

in bacteria, and some are implicated in bacterial virulence². For example, a life-threatening bacterium, *Salmonella typhimurium*, possesses the PhoP/PhoQ phosphorelay system which seems to be essential for its virulence in host organisms¹⁵. As all of these phosphorelay systems contain a conserved histidine kinase domain, which has not been found in mammalian cells, histidine kinases are excellent targets for antimicrobial action¹⁶. Members of a family of hydrophobic tyramines have already been identified as histidine kinase inhibitors; the half-maximal inhibitory concentration of these inhibitors is 2–600 μM (ref. 17). Our structure of the *E. coli* EnvZ histidine kinase catalytic domain will provide a vital foundation for the rational design of potent antibiotics that specifically inhibit multiple histidine kinases in this and other microbial species. □

Methods

Sample preparation. Recombinant uniformly ¹⁵N- and ¹³C-labelled EnvZ was expressed in overproducing *E. coli* strain pET11a/BL21(DE3) grown in M9 minimal medium, and was purified as described⁴. NMR samples contained 1.0–1.5 mM uniformly ¹⁵N- or ¹³C/¹⁵N-labelled, or unlabelled, protein in either 95% H₂O/5% ²H₂O or 99.996% ²H₂O containing 20 mM sodium phosphate, 50 mM KCl, 0.5 mM 4-(2-aminoethyl)-benzenesulphonyl fluoride, 50 μM sodium azide and 5 mM MgCl₂, pH 7.0, with 5 mM unlabelled or ¹⁵N/¹³C-labelled AMP-PNP.

NMR spectroscopy. All NMR spectra were acquired at 25 °C on Varian UNITY-plus 500, UNITY-600, and Bruker DMX750 spectrometers. ¹H, ¹³C, and ¹⁵N resonances of the backbone were assigned by analysing four triple-resonance experiments, (HB)/CBCA(CO)NNH, HNCACB, (HB)/CBCACO(CA)HA, and HNCO. Assignment of the sidechain resonances was done using the three-dimensional HCCH-TOCSY and ¹³C-edited NOESY-HMQC spectra. ³J_{NH α coupling constants were measured from ¹H/¹⁵N HMQC-J spectrum, and slow-exchanging amide protons were identified by recording a series of gradient-enhanced ¹H-¹⁵N HSQC spectra at different time points immediately after the H₂O buffer was changed to a ²H₂O buffer.}

Structure calculation. Structure calculations were done using a restrained molecular dynamics simulated annealing protocol¹⁸ within X-PLOR¹⁹. For structure calculations we used 1,782 interproton distance restraints (comprising 556 intraresidue, 440 sequential, 260 short-range, 507 long-range, 13 protein-ATP analogue, and 10 intra-analogue distances) obtained from heteronuclear three-dimensional NOE spectra. In addition to the NOE-derived distance restraints, 92 distance restraints for 46 hydrogen bonds and 122 dihedral angle restraints were included in the structure calculation. The average root mean square deviation (r.m.s.d.) values from idealized geometry for bonds, angles and improper are 0.005 Å, 0.61° and 0.39°, respectively. The total and Lennard-Jones potential energies are 516 ± 70 and -134 ± 29 kcal mol⁻¹, respectively (calculated with the use of square-well potentials for the experimental distance empirical energy term with a force constant of 50 kcal mol⁻¹ Å⁻²). None of the structures has violations greater than 0.40 Å (for distance restraints) and 3.0° (for dihedral angle restraints).

Ultraviolet-crosslinking assay. Purified proteins were each incubated in crosslinking buffer (20 mM Tris, pH 7.8, 50 mM KCl and 5% glycerol) containing [α -³²P]ATP (800 Ci mmol⁻¹, 10 mCi ml⁻¹) or [γ -³²P]ATP (3,000 Ci mmol⁻¹, 10 mCi ml⁻¹) in a final volume of 25 μl for 15 min at 4 °C. Proteins were crosslinked by ultraviolet irradiation at 254 nm at a height of 4.5 cm for 5 min on ice using an ultraviolet lamp (Model UVG-45, 115 V, 60 Hz, 0.16 A, UVP Inc., California). Crosslinked protein was visualized by autoradiography after analysis by 17.5% SDS-PAGE. For competition experiments, non-radioactive ATP was added in the reaction mixture.

Received 19 June; accepted 12 August 1998.

- Parkinson, J. S. in *Two-Component Signal Transduction* (eds Hock, J. A. & Silhavy, T. J.) 9–23 (ASM, Washington, DC, 1995).
- Egger, L. A., Park, H. & Inouye, M. Signal transduction via the histidyl-aspartyl phosphorelay. *Genes Cells* **2**, 167–184 (1997).
- Wirgler-Murphy, S. M. & Saito, H. Two-component signal transducers and MAPK cascades. *Trends Biochem. Sci.* **22**, 172–176 (1997).
- Park, H., Saha, S. K. & Inouye, M. Two-domain reconstitution of a functional protein histidine kinase. *Proc. Natl Acad. Sci. USA* **95**, 6728–6732 (1998).
- Branden, C. & Tooze, J. *Introduction to Protein Structure* (Garland, New York, 1991).
- Stock, J. B., Surette, M. G., Levit, M. & Park, P. in *Two-Component Signal Transduction* (eds Hock, J. A. & Silhavy, T. J.) 25–51 (ASM, Washington, DC, 1995).
- Yang, Y. & Inouye, M. Requirement of both kinase and phosphatase activities of an *Escherichia coli*

