Supplementary information for manuscript 'Amphiphilic nanotubes in the crystal structure of a biosurfactant protein hydrohpobin HFBII'.

Interactions in 3QQT, amphihilic tube structure.

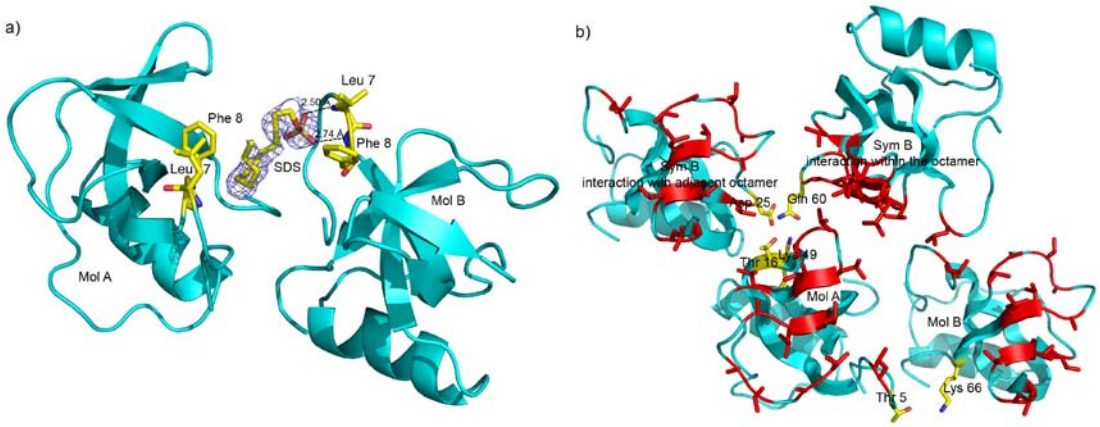
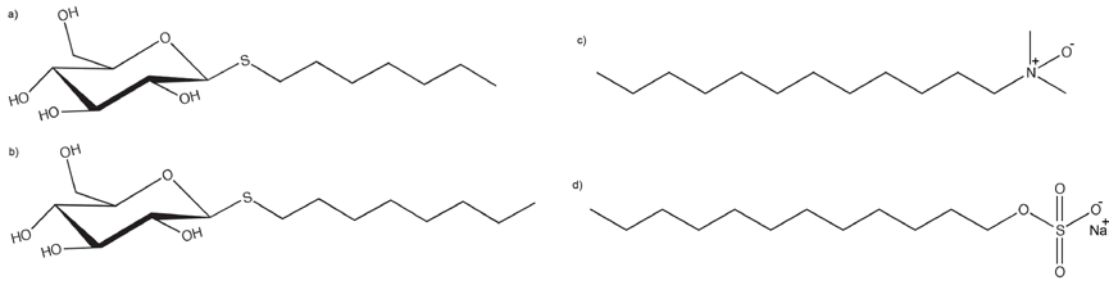


Figure A. For structure 3QQT: a) The OMIT map of detergent SDS drawn at 1.0 σ. Interacting residues Leu7 and Phe8 are labeled. b) The hydrogen bonding of molecule A to molecule B of the asymmetric unit, molecule B within the octamer and molecule B in the adjacent octamer.

Previously determined structures of hydrophobins.

We have previously determined the structure of class II hydrophobins HFBI and HFBII without the presence of detergent interaction (PDB-codes 1R2M or 2B97, and 2FZ6) and in addition, three structures with detergent (PDB codes 2GVM, 2PL6, and 2PL7). These are the only structures available for hydrophobins, excluding the NMR structures for class I hydrophobin EAS from *Neurospora crassa* (PDB codes 2FCM and 2 K6A), for which the oligomerization state cannot be evaluated to the structure determination method. The hydrophobins structures are summarized in the table below and include the structure described in the manuscript. Detailed description of each structure follows.

Protein (class)	PDB	Space group	Resolution	Detergent	Biological unit
HFBII (II)	1R2M	C2	1.0 Å	None	Tetramer
EAS (I)	2FMC	n.a.	n.a.	None	n.a.
HFBII (II)	2B97	C2	0.75 Å	None	Tetramer
HFBI (II)	2GVM	$C222_1$	2.3 Å	LDAO	Octamer
HFBI (II)	2FZ6	C2	2.1 Å	(OSG)	Tetramer
HFBII (II)	2PL7	$P2_12_12_1$	1.0 Å	HSG	Fiber
HFBII (II)	2PL6	$P2_1$	2.2 Å	HSG	Layer
EAS (I)	2K6A	n.a.	n.a.	None	n.a.
HFBII (II)	3QQT	I222	1.9 Å	(SDS)	Tube



Scheme A. Chemical composition of detergents a) HSG, heptyl-β-D-thioglucoside b) OSG, octyl-β-D-thioglucoside c) LDAO, lauryldimethylamine oxide and d) SDS, sodium dodecyl sulphate

HFBII structure 1R2M and 2B97 without detergent interaction.

