doi:10.1016/j.jmb.2011.04.034 J. Mol. Biol. (2011) 409, 642–653

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

The Extra-Membranous Domains of the Competence Protein HofQ Show DNA Binding, Flexibility and a Shared Fold with Type I KH Domains

Michael Tarry[1]†, Mari Jääskeläinen[1,2]†, Annamari Paino[2], Heidi Tuominen[2], Riikka Ihalin[2]* and Martin Högbom[1]*

[1]*Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden*
[2]*Department of Biochemistry and Food Chemistry, University of Turku, Turku FI-20014, Finland*

Received 7 February 2011; received in revised form 8 April 2011; accepted 12 April 2011
Available online 20 April 2011

Edited by R. Huber

Keywords:
membrane protein;
DNA uptake;
natural competence;
secretin;
structural dynamics

Secretins form large oligomeric assemblies in the membrane that control both macromolecular secretion and uptake. Several *Pasteurellaceae* are naturally competent for transformation, but the mechanism for DNA assimilation is largely unknown. In *Haemophilus influenzae*, the secretin ComE has been demonstrated to be essential for DNA uptake. In closely related *Aggregatibacter actinomycetemcomitans*, an opportunistic pathogen in periodontitis, the ComE homolog HofQ is believed to be the outer membrane DNA translocase. Here, we report the structure of the extra-membranous domains of HofQ at 2.3 Å resolution by X-ray crystallography. We also show that the extra-membranous domains of HofQ are capable of DNA binding. The structure reveals two secretin-like folds, the first of which is formed by means of a domain swap. The second domain displays extensive structural similarity to K homology (KH) domains, including the presence of a GxxG motif, which is essential for the nucleotide-binding function of KH domains, suggesting a possible mechanism for DNA binding by HofQ. The data indicate a direct involvement in DNA acquisition and provide insight into the molecular basis for natural competence.

Introduction

Aggregatibacter (*Actinobacillus*) *actinomycetemcomitans* belongs to the family of Gram-negative *Pasteurellaceae* and is one of the major opportunistic pathogens in localized aggressive periodontitis,[1–3] an infection that attacks and destroys the supportive tissues of the teeth. In addition to causing damage to the oral cavity, DNA and viable cells of some periodontal pathogens, including *A. actinomycetemcomitans*, have been found in atherosclerotic lesions,[4,5] indicating a potential link to systemic diseases.

A. actinomycetemcomitans is one of at least 44 species of bacteria that are naturally competent for transformation,[6,7] allowing cells to take up external DNA. In *A. actinomycetemcomitans*, as in *Haemophilus influenzae* and *Neisseria gonorrhoeae*, effective uptake of DNA involves a species-specific DNA sequence, termed the uptake signal sequence (USS).[7–9] The USS, which is approximately 10 nucleotides long, is frequently repeated in the genomes of these species, resulting in a preference

*Corresponding authors. E-mail addresses: riikka.ihalin@utu.fi; hogbom@dbb.su.se.

† M.T. and M.J. contributed equally to this work.

Present address: M. Tarry, Department of Biochemistry, Francesco Bellini Building, McGill University, Montreal, Quebec, Canada H3G 0B1.

Abbreviations used: USS, uptake signal sequence; dsDNA, double-stranded DNA; TIISS, type II secretion system; TIIISS, type III secretion system; PDB, Protein Data Bank; BSA, bovine serum albumin; SeMet, selenomethionine; ESRF, European Synchrotron Radiation Facility; RT, room temperature; KH, K homology.

for homo-specific DNA uptake in natural transformation. In the majority of cases, the DNA that is taken up most likely serves as a source of nutrition and nucleotides, but if it contains regions homologous to the host genome, it may be incorporated into the host bacterium's genome. This mechanism could be of relevance in the acquisition of antibiotic resistance genes or virulence factors.[10,11] A recent study suggests that natural competence combined with transformation may stabilize the otherwise hyper-dynamic core genome in *Pasteurellaceae*.[12]