- receptor (Tazl) for ligand-dependent signal transduction. *J. Mol. Biol.* **231**, 335–342 (1993).
- Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
- Prodromou, C. et al. Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. *Cell* **90**, 65–75 (1997).
- Wigley, D. B. et al. Crystal structure of an N-terminal fragment of the DNA gyrase B protein. *Nature* **351**, 624–629 (1991).
- Dutta, R. & Inouye, M. Reverse phosphotransfer from OmpR to EnvZ in a kinase⁺/phosphatase⁻ mutant of EnvZ (EnvZ-N347D), a bifunctional signal transducer of *Escherichia coli*. *J. Biol. Chem.* **271**, 1424–1429 (1996).
- Min, K.-T. et al. Sigma F, the first compartment-specific transcription factor of *B. subtilis*, is regulated by anti-sigma factor that is also a protein kinase. *Cell* **77**, 735–742 (1993).
- Mizuno, T. Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. *DNA Res.* **4**, 161–168 (1997).
- Ota, I. M. & Varshavsky, A. A yeast protein similar to bacterial two-component regulators. *Science* **262**, 566–569 (1993).
- Soncini, F. C. & Groisman, E. A. Two-component regulatory systems can interact to process multiple environmental signals. *J. Bacteriol.* **178**, 6796–6801 (1996).
- Dziejman, M. & Mekalanos, J. J. in *Two-Component Signal Transduction* (eds Hock, J. A. & Silhavy, T. J.) 305–317 (ASM, Washington, DC, 1995).
- Barrett, J. F. et al. Antibacterial agents that inhibit two-component signal transduction systems. *Proc. Natl Acad. Sci. USA* **95**, 5317–5322 (1998).
- Nilges, M., Gronenborn, A. M., Brünger, A. T. & Clore, G. M. Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. *Protein Eng.* **2**, 27–38 (1988).
- Brünger, A. T. *X-PLOR Version 3.1: A system for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).
- Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
- Merritt, E. A. & Murphy, M. E. P. Raster 3D Version 2.0—a program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).
- Holdgate, C. A. et al. The entropic penalty of ordered water accounts for weaker binding of the antibiotic novobiocin to a resistant mutant of DNA gyrase: a thermodynamic and crystallographic study. *Biochemistry* **36**, 9663–9673 (1997).
- Orengo, C. A., Brown, N. P. & Taylor, W. R. Fast structure alignment for protein databank searching. *Proteins* **14**, 139–167 (1992).

Acknowledgements. We thank L. Kay for providing NMR pulse sequences, L. Pearl for the Hsp90-ATP coordinates before PDB release and discussions, and S. Bagby for comments on the manuscript. This work was supported by grants from JSPS (to T.T.) from CREST (to M.K.) from the NIH (to M. Inouye), and from HHMI (to M. Ikura). R.I. and D.L. acknowledge HFSP postdoctoral fellowships. M. Ikura is an HHMI International Research Scholar and MRCC Scientist.

Correspondence and requests for materials should be addressed to M. Ikura (e-mail: mikura@oci.utoronto.ca). Coordinates have been deposited in the Brookhaven Protein Data Bank.

Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus

Peter B. Rosenthal*†, Xiaodong Zhang*†, Frank Formanowski†‡, Wolfgang Fitz†§, Chi-Huey Wong§, Herbert Meier-Ewert‡, John J. Skehel|| & Don C. Wiley¶

* Department of Molecular and Cellular Biology, Harvard University, and † Howard Hughes Medical Institute, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

‡ Abteilung für Virologie, Technical University of Munich, Biedersteiner Strasse 29, 80802 Munich 40, Germany

§ Department of Chemistry, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037, USA

|| National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

The spike glycoproteins of the lipid-enveloped orthomyxoviruses and paramyxoviruses have three functions: to recognize the receptor on the cell surface, to mediate viral fusion with the cell membrane, and to destroy the receptor. In influenza C virus, a single glycoprotein, the haemagglutinin-esterase-fusion (HEF) protein, possesses all three functions (reviewed in ref. 1). In influenza A and B, the first two activities are mediated by haemagglutinin and the third by a second glycoprotein, neuraminidase. Here we report

† Present addresses: MRC Laboratory of Molecular Biology, Hills Road, Cambridge, England CB2 2QH, UK (P.B.R.); Molecular Structure and Function Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London, WC2A 3PX, UK (X.Z.); Cressier, Switzerland (E.F.). Quest International, Bussum, The Netherlands (W.F.).

the crystal structure of the HEF envelope glycoprotein of influenza C virus. We have identified the receptor-binding site and the receptor-destroying enzyme (9-*O*-acylesterase) sites, by using receptor analogues. The receptor-binding domain is structurally similar to the sialic acid-binding domain of influenza A haemagglutinin, but binds 9-*O*-acetylsialic acid. The esterase domain has a structure similar to the esterase from *Streptomyces scabies* and a brain acetylhydrolase^{2,3}. The receptor domain is inserted into a surface loop of the esterase domain and the esterase domain is inserted into a surface loop of the stem. The stem domain is similar to that of influenza A haemagglutinin, except that the triple-stranded, α -helical bundle diverges at both of its ends, and the amino terminus of HEF2, the fusion peptide, is partially exposed. The segregation of HEF's three functions into structurally distinct domains suggests that the entire stem region, including sequences at the amino and carboxy termini of HEF1 which precede the post-translational cleavage site between HEF1 and HEF2, forms an independent fusion domain which is probably derived from an ancestral membrane fusion protein.

