



HAL
open science

Anionic lipids induce a fold-unfold transition in the membrane-translocating Engrailed homeodomain

Ludovic Carlier, Damien Samson, Lucie Khemtemourian, Alain Joliot, Patrick F.J. Fuchs, Olivier Lequin

► To cite this version:

Ludovic Carlier, Damien Samson, Lucie Khemtemourian, Alain Joliot, Patrick F.J. Fuchs, et al.. Anionic lipids induce a fold-unfold transition in the membrane-translocating Engrailed homeodomain. *Biochimica et Biophysica Acta: Biomembranes*, 2022, 1864 (11), pp.184030. 10.1016/j.bbamem.2022.184030 . hal-03897252

HAL Id: hal-03897252

<https://hal.sorbonne-universite.fr/hal-03897252v1>

Submitted on 13 Dec 2022

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.

Anionic lipids induce a fold-unfold transition in the membrane-translocating Engrailed homeodomain

Ludovic CARLIER^{a,*}, Damien SAMSON^a, Lucie KHEMTEMOURIAN^{a,‡}, Alain JOLIOT^b, Patrick F. J. FUCHS^{a,c}, Olivier LEQUIN^{a,*}

^a Sorbonne Université, Ecole Normale Supérieure, PSL University, CNRS, Laboratoire des Biomolécules, 4 place Jussieu, F-75005 Paris, France

^b INSERM U932, Institut Curie Centre de Recherche, PSL University, France

^c Université Paris Cité, UFR Sciences du Vivant, F-75013 Paris, France

‡ Present address: Institute of Chemistry and Biology of Membranes and Nanoobjects, Institut Polytechnique Bordeaux, CNRS UMR 5248, Université de Bordeaux, Pessac, France

* Corresponding authors:

Laboratoire des Biomolécules, Boite 182, 4 place Jussieu, F-75005 Paris, France

E-mail address:

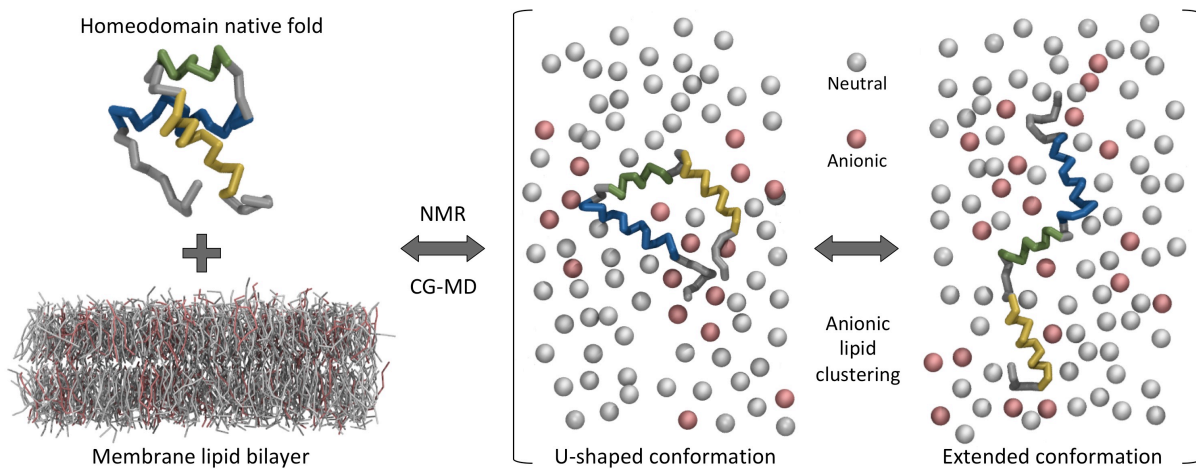
olivier.lequin@sorbonne.universite.fr; ludovic.carlier@sorbonne.universite.fr

Phone: +33-1-44276199

Highlights

- Engrailed homeodomain (EnHD) unfolds upon binding to anionic lipid membranes
- Near-native helical secondary structure of EnHD is maintained in membrane environment
- Global unfolding of EnHD leads to parallel insertion of helices within the membrane
- Anionic lipids are recruited and stabilize U-shaped EnHD conformations

Graphical abstract



Abstract

Homeoprotein transcription factors have the property of interacting with membranes through their DNA-binding homeodomain, which is involved in unconventional internalization and secretion. Both processes depend on membrane-translocating events but their detailed molecular mechanisms are still poorly understood. We have previously characterized the conformational properties of Engrailed 2 homeodomain (EnHD) in aqueous solution and in micelles as membrane-mimetic environments. In the present study, we used small isotropic lipid bicelles as a more relevant membrane-mimetic model to characterize the membrane-bound state of EnHD. We show that lipid bicelles, in contrast to micelles, adequately reproduce the requirement of anionic lipids in the membrane binding and conformational transition of EnHD. The fold-unfold transition of EnHD induced by anionic lipids was characterized by NMR using ^1H , ^{13}C , ^{15}N chemical shifts, nuclear Overhauser effects, residual dipolar couplings, intramolecular and intermolecular paramagnetic relaxation enhancements induced by site-directed spin-label or paramagnetic lipid probe, respectively. A global unpacking of EnHD helices is observed leading to a loss of the native fold. However, near-native propensities of EnHD backbone conformation are maintained in membrane environment, including not only the three helices but also the turn connecting helices H2 and H3. NMR and coarse-grained molecular dynamics simulations reveal that the EnHD adopts a shallow insertion in the membrane, with the three helices oriented parallel to the membrane. EnHD explores extended conformations and closed U-shaped conformations, which are stabilized by anionic lipid recruitment.

Abbreviations

CD, circular dichroism; CG, coarse-grained; CSD, chemical shift deviation; DHPC, 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; DMPG, 1,2-dimyristoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol); 5-doxyl-PC, 1-palmitoyl-2-stearoyl-(5-doxyl)-*sn*-glycero-3-phosphocholine; DPC, dodecylphosphocholine; DPPC, dipalmitoylphosphatidylcholine; DPPG, dipalmitoylphosphatidylglycerol; EnHD, chick Engrailed 2 homeodomain (residues 200–259); MD, molecular dynamics; NMR, nuclear magnetic resonance; POPG, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol); POPS, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine; PRE, paramagnetic relaxation enhancement; RDC, residual dipolar coupling.

Keywords

Homeodomain, cell-penetrating peptide, bicelle, fold-unfold transition

1 Introduction

Among the numerous proteins able to interact with biological membranes, only a few are able to directly translocate across the membrane. Protein translocation across the plasma membrane is involved in unconventional cell internalization [1] or cell secretion pathways [2]. This peculiar translocation property has been identified in different classes of proteins, in particular homeoprotein transcription factors [3], viral proteins such as Tat from HIV [4], and basic fibroblast growth factors [5]. Although structurally different, these proteins share some common features, namely the presence of motifs enriched in basic residues interacting with anionic ligands (nucleic acids, lipids, glycosaminoglycans).

Homeoproteins are a family of transcription factors characterized by a 60 amino acid DNA binding homeodomain adopting a stable three-helix compact fold in solution [6] (Fig. 1A). Their membrane translocation property has been ascribed to the evolutionarily conserved third helix within the homeodomain and turns out to be a ubiquitous property of homeoproteins in animals and plants [7,8]. We have chosen the homeodomain of Engrailed 2 (EnHD) as a model to investigate the structural determinants of membrane translocation. The structure of this protein has been studied in detail under different environments [9–11]. Engrailed homeoprotein can act as a neuroprotective factor preventing the apoptosis of dopaminergic neurons [12–15]. Its translocation property has been shown to play a role in neuronal communication, and is involved in brain patterning [16] and in axonal guidance during retinal cortex establishment [17–19].

Structure/function studies have led to the discovery of one of the first described cell-penetrating peptides (CPPs), penetratin, derived from the drosophila *Antennapedia* homeodomain third helix [20]. Biophysical and cellular studies of CPPs have aroused considerable interest, not only for deciphering the intriguing mechanisms of membrane translocation but also to design efficient cell-penetrating peptidic vectors able to deliver hydrophilic cargoes within cells. However, since their discovery in the 1990s, the mechanisms of cell internalization are still a matter of debate. Depending on the type of CPP sequence, cell line and experimental protocol to monitor uptake, direct membrane translocation pathways and vesicular transport routes have been identified, the two being not mutually exclusive [21].

Contrasting with the wealth of biophysical data obtained on the isolated third helix of homeodomain (penetratin, most often) [22], studies dedicated to the full homeodomain remain very limited. Nevertheless, deciphering the mechanism of translocation of homeoproteins requires the study of the native conformation of the protein, which determines the folding and exposure of critical residues involved in translocation. In addition, additional binding sites within the homeodomain can cooperate with the penetratin motif to promote membrane interaction.

We have previously demonstrated by circular dichroism (CD) and Trp fluorescence that the interaction with anionic lipids triggers conformational changes in EnHD [11]. By using solution NMR, we reported that EnHD loses its native fold in micellar environments (DPC micelles), but retains near-native helical secondary structures. The loss of the three-dimensional fold enables optimal interactions of EnHD cationic residues with polar headgroups and the burial of Trp48 and Phe49 residues in the hydrophobic interior of the membrane mimic. The tryptophan residue, which is buried in the hydrophobic core of the homeodomain in aqueous environment, has been previously recognized as a critical residue in the internalization process [23]. These results led us to propose that this conformational change acts as a molecular switch to initiate membrane binding and subsequent translocation of homeodomains [11].

The choice of detergent micelles in our previous study was driven by the size of the membrane mimics compatible with solution NMR studies. However, the high curvature of micelles, their dynamics and their possible denaturing character could alter the structure and interaction properties of protein, in comparison with natural lipid environments [24]. Another concern was that zwitterionic micelles induce large conformational changes in EnHD while zwitterionic lipid membranes do not. This underlines that micelle systems are not adequate to investigate the role of electrostatic interactions with anionic lipids, which is recognized as a key element to promote EnHD conformational transitions.

For these reasons, we turn in this study to another membrane-mimetic system, bilayered phospholipid micelles (bicelles). Bicelles are mixed micelles systems, composed of a mixture of short-chain and long-chain phospholipids [25,26]. They can form disk-shaped assemblies in which the long-chain phospholipid segregates in a flat bilayer region and are surrounded by a rim of short chain phospholipid detergent. The bicelle size can be controlled by the long chain to short chain lipid ratio (q). Values of q below 0.5 enable bicelles to tumble

isotropically in solution and to have a size compatible with solution NMR studies. We show herein that small bicelles adequately reproduce the properties of lipids, regarding the conformational effect induced by electrostatic interactions. This led us to carry out a high-resolution liquid NMR investigation, in combination with coarse-grained molecular dynamics simulations, to investigate how the water-soluble EnHD protein becomes associated with a negatively charged membrane surface and to characterize its membrane-bound state. Our study shows that EnHD bound to the membrane interconverts between extended conformations and more closed conformations allowing the protein to recruit a high number of anionic phospholipids at the membrane surface.

2 Materials and methods

2.1 Protein expression and purification

Recombinant EnHD proteins were produced in *E. coli* SE1 strain transformed with pSCherry vectors. Protein constructs were expressed as Cherry-EnHD fusion proteins comprising a (His)₆ tag at the *N*-terminus and a PreScission protease cleavage site upstream EnHD sequence. Proteins were produced in LB or in M9 minimal medium for ¹⁵N or ¹⁵N, ¹³C isotopic labeling, as described [27]. The EnHD construct corresponds to residues 200–259 of chicken Engrailed 2 homeoprotein (UniProt Q05917). Two Cys-containing mutants were also produced, in position 6 (T6C substitution) or at the C-terminus (added residue C61). The purification protocol of all EnHD protein constructs was similar to the one described for Cys mutants in reference [11]. Briefly, bacterial pellets were resuspended in a lysis buffer containing 50 mM sodium phosphate pH 8.0, 500 mM NaCl, 20 mM imidazole, and 1 mM DTT. After sonication, the soluble fraction was separated from cell debris by centrifugation and loaded on a Ni-NTA column (GE Healthcare) previously equilibrated with the lysis buffer. The protein was purified using IMAC standard protocols with a 20–300 mM imidazole gradient. Eluted fractions were pooled and dialyzed against PBS buffer, and PreScission protease (GE Healthcare) was added to cleave the fusion protein. A second step of purification was further performed using the Ni-NTA column to remove the cleaved (His)₆-Cherry fusion partner. The purity of protein fractions was controlled by SDS-PAGE and protein concentration was determined from the UV absorbance measured at 280 nm using a theoretical extinction coefficient of 6990 M⁻¹ cm⁻¹.

2.2 Samples preparation

DHPC, DMPC, DMPG and 5-doxyl-PC paramagnetic probe were obtained from Avanti Polar Lipids. Lipids were lyophilized before weighing. DHPC was first dissolved in the NMR buffer (30 mM sodium phosphate, pH 6.5, containing 10 % D₂O) and the DHPC solution was then added to DMPC and DMPG lipid powders for bicelle preparation. Bicelle homogeneity was obtained by repeating 3 cycles consisting in freezing in liquid nitrogen, incubation at 45 °C and centrifugation (5 min, 4000 g), with vigorous vortexing before and after centrifugation [28]. The protein solution was added to bicelles and freezing/incubation cycles were repeated until complete homogeneity. Bicelles incorporating 1 % 1-palmitoyl-2-stearoyl-(5-doxyl)-*sn*-glycero-3-phosphocholine (5-doxyl-PC) were prepared using the same protocol.

The paramagnetic spin label MTSL (*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonylthioate) was covalently attached to two EnHD mutants containing a single cysteine residue at position 6 or 61. Removal of excess label was performed through desalting using a PD-10 column (GE Healthcare) and through extensive buffer exchange using centrifugal filter units with a 5-kDa membrane cutoff (Amicon). The yield of labeling was evaluated to be >80 % from MALDI-TOF MS analysis. MTSL-labeled EnHD mutants were concentrated at 0.2 mM in the NMR buffer containing 30 mM sodium phosphate pH 6.5 and 70 mM NaCl. 2D ¹H-¹⁵N HSQC experiments were recorded at 40 °C before and after 2 h incubation with 5 mM ascorbic acid. The comparison of 2D ¹H-¹⁵N HSQC spectra of EnHD Cys mutants (obtained after reduction of the spin label) with that of native EnHD shows little chemical shift perturbation in aqueous and bicelle environments, indicating that both Cys mutations and probe attachment do not strongly perturb EnHD conformational space (see Fig. S1 for EnHD T6C mutant in the presence of bicelles).

2.3 Fluorescence spectroscopy

Fluorescence spectra were recorded at room temperature (25 °C) on a MOS-200 spectrophotometer, with an excitation wavelength of 280 nm, a spectral width between 295 and 440 nm, and a bandwidth of 5 nm. A quartz cell (Hellma) of 0.3 × 0.3 cm path length was used, using a protein concentration of 25 μM and a volume of 130 μL.

2.4 NMR spectroscopy

NMR experiments were acquired on a Bruker Avance III 500 MHz spectrometer equipped with a TCI $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ cryoprobe. NMR data were processed with TopSpin 3.1 (Bruker) or NMRPipe [29] programs and spectra were analyzed with NMRFAM-SPARKY [30]. 2D $^1\text{H}-^{15}\text{N}$ HSQC experiments were run between 30 °C and 40 °C on [^{15}N]-EnHD samples prepared in 5 mm Shigemi NMR tubes. The buffer used for NMR samples was 30 mM sodium phosphate, pH 6.5, 70 mM NaCl, and containing 10 % D_2O .

^1H , ^{13}C , ^{15}N NMR assignments of [^{15}N , ^{13}C]-EnHD in bicelles made of 50 mM DHPC and 25 mM DMPG were obtained from 3D HNCACB, CBCA(CO)HN, HBHA(CO)NH, HNCO, 3D ^{15}N TOCSY-HSQC and 3D ^{15}N NOESY-HSQC recorded at 37 °C or 40 °C. Chemical shifts have been deposited at BMRB under accession number 51516.

The ^{15}N longitudinal R_1 and transversal R_2 relaxation rate constants were determined by collecting a series of sensitivity enhanced $^{15}\text{N}-^1\text{H}$ HSQC spectra with different relaxation delays. The relaxation delays were set to 20, 50, 100, 135, 175, 200, 300, 500, 750, 1000, 2000, 5000 ms for R_1 measurements, and 17, 34, 51, 68, 85, 102, 136, 170, 187, 220, 237 ms for R_2 measurements. Steady state $^{15}\text{N}\{^1\text{H}\}$ NOE values were determined as the ratio of peak heights in $^{15}\text{N}-^1\text{H}$ spectra collected with and without proton saturation during the recycle delay. ^1H saturation was achieved by applying a train of 120° pulses spaced at 5 ms interval for a period of 4 s.

2D $^1\text{H}-^{15}\text{N}$ HSQC experiments of [^{15}N]-EnHD in the presence of 5-doxyl-PC lipid paramagnetic probe or MTSL covalently attached to Cys7 or Cys61 of EnHD mutant were recorded at 40 °C with a long relaxation time of 5 s. Reference diamagnetic HSQC experiments were recorded with identical acquisition parameters, in the absence of 5-doxyl-PC or after reduction of MTSL spin label with 5 mM ascorbic acid. Paramagnetic relaxation enhancements (PREs) were determined as the ratio of crosspeak intensities in the paramagnetic and diamagnetic conditions.