The functionality of hydrophobins arises from the hydrophobic residues on the protein surface, as presented below for structure 1R2M of HFBII from *Trichoderma reesei* (Figure A). The side chains of the residues contributing to the hydrophobic surface area are represented with sticks, labeled and colored red. Overlaid are the surface representation and the cartoon (helices and strands) representation of the protein. In the diagram Scheme 1 of the manuscript, the hydrophobin monomer is described for simplicity as a block with hydrophobic surface area in thick black. The highlighted part of the Scheme 1 in Figure B below corresponds to the monomeric species.

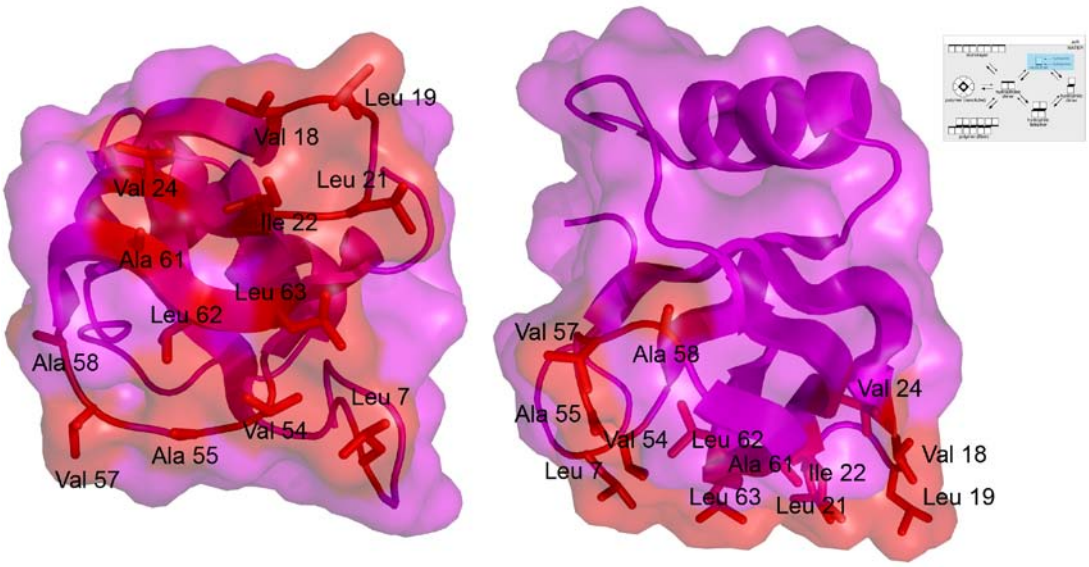


Figure B. The monomeric hydrophobin HFBII from *Trichoderma reesei* and the hydrophobic surface area.

In the solubilized from, hydrophobins appear as tetramers, formed by two dimers (structures 2FZ6 for HFBI and 1R2M for HFBII). In the determination of oligomeric states and assignment of intermolecular interactions, program PISA has been used (for citation, see manuscript). In structure 1R2M the two molecules present in the asymmetric unit (the smallest repeating unit in the

crystal structure) designated as mol A and mol B, form a dimer (Figure C). The symmetry related dimer (by symmetry operation –x, y, -z+1), packs to the groove in the dimer structure and the tetramer is formed. The hydrophobic surface areas are represented in red, however, the stick representation of residues of the hydrophobic patch is omitted in the tetramer structure for clarity.

