,22]. Interestingly, an orthologue of the IbeA protein (*Invasion of brain endothelium protein A*) was recently identified in AIEC [23]. IbeA has so far only been found in pathogenic *E. coli* and the literature available on its putative role suggests that it may serve as a virulence factor in various pathogenic *E. coli* including Neonatal Meningitis-causing *E. coli* (NMEC) and Avian Pathogenic *E. coli* (APEC). IbeA was originally identified in the NMEC K1 strain [24]. In this bacterium, IbeA was shown to be important for the crossing of the blood–brain barrier, using both an *in vitro* invasion assay on brain microvascular endothelial cells and a rat model of neonatal meningitis [25]. The name of the protein directly arose from these observations. Further

biochemical characterization of the protein using *in vitro* approaches was attempted but trials to produce a recombinant IbeA in bacteria yielded an insoluble protein expressed in inclusion bodies [26]. These results prompted the authors to categorize IbeA as an invasin embedded in the *E. coli* membrane, possibly serving as an anchoring point to the host cell surface. Putative host receptors for IbeA were proposed in follow-up publications, including vimentin, polypyrimidine tract-binding protein-associated splicing factor (PSF), and contactin-associated protein 1 (Caspr1) [27–30]. Other studies on the IbeA protein from APEC confirmed the importance of IbeA for the virulence of pathogenic *E. coli* [31–33]. They also demonstrated a role of the protein in the resistance to H<sub>2</sub>O<sub>2</sub> [34], suggesting that IbeA may also act as a regulatory factor to promote survival in the hostile oxidative environment encountered inside host phagocytic cells. Whether this function would be exerted by a membrane-bound IbeA or by a soluble form released in the phagosomal compartment of macrophages remains unclear.

More recently, the IbeA protein from AIEC was shown to be important for the invasion of intestinal epithelial cells and M cells, as well as for the survival of AIEC within macrophages [23]. Furthermore, infection of mice with an AIEC strain deleted of the *ibeA* gene led to a decreased inflammatory response and a less severe pathological outcome as compared to infection with the wild-type (WT) strain, highlighting the potential role of IbeA in the virulence of AIEC [23]. Based on these characteristics, IbeA appears as a genuine virulence factor and a promising drug target in several *E. coli* strains, including AIEC. Its exact nature, topology, and biological function are however still of debate and the discrepancy between the different architectural models proposed for IbeA, i.e. membrane-bound invasin versus soluble anti-oxidative factor, remains to be elucidated.

To gain insights into the biological function of IbeA, we undertook biochemical investigations on the full-length protein from the reference strain AIEC LF82. We first developed an efficient protocol for the production of a soluble form of recombinant IbeA. We then used a combination of size exclusion chromatography (SEC), mass spectrometry (MS), Tryptophan fluorescence experiments, 3D-modeling, *in silico* ligand docking, and molecular dynamics (MD) simulations, to unravel novel aspects of the biology of this protein. This led us to confirm its resemblance with FAD-dependent oxidoreductases and characterize it, for the very first time, at the molecular level. These results shed new light on the putative biological function of IbeA in modulating redox conditions within host phagocytic cells.

## Results

### Expression of recombinant IbeA as a soluble protein

To date, the sole attempt described in the literature to produce a recombinant IbeA was made on the protein from *E. coli* K1 and yielded an insoluble molecule expressed in inclusion bodies [26,30]. Although an apparently functional IbeA could be recovered following *in vitro* refolding or solubilization with harsh detergent conditions, we aimed at developing a more straightforward protocol for the expression and purification of recombinant IbeA from AIEC. After screening various constructs encompassing different affinity tags and providing distinct addressing strategies (periplasmic or native addressing), we managed to express a soluble form of full-length AIEC IbeA, simply bearing a cleavable N-terminal poly-histidine tag and no particular signal peptide sequence. Due to massive overexpression, a large amount of protein was detected in the insoluble fraction obtained after low-speed centrifugation of disrupted *E. coli* BL21 cells (Fig. 1A, sample P1). Nevertheless, a considerable amount of recombinant IbeA was also recovered in the soluble fraction (Fig. 1A, sample S1).

In unpublished thesis work, Cieza *et al.* [35] previously reported that IbeA is associated with the outer membrane of AIEC. To address whether this was also the case in our heterologous expression system, we further separated the disrupted external membranes of *E. coli* BL21 from the soluble cytosolic/periplasmic proteins by ultracentrifugation. As shown in Fig. 1A, IbeA is found in the pool of soluble proteins (sample S2). No protein can be convincingly detected in the membrane fraction (sample P2) or in a periplasmic extract of the *E. coli* BL21 cells (sample S3), as assessed by SDS/PAGE and Coomassie blue staining. Thus, in our heterologous system, IbeA appears to behave as a fully soluble protein expressed in the cytosol of *E. coli*, without any strong anchor to the external cell membrane. As these investigations were made on IbeA expressed with an N-terminal His<sub>6</sub> tag, we can however not rule out that the presence of the tag affects the subcellular localization of the recombinant protein.

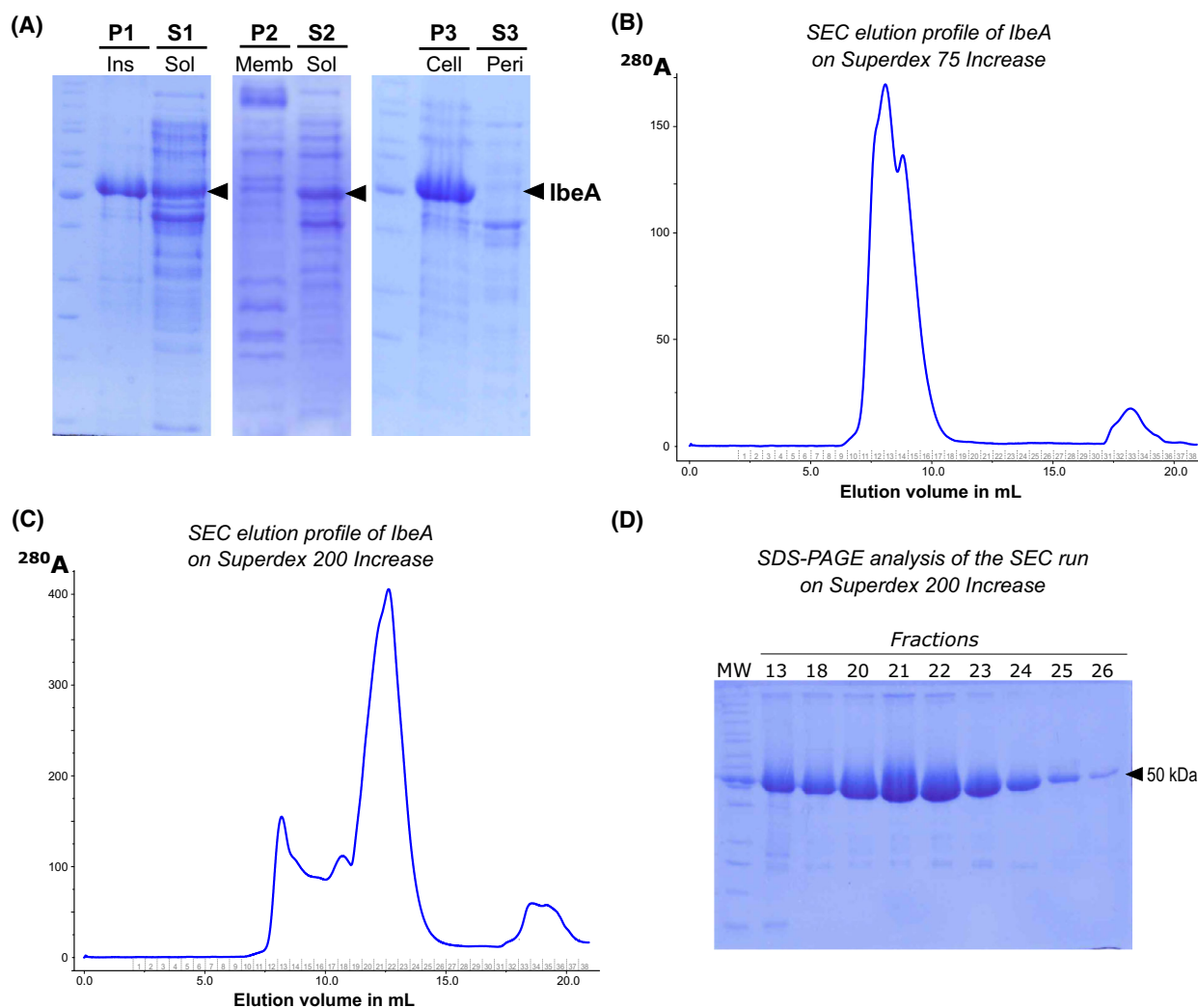
### IbeA is a homodimeric protein both under reducing and non-reducing conditions

To purify recombinant IbeA, we used a classical 3-step procedure consisting in a first Ni-based affinity chromatography, followed by proteolytic cleavage of the polyhistidine tag, counter-purification through affinity

chromatography to recover the cleaved recombinant IbeA, and final size-exclusion chromatography (SEC) (Fig. 1B–D). IbeA contains 10 cysteine residues, of which some could be important in stabilizing its tertiary or quaternary organization. We therefore pursued two parallel purification protocols, under either non-reducing or reducing conditions. Using these strategies, we routinely obtained 20 mg of pure protein per liter of *E. coli* BL21 culture.

Initial SEC data on a Superdex 75 Increase column showed that IbeA elutes almost in the void volume, indicating the presence of larger oligomer(s) than a simple monomer (Fig. 1B; expected size for monomeric IbeA: 49.6 kDa). Elution on a Superdex 200 Increase column confirmed the formation of larger species (Fig. 1C). The main elution peak for IbeA, centered at an elution volume of 13–14 mL depending on the redox conditions, revealed the presence of at least two IbeA forms: a predominant form eluting as the main, latest peak (F1 form), and a less abundant form eluting slightly earlier as a shoulder peak (F2 form). These two forms are detected both under reducing and non-reducing conditions (Fig. 2A,E), although the F2 form is much less prominent under reducing conditions. Both forms were attributed to IbeA since no other contaminant could be detected in the corresponding fractions following SDS/PAGE analysis (Fig. 1D).

The two forms of IbeA (F1 and F2) obtained under both non-reducing and reducing conditions were collected as separate fractions from the SEC run (Fig. 2A,E), and then analyzed by MS. Using native MS, we determined the oligomeric state of these different IbeA species. All these forms are dimeric, regardless of the redox conditions (Fig. 2B,C,F,G). We also used LC/ESI-TOF MS, a denaturing technique that only preserves covalent interactions such as disulfide bridges. Under non-reducing conditions, both the F1 and F2 forms of IbeA associate through disulfide crosslinking, as illustrated by the detection of dimers (D), and possibly trimers (T) and tetramers (Q) (Fig. 2D). This is not surprising since IbeA contains a high number of cysteines. Some of these residues may be exposed on the protein surface and react non-specifically with cysteines from neighboring IbeA molecules. Remarkably, the F2 form contains a predominant proportion of disulfide-crosslinked IbeA dimers, suggesting that these covalent dimers adopt a significantly distinct conformation as compared to the non-covalent dimers and therefore elute as a separate, earlier peak on the SEC column. Under reducing conditions, the LC/ESI-TOF MS spectra of both IbeA forms (Fig. 2H) show only monomers (M), confirming that the covalent forms are only arising from disulfide



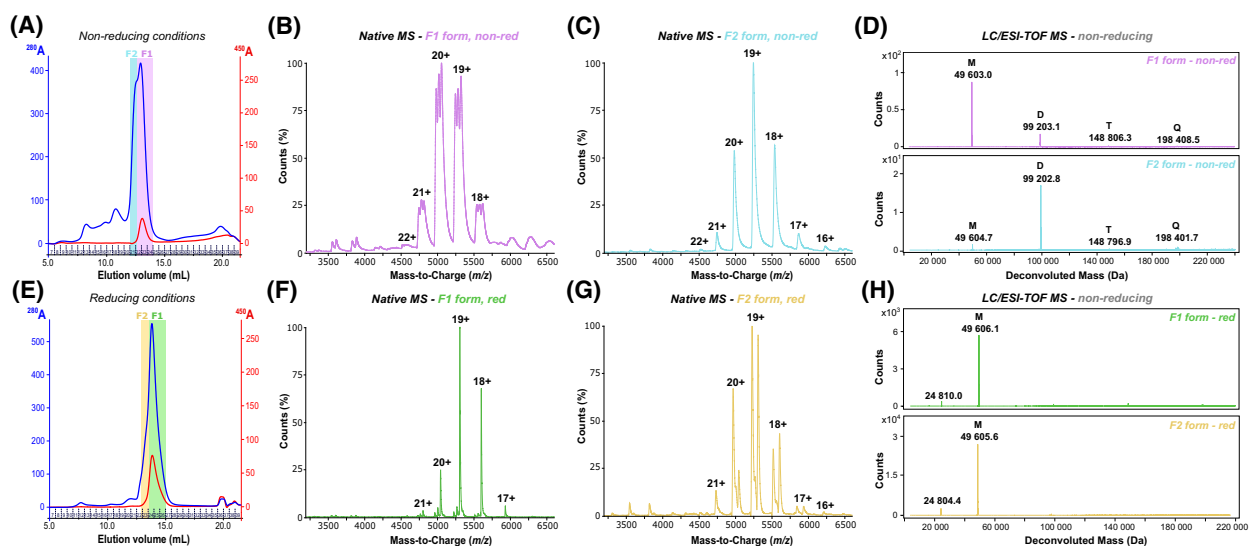
**Fig. 1.** Production of recombinant IbeA as a soluble and highly pure protein. SDS/PAGE gels and chromatograms are from a single, representative experiment but results were reproduced in at least three independent experiments ( $n = 3$ ). (A) SDS/PAGE analysis of the subcellular localization of recombinant IbeA expressed in BL21(DE3) cells. P1: insoluble fraction (cell debris and inclusion bodies); S1: soluble fraction; P2: membrane extract; S2: soluble proteins (cytosolic and periplasmic); P3: intact cell containing cytosolic and inner-membrane proteins, and insoluble fraction; S3: periplasmic extract. (B) Initial elution profile of IbeA on a Superdex 75 Increase column. IbeA elutes very early on that column, almost in the void volume. (C) Elution profile of IbeA on Superdex 200 Increase. Several peaks can be detected, corresponding to larger oligomeric assemblies than simple monomers. (D) SDS/PAGE analysis of the SEC purification step for IbeA on a Superdex 200 Increase column. Fraction numbers corresponding to the chromatogram displayed in panel C are indicated above the lanes. MW: molecular weight marker. Each fraction contains only highly pure IbeA.

crosslinking. Taken together, these MS data demonstrate that IbeA is a homodimeric protein, regardless of its redox state.

