

# Disulfide Bond Formation during the Folding of Influenza Virus Hemagglutinin

Mark S. Segal,\* Jacqueline M. Bye,\* Joseph F. Sambrook,\* and Mary-Jane H. Gething\*‡

\*Department of Biochemistry and ‡Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas, 75235

**Abstract.** To study the importance of individual sulfhydryl residues during the folding and assembly in vivo of influenza virus hemagglutinin (HA), we have constructed and expressed a series of mutant HA proteins in which cysteines involved in three disulfide bonds have been substituted by serine residues. Investigations of the structure and intracellular transport of the mutant proteins indicate that (a) cysteine residues in the ectodomain are essential both for efficient folding of HA and for stabilization of the folded molecule; (b) cysteine residues in the globular portion of the ectodomain are likely to form native

disulfide bonds rapidly and directly, without involvement of intermediate, nonnative linkages; and (c) cysteine residues in the stalk portion of the ectodomain also appear not to form intermediate disulfide bonds, even though they have the opportunity to do so, being separated from their correct partners by hundreds of amino acids including two or more other sulfhydryl residues. We propose a role for the cellular protein BiP in shielding the cysteine residues of the stalk domain during the folding process, thus preventing them from forming intermediate, nonnative disulfide bonds.

IT is well established that disulfide bonds contribute significantly to the stability of the native state of proteins that contain them (10, 38, 43, 51), but the role of disulfide bond formation during the process of protein folding is less certain. Creighton (10) has stated that disulfide bonds should not be considered as determinants of the folded conformations of proteins; rather, he suggests that disulfide bonds are formed spontaneously (if the redox conditions are appropriate) and specifically only when the protein conformation favors them. On the other hand, Kim and Baldwin (31) propose that disulfide bond formation can be regarded as an adhesion step, a structure-determining event that locks a region of a protein into a particular folded state. The apparent contradiction in these statements is reconciled by the reversibility of the disulfide bond "locks" (31). As both Creighton (10) and Kim and Baldwin (31) point out, the pairing of cysteines to form disulfide bonds is approximately random during the early stages of protein refolding. However, as folding proceeds, acquisition of secondary structure and disulfide bond formation become mutually cooperative processes. Thus, if a particular bond allows a folded structure to be formed stably, then that bond will be retained preferentially. Conversely, the acquisition of elements of secondary structure will energetically favor the formation of certain disulfide bonds.

Both scenarios predict that "non-native" disulfide bonds, i.e., bonds that are not present in the final native structure, may be formed transiently during protein folding. This is particularly true during protein refolding in vitro, when all

the cysteine residues in a polypeptide chain are immediately available for disulfide bond formation. Extensive studies by Creighton and co-workers of the in vitro refolding of bovine pancreatic trypsin inhibitor (BPTI;<sup>1</sup> reviewed in 11) indicated that significant quantities of BPTI molecules containing such non-native bonds are formed during refolding and that these molecules are obligate intermediates in the folding pathway. However, more recent investigations by Weissman and Kim (54) suggest that the well-populated intermediates present during the refolding of BPTI contain only native disulfide bonds. Nevertheless, rearrangement of these molecules to the final native form must proceed through the formation of transient intermediates that contain non-native disulfide bonds (54). The question remains as to whether the obligatory formation of non-native bonds is a common occurrence during protein refolding, and whether this mechanism is relevant during folding in vivo in eukaryotic cells, where protein disulfide isomerase (PDI) and the effective thiol/disulfide redox couple, reduced and oxidized glutathione, are present to assist disulfide rearrangements (17).

The majority of proteins containing disulfide bonds are transmembrane or secretory proteins that are either transported along the exocytic pathway of eukaryotic cells or exported directly through the plasma membrane of prokaryotic cells (51). In vivo, folding of these protein can begin when only a portion of the nascent polypeptide chain has been

1. *Abbreviations used in this paper:* BPTI, bovine pancreatic trypsin inhibitor; HA, hemagglutinin.

translocated across the lipid bilayer from its site of synthesis in the cytoplasm, and it has been well documented that native disulfide bonds can be formed before the completed polypeptide is released from the ribosome (3, 46). This situation could decrease the likelihood of formation of non-native disulfide bonds during folding of molecules such as immunoglobulin light chains (3) or serum albumin (7) in which the cysteine partners in native bonds are located either sequentially or close to one another in the amino acid sequence. However, the probability of formation of non-native bonds may be increased when the proper cysteine partners are separated in the polypeptide chain by many amino acids including other cysteine residues, as is the case for many disulfide-bonded proteins (51).

Protein synthesis on membrane-bound polysomes in eukaryotic cells occurs at a rate of  $\sim 300$  amino acids per minute (6, 29, 32, 52) and disulfide bond formation in the ER may be completed within seconds of termination of translation and translocation of the polypeptide chain (3, 39, 46). The rapidity of these processes makes it extremely difficult to investigate whether intermediate non-native disulfide bonds play a significant role during the folding of membrane and secretory proteins *in vivo*. It is, however, possible to use site-directed mutagenesis to substitute individual cysteine residues in a protein and thus to test their importance during folding and in the stability of the native molecule.

The protein chosen for this investigation was the hemagglutinin (HA) glycoprotein of influenza virus, whose folding pathway and final three-dimensional structure have been the subjects of extensive study (8, 9, 21, 22, 25, 55, 56). The HA molecule from the A/Japan/305/57 strain contains 15 cysteine residues (24). Twelve of these form six disulfide bridges in the HA ectodomain (see Fig. 1 A), while the remaining three are located within the transmembrane and cytoplasmic domains of the molecule (53). Comparison of the amino acid sequences of different HA proteins from more than 20 influenza virus subtypes (1) reveals that the positions of the 12 cysteines in the ectodomain are completely conserved, suggesting that all the disulfide bonds are important for the folding and/or stability of HA. A similar degree of conservation of a particular amino acid residue is rare at other positions in the molecule. Cysteine residues that are either embedded in the lipid bilayer or located in the reducing environment of the cytoplasm are not conserved (24), presumably because they do not form disulfide bonds.

We chose for mutagenesis cysteines involved in three disulfide bonds in the Japan HA ectodomain (see Fig. 1 A) that may, as discussed below, have differing propensities to form non-native bonds during the folding process. One bond, between Cys<sub>4</sub> and Cys<sub>462</sub>, links the NH<sub>2</sub> and COOH termini of the ectodomain. The second, between Cys<sub>42</sub> and Cys<sub>273</sub>, joins the NH<sub>2</sub> and COOH termini of the globular domain of the molecule. The third bond, between Cys<sub>55</sub> and Cys<sub>67</sub>, closes a small loop of only 11 amino acids on the surface of the globular domain.

During the biosynthesis and translocation of the Japan HA polypeptide, Cys<sub>4</sub> is the first sulfhydryl residue to enter the ER lumen. This residue has the opportunity to interact with eight other cysteine residues in the HA sequence before its proper partner, Cys<sub>462</sub>, is synthesized and translocated (see Fig. 1 B). Formation of the stable bond between Cys<sub>4</sub> and Cys<sub>462</sub> may be delayed even after Cys<sub>462</sub> is translocated into

the lumen, since that bond links adjacent chains of a five-stranded  $\beta$ -sheet (Fig. 1 A) that is not thought to form until the HA trimer assembles (56), an event that occurs with a half-time of between 7 and 30 min, depending on the rate of synthesis of HA monomers (6, 21, 25). The formation of the Cys<sub>4</sub>-Cys<sub>462</sub> disulfide bond emphasizes the hairpin topology of each HA subunit, with both the NH<sub>2</sub> and COOH termini lying adjacent to the membrane (see Fig. 1 A). Wilson et al. (56) inferred from this topology that the polypeptide may fold while still attached to the membrane not only by the COOH-terminal transmembrane sequence but also by the NH<sub>2</sub>-terminal hydrophobic signal sequence, whose cleavage would occur after the majority of the polypeptide chain had been synthesized. Delayed cleavage of the signal peptide would hold the NH<sub>2</sub>-terminus close to the membrane and might thus limit the opportunity for formation of non-native disulfide bonds involving Cys<sub>4</sub>.

The second cysteine residue in the Japan HA sequence, Cys<sub>42</sub>, is separated from its eventual partner, Cys<sub>273</sub>, by 231 amino acids, including four other cysteine residues (see Fig. 1 B). Since Cys<sub>42</sub> lies distant from the membrane in the native molecule (see Fig. 1 A), it should not be sterically hindered from forming non-native bonds.

Finally, Cys<sub>55</sub> and Cys<sub>67</sub> were selected because we wanted to study adjacent cysteine residues that may form a disulfide bond rapidly and directly without forming any non-native intermediates. These two cysteines are very close in the polypeptide chain, and are not separated by any other sulfhydryl residues (see Fig. 1, A and B). Furthermore, they are located within the globular domain of the HA monomer, which is thought to fold very rapidly, independent of the other portions of the molecule and well before assembly of the trimer (56). The remaining three disulfide bridges, one within the globular domain and two within the stalk domain, are similar to the Cys<sub>55</sub>-Cys<sub>67</sub> bond in that they also close small peptide loops that do not contain other cysteine residues (see Fig. 1 B).

## Materials and Methods

### Recombinant DNA Techniques

Buffers and reaction conditions for restriction enzymes, T4 DNA ligase, T4 polynucleotide kinase, and the Klenow fragment of *E. coli* DNA Polymerase I were those listed by the commercial source, Boehringer Mannheim Biochemicals (Indianapolis, IN). Isolation of DNA fragments, preparation of plasmid DNAs, and other standard recombinant DNA techniques were carried out as described in Sambrook et al. (48). Recombinant plasmids were transformed into *E. coli* strain DH-1 or TG1 cells by the method of Hanahan (27).

### Site-specific Mutagenesis of HA cDNA Using Mismatched Oligonucleotides

Plasmid pSVEHA3, which contains a full-length cDNA copy of the HA gene from influenza virus strain A/Japan/305/57 (20), was digested with KpnI and BamHI restriction endonucleases, and a 1,731-bp DNA fragment containing the entire HA coding sequence was purified by gel electrophoresis. This fragment was inserted between the KpnI and BamHI sites of the double-stranded replicative form of M13mp18 bacteriophage DNA, and *E. coli* TG1 cells were transfected with the ligated DNA molecules. A recombinant phage was selected in which the single-stranded DNA purified from the M13 virions secreted from infected bacterial cells contained the noncoding complementary sequences of HA cDNA. Six 17- or 18-mer oligonucleotides were synthesized to be complementary, except for a single mismatch, to sequences encoding regions of the protein where a serine resi-

due was to be substituted for a cysteine residue (Fig. 1 C). The procedures used to carry out the oligonucleotide-directed mutagenesis and identification of the desired mutants have been described in detail by Zoller and Smith (57, 58). Once the desired base substitutions were confirmed by the technique of chain termination sequence analysis (49), the double-stranded replicative form of the bacteriophage DNAs containing each mutant HA sequence was prepared and purified by centrifugation on CsCl<sub>2</sub> gradients. Convenient restriction endonuclease sites were used to isolate DNA fragments containing each altered sequence. These were used to replace the corresponding wild-type sequences in pSVEHA3. To create the cDNAs encoding the double mutants C4S/C462S and C42S/C273S, two restriction fragments, each containing a mutated codon, jointly replaced the analogous wild-type sequences within pSVEHA3. Because there are no unique restriction endonuclease sites between the codons specifying Cys<sub>55</sub> and Cys<sub>67</sub>, the double mutant C55S/C67S was constructed using, as template, single-stranded DNA isolated from M13 bacteriophage containing the C55S mutant cDNA.

### Generation of SV40-HA Virus Stocks and Infection of Simian CV-1 Cells

SV40-HA recombinant genomes containing wild-type and mutant forms of the Japan HA cDNA were transfected into CV-1 cells using DEAE-dextran and chloroquine as previously described (14). High-titer virus stocks were developed and used to infect fresh monolayers of CV-1 cells for analysis of the biosynthesis and intracellular transport of the wild-type and mutant HAs. Conditions for growth and infections of CV-1 cells were as described previously (14).