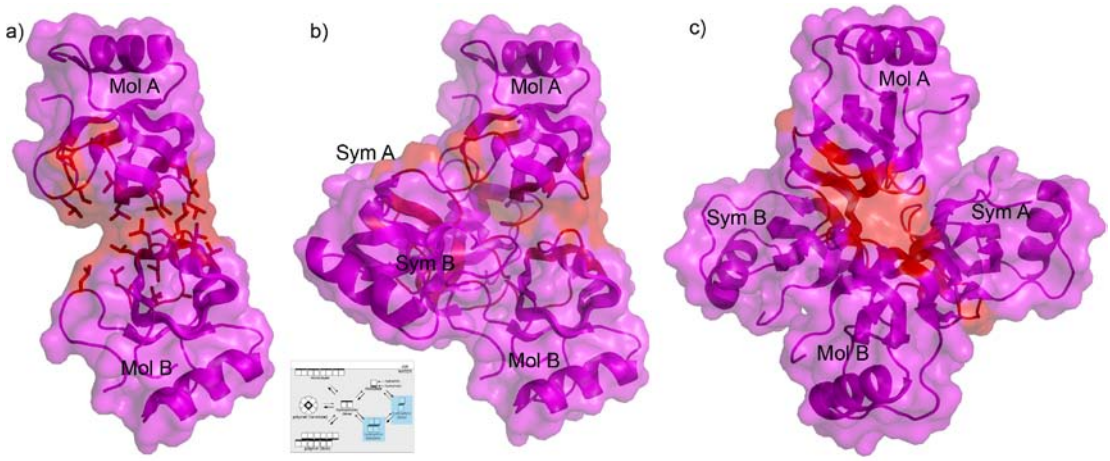


Figure C. a) Dimeric structure of HFBII as found in the asymmetric unit of structure 1R2M b) Tetrameric packing via symmetry-related dimer completes the biological unit c) Tetrameric structure rotated 90° in relation to b).

HFBI structure 2FZ6 without detergent interaction.

The asymmetric unit of HFBI from *Trichoderma reesei* is composed of four protein molecules designated as mol A, mol B, mol C and mol D. In the case of HFBI, detergent (octyl-β-D-thioglucoside, OSG) was added to the crystallization solution, however, it is not found coordinated to the protein structure. Molecules B and D host a conformational change in the area of the hydrophobic surface, demonstrating the plasticity and adaptability of the protein. The biological unit is a tetramer, composed of dimers AB and CD. The assembly and interactions are not strictly the same as in HFBII, however, the oligomerization modes correspond (Figure D).

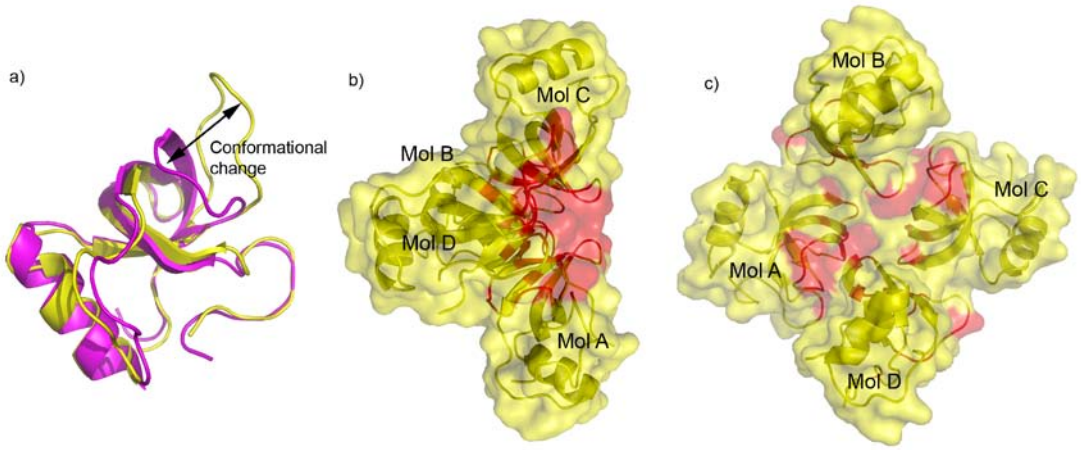


Figure D. a) HFBI mol B (yellow) and HFBII (magenta) superimposed to demonstrate the conformational change. The tetrameric array of the biological unit of HFBI in described in b) and rotated in c).

HFBII structure 2PL6 with detergent interaction, layer structure.

Detergents were used in the crystallization of HFBII structures 2PL6 and 2PL7 and HFBI structure 2GVM. The detergents used were heptyl-β-D-thioglucoside (HSG) in 2PL6 and 2PL7, and lauryldimethylamine oxide (LDAO) in 2GVM. The chemical composition of these detergents are presented in Scheme A.