Residual dipolar couplings were extracted from 2D IPAP $^1\text{H}-^{15}\text{N}$ HSQC experiments recorded on isotropic and anisotropic samples. The alignment was induced by a radially compressed negatively charged polyacrylamide gel. The EnHD/bicelle solution was soaked into a 6 % (w/v) polyacrylamide gel (acrylamide/bisacrylamide ratio of 39:1 w/w) doped with 5 % 2-acrylamido-2-methyl-1-propanesulfonic acid (AMPS). The gel of 6.1 mm initial diameter was then inserted into an NMR tube of 4.2 mm internal diameter.

2.5 NMR structure calculations

NMR-based models of EnHD in bicelles were calculated with XPLOR-NIH 2.52 program. Dihedral angle restraints were defined using TALOS-N program [31]. The ϕ , ψ angles of 46 residues were restrained to TALOS-N predicted values, the predicted range being increased by a factor of 3, with a minimal $\pm 15^\circ$ range. The ϕ angle of all other residues was restrained to negative values (-180° to 0°). Distance restraints were derived from 243 NOE cross-peaks identified on 2D, 3D ^{15}N -edited, and 3D ^{13}C -edited NOESY experiments. Positional restraints were applied on the Z coordinates of atoms using distance restraints to planes perpendicular to the Z axis [32]. First, the $\text{C}\alpha$ atom positions of all residues were restrained to $22 \text{ \AA} \pm 7 \text{ \AA}$, assuming that all protein atoms lie close to the water-membrane interface. From intermolecular paramagnetic relaxation enhancements induced by 5-doxyl-PC, two sets of restraints were applied on the Z coordinates of the N atoms belonging to surface residues ($25.5 \text{ \AA} \pm 3.5 \text{ \AA}$) and to buried residues ($18.5 \text{ \AA} \pm 3.5 \text{ \AA}$). Finally, the positions of aromatic protons involved in intermolecular NOEs with internal methylenic protons of lipid acyl chains were restrained to $14 \text{ \AA} \pm 5 \text{ \AA}$. A set of 200 structures was calculated by torsion angle dynamics at 3500 K followed by simulated annealing to 25 K and minimization, using parallhdg5.3 forcefield [33]. The 25 lowest energy structures were then refined in a second stage of torsion angle dynamics (dynamics at 3000 K, simulated annealing to 25 K and minimization) using an implicit membrane potential implemented in EEFx2 forcefield [34,35]. The best 10 conformers were selected to represent the final conformational ensemble. The coordinates of these models are available upon request to the authors.

2.6 Molecular dynamics

Coarse-grained (CG) molecular dynamics (MD) simulations of the EnHD protein in interaction with a lipid bilayer were carried out using MARTINI force field version 2.2 [36,37] and GROMACS 4.5.3 [38].

2.6.1 System setup. To get insights into the early steps of protein binding to the membrane, a first system containing one protein in the aqueous phase far enough ($\sim 35 \text{ \AA}$) from the bilayer was built, so that it had time to freely diffuse before encountering the membrane. This first simulation box had a volume of $12 \times 12 \times 15 \text{ nm}^3$ and contained a lipid bilayer made of 408 DPPC and 104 DPPG lipid molecules (giving a DPPC/DPPG 80:20 M ratio), 13,054

water particles (each water particle corresponds to 4 water molecules), one protein NMR conformer (taken from 3ZOB structure family) [11] and 96 Na⁺ ion particles to ensure system neutrality. The whole system contained 19,432 particles, which translates to a water thickness of ~10 nm. Three MD trajectories of 2 μ s length were run, starting with different initial velocities (simulations 1.1, 1.2 and 1.3). To keep the tertiary fold and the secondary structures intact, an elastic network was used [39]. This strategy allowed us to predict the early steps of the protein binding to the membrane, prior to the unfolding stage and full insertion. For each trajectory, protein binding to the membrane was found to be irreversible (no unbinding event was observed) and involved the same face of the protein.

Since the protein was irreversibly bound to the bilayer at this stage, a large water thickness was no longer necessary. To alleviate the computational effort, a second system was built (from the reproducible bound conformation), by reducing the size of the water layer (5717 beads, water thickness of ~4 nm), giving a total of 12,095 particles and a box volume of 12 \times 12 \times 9 nm³. The lipid composition and number of ions were identical to the first simulations. Two long MD trajectories of 60 μ s were run, started with different initial velocities (simulations 2.1 and 2.2). The elastic network was released in order to let the protein free to fully insert in the bilayer. Importantly, the regular helical secondary structures were enforced, as is standard practice for MARTINI membrane protein simulations without elastic network [37].

2.6.2 Simulation parameters. Simulations were run at 300 K using the velocity rescaling thermostat [40] and at 1 bar using the Parrinello-Rahman barostat [41]. A 1.2-nm cutoff was used for non-bonded interactions. The integration time step was 20 fs. Additional details are given in Supplementary Information.

2.6.3 Trajectory analysis. Details on the analysis of protein partitioning within the bilayer, protein conformation clustering and neighboring lipids are given in Supplementary Information.

3 Results

3.1 Investigation of EnHD conformational changes using small isotropic bicelles

To gain insight on the influence of membrane environments on EnHD conformational changes, we first conducted intrinsic tryptophan fluorescence experiments (Fig. 1A). EnHD

contains a single Trp residue (Trp48) whose burial in the hydrophobic core of the homeodomain makes it a good reporter of EnHD conformational transitions. Furthermore, this residue is known to be critical in lipid interactions that mediate membrane translocation [23]. The Trp fluorescence emission spectrum of EnHD is strongly quenched in the native state, owing to aromatic interactions within the hydrophobic core [42] or to a dipolar interaction with an immobilized water molecule [10]. We previously reported that the fluorescence spectra of EnHD are virtually identical in aqueous solution and in the presence of large unilamellar vesicles (LUVs) made of zwitterionic DMPC, whereas the addition of anionic DMPG lipids in LUVs leads to an increase in fluorescence intensity [11]. We ascribed these spectroscopic changes to a large conformational transition affecting EnHD three-dimensional structure concomitant with an insertion of tryptophan in the lipid bilayer [43].

We analyzed the effects on EnHD of bicelles made of short chain DHPC and long chain DMPC or DMPG phospholipids, using a q ratio of 0.5. Interestingly, the effects observed with bicelles match those induced by LUVs: the Trp environment is marginally affected by zwitterionic bicelles, while bicelles containing anionic DMPG induce large environment changes, similarly to LUVs (Fig. 1A). This indicates that the composition of long chain lipids (DMPC/DMPG) in bicelles influences Trp environment in EnHD to a similar extent as in LUVs. As a control, we investigated the effect of DHPC alone, at a concentration above the critical micellar concentration (CMC). The DHPC micelle induces large environment change in EnHD, in the absence of any anionic lipid, showing that DHPC detergent behaves differently, depending whether it is micellar or associated within lipid bicelles. It should be noted that the fluorescence emission of Tyr25 may partly overlap with that of Trp if the tyrosinate form is stabilized, which is unlikely in the presence of anionic lipids [44,45].

2D ^1H - ^{15}N HSQC NMR experiments were next recorded to analyze the effects of the bicelle composition on EnHD conformation and interaction (Fig. 1B). The ^1H - ^{15}N HSQC spectra of EnHD are very similar in aqueous solution and in the presence of zwitterionic bicelles ($q = 0.5$), as already observed [11]. The presence of zwitterionic lipids has minor effects on ^1H , ^{15}N chemical shifts and peak intensities. These weak perturbations can be interpreted as transient interactions between EnHD and lipid polar head groups, which do not affect EnHD three-dimensional structure. The effect of DHPC alone is markedly different at a concentration above the CMC, as it induces dramatic chemical shift perturbations and narrower chemical shift dispersion, indicative of a loss of EnHD three-dimensional structure

(Fig. 1B). The titration with bicelles containing increased proportions of anionic DMPG phospholipid (Fig. 1C) leads to large environment changes in EnHD in terms of chemical shift perturbation and peak intensities, confirming the effects observed in Trp fluorescence experiments. At DMPG ratios around 25–50 %, a severe broadening of resonances is observed for many residues, including Trp48 side chain HN ϵ 1 resonance. This effect could be due to the size of the protein/bicelle complex and/or to intermediate exchange regime between the native form in solution and the lipid-bound state. However, the titration with higher concentrations of DMPG (75–100 %) enables recovering spectra of better quality, indicating that an intermediate exchange regime of interaction with DMPG is probably at stake. We therefore chose to characterize the DMPG-bound state of EnHD using bicelles made of 50 mM DHPC and 25 mM DMPG, a condition that yields a spectrum quality compatible with high-resolution solution NMR studies.

3.2 NMR characterization of EnHD in complex with anionic bicelles

The 2D ^1H – ^{15}N HSQC spectrum of EnHD in DHPC/DMPG bicelles strongly differs from that in aqueous solution and exhibits reduced chemical shift dispersion. These features indicate that bicelle-bound EnHD does not adopt the native globular fold. To further characterize the conformation of EnHD, ^1H , ^{15}N , and ^{13}C resonances were assigned using double and triple resonance experiments. The analysis of these chemical shifts provides valuable information to characterize the conformation at the residue level. The chemical shift deviations (CSDs) of $^{13}\text{C}\alpha$ resonances with respect to random coil values (Fig. 2A) show that the helical secondary structure of EnHD is largely preserved. Three stretches of residues exhibiting positive CSD values (> 1.5 ppm) are observed, supporting the presence of three helices. The CSD profile is close to that observed in aqueous environment (Fig. S2), but slightly differs regarding the length of helices and in the connecting loops. A more detailed analysis of residue backbone conformation was provided by TALOS-N program, based on ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, $^{13}\text{C}'$, and $^1\text{H}\alpha$ chemical shifts [31]. Conformational predictions could be made by TALOS-N for 46 residues ranging from T9 to K55 (with the exception of L26), other residues in the N- and C-tails being predicted as dynamic by the program (Fig. S3). The three helical segments can be precisely delineated: H1 extends from A10 to Y25, H2 from E28 to E37, and H3 from E42 to K55. This definition of helices is remarkably similar to that in aqueous solution. A slight difference can be observed for H1 helix, which is one turn shorter in its C-terminal part in water (A10 to

E22). Interestingly, several residues in the helix connecting segments adopt backbone conformations in the same regions of the Ramachandran diagram as in aqueous solution: extended basin (β/P_{II}) for T9, T27, L40, N41; α_R region for L38 and α_L for G39. Thus, the overall secondary structure of EnHD tends to be very similar in water and in bicelle environment. The combined analysis of ^1HN , ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}'$ chemical shift variation [46] induced by the bicelles shows that perturbations are observed throughout the sequence, and are the strongest in the loops connecting α -helices (Fig. S4).

The assignment of NOE correlations in a 3D ^{15}N - and 3D ^{13}C -edited NOESY experiment (Fig. S5) further supports the chemical shift-based analysis. Indeed, the comparison of the intensities of intraresidual $\text{H}\alpha_i\text{-HN}_i$, sequential $\text{H}\alpha_i\text{-HN}_{i+1}$, and sequential $\text{HN}_i\text{-HN}_{i+1}$ NOEs is in good agreement with predicted backbone propensities at the residue level. Importantly, the observation of several medium range $\text{H}\alpha_i\text{-HN}_{i+2}$, $\text{H}\alpha_i\text{-HN}_{i+3}$, $\text{H}\alpha_i\text{-H}\beta_{i+3}$ and $\text{H}\alpha_i\text{-HN}_{i+4}$ NOEs demonstrates the persistent formation of helices in the bicelle environment. Nevertheless, no long-range NOEs could be detected between helices, suggesting that the helices are probably not tightly packed in a globular structure.

To gain insight on EnHD backbone dynamics, heteronuclear $\{^1\text{H}\}^{15}\text{N}$ NOEs were measured (Fig. 2B), together with ^{15}N R_1 and R_2 relaxation rates (Fig. S6). Residues in the core of the homeodomain (T9-K55) have heteronuclear $\{^1\text{H}\}^{15}\text{N}$ NOE values around 0.5 at 500 MHz. These values are lower than those previously determined in aqueous solution (around 0.8 at 500 MHz) [11], reflecting a higher mobility on the picosecond to nanosecond timescale in comparison with the globular state in water. However, they indicate that the helices are stable on this timescale. In contrast, residues outside the homeodomain core are highly disordered, confirming the analysis based on chemical shifts. Residues in the segments connecting H1-H2 and H2-H3 exhibit values that are only slightly smaller than those observed in the central part of the helices, indicating that they also experience restricted motions. ^{15}N R_1 and R_2 relaxation rates show quite uniform values in the homeodomain core (Fig. S6). The N- and C-terminal tails show decreased R_2 values, in agreement with the higher mobility of these segments inferred from heteronuclear NOE values. Moreover, the ^{15}N R_1/R_2 ratio measured in the helical segments provided an estimation of the rotational correlation time of the protein-bicelle complex to 15 ns, almost 4 times longer than that of EnHD in solution (~ 4 ns). This increase supports a tight binding with anionic bicelles.

One-bond residual dipolar couplings (RDCs) were measured to get information on the relative orientation of helices. In an anisotropic medium, internuclear dipolar couplings no longer average to zero and provide information on the average angular orientation of bond vectors relative to the magnetic field. The alignment was induced by an anisotropically compressed polyacrylamide gel, a medium compatible with bicelle formation [47]. Only one-bond ^1H – ^{15}N dipolar couplings could be successfully measured from 2D IPAP HSQC experiments, because 3D experiments used to measure other one-bond couplings had a low signal to noise ratio, under our experimental conditions. ^1H – ^{15}N RDCs measured in the different helix segments clearly show distinct ranges of values (Fig. 2C), suggesting that the helices do not behave as independent rods with random orientations.

Since the aforementioned NMR parameters provide mainly local information, we next measured paramagnetic relaxation enhancements (PREs) induced by intramolecular site-directed spin labeling. The PRE analysis is a convenient method to detect transient long-range contacts between distant regions of non-globular proteins, owing to the strong dipolar interaction between radical electron spin and nuclear spins. A nitroxide radical was covalently attached by Cys linkage chemistry to EnHD mutated with a cysteine at position 6 (upstream of helix H1) or 61 (C-terminal extremity). The PRE profiles measured for EnHD in bicelle environment are markedly different from those obtained in aqueous solution (Fig. 3). The introduction of a spin label at position 6 yields strong PREs in distant parts of EnHD in aqueous solution (Fig. 3B), with cancellation of residue intensity in helix H3, as expected from the globular 3D structure (Fig. 3A). These long-range effects are no longer observed in bicellar environment and prove the loss of the native fold (Fig. 3C). A weak effect is observed in the segment connecting H2 and H3 helices, indicating transient contacts. The PRE profiles obtained with the spin label grafted at the C-terminus yield a similar conclusion: the close spatial proximity between the C-terminus and the loop between H1 and H2 in the native fold is lost in lipid environment (Fig. 3D,E). Nevertheless, PREs are observed for residues 8–9 preceding H1 helix, suggesting the presence of long-range transient contacts between the N- and C-terminal parts of the protein in the bicelle-bound state.

3.3 Position of EnHD in anionic bicelles

The position of EnHD with respect to the bicelle surface was examined using intermolecular PREs induced by the lipophilic paramagnetic probe 5-doxyl-PC (Fig. 2D). Most EnHD residues

are affected by the probe, except for a few residues at the N- and C-termini. This indicates a strong association of EnHD core with anionic bicelles, whereas the disordered tails remain flexible in bulk solution. Periodic variation of PREs are observed along the three helices. These paramagnetic waves are typical of parallel orientation of the helices with respect to the bicelle surface. Within each helix, the most affected residues have a hydrophobic character (L13, L16, F20 in helix H1, L34 in helix H2, W48, F49, A54 in helix H3), corresponding to an orientation of the hydrophobic face of helices within the hydrocarbon phase. The most affected residues in EnHD are A7 and F8, which lie in a flexible segment preceding helix H1. This segment is expected to be embedded within the bicelle interior, on the basis of side chain hydrophobicity. However, considering the strong signal attenuation of ^1H - ^{15}N resonances, we cannot exclude a specific interaction of these residues with the paramagnetic doxyl group.

The analysis of NOESY experiments also revealed the presence of intermolecular NOEs between aromatic protons of EnHD (F8, F20, F49, W48, Y25) and lipid protons around 1.25–1.3 ppm that correspond to internal methylenic groups of lipid acyl chains. These intermolecular NOEs indicate that the aromatic side-chains of EnHD are embedded in the hydrophobic interior of bicelles.