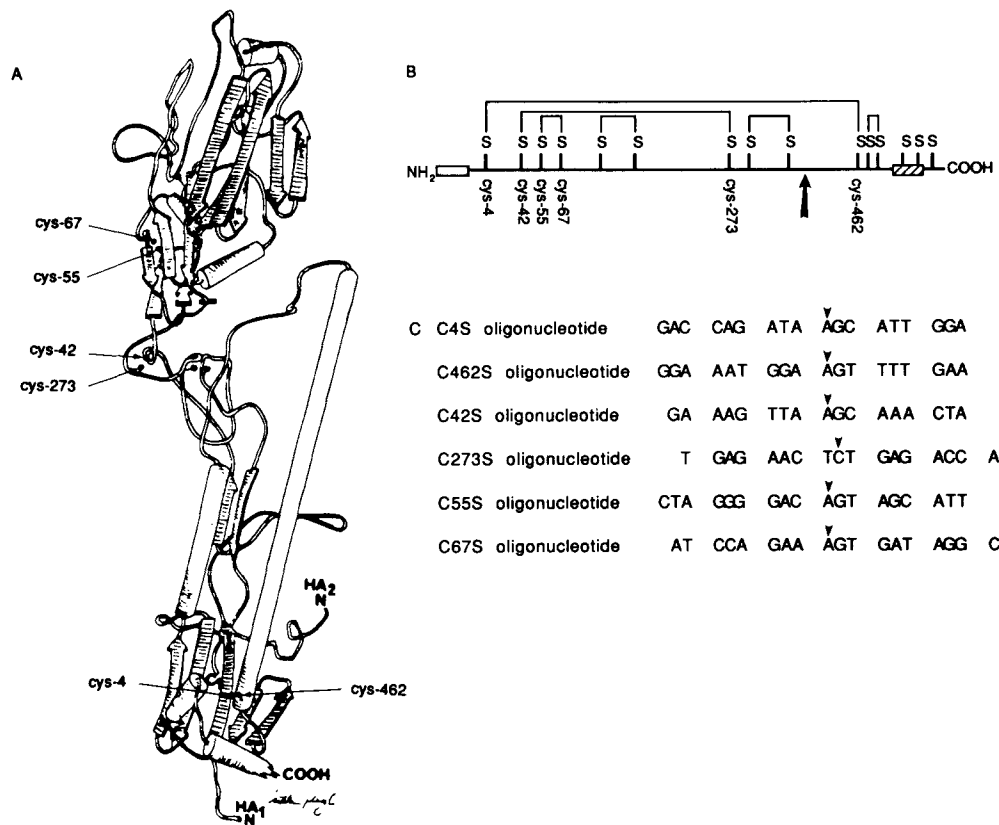
### Characterization of Wild-type and Mutant HAs Expressed in CV-1 Cells

Analysis of the biosynthesis of wild-type and mutant HA proteins was car-

ried out by pulse-chase radiolabeling of infected cells with [<sup>35</sup>S]methionine as described previously (25). To analyze the proportion of HA that was transported to the cell surface, trypsin (10 μg/ml) was included in the medium for the last 15 min of the chase period. Separation of monomeric and oligomeric forms of HA was performed by sedimentation velocity centrifugation on sucrose gradients as described previously (25), except that the buffer contained lauryl maltoside (6 mM; Boehringer Mannheim Biochemicals) instead of octyl glucoside. To analyze the protease sensitivity of HA molecules, cell extracts were incubated on ice with trypsin (100 μg/ml) for 30 min and then soybean trypsin inhibitor (200 μg/ml) was added to terminate the digestion. HA proteins were immunoprecipitated and analyzed by SDS-PAGE and autoradiography as previously described (25). Digestion of immunoprecipitated proteins with endo-N-acetylglucosaminidase H (endo H, a kind gift from Dr. Robert Trimble, NY Department of Health, Albany, NY) was carried out as described by Owen et al. (42). The cellular location of wild-type and mutant forms of HA was determined by indirect immunofluorescence of infected cells (2, 14) using the polyclonal and monoclonal anti-HA antibodies described in (25).

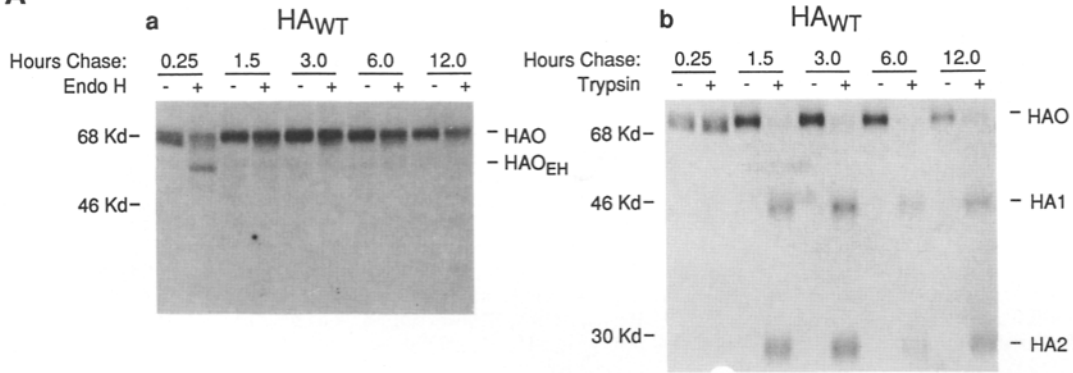
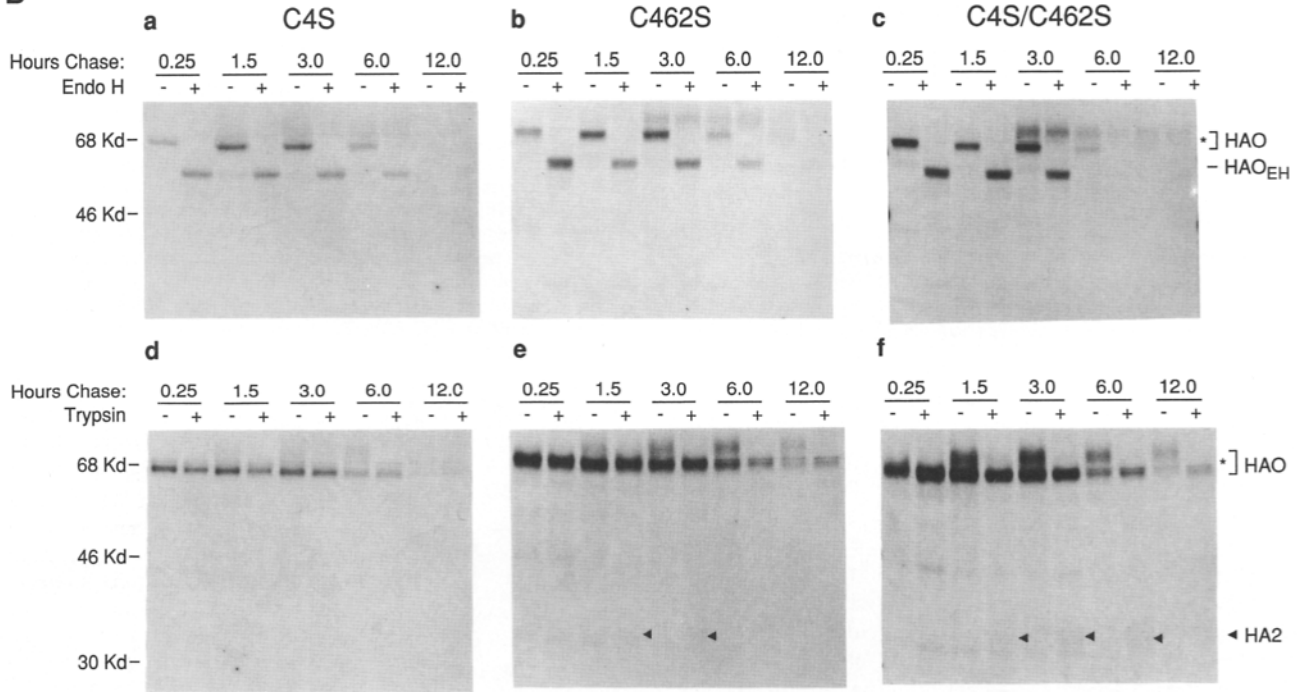
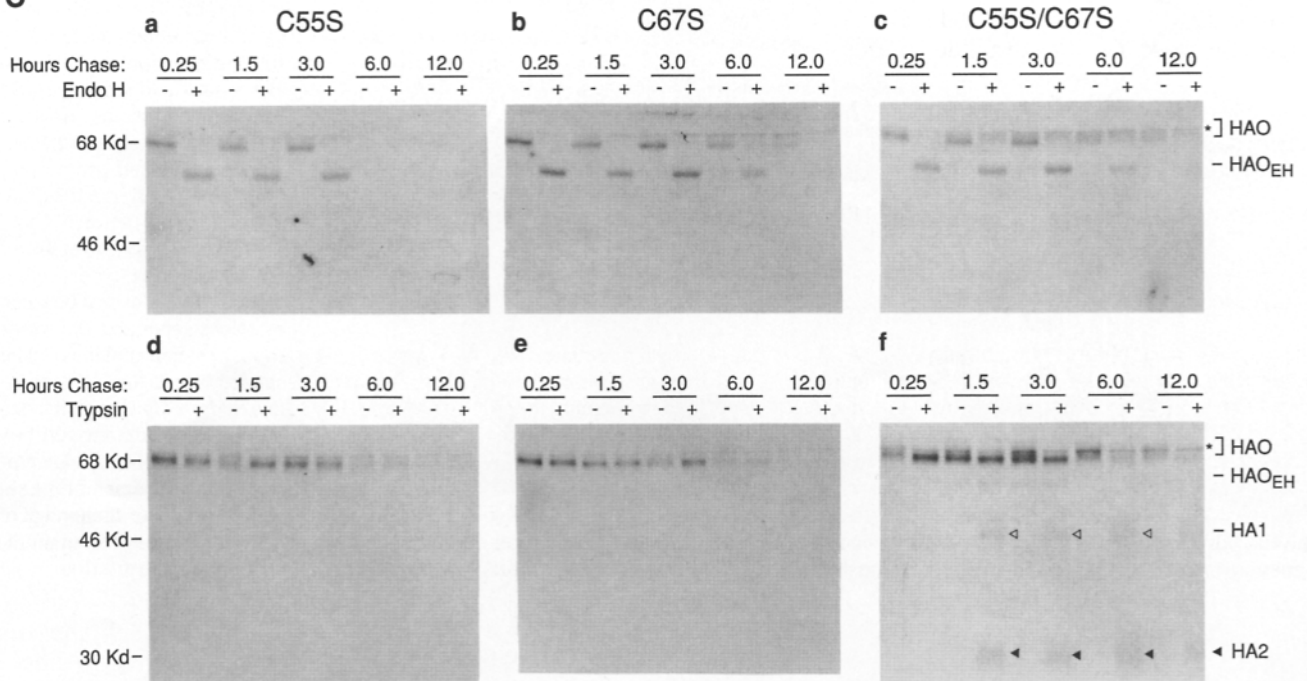
### Isolation of HA Proteins Using Ricin-Sepharose

Precipitation with Sepharose beads conjugated with ricin B chains was performed to isolate those proteins that have galactose residues added to their oligosaccharide side chains. 5.2 μg of ricin B chain, conjugated to Sepharose beads (Sigma Chemical Co., St. Louis, MO) was incubated with aliquots of sucrose gradient fractions (~250 μg of HA protein/1 mg of lectin). After 1 h of rocking at 4°C, the ricin B chain Sepharose-protein complexes were pelleted by centrifugation for 1 min at 12,000 g. The pellets were washed three times with NET-gel buffer (25). Subsequently, the HA proteins were immunoprecipitated as described previously (25) except that the precipitation buffer contained 0.2 M galactose (Sigma Chemical Co.) to elute the glycoproteins from the Sepharose-conjugated ricin B chain.



**Figure 1.** Illustration of the positions and disulfide bonding patterns of the cysteine residues in A/Japan/HA. (A) The positions in the HA from the A/Japan/305/57 strain of influenza virus of the cysteine residues selected for substitution by serine are shown on a schematic representation of the ectodomain of the HA from the A/X-31 influenza virus strain (56). The cysteine residues are numbered according to their position in the Japan HA amino acid sequence. (B) Linear map of the A/Japan HA showing the pattern of disulfide bonds in the molecule. Notice that only the disulfide bonds formed between Cys<sub>4</sub>-Cys<sub>462</sub> and Cys<sub>55</sub>-Cys<sub>67</sub> are of the nonadjacent type (i.e., a disulfide bond formed between cysteines separated by two or more cysteine residues). The shaded box at the NH<sub>2</sub>-terminus represents the signal sequence while the striped box near the COOH-terminus represents the transmembrane sequence. The site of cleavage of

the HA polypeptide into HA1 and HA2 chains is indicated by an arrow. (C) Sequences of the oligonucleotides used for site-directed mutagenesis of the selected cysteine residues in the Japan HA molecule. The arrowheads show the site of the nucleotide substitution.

**A****B****C**

## Results

### Construction and Expression of Mutant HA cDNAs Encoding Proteins with Cysteine Residues Substituted by Serine

Oligonucleotide-directed mutagenesis of the A/Japan HA cDNA (see Materials and Methods) was used to investigate the role of six cysteine residues that make up three of the six disulfide bonds present in the folded HA molecule (Fig. 1 A). To study the role of the individual residues in the folding pathway of the HA monomer, single substitution mutants were constructed, each with one of the cysteine codons of the targeted disulfide bond individually mutated to a serine codon. To study the contribution of the three disulfide bonds to the stability of the molecule, the codons specifying both cysteine residues involved in each bond were mutated jointly. We chose to mutate cysteine residues to serine since this change is likely to maintain the polarity and size of the amino acid side chains as closely as possible (26).

SV40-HA recombinant genomes were constructed in which either the wild-type Japan HA gene or the mutant genes replaced the coding sequences of the late region of SV40 DNA (Materials and Methods). These recombinant genomes were transfected into simian CV-1 cells and high-titer recombinant virus stocks were developed (14). After infection of CV-1 cells with these stocks, assays were performed to analyze the structure and intracellular transport of the wild-type and mutant HA proteins. The time course and efficiency of movement of the proteins from the ER to the Golgi apparatus were measured by assaying the rate and extent of acquisition of resistance to cleavage of oligosaccharide side chains by endoglycosidase H (33), while the arrival of HA molecules at the cell surface was monitored by assaying their accessibility to proteolytic cleavage into HA1 and HA2 polypeptides by trypsin added to the external medium (15). The state of folding or aggregation of the proteins was determined using conformation-sensitive antibodies (25), by measuring their sensitivity to proteases added to cell extracts, and by separating monomeric and oligomeric forms of HA on sucrose density gradients (25) before analysis by SDS-PAGE under reducing and nonreducing conditions. The results obtained are shown in Figs. 2-7, and are discussed first for the wild-type protein, and then in turn for the mutants of each of the three disulfide bonds.

