Structure of influenza haemagglutinin at the pH of membrane fusion

Per A. Bullough*†, Frederick M. Hughson*, John J. Skehel‡ & Don C. Wiley*§

* Department of Biochemistry and Molecular Biology and § Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA
‡ National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Low pH induces a conformational change in the influenza virus haemagglutinin, which then mediates fusion of the viral and host cell membranes. The three-dimensional structure of a fragment of the haemagglutinin in this conformation reveals a major refolding of the secondary and tertiary structure of the molecule. The apolar fusion peptide moves at least 100 Å to one tip of the molecule. At the other end a helical segment unfolds, a subdomain relocates reversing the chain direction, and part of the structure becomes disordered.

INFECTION by enveloped viruses involves fusion of viral and cellular membranes with subsequent transfer of viral genetic material into the cell. The virus components that mediate fusion are membrane glycoproteins, among the best characterized of which is the influenza virus haemagglutinin (HA)[1]. The HA is also responsible for binding influenza viruses to their sialylated cell-surface receptors, following which bound virus is internalized by endocytosis. At the low pH of endosomes, between pH 5 and pH 6, the fusion potential of the HA is activated[2-4] in a process requiring structural changes in HA[5,6] (reviewed in ref. 1). We report here the results of crystallographic analyses of a soluble fragment from low-pH-treated HA which indicate that the fusion-pH-induced conformation is substantially different from the neutral pH conformation.

Native HA has a relative molecular mass of 220K and is a trimer of identical subunits, each of which comprises two glycopolypeptides linked by a single disulphide bond, HA_1 (328 residues) and HA_2 (221 residues). Each HA_2 chain is anchored at its carboxy terminus in the viral membrane. Soluble trimers can be released by treating virus with the protease bromelain, which cleaves each HA_2 chain once after residue 175 (refs 7, 8). The structure[9,10] of this trimer, BHA (Fig. 1*a*), reveals that the HA_2 chains are major components of a mainly α-helical stem domain which forms the centre of the molecule. The HA_1 chains also contribute to the stem structure but primarily form three membrane distal globular domains containing the receptor binding sites. This structure poses two difficulties in understanding the role of the HA in membrane fusion[9]. First, the "135 Å" length of the molecule seems to restrict membranes from any closer approach. Second, the conserved, hydrophobic sequence at the N terminus of HA_2, the so called 'fusion peptide'[11,12], is ~100 Å from the distal tip and ~35 Å from the viral membrane end of the molecule (Fig. 1*b*); a major conformational change would be required for this segment to approach either the virus or host cell membrane.

When incubated at the pH of fusion, BHA undergoes structural changes, two consequences of which are particularly important for investigations of the fusion-activated molecule. First, it aggregates through exposure of the fusion peptide[11,13,14]. Such aggregates are unsuitable for crystallographic studies. Second, it becomes highly susceptible to proteases. In particular, thermolysin cleaves near the fusion peptide and the removal of this apolar segment solubilizes the fusion-pH-induced protein aggregate[13]. The soluble trimeric fragment considered here, called $TBHA_2$, is prepared from BHA at pH 5.0 by successive digestion with trypsin and thermolysin[13,15,16]. Each monomer includes residues 38–175 of the HA_2 chain disulphide linked to residues 1–27 of the HA_1 chain.

Dramatic differences are observed between the BHA and $TBHA_2$ structures. Large changes occur at both ends of the HA_2 chain. The N terminus is displaced 100 Å, a movement which in the intact molecule could transport the fusion peptide 150 Å or more. This movement appears to be driven by recruitment of at least 36 additional residues to the viral membrane-distal end of the triple-stranded α-helical coiled coil of BHA. At the C-terminal end, seven residues in the middle of the long BHA α-helix refold to form a bend, allowing the remainder of the helix and three short β-strands packed against it to jackknife back and pack against the start of the helix, a reorientation of 180°. These two structural changes displace most of the residues against which the C-terminal region of the HA_2 chain, residues 140–175, was packed in native BHA, apparently extruding it into an extended and partially disordered structure.

Structure determination

The $TBHA_2$ structure was determined using crystals cooled to −170 °C in cryoprotectant buffer[16] (Table 1). Briefly, a multiple isomorphous replacement (MIR) map calculated to 5.0 Å resolution was subjected to iterative cycles of threefold non-crystallographic symmetry averaging and phase extension; single isomorphous replacement (SIR) phases from the platinum derivative were combined with extended phases at each resolution step to 3.5 Å. An initial model built into this MIR/SIR map was improved by further rounds of building, averaging, and refinement at progressively higher resolution, guided chiefly by improvement in the free R factor[17]. Details are given in Table 1.

The current model (Fig. 2) includes residues 12_1–16_1 and 40_2–153_2 of the first monomer, residues 11_1–16_1 and 40_2–162_2 of the second monomer, residues 10_1–17_1 and 40_2–162_2 of the third monomer, and 37 water molecules (numbering scheme is that for intact HA: HA_1 and HA_2 residues are distinguished by the appropriate subscript). The chain trace is clear throughout this modelled region, as is electron density for most (89%) of the side chains, as judged by simulated annealing omit maps. However, a significant fraction (23%) of the molecule, including from two to 22 residues at each of the twelve chain termini, has not been modelled for lack of connected electron density. These apparently disordered or partially ordered regions pose difficulties for further refinement. In at least one case (residues 154_2–162_2), residues that are well-ordered in two monomers are

† Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

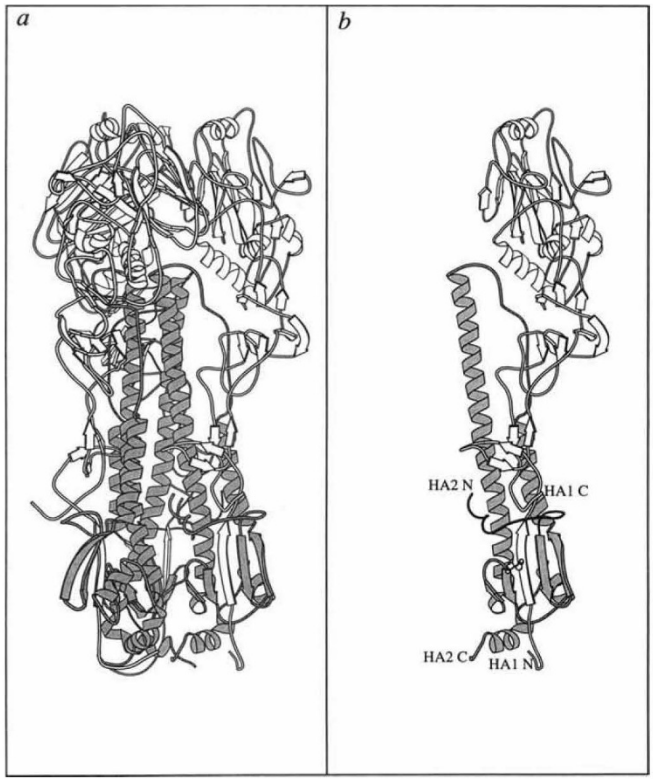

FIG. 1 BHA structure[10]. *a*, The BHA trimer and *b*, monomer are shown, with the HA_1 chains white and the HA_2 chains shaded. In *b*, the locations of the fusion peptide (black) and the N and C termini of the chains are indicated. Also shown is the interchain disulphide bond. The viral membrane end of the molecule is at the base, near the HA_2 C terminus generated by proteolytic removal of the transmembrane anchor sequence; the viral membrane distal globular domains are at the top. Figs 1, 2, 3*c* and *d* were generated using MOLSCRIPT[49]

prevented from adopting this conformation in the third monomer by a crystal contact, and indeed the electron density for these residues is uninterpretable.