### IbeA is a flavoprotein that can bind two FAD molecules per homodimer

Interestingly, after the first Ni-affinity chromatography, a bright yellow coloration was detected in the IbeA sample (Fig. 3A). This coloration was accompanied by

a marked absorbance of UV-visible light at 450 nm, as displayed on the SEC chromatograms of IbeA (Fig. 2A,E, red curves). These observations led us to hypothesize that a flavin cofactor had been captured by IbeA during heterologous expression in bacteria. To identify this compound, we performed direct-infusion MS (DIMS), a technique that allows the detection of small molecules. A peak at 786.16  $m/z$  was observed for the F1 form of IbeA, both under reducing and non-reducing conditions (Fig. 3B,C, middle spectra), and for the fully reduced

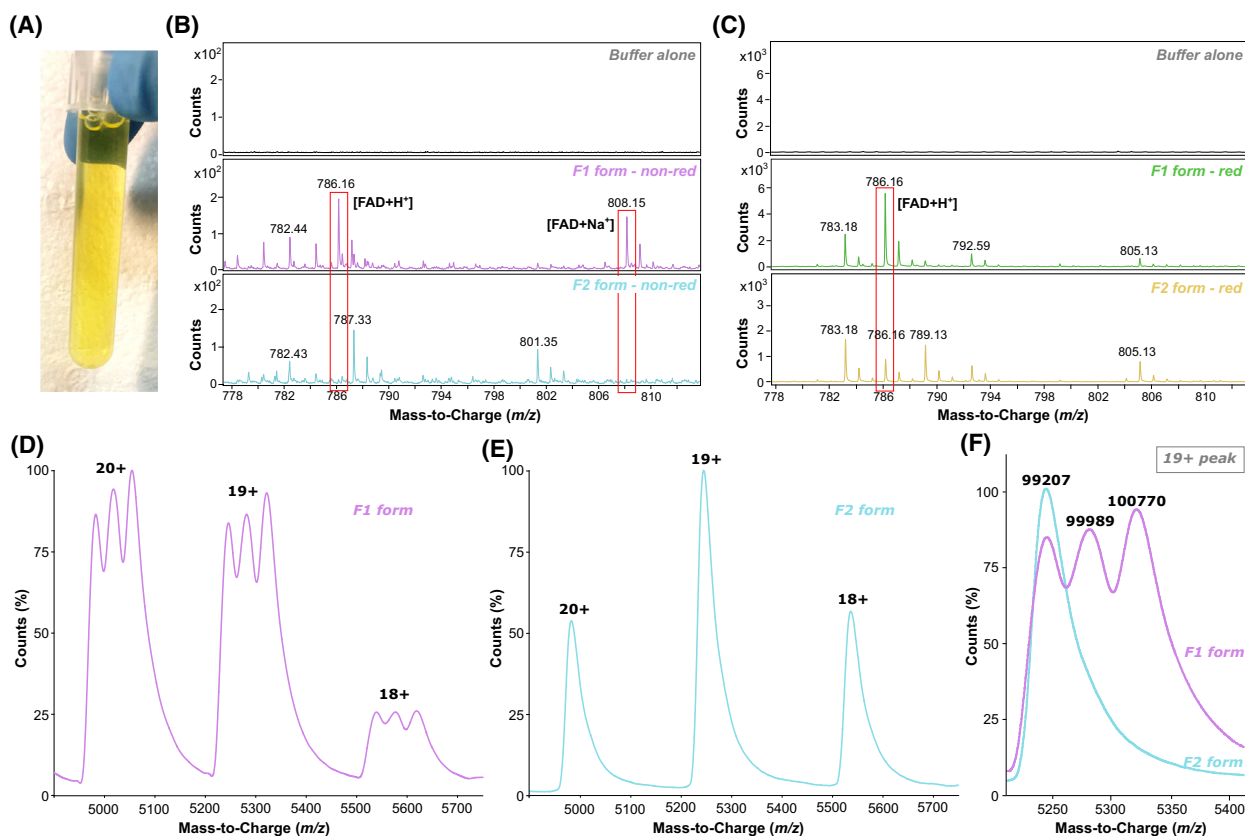


**Fig. 2.** IbeA is a homodimeric protein. Chromatograms and MS spectra are from a single, representative experiment but results were reproduced in at least two independent experiments ( $n = 2$ ). (A) SEC chromatogram of IbeA on a Superdex 200 Increase column under non-reducing conditions. Absorbance at 280 nm (blue curve) and at 450 nm (red curve) is reported. The fractions pooled to obtain the F1 and F2 forms are highlighted in purple and light blue, respectively. (B, C) Native MS spectra of the F1 (B) and F2 (C) forms of IbeA under non-reducing conditions. Charge states are indicated above each peak. Peaks corresponding to the different charged states are tripled for F1 but single for F2. (D) Liquid/chromatography coupled to electrospray ionization with time-of-flight (LC/ESI-TOF) MS deconvoluted spectra of the F1 (upper) and F2 (lower) forms of IbeA under non-reducing conditions. The oligomeric state is indicated above each peak (M: monomer; D: covalent dimer; T: covalent trimer; Q: covalent tetramer). (E) Same as panel A but under reducing conditions. (F, G) Native MS spectra of the F1 (F) and F2 (G) forms of IbeA under reducing conditions. Peaks corresponding to the different charged states are tripled for both F1 and F2. (H) LC/ESI-TOF MS deconvoluted spectra of the F1 (upper) and F2 (lower) forms of IbeA under reducing and denaturing conditions. Only monomers are present.

F2 form (Fig. 3C, lower spectrum). Importantly, this peak was not detected in the control sample containing only buffer (upper spectra). The flavin cofactors most commonly bound to proteins are flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). Given that FAD has a theoretical molecular weight of 785.55 Da and that the accuracy of the mass measurement by DIMS is 0.5  $m/z$  unit, we postulated that the observed peak at 786.16  $m/z$  corresponds to  $[\text{FAD}+\text{H}]^+$ . Under non-reducing conditions, an additional peak at 808.15  $m/z$  is also detected in the F1 form that could correspond to a  $[\text{FAD}+\text{Na}]^+$  adduct.

Looking more closely at the native MS spectra of F1, we observed that all peaks are tripled (Figs 2B,F and 3D,F, Table 1). Mass determination suggests the presence of distinct forms of IbeA, with a mass difference ranging from 774 to 798 Da between two consecutive peaks. We therefore concluded that the IbeA F1 samples contain three distinct species: (i) apo-dimers of IbeA (experimental mass: 99 202–99 207 Da; theoretical mass: 99 202 Da); (ii) IbeA dimers bound to one molecule of FAD (experimental mass: 99 989–100 000 Da; theoretical mass: 99 988 Da); and (iii) IbeA dimers carrying two molecules of FAD

(experimental mass: 100 770–100 780 Da; theoretical mass: 100 774 Da) (Table 1). Interestingly, the F2 form of IbeA obtained under non-reducing conditions contains no peak at 786.16 or 808.15  $m/z$  in the DIMS spectra (Fig. 3B, lower spectrum). Consistently, native MS peaks are all single (experimental mass of 99 201 Da; Figs 2C and 3E; Table 1). This indicates that under non-reducing conditions, the F2 form is an apo-homodimer of IbeA. This is in agreement with the SEC elution profile of IbeA under non-reducing conditions. Indeed, the F2 form does not seem to absorb light at 450 nm (Fig. 2A). On the contrary, a peak of absorbance at 450 nm aligns with the peak of absorbance at 280 nm for the F1 form, highlighting that this form corresponds to FAD-bound IbeA. Finally, under reducing conditions, the native MS spectrum of F2 also shows tripled peaks (Fig. 2G, Table 1). The reduced F2 form thus seems to contain all three forms of IbeA (apo-dimer, dimer with 1 FAD, and dimer with 2 FAD). However, as seen in Fig. 2E, the SEC elution peaks of the F1 and F2 forms are very close under reducing conditions. Therefore, the reduced F2 form may not be fully separated from the F1 form. No firm conclusion can thus be drawn on the FAD content of



**Fig. 3.** IbeA from AIEC is a flavoprotein that binds up to two FAD molecules per homodimer. MS spectra are from a single, representative experiment but results were reproduced in at least two independent experiments ( $n = 2$ ). (A) Bright yellow coloration of the IbeA sample after purification on nickel column (Ni-NTA). (B) DIMS spectra of the buffer (upper), F1 form (middle) and F2 form (lower) of IbeA under non-reducing conditions. A peak at 786.16  $m/z$ , corresponding to a FAD cofactor, is exclusively found in the F1 form. The peak at 808.15  $m/z$  is attributed to a FAD-Na adduct. The F2 form does not contain FAD cofactor. Signal at 787.33  $m/z$  is considered meaningless due to its large shift from expected FAD  $m/z$ . (C) Same as panel B but under reducing conditions. FAD is detected in the spectra of both the F1 and F2 forms. (D) Zoom on the peaks of charge state 18+, 19+, and 20+ in the native MS spectrum of the F1 form of IbeA under non-reducing conditions. All peaks are in fact tripled peaks, suggesting the presence of three distinct forms of IbeA. (E) Zoom on the peaks of charge state 18+, 19+, and 20+ in the native MS spectrum of the F2 form of IbeA under non-reducing conditions. All peaks are single, only one IbeA form is present (apo form). (F) Superimposition of the peak of charge state 19+ from the native MS spectra of the F1 (purple curve) and F2 (blue curve) forms of IbeA under non-reducing conditions, highlighting the different forms (apo/holo) of IbeA present in these samples. The F2 form only contains the apo-dimer (experimental mass 99 201 Da; theoretical mass: 99 202 Da) while the F1 form contains both the apo-dimer (experimental mass: 99 207 Da; theoretical mass: 99 202 Da), a dimer bound to one molecule of FAD (experimental mass: 99 989 Da; theoretical mass: 99 988 Da), and a dimer of IbeA carrying two FAD molecules (experimental mass: 100 770 Da; theoretical mass: 100 774 Da).

the F2 form under reducing conditions. Overall, these data demonstrate that IbeA is a FAD-binding flavoprotein that binds up to two FAD molecules per homodimer.

### Reducing conditions are required for optimal stability and FAD binding

We next investigated the effect of redox conditions on the stability and FAD-binding capacity of IbeA. First,

the stability of our recombinant IbeA over time was tested by pooling the whole peak encompassing both F1 and F2 forms and running it again on the SEC column after incubation on ice for 4 days. As shown in Fig. 4A, the protein in non-reducing buffer tends to form larger oligomers and aggregates over time, possibly through surface disulfide crosslinking between adjacent IbeA molecules. Furthermore, absorbance at 450 nm is almost completely lost, suggesting that the bound FAD is released over the incubation period. In

**Table 1.** Masses of the different IbeA species detected by native MS. Masses for the different forms (apo/holo) of dimeric IbeA, as extrapolated from the peaks corresponding to charged state 19+ are reported for the F1 and F2 forms, both under reducing and non-reducing conditions.

IbeA pool from SEC	Detected masses (Da)	Relative intensity	Putative composition	Expected mass (Da)
F1 form – non-red	99 207 ± 7	Important signal	Apo-IbeA dimer	99 202
	99 989 ± 6	Important signal	IbeA dimer + 1 FAD	99 988
	100 770 ± 4	Important signal	IbeA dimer + 2 FAD	100 774
F2 form – non-red	99 201 ± 6	Main signal	Apo-IbeA dimer	99 202
F1 form – red	99 202 ± 6	Low abundance	Apo-IbeA dimer	99 202
	100 000 ± 12	Low abundance	IbeA dimer + 1 FAD	99 988
	100 780 ± 8	Main signal	IbeA dimer + 2 FAD	100 774
F2 form – red	99 208 ± 7	Important signal	Apo-IbeA dimer	99 202
	99 999 ± 10	Low abundance	IbeA dimer + 1 FAD	99 988
	100 773 ± 5	Important signal	IbeA dimer + 2 FAD	100 774

contrast, the chromatogram obtained after 4-day incubation of the protein in reducing buffer does not show formation of large aggregates, one can only note a shift of the elution peak toward the F2 form, associated with a partial loss of absorbance at 450 nm (Fig. 4B). Reducing conditions thus appear to increase the stability of the protein by preventing its aggregation into large oligomers.

Estimation of the respective IbeA and FAD concentrations in the F1-F2 pool revealed an average stoichiometry of 0.35 to 0.45 molecule of FAD per molecule of IbeA (see [Materials and Methods](#) for calculation details). As native MS indicates that an IbeA : FAD stoichiometry of up to 2 : 2 can be reached, we next investigated whether more FAD could be bound by our sample, following incubation with exogenous FAD. As shown in Fig. 4C, no additional FAD can be captured by the IbeA homodimer in non-reducing conditions. In contrast, under reducing conditions, a dose-dependent increase in absorbance at both 280 and 450 nm is observed for the IbeA elution peak following 2 h incubation at 25 °C with different molar ratios of FAD, indicating binding of exogenous FAD to IbeA (Fig. 4D).

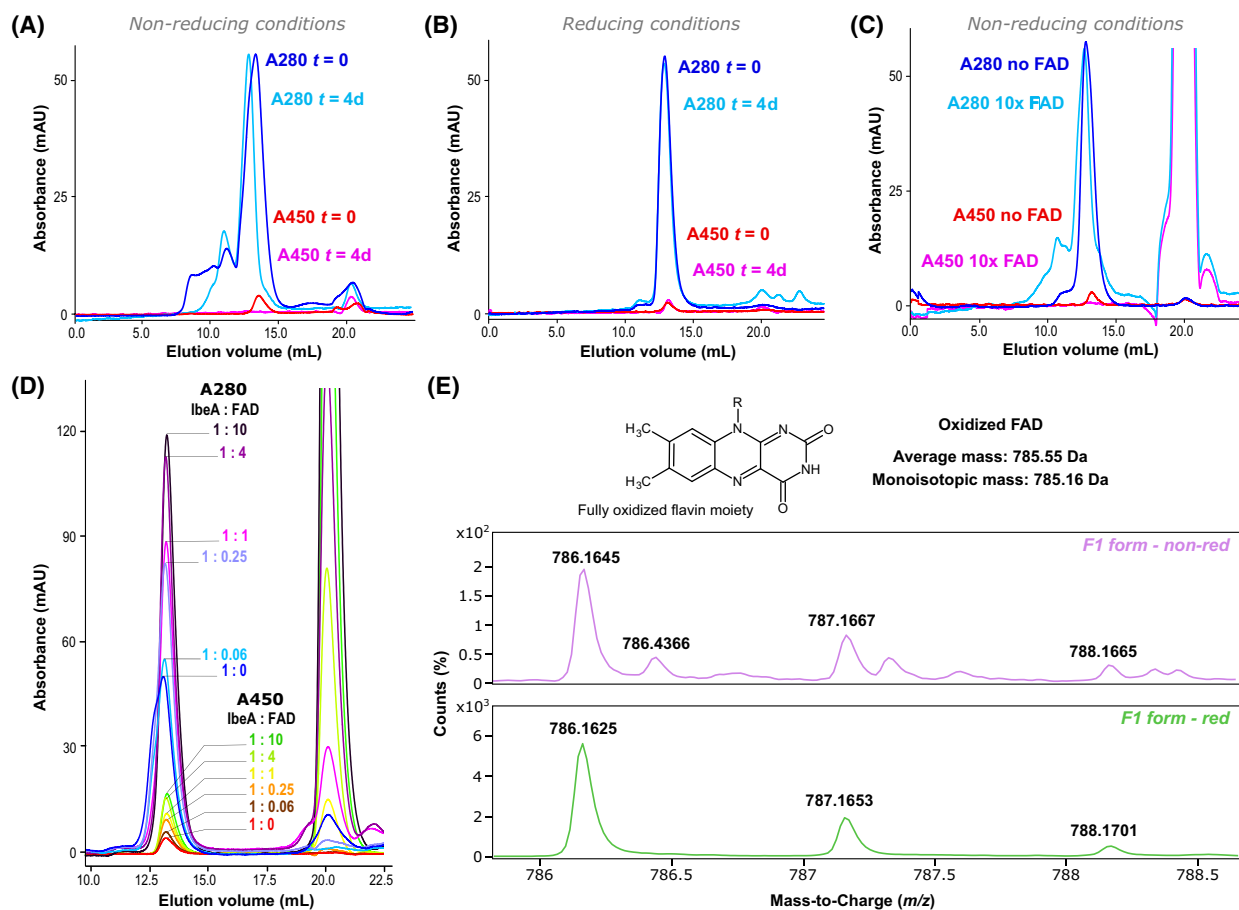
Reducing conditions are thus required for optimal stability and improved FAD-binding capacity of IbeA. The latter property could be linked to the redox state of FAD and/or to that of IbeA. When zooming on the FAD peak in the DIMS spectrum of the F1 form, we detected a predominant, singly-charged peak at 786.16 *m/z*, and two secondary peaks of lower abundance at 787.16 and 788.16 *m/z*, respectively (Fig. 4E). This suggests that the main form bound to IbeA is fully oxidized FAD (expected monoisotopic MS signal after protonation: 786.16 *m/z*), the two secondary

peaks reflecting the natural abundance of isotopes (e.g. <sup>2</sup>H, <sup>13</sup>C) within FAD.