In each of these structures, the detergent interacts with the hydrophobic surface area of the protein. In 2PL6, the asymmetric unit is composed of eight molecules, designated as A-H. Four of these molecules host a conformational change, similar to that in HFBI (2FZ6). The molecules in the asymmetric unit are in two layers and viewed from the top (along y-axes), form a horseshoe like array. The hydrophobic surface areas in each layer are arranged side by side (Figure E) The hydrophobic carbon chains of the detergent pack against the hydrophobic surface areas while the hydrophilic sugar head of the detergent is oriented towards the solvent. When the symmetry molecules are taken into account, two more molecules originating from the crystal symmetry complete the horseshoe arrangement into a six-membered ring structure. These hexameric rings have also been observed for natively assembled hydrophobin HFBII films on air-water interface by AFM, with corresponding dimensions (see citation 2 in manuscript for details) and thus we have concluded that the formation of these six-membered rings mimics the natural assembly mode. It is also noteworthy, that the interaction between the molecules in the asymmetric unit in the 3QQT structure of the amphiphilic tubes is the same as the interaction between molecules G and F in the six-membered ring in 2PL6.

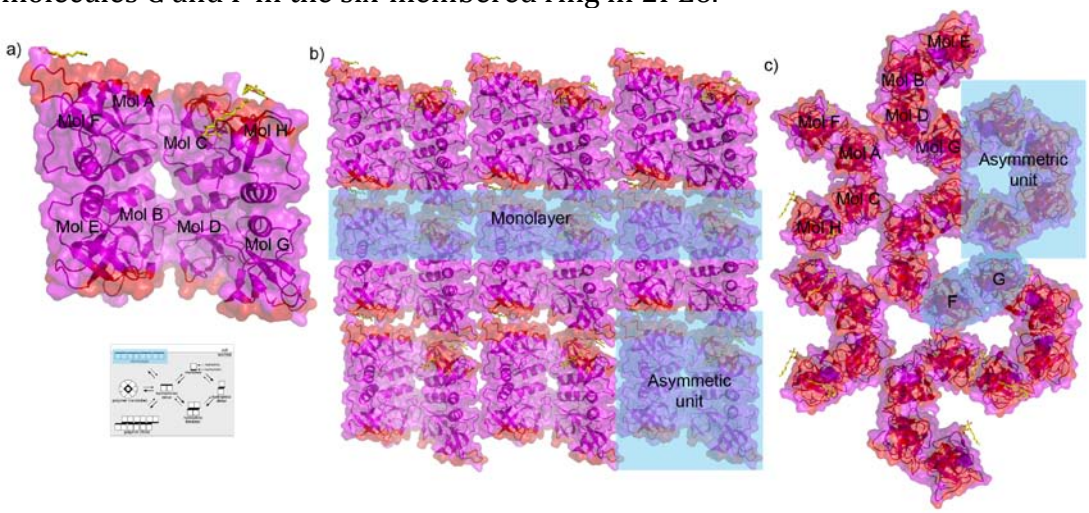


Figure E. a) Contents of the asymmetric unit in structure 2PL6 b) Arrangement of molecules in layer structure by crystal packing – monolayer is thought to correspond to the naturally occurring monolayer on air-water interface. c) Crystal packing showing the six-membered rings formed by neighboring asymmetric units, resembling the porous structure of hydrophobin monolayers on air-water interface viewed perpendicular to water surface. Highlighted is the

interaction between molecules F and G from neighboring asymmetric units, which corresponds exactly to interaction between the molecules of the asymmetric unit in structure 3QQT.

HFBII structure 2PL7 with detergent interaction, fiber structure.

In 2PL7, there are two molecules in the asymmetric unit, designated as A and B (Figure F). These molecules form a dimer with hydrophobic surface areas on one side of the dimer, corresponding to amphiphilic dimer in Scheme 1. The amphiphlic dimer is formed (with modest movement of the position of molecules) by the soluble tetramer splitting after having met a hydrophobic-hydrophilic interface and the dimeric species is the basic building block for the formation of the layer structure (2PL6 and naturally occurring monolayer films) and the fibrils structures observed in 2PL7 and 3QQT. The detergent in 2PL7 coordinates to hydrophobic surface area via the hydrophobic carbon tail, again leaving the sugar head group towards solvent. The detergents rather sit on top of the hydrophobic surface, while in 3QQT the detergent is incorporated between the protein molecules. The Phe8 residues are mostly buried in the 2PL7 structure. The fibril formation is via the hydrophobic surface areas, with dimers arranging side-by-side, hydrophobic surface areas to uniform direction, and a corresponding row of dimers with 180° rotation facing this row. Thus each molecule A is in contact with molecule B via the hydrophobic surface area. The formation of the fibrilar array conceals the hydrophobic surface areas from the solvent almost completely, resulting into a mostly hydrophilic fibril. Hydrophobins are known to form fibrils also in nature, and for class I hydrophobins the fibrils are amyloidal. We have suggested that the fibril formation occurs when there is excess amount of hydrophobin available in relation to demand to cover the hydrophobic-hydrophilic interface.

