Hydrophobins are proteins specific to filamentous fungi. Hydrophobins have several important roles in fungal physiology, for example, adhesion, formation of protective surface coatings, and the reduction of the surface tension of water, which allows growth of aerial structures. Hydrophobins show remarkable biophysical properties, for example, they are the most powerful surface-active proteins known. To this point the molecular basis of the function of this group of proteins has been largely unknown. We have now determined the crystal structure of the hydrophobin HFBII from *Trichoderma reesei* at 1.0 Å resolution. HFBII has a novel, compact single domain structure containing one α-helix and four antiparallel β-strands that completely envelop two disulfide bridges. The protein surface is mainly hydrophilic, but two β-hairpin loops contain several conserved aliphatic side chains that form a flat amphiphilic patch that makes the molecule amphiphilic. The amphiphilicity of the HFBII molecule is expected to be a source for surface activity, and we suggest that the behavior of this surfactant is greatly enhanced by the self-assembly that is favored by the combination of size and rigidity. This mechanism of function is supported by atomic force micrographs that show highly ordered arrays of HFBII at the air water interface. The data presented show that much of the current views on structure function relations in hydrophobins must be re-evaluated.

Fungi are found all around us in nature in decaying plant litter, as mushrooms that grow in the forests, or as molds that spoil foods. They have a remarkable ability to adapt to a wide variety of environmental conditions and ecosystems (1). They grow by spreading their hyphae, which can penetrate substrates upon which they grow, they can grow upwards into the air to spread spores, or they can attach to solid supports. One group of proteins that have been found to have an important role for fungal growth and development are the hydrophobins. They seem to be unique to fungi and have not been found in any other organisms. A common property in hydrophobins is that they seem to have a wide range of functions that are all apparently related to surface activity in one way or the other (2). Here, surface activity means the tendency to adsorb at interfaces and surfaces. The term interface means the boundary between, for example, air and liquid, or liquid and solid. Another word for a surface-active compound is “surfactant” (3). When a surfactant compound migrates to the air-water interface, it lowers the surface tension, i.e. the cohesive forces between water molecules at the surface.

Hydrophobins are, for example, involved in the formation of aerial structures in fungi (4–6). Here, the hydrophobins seem to serve a dual role by first lowering the water surface tension, which allows the fungus to penetrate this barrier, and then by forming a protective coating on the aerial structures and spores. Fungi also rely on hydrophobins for attachment to surfaces such as plant leaves or insect cuticles. This property has been demonstrated to be essential for the infectivity of some pathogenic fungi, which cause Dutch elm disease, rice blast, and chestnut blight (7). It is quite possible that the properties of hydrophobins have played an important role during evolution in enabling filamentous fungi to occupy their ecological niches.

Hydrophobins are among the most surface-active biomolecules known (2). They have a size of about 100 amino acids, and are also remarkably stable and can withstand temperatures near the boiling point of water. Consequently, they have potential for several applications. Hydrophobins can be used as adhesion domains for the immobilization of proteins to solid supports (8, 9) and as tags in fusion proteins for affinity purification (10). The latter method relies on the very high partitioning of hydrophobins in surfactant-water two phase systems (11). Hydrophobins are also highly efficient foam-forming agents. This property can be illustrated by a practical example. It has been found that microgram amounts of hydrophobin in a bottle of beer causes an explosive gush of foam upon opening the bottle. The contaminating hydrophobins probably originated from fungal growth during the malting or earlier stages (12).

A large number of genes for hydrophobins have been identified, and they all share a characteristic pattern of Cys residues in their primary sequence (Fig. 1) (13). Hydrophathy patterns can be used to divide them into classes I and II (14). The proteins in the two classes seem to differ in the solubility of the aggregates that they form. The sequence similarity of hydrophobins is in general weak, making it sometimes difficult to define them based on primary sequence. However, the characteristic and unique pattern of Cys residues is conserved in all hydrophobins. An essential feature of hydrophobins is that they form different supramolecular assemblies. It has been shown by atomic force microscopy that the class II hydropho-
bins HFBI and HFBII from Trichoderma reesei form highly ordered monolayer films (15, 16). The same proteins also form crystalline fibrils (17). On the other hand, the class I hydrophobin SC3 from Schizophyllum commune forms a "rodlet coating" where nanometer size rods are formed on interfaces by self-assembly (18). These and other class I hydrophobin assemblages are remarkable because of their insolubility. They are insoluble even in hot solutions of sodium dodecyl sulfate and can only be dissolved in some strong acids, such as trifluoroacetic acid. After evaporation of the acid, it is possible to again solubilize the protein in water and to make new coatings that are indistinguishable from the original ones. For some insoluble class I members, it has been shown that aggregates formed have similarities to amyloid fibrils based on reactions with the stain Congo Red (19, 20).