The major feature of the $TBHA_2$ structure (Fig. 2) is the three-stranded α-helical coiled coil, approximately 100 Å long, comprising one long α-helix from each monomer (residues 40_2–105_2). This is followed in each monomer by a small connecting loop and a shorter helix (113_2–129_2) which packs in an antiparallel orientation against the long helix of the same monomer and that of another monomer. A β-hairpin (131_2–140_2), together with an additional strand contributed by HA_1 residues (11_1–16_1), forms a small antiparallel β-sheet. HA_1 is connected to HA_2 by a disulphide bond between residues 14_1 and 137_2. An additional short helix (146_2–153_2) packs against the coiled coil. Beyond residue 153_2 for one chain, and beyond residue 162_2 for the other two chains, the electron density is not interpretable, presumably indicating disorder.

BHA and $TBHA_2$ compared

Only thirty residues (76_2–105_2; C in Fig. 3*a*, *b*) have the same structure in BHA and $TBHA_2$ (root mean square difference, 0.6 Å). In both molecules, these residues form part of the central triple-stranded α-helical coiled coil. For the purpose of comparing the two structures, BHA and $TBHA_2$ were overlaid by aligning these residues.

Secondary structure changes are summarized in Fig. 3*a*. Four α-helical segments (A, C, D, G) and three β-strands (1, E, F) are conserved in both structures (Fig. 3). An irregular segment of BHA, the extended loop B between helices A and C, adopts a helical conformation in $TBHA_2$. Conversely, a segment in BHA in the middle of the long α-helix (106_2–112_2) refolds into a loop, allowing a 180° bend between helices C and D in $TBHA_2$. An α-helix, H, in BHA (159_2–170_2) and five residues beyond it, 171_2–175_2, appear to be disordered in $TBHA_2$. The overall helix content agrees well with that measured by circular dichroism[18] and, because of compensating changes, is very similar to that of the corresponding portions of the BHA structure.

TABLE 1 Crystallographic and refinement statistics

Data collection and MIR statistics

Parameter	Native	TAMM*	K_2PtCl_4
Resolution (Å)	15.0–2.5	15.0–5.0	15.0–3.2
Completeness ($I > \sigma(I)$)	0.86	0.78	0.92
R_{merge}†	0.065	0.083	0.088
R_{iso}‡	—	0.28	0.15
Number of sites	—	2	3
R_c§	—	0.59	0.60
Phasing power‖	—	0.97	1.16

Refinement statistics (6.0–2.50 Å)

Number of atoms¶	3,135	
Number of reflections (working set)	25,888	
R_{cryst}#	0.222	
Number of reflections (free set)	1,905	
R_{free}#	0.291	
Deviations from ideal geometry	0.015 Å	(bonds)
	2.17°	(angles)
Real space fit**	0.88 ± 0.09	(main chain)
	0.82 ± 0.12	(side chain)

Crystals in space group $C222_1$ ($a = 168.7$ Å $b = 231.7$ Å $c = 53.8$ Å) were grown as described[16]. Native X-ray data were collected on phosphor image plates at the Cornell High Energy Synchrotron Source (CHESS) from a single crystal in buffer plus 20% glycerol[16] and integrated using DENZO (Z. Otwinowski, personal communication). Derivative data were collected from two crystals soaked in buffer plus glycerol and 0.1 mM TAMM or 1.0 mM K_2PtCl_4 for one day. Data were recorded using an Elliot GX-13 X-ray source and either a Xentronics area detector (TAMM derivative) or a Mar Research Image Plate (K_2PtCl_4 derivative). Data were integrated with BUDDHA[38] (TAMM derivative) or MOSFLM[39] (K_2PtCl_4 derivative). All data were reduced and scaled using CCP4 programs[40]. One Pt site was located in a difference Patterson map; two additional Pt sites and two Hg sites were observed in difference Fourier maps and confirmed in difference Patterson maps. Heavy-atom sites were refined using HEAVY[41]. The three Pt sites defined the approximate position and orientation of one non-crystallographic molecular three-fold symmetry axis in each asymmetric unit, which was refined using CYLINDER[42]. Iterative averaging/solvent flattening[43] produced a map which was partially fitted with largely α-helical segments of polyalanine using the programme O[44]. Heavy atom positions were re-refined against model phases to generate an improved MIR map, calculated at 5 Å, which was extended to 3.5 Å by means of real-space averaging, phase extension, and phase combination (see text). Cycles of rebuilding, positional refinement with strict non-crystallographic symmetry (using XPLOR[45]), and averaging (using RAVE and associated programmes (G. J. Kleywegt and T. A. Jones, personal communication)) at progressively higher resolution were used to improve the model as judged by the free R factor[17]. Individual isotropic B factors were introduced and refined. When no further improvement could be obtained, the model was rebuilt into unaveraged maps. Cycles of building and refinement of the model without non-crystallographic symmetry constraints were performed (r.m.s. differences between monomer pairs (all atoms) are currently 1.9, 1.7 and 1.7 Å). Thirty-seven water molecules per trimer were added (4σ difference density, $B < 80$ Å^2, good geometry). We have not been able to extend the model into the disconnected density visible near some chain termini. Despite high B factors (Fig. 2 legend), the chain trace is clear throughout in simulated annealing omit maps, although one slight break in main chain density is seen (in the second monomer at the $Phe\,138_2$ nitrogen) and density for five side chains is absent. Neither three-dimensional profile analysis[46] nor analysis of nonbonded atomic interactions[47] identifies any significant problem with the current model, and less than 3% of the (ϕ, ψ) angles lie outside allowed regions in a Ramachandran plot. Refined coordinates will be deposited in the Protein Data Bank (Brookhaven National Laboratory, Upton, NY).

* TAMM, tetrakis(acetoxymercuri)methane.

† $R_{merge} = \sum\sum_j |I_j - \langle I \rangle| / \sum |\langle I \rangle|$ where I_j is the intensity measurement for reflection j and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

‡ $R_{iso} = \sum ||F_{ph}| - |F_p|| / \sum |F_p|$ where F_{ph} and F_p are the derivative and native structure factors respectively.