### Analysis of the Folding, Assembly, and Intracellular Transport of Wild-type Japan HA

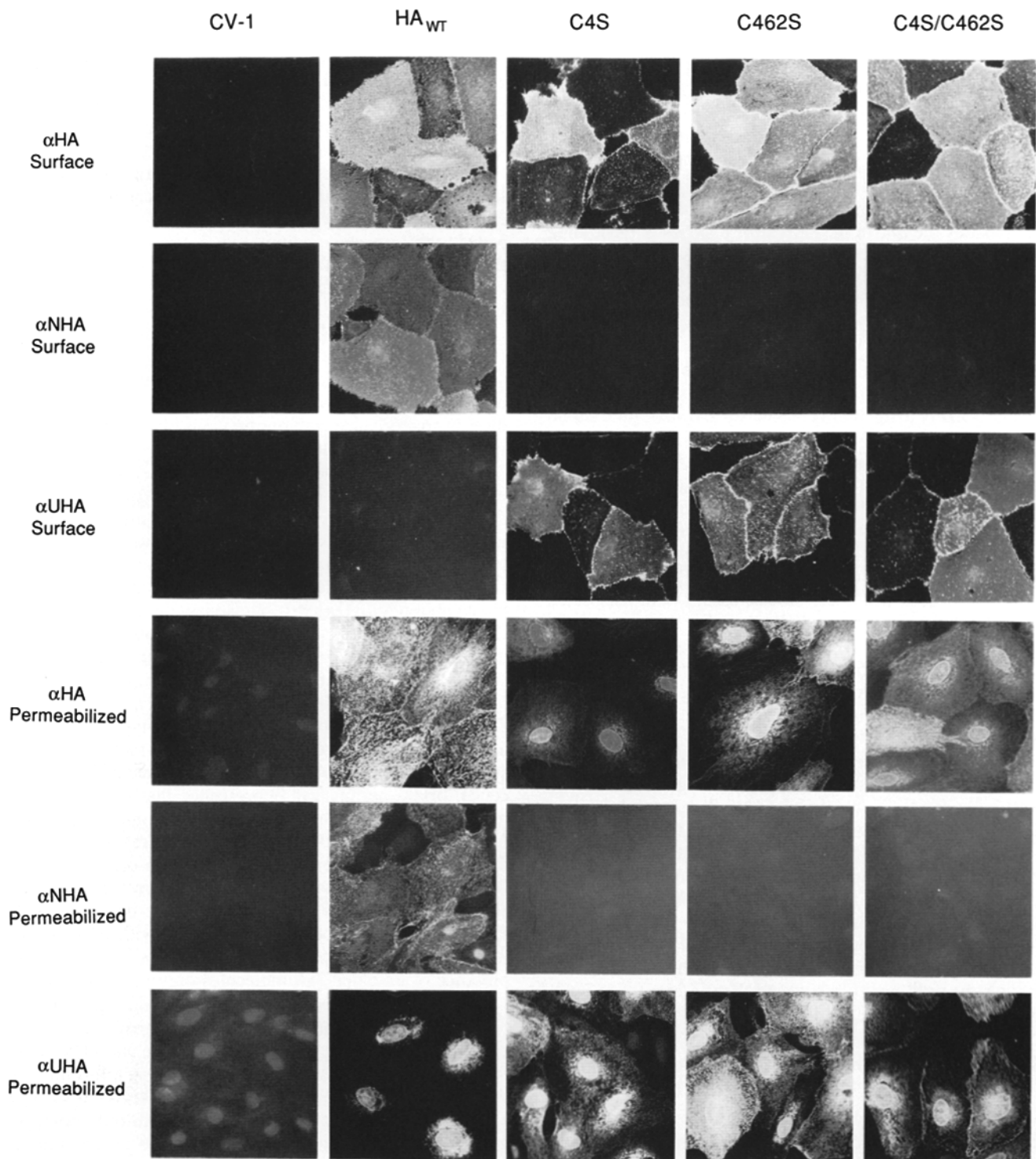
As reported previously (25), wild-type Japan HA is rapidly and efficiently transported from the ER to the Golgi appara-

tus and then to the cell surface. In the experiment shown in Fig. 2 Aa, cells expressing wild-type HA were labeled for 10 min with [<sup>35</sup>S]methionine and then chased in non-radioactive medium for periods up to 12 h. After 15 min of chase, ~50% of the labeled HA0 molecules remained in the core-glycosylated form ( $M_r$  68), containing oligosaccharide side chains that are sensitive to cleavage by endo H, while the remainder had been processed to the terminally-glycosylated form ( $M_r$  71) that is resistant to side-chain cleavage by endo H. None of these molecules were accessible to trypsin added to the external medium (Fig. 2 Ab), indicating that although they had been transported to or through the medial cisterna of the Golgi apparatus (45), they had not yet reached the plasma membrane. By 1.5 h of chase, however, the great majority of the labeled HA0 molecules had acquired resistance to endo H and could be quantitatively cleaved by trypsin into HA1 and HA2 polypeptides, demonstrating that they had been transported via the Golgi apparatus to the cell surface. In concordance with our previous observations that the rate of degradation of wild-type HA is very slow ( $t_{1/2}$  = 12-20 h) (28), disappearance of labeled HA molecules was evident only after 12 h of chase (Fig. 2 Aa).

The conformation of both intracellular and cell surface forms of HA was assessed by immunofluorescence using structure-specific antibodies (Fig. 3). Intact cells displayed similar patterns of diffuse surface staining with either a polyclonal antibody ( $\alpha$ HA) that binds virtually all forms of HA (25) or a monoclonal antibody ( $\alpha$ NHA) that recognizes only the correctly folded, trimeric forms of the protein (25). No staining of intact cells was obtained using an antibody ( $\alpha$ UHA) that recognizes only unfolded forms of HA (25). These results demonstrate that all the wild-type HA present at the cell surface is correctly folded. Permeabilized cells stained with  $\alpha$ HA displayed in addition to the diffuse fluorescence, bright perinuclear fluorescence that represents HA molecules present in the ER and Golgi apparatus (14). Staining of permeabilized cells with  $\alpha$ NHA revealed similar diffuse fluorescence but a different pattern of juxtannuclear fluorescence, a result consistent with the presence of folded HA molecules in the Golgi apparatus but not in the ER. Finally, staining of permeabilized cells with  $\alpha$ UHA was limited to reticular perinuclear structures indicating the presence of unfolded HA molecules only in the ER (25).

To analyze the state of assembly of the wild-type HA proteins, centrifugation of cell extracts on sucrose density gradients was used to separate the monomeric and oligomeric forms of newly synthesized HA labeled during a 30-min continuous pulse with [<sup>35</sup>S]methionine (Fig. 4 A). Wild-type HA polypeptides were located in two regions of the sucrose gradient. A peak centered on fractions 4-6 contains unas-

**Figure 2.** Analysis of intracellular transport of wild-type and mutant HA proteins. 36 h after infection with SV40-HA recombinant viruses, CV-1 cells were pulse labeled with [<sup>35</sup>S]methionine for 5 min and then incubated in an excess of nonradioactive methionine for the indicated times. Intact monolayers were treated with DME (-) or DME containing 5  $\mu$ g/ml of trypsin (+) during the last 15 min of the chase period. Cell extracts were then prepared and HA polypeptides were immunoprecipitated with  $\alpha$ HA serum. Precipitated proteins, obtained from cells not exposed to trypsin, were treated with endo H (+) or mock treated (-). All of the samples were analyzed by SDS-PAGE and visualized by autoradiography. HA0 represents uncleaved HA polypeptides; HA1 and HA2 are the two polypeptide chains that result from cleavage of HA0 with trypsin. The asterisks indicate the position of the BiP polypeptide which migrates slightly slower than the core-glycosylated HA0 species. The arrowheads in panels Be and Bf indicate the position of the precipitated HA2 polypeptide. The open and closed arrowheads in panel Cf indicate the positions of the HA1 and HA2 species, respectively.



**Figure 3.** Intracellular localization and antigenic characterization of wild-type and mutant HA proteins expressed in simian CV-1 cells. 48 h after infection with SV40-HA recombinant viruses or mock infection, CV-1 cells grown on glass slides were fixed with paraformaldehyde as described previously (14) before (intact cells showing surface expression) or after (permeabilized cells) treatment with 0.1% Triton X-100. Indirect immunofluorescent staining was carried out as described (14) using as primary antibody either a polyclonal rabbit antibody ( $\alpha$ HA) that binds virtually all forms of HA, a mAb ( $\alpha$ NHA) that recognizes only the correctly folded, trimeric forms of the protein, or a second polyclonal rabbit antibody ( $\alpha$ UHA) that recognizes only unfolded forms of HA (25). Secondary antibodies were either goat anti-rabbit or goat anti-mouse IgGs conjugated with FITC.



sembled monomers (25) that migrate as a single, sharp band after reduction with DTT (Fig. 4 A, a), but as a doublet of bands under nonreducing conditions (Fig. 4 A, b). These species probably represent populations of core-glycosylated molecules with distinct disulfide bonding patterns. Braakman et al. (6) have recently reported that newly synthesized HA polypeptides from the A/X31 and A/Japan influenza virus strains migrate under nonreducing conditions as three species, two of which (IT1 and IT2) are transient intermediates with incomplete disulfide bonds. IT1 was consistently the most slowly migrating species (6) and probably corresponds to the upper band of the monomer doublet (Fig. 4 b, lanes 4–6). Braakman et al. (6) reported that the IT2 species of Japan HA was difficult to discern because it migrates very close to the fully disulfide-bonded (NT) species. It is possible that IT2 and NT are also not resolved in the autoradiograph shown in Fig. 4; alternatively, the IT2 intermediate may not be present in significant amounts following the longer pulse used in our experiments. In the sucrose gradients shown in Fig. 4 A a second peak located in fractions 8 and 9 contains trimeric HA molecules (25) that are fully dissociated into monomers by SDS after reduction (Fig. 4 a). Although the trimers contain no intermolecular disulfide links (12), they are only partially dissociated by SDS into dimeric and monomeric species under nonreducing conditions (Fig. 4 b). These wild-type HA molecules also migrate as a doublet on SDS-PAGE, but in this case because of a difference in the degree of oligosaccharide processing rather than in the extent of disulfide bond formation. Thus, the more diffuse upper band of the doublet corresponds to molecules whose oligosaccharides have been terminally processed; the lower band corresponds to molecules whose oligosaccharide side-chains have undergone trimming by enzymes of the ER and Golgi apparatus (25). This was confirmed by performing the pulse-chase analysis shown in Fig. 5 A, which reveals the time course of trimming and terminal glycosylation of the oligosaccharide side chains on the HA molecules, and by incubating aliquots of the gradient fractions with Sepharose beads conjugated to ricin B chains, which recognize galactose residues added to the oligosaccharide side chains in the *trans*-cisternae of the Golgi complex (33). Those glycoproteins that bound to ricin-Sepharose were released with galactose before immunoprecipitation with  $\alpha$ HA serum and analysis by SDS-PAGE under nonreducing conditions (Fig. 4 A, c). As expected, galactose-containing HA molecules were present only in fractions 8 and 9, migrating under nonreducing conditions as undissociated trimers, partially dissociated dimers, and fully dissociated monomers. These terminally glycosylated molecules appear as single bands that migrate, in the case of the monomer and dimer species, with the mobility of the upper bands of the doublets seen in Fig. 4 A, b. Thus, all of the wild-type HA monomer subunits that have assembled into trimers and have been transported to the *trans*-cisternae of the Golgi apparatus appear to have the same arrangement of disulfide bonds.

As a final measure of the state of folding of the wild-type HA proteins we determined their sensitivity to proteolysis. Correctly folded HA trimers are specifically cleaved by trypsin into HA1 and HA2 polypeptides, while monomeric or improperly folded HA molecules are further degraded by the protease (12, 25, 50). When cell extracts containing wild-

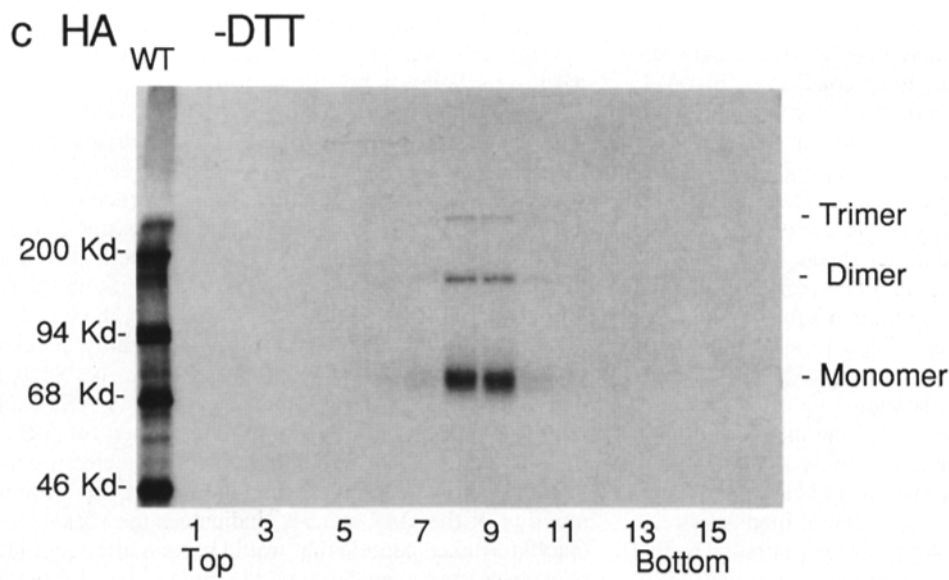
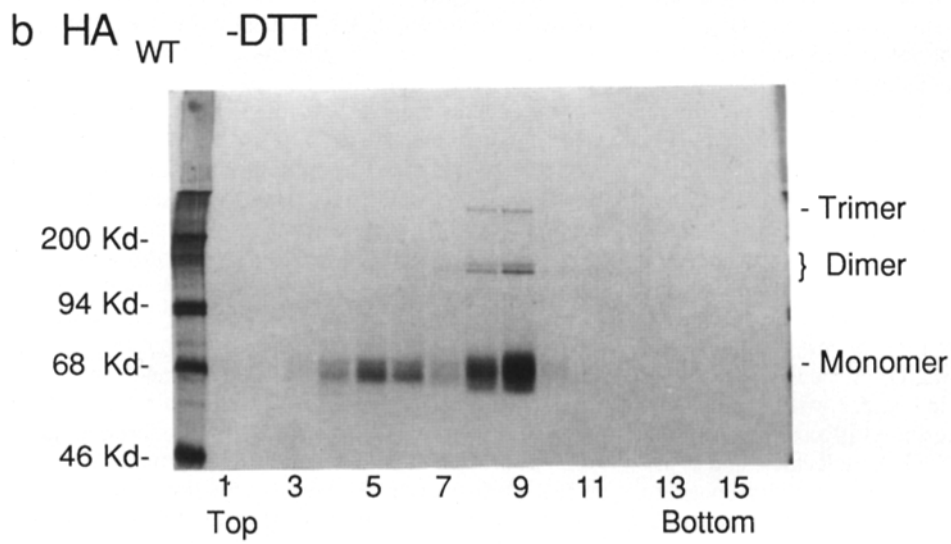
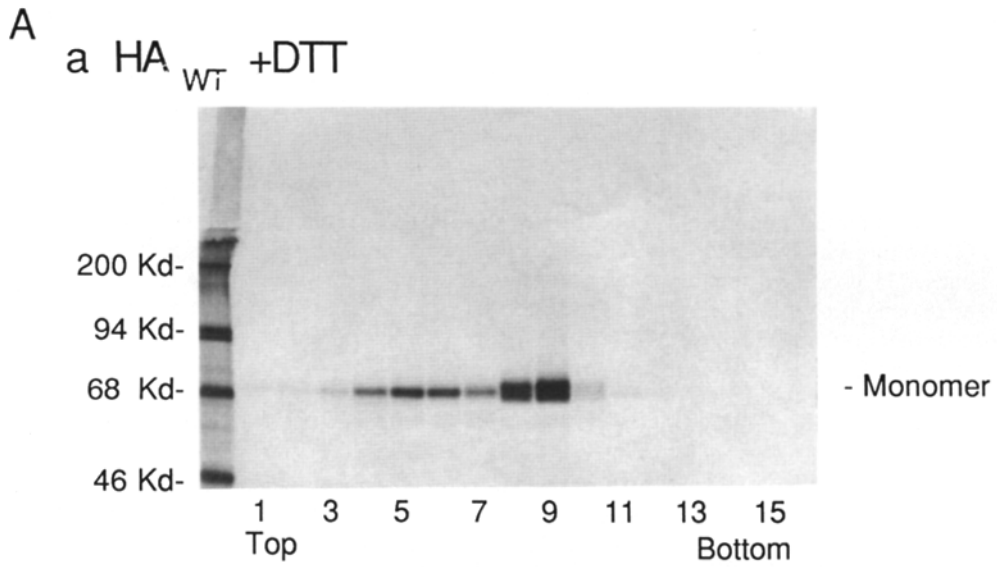
type HA labeled for 2 h with [ $^{35}$ S]methionine were incubated with trypsin (100  $\mu$ g/ml) for 30 min at 4°C, the protein was cleaved into HA1 and HA2 polypeptides (Fig. 6), confirming that newly synthesized wild-type HA is rapidly and quantitatively converted into correctly folded trimers.