Gram-negative bacteria need a complex machinery to move DNA from outside of the cell into the bacterial cytoplasm. The DNA must be bound from the extracellular milieu and traverse the outer membrane, followed by the cell wall and the cytoplasmic membrane.[13] Uptake is facilitated by a macromolecular transport system composed of many different proteins that display similarities to those utilized in assembly of type IV pili secretion systems and type II secretion systems (TIISS).[14] Double-stranded DNA (dsDNA) is transported across the outer membrane through a channel into the periplasm, where it becomes DNase resistant prior to subsequent transport of single-stranded DNA across the cytoplasmic membrane.[15] *pilABCD* were the first genes to be identified as being required for competence in *A. actinomycetemcomitans* and belong to the type IV pilus-like machinery.[16] These genes most likely code for prepilin (PilA), two transport proteins (PilB and PilC) and prepilin peptidase (PilD).[16] No *A. actinomycetemcomitans* outer membrane proteins that could be responsible for DNA binding and translocation have yet been reported. The *H. influenzae* outer membrane contains a DNA uptake system composed of multiple copies of the secretin ComE, which has been shown to be essential for DNA uptake.[17,18] The *A. actinomycetemcomitans* protein HofQ (UniProtKB accession code: C9R226) shows some 70% sequence identity to *H. influenzae* ComE (Supplementary Fig. 1) and is thought to perform the same function.[17,19]

Secretins are large homo-oligomeric assemblies built up of 50- to 70-kDa subunits, with 12–14 subunits forming a ring structure of approximately 100–150 Å in diameter.[20–28] They comprise a superfamily[29,30] and form components of several distinct secretion systems in the outer membrane, including TIISS,[31,32] type III secretion systems (TIIISS)[26] and type IV pilus biogenesis systems.[30,33,34] All secretins consist of two major regions. The C-terminal region, or β-domain, is believed to be the major determinant in oligomer formation and stability.[35] It is highly conserved and predicted to contain several β-strands[36] embedded in the outer membrane to form the actual pore through which transport occurs.[21] The N-terminal periplasmic region displays conservation only in secretins from related secretion pathways[29] and may be involved in substrate recognition,[28,37] gating of the proposed channel[35,38] and DNA binding[39] as well as contributing to subunit oligomerization.[35,40]

The atomic structures of periplasmic fragments from two secretins, EscC from enteropathogenic *Escherichia coli* and GspD from enterotoxigenic *E. coli*, have previously been determined.[37,40] EscC and GspD belong to TIIISS and TIISS, respectively, and share little sequence conservation in their N-terminal regions. In spite of this, the two structures display remarkable structural similarities, highlighted by the presence of conserved N-terminal secretin-like folds in both of them.[37,40]

In this study, we cloned and expressed the extramembranous domains of HofQ (emHofQ) and determined the structure at 2.3 Å. We also show that emHofQ is capable of binding dsDNA. The structure revealed two secretin-like folds, one of which displays extensive structural homology to K homology (KH) domains involved in DNA binding, suggesting a possible structural basis for HofQ's role in natural competence. The structure also exhibits far-reaching structural flexibility with one of the N-terminal secretin-like folds formed via a domain swap.

Results

Overall structure

emHofQ crystallized as a dimer in the asymmetric unit (Fig. 1a). Analysis of the molecular interactions using the PDBe-PISA server[41] suggests that this dimeric assembly of emHofQ is the stable form in solution (buried area = 4550 Å^2, $\Delta G_{int} = -15.3$ kcal/mol, $\Delta G_{diss} = 18.1$ kcal/mol). In addition, the protein migrated at an apparent molecular mass of 41 kDa on a calibrated size-exclusion column during purification, also suggesting a dimeric assembly (the molecular mass of the emHofQ construct is 19 kDa). Still, a conclusive assignment is not possible because elongated molecules are expected to migrate at an apparently larger molecular mass in size-exclusion chromatography.

Each emHofQ polypeptide chain consists of two globular domains, containing a total of seven β-strands and four α-helices (Fig. 1b and c). The two domains are linked by approximately 25 residues, which are disordered in the structure, indicated by broken lines in Fig. 1a and c. Because of the length of the disordered segment, there is a theoretical possibility for an alternative connection within the crystal lattice. However, this would produce a very extended structure with protein molecules weaved together throughout the crystal lattice. Though this may be theoretically possible, we find this arrangement very unlikely. The two chains, referred to here as A and B, can be superimposed with an rmsd of 0.98 Å between