§ $R_c = \sum ||F_{ph} \pm F_p| - |F_{h,c}|| / \sum |F_{ph} \pm F_p|$ for centric reflections where $F_{h,c}$ is the calculated heavy-atom structure factor.

‖ Phasing power $= \langle F_h \rangle / E$ where $\langle F_h \rangle$ is the root-mean-square heavy-atom structure factor and E is the residual lack of closure error.

¶ Non-hydrogen atoms included in the current model.

$R_{cryst,free} = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$, where the crystallographic and free R factors are calculated using the working and free reflection sets respectively ($F_{obs} > 2\sigma(F_{obs})$). The free reflections (about 7% of the total) were chosen before averaging and phase extension and were held aside throughout refinement.

** Real-space fit[44] of the current model to a SIGMAA-weighted $2F_{obs} - F_{calc}$ map[48] using phases calculated from the current model.

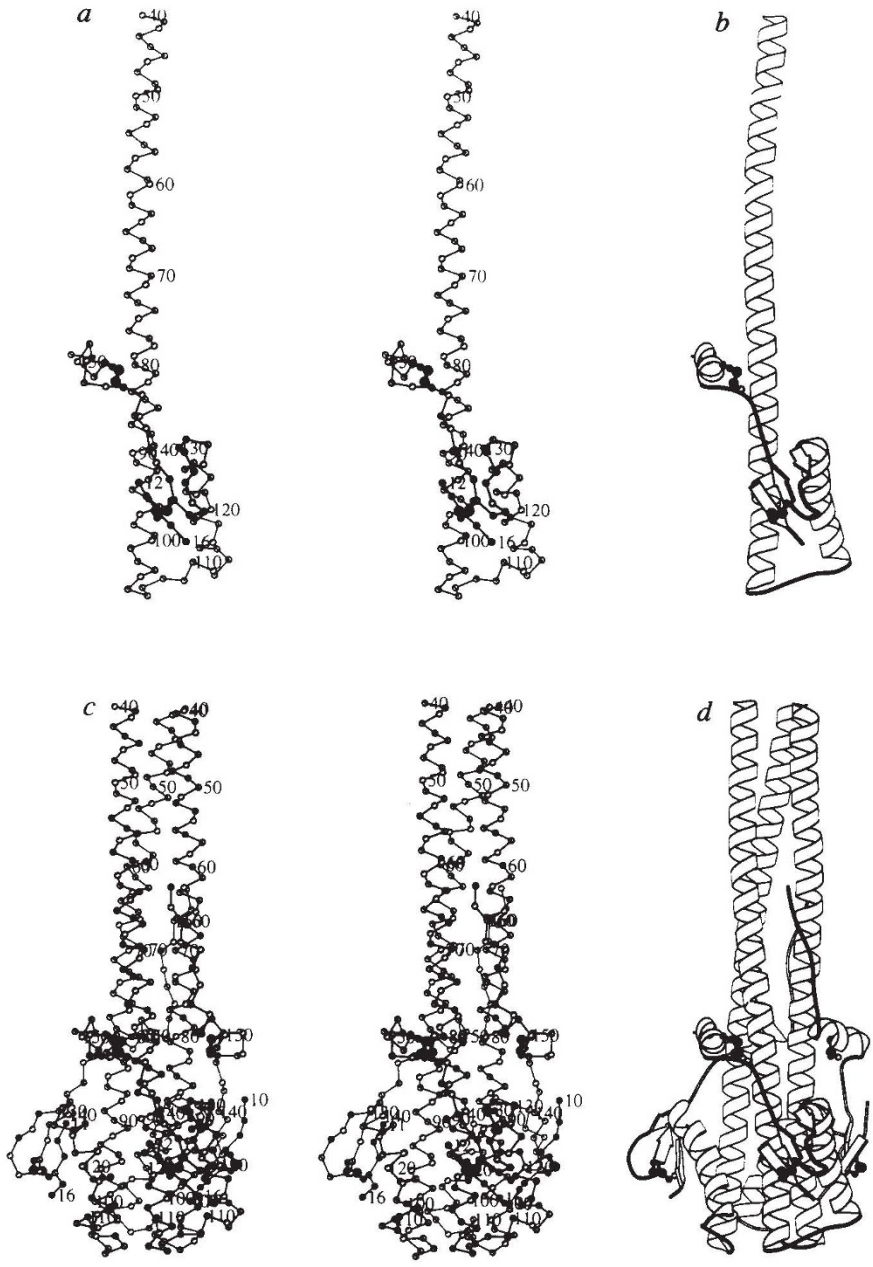

FIG. 2 $TBHA_2$ structure. *a*, Stereo image and *b*, ribbon diagram of the monomer, showing residues 12–16 of HA_1 and residues 40–153 of HA_2. *c*, Stereo image and, *d*, ribbon drawing of the trimer. The first monomer is as in *a* and *b*; the second contains HA_1 residues 11–16 and HA_2 residues 40–162, and the third contains HA_1 residues 10–17 and HA_2 residues 40–162. In *a* and *c*, HA_1 is shown with filled balls and HA_2 with open balls. Side chains are shown for cysteine residues, all of which participate in disulphide bonds. The mean *B* factor for all model atoms is 60 $Å^2$, in agreement with the *B* factor calculated from a Wilson plot (58 $Å^2$). Two short regions, 130_2–136_2 and 150_2–160_2, show particular evidence of disorder, with mean $B > 75$ $Å^2$ for each residue.

Tertiary structure refolding

N-terminal. Major refolding of secondary structure elements occurs at both ends of the $TBHA_2$ molecule. Helix A (38_2–55_2), which packs against the central helix CD in BHA, becomes the top of the triple-stranded α-helical coiled coil at the central core of $TBHA_2$ (Fig. 3*b*–*d*). Although the N-terminal residues (1_2–37_2) have been removed in the preparation of $TBHA_2$, the location of helix A suggests that the fusion peptide (1_2–23_2) and two β-strands (24_2–37_2) would also relocate to this end of the intact molecule, a distance of 100 Å or more. Loop B of BHA, which was docked against the C region of the long helix and made extensive contacts with HA_1, adopts an α-helical conformation and joins the triple-stranded coiled coil in $TBHA_2$ (Fig. 3*b*–*d*). This transition was anticipated[19] by the observation that residues 38_2–125_2 contain heptad repeats of hydrophobic residues which could potentially form a coiled coil[19,20]. It has been shown[19] that peptides corresponding to B and the first four turns of C, with or without A, form α-helical trimers, demonstrating the energetic preference of this segment to be a coiled coil.