The structure of the HEF trimer is shown in Fig. 1. The HEF monomer is composed of three domains: an elongated stem (red in Fig. 2a) active in membrane fusion (F), a receptor-destroying esterase domain (E) (green in Fig. 2a), and a receptor-binding domain (R) (blue in Fig. 2a). Two of these compact domains are made from non-contiguous segments of amino-acid sequence (Fig. 2b): the stem domain F consists of the amino-terminal amino acids 1–40 and carboxy-terminal residues 367–432 of HEF1 and all of HEF2 (labelled F1, F2, F3 in Fig. 2a, b); the esterase domain E consists of HEF1 segments comprising residues 41–150, which precede the receptor domain, and residues 311–366, which follow R (labelled E1, E' and E2 in Fig. 2a, b). The single-segment R domain is inserted into a surface loop of the esterase domain, and the esterase domain is inserted into a surface loop near the top of the stem domain F (Fig. 2c). The R and E domains of HEF are both compact, having their N and C termini within a few angströms of each other, so that they can be accommodated into a pre-existing

protein at surface loops without disrupting either protein's structure or function.

Alignment of the amino-acid sequences of HEF and HA based on their three-dimensional structures indicates that they have 12% sequence identity (the alignment is available from the authors). Nevertheless, both the overall structure (compare Fig. 2a and d) and the detailed folds of individual segments (Fig. 2a) of HEF and HA are remarkably similar. This is true for the globular R domain, as well as for the highly extended segments F1 and F2 (Fig. 2a), and for the similar helical-hairpin and membrane-proximal five-stranded β -sheet forming HEF2 and HA2 (F3 in Fig. 2a). Two of the segments of the enzyme domain, E1 and E2, are not found in HA (compare Fig. 2b and e), which is consistent with its lack of enzyme activity. If the E' domain of HA is a vestigial fragment of E, then HA may have evolved by the deletion of segments E1 and E2 from an ancestral gene similar to HEF. Segments E1, E', R, E2, and F2 are present, with 30% sequence identity, in the haemagglutinin-esterase (HE) found in coronaviruses⁴. (Data concerning the antigenicity of HEF and the structural relation between HEF and HE will be published elsewhere.)

The R domains of both HEF and HE are similar eight-stranded 'Swiss rolls' of β -sheet (Fig. 2a, d) containing the receptor-binding sites bounded by an α -helix, a loop, and an extended strand (superimposed in Fig. 3b) (except for one residue, Tyr 127 from E'). The structures of the complexes of HEF with two receptor analogues^{5,6}, 9-acetamidodialic acid α -methylglycoside (Fig. 3a) and 9-acetamidodialic acid α -thiomethylmercuryl glycoside⁷, show that sialosides bind similarly to HEF and to HA, although they are recognized by different amino-acid side chains (see ref. 8 and references therein). Sialic acid linkage specificity (α (2,3) vs α -(2,6)) has been attributed to residues near amino acid 226 in HA^{8–10}. The homologous loop (near HEF 270) is truncated in HEF (Fig. 2b), consistent with the lack of linkage specificity of influenza C virus¹¹. The HEF receptor differs from the HA receptor by the addition of an acetyl group at the 9-*O* position of the glycerol side chain. The acetyl methyl group binds in a nonpolar pocket unique to HEF (Fig. 3a).

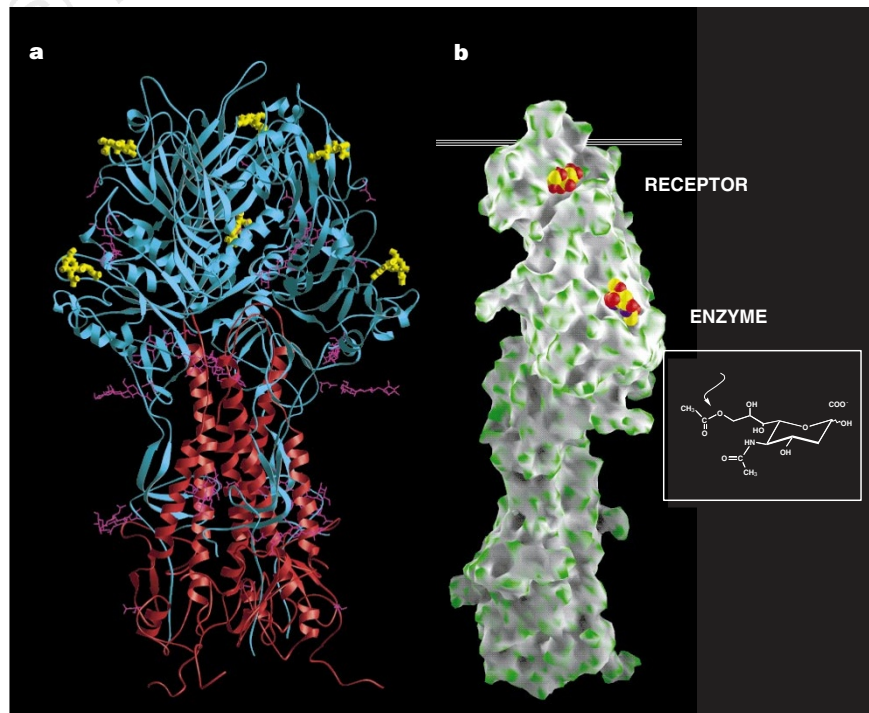


Figure 1 Haemagglutinin-esterase-fusion glycoprotein structure. **a**, The structure of the HEF trimer. HEF1 (blue), HEF2 (red), receptor analogue and enzyme inhibitor ligands, yellow; *N*-linked carbohydrate ball-and-stick (purple). HEF1 is linked to HEF2 by a disulfide bond from Cys 6 of HEF1 to Cys 137 of HEF2. **b**,