To this point, the properties of hydrophobins such as aggregation, adhesion, and surface activity have hampered structural studies. For example, as is shown in the current work, the widely used data on disulfide pairing (21) were incorrect. An NMR study of the EAS hydrophobin from Neurospora crassa conidia (class I) suggested that the protein is largely unstructured in solution, showing only a small region of β-sheet structure (19). Studies on the SC3 hydrophobin have suggested that the protein has different conformational states and different secondary structure contents when in solution or at a hydrophobic-water interface (20, 22). It has also been proposed that the protein refolds at hydrophobic-hydrophilic interfaces (23).

In this work, we present the first three-dimensional structure of a hydrophobin. We also present new atomic force microscopy (AFM) data showing that this hydrophobin forms highly ordered self-assembled layers at the air/water interface. Together, these data suggest a mechanism for how this protein functions.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—HFBII was produced using the T. reesei strain (VTT number D-99745) and grown as described in Ref. 24. To one liter of the centrifuged (4000 × g, 20 min) culture supernatant 2 g Berol 325 (Akzo Nobel) was added. The solution was allowed to settle in a separation funnel after mixing and the Berol phase was collected. The Berol was extracted with a 100 ml acetate buffer (50 mM, pH 5.0) and 15 ml of isobutanol with a subsequent centrifugation (3220 × g, 15 min). The protein was further purified by preparative reversed phase chromatography using a Vydac C4 (1 × 20 cm) column and a gradient elution from 0.1% trifluoroacetic acid to 100% acetonitrile containing 0.1% trifluoroacetic acid. Peak fractions were pooled and lyophilized. The protein identity was confirmed using mass spectrometry and antibodies (11, 25).

TABLE I

| Data collection and refinement statistics | Values in parenthesis are for the highest resolution shell (1.00–1.02). |
|----------------------------------------|-------------------------------------------------------------------|
| Space group                            | C2                                                                |
| Unit cell dimensions (Å)               | 78.66, 46.31, 34.59                                               |
| Unit cell angles (°)                   | 90, 112.16, 90                                                    |
| Resolution range (Å)                  | 25.1-1.0                                                          |
| Completeness (%)                      | 99.8 (99.6)                                                       |
| Rmerge (%)                            | 6.3 (23.5)                                                        |
| Rfree (%)                             | 7.6 (3.2)                                                         |
| Number of observations                 | 214050                                                            |
| Unique reflections                     | 61978                                                             |
| Temperature (°C)                      | 100 K                                                             |
| R-factor (%)                          | 15.8 (17.2)                                                       |
| Rfree (%)                             | 15.5 (18.8)                                                       |
| RMSD bond lengths (Å)                 | 0.007                                                             |
| RMSD bond angles (°)                  | 1.394                                                             |
| Total number of atoms                  | 1254                                                              |
| Number of protein atoms                | 1011                                                              |
| Number of water molecules             | 242                                                               |