C-terminal. Nearer the C-terminal end of the molecule, a portion of helix CD (106_2–112_2) that made a number of contacts with the fusion peptide in BHA[9,10] refolds to form a loop (Fig. 3*b*–*d*). In BHA, the bottom half (D) of the long α-helix diverges from the central axis of the molecule to form a tripod-like structure apparently stabilized by polar and nonpolar interactions with the fusion peptide. Below residue 112_2, the residues on the inner face of the long helices which surround the trimer axis are exclusively polar and/or charged. The relocation of the fusion peptide from the middle region of the CD helix may destabilize this tripod arrangement. Hence, although helix D (113_2–126_2) remains intact, it folds back together with the small antiparallel β-sheet (1, E, F in Fig. 3*b*) to pack against the triple-stranded coiled coil in an inverted orientation (Fig. 3*b*–*d*). There are many small changes in the architecture of this subdomain (1, D, E, F in Fig. 3*b*; blue in Figs 3*c*, *d*); its root-mean-square Cα coordinate difference versus BHA is over 3 Å. An extensive new hydrophobic core, which presumably stabilizes the fusion-pH-induced conformation, is formed as a result of this subdomain movement (Fig. 4*a*). The single Trp in each monomer (residue 92_2) is

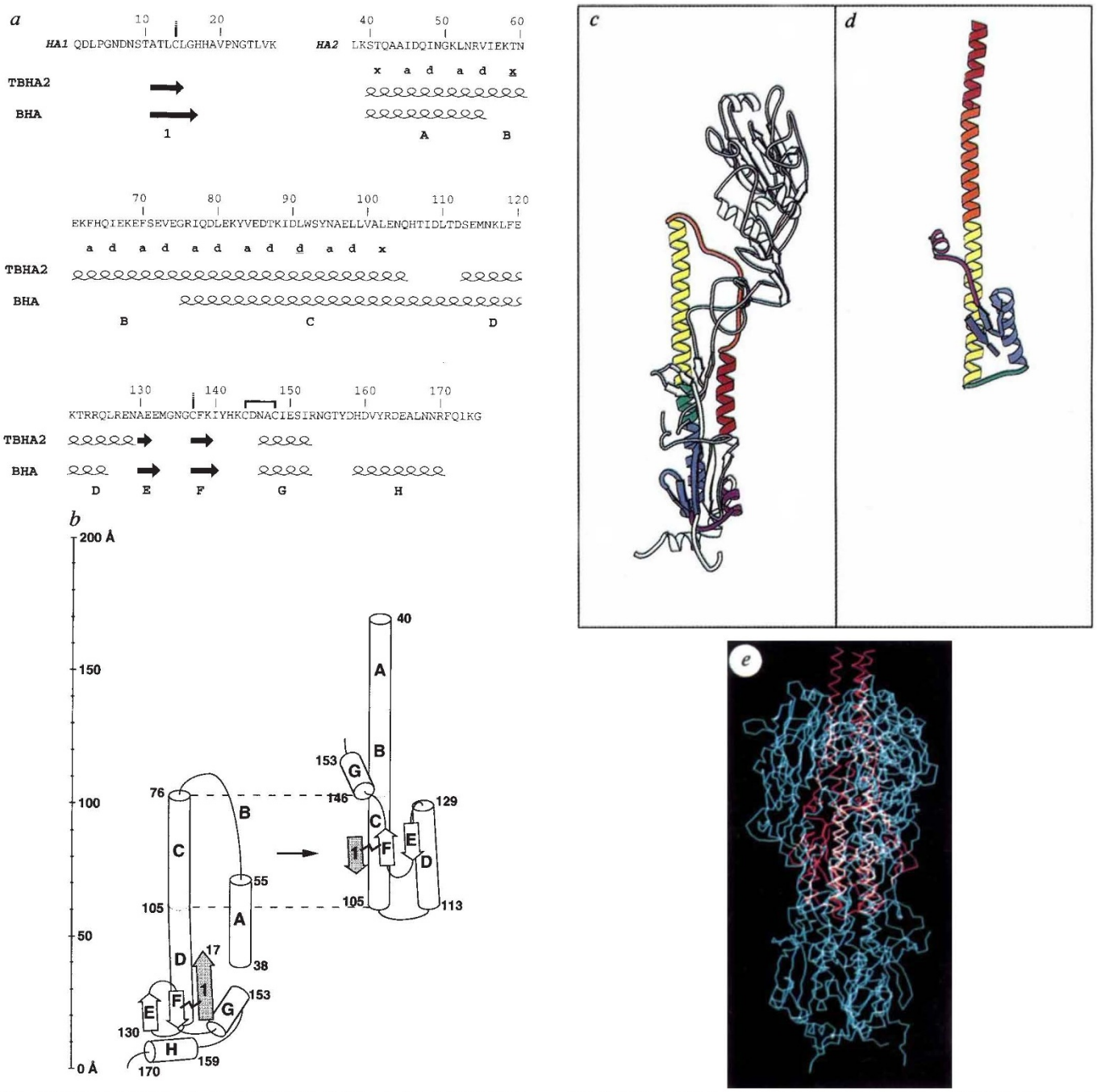

FIG. 3 The low-pH-induced conformational change. *a*, Secondary structure for $TBHA_2$ and the corresponding portion of BHA. α-helices and β-strands (arrows) are indicated below the amino-acid sequence. Letters A–H correspond to those in *b*. Core residue positions (a, d or x) for the $TBHA_2$ triple-stranded α-helical coiled coil are also shown (see text and Fig. 5); shifts in heptad register are centred on the underlined positions, as explained in the legend to Fig. 5. *b*, The structure of a $TBHA_2$ monomer (right) in schematic form is compared to the corresponding region of a BHA monomer (left). Consecutive regions of the HA_2 chain are labelled A–H (see text); H is apparently disordered in $TBHA_2$. The first β-strand of HA_1 is also shown, labelled '1'; both ends of this strand are apparently disordered in $TBHA_2$. The disulphide bond linking 14_1 and 137_2 is indicated. The two structures are aligned on the C region, which is unaffected by the conformational change. *c*, *d*, The structure of a $TBHA_2$ monomer (*d*) is compared to that of an intact BHA monomer (*c*). Structural elements of the HA_2 chain of $TBHA_2$ are coloured in rainbow order from amino to carboxy terminus; corresponding regions of BHA are coloured identically. In addition, the HA_1 β-strand which is disulphide-linked to HA_2 is blue; the blue elements (1, D, E, F in *b*) move approximately as a subdomain in the conformation change (see text). Regions of BHA which were proteolytically removed to generate $TBHA_2$ are shown in grey. Regions of BHA corresponding to apparently disordered regions of $TBHA_2$ are shown in white. *e*, $TBHA_2$ (red) overlaid on BHA (blue). The two structures are aligned using the C region of the coiled coil (*a*, *b*). Figs 3e and 5 were generated using the program O[44].

buried, consistent with fluorescence measurements of $TBHA_2$ in solution[18].