### IbeA shares structural homology with FAD-dependent oxidoreductases

As no structural data are so far available on IbeA and since the current literature remains controversial on its biological role (membrane-embedded invasin versus soluble anti-oxidative factor), we used prediction software to model the tertiary fold of IbeA (Fig. 5). As shown in Fig. 5A, the AlphaFold (AF) structural model obtained for IbeA reveals a globular architecture with two domains separated by a hinge region that may provide flexibility between them. Domain A (blue cartoons) corresponds to amino acids Leu14 to Arg177 and Arg332 to Phe456. It contains ten  $\alpha$ -helices, among which four are only one-turn long, and five  $\beta$ -sheets, of which three contains two antiparallel  $\beta$ -strands while the two others contain four and five  $\beta$ -strands, respectively. Domain A forms a classical FAD-binding domain. It bears a typical Rossmann fold (Fig. 5B) that is generally involved in the binding of the nucleotide moiety of biological cofactors such as FAD or NADH. The topology of the Rossmann fold predicted for IbeA (Fig. 5B, lower part) closely resembles that of glutathione reductases (GRs) [36]. Helix  $\alpha$ 2 is however much shorter in IbeA than in canonical GRs, which makes it an integral part of the Rossmann fold, while in GRs it extends further away from the core motif [37]. A GxGxxG(x)<sub>17</sub>E consensus sequence found in all members of the GR family is also present in the IbeA region encompassing strands  $\beta$ 1,  $\beta$ 2 and helix  $\alpha$ 1, and ranges from Gly23 to Glu46. Domain B of IbeA (red cartoons, amino acids Glu178 to Ser331) is composed of an  $\alpha\beta$  sandwich with a

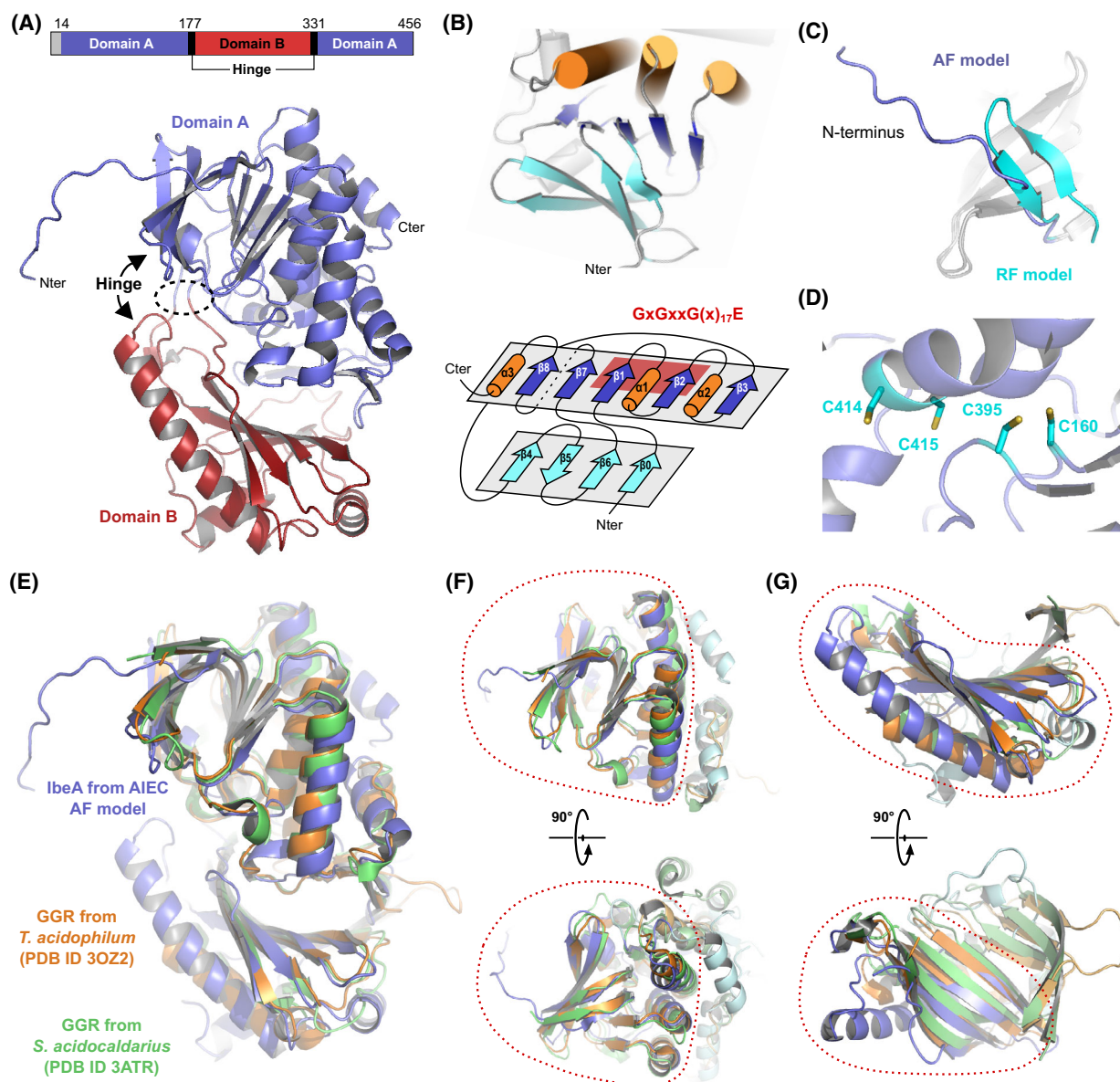




**Fig. 4.** Stability and FAD-properties of recombinant IbeA depending on the redox conditions. Chromatograms are from a single, representative experiment but results were reproduced in at least three independent experiments ( $n = 3$ ). (A) SEC analysis of the stability of recombinant IbeA over time under non-reducing conditions. The IbeA sample corresponding to the F1-F2 pool is rerun onto the SEC column either immediately after purification ( $t = 0$ ) or after 4-day incubation on ice ( $t = 4d$ ). Absorbance at both 280 and 450 nm is displayed. (B) Same as in panel A but under reducing conditions. (C) Capacity of recombinant IbeA to bind exogenous FAD under non-reducing conditions. The IbeA sample corresponding to the F1-F2 pool is rerun onto the SEC column following 2 h incubation at 25 °C with either buffer or a 10-fold molar excess of FAD. (D) Capacity of recombinant IbeA to bind exogenous FAD under reducing conditions. The IbeA sample corresponding to the F1-F2 pool is rerun onto the SEC column following 2 h incubation at 25 °C with either buffer or increasing concentrations of FAD (from 0.06 to 10-fold molar excess). A dose-dependent binding of exogenous FAD is observed. (E) Zoom on the DIMS spectra of the F1 form of IbeA under non-reducing (upper) or reducing (lower) conditions to identify the FAD form preferentially bound to IbeA. A main singly-charged signal is detected at 786.16  $m/z$ , corresponding to the fully oxidized form of FAD. The secondary peaks at 787.16 and 788.16  $m/z$  reflect the natural abundance of isotopes within FAD.

four-strand antiparallel  $\beta$ -sheet and a long  $\alpha$ -helix ( $\alpha_1\beta_4$  sandwich). The rest of the domain is predicted to be less structured and contains only three short  $\alpha$ -helices as secondary structure elements. The IbeA model predicted by RoseTTaFold (RF model) is highly similar to the AF model, with an overall root-mean-square deviation (RMSD) on  $C\alpha$  atoms of 1.4 Å, except for the first 15 N-terminal residues, which are relatively unstructured and adopt distinct conformations depending on the software used for prediction. Modeling with RoseTTaFold proposes these residues to fold into a

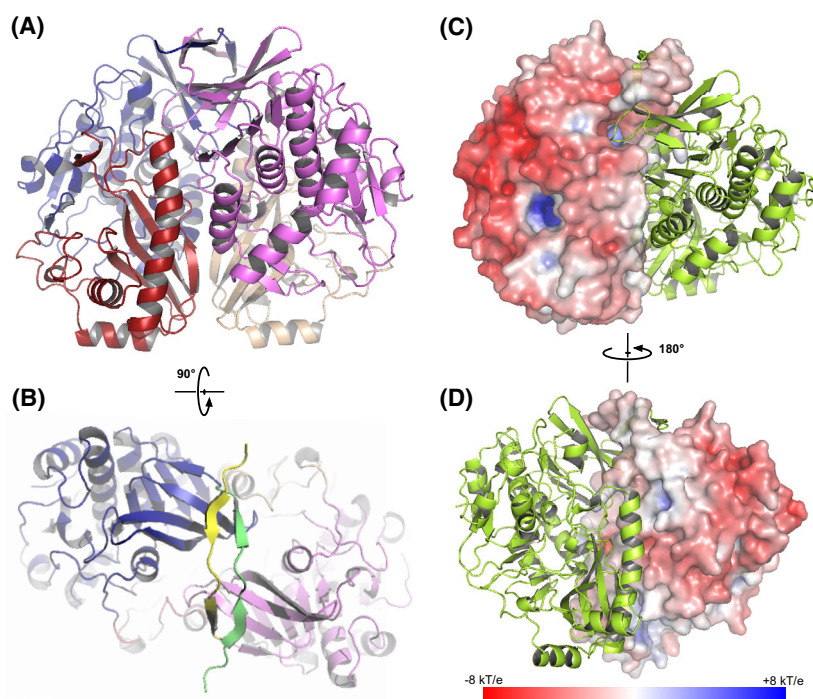
two-strand  $\beta$ -sheet while AF predicts an unstructured motif (Fig. 5C). We ruled out that this N-terminal region could be a signal peptide by analyzing the IbeA sequence through the online prediction server SIGNALP 6.0 [38]. Finally, IbeA contains ten cysteine residues, of which six are located on the surface of the proposed structural model. This may explain the propensity of IbeA to cluster through non-specific disulfide crosslinking under non-reducing conditions. Interestingly, the four other cysteines are buried within the IbeA structure and are all located in a nearby region (Fig. 5D).



**Fig. 5.** 3D-modeling of the IbeA protomer from AIEC reveals structural homology with FAD-dependent oxidoreductases. (A) IbeA 3D-model predicted by AlphaFold [74] (AF-model). IbeA is composed of two globular domains, domain A that corresponds to a classical FAD-binding domain, and domain B of unknown function. The two domains are connected by a hinge region that may provide relative flexibility. N-ter: N terminus; C-ter: C terminus. (B) Domain A of IbeA contains a classical Rossmann fold found in all FAD-dependent oxidoreductases from the GR family. It encompasses the GxGxxG(x)<sub>17</sub>E consensus motif involved in the stabilization of the adenine base of FAD. (C) Different models for the N terminus of IbeA suggest a highly flexible region (first 15 amino acids). RF model: RoseTTaFold model. (D) Buried deep within the IbeA structure, four cysteine residues are in close proximity and may form disulfide bridges or a putative ion binding site upon local conformational rearrangement. (E) Structural alignment of the AF-model of IbeA (blue cartoons) with the structure of the closest homologs identified with the Dali server [39]: the GGRs from *T. acidophilum* (orange cartoons; PDB: 3OZ2 [41]) and *S. acidocaldarius* (green cartoons; PDB: 3ATR [42]). AIEC: adherent invasive *E. coli*. (F) Zoom on domain A from the structural alignment displayed in panel E to highlight the high similarity between the Rossmann fold of IbeA and that of the two GGRs. (G) Zoom on domain B from the structural alignment displayed in panel E to highlight similarities and differences between the  $\alpha\beta$  sandwich of IbeA and that of the two GGRs.

Small structural rearrangements would be sufficient to bring them into close proximity and enable the formation of stabilizing intramolecular disulfide bridges.

Alternatively, they could participate in the binding of the FAD molecule or in the joint chelation of an internal cation not yet identified.

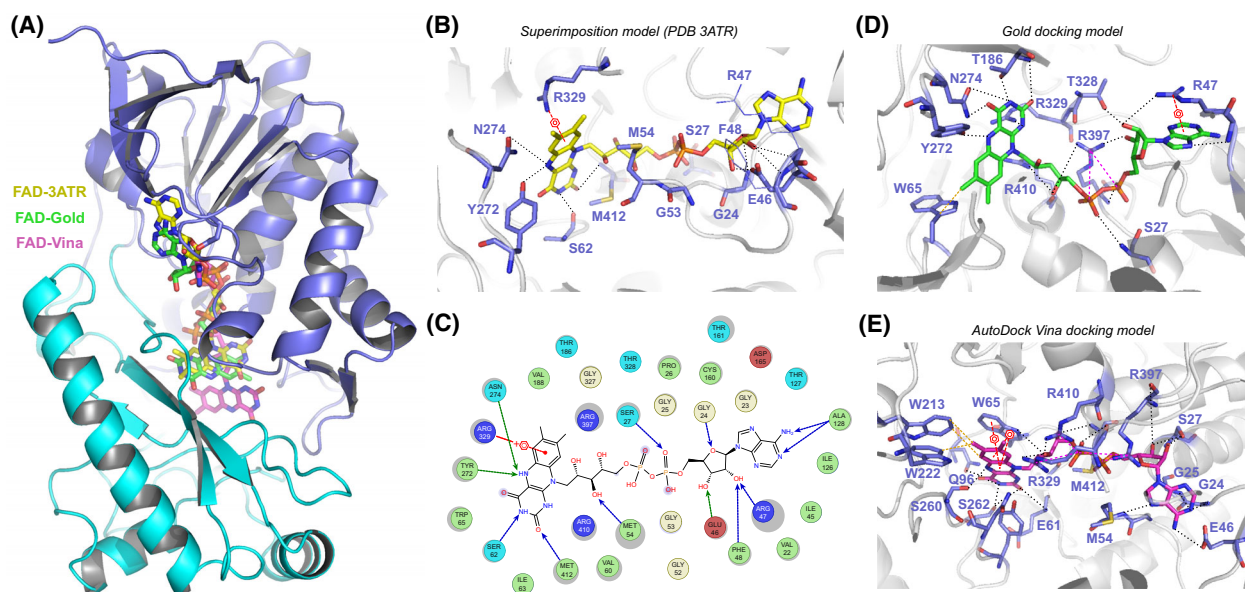


**Fig. 6.** 3D-modeling of the IbeA homodimer. (A) IbeA 3D-model predicted by AlphaFold-Multimer [43] (AF-model). The homodimer adopts a 2-fold symmetry along its longitudinal axis. (B) The flexible N termini are proposed to form an antiparallel  $\beta$ -sheet that would close up the top of the dimer interface. (C) Electrostatic potential surface representation for the IbeA dimer. These calculations reveal that the IbeA surface is highly negatively charged. (D) Same as panel C but with view on the dimer interface. The dimer interface is much more apolar except for the regions of close contact between the two IbeA protomers.

Comparison of the predicted structural model of IbeA with known structures from the Protein Data-Bank (PDB) using the Dali server [39] or PDBeFold [40] reveals strong structural homology with FAD-dependent oxidoreductases, as already hinted from sequence comparisons by Flechard *et al.* [34]. The strongest structural similarity is obtained with members of the geranylgeranyl reductase (GGR) family, in particular the geranylgeranyl glycerophospholipid reductase from *Thermoplasma acidophilum* (PDB: 3OZ2) [41] and the GGR from *Sulfolobus acidocaldarius* (PDB: 3ATR) [42]. The AF model of IbeA aligns well with the structures of both GGRs (Fig. 5E), with an overall RMSD on C $\alpha$  atoms of 2.42 and 2.90 Å between the IbeA model and the structures of *T. acidophilum* and *S. acidocaldarius* GGRs, respectively. In particular, the central core of the putative FAD-binding domain of IbeA encompassing the Rossmann fold superimposes well with the corresponding domains in both GGRs (Fig. 5F). Domain B of IbeA, which in FAD-oxidoreductases would correspond to the substrate-binding domain, does not present any obvious homology with known structures in the PDB. Nevertheless, it shares partial structural homology with the second domain of GGRs, in particular for the  $\alpha\beta$  sandwich region (Fig. 5G). The  $\alpha$ -helix is however longer in IbeA, and in the GGRs, two additional  $\beta$ -strands are provided by the FAD-binding domain to close up the

substrate binding pocket around the catalytic center for electron transfer. Our modeling analyses thus demonstrate that IbeA shares structural homology with FAD-dependent oxidoreductases, but the function of the second domain of IbeA and the nature of the putative substrate for oxidoreduction remain uncertain.

As our data show that IbeA is a dimeric protein, we also used AlphaFold-Multimer [43] to predict the quaternary architecture of the IbeA homodimer. The proposed model reveals a tightly packed dimeric assembly involving both domain A – domain A and domain B – domain B contacts, along the longitudinal side of the protein (Fig. 6A,B). As a result, the homodimer displays a 2-fold symmetry. Interestingly, the N termini of both IbeA protomers are proposed to strengthen the dimer packing by forming two antiparallel two-strand  $\beta$ -sheets that close up the interface at the top of the dimer (Fig. 6B). The surface of the IbeA dimer bears large patches of negatively charged residues as seen from electrostatic potential surface representation (Fig. 6C). In contrast, the protomer-protomer interface is much more apolar except for the regions of close contact, i.e. the  $\beta_4$ - $\beta_6$   $\beta$ -sheet and the  $\alpha_2$  helix from the Rossmann fold in domain A, and the long  $\alpha$  helix from the  $\alpha\beta$  sandwich in domain B (Fig. 6D). The interface area calculated by PISA [44] covers 3160 Å<sup>2</sup>, corresponding to 80 interfacing residues from each protomer, and with an overall solvation free energy



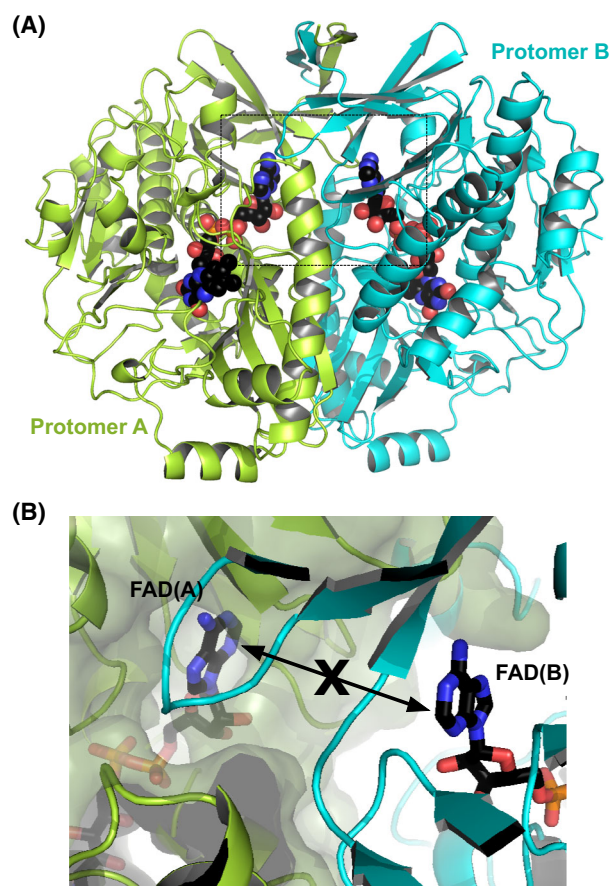
**Fig. 7.** Identification of the FAD binding pocket within IbeA. (A) Positioning of the FAD molecule within the IbeA protomer either through simple superimposition with the FAD-bound structure of the GGR from *S. acidocaldarius* (yellow sticks) in CoOt [79] or by use of the docking software GOLD (green sticks; [75]) or AutoDock VINA (pink sticks; [77]). (B) Key residues of IbeA proposed to interact with FAD using the simple superimposition model in CoOt [79]. Hydrogen bonds are shown with black dotted lines.  $\pi$ -stacking interactions are marked by red dotted lines with an aromatic ring. (C) Same as panel B but using a 2D ligand-protein interaction diagram generated with PLIP and the Lidia ligand builder from CoOt [79,80]. Residue color code: beige for apolar, green for hydrophobic, light blue for polar, dark blue for basic, dark red for acidic. Gray and blue circles surrounding the residues and FAD atoms, respectively, indicate the degree of exposure to the solvent (the larger the circle, the more exposed the residue/atom is). Blue and green arrows indicated hydrogen bonds via main and side chain atoms, respectively. Red arrows indicate  $\pi$ -stacking. (D, E) Key residues of IbeA proposed to interact with FAD using the GOLD [75] (D) or AutoDock VINA [77] (E) docking models. Hydrogen bonds are indicated with black dotted lines, hydrophobic interactions are highlighted by yellow dotted lines, pi-stacking interactions are marked by red dotted lines with an aromatic ring, and salt bridges are depicted with magenta dotted lines.

gain of  $-31.5 \text{ kcal}\cdot\text{mol}^{-1}$ , which gives good confidence in this dimer model.