In summary, newly synthesized wild-type HA is rapidly and efficiently transported along the secretory pathway to the cell surface. Before their assembly into trimers in the ER, monomers are present in at least two distinct disulfide bonding patterns. However, both core-glycosylated and terminally glycosylated trimeric molecules have only a single disulfide bonding configuration.

#### *Analysis of the Folding and Intracellular Transport of Mutants of the 4-462 Disulfide Bond*

All of the mutants of the 4-462 disulfide bond display the same transport behavior, which is quite distinct from that of the wild-type protein. A significant proportion of each of the three mutants remain sensitive to endo H even after 6 h of chase (Fig. 2 Ba, b and c), suggesting that the majority of these polypeptides are retained in the ER. Acquisition of endo H resistance by a minority of the mutant polypeptides does not occur until after 1.5–3.0 h of chase. Densitometric measurements indicate that for each mutant, approximately the same percentage (10–20%) of the nascent polypeptides are transported and become terminally glycosylated. These terminally glycosylated forms of the mutant HA0 molecules migrate more diffusely and with higher molecular mass than the similarly processed wild-type HA polypeptides. Since the core-glycosylated, endo H-sensitive polypeptides migrate with the same mobility as the core-glycosylated wild-type HA0 species, the differences in migration are not due to alterations in the polypeptide backbone or in the number of glycosylation sites utilized. Rather, this less uniform and more extensive processing is due to an increased modification of some or all of the oligosaccharide moieties during terminal glycosylation. Hyperglycosylation has also been observed for other HA mutants (19), as well as for mutants of human chorionic gonadotropin in which abnormal addition of polylactosaminoglycan moieties has been demonstrated to occur (18). Whether a decreased rate of transport of mutants through the Golgi cisternae is the cause of the aberrant terminal glycosylation is not yet known.

The terminally glycosylated forms of the mutant HA proteins are efficiently transported to the cell surface (Fig. 2 Bd, e and f) since at every time point, they are quantitatively susceptible to cleavage by trypsin. In the absence of the 4-462 disulfide bond, the HA1 polypeptide does not remain associated with HA2, which is the only species precipitable from the cell lysates (Fig. 2 Bd, e and f). Protein fragments immunoprecipitated from the conditioned media of these monolayers (data not shown) are presumably breakdown products of the released HA1 polypeptide. No such fragments are released by equivalent treatment of cells expressing the wild-type protein. In addition, when the 3-h chase sample shown in Fig. 2 Bf was subjected to electrophoresis under nonreducing conditions, there was no alteration in the mobility of the HA2 species, indicating the absence of the disulfide-linked peptide that would remain after degradation of a disulfide-linked HA1 polypeptide. The ability of the



**Figure 4.** Separation of monomeric, trimeric, and aggregated forms of wild-type and mutant HA by sucrose gradient centrifugation. 36 h after infection with SV40-HA recombinant viruses, CV-1 cells were labeled with [<sup>35</sup>S]-methionine for 30 min. An aliquot of the cell extract, prepared in a lauryl maltoside lysis buffer, was fractionated by velocity centrifugation on a 5–25% sucrose gradient (see Materials and Methods). Aliquots of each fraction were immunoprecipitated with  $\alpha$ HA alone or with HA followed by immunoprecipitated with ricin B chains conjugated to Sepharose beads (A, section c). The immunoprecipitated proteins were analyzed by SDS-PAGE under reducing (+DTT) or nonreducing (-DTT) conditions and visualized by autoradiography. The stacking portions of the gels are present in the autoradiographs.



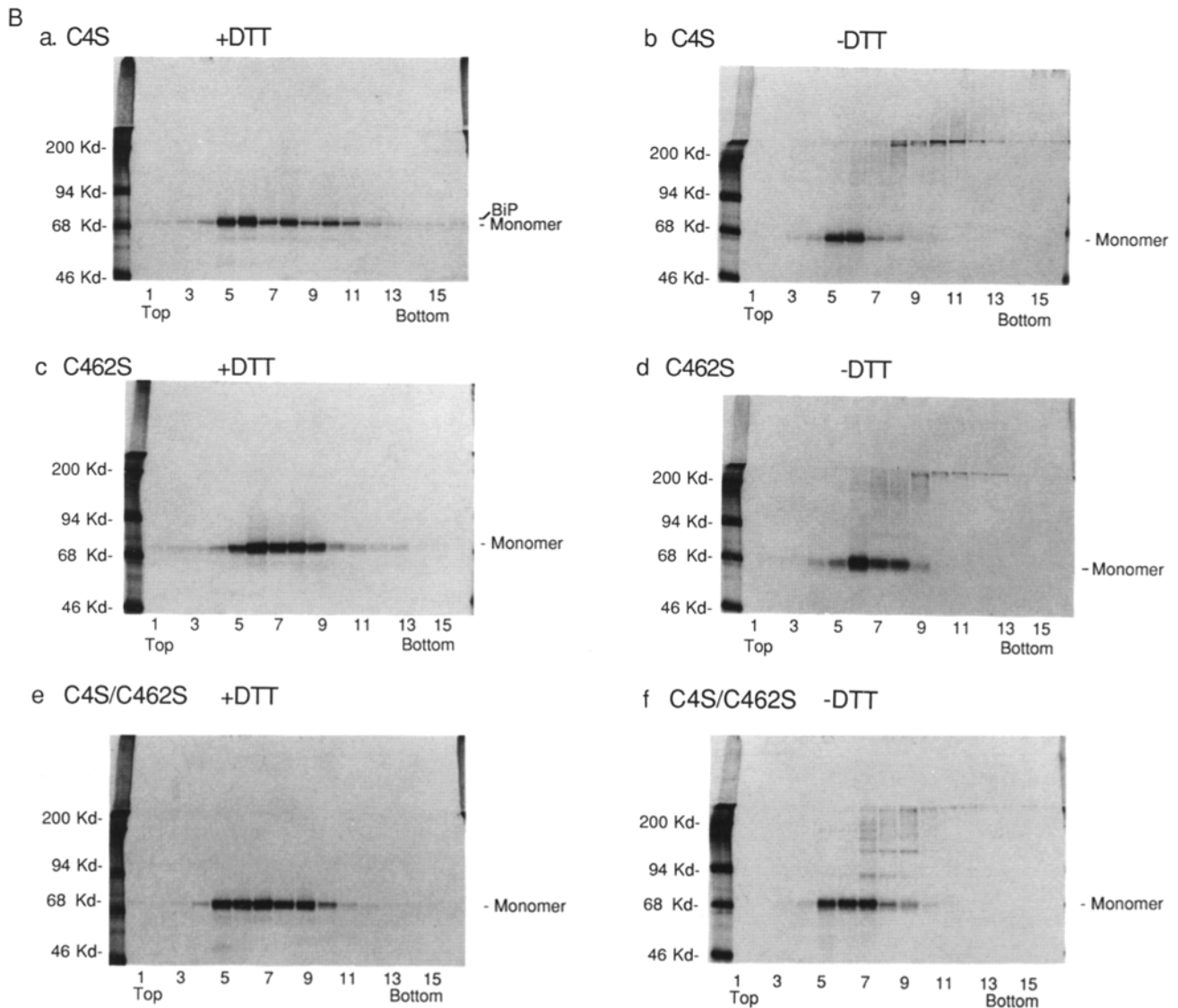


Figure 4.

HA1 chains of these mutants to dissociate from HA2 indicates that no other intermolecular disulfide bond is formed in lieu of the 4-462 bond. Thus, the Cys<sub>4</sub> residue, present in the C462S mutant, does not form disulfide bonds with the other cysteines of HA2 present near the base of the molecule (see Fig. 1 A). Both populations of mutant proteins, i.e., the fraction that is transported to the cell surface as well as the larger fraction that remains in the ER, appear to have a half-life of ~6 h (Fig. 2 B), somewhat shorter than the half-life of >12 h observed for wild-type HA.

The autoradiographs shown in Fig. 2 B also reveal a protein species migrating at ~77 kD that coprecipitates with the HA protein. This band has previously been identified as the Ig heavy chain binding protein BiP (5) and can be seen in the immunoprecipitations of all of the mutant proteins discussed in this paper. The significance of coprecipitation of BiP with these mutants will be discussed later.

Immunofluorescence studies of intact cells expressing the mutants of the 4-462 series using  $\alpha$ HA antibodies (Fig. 3) confirmed that at least a proportion of the mutant proteins

were transported to the cell surface. However, these cell surface molecules were in a conformation distinct from that of the wild-type HA protein since they were recognized by the  $\alpha$ UHA antibody that recognizes unfolded forms of HA but not by the  $\alpha$ NHA antibody specific for the correctly folded protein.

The sucrose density gradient sedimentation pattern of mutant polypeptides of the 4-462 series synthesized during a 30-min labeling period (Fig. 4 B) was quite distinct from that of wild-type HA. However, the three mutants behaved in a qualitatively similar manner. Analyses performed under nonreducing conditions (Fig. 4 B, b, d, and f) demonstrated that the monomeric or dissociated HA molecules in all fractions migrate as a single species, suggesting that the mutant HA monomers all have the same configuration of intramolecular disulfide bonds. As discussed earlier, monomeric wild-type HA present in the less dense portion of the gradient migrated as a doublet (Fig. 4 A, b, fractions 3-6). Taken together, these results suggest that the two configurations of disulfide bonds present in monomeric wild-type HA