Crystallization and Structure Determination—The protein crystals grew in hanging drops using 25% polyethylene glycol 2000, 0.2 M lithium sulfate, 10 mM MnCl₂, and 0.1 M Na-HEPES (pH 7.5). Crystallization drops were prepared by mixing equal amounts of protein solution (8 mg/ml) and crystallization agent. The data set was collected at 1.5 Å resolution using a copper rotating anode as an x-ray source (wavelength 1.54 Å) and at 1.0 Å resolution at the EMBL X11 beamline at the DORIS storage ring, DESY (wavelength 0.8126 Å) and processed with DENZO (26). The anomalous difference Patterson map calculated with the data measured at the home laboratory showed a clear single peak at Harker sections, suggesting that one manganese ion would have bound to the protein. The coordinates of this position were used in the program ACORN (27) as a starting point for the calculation of phases. The refined phases were further used in the program ARP/WARP (28), which automatically built the preliminary model for two protein molecules in the asymmetric unit. The structure was refined by iterative cycles of manual refitting with O (29) and positional refinement with REFMAC. The final model contains two protein molecules in a asymmetric unit. The refined molecules contain residues 1–70. The last residue Phe-71 was not visible and not included in the model. The figures were drawn with Setor (30) and PyMOL. The details of data collection and structure refinement are provided in Table I. Coordinates have been deposited in the Protein Data Bank with the code 1R2M. A 10 µl droplet of 10 µg/ml solution of HFBII in double distilled water was dried on highly oriented pyrolytic graphite (HOPG) in a vacuum desiccator. Images were acquired under ambient conditions in tapping mode using a NanoScopeIIIa Multimode AFM (Digital Instruments) equipped with an “E” scanner. The tapping mode was used with scan rates 0.6–1 Hz and as low a force as possible. Non-contact silicon cantilevers (NCH, Nano-Sensors) with the nominal resonance frequency of around 300 kHz and tip radius ca. 10 nm were used. A scanning probe image processor (SPIP, Image Metrology, Denmark) was used in image processing, which only included flattening in order to remove possible tilt in the image data, and in image analysis.

Fig. 1. Amino acid sequence comparison of class II hydrophobins. The secondary structure elements are given in the first row. The Cys residues are in black, and the residues found in the hydrophobic patch are in red. Protein identification related to the abbreviations used are: HFBII, T. reesei (accession P79073); HFBII, T. reesei (accession P52754); SRHII, Trichoderma harzianum (accession Y11641); QID3, Trichoderma harzianum (accession P52755); TR11, TR12, and TR5, the three segments of TH1 Claviceps purpurea (accession Q9UV14); CPPH1, CPPH1_2, CPPH3, CPPH4, and CPPH5, the five segments of cpph1 Claviceps purpurea (accession AJ418045); CRYP, Cryparin Cryphonectria parasitica (accession P52755); CU, Cerato-ulmii Ophiostoma ulmii (accession Q06153); MAG, Magnaporthe Magnaporthe grisea (accession AF128872); HCF5, Cladosporium fumatum A1335763; HCF6, C. fumatum (accession CAC27408).

The abbreviation used is: AFM, atomic force microscopy.
HFBII residues Val-18, Leu-19, Leu-21, Ile-22, and Val-24 are linked to an antiparallel $\beta$-strand. The first $\beta$-hairpin is in red and the second in purple. The central $\beta$-barrel consists of two $\beta$-hairpins that interlock in the same way as the leather pieces of a baseball.

FIG. 2. Topology and structure of HFBII. A, topology of HFBII. The arrows mark $\beta$-strands and the rectangle the $\alpha$-helix. B, three-dimensional structure of HFBII. The first $\beta$-hairpin is in red and the second in purple. The central $\beta$-barrel consists of two $\beta$-hairpins that interlock in the same way as the leather pieces of a baseball (C).

We have crystallized a member of the class II hydrophobins, HFBII from T. reesei, and determined its three-dimensional structure at 1.0 Å resolution. HFBII is a single domain protein with the approximate dimensions of $24 \times 27 \times 30$ Å. The structure represents a fold not previously described. It consists of one $\beta$-hairpin motif (strands S1 and S2) linked to an $\alpha$-helix that is linked to the second $\beta$-hairpin motif (strands S3 and S4). These two motifs are arranged together in an unusual way (31). They form a barrel consisting of four antiparallel $\beta$-strands in the order S4-S2-S1-S3. The two $\beta$-hairpins interlock together in much the same way as the two leather pieces of a baseball are sewn together (Fig. 2).

FIG. 3. The structure of HFBII in stereo showing $\beta$-barrel and four symmetrically arranged disulfide bridges in yellow. $\beta$-strands and Cys residues are labeled.