Residues 141_2–175_2, including helices G and H, which in BHA form a compact unit adjacent to the five-stranded β-sheet and part of helix D, adopt in $TBHA_2$ a more extended, partially disordered conformation running antiparallel to the coiled coil (Fig. 3*b*–*d*). This region (yellow in Fig. 4*b*) docks against residues (other colours in Fig. 4*b*) in BHA, many of which become buried, move to distant locations, or rearrange relative to one another in $TBHA_2$ (Fig. 4*b* legend). The destruction of the site which presumably had stabilized the 141_2–175_2 unit may account for the partial disordering of this

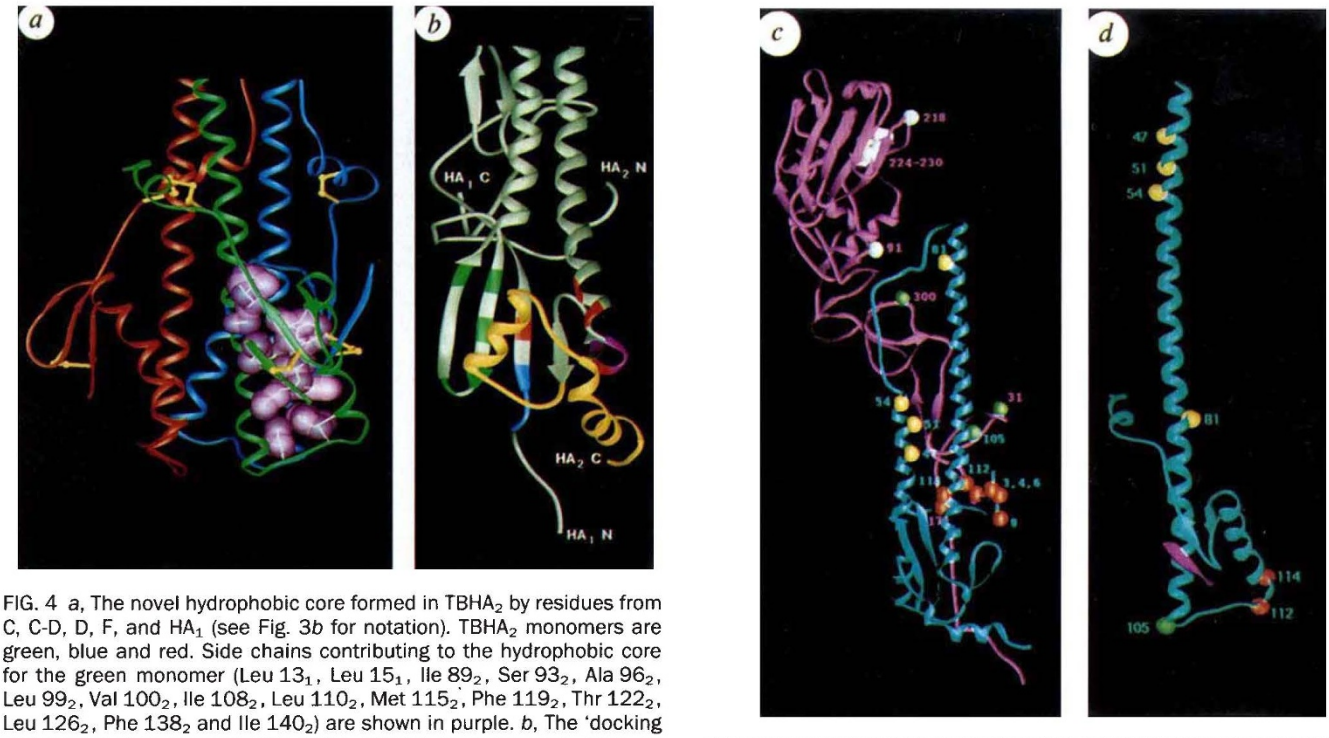

FIG. 4 *a*, The novel hydrophobic core formed in $TBHA_2$ by residues from C, C-D, D, F, and HA_1 (see Fig. 3*b* for notation). $TBHA_2$ monomers are green, blue and red. Side chains contributing to the hydrophobic core for the green monomer (Leu 13_1, Leu 15_1, Ile 89_2, Ser 93_2, Ala 96_2, Leu 99_2, Val 100_2, Ile 108_2, Leu 110_2, Met 115_2, Phe 119_2, Thr 122_2, Leu 126_2, Phe 138_2 and Ile 140_2) are shown in purple. *b*, The 'docking site' for the C-terminal portion of HA_2 in BHA. Residues 141_2–175_2 in BHA, including the G and H helices (Fig. 3*b*), are shown in yellow. Residues that make contacts with this region are green, red, magenta or blue; the remainder of the BHA monomer is grey. Green residues, in the N-terminal strands of HA_2, presumably move at least 100 Å as a result of the distal extension of the coiled coil at low pH. Red residues are largely buried in $TBHA_2$, inaccessible for interaction with the C-terminal region of HA_2. Magenta residues, although largely surface-exposed in $TBHA_2$, rearrange relative to one another. Blue residues are apparently disordered in at least one $TBHA_2$ monomer. *c*, *d*, Ribbon diagrams of BHA (*c*) and $TBHA_2$ (*d*) monomers showing the positions of mutations that raise the midpoint of the low-pH-induced conformational change[25,31]. HA_1 (magenta) and HA_2 (blue) chains are shown. Mutations are grouped into four classes: those in or near the fusion peptide (red), those affecting interactions between the long 'CD' helix and the outer helix 'A' or loop 'B' (yellow), those affecting HA_1–HA_2 contacts (green), and those affecting HA_1–HA_1 contacts (white). Most mutations were identified by selecting for mutant viruses able to grow in cells treated with amantidine hydrochloride, which at high concentrations raises the pH of the endosome. In all cases, the pH dependence of the conformational change and of fusion activity are very similar[25,31]. Images were generated using RIBBONS[50].

region in $THBA_2$. Helix G (146_2–153_2) remains intact in $TBHA_2$, probably stabilized in part by the disulphide 144_2–148_2, but it docks against the C region of the long helix, approximately where the B loop had docked in BHA (Fig. 3*b*–*d*). Residues beyond 153_2 in one subunit and 162_2 in the others, including helix H, were not located in our electron density maps, presumably because they are disordered in $TBHA_2$.