Monomer surface of HEF (Grasp²⁷) showing 9-*O*-acetylsialoside receptor binding site (top) and 9-*O*-acylesterase site (bottom). Inset, the esterase removes the acetyl group of 9-*O*-acetylsialic acid (see arrow).

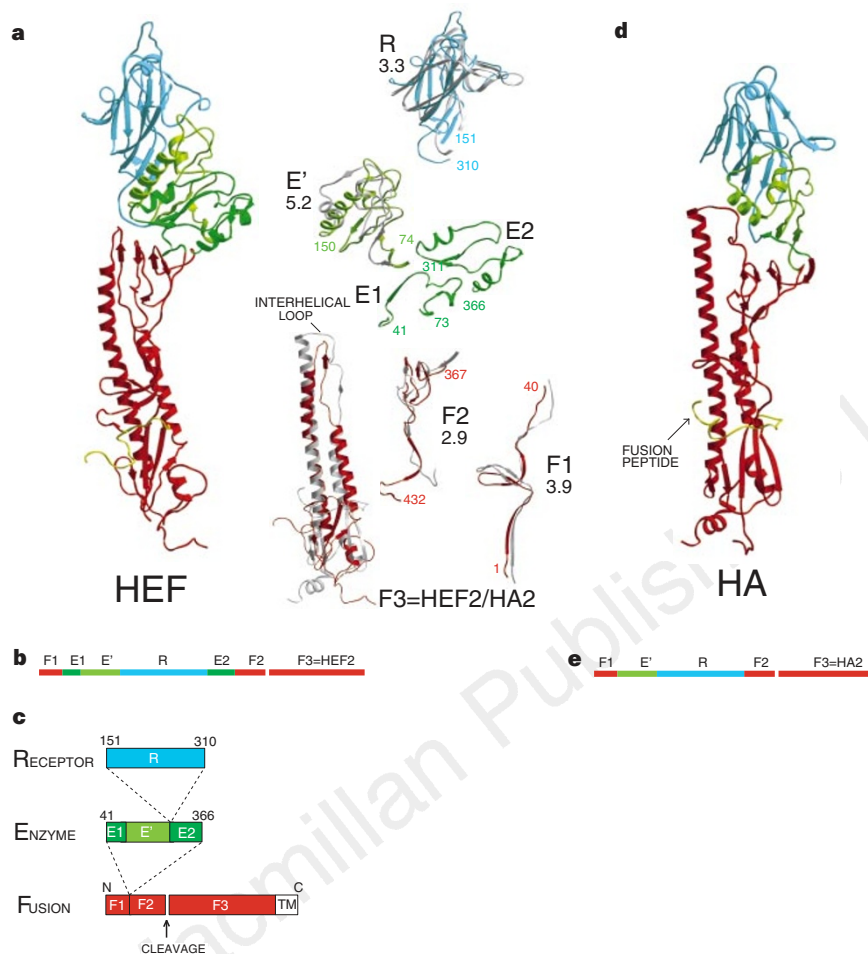


Figure 2 Comparison of HEF and HA monomers. **a**, HEF monomer structure and exploded views of the sequence segments (coloured by domain). Superposition of corresponding HA (X:31 strain; PDB code 1hge) segments (grey) are shown with r.m.s. values below the segment name. Fusion peptides are yellow. **b**, Linear order of the sequence segments in HEF, coloured by domains. Red segments: F1,

F2, F3; green: E1, E2; light green: E'; blue: R. **c**, Topological relationship of the compact domains in HEF. **d**, HA monomer structure with sequence segments coloured by domains. **e**, Linear order of the sequence segments in HA, coloured by domain. Red segments: F1, F2, F3; light green: E'; blue: R.

Complexes of HEF with the two non-hydrolysable receptor analogues described above also show how substrate interacts with the novel 9-*O*-acetyl esterase active site (Fig. 4a). Ser 57 in the catalytic triad (Ser 57 from E1, His 355 and Asp 352 from E2) is positioned for nucleophilic attack on the carbonyl carbon of the 9-*O* acetyl group of the bound sialoside. The carbonyl oxygen points into an 'oxyanion hole' formed by the side chain of Asn 117, and the NH groups of Gly 85 and Ser 57 (Fig. 4a). The ligand interactions and the structure of the enzyme site are completely different from those of the receptor-binding site.

