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). 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 activated2,3 in a process requiring structural changes in HA.4,5,6 (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 glycopolyptides linked by a single disulphide bond, HA1 (328 residues) and HA2 (221 residues). Each HA2 chain is anchored at its carboxy terminus in the viral membrane. Soluble trimers can be released by treating virus with the protease trypsin, which cleaves each HA2 chain once after residue 175 (refs 7, 8). The structure9,10 of this trimer, BHA1 (Fig. 1a), reveals that the HA2 chains are major components of a mainly α-helical stem domain which forms the centre of the molecule. The HA1 chains also contribute to the stem structure but primarily form three membrane distal globular domains containing the receptor binding sites. This structure poses difficulties in understanding the role of the HA in membrane fusion. 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 HA1, 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. 1b); 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 peptide1,2,3,13,14. 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 aggregate15. The soluble trimeric fragment considered here, called TBHA1, is prepared from BHA at pH 5.0 by successive digestion with trypsin and thermolysin15,16. Each monomer includes residues 38–175 of the HA2 chain disulphide linked to residues 1–27 of the HA1 chain.

Dramatic differences are observed between the BHA and TBHA1 structures. Large changes occur at both ends of the HA2 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 HA2 chain, residues 140–175, was packed in native BHA, apparently extruding it into an extended and partially disordered structure.

Structure determination

The TBHA1 structure was determined using crystals cooled to ~170 °C in cryoprotectant buffer18 (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 factor19. Details are given in Table 1.

The current model (Fig. 2) includes residues 12–16, and 40–153 of the first monomer, residues 11–16, and 40–162 of the second monomer, residues 10–17, and 40–162 of the third monomer, and 37 water molecules (numbering scheme is that for intact HA: HA1 and HA2 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 154–162), residues that are well-ordered in two monomers are
FIG. 1. BHA structure. a, The BHA trimer and b, monomer are shown, with the HA$_2$N chains white and the HA$_3$C 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 disulfide bond. The viral membrane end of the molecule is at the base, near the HA$_2$N C terminus generated by proteolytic removal of the transmembrane anchor sequence; the viral membrane distal globular domains are at the top. Figs 1, 2, 3c and d were generated using MOLSCRIPT.

The major feature of the TBHA$_2$ structure (Fig. 2) is the three-stranded 3-helical coiled coil, approximately 100 Å long, comprising one long 3-helix from each monomer (residues 40-105). This is followed in each monomer by a small connecting loop and a shorter helix (113-129) which packs in an antiparallel orientation against the long helix of the same monomer and that of another monomer. A 3-hairpin (131-140), together with an additional strand contributed by HA$_2$ residues (11-16), forms a small antiparallel 3-sheet. HA$_3$C is connected to HA$_2$ by a disulfide bond between residues 14 and 137. An additional short helix (146-153) packs against the coiled coil. Beyond residue 153, for one chain, and beyond residue 162 for the other two chains, the electron density is not interpretable, presumably indicating disorder.

**BHA and TBHA$_2$ Compared**

Only thirty residues (76-105; C in Fig. 3a, 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 3-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. 3a. Four 3-helical segments (A, C, D, G) and three 3-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 3-helix (106-112) refolds into a loop, allowing a 180° bend between helices C and D in TBHA$_2$. An 3-helix, H, in BHA (159-170) and five residues beyond it, 171-175, appear to be disordered in TBHA$_2$. The overall helix content agrees well with that measured by circular dichroism and, because of compensating changes, is very similar to that of the corresponding portions of the BHA structure.
**Tertiary structure refolding**