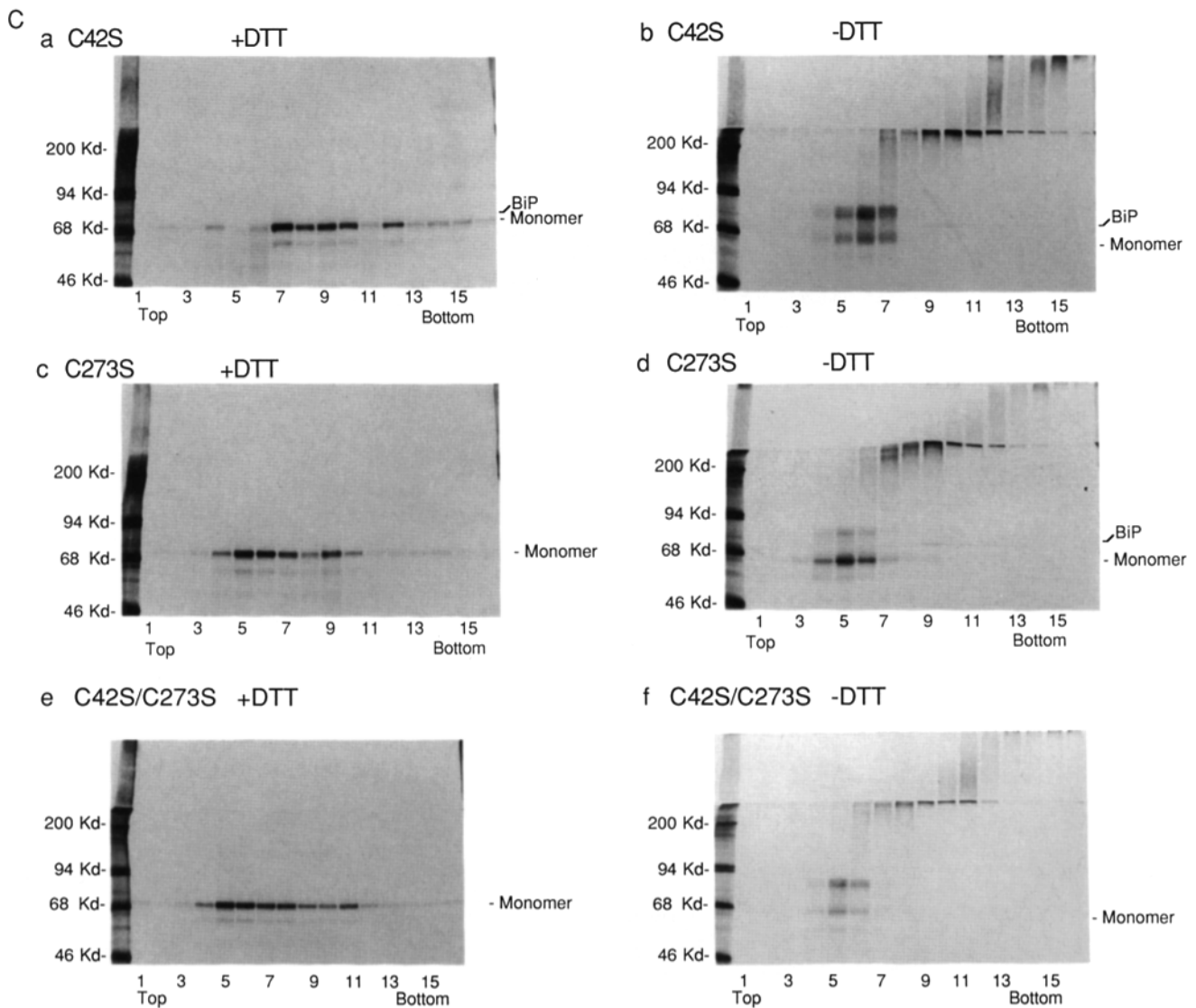
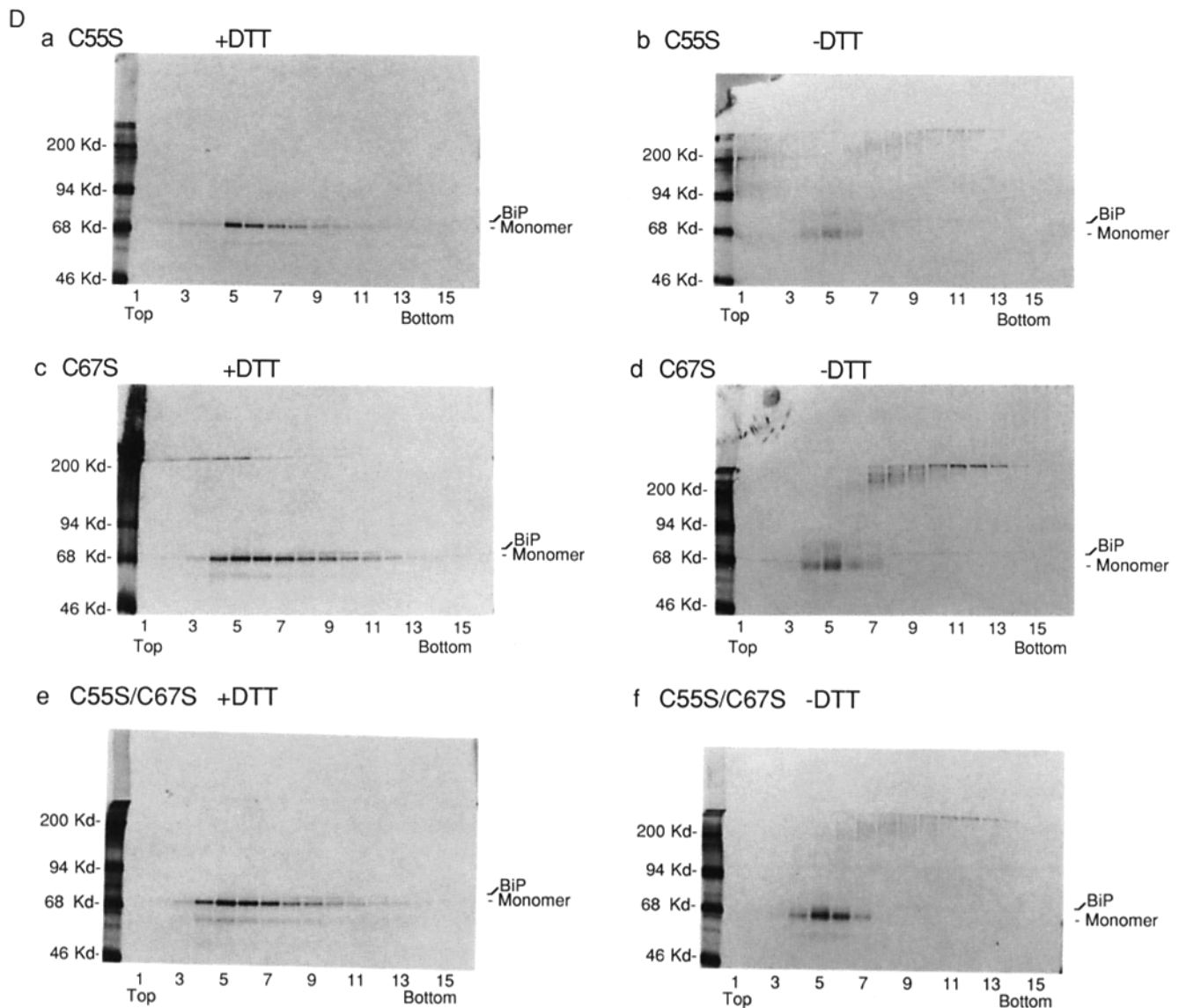


Figure 4.

molecules may differ by the absence or presence of the 4-462 disulfide bond.

Mutant proteins of the 4-462 series are unique in that fractions 6-9 from the middle of the gradient contain an array of HA species that migrate under nonreducing conditions with molecular masses that vary between 68 and 200 kD (Fig. 4 B, b, d, and f). To determine whether any of these bands represent aggregates of proteolytic breakdown products, the labeling of C4S/C462S and subsequent sucrose-gradient analysis was repeated with the inclusion of protease inhibitors in the lysis buffer and sucrose gradient solutions (data not shown). The bands between  $\sim 70$  and  $\sim 140$  kD were no longer present indicating that they were likely to have been the result of proteolysis. However, a number of HA species that migrate with molecular masses between  $\sim 140$  and  $\sim 200$  kD were still present in fractions 6-9. The molecular masses of these species are intermediate between the molecular mass observed for the HA dimer and trimer suggesting that they may correspond to a heterogeneous population of mutant HA dimers and trimers containing a variety of disulfide bond configurations. Finally, significant

amounts of the mutant proteins sedimented in the denser regions of the gradients (Fig. 4 B, a-f, fractions 11-16) where little or no wild-type protein was found (Fig. 4 A, a and b). The mutant polypeptides contained in these dense fractions, when analyzed under nonreducing conditions (Fig. 4 B, b, d, and f), were resistant to dissociation by SDS and migrated more slowly during SDS-PAGE than do wild-type HA trimers, indicating that the HA molecules are present in aggregates that may be covalently linked by intermolecular disulfide bonds. The aggregates observed are not an artefact of oxidation following detergent solubilization because their occurrence was not prevented by the inclusion of iodoacetamide in the lysis buffer. Analysis of the time course of oligomerization and aggregation of the 4-462 mutants (Fig. 5 B) indicated that SDS-resistant, high molecular mass species were present as early as 3-6 min after synthesis. Over a period of 3 h the majority of the labeled molecules became incorporated into these aggregates. A distinct species of molecular mass  $\sim 130$  kD is present in all lanes in Fig. 5 B, and can also be seen, albeit less distinctly, in the autoradiographs shown in Fig. 4 B, b, d, and f. The size of this species is not



**Figure 4.**

the same as that of dimeric wild-type HA (Fig. 5 A). We tentatively identified these molecules as SDS-resistant heterodimers of mutant HA and BiP polypeptides, but have so far been unable to obtain unequivocal evidence for this assignment.