A search for proteins with structural similarity to the HEF enzyme domain identified the $\alpha 1$ subunit of platelet-activating factor acetylhydrolase (Ib) from bovine brain (PAF-AH)² and the esterase from *Streptomyces scabies* (SsEst)¹ (Fig. 4b). All three proteins have a similar topology, despite sharing only 13% sequence identity, and their core residues superimpose (r.m.s. ~ 3 Å), including the central five-stranded β -sheet and long flanking helices (blue in Fig. 4b). When the core folds are superimposed, the catalytic triads (green in Fig. 4b) and oxyanion-hole residues of each protein overlap.

Interactions between the α -helices in the stem of the HEF trimer and the packing of the N-terminal fusion peptide of F3 have implications for low-pH-induced conformational change in HEF and the mechanism by which this induces membrane fusion at low pH. The F3 segments of HA and HEF are very similar in structure (Fig. 2a): each monomer has a central α -helix along the threefold axis and a smaller N-terminal helix packed antiparallel on the

outside and connected by an interhelical loop (Fig. 2a). In HEF2, although the central helices interact closely in the middle like HA, they diverge from the trimer axis at both ends (Fig. 5a). At the top, the interhelical loops interpose between the first five turns of the long helices (residues 80–97), where loop residues HEF2 Arg 69 and central helix residues HEF2 Glu 95 form salt bridges and contact an unidentified ion (possibly a sulphate) on the trimer axis. Although HEF2 has a sequence deletion of seven residues in the loop region, this difference would preserve the register of the heptads during the formation of an extended triple-stranded coil in the low-pH conformation, like that in HA¹². Unlike HA, however, the top third of the triple-stranded helical bundle must first make interactions on the trimer axis after the removal of the interhelical loop, before the N-terminal coiled-coil extension can form.

Three tryptophans (HEF2 116) form the last interaction on the trimer axis (Fig. 5a), below which the helices diverge as in HA. Unlike HA, in which residues 2 (Leu) and 3 (Phe) of the HA2 N-terminal 'fusion peptides' interact across the trimer axis, HEF2 residues Val 11 and Leu 12 are closest to the trimer axis but further penetration is blocked by tryptophans at position 116. Residues N-terminal to residue 10 fold back out to the surface of the protein, where aspartic acids at positions 5 and 6 of HEF2 can interact with Arg 29 and Lys 30 of HEF2 and Lys 4 of HEF1. Residues 1–4 of HEF2 appear to be disordered on the surface of the molecule. In HEF, the residues that are functionally analogous to the fusion peptide of HA, namely the buried residues whose exposure would convert a soluble

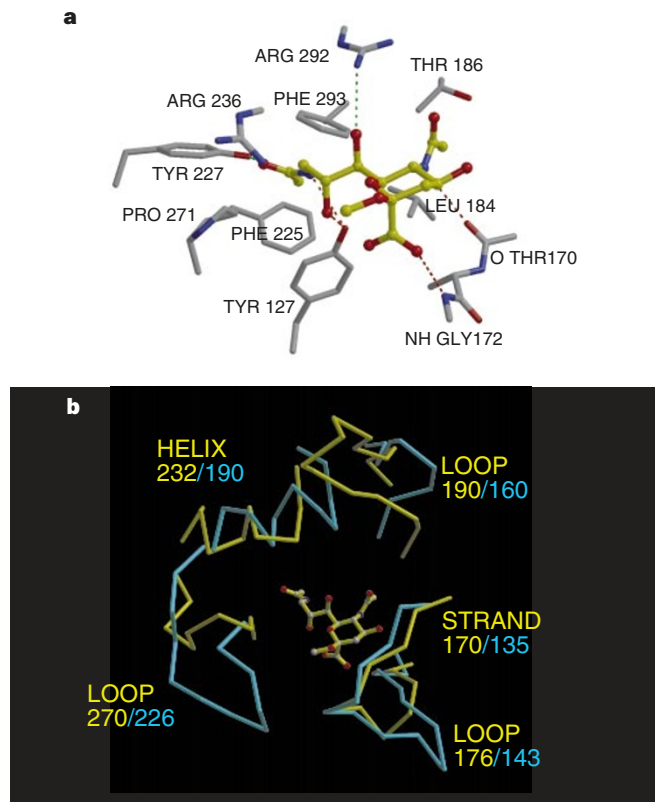


Figure 3 HEF receptor binding. **a**, Ligand bound to the receptor-binding site. Potential hydrogen bonds are indicated in green or red for those conserved in HA ligand binding. Four polar contacts are formed with the ligand identically in HEF and HA: two from the hydroxyl group of HEF1 Tyr127 (Y98 in HA1) to the 8-hydroxyl and 9-amide of the ligands and two from main-chain atoms: the carbonyl oxygen of HEF1 residue 170 (135 in HA1) to the 5-amide of the ligand and the amide of HEF1 172 (137 in HA1) to the carboxylate of the ligand. The acetyl methyl group binds in a nonpolar pocket unique to HEF, formed by Phe 225 and 293, and Pro 271; the acetyl carbonyl oxygen contacts the hydroxyl group of Tyr 224 and the guanidino group of Arg 236. **b**, Comparison of architecture of the HEF and HA binding sites. The figure was constructed by superimposing the common parts of the ligands in HA and HEF.

protein into a lipophilic one, are displaced along the sequence by six residues. HEF may therefore be regarded as having an internal fusion peptide, similar to virus fusion proteins that do not require cleavage activation. The conformation of the N terminus of the HEF2 fusion peptide may indicate that fusion peptides can insert into membranes as loops.