Disulfide Bridges—A distinctive feature of hydrophobins is that they contain eight Cys residues that form four disulfide bridges (Fig. 3). The current structure shows that bridge 14–26 connects the strands of the first $\beta$-hairpin together and that bridge 53–64 connects the strands of the second $\beta$-hairpin. The four sulfur atoms of these two bridges are completely inside the $\beta$-barrel. The remaining two bridges are outside the $\beta$-barrel. The bridge 3–52 connects the N-terminal loop to $\beta$-strand S3, and the bridge 13–43 connects the $\beta$-strand S1 and the helix together. The disulfide bridges are symmetrically located in almost the same plane in the structure (Fig. 3). They form a disulfide bridge network that efficiently spans the entire molecule thus creating a very stable and compact structure.

Hydrophobic Surface Patch—Hydrophobins are highly surface-active and interact strongly with hydrophobic surfaces. Surface features of the protein are expected to be important for this function. The amino acid sequence comparison of class II hydrophobins revealed a number of conserved and exposed hydrophobic residues (Figs. 1 and 4). Interestingly, these residues are all aliphatic and are located in two $\beta$-hairpins. The HFBII residues Val-18, Leu-19, Leu-21, Ile-22, and Val-24 are located in the first hairpin (loop 1). Residues Val-54, Val-57, Ala-58, Ala-61, and Leu-63 are located in the second hairpin (loop 2). A hydrophobic surface patch is thus formed and is completed by Leu-7 in the N-terminal loop. As we examine the overall shape of the protein, we note that the surface at the hydrophobic patch is relatively flat. It also seems that loop 1 has a pronounced role in the formation of this hydrophobic patch. The total surface area of a HFBII monomer is about 3200 Å$^2$, and the estimated area of the conserved hydrophobic patch is about 400 Å$^2$ (12% of the total).

Self-assembled Layers—AFM images show polycrystalline structures consisting of two-dimensional single crystalline domains in the surface layers of HFBII. In Fig. 5, an image of $100 \times 100$ nm$^2$ in size is shown. The crystallinity of the protein layer was confirmed and the lattice constants were determined by Fourier transformation of the image data. An oblique structure with lattice constants $a = 67.1$ Å, $b = 44.8$ Å, $\gamma = 139.2$ degree was obtained. The lattice constants equal qualitatively to the lattice constants found in Langmuir-Blodgett films of HFBII (15). The slight distortion observed in the AFM images, and hence in the transformed images also, is due to scanner hysteresis, creep, and drift in the microscope. These problems are important to consider because of the relatively slow scan speeds and because images of the first full scan were captured in order to obtain images of the soft protein surfaces with a non-contaminated tip. The way in which the sample was prepared and the similarity to the previously studied Langmuir-Blodgett films suggest that the image most likely represents the air-water interface that is deposited on the solid support as the water has evaporated.

DISCUSSION

The presented structure of HFBII offers new perspectives for understanding and exploiting the biophysical properties of hydrophobins, such as their strongly amphiphilic nature, self-assembly, and aggregation. HFBII has a globular structure that is stabilized by a network of disulfide bonds. On the protein surface, there is distinct patch of aliphatic side chains that are conserved in homologous proteins. We suggest that this patch makes the hydrophobin molecule amphiphilic and is therefore highly important to the function of the protein. One
side of the molecule is thus easily solvated by water, whereas
the other tries to escape it. However, the exposed hydrophobic
side chains must be prevented from turning to the inside of the
protein hydrophobic core to escape water. Therefore, the rigid-
ity of the HFBII molecule may assist in keeping the side chains
of the hydrophobic patch exposed to solvent.

In addition to the atomic resolution structure of HFBII, we
also present here atomic force micrographs of HFBII that show
molecular films with highly regular patterns. We previously
have shown that ordered films can be made using the Lang-
muir-Blodgett technique (15), but by refining the technique we
were now able to show that such ordered interface layers are
formed spontaneously at the surface of water droplets. We thus
have two views, i.e. the structure of the individual molecule,
and the structure of an assembly that they form. Together,
these data can be used to understand how this protein
functions.