### Identification of the FAD binding pocket within IbeA

To gain insights into the mode of binding of FAD by IbeA, we next generated a docking model of oxidized FAD within the AF-predicted 3D-structure of IbeA. Two docking software were used that gave slightly distinct models (Fig. 7A, green and pink sticks). In addition, we also employed manual docking by simple superimposition of the FAD molecule present in the FAD-bound structure of the GGR from *S. acidocaldarius* (Fig. 7A, yellow sticks). The overall FAD-binding pocket is globally the same in all models, despite subtle differences in the positioning of the FAD molecule. As expected from sequence and structural homology, FAD is located within domain A, with the adenine base being held in place by the Rossmann fold, while the flavin moiety is embedded in

a central pocket at the interface between domains A and B, consistent with what is observed in FAD-dependent oxidoreductases [36]. The docking model proposed by Gold places the FAD molecule at the same height within the IbeA central cavity as in the model obtained by simple superimposition with FAD-bound GGRs (Fig. 7B–D). The flavin moiety is however shifted by  $180^\circ$ , which leads the adenine base to be pushed further up toward the inner core of the Rossmann fold than in the superimposition model based on PDB entry 3ATR. In contrast, the docking model proposed by AutoDock VINA places the flavin moiety further down in the central cavity, closer to domain B, and with a stabilizing  $\pi$ - $\pi$  stacking with Trp65 (Fig. 7E). This alternative positioning however partly disrupts the proper anchoring of the adenine moiety within the Rossmann fold. Docking of FAD into our dimer model is consistent with a 2 : 2 FAD: IbeA complex, i.e. one molecule of FAD per each protomer (Fig. 8A). Whether cooperativity between the two protomers is required for optimal FAD binding



**Fig. 8.** Docking model of FAD into the IbeA homodimer. (A) Docking model with two FAD molecules docked per IbeA homodimer, one in each IbeA protomer (protomer A or B) using GOLD [75]. (B) Zoom on the closest region between the two FAD molecules, at the dimer interface. The adenine ring of each molecule is inserted in a pocket closed by  $\beta$ -sheets from the Rossmann fold, thus preventing direct connection between the two binding cavities within each protomer.

remains unknown, but the binding pockets from the two protomers do not appear to be connected in our dimer model, suggesting that each FAD molecule can bind independently of the other to each IbeA protomer (Fig. 8B).

Despite subtle differences, all docking models point out at a similar binding pocket lined up by the same residues potentially involved in the coordination of different parts of the FAD molecule (Fig. 7B–E). Key residues found to interact with FAD in all docking models include Glu46, the final glutamate from the conserved GxGxxG(x)<sub>17</sub>E motif within the Rossmann fold, and Arg47, which stabilize the adenosine moiety of FAD (adenine and/or ribose groups), possibly through a  $\pi$ -stacking interaction between with the adenine rings. Gly24 and Gly25, the second and third

residues from the consensus GxGxxG(x)<sub>17</sub>E sequence, also connect to the adenosine part through main chain hydrogen bonds. Ser27 contacts the hydroxyl groups of the ribose and the pyrophosphate moieties. Ser62 and Trp65 stabilize the isoalloxazine ring, one docking model even proposing a strong  $\pi$ -stacking between Trp65 and the flavin ring. Tyr272, Asn274, and Arg329 also interact with the polar groups of the isoalloxazine ring, while Arg397 and Arg410 may rather stabilize the hydroxyl groups from the pyrophosphate and D-ribitol moieties. Depending on the docking model, other residues are predicted to interact directly with FAD, including Phe48, Gly53, Met54, Glu61, Gln96, Ala128, Thr186, Trp213, Trp222, Ser260, Ser262, Thr328, and Met412, through hydrogen bonds, hydrophobic contacts or salt bridges. Finally, the four internal cysteines of IbeA described in Fig. 5D are also in close proximity to the FAD molecule. In particular, Cys160 and Cys415 could come within hydrogen bonding distance from the pyrophosphate groups, providing small conformational rearrangements.

Interestingly, quite a few tryptophan residues are located nearby the FAD molecule in both docking models. We took advantage of this feature to measure the affinity of the interaction between IbeA and FAD, with help of Trp fluorescence quenching experiments. As shown in Fig. 9A, the Trp fluorescence of IbeA is quenched in the presence of FAD, in a dose-dependent manner. Taking into account the primary and secondary inner filter effects (IFE) generated by FAD itself (due to primary absorbance of the excitation light at 280 nm and secondary absorbance of the Trp-emitted light between 300 and 400 nm by FAD; see [Materials and Methods](#) and Fig. 9B–E for calculation details), we could derive an apparent dissociation constant ( $K_{d-app}$ ) of  $127 \pm 46$  nM for the interaction of FAD with WT IbeA (Fig. 9F; Table 2), using a quadratic regression to account for FAD depletion (see [Materials and Methods](#)).

To validate the docking models, we next generated single- and double-point mutants (Ala substitution) for the residues lining up the putative FAD binding pocket. As shown in Table 2 and Fig. 10A, all mutated residues except R329 show significantly decreased affinity for FAD, the most destabilizing mutations being R397A and S62A-W65A. We also tried to mutate Glu46 and Arg47 into alanine but these yielded an instable protein that could not be purified to homogeneity. The moderate loss of affinity most probably reflects the large size of the binding pocket that a single mutation cannot compromise severely. Nevertheless, the effect of these mutations on the  $K_{d-app}$  values clearly demonstrates that these

residues are in close proximity to the FAD molecule within FAD-bound IbeA, thus validating the proposed binding pocket.

Flavin-dependent oxidoreductases generally display specificity toward a single flavin derivative. In our case, the structural homology with FAD-oxidoreductases points out at FAD as being the preferred flavin compound for IbeA. In agreement, no significant interaction between IbeA and FMN could be detected using Trp fluorescence quenching measurement (Fig. 10B; Table 2). We also tested whether IbeA could bind to NAD<sup>+</sup>/NADH or NADP<sup>+</sup>/NADPH, but no significant changes in Trp fluorescence could be detected in the presence of these nicotinamide cofactors (Fig. 10B; Table 2). This tends to suggest that they can probably not replace FAD in the binding pocket we identified. However, they could still bind to IbeA at a distinct site and/or be used by IbeA as substrates for the reduction of FAD. To test this hypothesis, we incubated our IbeA sample (F1 form) in the presence of exogenously added FAD and NADH or NADPH, under anaerobic conditions, and monitored the absorbance at 450 nm over time. As shown in Fig. 10C,D, no change in the absorbance at 450 nm is detected following 2 h or 6 h incubation in the presence of either nicotinamide cofactor, which demonstrates that in these *in vitro* conditions, neither NADH nor NADPH can be used by IbeA to reduce FAD.

### Conformational changes upon FAD binding

Our SEC experiments combined to MS analyses show that the apo-form of the IbeA homodimer (F2 form) has a distinct behavior on the SEC column. Indeed, it elutes at least one fraction earlier on the SEC column than FAD-bound IbeA (F1 form) (Fig. 11A–C). This differential behavior is also illustrated by the fact that addition of exogenous FAD triggers the disappearance of the shoulder peak corresponding to the F2 form on the SEC elution profile, yielding a single and homogeneous peak of absorbance at 280 nm corresponding to the sole F1 form and coinciding perfectly with the peak of absorbance at 450 nm (Fig. 11B). Native MS confirmed that there is no change in the oligomeric state of IbeA upon FAD binding since both the F1 and F2 forms are almost exclusively dimeric. Taken together, these observations suggest that substantial structural differences exist between the apo- and FAD-bound IbeA dimers and that FAD binding induces non-negligible conformational rearrangements within IbeA, either at the protomer or at the dimer level.

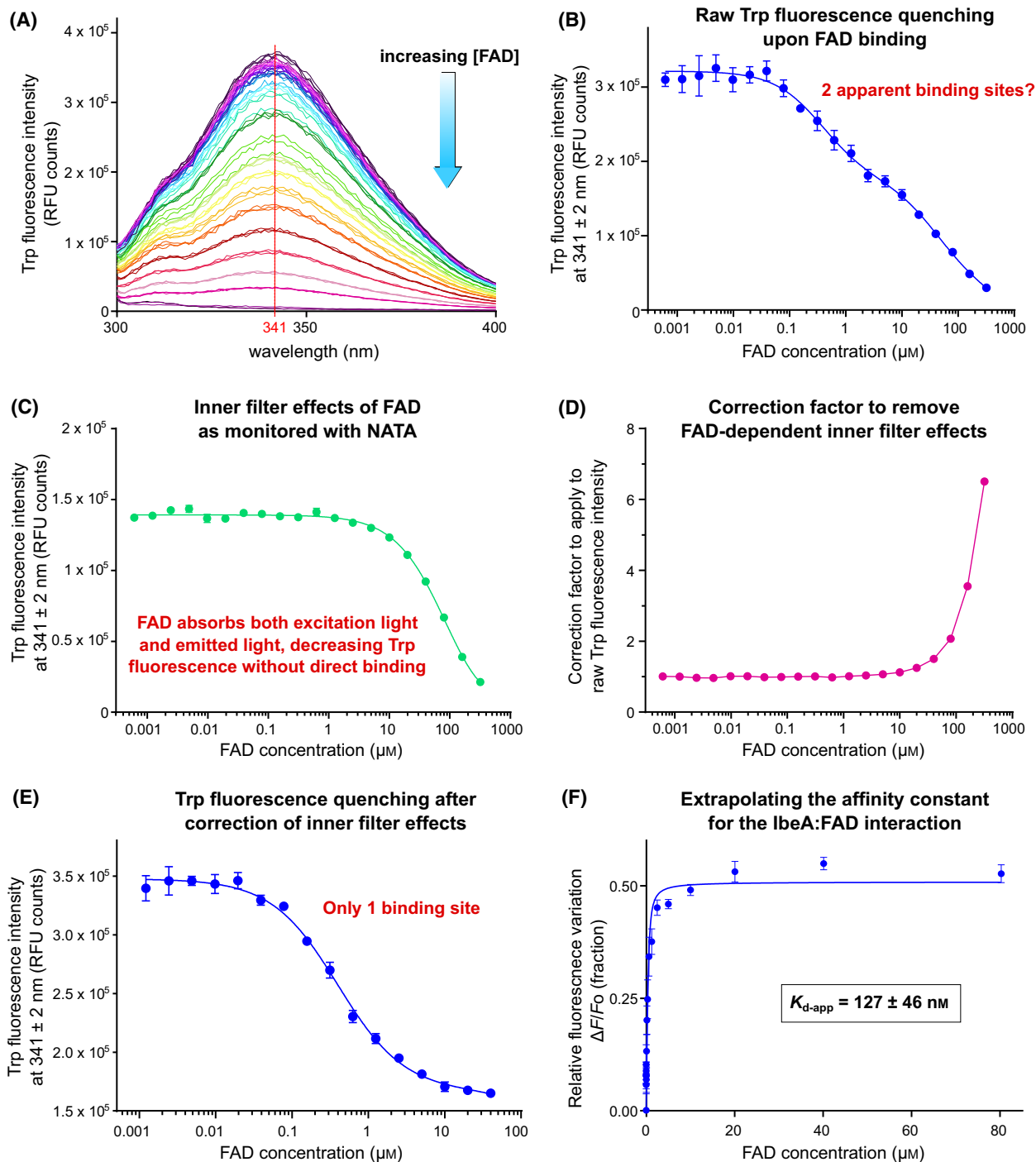
To gain insights into the possible conformational changes occurring upon FAD binding, we performed

MD simulations on the AF model of apo-IbeA and the FAD docking models generated previously. We first ran these simulations on the monomer models. As shown in Fig. 11D,E, no major differences are observed between the apo- and FAD-bound models, both for the overall radius of gyration and for the solvent-accessible surface area. We also performed these simulations on the homodimer models of apo- and FAD-bound IbeA. Docking of two FAD molecules within the IbeA homodimer using Gold yielded a positioning very similar to that obtained by simple superimposition with the FAD molecule present in the structure of GGR from *S. acidocaldarius*. We therefore only display the results of the MD simulations obtained with the Gold docking model. Although no major differences could be observed by direct visualization of the MD trajectories between the apo or FAD-bound dimer, a clear decrease of both radius of gyration and surface area values was observed throughout the duration of the 600 ns long simulations (Fig. 11F,G). These metrics suggest that presence of FAD promotes some compaction of the dimer or at least induces more constraints on the domains or loops motions. These movements remain however relatively small and do not seem to be sufficient to explain the large shift in elution volume observed in our SEC experiments between apo- and FAD-bound IbeA.

While MD simulations revealed no major conformational rearrangements of IbeA upon binding to FAD, smaller movements are nevertheless observed (Figs 12 and 13). First, while apo and FAD-bound IbeA dimers adopt a similar behavior during MD simulations, the convergence toward a stable conformational state is reached faster for the FAD-bound protein, indicating a higher conformational stability in presence of FAD (Fig. 12A,B). Furthermore, a certain degree of flexibility for the conformation and orientation of the isoalloxazine ring of FAD is observed during the simulations while the adenosine moiety stays more rigidly anchored within the Rossmann fold (Fig. 13A). These adjustments for the FAD molecule are accompanied by significant loop motions, in particular for loop Gly89-Ala104 in domain A and loop Thr206-Val226 in domain B (Fig. 12D). These movements are also observed in the apo-dimer and are even more important in the absence of FAD (Fig. 12C). Of interest, the following regions undergo significant rearrangements in both apo and FAD-bound IbeA during MD simulations: (a) the small  $\alpha$ -helix turn following helix  $\alpha$ 1 ( $\alpha$ 1' turn) and helix  $\alpha$ 2 from the Rossmann fold in domain A move into (apo protein, green cartoon) or away from (FAD-bound protein, blue cartoon) the FAD binding pocket, depending on whether

FAD is absent or present, respectively (Fig. 13B), suggesting that these motifs may act as a switch to stabilize the central cavity in the absence of ligand; (b) the loop encompassing residues Gly89 to Ala104, just below the region described in (a), moves toward (FAD-bound protein, blue cartoon) or away from

(apo protein, green cartoon) a putative entry channel, concomitantly with helix Lys196-Thr206 from domain B, possibly serving as a gate for ligand entry (Fig. 13C); (c) the loop encompassing residues Thr206 to Val226 within domain B also moves toward (apo protein, green cartoon) or away from (FAD-bound



**Fig. 9.** Binding of FAD to recombinant IbeA using Trp fluorescence quenching assay. Data were generated from three independent experiments, with technical triplicates for each experiment ( $n = 3$ ). (A) Trp fluorescence emission for the IbeA protein incubated with increasing concentrations of FAD recorded between 300 and 400 nm following excitation at 280 nm (bandwidth of the excitation: 5 nm; bandwidth of the emission: 10 nm; measurements every 1 nm). FAD concentrations varied between 6 nM and 320  $\mu$ M. Presence of FAD induces a dose-dependent quenching of the fluorescence, as monitored by the decreased fluorescence of the peak maximum at 341 nm. (B) Variation of the maximum Trp fluorescence intensity (at  $341 \pm 2$  nm) of IbeA upon incubation with increasing concentrations of FAD (reported with a log10 scale). Error bars represent the standard deviation (SD). The observed quenching seems to follow a biphasic sigmoidal curve, suggesting two non-equivalent binding sites (one high and one low affinity binding sites). (C) Variation of the maximum Trp fluorescence intensity (at  $341 \pm 2$  nm) of NATA (N-Acetyl-L-tryptophanamide, analog of Tryptophan) upon incubation with increasing concentrations of FAD (reported with a log10 scale). High concentrations of FAD induce a decrease of the NATA fluorescence although NATA does not bind to FAD. This decrease is due to primary and secondary IFE, i.e. FAD absorbs both the excitation light at 280 nm and the fluorescence emitted by the Trp analog. (D) The NATA fluorescence decrease induced by FAD allows to calculate a correction factor for each FAD concentration employed in the assay. The raw fluorescence measured in panel B must be multiplied by this correction factor to correct the IFE. (E) Variation of the maximum Trp fluorescence intensity (at  $341 \pm 2$  nm) of IbeA upon incubation with increasing concentrations of FAD (reported with a log10 scale), after correction of the IFE. Error bars represent the SD. This time, the quenching curve is monophasic, suggesting a single binding site as expected from the docking models. (F) Relative Trp fluorescence variation (at  $341 \pm 2$  nm) of IbeA upon incubation with increasing concentrations of FAD (reported with a linear scale). Error bars represent the SD. Using a quadratic regression to account for ligand depletion, an apparent dissociation constant ( $K_{d-app}$ ) of  $127 \pm 46$  nM can be extrapolated for the interaction of WT IbeA with FAD.