HA_1–HA_2 interactions. In $TBHA_2$, HA_2 has refolded so extensively that many of the HA_1–HA_2 interactions seen in BHA would no longer be possible even if the bulk of the HA_1 chain were still present. In BHA, loop B (orange in Fig. 3*c*) makes all of the HA_1–HA_2 contacts involving the globular head domain of HA_1, and many involving the descending C-terminal HA_1 strand. Having refolded into a helical conformation in $TBHA_2$, the B region cannot participate in these interactions. Indeed, chemical crosslinking, antibody binding, proteolysis, electron microscopy, mutation, and site-specific modification data all indicate that the HA_1 globular domains in BHA or in the intact HA are displaced in the fusion pH conformational change[5,6,13–15,21–25]. A second site (the C-D loop; green in Fig. 3*c*, *d*) which refolds significantly in $TBHA_2$ was the docking site for a loop (residues 24_1–37_1) in the ascending N-terminal HA_1 strand in BHA. Site-specific antipeptide antisera binding and proteolytic susceptibility at residue 27_1 indicate that the environment of this HA_1 loop rearranges in the intact protein at the pH of fusion[5,11].

The observed irreversibility of the fusion-pH-induced conformational transition[11,26] and the increased thermostability of $TBHA_2$ compared to BHA even at neutral pH[27] suggest that the low-pH-induced structure is more thermodynamically stable than the neutral pH structure, that is, that the neutral pH structure is metastable[19,27]. Kinetically controlled activity and conformational refolding have also been observed in the serpin family of protease inhibitors (reviewed in ref. 28).

$TBHA_2$ coiled coil

The triple-stranded coiled coil of $TBHA_2$ displays, with few exceptions, the knobs-into-holes packing predicted by Crick[29] (Fig. 5*a*, *b*). The residues that compose the core are mostly hydrophobic and recur, in general, with a heptad (3–4) periodicity, where heptad positions are denoted a–g (Fig. 3*a*). At two positions along the coil an unusual 3–4–4–3 periodicity is observed (Figs 3*a*, 5*c*), as predicted in part by Ward and Dopheide[20] but not by others[19], who assumed a single unbroken register. A small number of single-residue insertions or 'skips' are also observed in the heptad repeat sequences of proteins, such as myosin, which contain long double-stranded coiled coils[30]. In $TBHA_2$, the net result of these shifts is to underwind the coiled coil to an average pitch of ~300–400 Å.

pH mutants

Mutant viruses were selected by Daniels *et al.* that undergo the conformational change and mediate fusion at elevated pH[25,31]. The distribution of the mutations throughout the HA molecule, coupled with their effects on the pH of fusion, provide support both for the structural transition proposed here and for its functional significance in promoting membrane fusion. Specifically, these mutations can be seen to identify interactions which are modified in the transition state to the fusion-active structure. They have been placed into four groups (Fig. 4*c*, *d*). The first group (red in Fig. 4*c*, *d*) includes residues that appear to stabilize

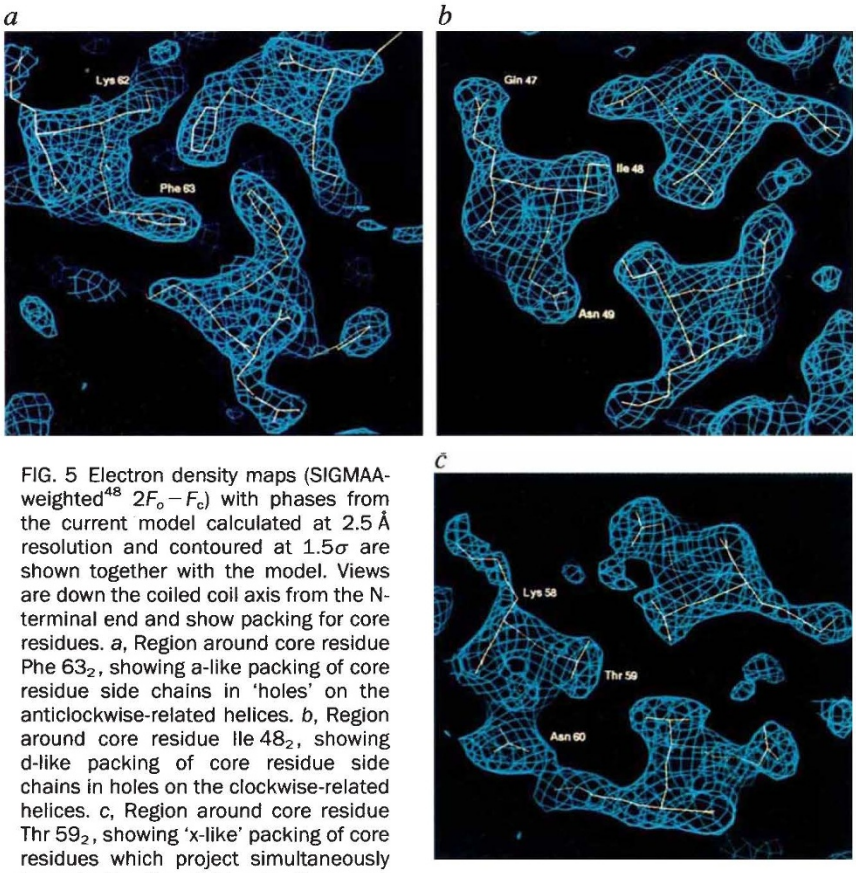

FIG. 5 Electron density maps (SIGMAA-weighted[48] $2F_o - F_c$) with phases from the current model calculated at 2.5 Å resolution and contoured at 1.5σ are shown together with the model. Views are down the coiled coil axis from the N-terminal end and show packing for core residues. *a*, Region around core residue Phe 63_2, showing a-like packing of core residue side chains in 'holes' on the anticlockwise-related helices. *b*, Region around core residue Ile 48_2, showing d-like packing of core residue side chains in holes on the clockwise-related helices. *c*, Region around core residue Thr 59_2, showing 'x-like' packing of core residues which project simultaneously towards the threefold axis. Core residues are unambiguously a-like or d-like except in three places (Fig. 3a). Two x-like residues are seen near the ends of the coiled coil. The third, shown here, lies in the midst of an otherwise unbroken 3–4 repeat extending from residue 45_2 to 87_2; the spacing for core residues surrounding Thr 59_2 is 3-4-4-3 (Fig. 3a). Thus, it serves as the crossover point between two different but related heptad registers, thereby maximizing the hydrophobicity of the core residues. Another register shift occurs at Leu 91_2, a d-like residue.

the buried location of the N-terminal fusion peptide. The fusion peptide, then, is apparently partially or fully released in the transition to the fusion-active state. A second group (yellow) identifies residues that appear to stabilize the short helix A and the loop B against the long helix CD, suggesting that the transition involves rearrangements of helix A and loop B, perhaps as the residues preceding C are recruited into the coiled coil. The third group (green) are residues that appear to stabilize contacts between HA_1 and B and CD of HA_2, consistent with the proposal that these contacts break as the B loop refolds into a helical conformation and the middle of helix CD refolds into a loop (Fig. 3).