As surfactants (surface-active agents), the hydrophobins are
undoubtedly interesting molecules. HFBII is able to reduce the
surface tension of water from 72 mJ/m^2 to 28 mJ/m^2 at a
concentration of 20 μg/ml. One major difference, compared
with most other highly surface-active molecules, is that the
relatively large size of HFBII is combined with rigidity in the
molecule, i.e. it is a shape-persistent molecule. The question
arises how shape persistency affects amphiphilicity compared
with more small or flexible surface-active molecules. A possible
effect is that by reducing the conformational degrees of free-
dom, the mixing entropy in solutions can be drastically re-
duced, which promotes aggregation tendency and reduces sol-
ubility (see Ref. 32). This may be important in surface
attachment in order to guarantee that the hydrophobins ad-
here well and are not lost due to redissolution. That is, the
highly ordered surface layer is formed because the individual
molecules are rigid, amphiphilic, and large enough to make
favorable intermolecular contacts. Therefore, favorable lateral
contacts can enhance the surface activity of hydrophobins. The
role of these factors on the origin of the high surface activity of

---

\(^2\) S. Askolin, manuscript in preparation.
hydrophobins could be tested by mutating amino acids, such as the charged residues, putatively forming lateral interactions.

Previously, it has been shown that hydrophobins can have a complex multimerization behavior in solutions and at interfaces (16). HFBII probably forms tetramers at high protein concentrations, but dimers or monomers occur at lower concentrations (17). The wild-type HFBII hydrophobin behaved similarly, but formed aggregates approximately the size of decamers when it was genetically fused to a 40-kDa cellulase enzyme (9). In the crystal structure of HFBII, the asymmetric unit contains two protein molecules that form a dimer (Fig. 6A). The hydrophobic patches of both molecules are shown in yellow and orange. B, a model for the assembly of HFBII. There is equilibrium between the monomeric, oligomeric, and monolayer state. Amphiphilicity drives a monomeric hydrophobin to oligomers or monolayers. The monolayer has a highly crystalline structure (see Fig. 5), indicating favorable lateral interactions that further stabilize it.

Fig. 6. Multimerization and assembly of HFBII. A, the observed structure of a HFBII dimer. Different HFBII monomers are colored in blue and red. The hydrophobic patches of both molecules are shown in yellow and orange. B, a model for the assembly of HFBII. There is equilibrium between the monomeric, oligomeric, and monolayer state. Amphiphilicity drives a monomeric hydrophobin to oligomers or monolayers. The monolayer has a highly crystalline structure (see Fig. 5), indicating favorable lateral interactions that further stabilize it.
Acknowledgments—We thank EMBL/DESY, Hamburg and its staff for help with the data collection. We also thank O. Ikkala and J. Lendesborough for discussions and useful comments on the manuscript.