Because the influenza C virus esterase domain E is folded like other esterases (SsEst and PAF-AH), and the R domain present in HEF and HA is a ubiquitous folding unit also found as a receptor-binding domain in the orbivirus BTV¹³, we conclude that HEF must have evolved from functional domains (Fig. 2c). A precursor to HEF may have evolved by recombination events that resulted in the insertion of R into a surface loop of E, and E into the interhelical loop of the stem domain F (Fig. 2c). In such a scheme, the trimeric F domain (Fig. 5b) may have been an ancestral membrane-fusion protein analogous to the single-function fusion proteins of paramyxoviruses such as Sendai¹⁴. Similar modular structures of the envelope glycoproteins of retroviruses are suggested by biochemical data. The first 62 and the last 20 residues of gp120 from HIV-1 can be removed, retaining receptor binding of the fragment¹⁵, indicating that those terminal segments might be analogous to F1 and F2, forming with gp41 (F3) the stem of gp160 (refs 15–17).

The HEF structure implies that the membrane fusion domains of HEF and HA consist of F1 and F2, in addition to the segment F3 = HEF2, suggesting that F1 and F2 may play a part in membrane

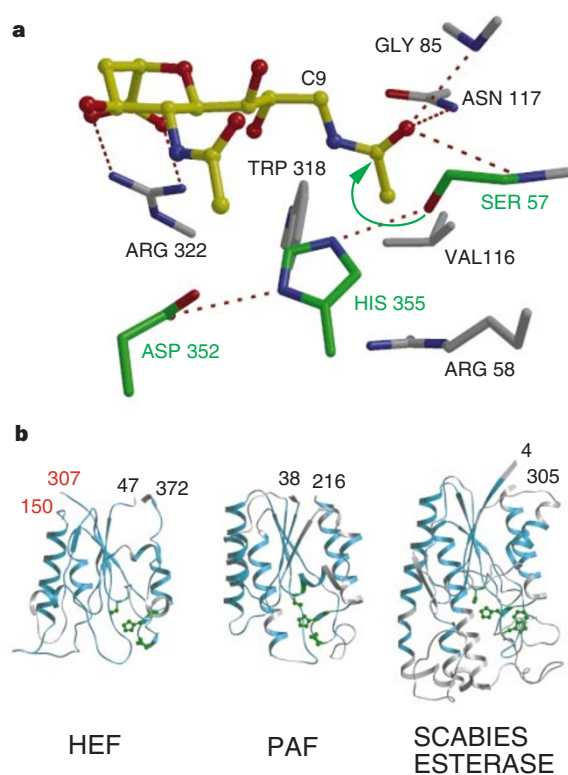


Figure 4 HEF enzyme active site. **a**, 9-Acetamidosalic acid α -methyl glycoside ($K_i = 2$ mM) or 9-acetamidosalic acid α -thiomethylmercury glycoside ($K_i = 4$ mM) shown in yellow. Catalytic triad is shown in green. Oxyanion-hole hydrogen bonds are on the right. Arg 322 forms two hydrogen bonds with the sialoside carboxylate group. **b**, Esterases with structural similarity to the HEF enzyme domain (E1, E', E2; Fig. 2a). PAF (platelet-activating factor acetylhydrolase Ib) (Z-score = 8.0, sequence identity 13%, 129 aligned residues) and SsEst (esterase from *Streptomyces scabies*) (Z-score = 7.9, sequence identity 10%, 146 aligned residues). Side chains for the catalytic triads are shown in green. N and C termini of the domains are numbered (black). The HEF receptor-binding domain (R) inserts at residues marked in red. Blue ribbon indicates similar structure identified by the Dali program²⁴. HEF and PAF-AH lack the extended Ω -loops of SsEst (bottom).

fusion, either by controlling the low-pH-induced conformational change required for fusion or during the formation of a fusion pore. In both HEF and HA, F2 packs against the interhelical loop of F3, which refolds to a helix at low pH, but the position of F2 after refolding is unknown. The β -hairpin of F1 (residues 15–32 of HEF1; Fig. 2a) has already been implicated in the low-pH-induced conformational change of HA by proteolytic susceptibility¹⁸, and by its location adjacent to the site of the helix-to- β -turn refolding and chain-direction reversal on the long HA2 helix¹². □

Methods

Structure determination. HEF was obtained by bromelain digestion of C/Johannesburg/1/66 virions and two crystal forms were characterized as described previously¹⁹. Crystals in harvest buffer (60% saturated ammonium sulphate, 50 mM Tris-HCl pH 7.1, 140 mM NaCl) were transferred to a ligand soak solution (40 mM 9-acetamidosalic acid α -thiomethylmercury glycoside, 300 mM MOPS pH 7.1, 140 mM NaCl) for 2 h and then cryoprotected by serial transfer through ligand soak solution containing 5–25% glycerol in 5% steps. Complexes of 9-acetamidosalic acid α -methyl glycoside ($K_i = 2$ mM) with HEF were prepared by soaking form I crystals in 40 mM ligand using the same procedure but with a slightly different soak solution (40 mM 9-acetamidosalic acid α -methyl glycoside, 100 mM Tris-HCl pH 7.1, 140 mM NaCl). 9-acetamidosalic acid α -methyl glycoside was a gift from J. Hanson.