3.4 NMR structure calculation of EnHD membrane-bound state

The structure of EnHD bound to anionic bicelles appears to be more dynamic than the native state in aqueous solution but clearly shows important degree of order at the backbone level, with stable helix structuring and positioning in the bicelle. To describe the conformational space of the EnHD membrane-bound state, models were calculated by restrained molecular dynamics using the available NMR data. Experimental NMR information includes 46 (ϕ , ψ) dihedral angle restraints obtained from TALOS-N analysis, 243 distance restraints from assigned NOEs. To further take into account the parallel positioning of helices with respect to the bicelle surface, we used an atom to plane distance potential available in XPLOR-NIH structure calculation program [32]. These positional restraints were applied to N atoms and were defined on the basis of intermolecular PREs induced by the 5-doxyl-PC probe. Finally positional restraints were applied to the aromatic sidechains, on the basis of observed intermolecular NOEs with lipids. No RDCs were used in the structure calculation because of the limited experimental data set (single alignment condition).

The structures were calculated with a high temperature simulated annealing protocol using torsion angle dynamics in XPLOR-NIH and were refined using an implicit membrane potential (EEFx2 forcefield) to provide a description of the anisotropic water-membrane interface. The final NMR ensemble is shown in Fig. 4. The relative orientation between the three helices is not well-defined, reflecting protein flexibility in the bicelle environment. However, helices H2 and H3 adopt a privileged helix-turn-helix topology, while helix H1 has a looser orientation. Some conformers show close proximity of the N- and C-terminal regions, in agreement with transient contacts detected in PRE experiments.

3.5 Coarse-grained molecular dynamics simulations of EnHD with lipid membranes

To get more detailed insights on EnHD interaction with lipid membranes, we performed coarse-grained (CG) molecular dynamics (MD) simulations of EnHD in an explicit membrane made of 80 % DPPC and 20 % DPPG. We chose to use a binary lipid composition that differed from the purely anionic bicelle condition used in the NMR study to get information on EnHD - lipid interactions in a more realistic membrane environment made of both zwitterionic and anionic lipids. In a first set of simulations, three MD trajectories of 2 μ s each were run with a hydrated phospholipid bilayer. An elastic network was applied to keep the protein fold intact. The protein was initially positioned in the aqueous phase, far apart from the membrane bilayer (Fig. S7). The protein starts to diffuse in water in a first phase, which is then followed by an attraction towards the membrane. As observed previously on other peripheral proteins such as phospholipase A2 [48] or HIV-1 Matrix Protein [49], the protein undergoes multiple unsuccessful collision events on the membrane. The protein starts rapidly (after a few tens of ns) to “sense” the membrane via its N-ter and C-ter basic residues. At this stage, this is not a tight binding event, as no residue is buried within the membrane, but rather a weak adsorption driven by electrostatic interactions. Eventually, EnHD binds irreversibly to the membrane surface, after 0.4 to 0.7 μ s, with a similar scenario in the 3 simulations (Fig. S8). This binding occurs between the basic C-terminus of helix H3 (K₅₅IKKAT₆₀) and the membrane, mainly via stronger electrostatic interactions with DPPG lipids. The binding of the C-terminal part of EnHD provides a “soft” anchoring to the polar heads, while permitting the protein to rotate and engage transiently other contact surfaces. This C-terminal anchoring eventually leads to a more stable orientation, which is reproducible over the 3 simulations. Multiple contact points can be observed in Fig. S9,

involving not only the C-terminus of helix H3 (K₅₅IKK₅₈), but additional residues in the cationic N-tail of EnHD (K₂RPR₅) as well as at the end of helix H2 (Q₂₁, T₂₂). In this bound state, helix H3 has the deepest insertion and is oriented perpendicular to the membrane surface. The folded EnHD structure prevents an optimal binding of helices H1 and H2 with the membrane surface. Therefore, the protein is not deeply buried inside the membrane and only the C-terminal residues partition below the phosphate groups of lipids (Fig. S9).

In a second step consisting in two MD simulations of 60 μ s duration (simulations 2.1 and 2.2), the elastic network was released in order to allow the homeodomain to unfold and insert deeper in the membrane. It is important to bear in mind that the regular helical secondary structure of EnHD, inferred by NMR on the bicelle-bound form of EnHD, was enforced at this stage, as is standard practice with MARTINI simulations of membrane proteins [37]. A partial unfolding of EnHD is observed within the first few μ s, concomitant to the insertion of helices H1 and H3 in the membrane, with orientations parallel to the membrane surface. Complete unpacking of EnHD helices occurs at 5.6 μ s and 13.7 μ s (for simulations 2.1 and 2.2, respectively), leading to the embedding of the three helices within the membrane. All figures and statistics described hereafter consider the range 15–60 μ s after the full unfolding/insertion have occurred for both simulations. The position of the helices within the membrane then becomes very stable, with a parallel positioning and comparable immersion depths over the 2 simulations (Fig. 5). The protein is much more buried inside the membrane compared to the first phase (Fig. S9). Indeed, many hydrophobic residues are deeply inserted below the phosphate level all along the protein sequence. This parallel orientation and deep insertion of some residues for the 3 helices is possible thanks to their non-negligible amphipathic character.

Once inserted, EnHD structure is highly dynamic and explores various orientations between its helices in the membrane plane. The conformational space was analyzed by clustering using a cutoff of 8 Å (Fig. 6). The first populated cluster (~45 % of all conformations) corresponds to an extended thread-like structure in which the three helices are roughly parallel to one another. The second cluster (21 %) displays a more compact U-shaped structure with turn conformations on either side of helix H2. This conformation leads to closer distances between the N- and C-termini and could account for the transient contacts detected by PRE NMR experiments. Clusters 3 and 4 have intermediate shapes and are probably transition conformations between clusters 1 and 2. In the remaining of the text, we

will call the two main clusters 1 and 2 “extended cluster” and “U-shaped cluster” respectively.

Next, we examined the dynamics between the extended and U-shaped structures. The distance fluctuation between residues 5 and 60 is shown in Fig. 7A for simulation 2.1 (and Fig. S10 for simulation 2.2). A low value, below 2.5 nm, is specific to the U-shaped conformation (cluster 2), while a value higher than 2.5 nm is represented by the extended structures (cluster 1) and the intermediate clusters 3 and 4. Extended conformations are dominant, although several short excursions to a low distance are observed, which go back rapidly to extended conformations. Interestingly, the U-shaped is stabilized around 40 μ s for about 6–7 μ s. Similarly, in simulation 2.2 (Fig. S10), a long-lived state of the U-shaped conformation is observed (around 20 μ s for roughly 5 μ s). A life time of several μ s implies that some interactions should stabilize this U-shaped conformation. The number of neighboring lipids around EnHD protein is shown in Fig. 7B. For extended conformations, there are around 90 DPPC lipids and 60 DPPG lipids having a contact with the protein. Interestingly, the number of DPPC goes down to 60 lipids around 40 μ s, while the number of DPPG remains constant at 60 lipids. This reduction in the number of DPPC near the protein is concomitant with the occurrence and stabilization of the U-shaped conformation. The fraction of DPPG around EnHD protein is significantly enriched in comparison to the total DPPG fraction in the whole box (Fig. 7B). This indicates that DPPG clusters around the positively charged EnHD. The fraction of bound DPPG amounts to \sim 40 % for extended conformations, in comparison to 20 % for the whole box. Interestingly, the fraction of DPPG is higher for the U-shaped cluster, reaching \geq 50 %. The snapshots shown in Fig. 7D provide a molecular basis for the stabilization of the U-shaped conformation. Indeed, this conformation is possible thanks to the recruitment of negatively charged DPPG lipids, allowing the positively charged N- and C-terminal parts of the protein to approach each other. The U-shaped conformation is more compact than the extended one, thus some lipids need to be expelled from the protein surface to interconvert from U-shaped to extended. These expelled lipids turn out to be solely DPPC.

4 Discussion

This combined NMR/MD study provides for the first time a detailed analysis of EnHD conformational space in lipid environments, which is a prerequisite to decipher at the molecular level the mechanisms of membrane translocation of native homeoproteins. We had previously characterized by NMR the conformation of EnHD in two detergent micelles (zwitterionic DPC and anionic SDS) [11]. However, the results gathered in detergent environments should be interpreted with caution because the potential protein denaturing effects of micelles may affect the stability and dynamics of helical secondary structures, as well as alter the interactions between helices, as described for several transmembrane [24] or surface-bound proteins [50]. In this respect, we observed that zwitterionic phosphocholine detergent micelles (DPC [11] and DHPC in this study) cause EnHD unfolding whereas zwitterionic lipid bilayers do not. The incorporation of anionic lipids in bilayer membranes appears to be a critical factor to induce a major conformational transition of homeodomain. Moreover, membrane negative charge density is a key element for in vitro observation of direct translocation of cationic cell-penetrating peptides in liposomes [51]. Therefore, it is important that membrane mimics used for structural studies adequately reproduce the conformational effects triggered by protein-lipid electrostatic interactions.

We found that small bicelles behave like liposomes, regarding the membrane charge effect on EnHD conformational transition. Low q ratios had to be chosen to form fast-tumbling bicelles compatible with solution NMR, requiring the use of high DHPC detergent concentrations. One possible concern with such small bicelles ($q < 1$) is the incomplete lipid segregation from detergent which may result in mixed micelle behavior rather than true lipid bilayer formation [52]. However, despite the high proportion of DHPC detergent and low q ratio, we show that zwitterionic phosphatidylcholine bicelles do not induce any EnHD protein unfolding, in contrast with pure DHPC micelles. The incorporation of anionic DMPG lipids in bicelles has dramatic effects on ^1H - ^{15}N NMR spectra of EnHD, indicative of strong binding to anionic lipids associated with major conformational changes in EnHD. Significant line broadening is observed at intermediate DMPG concentrations (25–50 %), presumably arising from the exchange between the solution and lipid-bound state or between multiple lipid-bound conformers. Importantly, little chemical shifts variation is observed for DMPG ratios beyond 50 %, while significant line narrowing enables recovery of all ^1H - ^{15}N correlations. This can be explained by a shift of the equilibrium towards a major lipid-bound state of EnHD, reducing exchange-broadening contributions. The weak chemical shifts

variations between 75 % and 100 % DMPG suggest that the structure of the bicelle-bound state does not vary much in this range of high DMPG ratio. Therefore, we chose to use the highest DMPG ratio to characterize in detail the structure of the anionic bicelle-bound state. The rotational correlation time determined from ^{15}N relaxation provides an estimation of the size of EnHD/anionic bicelle complex, which is roughly twice as big as that of previously characterized EnHD/micelles complexes [11].

To get more insight on the protein-lipid interactions, we performed molecular dynamics simulations of EnHD in an explicit lipid membrane bilayer. Since previous experimental kinetics experiments showed that the homeodomain folds or unfolds on a μs time scale in aqueous solution [53,54], we anticipated an even slower kinetics in a lipid environment, precluding the use of all-atom simulations. Therefore, we used here a coarse-grained representation with the MARTINI force field. This latter has been parameterized to accurately reproduce the partition of molecules between polar and apolar phases but does not handle the full backbone flexibility. In order to observe the unfolding transition after the protein insertion, we used the trick of releasing the elastic network while maintaining helices in place. Accordingly, we do not expect to get the exact mechanism of the coupling between membrane binding and protein unfolding. The real kinetics between the different protein conformational states is also overestimated in our CG simulations, because of the inherent acceleration due to the CG representation. However, CG simulations are extremely useful to study how some lipids are specifically enriched near the protein [55], as shown herein for Engrailed interaction with PG. Furthermore, our strategy allowed us to sample how the 3 helices of EnHD partition into the membrane and interact with lipids at a near atomic resolution.

NMR data show that homeodomain binding to anionic lipid membrane leads to a loss of the globular native fold, while preserving the secondary structure. The three helical segments of EnHD have stable conformations in the lipid environment, as evidenced by chemical shift and NOE analysis. A similar observation had also been made in DPC and SDS micelles. Thus, the three helical regions of the homeodomain retain their high helical propensity in a wide variety of environments, with only minor adjustments of helices length. The main differences concern helix H1, which tends to be one turn longer in its C-terminal part in both micelles and lipid environments. The length of H3 helix is also variable on its C-terminal side. The 56–60 region exhibits helical propensity and is stabilized to different extent, depending

on the environment or upon complexation to its DNA ligand [56]. Interestingly, the NMR chemical shift analysis also reveals that residues outside helical segments explore a conformational space which is close to the native one, in particular in the H2-H3 turn (residues 38–41).

The three helical segments of EnHD are inserted in the membrane with parallel orientations to the membrane surface, as revealed by positioning experiments using a lipid paramagnetic probe. The analysis of helix positioning in simulations is also in excellent agreement with NMR results. MD trajectories underline that the three helices are not deeply buried in the lipid membrane and are inserted at the level of lipid head groups. Most residues on the immersed helix faces have aliphatic or aromatic side chains (L13, L16, F20 in helix H1, L34 in helix H2, I45, W48, F49 in helix H3). Segments connecting helices are also anchored in the membrane via hydrophobic contacts (P4 and F8 in the N-tail, Y25 and L26 in the H1-H2 connecting segment, L38 and L40 in the H2-H3 connecting segment, I56 in the C-tail). These anchoring points may contribute to the restriction of the available conformational space, as inferred from NMR chemical shift and relaxation analysis, not only in the three helices but also in helix-connecting segments. There do not appear to be strong differences in the positioning of the three helices. Yet, the membrane translocation property is specifically mediated by helix H3 of homeodomains. This helix is not deeply buried in the lipid bilayer and this observation is in agreement with data obtained on cell-penetrating peptides derived from H3 helices of different homeodomains [22]. Helix H3 sequence departs from the other two helices by a higher number of hydrophobic residues and, most importantly, by a high density of cationic residues. Both types of residues have been recognized as critical for the internalization. The helix structuring and positioning leads to the expected partitioning of most hydrophobic and polar residues in the lipid apolar phase and in the lipid polar head group/water interface, respectively. However, helix H3 is not perfectly amphipathic, so that K52 side chain is oriented on the lipid-embedded face, interacting with W48/F49 hydrophobic core. Positions 48 and 49 appear to be strictly conserved in homeodomain sequences while a cationic residue is most often found at position 52 [6]. It would be interesting to investigate if these conserved aromatic/cationic interactions play a role in the membrane translocation activity of homeodomains, as proposed for ion pair- π interactions described in Arg, Trp-rich cell-penetrating peptides [57].

Although EnHD loses its native globular fold upon interaction with bicelles, its conformational space cannot be merely described by extended conformations in which the three helices have random orientations with respect to one another. Indeed, the RDC profile indicates that helices H2 and H3 have distinct orientations in the bicelle complex. The NMR ensemble shows that the segment connecting H2 and H3 has significant turn propensity. In addition, PRE data reveal the presence of transient long-range contacts in EnHD between the N- and C-tails. Such non-extended structures might be experimentally induced by the small size of the bicellar membrane-mimetic system. However, we found that U-shaped conformers were also encountered in MD simulations. The lipid/water box dimensions undoubtedly provide a much larger lipid bilayer surface to accommodate EnHD, in comparison with experimental small bicelles. The long lifetime of such U-shaped states, reaching several μs in both simulations, also argues for the relevance of such conformations and the presence of stabilizing interactions. Close inspection of protein/lipid interactions in these U-shaped states provides a basis for such stabilization. Indeed, the U-shaped conformations are characterized by the entrapment of anionic lipids between the cationic N- and C-tails. Such conformations correspond to the second most-populated conformational cluster in the two MD simulations. However, accurate estimation of their population would require free energy calculations using enhanced sampling techniques. The transition between extended and U-shaped conformations is rather slow since we observe only one long excursion in both simulations 2.1 and 2.2. The real kinetics between both main conformations is overestimated in our simulations, due to the CG representation, but it probably reaches tens to hundreds of μs , or possibly the ms time scale.

The calculated conformational ensemble of EnHD bound to anionic bicelles has similarities with the structure of an Engrailed folding intermediate in aqueous solution, which was characterized by engineering a destabilizing mutation in homeodomain hydrophobic core (L16A) [54]. The L16A mutant of *Drosophila* Engrailed homeodomain has near-native secondary structure around H1, H2 and H3 helices, with a lengthening of helix H1 and a shortening of helix H3 in their C-terminal parts, in comparison with the wild-type protein. The L16A mutant exhibits a U-shaped structure, in which helices H2 and H3 form a well-defined, near-native helix-turn-helix structure while helix H1 does not interact with H2-H3 core. In comparison with this folding intermediate, bicelle-bound EnHD structure is less well-defined in the H2-H3 turn segment and the orientation between H2 and H3 helices is altered

to ensure more favorable parallel helix positioning to the membrane surface and anchoring of side-chain hydrophobic residues within the membrane.

Our structural data were obtained using a high proportion of anionic DMPG, a lipid that is not found in eukaryotic plasma membranes but is commonly used in biophysical studies of membrane-active peptides. Importantly, the translocation of cell-penetrating peptides can be experimentally observed *in vitro*, only when membrane systems incorporate a high density of anionic lipids [51,58]. For instance, the translocation of penetratin across droplet interface bilayers requires the presence of at least 40 % POPG and is not observed if POPG is replaced by POPS, another anionic lipid. This indicates that the nature of the lipid greatly influences the formation of translocation-active peptide-lipid complexes. Among the physiologic anionic lipid species that could be recruited by homeodomain, polyphosphoinositides appear to be good candidates. Indeed, it was recently shown that phosphatidylinositol-4,5-bisphosphate (PIP₂) is required for Engrailed 2 homeoprotein translocation, in both secretion and internalization transfers [59].