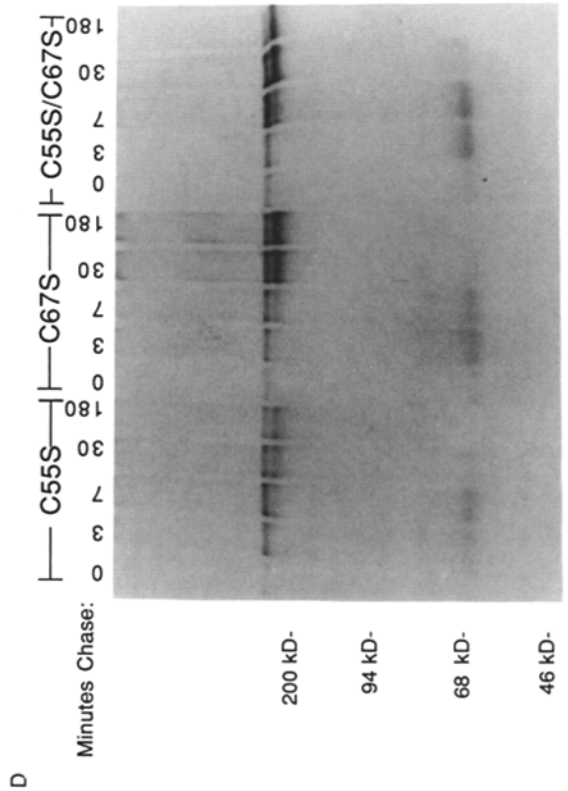
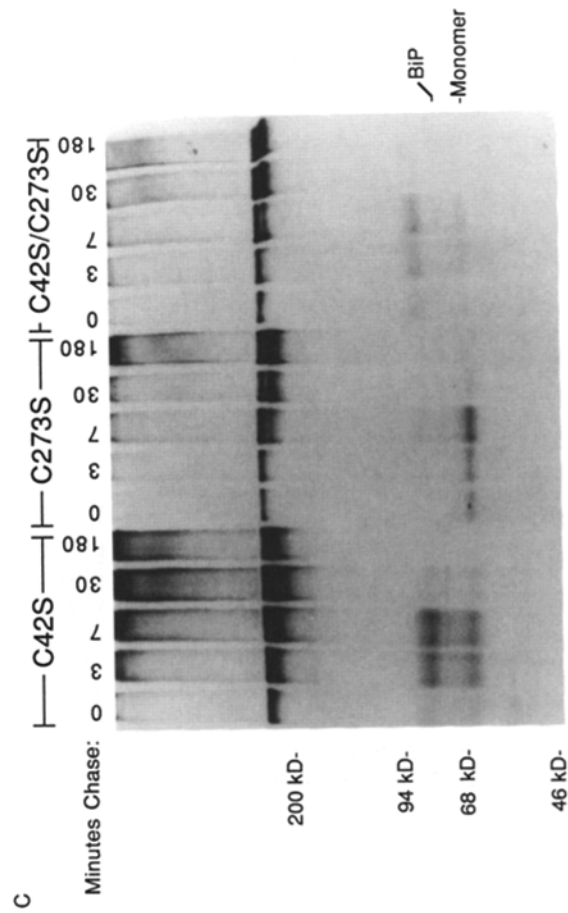
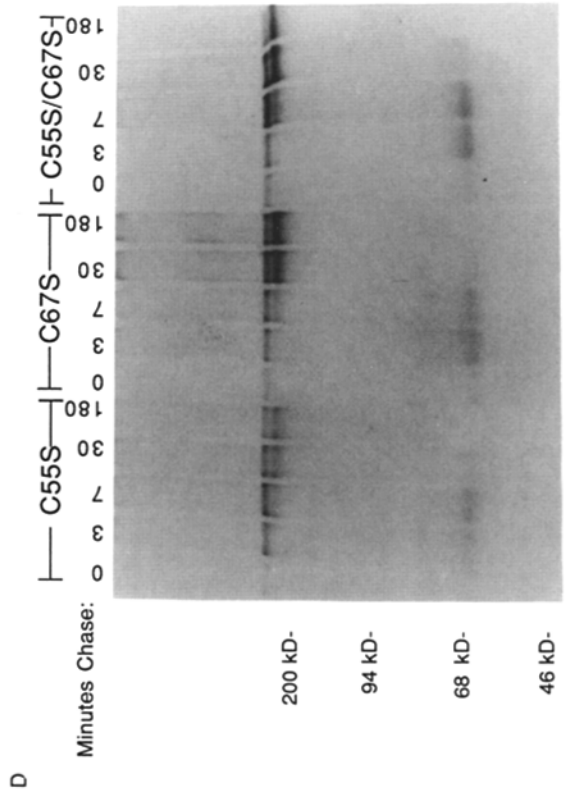
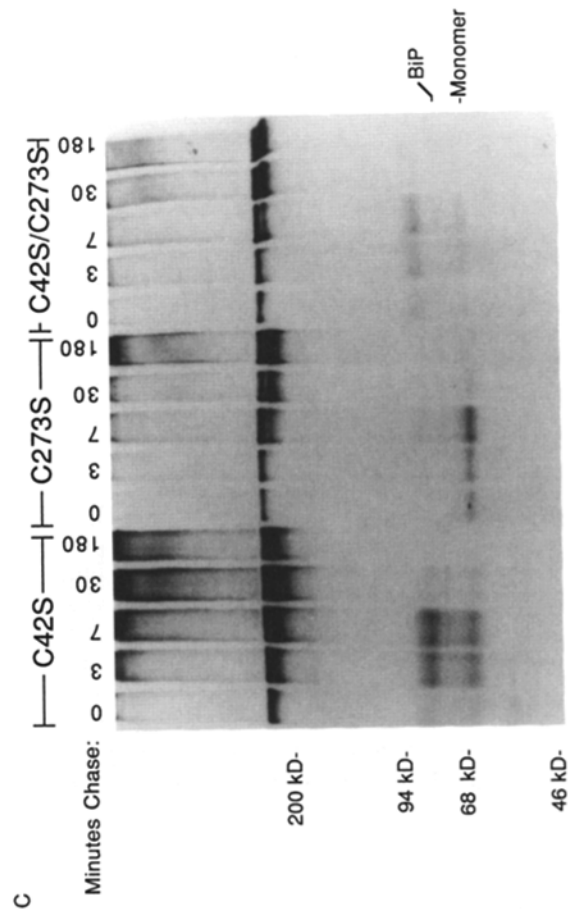
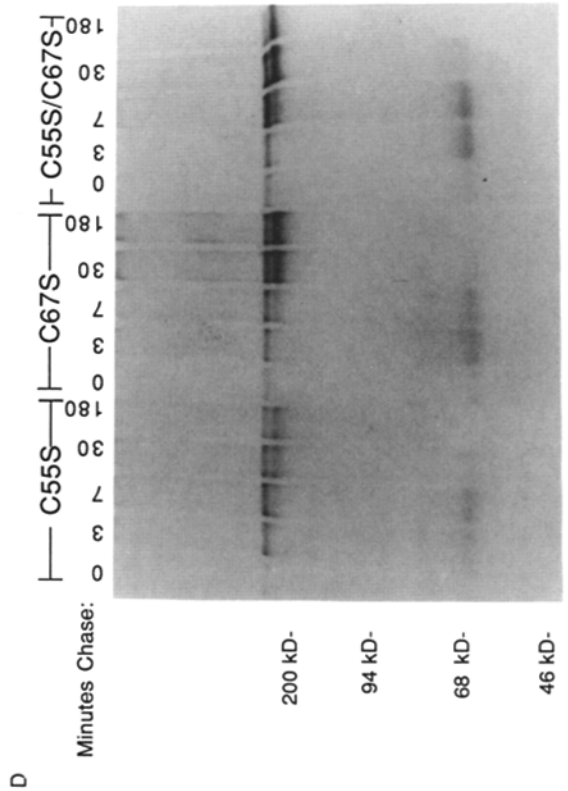
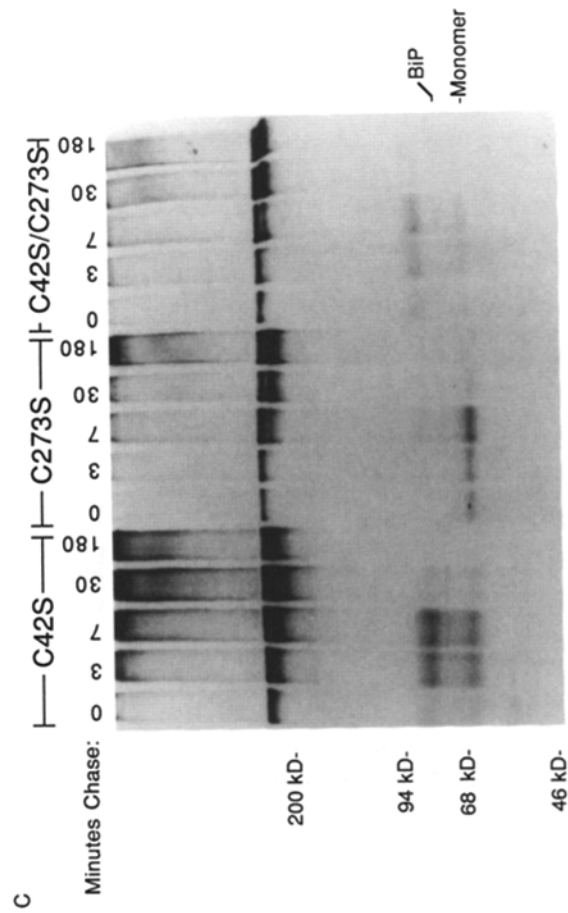
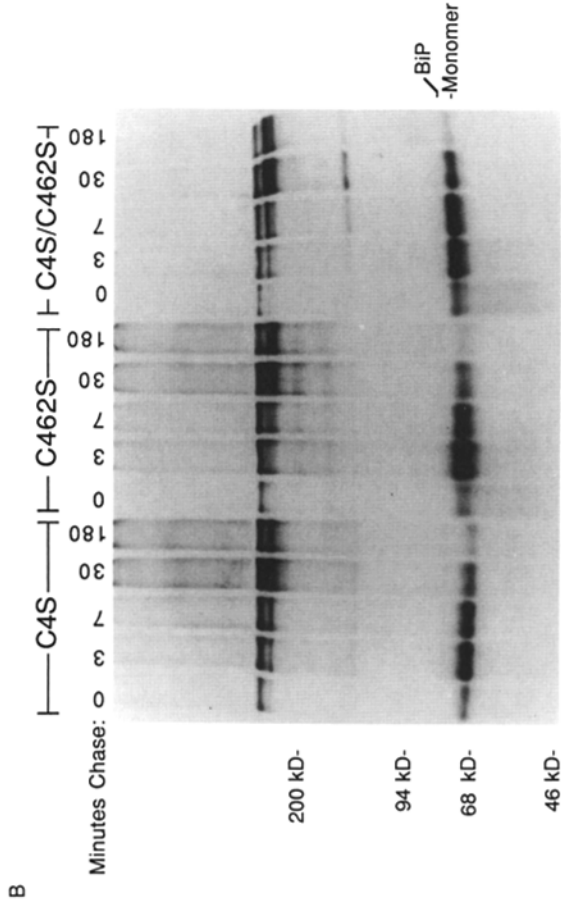
Finally, analysis of the protease sensitivity of the mutants of the 4-462 series (Fig. 6) showed that molecules synthesized during a 2-h pulse with [<sup>35</sup>S]methionine were more susceptible to degradation by trypsin than the wild-type protein. All three mutants showed similar patterns of proteolytic fragmentation, suggesting that these three proteins have a similar conformation. The pattern of digestion of each of the 4-462 mutants, as well as that of all the other serine substitution mutants (also shown in Fig. 6) suggests that in every case the HA1 polypeptide is profoundly more sensitive to protease than the HA2 polypeptide. These results may indicate that all of the mutant molecules contain a malformed globular domain and a correctly folded stalk domain. However, a more likely explanation is that due to the amphipathic nature of the  $\alpha$ -helices that are formed from a significant portion of the HA2 chain (56), the mutant proteins form non-native

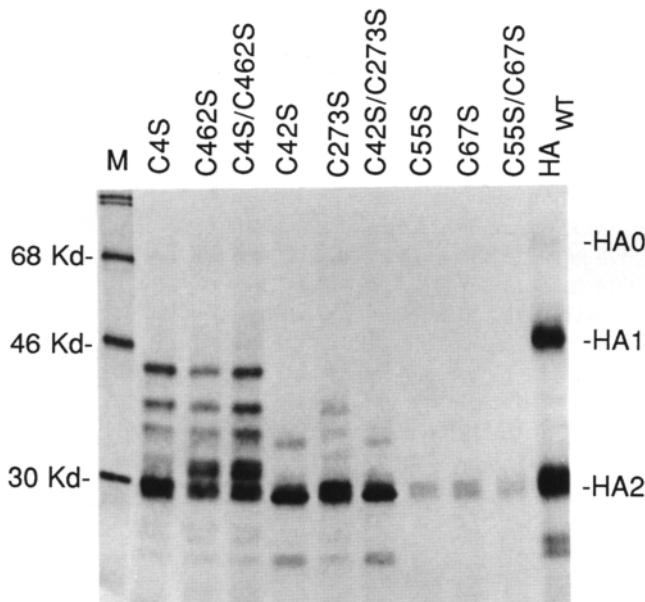
structures involving inter- or intra-molecular hydrophobic interactions (or similar interactions with BiP—see below) that shield protease-sensitive sites in the HA2 polypeptide.

In summary, the results of the analyses of the 4-462 series of mutants suggest that: (a) all three mutants show the same, mostly aberrant, pattern of folding; (b) unlike the wild-type HA protein, the mutant monomers may assume only one configuration of disulfide bonds; (c) the majority of the mutant monomers (>80–90%) rapidly aggregate to form large oligomers that remain in the ER; the disulfide bonds in these aberrant oligomers may shuffle into new intra- and intermolecular configurations; and (d) a minority of the mutant monomers slowly assemble into forms that are transported along the exocytic pathway to the cell surface; although these molecules are probably trimeric, they have an altered conformation and are less stable than the wild-type protein.

#### ***Analysis of the Folding and Intracellular Transport of Mutants of the 42-273 Disulfide Bonds***

The single and double substitution mutants of the cysteines





**Figure 6.** Protease sensitivity of wild-type and mutant HA proteins. Simian CV-1 cells infected with the SV40-HA recombinant vectors were labeled with [<sup>35</sup>S]methionine for 2 h and chased with an excess of nonradioactive methionine for 2 h. At the end of the chase cell lysates were prepared. Trypsin was added to a final concentration of 100 μg/ml to aliquots of the cell extracts which were then incubated on ice for 30 min. After addition of soybean trypsin inhibitor to a final concentration of 200 μg/ml, the HA polypeptides were immunoprecipitated, analyzed by SDS-PAGE, and visualized by autoradiography. The two major bands in the HA<sub>WT</sub> lane represents HA1 (~46 kD) and HA2 (~30 kD); the lower bands are unidentified minor proteolytic products.

involved in the 42-273 disulfide bond displayed transport phenotypes similar to those of the mutants of the 4-462 series, except that the 42-273 mutants were more defective in assembly and transport. Again, all three mutants showed very similar behavior, indicating that the loss of both partners of the 42-273 disulfide bond was neither more problematic nor more advantageous than losing either of the component cysteine residues.

The results obtained in the transport studies are summarized in Table I. The majority of the newly synthesized mutant polypeptides remained in the ER and were not transported. However, a small proportion (~10%) was transported to the Golgi apparatus and became resistant to endo H after 3 h of chase. These terminally glycosylated proteins were susceptible to exogenously added trypsin, indicating that all the mutant polypeptides that reached the medial cisternae of the Golgi apparatus were transported to the cell surface. Immunofluorescence studies such as those illustrated in Fig. 3 yielded results identical to those obtained with the mutants of the 4-462 series, suggesting that the mutant proteins that

**Table I. Summary of the Transport Characteristics of Wild-type and Mutant HAs**

	After 1.5 h of Chase		Maximum % of HA transported from ER <sup>‡</sup>
	% Endo H resistant	% Transported to cell surface*	
JHA	>90	>90	>90
C4S	~10	~10	~20
C462S	~10	~10	~20
C4S/C462S	~10	~10	~20
C42S	~5	~5	~10
C273S	~5	~5	~10
C42S/C273S	~5	~5	~10
C55S	<5	<5	<5
C67S	<5	<5	<5
C55S/C67S	~25	~25	~25

\* As determined by the ability of exogenous trypsin to cleave labeled HA molecules on the surface on intact cells.

‡ As determined by maximum percentage of labeled HA achieving resistance to endo H and/or becoming sensitive to exogenous trypsin over a 12-h time course.

reached the cell surface were in a conformation different from that of wild-type HA.

The behavior during sucrose gradient fractionation of mutant polypeptides of the 42-273 series synthesized during a 30-min labeling period is shown in Fig. 4 C. The three mutants showed very similar patterns of sedimentation which again differed significantly from that observed for the wild-type protein. The mutant proteins were detected throughout the gradients (Fig. 4 C, a-f, fractions 2-16), indicating the presence of large aggregates in the denser fractions where wild-type molecules are not found. When analyzed under reducing conditions (Fig. 4 C, a, c, and e) all the mutant HA molecules migrated predominantly as sharp bands corresponding to monomeric species with molecular masses of ~69 kD. A small amount of a proteolytic breakdown product (~60 kD) that has been observed previously with other mutants (14) is also present in each fraction that contains the full-length molecules. When analyzed under nonreducing conditions, the three mutants showed interesting differences in their migration patterns. In every case, the mutant HA molecules present in the less dense fractions (4-7) migrated as three or more species when analyzed by SDS-PAGE under nonreducing conditions (Fig. 4 C, b, d, and f). These species do not differ as the result of proteolytic degradation because they migrate as a single band under reducing conditions. Therefore, the monomeric forms of the mutant polypeptides exist in at least three distinct disulfide-bonded configurations. As with the wild-type protein (see above), two of these configurations may correspond to molecules that have either formed or still lack the disulfide bond between cysteines 4 and 462, which are unaltered in mutants of the 42-273 series.

**Figure 5.** Analysis of the time course of oligomerization of wild-type and mutant HA proteins. 36 h after infection with SV40-HA recombinant viruses, CV-1 cells were pulse-labeled with [<sup>35</sup>S]methionine for 3 min and chased for the indicated time in an excess of nonradioactive methionine. Cell extracts were then prepared using lauryl maltoside as the detergent and the HA polypeptides were immunoprecipitated with αHA, analyzed by SDS-PAGE under nonreducing conditions and visualized by autoradiography. The stacking portions of the gels are present in the autoradiographs.

Monomers of the C42S mutant migrate predominantly as two widely spaced doublets with all four bands having approximately equal intensities (Fig. 4 C, b), while C273S monomers migrate as at least three species, two of high and one of low mobility, of which the most rapidly migrating is the most populated (Fig. 4 C, d). Finally, the C42S/C273S double substitution mutant also migrates as three species but their distribution differs from that seen with the C273S mutant in that only one high mobility species is present and the intensities of the bands are approximately equal (Fig. 4 C, f). It should be noted that the electrophoretic mobilities of the faster moving species are similar to those of the wild-type HA monomer doublet (Fig. 4 A, b), while the mobilities of the other species are very much slower than that of either band of the HA doublet.

As was observed for the mutants of the 4-462 disulfide bond, significant amounts of the 42-273 mutant proteins sedimented in the denser regions of the gradients (Fig. 4 C, a-f, fractions 11-16). The mutant polypeptides contained in these dense fractions, when analyzed under nonreducing conditions (Fig. 4 C, b, d, and f), were resistant to dissociation by SDS and migrated during SDS-PAGE as very high molecular mass aggregates. It is likely that the HA molecules within these aggregates are covalently linked by intermolecular disulfide bonds. Analysis of the time course of oligomerization and aggregation of these mutants (Fig. 5 C) indicated that SDS-resistant, high molecular mass species were present as early as 3 min after synthesis of the mutant proteins. Over a period of 30 min the majority of the labeled molecules became incorporated into the aggregates.