REFERENCES

1. Gow, N., and Gadd, G. (1995) *The Growing Fungus*, Chapman & Hall, London
2. Wosten, H. A., and de Vocht, M. L. (2000) *Biochim. Biophys. Acta* 1469, 79–86
3. Jonsson, B., Lindman, B., Holmberg, H., and Kronberg, B. (1998) *Surfactants and Polymers in Aquous Solution*, John Wiley & Sons, Chichester, UK
4. Talbot, N. J. (1997) *Curr. Biol.* 7, R78–R81
5. Talbot, N. J. (1999) *Nature* 398, 295–296
6. Wosten, H. A., van Wetter, M. A., Luigenes, L. G., van der Mei, H. C., Busscher, H. J., and Wessels, J. G. (1999) *Curr. Biol.* 9, 85–88
7. Tucker, S. L., and Talbot, N. J. (2001) *Biochim. Biophys. Acta* 1569, 139–150
8. Linder, M., Selber, K., Nakari-Setälä, T., Söderlund, H., and Penttilä, M. (2002) *Protein Sci.* 11, 2257–2266
9. Collen, A., Persson, J., Linder, M., Nakari-Setälä, T., Penttilä, M., Tjerneld, F., and Sivars, U. (2002) *Biochim. Biophys. Acta* 1569, 139–150
10. Linder, M., Selber, K., Nakari-Setälä, T., Qiao, M., Kula, M.-R., and Penttilä, M. (2001) *Biomacromolecules* 2, 511–517
11. Kleemola, T., Nakari-Setälä, T., Linder, M., Penttila, M., Olkku, J., and Haikara, A. (2002) *Proceedings of the European Brevity Convention 2001*, pp. 127–138
12. Wosten, H. A. B. (2001) *Annu. Rev. Microbiol.* 55, 625–648
13. Wosten, H. A. B. (1994) *Annu. Rev. Phytopathol.* 32, 413–437
14. Pananan, A., Vuorimaa, E., Torkkeli, M., Penttila, M., Kauranen, M., Ikka, O., Lemmetyinen, H., Serimaa, R., and Linder, M. (2003) *Biochemistry* 42, 5253–5258
15. Serimaa, R., Torkkeli, M., Pananan, A., Linder, M., Kisko, K., Knapila, M., Ikka, O., Vuorimaa, E., and Lemmetyinen, H. (2003) *J. Appl. Cryst.* 36, 499–502
16. Torkkeli, M., Serimaa, R., Ikka, O., and Linder, M. (2002) *Biophys. J.* 83, 2249–2247
17. Wosten, H. A. B., de Vries, O. M. H., and Wessels, J. G. H. (1993) *Plant Cell* 5, 1567–1574
18. Mackay, J. P., Matthews, J. M., Winefield, R. D., Mackay, L. G., Haverkamp, R. G., and Templeton, M. D. (2001) *Structure* 9, 83–91
19. de Vocht, M. L., Reviakine, I., Wosten, H. A., Brisson, A., Wessels, J. G., and Robillard, G. T. (2000) *J. Biol. Chem.* 275, 28428–28432
20. Yaguchi, M., Puszari-Carey, M., Roy, C., Surewicz, W. K., Carey, P. R., Stevenson, K. J., Richards, W. C., and Takai, S. (1995) in *Dutch Elm Disease Research, Cellular and Molecular Approaches* (Stucklin, M., and Sherwin, J. L., eds) pp. 152–170, Springer Verlag, New York
21. de Vocht, M. L., Reviakine, I., Ulrich, W. P., Bergsma-Schutter, W., Wosten, H. A., Vogel, H., Brisson, A., Wessels, J. G., and Robillard, G. T. (2002) *Protein Sci.* 11, 1199–1205
22. Zangi, R., De Vocht, M. L., Robillard, G. T., and Mark, A. E. (2002) *Biophys. J.* 83, 112–124
23. Bailey, M., Askolin, S., Horhammer, N., Tenkanen, M., Linder, M., Penttilä, M., and Nakari-Setälä, T. (2002) *Appl. Microbiol. Biotechn.* 58, 721–727
24. Nakari-Setälä, T., Aro, N., Ilmen, M., Munoz, G., Kalkkinen, N., and Penttilä, M. (1997) *Eur. J. Biochem.* 245, 415–423
25. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* 276, 307–326
26. Foadi, J., Woolfson, M., Dodson, E. J., Wilson, K. S., Jia-xing, Y., and Chao-de, Z. (2000) *Acta Cryst.* D56, 1137–1147
27. Perrakis, A., Morris, R. M., and Lannin, V. S. (1999) *Nature Struct. Biol.* 6, 458–463
28. Jones, A. T., Zhou, J. Y., Cowan, C. W., and Kjeldgaard, M. (1991) *Acta Cryst.* A47, 110–119
29. Evans, S. V. (1993) *J. Mol. Graph.* 11, 134–138
30. Branden, C., and Tooze, J. (1998) *Introduction to Protein Structure*, pp. 31, Garland Publishing, New York
31. Gedde, U. W. (1995) *Polymer Physics*, Chapman & Hall, London
32. Scholtmeijer, K., Wessels, J. G. H., and Wosten, H. A. B. (2001) *Appl. Microbiol. Biotechn.* 56, 1–8
33. de Vocht, M. L., Scholtmeijer, K., van der Vege, E. W., de Vries, O. M., Sonneveld, N., Wosten, H. A., Ruyschaert, J. M., Hadziloannou, G., Wessels, J. G., and Robillard, G. T. (1998) *Biophys. J.* 74, 2059–2068
34. Nakari-Setälä, T., Aro, N., Kalkkinen, N., Alatalo, E., and Penttilä, M. (1996) *Eur. J. Biochem.* 235, 248–255
35. Lehn, J. M. (2002) *Science* 295, 2400–2403