The negatively charged lipids (phosphatidylserine, PIP₂) are mostly found in the cytosolic leaflet of plasma membrane, making them privileged binding partners involved in the extracellular secretion of basic proteins. Accordingly, PIP₂ was shown to play a role in the unconventional secretion pathways of FGF2 [5] and Engrailed 2 proteins [59]. A recent study demonstrated that PIP₂ was also involved in the internalization of Penetratin peptide [60]. Interestingly, both inner leaflet and outer leaflet pools of PIP₂ modulate penetratin internalization [60]. The interaction of cell-penetrating peptides with membranes may also cause a redistribution of lipids across leaflets, as evidenced for Tat peptide and phosphatidylserine [61].

The electrostatic interactions between EnHD protein and lipids appear to be important not only to increase membrane binding affinity but also to induce EnHD unfolding. Our MD simulations show a clustering of anionic lipids around the positively charged EnHD, with a ~2 fold DPPG enrichment in the lipid fraction in contact with EnHD. The anionic lipid clustering is a common feature of cationic membrane-active peptides [62,63]. *In vivo*, this property could be all the more important as native cellular membranes intrinsically contain low amounts of negatively charged lipids.

5 Concluding remarks

We have shown that a global fold-unfold transition is involved in the membrane binding of EnHD and we have characterized EnHD membrane-bound state by NMR and coarse-grained MD simulations. The homeodomain may thus be viewed as an unfoldon, i.e. a folded structure that undergoes functional unfolding triggered by anionic lipid binding [64]. Unfolding transitions induced by membrane binding are not uncommon and have been characterized in a variety of proteins, including cytochrome c, translocating toxins, viral proteins, lipid-binding proteins and apolipoproteins [65]. This fold-unfold transition endows the homeodomain with a conformational plasticity enabling it to bind to partners as diverse as DNA and anionic lipid membranes.

CRedit author statement

L.C.: Conceptualization, Investigation, Writing-Review & Editing. D.S.: Investigation. L.K.: Investigation. A.J.: Writing-Review & Editing. P.F.: Conceptualization, Investigation, Writing-Review & Editing. O.L.: Conceptualization, Investigation, Writing-Original Draft, Review & Editing.

Declaration of competing interest

The authors declare that they have no competing financial interests or personal relationships which may be considered as potential competing interests.

Acknowledgments

The financial contribution of the Agence Nationale de la Recherche (grant N° ANR-BLANC 2017-CROSS) is gratefully acknowledged. D.S. was supported by a Ph.D. grant from the French Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI).

Appendix A Supplementary data

Molecular dynamics simulations parameters and analysis

Table S1 NMR and structural statistics for the ensemble of 10 EnHD models

Fig. S1 Overlay of 2D ^1H - ^{15}N HSQC spectra of wild-type EnHD and T6C mutant grafted with the probe in the diamagnetic state, in the presence of anionic bicelles

Fig. S2 Comparison of EnHD $^{13}\text{C}\alpha$ CSD in water and in DMPG bicelles

Fig. S3 TALOS prediction of ϕ , ψ angles of EnHD in water and in DMPG bicelles

Fig. S4 Combined ^1HN , ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}'$ chemical shift variation of EnHD residues in water and DMPG bicelles

Fig. S5 NOE diagram of EnHD in DMPG bicelles

Fig. S6 ^{15}N R_1 and R_2 relaxation rates of EnHD in DMPG bicelles

Fig. S7 Snapshots explaining the general MD strategy used in this paper

Fig. S8 Minimum distance between the protein and the phosphate beads in simulations 1.1, 1.2 and 1.3

Fig. S9 Mean partitioning of protein backbone beads in simulations 1.1, 1.2 and 1.3

Fig. S10 Evolution of EnHD structure compactness and its enrichment in lipids in simulation 2.2

References

- [1] S. Sagan, F. Burlina, I.D. Alves, C. Bechara, E. Dupont, A. Joliot, Homeoproteins and homeoprotein-derived peptides: going in and out, *Curr. Pharm. Des.*, 19 (2012) 2851–2862.
- [2] C. Rabouille, Pathways of Unconventional Protein Secretion, *Trends Cell Biol.*, 27 (2017) 230–240.
- [3] J. Spatazza, E. Di Lullo, A. Joliot, E. Dupont, K.L. Moya, A. Prochiantz, Homeoprotein signaling in development, health, and disease: a shaking of dogmas offers challenges and promises from bench to bed, *Pharmacol. Rev.*, 65 (2013) 90–104.
- [4] M. Green, P.M. Loewenstein, Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein, *Cell*, 55 (1988) 1179–1188.
- [5] J.P. Steringer, S. Lange, S. Čujová, R. Šachl, C. Poojari, F. Lolicato, O. Beutel, H.-M. Müller, S. Unger, Ü. Coskun, A. Honigmann, I. Vattulainen, M. Hof, C. Freund, W. Nickel, Key steps in unconventional secretion of fibroblast growth factor 2 reconstituted with purified components, *ELife*, 6 (2017).
- [6] S. Banerjee-Basu, A.D. Baxevanis, Molecular evolution of the homeodomain family of transcription factors, *Nucleic Acids Res.*, 29 (2001) 3258–3269.
- [7] M. Tassetto, A. Maizel, J. Osorio, A. Joliot, Plant and animal homeodomains use convergent mechanisms for intercellular transfer, *EMBO Rep.*, 6 (2005) 885–890.
- [8] E.J. Lee, N. Kim, J.W. Park, K.H. Kang, W.-I. Kim, N.S. Sim, C.-S. Jeong, S. Blackshaw, M. Vidal, S.-O. Huh, D. Kim, J.H. Lee, J.W. Kim, Global Analysis of Intercellular Homeodomain Protein Transfer, *Cell Rep.*, 28 (2019) 712-722.e3.
- [9] T.L. Religa, Comparison of multiple crystal structures with NMR data for engrailed homeodomain, *J. Biomol. NMR*, 40 (2008) 189–202.
- [10] E.J. Stollar, U. Mayor, S.C. Lovell, L. Federici, S.M.V. Freund, A.R. Fersht, B.F. Luisi, Crystal structures of engrailed homeodomain mutants: implications for stability and dynamics, *J. Biol. Chem.*, 278 (2003) 43699–43708.
- [11] L. Carlier, S. Balayssac, F.-X. Cantrelle, L. Khemtémourian, G. Chassaing, A. Joliot, O. Lequin, Investigation of homeodomain membrane translocation properties: insights from the structure determination of engrailed-2 homeodomain in aqueous and membrane-mimetic environments, *Biophys. J.*, 105 (2013) 667–678.
- [12] H.H. Simon, H. Saueressig, W. Wurst, M.D. Goulding, D.D. O’Leary, Fate of midbrain dopaminergic neurons controlled by the engrailed genes, *J. Neurosci.*, 21 (2001) 3126–3134.
- [13] L. Sonnier, G. Le Pen, A. Hartmann, J.-C. Bizot, F. Trovero, M.-O. Krebs, A. Prochiantz, Progressive Loss of Dopaminergic Neurons in the Ventral Midbrain of Adult Mice Heterozygote for Engrailed1, *J. Neurosci.*, 27 (2007) 1063–1071.
- [14] H. Rekaik, F.-X. Blaudin de Thé, A. Prochiantz, J. Fuchs, R.L. Joshi, Dissecting the role of Engrailed in adult dopaminergic neurons - Insights into Parkinson disease pathogenesis, *FEBS Lett.*, 589 (2015) 3786–3794.
- [15] F. Blaudin de Thé, H. Rekaik, E. Peze-Heidsieck, O. Massiani-Beaudoin, R.L. Joshi, J. Fuchs, A. Prochiantz, Engrailed homeoprotein blocks degeneration in adult dopaminergic neurons through LINE-1 repression, *EMBO J.*, 37 (2018).
- [16] C. Rampon, C. Gauron, T. Lin, F. Meda, E. Dupont, A. Cosson, E. Ipendey, A. Frerot, I. Aujard, T. Le Saux, D. Bensimon, L. Jullien, M. Volovitch, S. Vriza, A. Joliot, Control of brain patterning by Engrailed paracrine transfer: a new function of the Pbx interaction domain, *Development*, 142 (2015) 1840–1849.
- [17] I. Brunet, A.A. Di Nardo, L. Sonnier, M. Beurdeley, A. Prochiantz, The topological role

of homeoproteins in the developing central nervous system, *Trends Neurosci.*, 30 (2007) 260–267.

[18] A. Wizenmann, I. Brunet, J.S.Y. Lam, L. Sonnier, M. Beurdeley, K. Zarbalis, D. Weisenhorn-Vogt, C. Weinl, A. Dwivedy, A. Joliot, W. Wurst, C. Holt, A. Prochiantz, Extracellular Engrailed participates in the topographic guidance of retinal axons in vivo, *Neuron*, 64 (2009) 355–366.

[19] A. Wizenmann, O. Stettler, K.L. Moya, Engrailed homeoproteins in visual system development, *Cell. Mol. Life Sci.*, 72 (2015) 1433–1445.

[20] D. Derossi, A.H. Joliot, G. Chassaing, A. Prochiantz, The third helix of the Antennapedia homeodomain translocates through biological membranes, *J. Biol. Chem.*, 269 (1994) 10444–10450.

[21] W.B. Kauffman, T. Fuselier, J. He, W.C. Wimley, Mechanism Matters: A Taxonomy of Cell Penetrating Peptides, *Trends Biochem. Sci.*, 40 (2015) 749–764.

[22] S. Balayssac, F. Burlina, O. Convert, G. Bolbach, G. Chassaing, O. Lequin, Comparison of penetratin and other homeodomain-derived cell-penetrating peptides: interaction in a membrane-mimicking environment and cellular uptake efficiency, *Biochemistry*, 45 (2006) 1408–1420.

[23] I. Le Roux, A.H. Joliot, E. Bloch-Gallego, A. Prochiantz, M. Volovitch, Neurotrophic activity of the Antennapedia homeodomain depends on its specific DNA-binding properties, *Proc. Natl. Acad. Sci.*, 90 (1993) 9120–9124.

[24] C. Chipot, F. Dehez, J.R. Schnell, N. Zitzmann, E. Pebay-Peyroula, L.J. Catoire, B. Miroux, E.R.S. Kunji, G. Veglia, T.A. Cross, P. Schanda, Perturbations of Native Membrane Protein Structure in Alkyl Phosphocholine Detergents: A Critical Assessment of NMR and Biophysical Studies, *Chem. Rev.*, 118 (2018) 3559–3607.

[25] I. Marcotte, M. Auger, Bicelles as model membranes for solid- and solution-state NMR studies of membrane peptides and proteins, *Concepts Magn. Reson. Part A*, 24A (2005) 17–37.

[26] U.H.N. Dürr, R. Soong, A. Ramamoorthy, When detergent meets bilayer: Birth and coming of age of lipid bicelles, *Prog. Nucl. Magn. Reson. Spectrosc.*, 69 (2013) 1–22.

[27] R. Augustyniak, S. Balayssac, F. Ferrage, G. Bodenhausen, O. Lequin, ^1H , ^{13}C and ^{15}N resonance assignment of a 114-residue fragment of Engrailed 2 homeoprotein, a partially disordered protein, *Biomol. NMR Assign.*, 5 (2011) 229–231.

[28] G. Raffard, S. Steinbruckner, A. Arnold, J.H. Davis, E.J. Dufourc, Temperature–Composition Diagram of Dimyristoylphosphatidylcholine–Dicaproylphosphatidylcholine “Bicelles” Self-Orienting in the Magnetic Field A Solid State ^2H and ^{31}P NMR Study, *Langmuir*, 16 (2000) 7655–7662.

[29] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR*, 6 (1995) 277–293.

[30] W. Lee, M. Tonelli, J.L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy, *Bioinforma. Oxf. Engl.*, 31 (2015) 1325–1327.

[31] Y. Shen, A. Bax, Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks, *J. Biomol. NMR*, 56 (2013) 227–241.

[32] C. Xu, E. Gagnon, M.E. Call, J.R. Schnell, C.D. Schwieters, C.V. Carman, J.J. Chou, K.W. Wucherpennig, Regulation of T Cell Receptor Activation by Dynamic Membrane Binding of the CD3 ϵ Cytoplasmic Tyrosine-Based Motif, *Cell*, 135 (2008) 702–713.

[33] G.A. Bermejo, C.D. Schwieters, Protein Structure Elucidation from NMR Data with the

Program Xplor-NIH, in: R. Ghose (Ed.), *Methods Mol Biol*, Springer New York, New York, NY, 2018: pp. 311–340.

[34] Y. Tian, C.D. Schwieters, S.J. Opella, F.M. Marassi, A Practical Implicit Membrane Potential for NMR Structure Calculations of Membrane Proteins, *Biophys. J.*, 109 (2015) 574–585.

[35] Y. Tian, C.D. Schwieters, S.J. Opella, F.M. Marassi, High quality NMR structures: a new force field with implicit water and membrane solvation for Xplor-NIH, *J. Biomol. NMR*, 67 (2017) 35–49.

[36] S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations, *J. Phys. Chem. B*, 111 (2007) 7812–7824.

[37] L. Monticelli, S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, S.-J. Marrink, The MARTINI Coarse-Grained Force Field: Extension to Proteins, *J. Chem. Theory Comput.*, 4 (2008) 819–834.

[38] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation, *J. Chem. Theory Comput.*, 4 (2008) 435–447.

[39] X. Periole, M. Cavalli, S.-J. Marrink, M.A. Ceruso, Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition, *J. Chem. Theory Comput.*, 5 (2009) 2531–2543.

[40] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.*, 126 (2007) 014101.

[41] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.*, 52 (1981) 7182–7190.

[42] V. Nanda, L. Brand, Aromatic interactions in homeodomains contribute to the low quantum yield of a conserved, buried tryptophan, *Proteins*, 40 (2000) 112–125.

[43] B. Christiaens, S. Symoens, S. Verheyden, Y. Engelborghs, A. Joliot, A. Prochiantz, J. Vandekerckhove, M. Rosseneu, B. Vanloo, S. Vanderheyden, Tryptophan fluorescence study of the interaction of penetratin peptides with model membranes, *Eur. J. Biochem.*, 269 (2002) 2918–2926.

[44] J.R. Lakowicz, ed., *Protein Fluorescence*, in: *Princ. Fluoresc. Spectrosc.*, Springer US, Boston, MA, 2006: pp. 529–575.

[45] J.A. Poveda, M. Prieto, J.A. Encinar, J.M. González-Ros, C.R. Mateo, Intrinsic Tyrosine Fluorescence as a Tool To Study the Interaction of the Shaker B “Ball” Peptide with Anionic Membranes, *Biochemistry*, 42 (2003) 7124–7132.

[46] F.H. Schumann, H. Riepl, T. Maurer, W. Gronwald, K.-P. Neidig, H.R. Kalbitzer, Combined chemical shift changes and amino acid specific chemical shift mapping of protein–protein interactions, *J. Biomol. NMR*, 39 (2007) 275–289.

[47] J.J. Chou, J.D. Kaufman, S.J. Stahl, P.T. Wingfield, A. Bax, Micelle-Induced Curvature in a Water-Insoluble HIV-1 Env Peptide Revealed by NMR Dipolar Coupling Measurement in Stretched Polyacrylamide Gel, *J. Am. Chem. Soc.*, 124 (2002) 2450–2451.

[48] D. Bucher, Y.-H. Hsu, V.D. Mouchlis, E.A. Dennis, J.A. McCammon, Insertion of the Ca²⁺-Independent Phospholipase A2 into a Phospholipid Bilayer via Coarse-Grained and Atomistic Molecular Dynamics Simulations, *PLoS Comput. Biol.*, 9 (2013) e1003156.

[49] L. Charlier, M. Louet, L. Chaloin, P. Fuchs, J. Martinez, D. Muriaux, C. Favard, N. Floquet, Coarse-Grained Simulations of the HIV-1 Matrix Protein Anchoring: Revisiting Its Assembly on Membrane Domains, *Biophys. J.*, 106 (2014) 577–585.