Data collection and processing. X-ray diffraction data were collected at the Cornell High Energy Synchrotron Source by flash-cooling crystals and

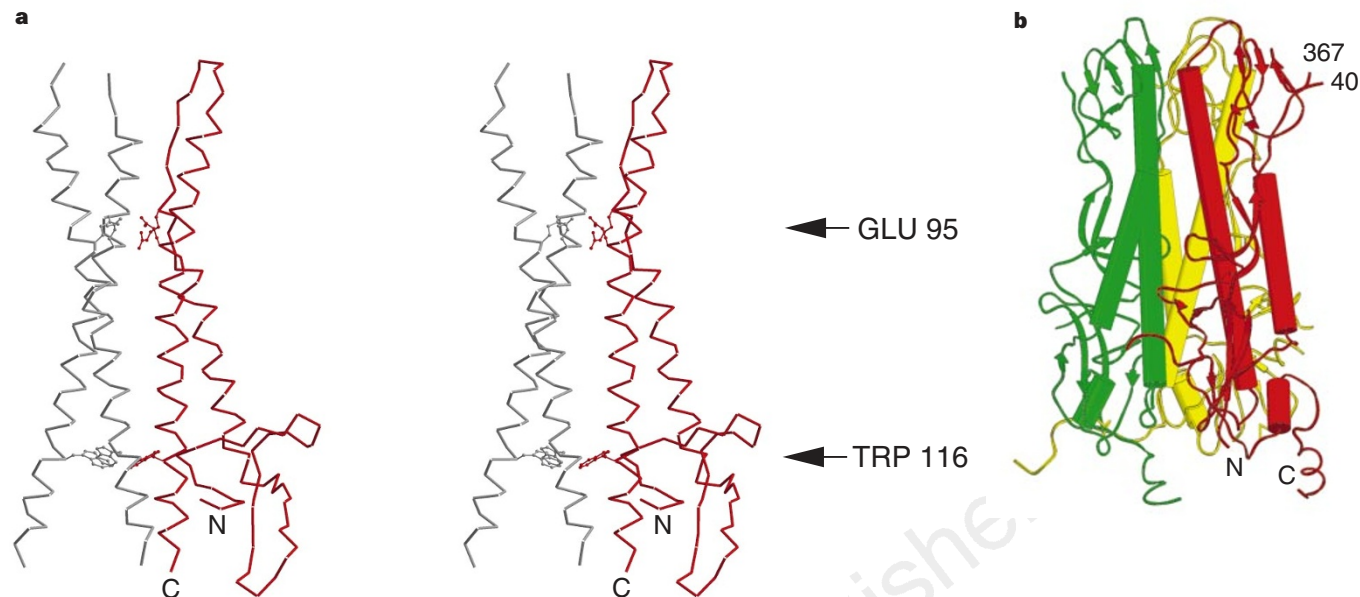


Figure 5 α -Helical interactions and the fusion peptide in HEF2. **a**, Stereo diagram of the triple-stranded α -helical bundle of HEF (79–126 of HEF2). Only the region 98–113 interacts across the trimer axis. The red monomer shows residues 4–126

(labelled N, C). **b**, Trimeric fusion domain of HEF consisting of segments F1, F2, and F3. N terminus of F1 and C terminus of F3 are indicated.

phosphorimage plate detection as described¹⁹. A 6.5 Å derivative dataset of form I crystals complexed with 9-acetamidiosialic acid α -thiomethylmercury glycoside collected on a Mar scanner and Elliot GX-13 rotating anode provided initial phases. Data were processed using Denzo, Scalepack²⁰, and programs from the CCP4 suite²¹.

Phasing, model building and refinement. 6.5 Å resolution SIR phases were extended to 3.5 Å resolution in form I crystals by iterative solvent flattening, histogram matching, and non-crystallographic symmetry averaging about the molecular three-fold axis using the program DM²². Details of the phase extension, model building, two-crystal-averaging, and refinement to 3.2 Å resolution will be described elsewhere (X.Z. *et al.*, manuscript submitted). The current model ($R_{\text{free}} = 26.7\%$, $R_{\text{work}} = 22.3\%$) contains HEF1 residues 1–427 (out of 432), residues 4–165 (out of 175) of HEF2, and no solvent molecules.

Core oligosaccharide (MAN-NAG-NAG) was built at 5 of 8 potential *N*-linked glycosylation sites (indicated in Fig. 2). No evidence for CHO was found at HEF1 Asn 117, occurring at the rarely glycosylated sequence NWSP. For the liganded HEF structures, the HEF model was subjected to rigid body and positional refinement against ligand data to 3.5 Å. Ligands were built into iteratively averaged $2F_o - F_c$ density phased with the HEF model, omitting residues within 5 Å of the binding sites.

Structure analysis. A Go plot²³ was used to help identify domains in HEF. The Dali program and database were used to find structurally similar domains²⁴. Least-squares superpositions were performed using the program Lsqman²⁵. RIBBONS²⁶ was used to produce Fig. 1a; GRASP²⁷ for Fig. 1b; BOBSCRIPT²⁸ and Raster3D²⁹ for Figs 3a and 4a; and SETOR³⁰ for Figs 2a, 3b, 4b and 5.

Received 2 July; accepted 4 September 1998.