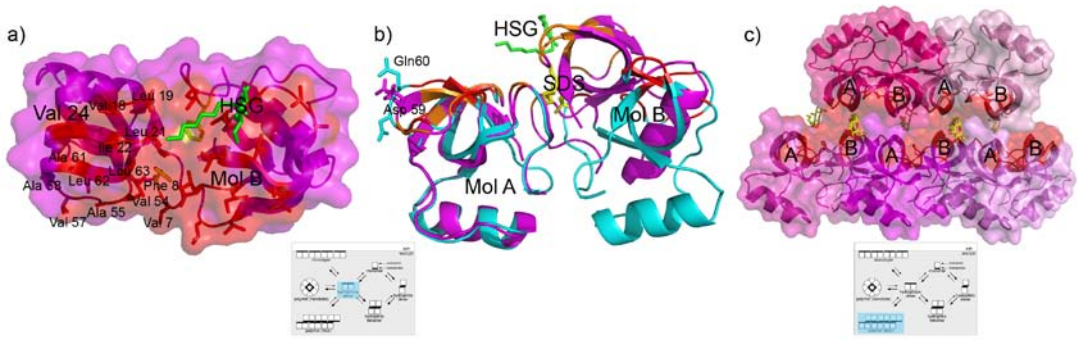


Figure F. a) The amphiphilic dimer in the crystal structure 2PL7 of HFBII. The hydrophobic surface area in red and the contributing residues labeled in molecule A. In yellow, Phe8 residues in the vicinity of hydrophobic area but in this structure, mostly buried. The detergent molecule in green. B) The amphiphilic dimers of structures 2PL7 (fibril) in magenta and 3QQT (amphiphilic tube) in cyan superimposed for molecule A. Detergents SDS and HSG in yellow and green, respectively. The area of conformational differences (Asp59 to Gln60) shown in stick model. c) Five asymmetric units in different shade of magenta showing the formation of hydrophilic fibril. Molecules A and B are labeled and detergents are indicated in yellow.

HFBI structure 2GVM with detergent interaction, octamer structure

2GVM is a mutant structure of HFBI, in which two hydrophobin molecules were joined together by a linker. The linker region does not appear in the electron density either because it is too flexible, or because the crystallized from does not contain the linker. In the asymmetric unit, there are four protein molecules, designated Mol A-D (Figure G). The asymmetric unit also contains 10 LDAO-detergent molecules. The four molecules in the asymmetric unit are arranged so, that the hydrophobic surface areas align partially on the tetramer interface and partially on one side of the tetramer. Two detergent molecules are occupying the space at the tetramer interface with hydrophobic carbon chain in contact with the hydrophobic surface area. The rest of the detergent molecules are located on one side of the tetramer again in contact with the hydrophobic surface areas of the protein. Two neighboring asymmetric units form a sandwich-like structure with majority of detergents in the middle and hydrophobin tetramers on each side. The crystal packing, as also the detergent structure, resembles of that in 3QQT, however no tubes are formed and the structure is more porous and contains much more solvent.

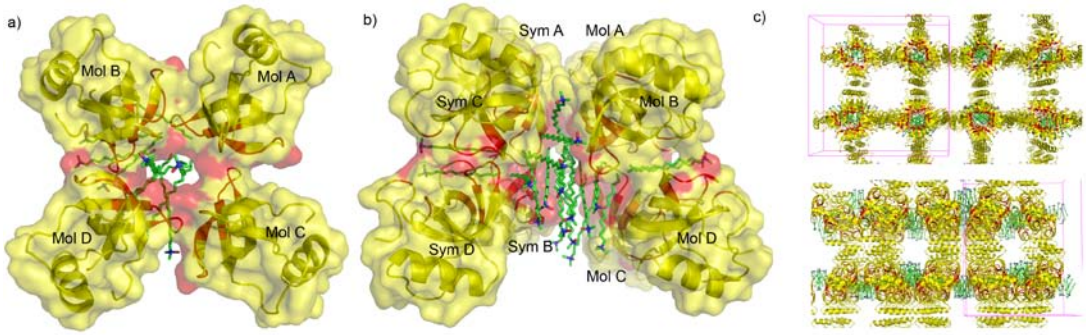


Figure G. a) The asymmetric unit of HFBI structure 2GVM, with hydrophobic surface areas in red and the detergents in green tone. b) The functional unit is a sandwich-like structure composed of two neighboring asymmetric units c) The crystal packing of 2GVM resembles that of 3QQT, however, to tubes are formed.