- [50] J.V. Vermaas, E. Tajkhorshid, Conformational heterogeneity of α -synuclein in membrane, *Biochim. Biophys. Acta BBA - Biomembr.*, 1838 (2014) 3107–3117.
- [51] A. Walrant, L. Matheron, S. Cribier, S. Chaignepain, M.-L. Jobin, S. Sagan, I.D. Alves, Direct translocation of cell-penetrating peptides in liposomes: a combined mass spectrometry quantification and fluorescence detection study, *Anal. Biochem.*, 438 (2013) 1–10.
- [52] M. Beaugrand, A.A. Arnold, J. Hénin, D.E. Warschawski, P.T.F. Williamson, I. Marcotte, Lipid Concentration and Molar Ratio Boundaries for the Use of Isotropic Bicelles, *Langmuir*, 30 (2014) 6162–6170.
- [53] U. Mayor, N.R. Guydosh, C.M. Johnson, J.G. Grossmann, S. Sato, G.S. Jas, S.M.V. Freund, D.O.V. Alonso, V. Daggett, A.R. Fersht, The complete folding pathway of a protein from nanoseconds to microseconds, *Nature*, 421 (2003) 863–867.
- [54] T.L. Religa, J.S. Markson, U. Mayor, S.M.V. Freund, A.R. Fersht, Solution structure of a protein denatured state and folding intermediate, *Nature*, 437 (2005) 1053–1056.
- [55] V. Corradi, B.I. Sejdiu, H. Mesa-Gallosa, H. Abdizadeh, S.Yu. Noskov, S.J. Marrink, D.P. Tieleman, Emerging Diversity in Lipid–Protein Interactions, *Chem. Rev.*, 119 (2019) 5775–5848.
- [56] C.R. Kissinger, B. Liu, E. Martin-Blanco, T.B. Kornberg, C.O. Pabo, Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions, *Cell*, 63 (1990) 579–590.
- [57] A. Walrant, A. Bauzá, C. Girardet, I.D. Alves, S. Lecomte, F. Illien, S. Cardon, N. Chaianantakul, M. Pallerla, F. Burlina, A. Frontera, S. Sagan, Ionpair- π interactions favor cell penetration of arginine/tryptophan-rich cell-penetrating peptides, *Biochim. Biophys. Acta BBA - Biomembr.*, 1862 (2020) 183098.
- [58] P. Gehan, S. Kulifaj, P. Soule, J.B. Bodin, M. Amoura, A. Walrant, S. Sagan, A.R. Thiam, K. Ngo, V. Vivier, S. Cribier, N. Rodriguez, Penetratin translocation mechanism through asymmetric droplet interface bilayers, *Biochim. Biophys. Acta BBA - Biomembr.*, 1862 (2020) 183415.
- [59] I. Amblard, E. Dupont, I. Alves, J. Miralvès, I. Queguiner, A. Joliot, Bidirectional transfer of homeoprotein EN2 across the plasma membrane requires PIP₂, *J. Cell Sci.*, 133 (2020) jcs244327.
- [60] L. Bechtella, E. Chalouhi, P. Milán Rodríguez, M. Cosset, D. Ravault, F. Illien, S. Sagan, L. Carlier, O. Lequin, P.F.J. Fuchs, E. Sachon, A. Walrant, Structural Bases for the Involvement of Phosphatidylinositol-4,5-bisphosphate in the Internalization of the Cell-Penetrating Peptide Penetratin, *ACS Chem. Biol.*, 17 (2022) 1427–1439.
- [61] V. Del Gaizo Moore, R.M. Payne, Transactivator of Transcription Fusion Protein Transduction Causes Membrane Inversion, *J. Biol. Chem.*, 279 (2004) 32541–32544.
- [62] W. Zhang, S.O. Smith, Mechanism of penetration of Antp(43–58) into membrane bilayers, *Biochemistry*, 44 (2005) 10110–10118.
- [63] P. Wadhvani, R.F. Epand, N. Heidenreich, J. Bürck, A.S. Ulrich, R.M. Epand, Membrane-Active Peptides and the Clustering of Anionic Lipids, *Biophys. J.*, 103 (2012) 265–274.
- [64] V.N. Uversky, Functional roles of transiently and intrinsically disordered regions within proteins, *FEBS J.*, 282 (2015) 1182–1189.
- [65] U. Jakob, R. Kriwacki, V.N. Uversky, Conditionally and Transiently Disordered Proteins: Awakening Cryptic Disorder To Regulate Protein Function, *Chem. Rev.*, 114 (2014) 6779–6805.

[66] X. Daura, K. Gademann, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, Peptide Folding: When Simulation Meets Experiment, *Angew. Chem. Int. Ed.*, 38 (1999) 236–240.

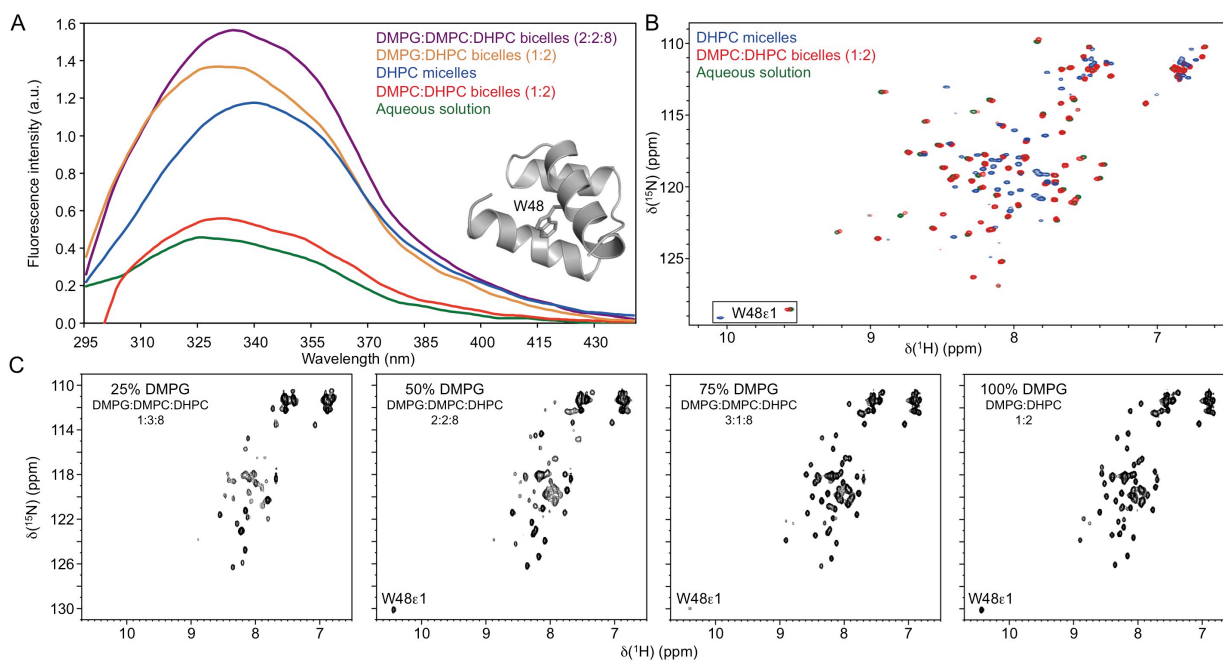
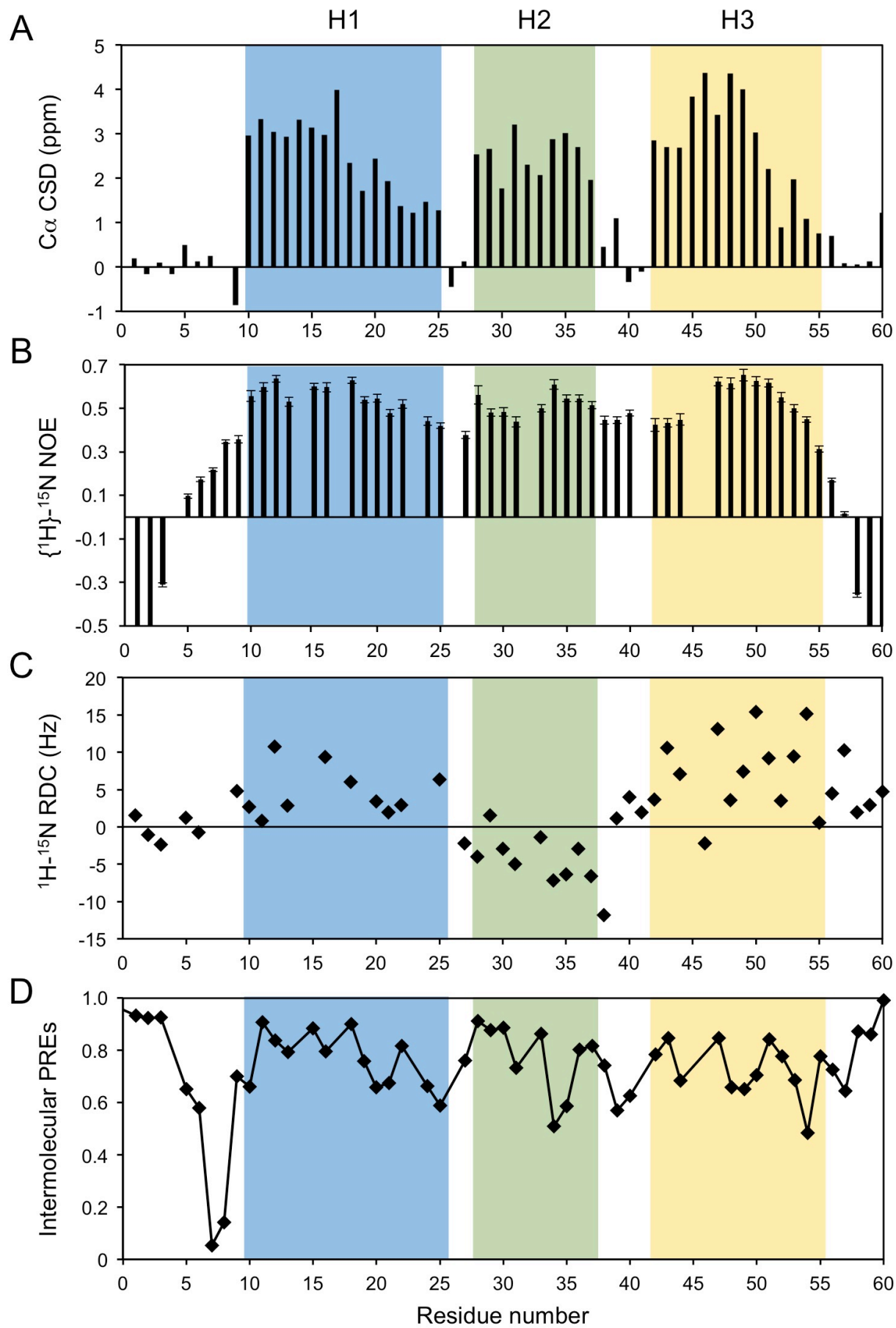


Fig. 1 Influence of membrane environments on EnHD conformation monitored by Trp fluorescence and 2D ^1H , ^{15}N NMR spectroscopy. (A) Trp fluorescence of 25 μM EnHD in aqueous solution, in DHPC micelles, in zwitterionic and in anionic bicelles ($q = 0.5$). (B) Superimposed 2D ^1H - ^{15}N HSQC spectra of 0.5 mM EnHD in aqueous solution, in 50 mM DHPC micelles, and in zwitterionic bicelles made of 50 mM DHPC and 25 mM DMPC ($q = 0.5$). (C) 2D ^1H - ^{15}N HSQC spectra of 0.5 mM EnHD in bicelles (50 mM DHPC, 25 mM DMPC/DMPG, $q = 0.5$) containing increasing amounts of anionic DMPG lipids (from left to right: 25, 50, 75 and 100 %).



(figure previous page)

Fig. 2 NMR parameters of EnHD in complex with anionic bicelles (0.5 mM EnHD, 50 mM DHPC, 25 mM DMPG, $q = 0.5$). (A) Chemical shift deviation (CSD) of $^{13}\text{C}\alpha$ resonances from random coil values. (B) Heteronuclear $\{^1\text{H}\}^{15}\text{N}$ NOEs measured at 500 MHz and 37 °C. (C) One-bond residual dipolar couplings (RDCs) between amide $^1\text{H}^{\text{N}}$ and ^{15}N , induced by a radially compressed negatively charged polyacrylamide gel. (D) Position of EnHD with respect to anionic bicelle surface inferred from intermolecular PREs. PREs were induced by the addition of 5-doxyl PC paramagnetic lipid in DHPC/DMPG bicelles.

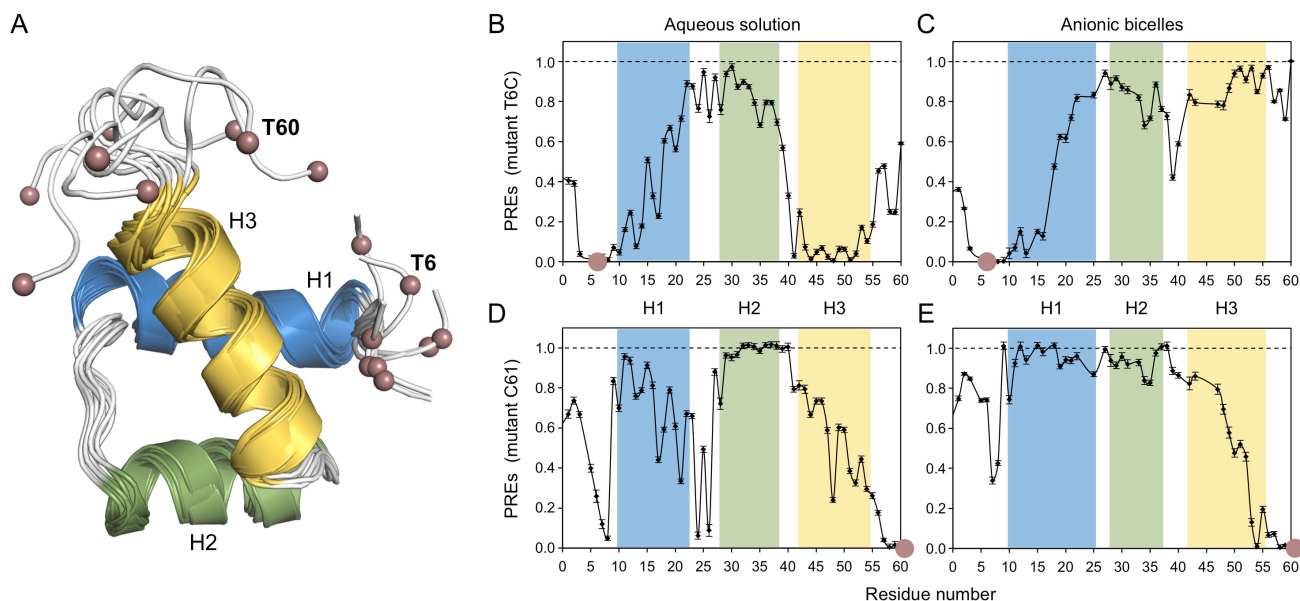


Fig. 3 Paramagnetic relaxation enhancement (PRE) induced by intramolecular spin labels. (A) NMR structure of EnHD in aqueous solution (PDB 3ZOB) showing the position of residues selected for spin label covalent attachment (T6C substitution mutation or C61 insertion mutation). 8 conformers are displayed to show the conformational space of the spin label probe. (B, C) PREs induced by the nitroxide radical attached to C6 position in aqueous solution (B) or in anionic bicelles (C). (D, E) PREs induced by the nitroxide radical attached to C61 position in aqueous solution (D) or in anionic bicelles (E). PREs were measured as the ratio of crosspeak intensities in the paramagnetic and diamagnetic conditions. A null residual intensity corresponds therefore to a strong PRE.

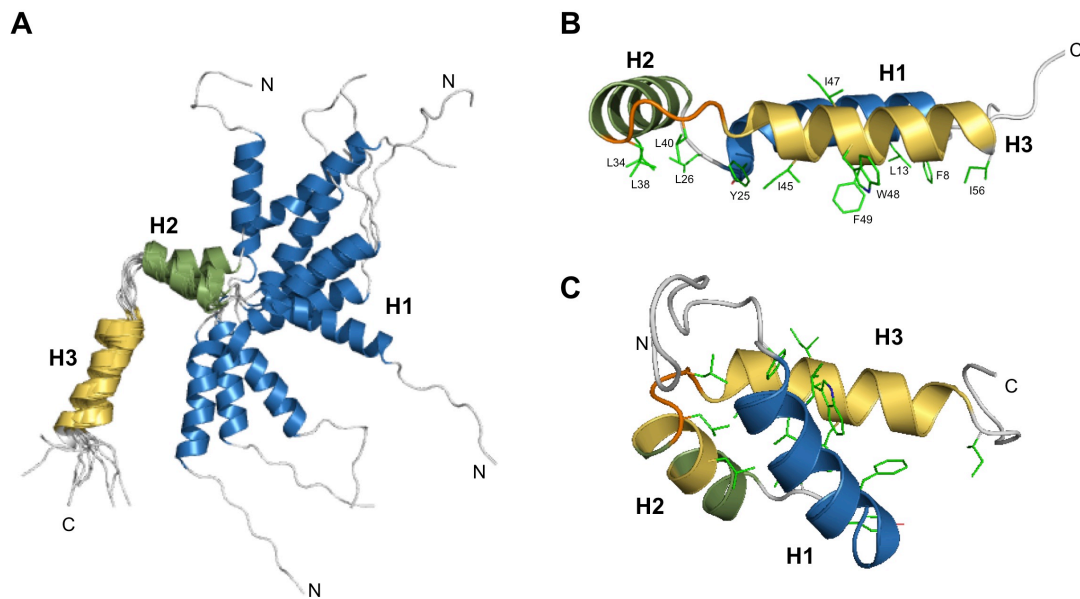


Fig. 4 Models of EnHD bound to anionic bicelles. (A) Conformational ensemble of 10 NMR structures superimposed over residues 28–55 (H2-H3 segment). The structures are viewed from the top of the membrane surface. (B) Perpendicular view of one selected model showing the orientation of hydrophobic side chains. The turn segment connecting H2 and H3 helices is shown in orange. (C) Corresponding view of EnHD native state (PDB 3ZOB) with the same orientation of H3 helix.