Mutations in the interface between adjacent HA_1 subunits form a fourth group (white, Fig. 4*c*). No new insights about this group are available directly from the $TBHA_2$ structure, from which HA_1 residues 28_1–328_1 have been proteolytically removed. As previously discussed[6,31], these mutations suggest that the trimeric assembly of HA_1 globular domains rearranges or dissociates at low pH. Movement of the globular domains is probably accompanied by the refolding of the B loop into a helical conformation, as the B loop is the site against which HA_1 is docked in BHA. The HA_1 globular domains themselves apparently retain a native-like structure as they interact with many anti-BHA antibodies[22,23,26], bind sialic acid receptor analogues with the same affinity as BHA[32], and display an appropriate circular dichroism spectrum[18].

Electron microscopy

The dimensions and shape of $TBHA_2$ revealed by the X-ray structure agree well with electron micrographs (EMs) of the molecule. In particular, EM images show molecules 105 ± 7 Å in length and many display a knob at one end, presumably corresponding to the bulky end seen in the X-ray structure[15] (bottom, Fig. 2*d*). In addition, EM images of an earlier soluble thermolysin digestion product, identical to $TBHA_2$ except that the predominant HA_2 N terminus is residue 24_2 rather than 38_2, show molecules about 130 Å long, again with terminal knobs[15]. This observation is consistent with the idea that residues 24_2–37_2 extend from the tip of the coiled coil, possibly as a continuation of the α-helical coiled coil (~21 Å length) or in the β-hairpin conformation they adopt in BHA (~22 Å length).

Trypsin-digested low-pH BHA, identical to $TBHA_2$ except that it retains the first 37 residues of the HA_2 chains including the fusion peptides, forms 'rosettes' of 8 ± 1 molecules with distal knobs[15]. Again, these knobs are presumably the bulky 'bottom' region seen in $TBHA_2$, whereas the fusion peptides mediate aggregation at the centre of the rosette. Virus particles or HA-containing liposomes incubated at pH 5 and then digested with trypsin reveal thin spikes projecting 105 ± 10 Å from the virus membrane or liposome surface[24]. These spikes have membrane-distal knobs. One model which explains this observation is that both the fusion peptides and the C-terminal transmembrane segments are inserted into the virus or liposomal membrane, causing inversion of the main body of the trimer relative to the membrane. This model is possible given the structure of $TBHA_2$, inasmuch as the apparently disordered C-terminal region of the HA_2 chains could, in extended conformations, span the distance to the membrane whatever the overall orientation of the trimer. Moreover, it provides a possible explanation for the observation that virus becomes inactivated for fusion when incubated at fusion pH in the absence of target membrane[33]: namely, that the fusion peptides insert non-productively into the viral membrane.

The removal of 28_1–328_1 and both termini of HA_2 in producing $TBHA_2$ might have removed residues which stabilize a different conformation, or might have generated additional degrees of freedom unavailable in the intact molecule. While the $TBHA_2$ structure is consistent with the mutagenic analysis and EM data discussed above, further experiments are required to assess the possibility that, for example, disorder at the chain termini is a result of proteolytic cleavage.

Implications for membrane fusion

The $TBHA_2$ structure suggests that the fusion-pH-induced conformational change delivers the fusion peptides at least 100 Å towards the target membrane (Fig. 3*e*), consistent with the expectation that HA might form a bridge between the viral and cellular membranes[1,11,19,33]. We cannot rule out, however, the possibility that the main body of the molecule inverts to place the fusion peptide in the viral membrane[15]. Such inverted molecules might also be observed in the post-fusion state, wherein the transmembrane anchor segments and the fusion peptides may be associated with the same (fused) membrane[34]. Thus, the low-pH-induced conformation observed here could, in one orientation, mediate membrane fusion and, in another, lead to inactivation.

The increased flexibility inferred from the apparent disorder in the C-terminal region of $TBHA_2$ suggests that the extra-

membrane domain of HA at the pH of fusion might be able to bend with respect to the plane of the membrane. Such bending could allow the opposed membranes to approach more closely than the length of $TBHA_2$. In addition, increased orientational flexibility of the HA extramembrane domain might facilitate higher-order assembly of trimers, as proposed to be necessary for HA-mediated fusion[35], perhaps in the formation of a fusion pore structure[36]. Studies to date have not ruled out the possibility that other parts of the polypeptide, such as the membrane-proximal region inferred to be disordered from the $TBHA_2$ structure, are directly active in membrane fusion, as suggested for the C-terminal anchor region[37]. □

Received 31 May; accepted 13 July 1994.