Analysis of the protease sensitivity of the mutants of the 42-273 series (Fig. 6) showed that molecules synthesized during a 2-h pulse with [<sup>35</sup>S] methionine were much more susceptible to degradation by trypsin than the wild-type protein. Mutants C42S and C42S/C273S showed similar patterns of proteolytic fragmentation, suggesting that these two proteins have a similar conformation. However, mutant C273S appears to be slightly more resistant to protease digestion, an observation that may correlate with the increased proportion of more fully disulfide-bonded monomers observed for this mutant (see above).

In summary, the results of the analyses of the 42-273 series of mutants suggest that: (a) monomers assume at least three distinct disulfide bond configurations, two of which probably correspond to molecules that have either formed or still lack the 4-462 disulfide bond; (b) the three mutants are equally and severely defective in assembly and transport with <10% of the molecules reaching the cell surface; and (c) the great majority of the newly synthesized polypeptides rapidly aggregate into large oligomers that are retained in the ER; the disulfide bonds present in these aberrant oligomers may shuffle into new intra- and intermolecular configurations.

#### ***Analysis of the Folding and Intracellular Transport of Mutants of the 55-67 Disulfide Bond***

Unlike the two sets of mutants just described, the double substitution mutant of the 55-67 disulfide bond has a transport phenotype distinct from that of the single substitution mutants. These mutants, C55S and C67S, never attain resistance to endo H even after 3 h of chase (Fig. 2 C, a and b). However, ~20-30% of the radiolabeled C55S/C67S polypeptides acquire resistance to the glycosidase during the

chase period (Fig. 2 C, c). These C55S/C67S polypeptides are transported from the ER at a rate significantly slower than that observed for wild-type HA, since only after 1.5 h of chase do any labeled mutant polypeptides become resistant to endo H. These molecules migrate with the same mobility as terminally glycosylated wild-type HA, thus, differing from the other mutants in not undergoing more extensive terminal processing. Consistent with these results, C55S or C67S polypeptides never become accessible to trypsin added to the medium (Fig. 2 C, d and e). However, the terminally glycosylated C55S/C67S molecules present after 1.5 or more hours of chase are accessible to the protease at the cell surface and are cleaved into HA1 and HA2 polypeptides (Fig. 2 C, f). C55S, C67S, and C55S/C67S polypeptides remaining within the ER all have a relatively rapid turnover time of ~6 h. However, those C55S/C67S molecules that are transported to the cell surface appear to be as stable as the wild-type HA molecules, having a  $t_{1/2}$  of degradation of >12 h.

Although the C55S/C67S double mutant differs from the single substitution mutants in that a proportion of the newly synthesized polypeptides are transported along the secretory pathway, all three mutant polypeptides synthesized during a 30-min labeling period showed very similar patterns of sedimentation on sucrose gradients (Fig. 4 D, a-f). In each case, the mutant HA proteins are detected throughout the gradients in fractions 2-16, indicating the presence of very large aggregates in the denser fractions where wild-type HA molecules are not found (Fig. 4 A). When analyzed under reducing conditions (Fig. 4 D, a, c, and e), the mutant polypeptides in every fraction migrated predominantly as a sharp band of monomers with a molecular mass of ~69 kD. A small amount of the ~60-kD proteolytic breakdown product (discussed above) is also present in each fraction that contains the full-length molecules. However, distinct differences were observed between the three mutants when the proteins in the gradient fractions were analyzed by SDS-PAGE under nonreducing conditions (Fig. 4 D, b, d, and f). In the case of the C55S single substitution mutant, the HA monomers present in the less dense fractions (4-7) migrate predominantly as three species, including a more rapidly migrating doublet whose mobilities are similar to those of wild-type HA monomers (Fig. 4 D, b). C67S monomers also migrate as two or three widely-spaced bands (Fig. 4 D, d). By contrast, monomers of the C55S/C67S double substitution mutant resemble wild-type monomers in migrating only as a closely-spaced doublet, whose faster migrating band predominates. Thus it appears that HA molecules that lack both partners of the 55-67 disulfide bond initially fold into a structure very similar to that of wild-type HA monomers. However, molecules that lack only one or other of the partners fold into a variety of conformations containing different disulfide bond configurations.

As was observed for the mutants of the other disulfide bonds, significant amounts of the 55-67 mutant proteins sedimented in the denser regions of the gradients (Fig. 4 D, a-f, fractions 11-16). The mutant polypeptides contained in these dense fractions, when analyzed under nonreducing conditions (Fig. 4 D, b, d, and f), were resistant to denaturation by SDS and migrated during SDS-PAGE as very high molecular mass aggregates. It is likely that the HA molecules within these aggregates are covalently linked by intermolecular disulfide bonds. Analysis of the time course of oligomer-



ization and aggregation (Fig. 5 C) indicated that SDS-resistant, high molecular mass species were present as early as 3 min after synthesis of the mutant proteins. Over a period of 30 min the majority of the labeled molecules became incorporated into these aggregates.

Proteolytic digestion of the mutants of the 55–67 series yields only two residual fragments that migrate as a doublet with molecular masses similar to that of wild-type HA2 (Fig. 6). Although this doublet probably represents the HA2 polypeptide and a proteolytically clipped HA2, we cannot rule out the possibility that one or both bands are a proteolytic product of the HA1 polypeptide. Upon much longer exposure of this autoradiograph there is a faint band, visible only in the C55S/C67S lane, that represents a protein that migrates with the molecular mass of HA1 (result not shown). This may represent the small portion of C55S/C67S that folds correctly and is transported along the exocytic pathway.

In summary, the results of the analyses of the 55–67 series of mutants suggest that: (a) unlike the mutants of the 4–462 and 42–273 disulfide bonds, the double substitution mutant C55S/C67S displays a folding and transport phenotype distinct from those of the single substitution mutants; (b) the majority of C55S/C67S molecules initially fold to form monomers with structures similar to that of the wild-type HA monomer; by contrast, the C55S and C67S mutants initially fold to form monomers having a variety of disulfide bonding patterns; (c) all three mutant proteins show a high propensity to aggregate into large oligomers that are retained in the ER and undergo disulfide rearrangements, however, only the C55S/C67S mutant also assembles (albeit with relatively low efficiency) into transport-competent molecules that move along the secretory pathway to the cell surface.

#### **Interaction of the ER Protein BiP with the Cysteine Substitution Mutants**

The ER-resident protein BiP interacts transiently with a variety of nascent wild-type exocytotic proteins (4, 5, 13, 37, 41), including HA (21), and more permanently with misfolded or unassembled proteins whose transport from the ER is blocked (5, 25, 30, 37). We have previously reported the association of BiP with a large number of substitution and deletion mutants of HA (19, 25) and have in addition shown that the accumulation in the ER of mutants such as C67S (previously known as HAJ567) causes the induction of synthesis of BiP mRNA and protein (34).

The autoradiographs presented in Figs. 2, 4, and 5 (B–D) show that a protein species of 77 kD, identified as BiP by immunoprecipitation studies (not shown), coprecipitates with the mutant HA proteins. The intensity of the sharp bands corresponding to BiP varies between different experiments, and the bands are sometimes partly or fully obscured by the broader bands of HA molecules. Nevertheless, all nine cysteine substitution mutants can be observed to interact stably with the ER protein.

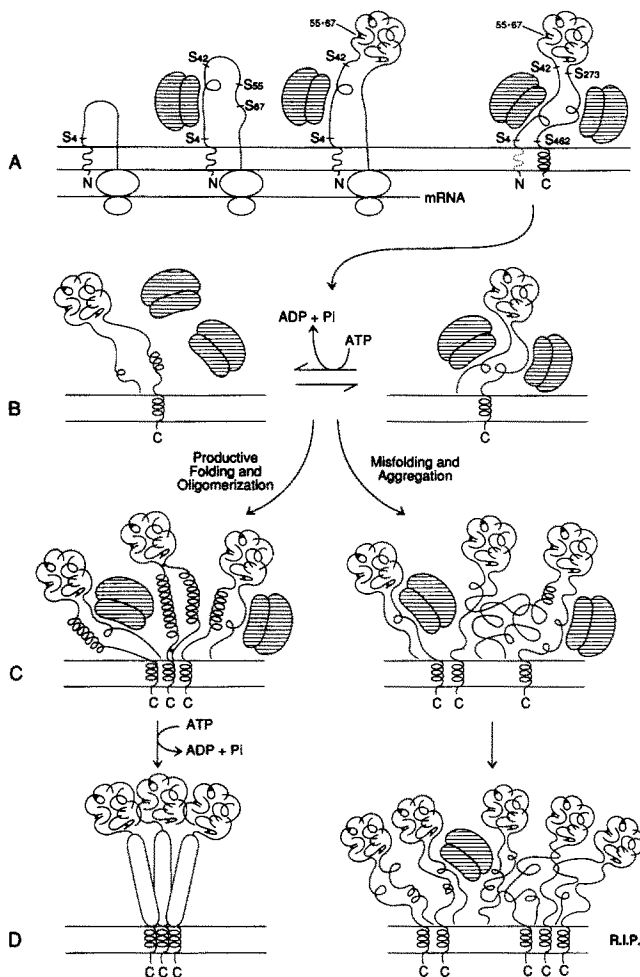
Interestingly, the anti-HA immunoprecipitations carried out after sucrose gradient fractionation demonstrate that BiP coprecipitates with mutant HA molecules present throughout the gradients, indicating that it associates both with monomeric HA molecules and with HA oligomers and aggregates (see for example Fig. 4 D, c). Electrophoretic analyses performed under nonreducing conditions reveal bands corresponding to monomeric BiP in immunoprecipitates

from the denser gradient fractions (9–13), although the mutant HA molecules present in these same precipitates are detected only as large aggregates (Fig. 4 D, b, d, and f). Thus, the BiP molecules are not disulfide bonded to the aggregated polypeptides and the association between BiP and HA molecules is sensitive at high temperature to disruption by SDS. This result confirms previous observations that BiP is non-covalently associated with large aggregates of misfolded X31 HA (30). Our view of the role played by BiP during its interaction(s) with HA is discussed below and illustrated in Fig. 7.

#### **Discussion**

During the folding of the wild-type HA molecule, individual cysteine residues may be required to transiently form non-native disulfide bonds; alternatively, these residues may normally pair productively only with a single partner forming bonds that stabilize native conformations both in folding intermediates and in the final protein structure. Our investigations of the effect of substitution by serine of six cysteines of the HA ectodomain suggest that the major effect of the loss of a disulfide bond through mutation of both of the component cysteines is to destabilize the normal conformation of the protein, leading to increased protein misfolding. Decreased stability of native protein structures has previously been observed during *in vitro* folding studies using partially reduced and carboxymethylated proteins lacking one or more disulfide bonds (43). The three disulfide bonds studied in this work differ in their apparent importance in stabilizing the native structure of the HA molecule. Thus the C55S/C67S double mutant, which lacks the bond that normally closes a small loop of 11 amino acids on the surface of the globular domain, has the least defective phenotype with ~30% of newly synthesized molecules folding into transport-competent structures. The C42S/C273S and C4S/C462S double mutants, which lack bonds normally closing loops of 230 amino acids (the globular domain) and 457 amino acids (almost the whole ectodomain), are significantly more defective with only 10–20% of molecules being transport-competent. When only one of the usual partners of a bond is present in a mutant HA polypeptide, the folding defect due to loss of the bond may be exacerbated by formation of intramolecular disulfide linkages that do not normally occur in the wild-type protein. In all cases, the misfolded mutant HA proteins are retained in the ER and rapidly aggregate to form large oligomers in which the intramolecular disulfide bonds may shuffle into new intramolecular configurations. However, the formation of disulfide-bonded aggregates is not a direct consequence of alteration of the normal complement of cysteine residues in the HA molecules because the accumulation of similar cross-linked aggregates has been observed when the number or location of N-linked oligosaccharide side chains is altered on HA (30) or on the G protein of vesicular stomatitis virus (36).