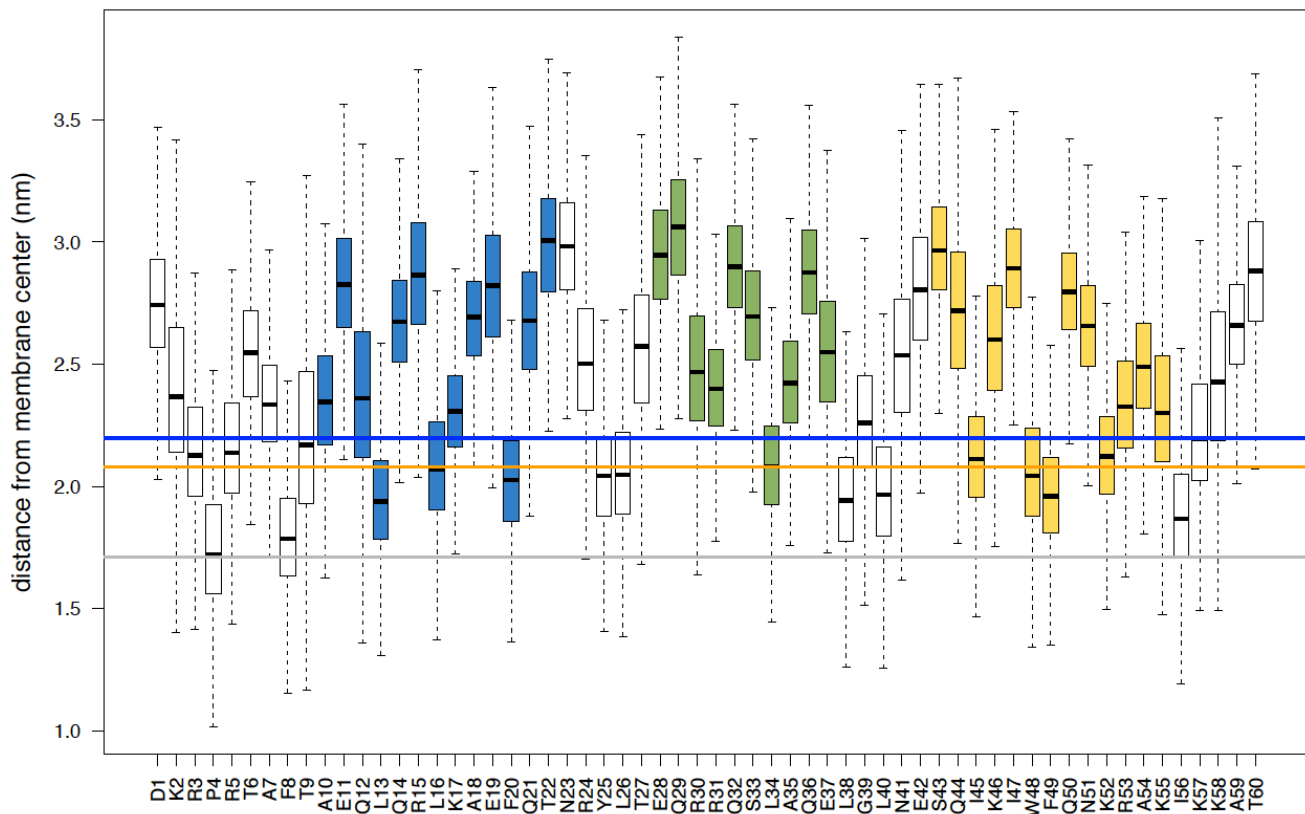


Fig. 5 Mean partitioning of EnHD side chain beads in simulations 2.1 and 2.2. The vertical axis corresponds to the partitioning in the z dimension of each bead with respect to the center of the bilayer (which is towards the bottom at a z of 0 nm). Each box and whiskers correspond to statistics over the 2 simulations 2.1 and 2.2, calculated over 15-60 μ s (after full insertion of the protein within the bilayer). The blue horizontal line represents the mean choline level of DPPC, the orange line the mean phosphate level (of both DPPC and DPPG), and the gray line the mean glycerol level (GL1 bead of both DPPC and DPPG) of the lipid leaflet to which the protein binds.

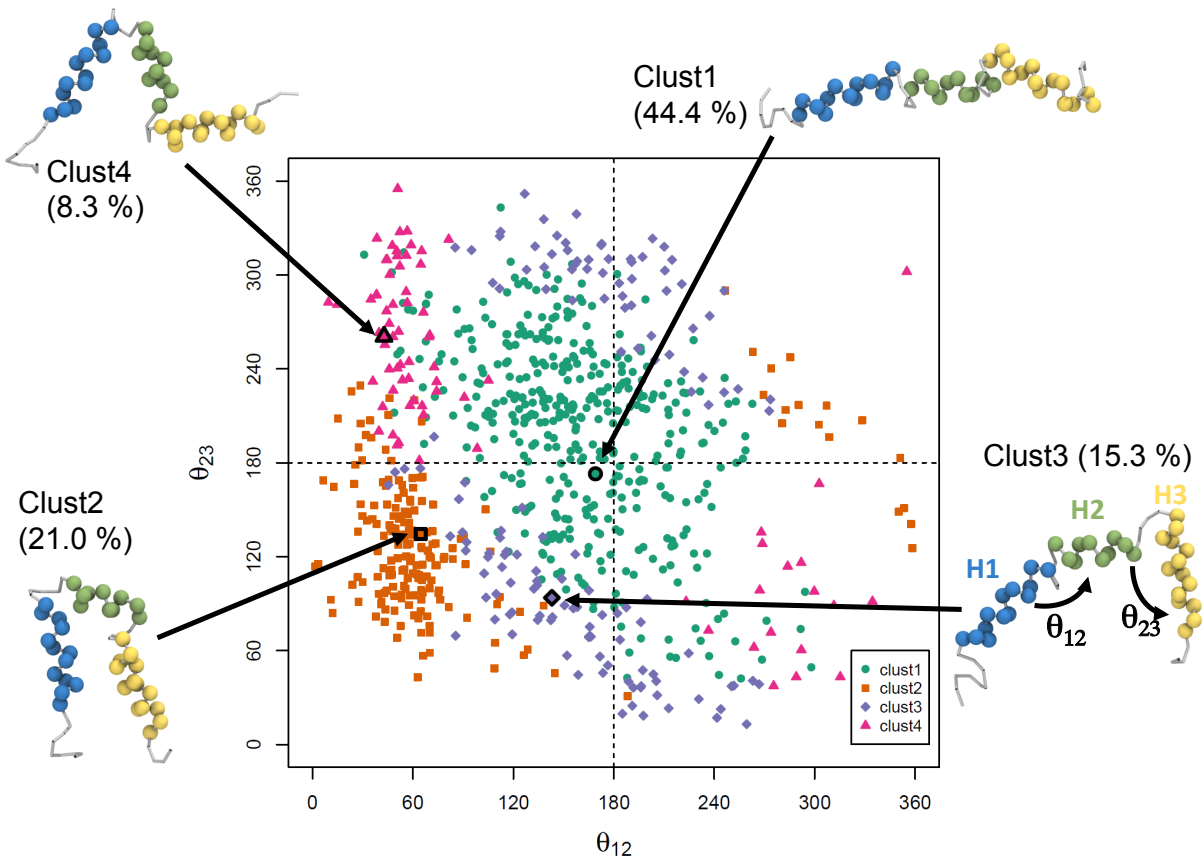


Fig. 6 Clustering analysis of the conformational space spanned by EnHD once inserted within the membrane. The clustering was performed with the GROMOS method [66] from the program *g_cluster* of GROMACS [38]. Only the 4 most populated clusters are shown, amounting to 89 % of all conformations. Each cluster center is represented on the side with the 3 helices colored blue, green and yellow for H1, H2 and H3, respectively. Since the helices are rigid bodies inserting in the plane of the membrane, the conformation of the protein can be conveniently represented by two angular parameters, θ_{12} and θ_{13} , which are the angles between H1 and H2, and H2 and H3 respectively. Each conformation of the trajectory is represented as a dot with a specific symbol and colour for each cluster. The cluster center is indicated with an arrow and a symbol circled with black.

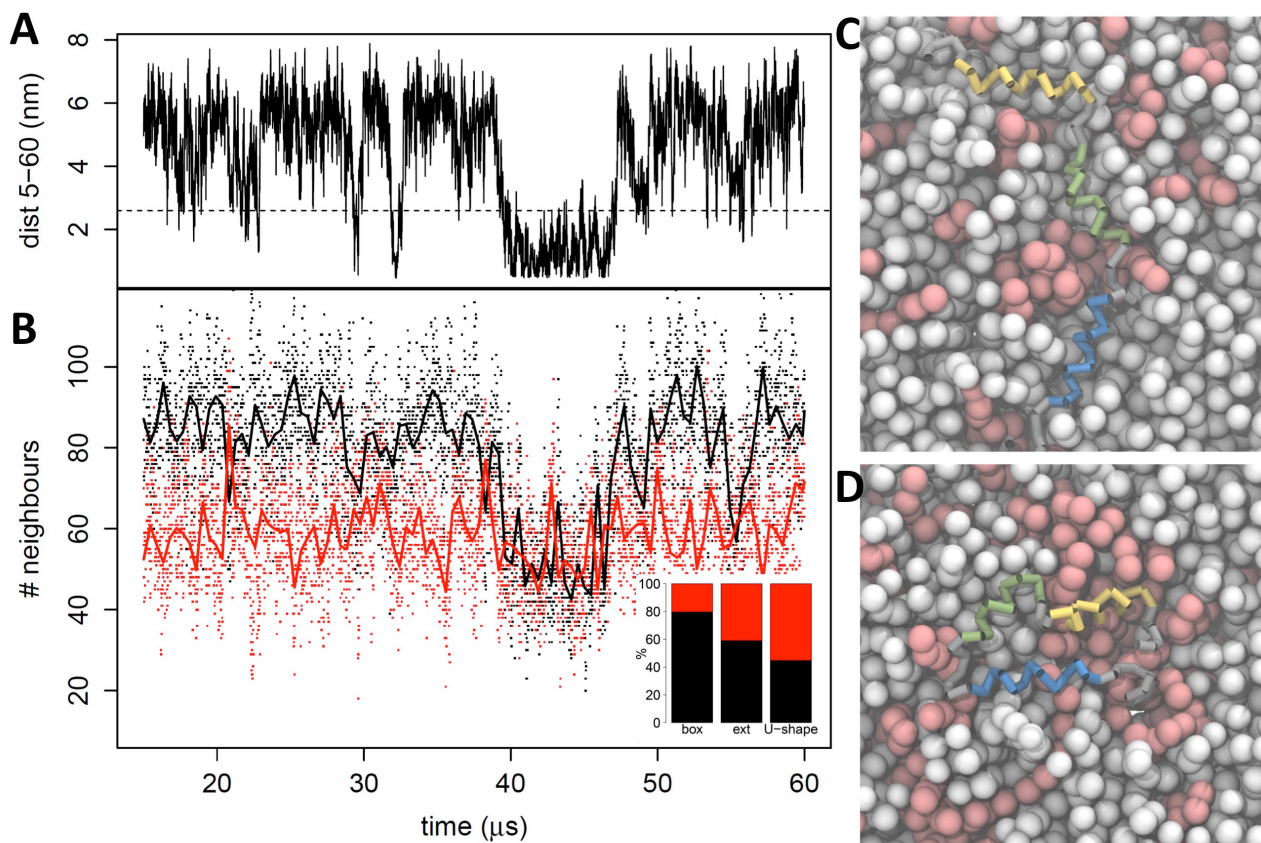


Fig. 7 Evolution of EnHD structure compactness and its enrichment in lipids in simulation 2.1. (A) Distance between the side chains of residues 5 and 60. A low value is indicative of a U-shaped conformation (cluster 2), while a high value corresponds to an extended conformation (cluster 1). (B) Number of neighboring lipids having a contact with the protein (black DPPC, red DPPG). Each dot represents a number of neighbors for a given conformation, the solid lines is a smoothing of all dots. The inset in the bottom right represents the molar ratio between DPPC and DPPG for the whole membrane, for the lipids close to the protein in an extended (ext) or U-shaped conformation. (C, D) Snapshots (top view) of EnHD in an extended (C) and U-shaped (D) conformation. DPPC lipids are represented with white spheres and DPPG with pale red spheres. The protein is represented with sticks between backbone beads. H1, H2 and H3 are represented in blue, green and yellow respectively, loops are in gray.

Supplementary Information

Anionic lipids induce a fold-unfold transition in the membrane-translocating Engrailed homeodomain

Ludovic CARLIER, Damien SAMSON, Lucie KHEMTEMOURIAN, Alain JOLIOT, Patrick F. J. FUCHS, Olivier LEQUIN

Molecular dynamics simulations parameters and analysis

Table S1. NMR and structural statistics for the ensemble of 10 EnHD models

Fig. S1. Overlay of 2D ^1H - ^{15}N HSQC spectra of wild-type EnHD and T6C mutant grafted with the probe in the diamagnetic state, in the presence of anionic bicelles

Fig. S2. Comparison of EnHD $^{13}\text{C}\alpha$ CSD in water and in DMPG bicelles

Fig. S3. TALOS prediction of ϕ , ψ angles of EnHD in water and in DMPG bicelles

Fig. S4. Combined ^1HN , ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}'$ chemical shift variation of EnHD residues in water and DMPG bicelles.

Fig. S5. NOE diagram of EnHD in DMPG bicelles

Fig. S6. ^{15}N R_1 and R_2 relaxation rates of EnHD in DMPG bicelles

Fig. S7. Snapshots explaining the general MD strategy used in this paper

Fig. S8. Minimum distance between the protein and the phosphate beads in simulations 1.1, 1.2 and 1.3

Fig. S9. Mean partitioning of protein backbone beads in simulations 1.1, 1.2 and 1.3

Fig. S10. Evolution of EnHD structure compactness and its enrichment in lipids in simulation 2.2

Molecular dynamics simulations parameters and analysis.

Simulation parameters. A 1.2-nm cutoff was used for non-bonded interactions with a shift on energies and forces from 0 to 1.2 nm and from 0.9 to 1.2 nm for electrostatic and Lennard-Jones interactions, respectively. The neighbor list was updated every 10 steps and the relative dielectric constant for the medium was set to 15. The temperature was kept at 300 K using the velocity rescaling thermostat [37] with a time constant of 1 ps. Pressure was kept at 1 bar in a semi-isotropic way (i.e. independently in the plane of the membrane and perpendicular to the membrane) using the Parrinello-Rahman barostat [38] with a time constant of 4 ps. The integration time step was 20 fs and structures were saved every ns for analysis. For simulations 1.1, 1.2, and 1.3, an elastic network (EN) was used to maintain the tertiary fold of the protein [36]. A force constant of $500 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$ was used between residues which were less than 0.8 nm apart. For simulations 2.1 and 2.2, no elastic network was used. The three helices of the protein were enforced using dihedral angle potentials between consecutive backbone beads [34] since no elastic network was used for these simulations.

Trajectory analysis. The simulations were analyzed in terms of partitioning of the protein within the bilayer. This was done by extracting the z coordinate of each backbone bead (or the center of mass of the side chain), and subtracting it to the center of mass of the bilayer. This latter was determined by considering the phosphate beads of both leaflets. The results are represented as boxplots on which the outliers were removed for a better visualization. The number of lipids around the protein was analyzed using the *mindist* program from GROMACS. A neighboring lipid (DPPC or DPPG) was considered when any of its particle was less than 6 Å apart of the protein.

In simulations 2.1 and 2.2, the protein was fully inserted after 5.6 and 13.7 μs , respectively. Both trajectories after the protein insertion (time window 15–60 μs) were concatenated to perform a clustering of the protein conformations. This was done using the program *g_cluster* from GROMACS using the GROMOS algorithm [39]. Only backbone beads were considered. The cutoff was optimized in order to get a reasonable number of clusters (around 10) and avoid too many small clusters of one element. A cutoff value of 0.8 nm was thus chosen.

Since the protein eventually inserted horizontally within the bilayer (each helix stayed in the plane of the bilayer) in simulations 2.1 and 2.2, the protein conformations could be simply described with the angles between helix H1 and helix H2, as well as between helix H2 and helix H3, called θ_{12} and θ_{23} , respectively (see figure 7 for their definition). For that, each helix axis vector (**h1**, **h2** and **h3**) was computed by considering the two extreme beads of each helix and the angles between each pair of vectors were extracted using the scalar product. This procedure returns an angle in the range $[0, 180]$ degrees. Since the helices stay roughly horizontal in the plane of the bilayer, the sign for each angle could be assigned by taking that of the vector product of each pair of helices (for example, the sign θ_{12} was that of the vector product **h1** \times **h2**). In the end these angles thus span a range of $[-180, 180]$ degrees.