**Table 2.** Apparent dissociation constants ( $K_{d-app}$ ) for the interaction of IbeA (WT or mutant proteins) with various nucleotide cofactors.  $K_{d-app}$  values were calculated from 3 independent experiments using nonlinear regression with a quadratic eq. (SD: standard deviation). The statistical significance of the differences observed in  $K_{d-app}$  values was analyzed in GRAPHPAD PRISM using unpaired  $t$ -tests with Welch's correction.  $P$ -values are reported as follow: \*  $0.05 > P > 0.01$ ; \*\*  $0.01 > P > 0.001$ ; \*\*\*  $0.001 > P > 0.0001$ ; \*\*\*\*  $P < 0.0001$ ; ns: not significant; nr: not relevant.

IbeA protein	Ligand	$K_{d-app} \pm$ SD (nM)	$P$ -value
WT	FAD	$127 \pm 46$	nr
S27A	FAD	$744 \pm 167$	*
S62A-W65A	FAD	$987 \pm 206$	*
Y272A-N274A	FAD	$536 \pm 73$	****
R329A	FAD	$210 \pm 60$	ns
R397A	FAD	$2107 \pm 339$	**
R410A-N411A	FAD	$422 \pm 76$	***
WT	FMN	No binding	nr
WT	NAD	No binding	nr
WT	NADP	No binding	nr

protein, blue cartoon) a second, putative entry channel (Fig. 13D), in an opposite manner to the movement described for loop Gly89-Ala104. Whether the movements observed in MD simulations may represent genuine conformational states adopted by the IbeA protein throughout the oxidoreduction process remains to be determined.

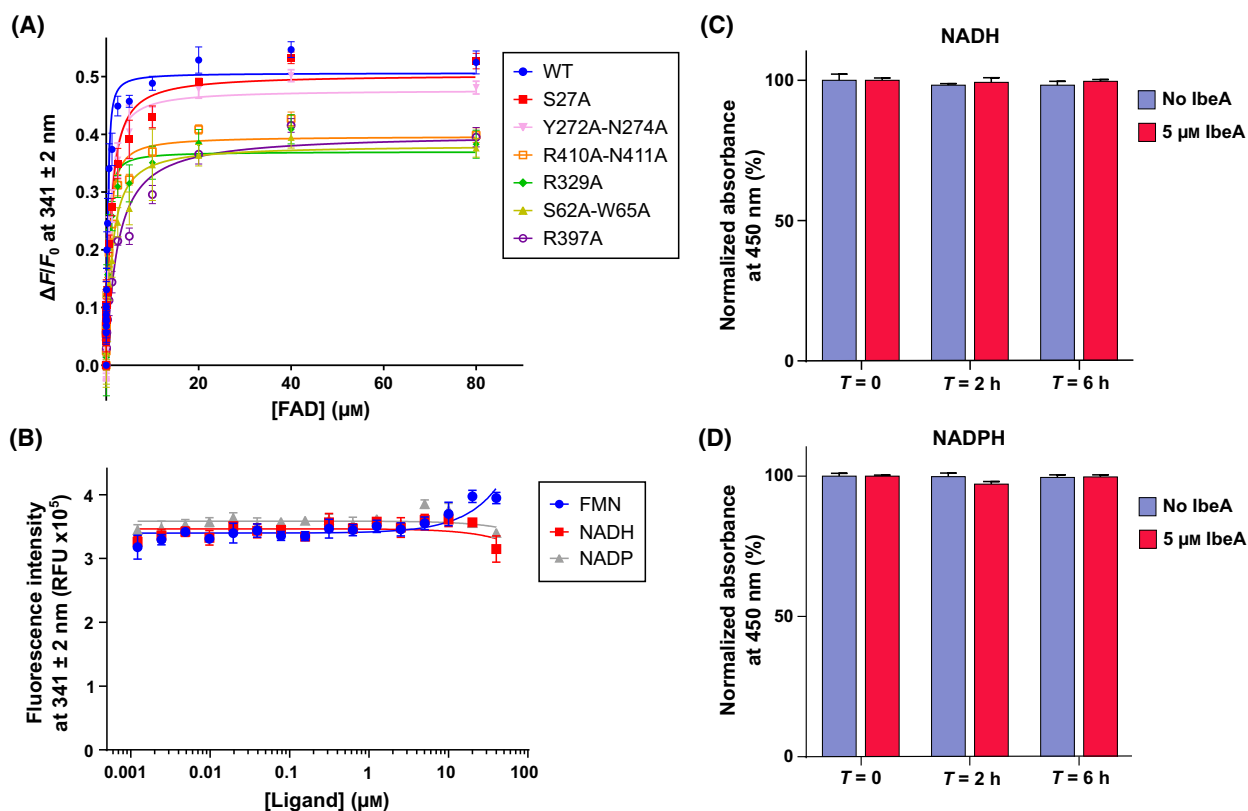
## Discussion

In this study, we provided the first biochemical characterization of the full-length IbeA protein from AIEC,

demonstrating that it behaves as a fully soluble factor expressed in the cytosol of *E. coli*. Using a combination of SEC, MS, and Trp fluorescence quenching assays, we showed that IbeA is a homodimeric flavo-protein that preferentially binds to FAD, with optimal stability and binding capacity under reducing conditions. We also reported a first, predictive 3D-structural model for the protein, revealing strong structural homology with FAD-dependent oxidoreductases. Thanks to *in silico* docking and mutation-based validation, we identified the FAD binding pocket, and using MD simulations, we proposed conformational rearrangements of the IbeA protomers upon FAD binding. Finally, our SEC experiments highlighted substantial conformational differences between the apo- and FAD-bound forms of dimeric IbeA that remain yet to be characterized at the atomic level.

While the role of IbeA as a genuine virulence factor in various pathogenic bacteria has been firmly established since its discovery in 1995 [24,32,35], its exact biological function remains controversial. Initial reports on the protein from the NMEC strain K1 described an insoluble protein, present in inclusion bodies during heterologous expression in bacteria [24–26]. This observation, together with the identified role of IbeA in promoting NMEC crossing through the blood–brain barrier, prompted the authors to categorize the protein as a membrane-embedded invasin possibly resembling outer membrane porins [26]. On the other hand, more recent work by Flechard *et al.* [34] proposed that the IbeA protein from APEC strain BEN2908 is a cytosolic protein involved in the resistance to  $H_2O_2$ -mediated oxidative stress and hinted for



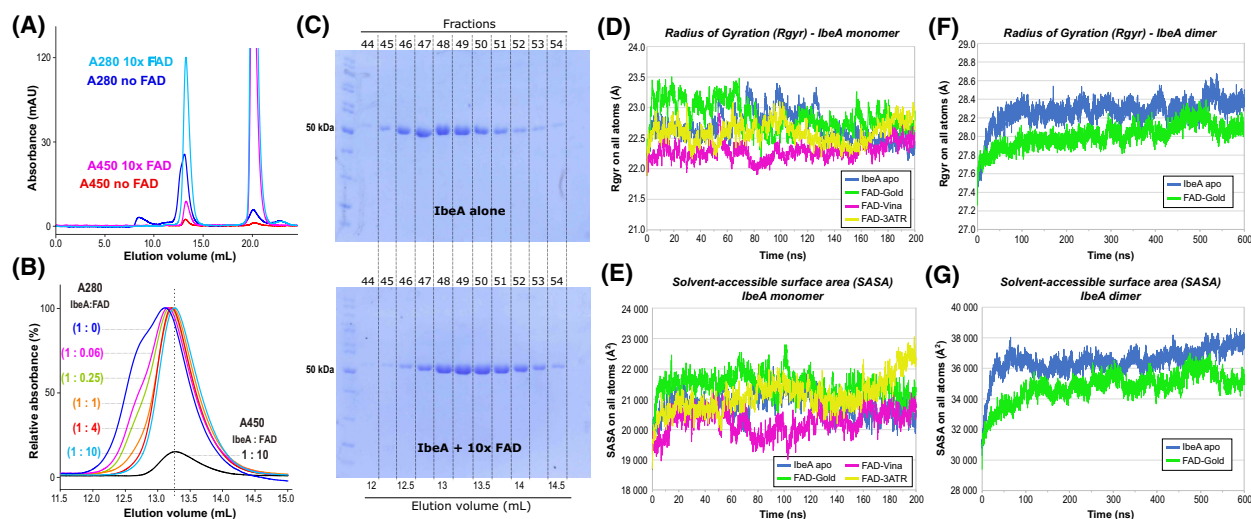


**Fig. 10.** Validation of the FAD binding pocket through mutational analysis and evaluation of nicotinamide cofactors as possible substrates of IbeA. (A) Trp fluorescence quenching assay to measure the interaction between WT or mutant IbeA proteins and FAD. The relative fluorescence variation ( $\Delta F/F_0$ ) is reported as a function of the FAD concentration. The data points arise from 3 independent measurements ( $n = 3$ ), with technical triplicates for each measurement. Error bars represent the standard deviation (SD). The apparent dissociation constants were calculated using a quadratic regression in GRAPHPAD PRISM. The fitting curves are also displayed. (B) Trp fluorescence quenching assay to measure the interaction between WT IbeA and FMN, nicotinamide adenine dinucleotide (NADH) or nicotinamide adenine dinucleotide phosphate (NADP). The fluorescence intensity at  $341 \text{ nm} \pm 2 \text{ nm}$ , corrected for IFE, is reported as a function of the FAD concentration (log10 scale). The data points arise from 3 independent measurements ( $n = 3$ ), with technical triplicates for each measurement. Error bars represent the SD. No significant variation of the fluorescence is observed upon addition of increasing ligand. (C) IbeA activity assay in the presence of NADH. Variation of the absorbance at  $450 \text{ nm}$  upon 2 h or 6 h incubation of IbeA ( $5 \mu\text{M}$ ; red histograms) or control buffer without IbeA (blue histograms) in the presence of exogenous FAD and NADH. The data points arise from 3 independent measurements ( $n = 3$ ) and error bars represent the SD. No decrease in absorbance at  $450 \text{ nm}$  is observed over time, demonstrating that FAD is not reduced in the presence of NADH. (D) Same as in panel C but in the presence of NADPH. No decrease in absorbance at  $450 \text{ nm}$  is observed over time, demonstrating that FAD is not reduced either in the presence of NADPH.

the first time a possible resemblance with FAD-oxidoreductases. Our data clearly corroborate this second hypothesis and confirm that IbeA is not an integral membrane protein. We cannot rule out that the proteins from the different *E. coli* strains adopt distinct 3D-architectures linked to distinct physiological functions. This hypothesis seems however unlikely as the proteins only differ by one or two amino acids over 456, and modeling of IbeA from NMEC and APEC with AlphaFold gave similar 3D-prediction models as for IbeA from AIEC (data not shown).

Although our data and those on the protein from APEC give strong evidence for a fully soluble protein,

Cieza *et al.* [35] reported a potential association of IbeA with the AIEC cell membrane. One possible explanation may be that while not directly embedded inside the membrane, IbeA could interact either constitutively or transiently with a membrane-bound binding partner. Interestingly, the *ibeA* gene is present on a genomic locus unique to extraintestinal pathogenic *E. coli* (GimA locus) and highly involved in the pathogenicity of these strains [45]. This locus encodes several virulence factors, of which some may be transmembranar. In particular, *ibeA* is part of a 3-gene operon (*ibeRAT*) that encodes a putative transmembrane transporter, IbeT, whose function seems to be tightly



**Fig. 11.** Compaction of IbeA upon FAD binding. Chromatograms and SDS/PAGE gels are from a single, representative experiment but results were reproduced in at least three independent experiments ( $n = 3$ ). (A) Comparison of the SEC elution profile of IbeA (F1-F2 pool) alone or incubated with a 10-fold molar excess of FAD under reducing conditions. Upon FAD binding, the absorbances at both 280 and 450 nm increase for the IbeA elution peak. Furthermore, the peak of absorbance at 280 nm corresponding to IbeA sharpens and aligns with the peak of absorbance at 450 nm, while the shoulder containing the F2 form disappears. (B) Zoom on the peak of absorbance at 280 nm corresponding to the IbeA homodimer from the experiment described in Fig. 3D, with all chromatograms scaled to the same height to visualize the dose-dependent sharpening and shift toward later elution volume of the peak. (C) SDS/PAGE analysis of the SEC experiments depicted in panel A. Fractions corresponding to elution volume 12 to 15 mL were loaded on the gel, for the chromatogram of IbeA alone (upper gel) or for the chromatogram of IbeA incubated with a 10-fold molar excess of FAD (lower gel). In the presence of saturating amounts of FAD, the elution of IbeA shift by at least one fraction, suggesting a more compact form for the FAD-bound IbeA homodimer. (D) Radius of gyration calculated using the *vMD* package [85] for apo- or FAD-bound IbeA (monomeric) from MD simulations. (E) Solvent-accessible surface area calculated using the *vMD* package [85] for apo- or FAD-bound IbeA (monomeric) following MD simulations. (F) Radius of gyration calculated using the *vMD* package [85] for apo- or FAD-bound IbeA (dimeric) from MD simulations. (G) Solvent-accessible surface area calculated using the *vMD* package [85] for apo- or FAD-bound IbeA (dimeric) from MD simulations.

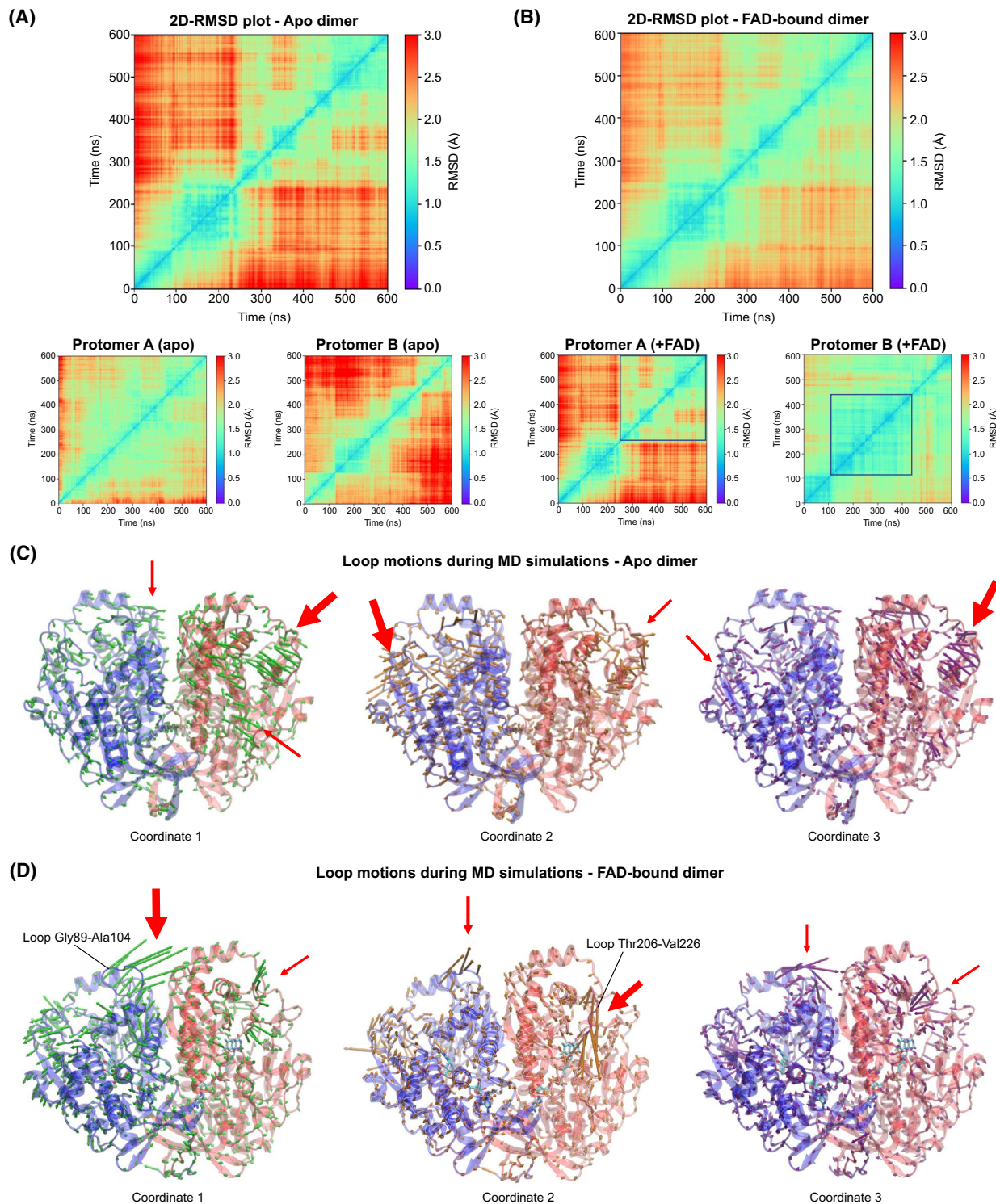
connected to that of IbeA [46]. Whether IbeA interacts with this transporter, or other membrane proteins absent in non-pathogenic *E. coli*, remains to be elucidated to specify the *in vivo* subcellular localization of the protein. Another possibility would be that IbeA has various functions throughout the life cycle of the bacteria. Hence, it could first be exposed on the bacterial surface, in order to interact with host cell surface receptors such as the ones identified on brain endothelial cells by the Huang group [27,28,30], thereby facilitating bacterial invasion. IbeA could then be released from the membrane and/or secreted within invaded host cells to achieve other functions possibly related to resistance to oxidative stress [34]. Interestingly, all the strains expressing IbeA are categorized as translocating *E. coli* (TEC) and recent studies have shown that the GimA locus is a hallmark of TEC strains [47]. Although the mechanisms of translocation are not all fully characterized, genomic studies showed that AIEC possess both Type IV and Type VI secretion systems (T4SS and T6SS) that are important for their pathogenicity [48–50]. Furthermore, in APEC strain K1,

mutations in T6SS2 result in a decreased expression of IbeA, thus establishing a clear link between IbeA and the secretion machinery of pathogenic *E. coli* [51].