**N-terminal.** Major refolding of secondary structure elements occurs at both ends of the TBHA₂ molecule. Helix A (38₋₋₅₅₂), 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 TBHA₂ (Fig. 3b−d). Although the N-terminal residues (1₋₋₃₇₂) have been removed in the preparation of TBHA₂, the location of helix A suggests that the fusion peptide (1₋₋₂₃₂) and two β-strands (2₄₋₋₃₇₂) 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 HA₁, adopts an α-helical conformation and joins the triple-stranded coiled coil in TBHA₂ (Fig. 3b−d). This transition was anticipated by the observation that residues 3₈₋₋₁₂₅₂ contain heptad repeats of hydrophobic residues which could potentially form a coiled coil.¹⁹ ²⁰ It has been shown 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₋₋₁₁₂₂) that made a number of contacts with the fusion peptide in BHA⁸ ¹⁰ refolds to form a loop (Fig. 3b−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 1₁₂₂, 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₋₋₁₂₆₂) remains intact, it folds back together with the small antiparallel β-sheet (1, E, F in Fig. 3b) to pack against the triple-stranded coiled coil in an inverted orientation (Fig. 3b−d). There are many small changes in the architecture of this subdomain (1, D, E, F in Fig. 3b; blue in Figs 3c, d); its root-mean-square Ca 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. 4a). The single Trp in each monomer (residue 9₂₂) is
FIG. 3 The low-pH-induced conformational change. a, Secondary structure for TBHA₂ 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 TBHA₂ 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 TBHA₂ monomer (right) in schematic form is compared to the corresponding region of a BHA monomer (left). Consecutive regions of the HA₃ chain are labelled A–H (see text); H is apparently disordered in TBHA₂. The first β-strand of HA₃ is also shown, labelled ‘1’. Both ends of this strand are apparently disordered in TBHA₂. The disulfide bond linking 14₁ and 137₂ is indicated. The two structures are aligned on the C region, which is unaffected by the conformational change. c, d, The structure of a TBHA₂ monomer (d) is compared to that of an intact BHA monomer (c). Structural elements of the HA₃ chain of TBHA₂ are coloured in rainbow order from amino to carboxy terminus; corresponding regions of BHA are coloured identically. In addition, the HA₁ β-strand which is disulphide-linked to HA₂ is blue; the blue elements (E, 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 TBHA₂ are shown in grey. Regions of BHA corresponding to apparently disordered regions of TBHA₂ are shown in white. e, TBHA₂ (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 ORTEP.

buried, consistent with fluorescence measurements of TBHA₂ in solution. Residues 141₁–175₂, 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 TBHA₂ a more extended, partially disordered conformation running antiparallel to the coiled coil (Fig. 3b–d). This region (yellow in Fig. 4b) docks against residues (other colours in Fig. 4b) in BHA, many of which become buried, move to distant locations, or rearrange relative to one another in TBHA₂ (Fig. 4b legend). The destruction of the site which presumably had stabilized the 141₁–175₂ unit may account for the partial disordering of this
FIG. 4. a, The novel hydrophobic core formed in TBH₂₆ by residues from C, D, E, F, and H; (see Fig. 3b for notation). TBH₂₆ monomers are green, blue, and red. Side chains contributing to the hydrophobic core for the green monomer (Leu 13a, Leu 15a, Ile 89a, Ser 93a, Ala 96a, Leu 99a, Val 100a, Ile 108a, Leu 110a, Met 115a, Phe 119a, Thr 122a, Leu 126a, Phe 138a, and Ile 140a) are shown in purple. b, The ‘docking site’ for the C-terminal portion of HA in BHA. Residues 141c–175c in BHA, including the G and H helices (Fig. 3b), 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 HA₃, 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 TBH₂₆, inaccessible for interaction with the C-terminal region of HA. Magenta residues, although largely surface-exposed in TBH₂₆, rearrange relative to one another. Blue residues are apparently disordered in at least one TBH₂₆ monomer. c, d, Ribbon diagrams of BHA (c) and TBH₂₆ (d) monomers showing the positions of residues that raise the midpoint of the low-pH-induced conformational change. h₄, (magenta) and h₃, (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 HA₃–HA₃ contacts (green), and those affecting HA₁–HA₃ contacts (white). Most mutations were identified by selecting for mutant viruses able to grow in cells treated with amantadine 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. Images were generated using RIBBONS.

region in TBH₂₆. Helix G (146c–153c) remains intact in TBH₂₆, probably stabilized in part by the disulfide 144c–148c, but it docks against the C region of the long helix, approximately where the B loop had docked in BHA (Fig. 3b–d). Residues beyond 153c in one subunit and 162c in the others, including helix H, were not located in our electron density maps, presumably because they are disordered in TBH₂₆.