Table S1. NMR and structural statistics for the ensemble of 10 EnHD models

Experimental restraints	
Number of NOE-derived distance restraints	243
Intra-residue	73
Sequential	113
Medium-range ($1 < i-j \leq 4$)	57
Long-range ($ i-j > 4$)	0
Number of plane distance restraints	104
Number of dihedral angle restraints	105
ϕ	59
ψ	46
Energies and structural statistics	
EEFx2 energy (van der Waals, electrostatic, implicit solvation)	-1024 ± 14
Torsion DB energy	1998 ± 39
RMS deviations from ideal geometry	
Bond (Å)	0.006
Angle (°)	0.80 ± 0.02
Improper (°)	0.39 ± 0.04
RMS deviations on restraints	
NOE-derived distances (Å)	0.02 ± 0.01
Plane distances (Å)	0.03 ± 0.03
Dihedrals (°)	0.63 ± 0.22
Structural dispersion: backbone atoms rmsd (Å)	
All residues (1-60), residues 8-55, residues 28-55 (H2-H3)	8.94, 5.81, 1.26
Helix residues 10-25 (H1), 28-37 (H2), 42-55 (H3)	0.14, 0.20, 0.29
Ramachandran plot (% residues per region)	
Most favored	95.2 %
Additionally allowed	4.6 %
Generously allowed	0 %
Disallowed regions	0.2 %
MolProbity clash score (Z-score)	1.21

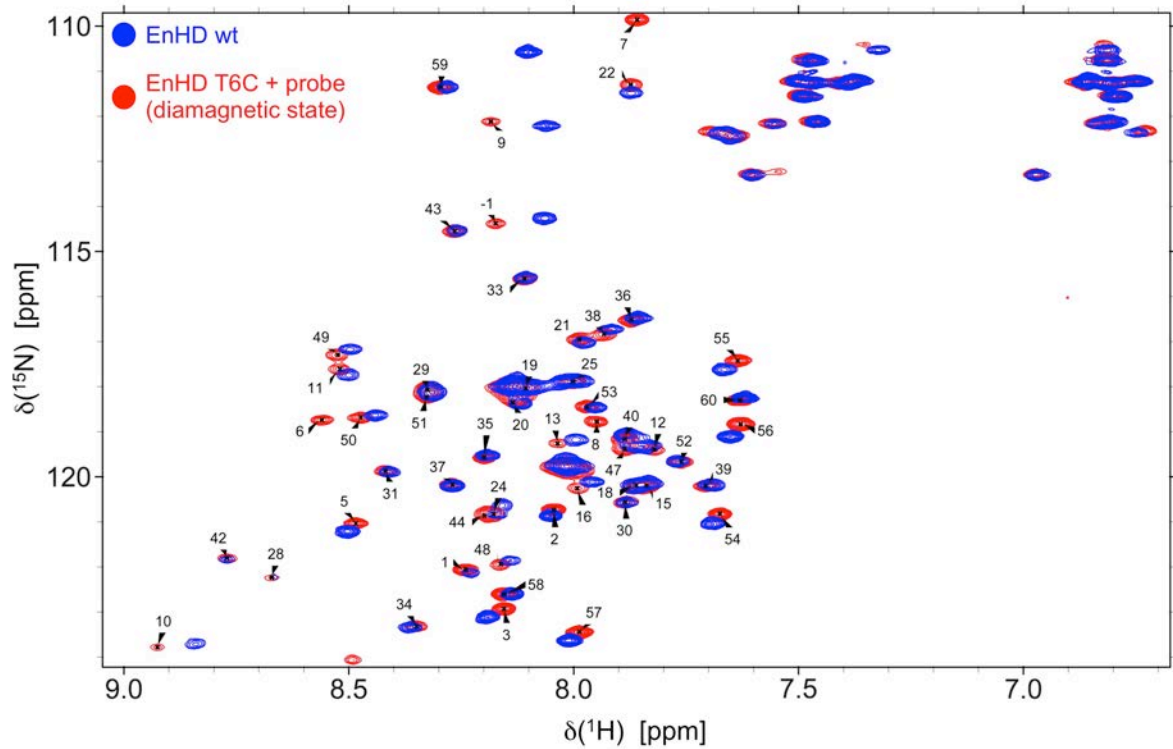


Fig. S1. Overlay of 2D ^1H - ^{15}N HSQC spectra of wild type EnHD (blue) and T6C mutant grafted with the probe in the diamagnetic state, in the presence of anionic bicelles. The assignments are indicated for the wild-type protein. Chemical shift changes are observed mainly around the site of mutation (residues 3-8) and in the C-terminal region (residues 54-58).

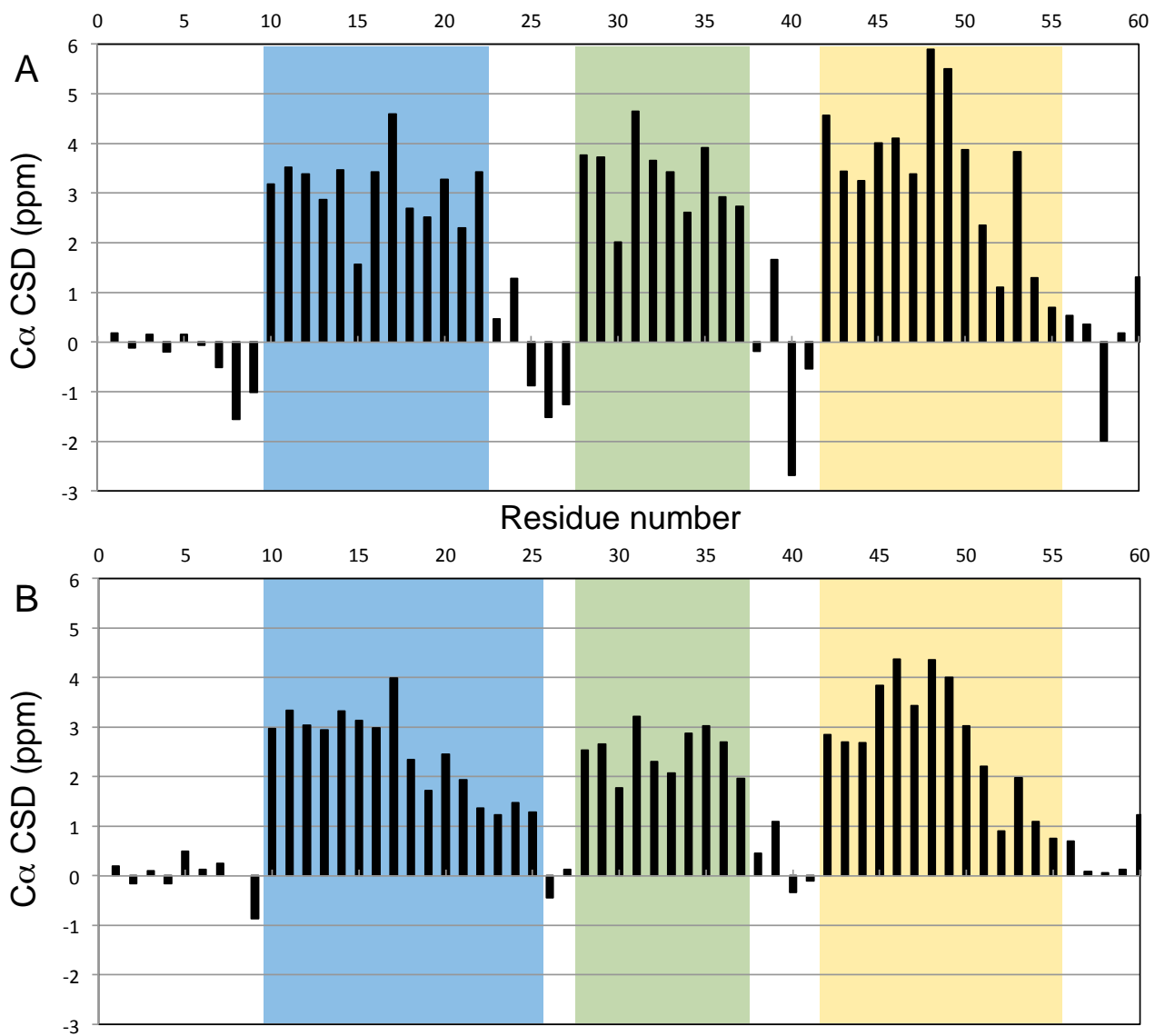


Fig. S2. Comparison of EnHD $^{13}C\alpha$ CSD in water (A) and in DMPG bicelles (B)

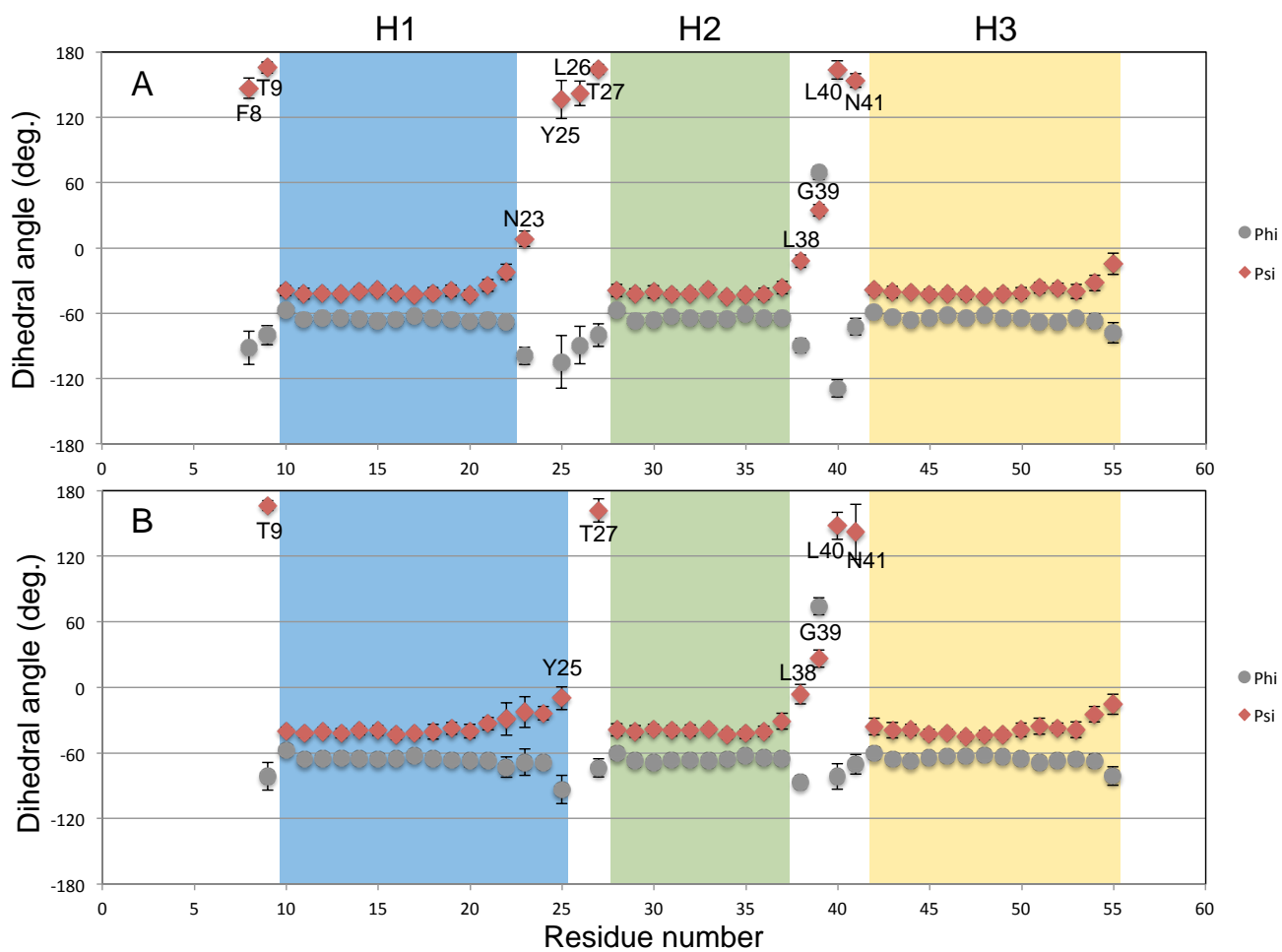


Fig. S3. TALOS prediction of ϕ , ψ angles of EnHD in water (A) and in DMPG bicelles (B)

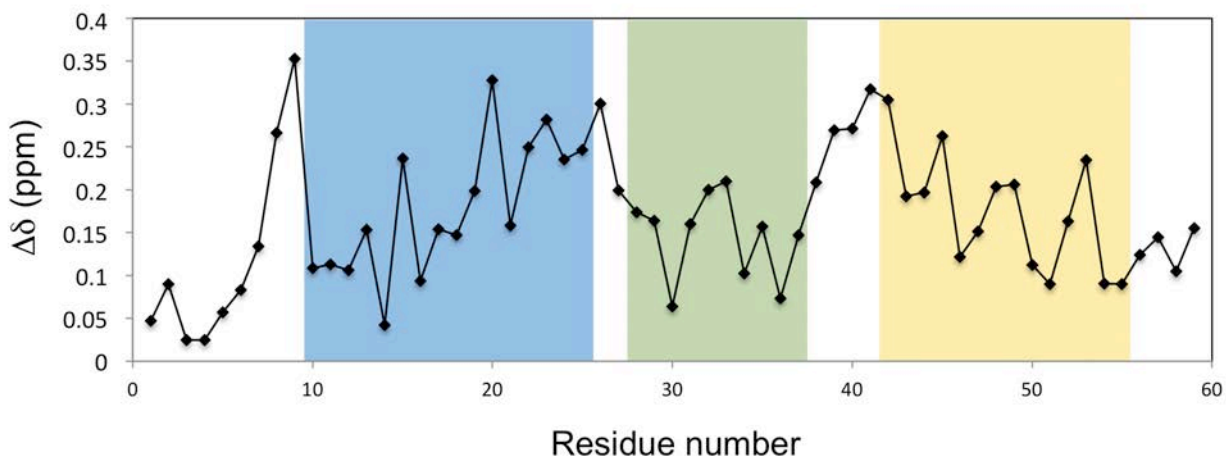


Fig. S4. Combined ^1HN , ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}'$ chemical shift variation of EnHD residues in water and DMPG bicelles. The combined chemical shift variation $\Delta\delta$ for each residue was calculated using the formula: $\Delta\delta = \frac{1}{4} ((\omega_{\text{HN}} \Delta\delta_{\text{HN}})^2 + (\omega_{\text{N}} \Delta\delta_{\text{N}})^2 + (\omega_{\text{C}\alpha} \Delta\delta_{\text{C}\alpha})^2 + (\omega_{\text{C}' } \Delta\delta_{\text{C}'})^2)^{1/2}$ where $\Delta\delta_{\text{HN}}$, $\Delta\delta_{\text{N}}$, $\Delta\delta_{\text{C}\alpha}$, $\Delta\delta_{\text{C}'}$ are the chemical shift variation between water and DMPG bicelles for ^1HN , ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}'$ nuclei, respectively, and ω_{HN} , ω_{N} , $\omega_{\text{C}\alpha}$, $\omega_{\text{C}'}$ are weighted chemical factors calculated from the spread in chemical shifts within the BMRB database [46]. The values of ω_{HN} , ω_{N} , $\omega_{\text{C}\alpha}$, $\omega_{\text{C}'}$ were 1.55, 0.236, 0.447 and 0.510, respectively [46].