- Herrler, G. & Klenk, H.-D. Structure and function of the HEF glycoprotein of influenza C virus. *Adv. Virus Res.* **40**, 213–234 (1991).
- Wei, Y. *et al.* A novel variant of the catalytic triad in the *Streptomyces scabies* esterase. *Nature Struct. Biol.* **2**, 218–223 (1995).
- Ho, Y. S. *et al.* Brain acetylcholinesterase that inactivates platelet-activating factor is a G-protein-like trimer. *Nature* **385**, 89–93 (1997).
- Luytjes, W., Bredenbeek, P. J., Noten, A. F., Horzinek, M. C. & Spaan, W. J. Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA recombination between coronaviruses and influenza C virus. *Virology* **166**, 415–422 (1988).
- Rosenthal, P. B. *Crystallographic Studies of the Influenza C Virus Surface Glycoprotein*. Thesis, Harvard Univ., 1996.
- Herrler, C. *et al.* A synthetic sialic acid analogue is recognized by influenza C virus as a receptor determinant but is resistant to the receptor-destroying enzyme. *J. Biol. Chem.* **267**, 12501–12505 (1992).
- Fitz, W., Rosenthal, P. B. & Wong, C.-H. Synthesis and inhibitory properties of a thiomethylmercuric sialic acid with application to the X-ray structure determination of 9-*O*-acetylsialic acid esterase from influenza C virus. *Biorg. Med. Chem.* **4**, 1349–1353 (1996).
- Eisen, M. B., Sabesan, S., Skehel, J. J. & Wiley, D. C. Binding of the influenza A virus to cell-surface

receptors: Structures of five hemagglutinin-sialyloligosaccharide complexes determined by X-ray crystallography. *Virology* **232**, 19–31 (1997).

- Weis, W. *et al.* Structure of the influenza haemagglutinin complexed with its receptor, sialic acid. *Nature* **333**, 426–431 (1988).
- Rogers, G. N. *et al.* Single amino-acid substitutions in influenza haemagglutinin change receptor binding specificity. *Nature* **304**, 76–78 (1983).
- Rogers, G. N., Herrler, G., Paulson, J. C. & Klenk, H.-D. Influenza C virus uses 9-*O*-acetyl-*N*-acetylneuraminic acid as a high affinity receptor determinant for attachment to cells. *J. Biol. Chem.* **261**, 5947–5951 (1986).
- Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37–43 (1994).
- Grimes, J., Basak, A. K., Roy, P. & Stuart, D. The crystal structure of bluetongue virus VP7. *Nature* **373**, 167–170 (1995).
- Scheid, A. & Choppin, P. Identification of biological activities of paramyxovirus glycoproteins. Activation of cell fusion, hemolysis and infectivity by proteolytic cleavage of an inactive precursor protein of Sendai virus. *Virology* **57**, 475–490 (1974).
- Pollard, S. R., Rosa, M. D., Rosa, J. J. & Wiley, D. C. Truncated variants of gp120 bind CD4 with high affinity and suggest a minimum CD4 binding region. *EMBO J.* **11**, 585–591 (1992).
- Wyatt, R. *et al.* Analysis of the interaction of the human immunodeficiency virus type 1 gp120 envelope glycoprotein with the gp41 transmembrane glycoprotein. *J. Virol.* **71**, 9722–9731 (1997).
- Helseth, E., Olshevsky, U., Furman, C. & Sodroski, J. Human immunodeficiency virus type 1 gp120 envelope glycoprotein regions important for association with the gp41 transmembrane glycoprotein. *J. Virol.* **65**, 2119–2123 (1991).
- Skehel, J. J. *et al.* Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proc. Natl. Acad. Sci. USA* **79**, 968–972 (1982).
- Rosenthal, P. B. *et al.* Crystallization and preliminary X-ray diffraction studies of the influenza C virus glycoprotein. *Acta Crystallogr. D* **52**, 1041–1045 (1996).
- Otwinowski, Z. & Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Collaborative Computational Project No. 4, the CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Cowan, K. D. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 34–38 (1994).
- Go, M. Correlation of DNA exonic regions with protein structural units in haemoglobin. *Nature* **291**, 90–92 (1981).
- Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).
- Kleywegt, G. J. & Jones, T. A. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 9–14 (1994).
- Carson, M. Ribbon models of macromolecules. *J. Mol. Graph. Modelg* **5**, 103–106 (1987).
- Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1991).
- Esnouf, R. M. An extensively modified version of Molscript which includes greatly enhanced coloring capabilities. *J. Mol. Graph. Modelg* **15**, 132–134 (1997).
- Merritt, E. A. & Bacon, D. J. Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **276**, 505–524 (1997).
- Evans, S. V. SETOR: Hardware-lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graph. Modelg* **11**, 134–138 (1993).

Acknowledgements. We thank M. Frayser and R. Crouse for technical assistance, and members of the Harrison-Wiley laboratory and the staff of the Cornell High Energy Synchrotron Source for assistance with data collection. P.B.R. received graduate research assistant support from the Department of Molecular and Cellular Biology and the committee on higher degrees in biophysics, Harvard University, as a student, and an NIH training grant as a postdoctoral fellow; X.Z. was supported by the HHMI. This work was supported by the NIH, the Deutsche Forschungsgemeinschaft, the MRC and the HHMI. D.C.W. is an investigator of the HHMI.

Correspondence and requests for materials should be addressed to D.C.W. Coordinates will be deposited in the Brookhaven Database and are available before release from <http://www.crystal.harvard.edu>.