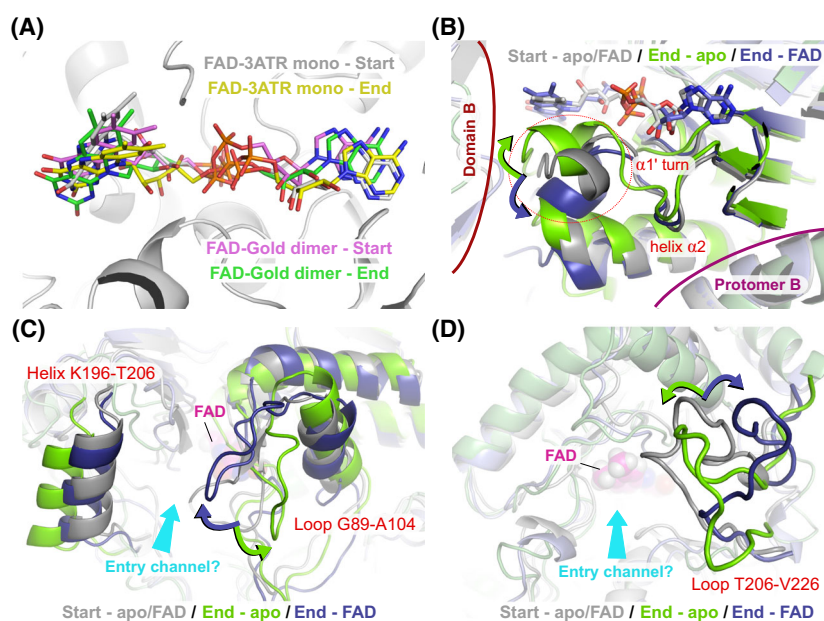
Our modeling studies show without ambiguity that IbeA shares strong structural homology with FAD-dependent oxidoreductases. The accuracy of our models is ascertained by the high confidence indexes we obtained for all our structure predictions. The predicted 3D-models propose IbeA to contain 32%  $\alpha$ -helices and 24%  $\beta$ -sheets. Using circular dichroism, Mendu *et al.* [26] previously reported that the IbeA protein from NMEC would bear 30%  $\alpha$ -helices and 34%  $\beta$ -sheets. Although the content of  $\beta$ -sheet elements is slightly lower in our model, it is still well in line with these values, taking into account the two different purification protocols (native purification versus *in vitro* refolding and subsequent solubilization with detergent). The two-domain organization we describe for IbeA is typical of FAD-oxidoreductases, with an N-terminal flavin-binding domain specific for FAD and a C-terminal domain resembling the substrate domain of known reductases. Consistently with the

literature [52], IbeA does not bind to FMN. It shows specificity toward FAD, bearing consensus sequence motifs involved in the chelation of this flavin derivative in GRs [36]. We also demonstrate that IbeA does

not use NADH or NADPH as substrates to reduce FAD, at least not in our *in vitro* activity assay. While many FAD-oxidoreductases can use nicotinamide cofactors as substrates for the reduction of FAD,



**Fig. 12.** Conformational stability and loop motions for IbeA dimer (apo or FAD-bound) during MD simulations. (A) 2D-RMSD plot for apo IbeA dimer during MD simulations, generated with the MDAnalysis PYTHON package [86]. The globally high RMSD values on backbone atoms (close to 3 Å) show that the protein dynamics were unstable before reaching a stable conformation (after 225 ns). (B) 2D-RMSD plot for FAD-bound IbeA dimer (Gold docking model) during MD simulations, generated with the MDAnalysis PYTHON package [86]. Due to lower RMSD values, the protein dynamics are more stable and conformational stability is reached faster than for the apo dimer (squares indicate areas for which IbeA has the same conformation). (C) Loop motions observed for the apo dimer during MD simulations computed from covariance analysis (principal component analysis/essential dynamics calculated using the PRODY package [87]) to identify correlated motions (three first components or coordinate 1 to 3). Red arrows indicate the regions where large movements occur, the width of the arrow is proportional to the amplitude of the movement. (D) Same as panel C but for the FAD-bound dimer. In both simulations, the loop with largest movements are loop Gly89-Ala104 in domain A and loop Thr206-Val226 in domain B. Overall, loop motions in the apo dimer are more important than in the FAD-bound dimer.



**Fig. 13.** Detailed conformational changes occurring within IbeA upon FAD binding. (A) Comparison of the position of the FAD molecule within the FAD binding pocket at the beginning and at the end of the MD simulations, for both monomeric and dimeric models of IbeA. The monomeric model used for comparison is the one based on simple superimposition of the FAD molecule from PDB entry 3ATR while the dimeric model is derived from docking with Gold. In both cases, the position of the adenine moiety varies only slightly between the start and the end of the simulation, while the position of the isoalloxazine ring is more flexible. (B) Comparison of the movements of the small  $\alpha_1'$  helix turn and helix  $\alpha_2$  during MD simulations between the apo dimer model (gray versus green cartoon) and the dimer model docked with two FAD molecules (gray versus blue cartoon). In the absence of FAD, these two helices move toward the empty FAD binding pocket, while they move away when FAD is present. (C) Comparison of the movements of the Gly89-Ala104 loop in domain A and helix Lys196-Thr206 in domain B during MD simulations between the apo dimer model (gray versus green cartoon) and the dimer model docked with two FAD molecules (gray versus blue cartoon). In the absence of FAD, the helix and loop move away from each other, opening a possible entry channel for ligand within each IbeA protomer. (D) Comparison of the movements of the Thr206-Val226 loop in domain B during MD simulations between the apo dimer model (gray versus green cartoon) and the dimer model docked with two FAD molecules (gray versus blue cartoon). In the presence of FAD, the loop moves away from the protein, opening a possible second entry channel for ligand within each IbeA protomer (substrate entry channel?).

including several members of the GGR family, other do not function with NADH or NADPH. This is for example the case for the menaquinone-specific prenyl reductase, a GGR from *Archaeoglobus fulgidus* [53]. Thus, it is possible that IbeA also displays an atypical mechanism of oxidoreduction that would not rely on nicotinamide cofactors.

The dimeric organization of IbeA is not uncommon for FAD-dependent oxidoreductases, including the well-studied GR family [54]. Dimerization of these enzymes has been proposed to increase their stability, their cooperativity for substrate binding, and/or their catalytic efficiency [55–57]. Many different types of quaternary architectures exist for dimeric oxidoreductases.

Despite extensive efforts, we could not find enzymes displaying a similar inter-subunit interface as the one depicted in the IbeA dimer model. However, the high strength of interaction between the two protomers calculated for our model give us strong confidence that this quaternary organization may correspond to the physiological one. The impact of this dimerization on IbeA function remains yet to be investigated.

The close resemblance of IbeA with GGRs, including domain B, is intriguing. These enzymes are known to hydrogenate polyprenyl substrates that are used for lipid biosynthesis in archaea [58] or chlorophyll biosynthesis in plants [59]. Although geranylgeranyl phospholipids are not commonly found in bacteria, this could hint for an IbeA substrate related to lipid metabolism. On the other hand, Xu *et al.* [60] identified a weak sequence homology between the region of IbeA encompassing residues 229–343 (roughly corresponding to the core of domain B in our model) and the Phox homology domain of cargo proteins, a phosphoinositide-binding motif involved in the targeting of proteins to cell membranes [61] and originally identified in the p40 and p47 subunits of the NADPH oxidase [62]. The resemblance between IbeA domain B and these different motifs remains low and cannot be used to predict a clear biological function or a possible substrate for the redox reaction. Overall, these observations suggest that although it has no direct membrane anchor, IbeA may be linked to membrane-related processes.

Taken together, our findings bring novel clues on the biological function of IbeA and help progressing toward understanding its role as a virulence factor in pathogenic *E. coli* by demonstrating that it most probably behaves as a FAD-dependent oxidoreductase. IbeA-expressing bacteria are known to invade and survive within host cells. AIEC, in particular, can reside within the phagolysosomal compartment of human macrophages, where they would encounter harsh acidic and oxidative conditions [14,15]. The putative redox function of IbeA may be employed to counteract host stresses in these intracellular compartments, in line with the role attributed to IbeA in enhanced survival and replication within host cells. More investigations will be required to fully address the biological function of this virulence factor and identify its physiological substrate and the redox process it may orchestrate.

## Materials and methods

### Bacterial strains and chemical reagents

*E. coli* BL21(DE3) cells were purchased from Invitrogen (Thermo Fisher Scientific, Waltham, MA, USA). AIEC LF82 reference strain was a generous gift from Nicolas Barnich

(M2iSH, Clermont Auvergne University, France). FAD, beta-nicotinamide adenine dinucleotide (NAD/NADH), and beta-nicotinamide adenine dinucleotide phosphate (NADP/NADPH) were purchased as disodium salts from Merck (Darmstadt, Germany). N-Acetyl-L-tryptophanamide (NATA) was purchased from Merck. Riboflavin 5'-monophosphate (FMN) was purchased as sodium salt from Cayman Chemical (Ann Arbor, MI, USA).

### Expression and purification of recombinant IbeA from AIEC

The IbeA full-length ORF was amplified from AIEC LF82 genomic DNA with the following forward and reverse primers, respectively:

5' CCCACTACTGAGAATCTTTATTTTCAGGGCAT GGAATTTTATCTGGAACCCGC 3' and 5' GAGCTCG AATTCCGGATCCGGTACCTTAAAAGACTTTTACGCCATTTTGC 3'. The resulting PCR fragment was then cloned in frame after the sequence coding for the Tobacco Etch Virus (TEV) protease cleavage site in pETM11 vector (EMBL vector collection), using restriction-free cloning [63]. Point mutants of the putative FAD-binding pocket were obtained by PCR-based site-directed mutagenesis using High Fidelity Phusion DNA Polymerase (New England Biolabs, Ipswich, MA, USA) and anti-complementary oligonucleotides bearing the mutation to introduce. The resulting constructs were used to transform heat-competent *E. coli* BL21 (DE3). After selection on LB agar plates supplemented with 50  $\mu\text{g}\cdot\text{mL}^{-1}$  of kanamycin, transformed cells were grown at 37 °C for 5 h, until they reached exponential phase of growth. Protein expression was then induced overnight at 18 °C by adding 1 mM IPTG in the culture medium. The next day, cells were harvested by centrifugation at 6000 *g* and disrupted by sonication. After clarification by centrifugation, the lysate containing soluble proteins was loaded onto a 5 mL HisTrap Ni-column (Cytiva, Marlborough, MA, USA) equilibrated with 40 mM Tris-HCl pH 7.5, 300 mM NaCl, 30 mM imidazole, 1 mM phenylmethylsulfonyl fluoride for purification in non-reducing conditions or with the same buffer supplemented with 10 mM  $\beta$ -mercaptoethanol for purification in reducing conditions. Non-specifically bound contaminants were removed by performing a high salt wash with a buffer containing 1 M NaCl. The protein of interest was finally eluted with a buffer containing 500 mM imidazole. The N-terminal polyhistidine tag was then removed overnight by incubating the sample at 4 °C with home-made recombinant TEV (rTEV) protease [64], using a protein to rTEV protease mass ratio of 50 : 1. A counter-purification was then performed on the His-Trap Ni-column to separate the cleaved IbeA protein, eluting in the flow-through, from the His<sub>6</sub>-tag and His-tagged rTEV protease that remained bound to the column. In a final step, the protein was purified by size-exclusion chromatography

(SEC) on a 24 mL Superdex 200 Increase 10/300 GL column (Cytiva) equilibrated with Buffer 1 (20 mM Tris-HCl pH 7.5, 150 mM NaCl) or Buffer 2 (same as Buffer 1 but with 10 mM  $\beta$ -mercaptoethanol). Fractions of interest were analyzed by SDS/PAGE, pooled and flash-frozen in liquid nitrogen for storage at  $-80^{\circ}\text{C}$ .

### Subcellular localization of recombinantly expressed IbeA from AIEC

Following expression of His<sub>6</sub>-tagged IbeA in BL21(DE3) cells as described above, the bacterial pellet from a 1.5 L culture was resuspended in 40 mL of 50 mM Tris-HCl pH 7.5, 300 mM NaCl and the cells were disrupted by sonication. The lysate was first subjected to low speed centrifugation (30 min at 30 000 *g* and 4  $^{\circ}\text{C}$ ). The resulting supernatant (S1) was then separated from the pellet corresponding to the insoluble fraction (cell debris + inclusion bodies; P1) and subjected to high speed centrifugation (2 h at 200 000 *g* and 4  $^{\circ}\text{C}$ ), in order to isolate the membrane fraction (recovered in the second pellet, P2) and the soluble fraction containing both cytosolic and periplasmic proteins (second supernatant, S2). To prepare a periplasmic extract, the bacterial pellet from a 1.5 L culture of IbeA-expressing BL21(DE3) cells was gently resuspended in ice-cold buffer containing 20 mM Tris-HCl pH 7.5, 5 mM EDTA, 25% (w/v) sucrose (4 mL of buffer/g of bacterial pellet), without vortexing, to avoid mechanical disruption of the cell membrane. The bacterial suspension was incubated on ice for 15 min and then centrifuged for 20 min at 4000 *g* and 4  $^{\circ}\text{C}$ . The pellet was separated from the supernatant and quickly resuspended in ice cold buffer containing 5 mM magnesium chloride (4 mL of buffer/g of bacterial pellet). The new suspension was incubated on ice for 1 h, to allow for outer membrane disruption through osmotic shock, and then centrifuged for 30 min at 30 000 *g* and 4  $^{\circ}\text{C}$ . The resulting supernatant (S3) corresponds to the periplasmic extract while the pellet (P3) contains the intact inner cell, including cytosolic and inner membrane-bound proteins. Fractions of interest were analyzed by SDS/PAGE using PageRuler™ Unstained Protein Ladder (Thermo Fisher Scientific) as molecular weight marker.

### Native mass spectrometry (native MS)

Prior to native MS analysis, buffer of the IbeA samples was exchanged to 250 mM ammonium acetate pH 7.0 (plus 10 mM dithiothreitol for reducing conditions) through SEC on a Superdex 200 Increase column. The samples were then analyzed by native MS as previously described [65–68]. Protein ions were generated using a nanoflow electrospray (nano-ESI) source. Nanoflow platinum-coated borosilicate electrospray capillaries were bought from Thermo Electron SAS (Courtaboeuf, France). MS analyses were carried out on a quadrupole time-of-flight mass spectrometer (Q-TOF Ultima, Waters Corporation, Manchester, UK). The instrument was modified for the detection of high masses [69,70].

The following instrumental parameters were used: capillary voltage = 1.2–1.3 kV, cone potential = 40 V, RF lens-1 potential = 40 V, RF lens-2 potential = 1 V, aperture-1 potential = 0 V, collision energy = 30–140 V, and microchannel plate (MCP) = 1900 V. All mass spectra were calibrated externally using a solution of Cesium Iodide (6 mg·mL<sup>-1</sup> in 50% isopropanol) and were processed with the MASSLYNX 4.0 software (Waters Corporation, Manchester, UK) and with MASSIGN software package [71].