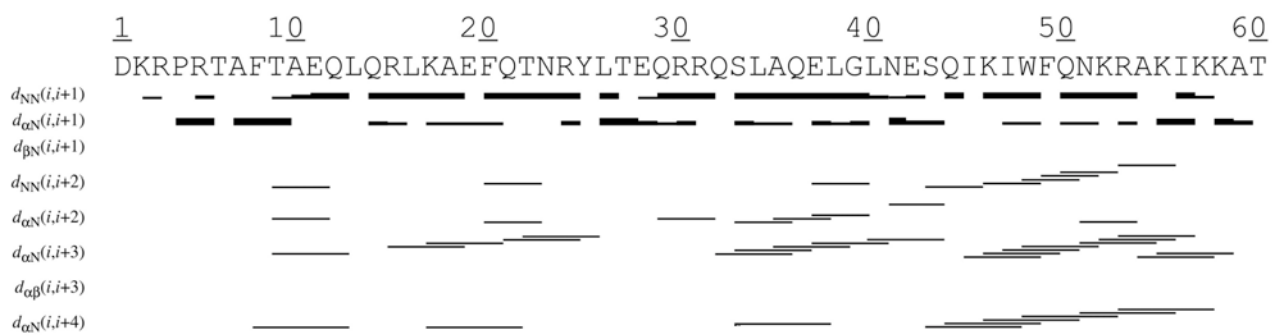


Fig. S5. NOE diagram of EnHD in DMPG bicelles

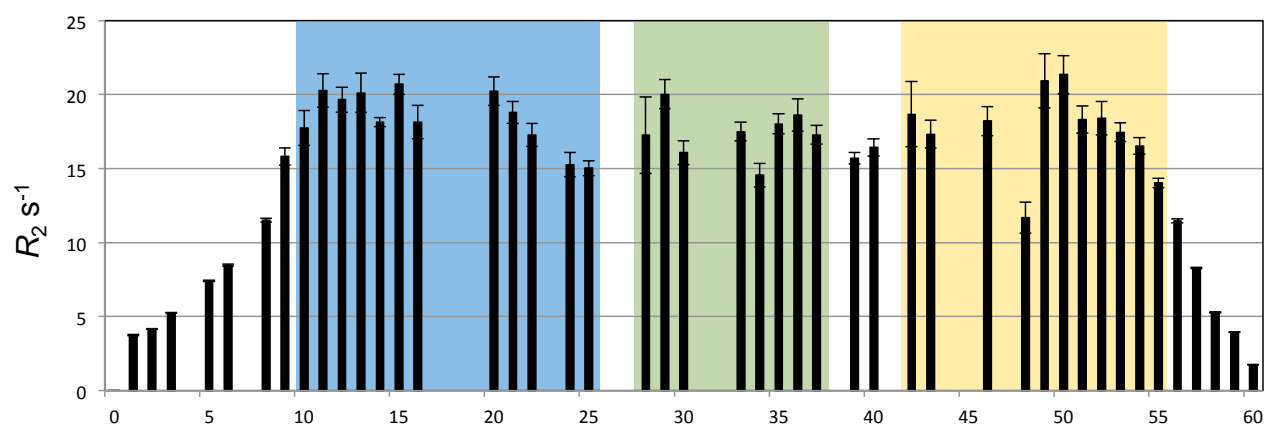
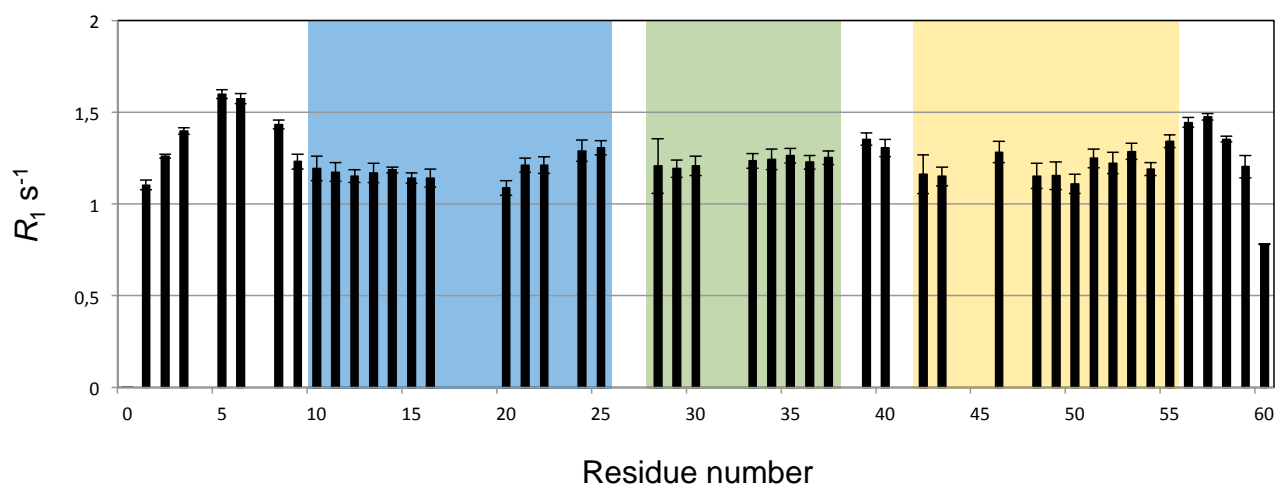


Fig. S6. ^{15}N R_1 and R_2 relaxation rates of EnHD in DMPG bicelles

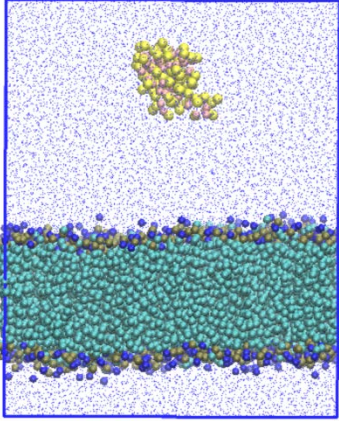
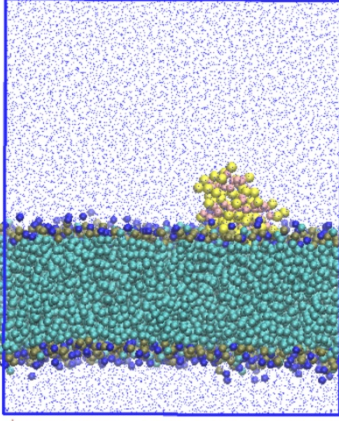
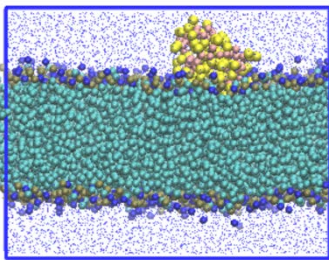
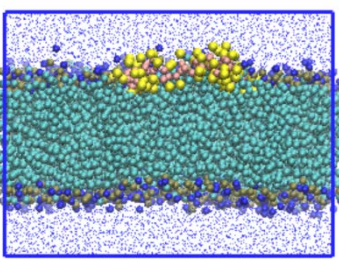
	Starting conformation	Bound conformation
Simulations 1.1, 1.2, 1.3 (with elastic network) Water layer 10 nm thickness		
	Starting bound conformation	Inserted unfolded conformation
Simulations 2.1, 2.2 (without elastic network) Water layer 4 nm thickness		

Fig. S7. Snapshots explaining the general MD strategy used in this paper (see text). Water beads are represented as blue dots. Protein and lipid beads are represented as spheres (backbone: salmon, side chain: yellow, lipids: cyan, ochre and blue).

In simulations 1.1, 1.2, and 1.3, the protein was initially positioned in the aqueous phase, far apart (~3.5 nm) from the membrane bilayer, the layer of water being thick enough (~10 nm) to let the protein freely diffuse before encountering the membrane. Simulations 2.1 and 2.2 were run with less water (~4 nm thickness) since the protein was irreversibly bound to the membrane.

Minimum Distance

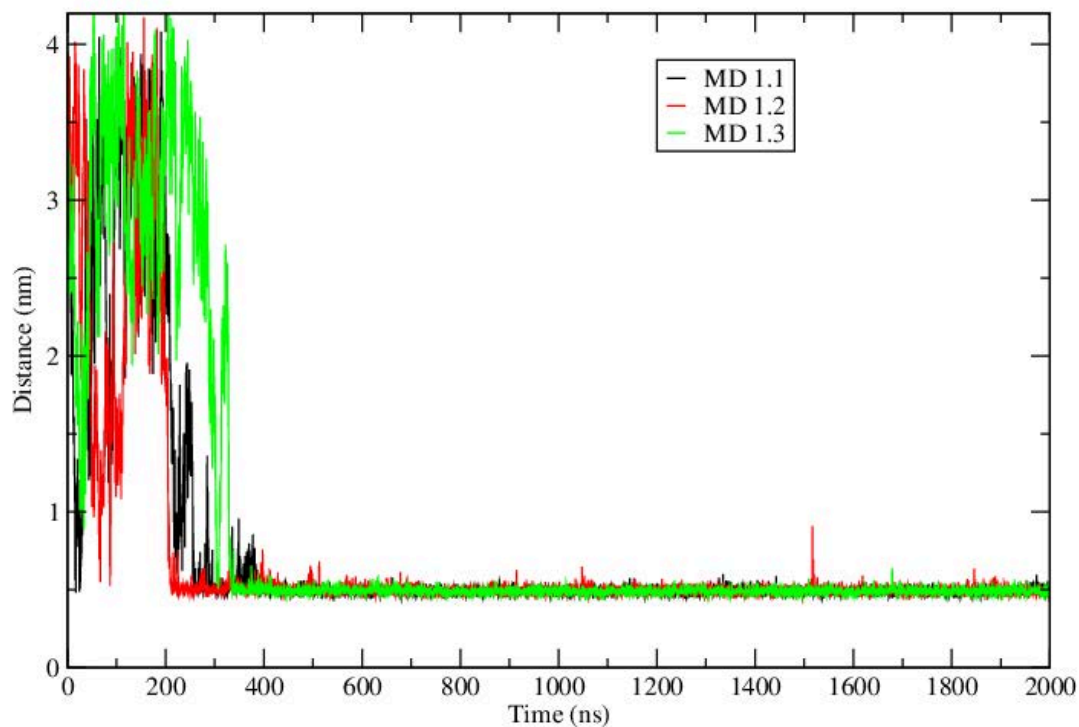


Fig. S8. Minimum distance between the protein and the phosphate beads in simulations 1.1, 1.2 and 1.3. Note that these simulations were performed with the elastic network on the protein. This figure illustrates the nearly irreversible protein binding after a few hundreds of ns (400, 720 and 425 ns for simulations 1.1, 1.2 and 1.3, respectively).

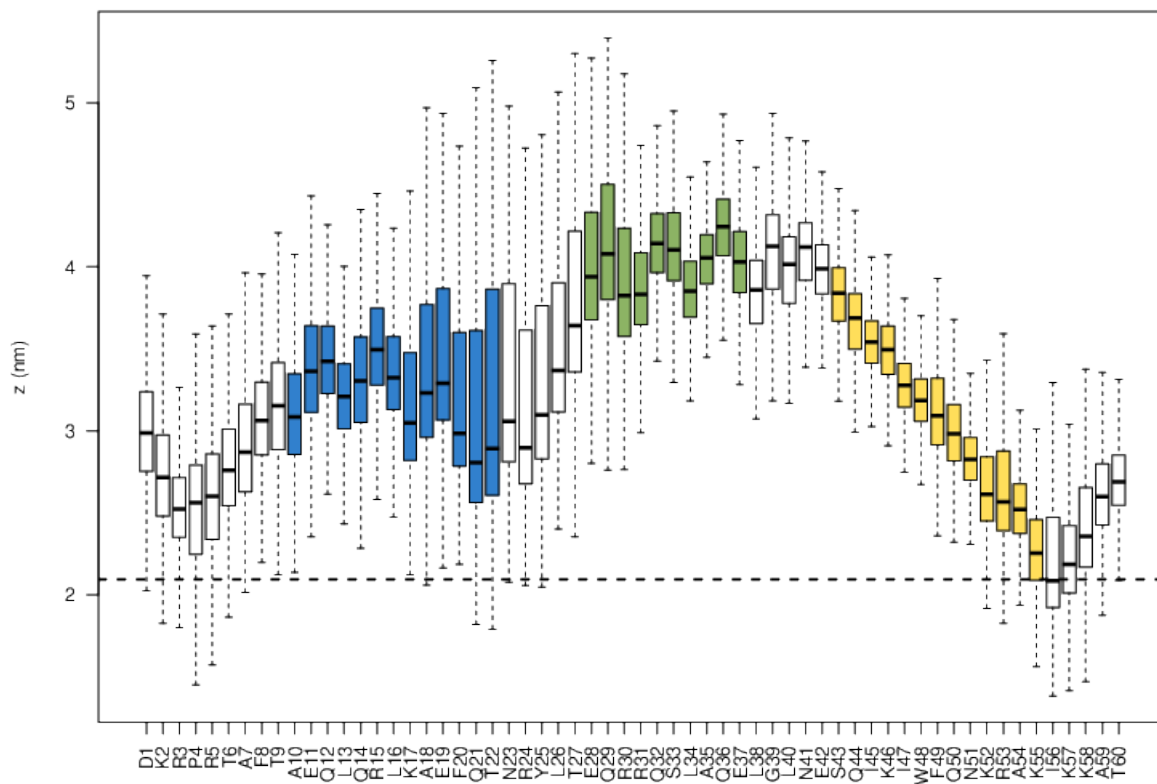


Fig. S9. Mean partitioning of protein backbone beads in simulations 1.1, 1.2 and 1.3. The y axis corresponds to the partitioning in the z dimension of each bead with respect to the center of the bilayer (which would be at the bottom a z of 0 nm). Each box and whiskers correspond to statistics over the 3 simulations, but only the part after 500 ns (once the protein is bound) was used for each trajectory. The protein bound to the upper leaflet in trajectory 1.1 and to the lower leaflet for trajectories 1.2 and 1.3. For simplicity we represent here the protein as if it were bound to the upper leaflet for the 3 trajectories. The dashed horizontal line represents the mean phosphate level of the leaflet to which the protein bounds.

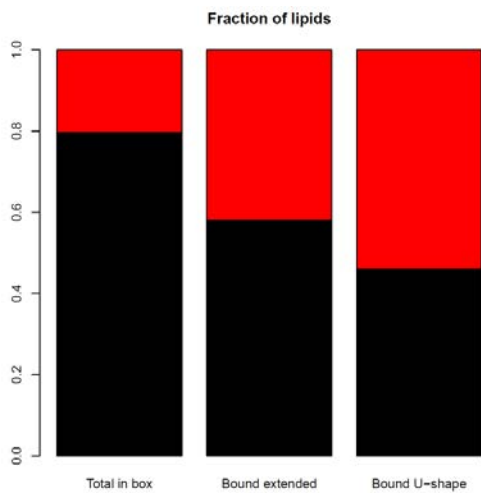
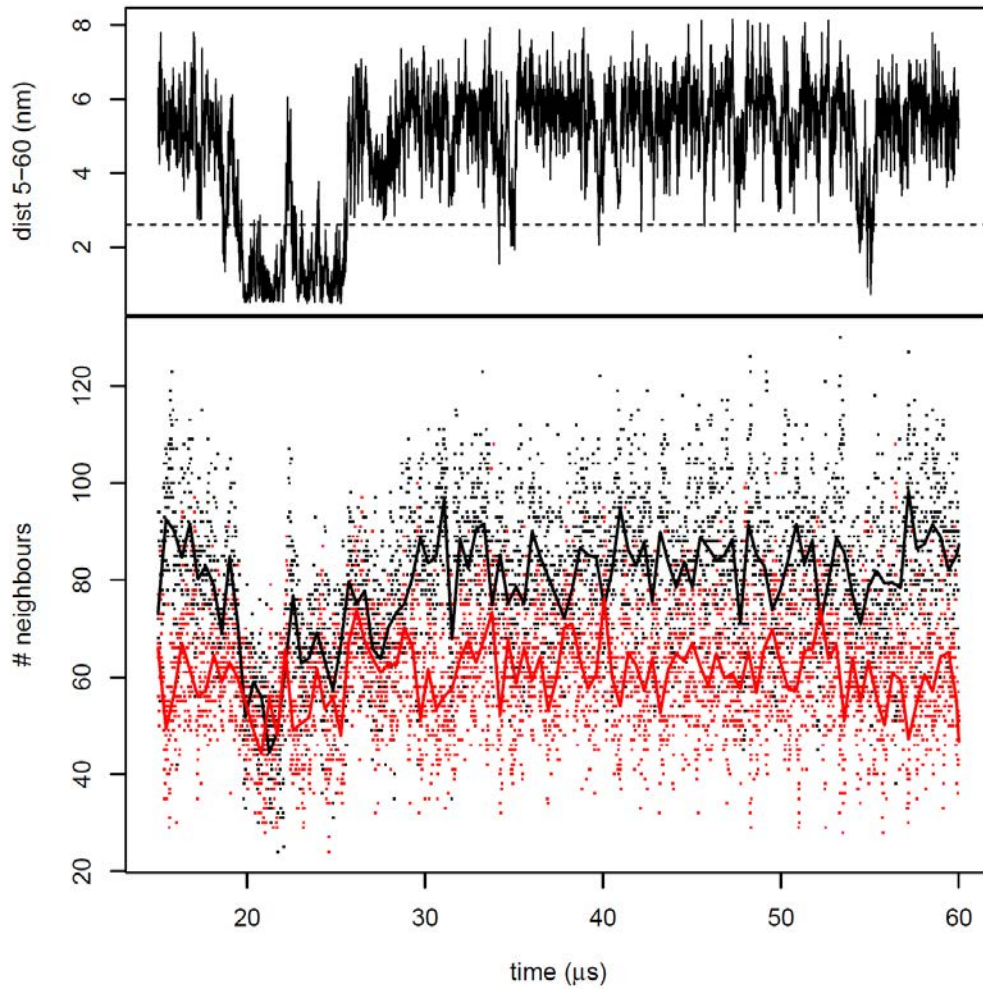


Fig. S10. Evolution of EnHD structure compactness and its enrichment in lipids in simulation 2.2. The top panel shows the distance between the side chain of residues 5 and 60. A low value is indicative of a U-shape conformation (cluster 2), while a high value corresponds to

an extended conformation (cluster 1). The medium panel shows the number of neighboring lipids having a contact with the protein (black DPPC, red DPPG). Each dot represents a number of neighbors for a given conformation, the solid lines represent a smoothing of all dots. The bottom panel shows the molar ratio between DPPC and DPPG for the whole membrane, for the lipids close to the protein in an extended or U-shape conformation.