HA₁–HA₃ interactions. In TBH₂₆, HA₁ has refolded so extensively that many of the HA₁–HA₃ interactions seen in BHA would no longer be possible even if the bulk of the HA₁ chain were still present. In BHA, loop B (orange in Fig. 3c) makes all of the HA₁–HA₃ contacts involving the globular head domain of HA₁ and many involving the descending C-terminal HA₁ strand. Having refolded into a helical conformation in TBH₂₆, 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 HA₁ globular domains in BHA or in the intact HA are displaced in the fusion pH conformational change. A second site (the C-D loop; green in Fig. 3c, d) which refolds significantly in TBH₂₆ was the docking site for a loop (residues 24c–37c) in the ascending N-terminal HA₁ strand in BHA. Site-specific antipeptide antisera binding and proteolytic susceptibility at residue 27 indicate that the environment of this HA₁ loop rearranges in the intact protein at the pH of fusion.

The observed irreversibility of the fusion-pH-induced conformational transition and the increased thermostability of TBH₂₆ compared to BHA even at neutral pH 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. Kinetically controlled activity and conformational refolding have also been observed in the serpin family of protease inhibitors (reviewed in ref. 28).

TBH₂₆ coiled coil

The triple-stranded coiled coil of TBH₂₆ displays, with few exceptions, the knobs-into-holes packing predicted by Crick (Fig. 5a, 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. 3a). At two positions along the coil an unusual 3–4–4–3 periodicity is observed (Figs 5a, 5c), as predicted in part by Ward and Dopheide but not by others, 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. In TBH₂₆, 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. 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. 4c, d). The first group (red in Fig. 4c, d) includes residues that appear to stabilize
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 HA1 and B and CD of HA2, 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 HA1 subunits form a fourth group (white, Fig. 4c). No new insights about this group are available directly from the TBHA2 structure, from which HA1 residues 28–328, have been proteolytically removed.

As previously discussed,6,31 these mutations suggest that the trimeric assembly of HA1 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 HA1 is docked in BHA. The HA1 globular domains themselves apparently retain a native-like structure as they interact with many anti-BHA antibodies22,23,26, bind sialic acid receptor analogues with the same affinity as BHA,35, and display an appropriate circular dichroism spectrum.36

Electron microscopy
The dimensions and shape of TBHA2 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 structure5 (bottom, Fig. 2d). In addition, EM images of an earlier soluble thermolysin digestion product, identical to TBHA2 except that the predominant HA2 N terminus is residue 24, rather than 36, show molecules about 130 Å long, again with terminal knobs15. This observation is consistent with the idea that residues 24–37 extend from the tip of the coiled coil, possibly as a continuation of the a-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 TBHA2 except that it retains the first 37 residues of the HA2 chains including the fusion peptides, forms ‘rosettes’ of 8±1 molecules with distal knobs17. Again, these knobs are presumably the bulky ‘bottom’ region seen in TBHA2, 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 inward inversion of the main body of the trimer relative to the membrane. This model is possible given the structure of TBHA2, inasmuch as the apparently disordered C-terminal region of the HA2 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 membrane34; namely, that the fusion peptides insert non-productively into the viral membrane.

The removal of 28–328, and both termini of HA2 in producing TBHA2 might have removed residues which stabilize a different conformation, or might have generated additional degrees of freedom unavailable in the intact molecule. While the TBHA2 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 TBHA2 structure suggests that the fusion-pH-induced conformational change delivers the fusion peptides at least 100 Å towards the target membrane (Fig. 3e), consistent with the expectation that HA might form a bridge between the viral and cellular membranes1,11,19,33. We cannot rule out, however, the possibility that the main body of the molecule inverts to place the fusion peptide in the viral membrane15. 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) membrane24. 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 TBHA2 suggests that the extra-
LETTERS TO NATURE

Measurement of the microwave background temperature at a redshift of 1.776


* 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 measurements1-4, 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 quasars5. 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 7.5 ± 0.8 K, assuming that no other significant sources of excitation are present. This agrees with the theoretical prediction of 5.58 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 method6 with fine structure lines in the cyanogen (CN) molecule, although it was not recognised as such until after Penzias and Wilson identified the CMBR7. Cyanogen excitation can now be used to measure T_{CMBR} very precisely. Roth et al.8, 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(±0.022) K, in agreement with the COBE result9 of 2.726 ± 0.010 K. Bahcall and Wolf10 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 suited11. 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...