### Liquid chromatography/Electrospray-ionization mass spectrometry (LC/ESI-TOF MS) under denaturing conditions

Presence of covalent oligomers in our IbeA samples was investigated by LC/ESI-TOF MS on a 6210 TOF mass spectrometer coupled to a high-performance liquid chromatography (HPLC) system (1100 series; Agilent Technologies, Santa Clara, CA, USA) using previously described procedure [72]. The mass spectrometer was calibrated with tuning mix (ESI-L, Agilent Technologies). The following instrumental settings were used: gas temperature (nitrogen) 300  $^{\circ}\text{C}$ , drying gas (nitrogen) 7 L·min<sup>-1</sup>, nebulizer gas (nitrogen) 10 psig,  $V_{\text{cap}}$  4 kV, fragmentor 250 V, skimmer 60 V,  $V_{\text{pp}}$  (octopole RF) 250 V. The HPLC mobile phases were prepared with HPLC grade solvents. Mobile phase A composition was: H<sub>2</sub>O 95%, ACN 5%, trifluoroacetic acid (TFA) 0.03%. Mobile phase B composition was: ACN 95%, H<sub>2</sub>O 5%, TFA 0.03%.

4  $\mu\text{L}$  of each protein sample at 5  $\mu\text{M}$  were first desalted online for 3 min with 100% of mobile phase A (flow rate of 50  $\mu\text{L}\cdot\text{min}^{-1}$ ), using a C8 reverse phase micro-column (Zorbax 300SB-C8, 5  $\mu\text{m}$ , 5  $\times$  0.3 mm, Agilent Technologies). The sample was then eluted with 70% of mobile phase B (flow rate of 50  $\mu\text{L}\cdot\text{min}^{-1}$ ) and MS spectra were acquired in the positive ion mode in the 300–3000 *m/z* range. Data were processed with MASSHUNTER (v. B.02.00, Agilent Technologies) and GPMW (v. 7.00b2, Lighthouse Data, Denmark) software.

### Direct infusion mass spectrometry (DIMS)

Direct infusion mass spectrometry requires the use of a syringe for delivering samples, directly into the electrospray source of the ESI-TOF MS on a 6210 TOF mass spectrometer. The flow rate was of 20  $\mu\text{L}\cdot\text{min}^{-1}$ . The following instrumental settings were used: gas temperature (nitrogen) 325  $^{\circ}\text{C}$ , drying gas (nitrogen) 5 L·min<sup>-1</sup>, nebulizer gas (nitrogen) 30 psig,  $V_{\text{cap}}$  3.5 kV, fragmentor 200 V, skimmer 65 V,  $V_{\text{pp}}$  (octopole RF) 250 V.

### Monitoring of the stability of IbeA over time by SEC

For stability assessment, fractions of purified IbeA encompassing both the F1 and F2 forms were pooled together following initial purification by SEC and were rerun either

immediately or after 4-day incubation at 4 °C (on ice) on a Superdex 200 Increase column (Cytiva) equilibrated with Buffer 1 (non-reducing conditions) or Buffer 2 (reducing conditions). The same experiments were also conducted with an F1-F2 pool directly stored at −80 °C after initial purification and then thawed for rerun onto the SEC column, either immediately or after 4-day incubation on ice. Both protocols yielded the same results.

### Binding of FAD to recombinant IbeA as monitored by SEC

500 µg of a freshly thawed aliquot of F1-F2 IbeA pool were incubated for 2 h at 25 °C with a 0.06 to 10-fold molar excess of FAD (or equivalent volume of ultrapure water as control) in a total volume of 450 µL (protein concentration: 1.1 mg·mL<sup>−1</sup> or 22.5 µM) in Buffer 1 (non-reducing conditions) or Buffer 2 (reducing conditions). The mix was then loaded through a 1 mL injection loop onto the Superdex 200 Increase 10/300 GL column (24 mL) equilibrated in either Buffer 1 or Buffer 2. The absorbance at both 280 and 450 nm was recorded through the entire elution. Fractions of 0.5 mL were collected during the entire run, except for the analytical run displayed in Fig. 11B,C for which 0.25 mL fractions were collected all throughout the elution.

The fraction of FAD bound per IbeA protomer within the F1-F2 pool was estimated as follows. Protein concentration was first measured using standard Bradford assay with the Coomassie (Bradford) Protein assay kit (Thermo Fisher Scientific), according to the manufacturer's instructions. Next the IbeA:FAD mix was denatured by incubation at 95 °C for 10 min in a heating block. Denatured IbeA was then removed by centrifugation at 20 000 *g* for 15 min in a table-top centrifuge cooled down at 4 °C. The supernatant containing the released FAD was then recovered and kept on ice. The concentration of FAD was determined by measuring the absorbance at 450 nm on a Nanodrop One/One<sup>C</sup> spectrophotometer (ThermoFisher Scientific), using the molar extinction coefficient of FAD in solution (11 300 M<sup>−1</sup>·cm<sup>−1</sup> at 450 nm).

### Tryptophan fluorescence quenching assay to measure IbeA-ligand interactions

All tryptophan fluorescence experiments were performed on a Spark<sup>®</sup> multimode microplate reader (TECAN, Männedorf, Switzerland). Recombinant WT or mutant

IbeA (F1-F2 pool) was incubated for 2 h in a water bath at 25 °C, at a concentration of 1 µM, with the ligand of interest (FAD, FMN, NADH or NADP) at concentrations ranging from 0.6 nM to 320 µM (2 : 2 dilution series from 320 µM; 20 concentrations plus the control without ligand). The mixes were then aliquoted in a Greiner Bio-One 96-well black plate with flat bottom (150 µL per well; 3 technical replicates) and fluorescence measurements were performed using the TECAN microplate reader, with temperature controlled at 25 °C by an external thermostatic water circulator. The excitation wavelength was fixed to 280 nm (bandwidth of 5 nm to avoid interfering with the emitted light) and Trp fluorescence emission spectra were collected every 1 nm between 300 and 400 nm (bandwidth of 10 nm). In each experiment, fluorescence of Buffer 2 was measured in parallel for background determination. As the wavelength corresponding to the fluorescence maximum may vary slightly from one data point to the other, an average value was taken over 5 different wavelengths centered at 341 nm (341 ± 2 nm). Data from three independent experiments were collected.

To correct for primary and secondary IFE of each nucleotide ligand, experiments were repeated in triplicates as described above but with 1 µM of N-Acetyl-L-tryptophanamide (NATA) instead of IbeA, as described by Yammine *et al.* [73]. Calculations of the IFE-corrected fluorescence intensity are detailed in Fig. 9. Briefly, a dose-dependent correlation between the Trp fluorescence decrease due to IFE and the ligand concentration could be derived from the experiments with NATA, yielding a correction factor for each ligand concentration that could then be applied to the Trp fluorescence measured in the presence of IbeA. This IFE-correction however yielded aberrant values at concentrations of ligand above 80 µM. Therefore, all data above 80 µM for the ligand concentration were truncated for the determination of the affinity constant.

To obtain the apparent dissociation constant ( $K_{d-app}$ ) for the various IbeA-ligand complexes, the relative fluorescence variation  $\Delta F/F_o$  was reported as a function of the ligand concentration, where  $\Delta F$  is the absolute value of the difference between Trp fluorescence of IbeA at a given ligand concentration and the Trp fluorescence of IbeA without ligand ( $F_o$ ). As  $K_d$  values are within the same range as the protein concentration (low micromolar), the data were fitted with a quadratic regression in GRAPH PAD PRISM (v 6.07; Boston, MA, USA), to account for FAD depletion upon complex formation. The following quadratic equation was employed for data fitting:

$$\frac{\Delta F}{F_o} = \frac{\Delta F_m}{F_o} \times \frac{([P_{tot}] + [L_{tot}] + K_d) - \sqrt{([P_{tot}] + [L_{tot}] + K_d)^2 - 4 \times [P_{tot}] \times [L_{tot}]}}{2 \times [P_{tot}]}$$

where  $\Delta F$  is the recorded fluorescence change at a given ligand concentration;  $F_0$  is the initial fluorescence in the absence of ligand;  $\Delta F_m$  is the maximal fluorescence change upon ligand binding;  $[P_{\text{tot}}]$  and  $[L_{\text{tot}}]$  are the total concentrations of protein and ligand, respectively; and  $K_d$  is the dissociation constant.

### **In vitro IbeA activity assay in the presence of nicotinamide cofactors**

All experiments were carried out under anaerobic conditions, in an inflatable polyethylene glove bag (Thermo Fisher Scientific) inflated with  $N_2$  gas and tightly closed with sealing clamps. Reaction mixtures were prepared in a final volume of 60  $\mu\text{L}$  and contained 150  $\mu\text{M}$  FAD, 150  $\mu\text{M}$  NADH or NADPH, and 5  $\mu\text{M}$  purified IbeA (F1 form prepared under reducing conditions) in 20 mM Tris-HCl pH 7.5, 150 mM NaCl. Prior to mixing, the stock solutions of FAD, NADH, NADPH and the buffers were bubbled with  $N_2$  gas to remove oxygen. All compounds were mixed together and the reaction was started by adding purified IbeA (or equivalent volume of IbeA buffer for control). The mixtures were then incubated at 25 °C for 2 or 6 h. Finally, absorbance at 450 nm was recorded to evaluate whether FAD reduction occurs (decrease in absorbance at 450 nm) or not (constant absorbance at 450 nm over time). The measured absorbances at 450 nm were normalized against the value at time  $t = 0$  (taken as the 100% absorbance reference), both for the control samples without IbeA and for the samples incubated in the presence of 5  $\mu\text{M}$  IbeA. Data from three independent experiments were collected.

### **3D-modeling and ligand docking**

The 3D structural models for the IbeA protomer were predicted using AlphaFold (v2.3.2) [74] from the online Colab notebook or RoseTTAFold [75] from the GitHub repository. The 3D-model for IbeA homodimer was generated using AlphaFold-Multimer [43]. Prior to ligand docking, the best-ranked AlphaFold model was subjected to energy minimization by applying 10 000 steps of conjugate gradient algorithm and using protein topology/parameters from the Charmm36m force field [76]. Docking of FAD onto the IbeA 3D-structure was performed using either AUTODOCK VINA (v1.1.2) [77] or GOLD v5.8 (Genetic Optimization for Ligand Docking) from the Cambridge Crystallographic Data Centre [78]. For docking with AUTODOCK VINA, the search was performed with a docking box centered as followed:  $x = 0.2378$ ,  $y = 0.537$ , and  $z = -0.1796$  with dimension ( $\text{\AA}$ )  $X = 99.11$ ,  $Y = 53.76$  and  $Z = 75.8$ , and exhaustiveness defined to 8. For docking with Gold, a target binding area was defined as a sphere of radius of 20  $\text{\AA}$  around the backbone nitrogen atom of Met54 residue (for each protomer) with the search option of cavity activated.

The search of best docking poses was obtained by executing 50 runs of genetic algorithms (search-based optimization technique based on genetics and natural selection) and the ranking of docking solutions was computed by the GoldScore fitness function by using the complete linkage clustering method from the RMSD matrix of generated solutions. A third docking model was generated by simple superimposition of the IbeA 3D-model with the X-ray structure of FAD-bound GGR from *Sulfolobus acidocaldarius* (PDB: 3ATR) in COOT [79] was also employed. All figures with structural models were made with the PYMOL Molecular Graphics System (version 0.99rc6; DeLano Scientific LLC, San Francisco, CA, USA). The 2D ligand-protein interaction diagrams were generated with PLIP and the Lidia ligand builder from COOT [79,80]. Electrostatic and solvation properties of the IbeA dimer surface were calculated using APBS [81] and represented as surface potentials ranging from  $-8$  to  $+8$  kT/e in PYMOL.

### **Molecular dynamics (MD) simulations**

Molecular dynamics simulations were performed using the NAMD package v3.0 $\alpha$ 12 [82] in conjunction with the Charmm36m force field [76]. The two 3D-models of IbeA homodimer (apo or FAD-bound) were immersed in a cubic explicit solvent box of 100  $\text{\AA}^3$  using the TIP3P water model [83] with periodic boundary conditions to simulate the biological environment realistically. Next,  $\text{Na}^+$  and  $\text{Cl}^-$  ions were added at a final concentration of 0.154 M to ensure neutrality of the periodic box. Each system was first energy minimized by performing 10 000 steps of conjugate gradients and then equilibrated for 250 ps. All simulations were carried out in the isobaric-isothermal ensemble, at constant temperature (310 K) and pressure (1 atm) using Langevin dynamics and Langevin piston. Long-range electrostatic interactions were treated using the particle-mesh Ewald approach [84]. Integration time step of the equation of motions was 2 fs and MD trajectories were recorded during 200 ns (every 100 ps one frame was saved leading to a total of 2000 frames) for IbeA monomeric forms and during 600 ns (6000 frames) for homodimers. Gyration radius ( $R_{\text{gyr}}$ ) and Solvent Accessible Surface Area (SASA) of protein models were computed using the VMD package [85] on the full-length trajectories while 2D-Root Mean Square Deviations (2D-RMSD) were computed on protein backbone atoms using the MDAnalysis PYTHON package [86]. Covariance analysis was processed by using the PRODY package [87] allowing principal component analysis or protein essential dynamics analysis.

### **Statistical analyses**

Statistical analyses to compare the  $K_{d\text{-app}}$  values obtained for the binding of FAD to IbeA WT and mutants were



performed in GRAPHPAD PRISM (version 6.07) using unpaired *t*-test with Welch's correction.

## Acknowledgements

We are grateful to Dr. Nicolas Barnich for providing the AIEC LF82 strain. We thank Dr. Corinne Lionne and Dr. Anne Blanc-Potard for fruitful discussion. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-05-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA). Financial support to access the ISBG platform was provided to LY by Instruct-ERIC (PID 17773).

## Conflict of interest

The authors declare no conflict of interest.

## Author contributions

LY conceived the study and planned the experiments; TP, AK, LS, MB, EBE, LC, and LY performed experiments; TP, MB, GL, EBE, LC, and LY analyzed the data and prepared the figures; LY wrote the original draft of the paper with help and proofreading from all authors; GL and LY secured funding.

## Peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/febs.16969>.

## Data availability statement

The data that support the findings of this study are available from the corresponding author ([laure.yatime@inserm.fr](mailto:laure.yatime@inserm.fr)) upon request.

## References

- Darfeuille-Michaud A, Boudeau J, Bulois P, Neut C, Glasser A-L, Barnich N, Bringer M-A, Swidsinski A, Beaugerie L & Colombel J-F (2004) High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. *Gastroenterology* **127**, 412–421.
- Rahmouni O, Vignal C, Titécat M, Foligné B, Pariente B, Dubuquoy L, Desreumaux P & Neut C (2018) High carriage of adherent invasive *E. coli* in wildlife and healthy individuals. *Gut Pathog* **10**, 23.
- Negrone A, Costanzo M, Vitali R, Superti F, Bertuccini L, Tinari A, Minelli F, Di Nardo G, Nuti F, Pierdomenico M *et al.* (2012) Characterization of adherent-invasive *Escherichia coli* isolated from pediatric patients with inflammatory bowel disease. *Inflamm Bowel Dis* **18**, 913–924.
- Elliott TR, Hudspeth BN, Wu G, Cooley M, Parkes G, Quiñones B, Randall L, Mandrell RE, Fagerquist CK, Brostoff J *et al.* (2013) Quantification and characterization of mucosa-associated and intracellular *Escherichia coli* in inflammatory bowel disease. *Inflamm Bowel Dis* **19**, 2326–2338.
- Lapaquette P, Bringer M-A & Darfeuille-Michaud A (2012) Defects in autophagy favour adherent-invasive *Escherichia coli* persistence within macrophages leading to increased pro-inflammatory response: autophagy controls AIEC replication within macrophages. *Cell Microbiol* **14**, 791–807.
- Elhenawy W, Tsai CN & Coombes BK (2019) Host-specific adaptive diversification of Crohn's disease-associated adherent-invasive *Escherichia coli*. *Cell Host Microbe* **25**, 301–312.e5.
- Zhang S, Morgan X, Dogan B, Martin F-P, Strickler S, Oka A, Herzog J, Liu B, Dowd SE, Huttenhower C *et al.* (2022) Mucosal metabolites fuel the growth and virulence of *E. coli* linked to Crohn's disease. *JCI Insight* **7**, e157013.
- Small C-LN, Reid-Yu SA, McPhee JB & Coombes BK (2013) Persistent infection with Crohn's disease-associated adherent-invasive *Escherichia coli* leads to chronic inflammation and intestinal fibrosis. *Nat Commun* **4**, 1957.
- Viladomiu M, Metz ML, Lima SF, Jin W-B, Chou L, Guo C-J, Diehl GE, Simpson KW, Scherl EJ & Longman RS (2021) Adherent-invasive *E. coli* metabolism of propanediol in Crohn's disease regulates phagocytes to drive intestinal inflammation. *Cell Host Microbe* **29**, 607–619.e8.
- O'Brien CL, Bringer M-A, Holt KE, Gordon DM, Dubois AL, Barnich N, Darfeuille-Michaud A & Pavli P (2017) Comparative genomics of Crohn's disease-associated adherent-invasive *Escherichia coli*. *Gut* **66**, 1382–1389.
- Camprubí-Font C, Ewers C, Lopez-Siles M & Martínez-Medina M (2019) Genetic and phenotypic features to screen for putative adherent-invasive *Escherichia coli*. *Front Microbiol* **10**, 108.
- Glasser A-L, Boudeau J, Barnich N, Perruchot M-H, Colombel J-F & Darfeuille-Michaud A (2001) Adherent invasive *Escherichia coli* strains from patients with Crohn's disease survive and replicate within