1. Wiley, D. C. & Skehel, J. J. *A. Rev. Biochem.* **56,** 365–394 (1987).
2. Maeda, T. & Ohnishi, S. *FEBS Lett.* **122,** 283–287 (1980).
3. Huang, R. T. C., Rott, R. & Klenk, H.-D. *Virology* **110,** 243–247 (1981).
4. White, J., Matlin, K. & Helenius, A. *J. Cell Biol.* **89,** 674–679 (1982).
5. White, J. M. & Wilson, I. A. *J. Cell Biol.* **105,** 2887–2896 (1987).
6. Godley, L. *et al. Cell* **68,** 635–645 (1992).
7. Brand, C. M. & Skehel, J. J. *Nature new Biol.* **238,** 145–147 (1972).
8. Skehel, J. J. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **72,** 93–97 (1975).
9. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289,** 366–373 (1981).
10. Weis, W. I., Brünger, A. T., Skehel, J. J. & Wiley, D. C. *J. molec. Biol.* **212,** 737–761 (1990).
11. Skehel, J. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **79,** 968–972 (1982).
12. Gething, M.-J., Doms, R. W., York, D. & White, J. *J. Cell Biol.* **102,** 11–23 (1986).
13. Daniels, R. S. *et al.* in *The Origin of Pandemic Influenza Viruses* (eds Laver, G. W. & Chu, C. M.) 1–7 (Elsevier–North Holland, New York, 1985).
14. Doms, R. W., Helenius, A. & White, J. *J. biol. Chem.* **260,** 2973–2981 (1985).
15. Ruigrok, R. W. H. *et al. J. gen. Virol.* **69,** 2785–2795 (1988).
16. Bullough, P. A. *et al. J. molec. Biol.* **236,** 1262–1265 (1994).
17. Brünger, A. T. *Nature* **355,** 472–475 (1992).
18. Wharton, S. A. *et al. J. biol. Chem.* **263,** 4474–4480 (1988).
19. Carr, C. M. & Kim, P. S. *Cell* **73,** 823–832 (1993).
20. Ward, C. W. & Dopheide, T. A. *Aust. J. biol. Sci.* **33,** 441–447 (1980).
21. Graves, P. N., Schulman, J. L., Young, J. F. & Palese, P. *Virology* **126,** 106–116 (1983).
22. Webster, R. G., Brown, L. E. & Jackson, D. C. *Virology* **126,** 587–599 (1983).
23. Daniels, R. S. *et al.* in *The Origin of Pandemic Influenza Viruses* (eds Laver, W. G. & Chu, C. M.) 8–18 (Elsevier–North Holland, New York, 1985).
24. Ruigrok, R. W. H. *et al. EMBO J.* **5,** 41–49 (1986).
25. Daniels, R. S. *et al. EMBO J.* **6,** 1459–1465 (1987).
26. Yewdell, J. W., Gerhard, W. & Bachi, T. *J. Virol.* **48,** 239–248 (1983).
27. Ruigrok, R. W. H. *et al. Virology* **155,** 484–497 (1986).
28. Goldsmith, E. J. & Mottonen, J. *Structure* **2,** 241–244 (1994).
29. Crick, F. H. C. *Acta crystallogr.* **6,** 689–697 (1953).
30. McLachlan, A. D. & Karn, J. *J. molec. Biol.* **164,** 605–626 (1983).
31. Daniels, R. S. *et al. Cell* **40,** 431–439 (1985).
32. Sauter, N. K. *et al. Biochemistry* **31,** 9609–9621 (1992).
33. Stegmann, T. & Helenius, A. in *Viral Fusion Mechanisms* (ed. Bentz, J.) 89–111 (CRC Press, Boca Raton, Florida, 1993).
34. Söllner, T., Bennett, M. K., Whiteheart, S. W., Scheller, R. H. & Rothman, J. E. *Cell* **75,** 409–418 (1993).
35. Ellens, H., Bentz, J., Mason, D., Zhang, F. & White, J. M. *Biochemistry* **29,** 9697–9707 (1990).
36. Spruce, A. E., Iwata, A., White, J. M. & Almers, W. *Nature* **342,** 555–558 (1989).
37. Kemble, G. W., Danieli, T. & White, J. M. *Cell* **76,** 383–391 (1994).
38. Blum, M., Metcalf, P., Harrison, S. C. & Wiley, D. C. *J. appl. Crystallogr.* **20,** 235–242 (1987).
39. Leslie, A. G. W. in *CCP4 and ESF-EACMB Newsletter on Protein Crystallography, No. 26* (Daresbury Laboratory, Warrington, 1992).
40. *CCP4, A Suite of Programs for Protein Crystallography* (SERC (UK) Collaborative Computing Project No. 4, Daresbury Laboratory, Warrington, 1979).
41. Terwilliger, T. C. & Eisenberg, D. *Acta crystallogr.* A**39,** 813–817 (1983).
42. Metcalf, P., Blum, M., Freymann, D., Turner, M. & Wiley, D. C. *Nature* **325,** 84–86 (1987).
43. Bricogne, G. *Acta crystallogr.* A**32,** 832–847 (1976).
44. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* A**47,** 110–119 (1991).
45. Brünger, A. T. *X-PLOR (Version 3.1): A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).
46. Lüthy, R., Bowie, J. U. & Eisenberg, D. *Nature* **356,** 83–85 (1992).
47. Colovos, C. & Yeates, T. O. *Prot. Sci.* **2,** 1511–1519 (1993).
48. Read, R. J. *Acta crystallogr.* A**42,** 140–149 (1986).
49. Kraulis, P. *J. appl. Crystallogr.* **24,** 924–950 (1991).
50. Carson, M. *J. molec. Graphics* **5,** 103–106 (1987).

ACKNOWLEDGEMENTS. P.A.B. and F.M.H. contributed equally to this work. We dedicate this structure determination to Max Perutz, FRS, a founder of our science, on the year of his eightieth birthday. We thank R. Daniels, R. Ruigrok, A. Treharne and J. Newman for earlier contributions; E. Collins, S. Garman, P. Rosenthal, L. Stern, S. Watowich, and the staff of the Cornell High Energy Synchrotron source MacCHESS for help with data collection; D. Stevens and A. Douglas for technical assistance; and our colleagues for advice and encouragement. F.M.H. was supported by a Helen Hay Whitney Foundation fellowship and the HHMI. P.A.B. was supported by the NIH. This work was funded by the NIH and the MRC. D.C.W. is an investigator with the Howard Hughes Medical Institute. Coordinates are available from the authors (email: hughson@XTAL∅.harvard.edu).

Measurement of the microwave background temperature at a redshift of 1.776

A. Songaila*, L. L. Cowie*, S. Vogt†, M. Keane†, A. M. Wolfe‡, E. M. Hu*, A. L. Oren‡, D.-R. Tytler‡ & K. M. Lanzetta§

* Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA
† Lick Observatory, University of California, Santa Cruz, California 95064, USA
‡ Department of Astrophysics and Space Science, University of California, San Diego, La Jolla, California 92093, USA
§ Astronomy Program, Department of Earth and Space Science, State University of New York, Stony Brook, New York 11794, USA

HOT Big Bang cosmology predicts that the temperature of the cosmic microwave background radiation will increase linearly with increasing redshift to early in the history of the Universe. The local background temperature (2.7 K) is known very accurately from direct measurements[1–3], but other techniques must be used to estimate it at non-zero redshifts. One way is to determine the excitation of atomic transitions in absorbing clouds along the lines-of-sight to distant quasars[4]. When the transitions are in equilibrium with the microwave background radiation, the radiation will populate the fine-structure levels of the ground states of certain atoms, and the relative populations of the levels can be used to calculate its temperature. Here we report the detection of absorption from the first fine-structure level of neutral carbon atoms in a cloud at a redshift of 1.776, towards the quasar Q1331 + 170. The population ratio yields a temperature of 7.4 ± 0.8 K, assuming that no other significant sources of excitation are present. This agrees with the theoretical prediction of 7.58 K.

The cosmic microwave background radiation (CMBR) will populate excited levels of atomic and molecular species when the energy separations involved are not too different from the CMBR peak frequency. The first measurement of the local CMBR temperature was in fact made using this method[5] with fine structure lines in the cyanogen (CN) molecule, although it was not recognised as such until after Penzias and Wilson identified the CMBR[6]. Cyanogen excitation can now be used to measure T_{CMBR} very precisely. Roth *et al.*[7], who measured the rotational excitation of CN toward five Galactic stars and carefully corrected for local sources of excitation, found a value of T_{CMBR} at 2.64 mm of $2.729(^{+0.023}_{-0.031})$ K, in agreement with the COBE result[2] of 2.726 ± 0.010 K.

Bahcall and Wolf[4] first suggested that the method could be extended to high redshift, where T_{CMBR} could be presumed to be larger, using atomic fine-structure transitions in absorbing clouds toward high-redshift quasars. Useful transitions for this purpose include those of C^0, C^+ and N^+, with C^0 being particularly well suited[4,8]. This measurement has been attempted several times, but has generally been limited by the resolution and signal-to-noise available in reasonable exposure times at the very faint magnitudes involved, and, in the case of C^0, the intrinsic weakness of the line. C^+ and N^+ have strong lines, common in the spectra of quasars, but their relatively high fine-structure