- macrophages without inducing host cell death. *Infect Immun* **69**, 5529–5537.
- 13 Dumych T, Yamakawa N, Sivignon A, Garenaux E, Robakiewicz S, Coddeville B, Bongiovanni A, Bray F, Barnich N, Szunerits S *et al.* (2018) Oligomannose-rich membranes of dying intestinal epithelial cells promote host colonization by adherent-invasive *E. coli*. *Front Microbiol* **9**, 742.
  - 14 Bringer M-A, Billard E, Glasser A-L, Colombel J-F & Darfeuille-Michaud A (2012) Replication of Crohn's disease-associated AIEC within macrophages is dependent on TNF- $\alpha$  secretion. *Lab Invest* **92**, 411–419.
  - 15 Prudent V, Demarre G, Vazeille E, Wery M, Quenech'Du N, Ravet A, Dauverd-Girault J, Van Dijk E, Bringer M-A, Descrimes M *et al.* (2021) The Crohn's disease-related bacterial strain LF82 assembles biofilm-like communities to protect itself from phagolysosomal attack. *Commun Biol* **4**, 627.
  - 16 Martinez-Medina M, Strozzi F, Ruiz Del Castillo B, Serrano-Morillas N, Ferrer Bustins N & Martínez-Martínez L (2020) Antimicrobial resistance profiles of adherent invasive *Escherichia coli* show increased resistance to  $\beta$ -lactams. *Antibiotics* **9**, 251.
  - 17 Chassaing B, Rolhion N, Vallée AD, Salim SY, Prorok-Hamon M, Neut C, Campbell BJ, Söderholm JD, Hugot J-P, Colombel J-F *et al.* (2011) Crohn disease-associated adherent-invasive *E. coli* bacteria target mouse and human Peyer's patches via long polar fimbriae. *J Clin Invest* **121**, 966–975.
  - 18 Dogan B, Suzuki H, Herlekar D, Sartor RB, Campbell BJ, Roberts CL, Stewart K, Scherl EJ, Araz Y, Bitar PP *et al.* (2014) Inflammation-associated adherent-invasive *Escherichia coli* are enriched in pathways for use of propanediol and iron and M-cell translocation. *Inflamm Bowel Dis* **20**, 1919–1932.
  - 19 Hase K, Kawano K, Nochi T, Pontes GS, Fukuda S, Ebisawa M, Kadokura K, Tobe T, Fujimura Y, Kawano S *et al.* (2009) Uptake through glycoprotein 2 of FimH+ bacteria by M cells initiates mucosal immune response. *Nature* **462**, 226–230.
  - 20 Agus A, Massier S, Darfeuille-Michaud A, Billard E & Barnich N (2014) Understanding host-adherent-invasive *Escherichia coli* interaction in Crohn's disease: opening up new therapeutic strategies. *Biomed Res Int* **2014**, 567929.
  - 21 Abdelhalim KA, Uzel A & Gülşen Ünal N (2020) The role of major virulence factors and pathogenicity of adherent-invasive *Escherichia coli* in patients with Crohn's disease. *Prz Gastroenterol* **15**, 279–288.
  - 22 Sivignon A, Yan X, Alvarez Dorta D, Bonnet R, Bouckaert J, Fleury E, Bernard J, Gouin SG, Darfeuille-Michaud A & Barnich N (2015) Development of heptylmannoside-based glycoconjugate antiadhesive compounds against adherent-invasive *Escherichia coli* bacteria associated with Crohn's disease. *MBio* **6**, e01298–15.
  - 23 Cieza RJ, Hu J, Ross BN, Sbrana E & Torres AG (2015) The IbeA Invasin of adherent-invasive *Escherichia coli* mediates interaction with intestinal epithelia and macrophages. *Infect Immun* **83**, 1904–1918.
  - 24 Huang SH, Wass C, Fu Q, Prasadarao NV, Stins M & Kim KS (1995) *Escherichia coli* invasion of brain microvascular endothelial cells in vitro and in vivo: molecular cloning and characterization of invasion gene *ibe10*. *Infect Immun* **63**, 4470–4475.
  - 25 Huang S, Wan Z, Chen Y, Jong AY & Kim KS (2001) Further characterization of *Escherichia coli* brain microvascular endothelial cell invasion gene *ibeA* by deletion, complementation, and protein expression. *J Infect Dis* **183**, 1071–1078.
  - 26 Mendu DR, Dasari VR, Cai M & Kim KS (2008) Protein folding intermediates of invasin protein IbeA from *Escherichia coli*: protein folding intermediates. *FEBS J* **275**, 458–469.
  - 27 Zou Y, He L & Huang S-H (2006) Identification of a surface protein on human brain microvascular endothelial cells as vimentin interacting with *Escherichia coli* invasion protein IbeA. *Biochem Biophys Res Commun* **351**, 625–630.
  - 28 Zou Y, He L, Wu C-H, Cao H, Xie Z-H, Ouyang Y, Wang Y, Jong A & Huang S-H (2007) PSF is an IbeA-binding protein contributing to meningitic *Escherichia coli* K1 invasion of human brain microvascular endothelial cells. *Med Microbiol Immunol* **196**, 135–143.
  - 29 Chi F, Bo T, Wu C-H, Jong A & Huang S-H (2012) Vimentin and PSF act in concert to regulate IbeA+ *E. coli* K1 induced activation and nuclear translocation of NF- $\kappa$ B in human brain endothelial cells. *PLoS One* **7**, e35862.
  - 30 Zhao W-D, Liu D-X, Wei J-Y, Miao Z-W, Zhang K, Su Z-K, Zhang X-W, Li Q, Fang W-G, Qin X-X *et al.* (2018) Caspr1 is a host receptor for meningitis-causing *Escherichia coli*. *Nat Commun* **9**, 2296.
  - 31 Vidotto MC, Queiroz MB, De Lima NCS & Gaziri LCJ (2007) Prevalence of *ibeA* gene in avian pathogenic *Escherichia coli* (APEC). *Vet Microbiol* **119**, 88–89.
  - 32 Germon P, Chen Y-H, He L, Blanco JE, Brée A, Schouler C, Huang S-H & Moulin-Schouleur M (2005) *ibeA*, a virulence factor of avian pathogenic *Escherichia coli*. *Microbiology* **151**, 1179–1186.
  - 33 Wang S, Niu C, Shi Z, Xia Y, Yaqoob M, Dai J & Lu C (2011) Effects of *ibeA* deletion on virulence and biofilm formation of avian pathogenic *Escherichia coli*. *Infect Immun* **79**, 279–287.
  - 34 Flechard M, Cortes MAM, Reperant M & Germon P (2012) New role for the *ibeA* gene in H<sub>2</sub>O<sub>2</sub> stress resistance of *Escherichia coli*. *J Bacteriol* **194**, 4550–4560.
  - 35 Cieza RJ (2015) The invasin IbeA and its role in adherent-invasive *Escherichia coli* (AIEC) pathogenesis. PhD Thesis, University of Texas Medical Branch, USA.

- 36 Dym O & Eisenberg D (2001) Sequence-structure analysis of FAD-containing proteins. *Protein Sci* **10**, 1712–1728.
- 37 Karplus PA & Schulz GE (1987) Refined structure of glutathione reductase at 1.54 Å resolution. *J Mol Biol* **195**, 701–729.
- 38 Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G & Nielsen H (2022) SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol* **40**, 1023–1025.
- 39 Holm L, Laiho A, Törönen P & Salgado M (2023) DALI shines a light on remote homologs: one hundred discoveries. *Protein Sci* **32**, e4519.
- 40 Krissinel E & Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256–2268.
- 41 Xu Q, Eguchi T, Mathews II, Rife CL, Chiu H-J, Farr CL, Feuerhelm J, Jaroszewski L, Klock HE, Knuth MW *et al.* (2010) Insights into substrate specificity of geranylgeranyl reductases revealed by the structure of digeranylgeranyl glycerophospholipid reductase, an essential enzyme in the biosynthesis of archaeal membrane lipids. *J Mol Biol* **404**, 403–417.
- 42 Sasaki D, Fujihashi M, Iwata Y, Murakami M, Yoshimura T, Hemmi H & Miki K (2011) Structure and mutation analysis of archaeal geranylgeranyl reductase. *J Mol Biol* **409**, 543–557.
- 43 Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Židek A, Bates R, Blackwell S, Yim J *et al.* (2021) *Protein complex prediction with AlphaFold-Multimer*. Bioinformatics.
- 44 Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
- 45 Homeier T, Semmler T, Wieler LH & Ewers C (2010) The GimA locus of extraintestinal pathogenic *E. coli*: does reductive evolution correlate with habitat and Pathotype? *PLoS One* **5**, e10877.
- 46 Cortes MAM, Gibon J, Chanteloup NK, Moulin-Schouleur M, Gilot P & Germon P (2008) Inactivation of *ibeA* and *ibeT* results in decreased expression of type 1 fimbriae in extraintestinal pathogenic *Escherichia coli* strain BEN2908. *Infect Immun* **76**, 4129–4136.
- 47 Bachmann NL, Katouli M & Polkinghorne A (2015) Genomic comparison of translocating and non-translocating *Escherichia coli*. *PLoS One* **10**, e0137131.
- 48 Elhenawy W, Hordienko S, Gould S, Oberc AM, Tsai CN, Hubbard TP, Waldor MK & Coombes BK (2021) High-throughput fitness screening and transcriptomics identify a role for a type IV secretion system in the pathogenesis of Crohn's disease-associated *Escherichia coli*. *Nat Commun* **12**, 2032.
- 49 Nash JH, Villegas A, Kropinski AM, Aguilar-Valenzuela R, Konczyk P, Mascarenhas M, Ziebell K, Torres AG, Karmali MA & Coombes BK (2010) Genome sequence of adherent-invasive *Escherichia coli* and comparative genomic analysis with other *E. coli* pathotypes. *BMC Genomics* **11**, 667.
- 50 Le TTH, Kellenberger C, Boyer M, Santucci P, Flaugnatti N, Cascales E, Roussel A, Canaan S, Journet L & Cambillau C (2023) Activity and crystal structure of the adherent-invasive *Escherichia coli* Tle3/Tli3 T6SS effector/immunity complex determined using an AlphaFold2 predicted model. *Int J Mol Sci* **24**, 1740.
- 51 Ma J, Bao Y, Sun M, Dong W, Pan Z, Zhang W, Lu C & Yao H (2014) Two functional type VI secretion systems in avian pathogenic *Escherichia coli* are involved in different pathogenic pathways. *Infect Immun* **82**, 3867–3879.
- 52 Hiraka K, Tsugawa W & Sode K (2020) Alteration of electron acceptor preferences in the oxidative half-reaction of Flavin-dependent oxidases and dehydrogenases. *Int J Mol Sci* **21**, 3797.
- 53 Hemmi H, Takahashi Y, Shibuya K, Nakayama T & Nishino T (2005) Menaquinone-specific prenyl reductase from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J Bacteriol* **187**, 1937–1944.
- 54 Karplus PA & Schulz GE (1989) Substrate binding and catalysis by glutathione reductase as derived from refined enzyme: substrate crystal structures at 2 Å resolution. *J Mol Biol* **210**, 163–180.
- 55 Loveridge EJ, Rodriguez RJ, Swanwick RS & Allemann RK (2009) Effect of dimerization on the stability and catalytic activity of dihydrofolate reductase from the hyperthermophile *Thermotoga maritima*. *Biochemistry* **48**, 5922–5933.
- 56 Sanchez JE, Gross PG, Goetze RW, Walsh RM, Peebles WB & Wood ZA (2015) Evidence of kinetic cooperativity in dimeric Ketopantoate reductase from *Staphylococcus aureus*. *Biochemistry* **54**, 3360–3369.
- 57 Martin LL, Holien JK, Mizrahi D, Corbin CJ, Conley AJ, Parker MW & Rodgers RJ (2015) Evolutionary comparisons predict that dimerization of human cytochrome P450 aromatase increases its enzymatic activity and efficiency. *J Steroid Biochem Mol Biol* **154**, 294–301.
- 58 Nishimura Y & Eguchi T (2006) Biosynthesis of archaeal membrane lipids: digeranylgeranyl glycerophospholipid reductase of the thermoacidophilic archaeon *Thermoplasma acidophilum*. *J Biochem* **139**, 1073–1081.
- 59 Proctor MS, Sutherland GA, Canniffe DP & Hitchcock A (2022) The terminal enzymes of (bacterio)chlorophyll biosynthesis. *R Soc Open Sci* **9**, 211903.
- 60 Xu X, Zhang L, Cai Y, Liu D, Shang Z, Ren Q, Li Q, Zhao W & Chen Y (2020) Inhibitor discovery for the *E. coli* meningitis virulence factor IbeA from homology

- modeling and virtual screening. *J Comput Aided Mol Des* **34**, 11–25.
- 61 Wishart MJ, Taylor GS & Dixon JE (2001) Phoxy lipids: revealing PX domains as phosphoinositide binding modules. *Cell* **105**, 817–820.
- 62 Ponting CP (1996) Novel domains in NADPH oxidase subunits, sorting nexins, and PtdIns 3-kinases: binding partners of SH3 domains? *Protein Sci* **5**, 2353–2357.
- 63 Bond SR & Naus CC (2012) RF-Cloning.org: an online tool for the design of restriction-free cloning projects. *Nucleic Acids Res* **40**, W209–W213.
- 64 Alsarraf HMAB, Laroche F, Spaink H, Thirup S & Blaise M (2011) Purification, crystallization and preliminary crystallographic studies of the TLDC domain of oxidation resistance protein 2 from zebrafish. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **67**, 1253–1256.
- 65 Boeri Erba E & Petosa C (2015) The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes: the emerging role of native mass spectrometry. *Protein Sci* **24**, 1176–1192.
- 66 Boeri Erba E, Signor L, Oliva MF, Hans F & Petosa C (2018) Characterizing intact macromolecular complexes using native mass spectrometry. In *Protein Complex Assembly* (Marsh JA, ed.), pp. 133–151. Springer, New York, NY.
- 67 Boeri Erba E, Signor L & Petosa C (2020) Exploring the structure and dynamics of macromolecular complexes by native mass spectrometry. *J Proteomics* **222**, 103799.
- 68 Pellegrini E, Signor L, Singh S, Boeri Erba E & Cusack S (2017) Structures of the inactive and active states of RIP2 kinase inform on the mechanism of activation. *PLoS One* **12**, e0177161.
- 69 Sobott F, Hernández H, McCammon MG, Tito MA & Robinson CV (2002) A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal Chem* **74**, 1402–1407.
- 70 Van Den Heuvel RHH, Van Duijn E, Mazon H, Synowsky SA, Lorenzen K, Versluis C, Brouns SJJ, Langridge D, Van Der Oost J, Hoyes J *et al.* (2006) Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal Chem* **78**, 7473–7483.
- 71 Morgner N & Robinson CV (2012) *Mass ign*: an assignment strategy for maximizing information from the mass spectra of heterogeneous protein assemblies. *Anal Chem* **84**, 2939–2948.
- 72 Signor L, Paris T, Mas C, Picard A, Lutfalla G, Boeri Erba E & Yatime L (2021) Divalent cations influence the dimerization mode of murine S100A9 protein by modulating its disulfide bond pattern. *J Struct Biol* **213**, 107689.
- 73 Yammine A, Gao J & Kwan A (2019) Tryptophan fluorescence quenching assays for measuring protein–ligand binding affinities: principles and a practical guide. *Bio-Protoc* **9**, e3253.
- 74 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.
- 75 Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN, Schaeffer RD *et al.* (2021) Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876.
- 76 Huang J & MacKerell AD (2013) CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* **34**, 2135–2145.
- 77 Trott O & Olson AJ (2009) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455–461.
- 78 Jones G, Willett P, Glen RC, Leach AR & Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* **267**, 727–748.
- 79 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of *Coot*. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.
- 80 Salentin S, Schreiber S, Haupt VJ, Adasme MF & Schroeder M (2015) PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Res* **43**, W443–W447.
- 81 Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, Brookes DH, Wilson L, Chen J, Liles K *et al.* (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci* **27**, 112–128.
- 82 Phillips JC, Hardy DJ, Maia JDC, Stone JE, Ribeiro JV, Bernardi RC, Buch R, Fiorin G, Hémin J, Jiang W *et al.* (2020) Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* **153**, 044130.
- 83 Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW & Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926–935.
- 84 Darden T, York D & Pedersen L (1993) Particle mesh Ewald: an  $N\log(N)$  method for Ewald sums in large systems. *J Chem Phys* **98**, 10089–10092.
- 85 Humphrey W, Dalke A & Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* **14**, 33–38.
- 86 Gowers R, Linke M, Barnoud J, Reddy T, Melo M, Seyler S, Domański J, Dotson D, Buchoux S, Kenney I *et al.* (2016) MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations, pp. 98–105. Austin, TX. doi: [10.25080/Majora-629e541a-00e](https://doi.org/10.25080/Majora-629e541a-00e)
- 87 Zhang S, Krieger JM, Zhang Y, Kaya C, Kaynak B, Mikulska-Ruminska K, Doruker P, Li H & Bahar I (2021) *ProDy* 2.0: increased scale and scope after 10 years of protein dynamics modelling with Python. *Bioinformatics* **37**, 3657–3659.