We obtained no evidence for an obligatory role for the formation of non-native disulfide bonds during the folding of HA *in vivo*. Nevertheless, such bonds do form in some, but not all, of the mutant proteins. As discussed earlier, the six cysteine residues studied in this work differ in the opportunities available for them to form non-native bonds during folding of the wild-type HA polypeptide. Thus Cys<sub>4</sub> would have



**Figure 7.** Proposed role of BiP in prevention of non-native disulfide bond formation. (Line A) Biosynthesis of HA occurs on membrane-bound polysomes, with membrane translocation being initiated by interaction of the NH<sub>2</sub>-terminal signal peptide (wavy line) with the membrane of the ER. As the polypeptide is transferred into the lumen of the ER, BiP (cross-hatched) binds to sequence(s) within the NH<sub>2</sub>-terminal 50 amino acids, isolating them from interacting with the next ~230 residues, which will fold to form the globular domain even before the synthesis and translocation of the polypeptide chain has been completed. The disulfide bond between Cys<sub>55</sub> and Cys<sub>67</sub> forms as the globular domain folds. The COOH-terminal sequences that, together with the NH<sub>2</sub>-terminal 50 amino acids, will form the stem domain will not take up their final conformation before the monomers assemble into trimers, and additional BiP binding site(s) will be exposed. (Line B) Release of BiP molecules from their binding sites on HA is promoted by hydrolysis by BiP of ATP (47). If other HA molecules are not available for oligomeric assembly, BiP, which is present at high concentration in the ER lumen, may rebind to the HA monomer and continue to stabilize the unfolded protein until another cycle of ATP hydrolysis again frees the HA molecule to continue the folding process. The Cys<sub>4</sub>-Cys<sub>462</sub> and Cys<sub>42</sub>-Cys<sub>273</sub> disulfide bonds may form in the monomer following BiP release, but quantitative formation of these bonds does not occur before trimerization. (Line C) If other HA monomers are available for assembly, release of HA from BiP allows either (a) the continuation of productive folding and oligomerization, or (b) misfolding and aggregation. In either situation, individual BiP binding sites may become inaccessible within the folded or misfolded structures. For the wild-type protein, productive folding is very efficient with >95% of the molecules being incorporated into native structures. Mutant HA molecules (or wild-type mole-

the opportunity to interact with eight other cysteine residues in the HA sequence before its final partner, Cys<sub>462</sub>, is translocated into the ER lumen. However, our results suggest that Cys<sub>4</sub> does not form any alternative disulfide pairings in newly synthesized mutant monomers or in the assembled molecules that are transported to the cell surface. Nor is it required to do so, because mutants lacking Cys<sub>4</sub> are no more defective in folding and transport than are the mutants lacking Cys<sub>462</sub>. Similarly, Cys<sub>42</sub> has the opportunity to interact with four other cysteine residues before its partner, Cys<sub>273</sub>, emerges into the ER lumen. Like Cys<sub>4</sub>, Cys<sub>42</sub> appears not to be required to form intermediate, non-native disulfide bonds since the presence or absence of this residue had little effect on the extent of folding and transport of the molecule. Nevertheless, mutants lacking either Cys<sub>4</sub> or Cys<sub>42</sub>, or both, appear eventually to form significant numbers of non-native intramolecular disulfide bonds. Finally, because Cys<sub>55</sub> and Cys<sub>67</sub> are very close in the polypeptide chain and are not separated by any other sulfhydryl residues, the bond between them seems likely to form rapidly and directly. Formation of non-native disulfide bond intermediates involving these residues can not be obligatory because the double mutant lacking both Cys<sub>55</sub> and Cys<sub>67</sub> was the least defective of all the cysteine substitution mutants. Both single substitution mutants were equivalently and severely defective in folding and transport suggesting that in the absence of their correct partners Cys<sub>55</sub> and Cys<sub>67</sub> form aberrant crosslinks that interfere with folding and assembly.

It is possible that because the Cys<sub>55</sub>-Cys<sub>67</sub> disulfide bond is normally formed so rapidly during folding of the HA monomer, it has not been necessary for the molecule to evolve so as to avoid formation of stable intermediates having non-native bonds involving these residues. However, Cys<sub>4</sub> and Cys<sub>42</sub> cannot immediately form a covalent linkage with their native partners, and yet do not appear to form bonds with other cysteines during the translocation and folding of the remainder of the molecule. This suggests that the NH<sub>2</sub>-terminal 50 residues of HA, which will eventually fold with residues 278 to 520 to form the stalk domain of the molecule, do not interact with residues 51-277 as they fold to form the globular domain. It seems that a mechanism must exist to isolate this region during the folding of the globular domain and, in consequence, prevent the formation of non-native disulfide bonds. A likely candidate for this role is the ER-resident protein BiP, which interacts transiently with wild-type HA subunits before the final formation of the correctly folded trimer (21).

As illustrated in Fig. 7, we propose that BiP binds to sequence(s) within the NH<sub>2</sub>-terminal 50 residues of the translocating HA polypeptide and shields this region while the remainder of the molecule is synthesized and transferred into

cules that are aberrantly glycosylated [19, 30]) misfold with a frequency that reflects on the importance of the altered component in the folding process. (Line D) Productive folding is completed and the native HA trimers can leave the ER and move along the secretory pathway. Misfolded molecules, which are retained in the ER, continue to aggregate and intramolecular disulfide bonds may form. BiP will continue to associate with these molecules as long as binding sites are exposed on the surface of the aggregate.

the ER lumen. In a separate study of deletion mutants of X31 HA we have shown that BiP binds with high affinity to sequences within the stalk domain, including at least one site within the NH<sub>2</sub>-terminal 48 residues of the protein, while it does not appear to bind stably to sequences that form the globular domain (M. S. Segal, J. F. Sambrook, and M. J. Gething, manuscript in preparation). Furthermore, in vitro experiments a synthetic peptide corresponding to a sequence from the NH<sub>2</sub>-terminus of X31 HA that includes the cysteine analogous to Cys<sub>4</sub> of Japan HA efficiently elicits the peptide-dependent ATPase activity (16) of purified murine BiP (Sylvie Blond-Elguindi, J. F. Sambrook and M. J. Gething, unpublished results). Thus, the binding of BiP to the NH<sub>2</sub>-terminal region of Japan HA may prevent Cys<sub>4</sub> and Cys<sub>42</sub> from forming a disulfide bond between themselves or with other cysteines as they are translocated into the lumen of the ER. We have at present no reason to believe that the cysteine residues themselves form part of the recognition sequence(s) for BiP binding.

In order for the native Cys<sub>4</sub>-Cys<sub>462</sub> and Cys<sub>42</sub>-Cys<sub>273</sub> disulfide bonds to form, BiP would have to be displaced, presumably as a natural consequence of cycles of ATP hydrolysis by BiP (16, 35, 40, 47). Rothman (47) has suggested that if the rate of ATP hydrolysis by BiP in vitro (~0.2 mol·min<sup>-1</sup> per mol of BiP) is similar to that in vivo, BiP would remain bound to its polypeptide substrates for ~5 min. This period of binding would appear sufficient to prevent the formation of non-native disulfide bonds, since synthesis of an HA polypeptide takes ~2 min (6), and short enough to be a component of the process of folding and assembly of the wild-type HA molecule since the observed *t*<sub>1/2</sub> of trimerization of HA is ~7–10 min under the experimental conditions used in this work (25). Release of BiP from the translocated HA polypeptide would free the molecule to continue the folding process (Fig. 7). Productive folding of HA would occlude the binding sites for BiP, which presumably recognizes sequences that are inaccessible in the native structure (23). If, as a consequence of mutation(s) in the polypeptide, this process does not occur rapidly, BiP, which is present at high concentration in the ER lumen, is likely to bind again and remain associated with the polypeptide until a second cycle of ATP hydrolysis-mediated release occurs. During each interval of release the polypeptide would have the opportunity to fold correctly and oligomerize or to form nonproductive inter- or intramolecular interactions yielding misfolded or aggregated molecules. The nature of the sequence alteration in each mutant would determine the likelihood of the polypeptide becoming irreversibly misfolded. Such a scenario would explain the variations between the cysteine substitution mutants in the proportion of newly synthesized polypeptides that fold productively, and also explain the significant delay in the assembly and transport of these molecules. Cycles of binding, release and re-binding of BiP to the misfolded HA proteins would continue unless or until all the binding sites on the molecules become occluded in aggregated structures.

We would like to thank Patricia Gallagher, Karen McCammon, Laurie Roman, Steve Lacey, and Ed Madison for many helpful discussions. We also thank Neal Robinson for suggesting the use of the detergent lauryl maltoside for preparation of cell extracts for gradient centrifugation experiments.

Received for publication 9 January 1992 and in revised form 23 April 1992.

## References

- Air, G. M. 1981. Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus. *Proc. Natl. Acad. Sci. USA.* 78:7639–7643.
- Ash, J. F., D. Louvard, and S. J. Singer. 1977. Antibody-induced linkages of plasma membrane proteins to intracellular actomyosin-containing filaments in cultured fibroblasts. *Proc. Natl. Acad. Sci. USA.* 74:5584–5588.
- Bergman, L. W., and W. M. Kuehl. 1979. Formation of an intrachain disulfide bond on nascent immunoglobulin light chains. *J. Biol. Chem.* 254:8869–8876.
- Blount, P., and J. P. Merlie. 1991. BiP associates with newly synthesized subunits of the mouse muscle nicotinic receptor. *J. Cell Biol.* 113:1125–1132.
- Bole, D. G., L. M. Hendershot, and J. F. Kearney. 1986. Posttranslational association of immunoglobulin heavy chain binding protein with nascent heavy chains in nonsecreting and secreting hybridomas. *J. Cell Biol.* 102:1558–1566.
- Braakman, I., H. Hoover-Litty, K. Wagner, and A. Helenius. 1991. Folding of influenza hemagglutinin in the endoplasmic reticulum. *J. Cell Biol.* 114:401–411.
- Brown, J. R. 1976. Structural origins of mammalian albumin. *Fed. Proc.* 35:2141–2144.
- Copeland, C. S., R. W. Doms, E. M. Bolzau, R. G. Webster, and A. Helenius. 1986. Assembly of influenza hemagglutinin trimers and its role in intracellular transport. *J. Cell Biol.* 103:1179–1191.
- Copeland, C. S., K.-P. Zimmer, K. R. Wagner, G. A. Healey, I. Mellman, and A. Helenius. 1988. Folding, trimerization and transport are sequential events in the biogenesis of influenza virus hemagglutinin. *Cell.* 53:197–209.
- Creighton, T. E. 1988. Disulphide bonds and protein stability. *BioEssays.* 8:57–63.
- Creighton, T. E. 1990. Protein folding. *Biochem. J.* 270:1–16.
- Doms, R. W., and A. Helenius. 1986. Quaternary structure of influenza virus hemagglutinin after acid treatment. *J. Virol.* 60:833–839.
- Dorner, A. J., D. G. Bole, and R. J. Kaufman. 1987. The relationship of N-linked glycosylation and heavy chain-binding protein association with the secretion of glycoproteins. *J. Cell Biol.* 105:2665–2674.
- Doyle, C., M. G. Roth, J. Sambrook, and M. J. Gething. 1985. Mutations in the cytoplasmic domain of the influenza virus hemagglutinin affect different stages of intracellular transport. *J. Cell Biol.* 100:704–714.
- Doyle, C., J. Sambrook, and M. J. Gething. 1986. Analysis of progressive deletions of the transmembrane and cytoplasmic domains of influenza hemagglutinin. *J. Cell Biol.* 103:1193–1204.
- Flynn, G. C., T. G. Chappell, and J. E. Rothman. 1989. Peptide binding and release by proteins implicated as catalysts of protein assembly. *Science (Wash. DC).* 245:385–390.
- Freedman, R. B., N. J. Bulleid, H. C. Hawkins, and J. L. Paver. 1989. Role of protein disulphide-isomerase in the expression of native proteins. *Biochem. Soc. Symp.* 55:167–192.
- Fukuda, M., J.-L. Guan, and J. K. Rose. 1988. A membrane-anchored form but not the secretory form of human chorionic gonadotropin- $\alpha$  chain acquires polylactosaminoglycan. *J. Biol. Chem.* 263:5314–5318.
- Gallagher, P., J. Henneberry, I. Wilson, J. Sambrook, and M. J. Gething. 1988. Addition of carbohydrate side chains at novel sites on influenza virus hemagglutinin can modulate the folding, transport and activity of the molecule. *J. Cell Biol.* 107:2059–2073.
- Gething, M.-J., and J. Sambrook. 1982. Construction of influenza haemagglutinin genes that code for intracellular and secreted forms of the protein. *Nature (Lond.).* 300:598–603.
- Gething, M.-J., and J. Sambrook. 1989. Protein folding and intracellular transport: studies on influenza virus haemagglutinin. *Biochem. Soc. Symp.* 55:155–166.
- Gething, M.-J., and J. F. Sambrook. 1990. Transport and assembly processes in the endoplasmic reticulum. *Sem. Cell Biol.* 1:65–72.
- Gething, M.-J., and J. F. Sambrook. 1992. Protein folding in the cell. *Nature (Lond.).* 355:33–45.
- Gething, M.-J., J. Bye, J. J. Skehel, and M. Waterfield. 1980. Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from H2 and H3 strains elucidates antigenic shift and drift in human influenza virus. *Nature (Lond.).* 287:301–306.
- Gething, M.-J., K. McCammon, and J. Sambrook. 1986. Expression of wild-type and mutant forms of influenza hemagglutinin: the role of folding in intracellular transport. *Cell.* 46:939–950.
- Goldenberg, D. P. 1988. Kinetic analysis of the folding and unfolding of a mutant form of bovine pancreatic trypsin inhibitor lacking the cysteine-14 and -38 thiols. *Biochemistry.* 27:2481–2489.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166:557–580.
- Hearing, J., E. Hunter, L. Rodgers, M. J. Gething, and J. Sambrook. 1989. Isolation of Chinese hamster ovary cell lines that are temperature-conditional for the cell-surface expression of integral membrane proteins. *J. Cell Biol.* 108:339–353.
- Horwitz, M. S., M. D. Scharff, and J. V. Maizel. 1969. Synthesis and assembly of adenovirus 2. I. Polypeptide synthesis, assembly of capsomers,

- and morphogenesis of the virion. *Virology*. 39:682-694.
30. Hurlley, S. M., D. G. Bole, H. Hoover-Litty, A. Helenius, and C. S. Copeland. 1989. Interactions of misfolded influenza virus hemagglutinin with binding protein (BiP). *J. Cell Biol.* 108:2117-2126.
  31. Kim, P. S., and R. L. Baldwin. 1990. Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* 59:631-660.
  32. Knopf, P. M., and H. Lamfron. 1965. Changes in the ribosome distribution during incubation of rabbit reticulocytes in vitro. *Biochem. Biophys. Acta*. 95:398-407.
  33. Kornfeld, R., and S. Kornfeld. 1985. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* 54:631-664.
  34. Kozutsumi, Y., M. Segal, K. Normington, M. J. Gething, and J. Sambrook. 1988. The presence of malfolded proteins in the endoplasmic reticulum signals the induction of glucose regulated proteins. *Nature (Lond.)*. 332:462-464.
  35. Kozutsumi, Y., K. Normington, E. Press, C. Slaughter, J. Sambrook, and M. J. Gething. 1989. Identification of immunoglobulin heavy chain binding protein as glucose regulated protein 78 on the basis of amino acid sequence, immunological crossreactivity and functional activity. *J. Cell Sci.* 1:115-137.
  36. Machamer, C. E., and J. K. Rose. 1988. Vesicular stomatitis virus G proteins with altered glycosylation sites display temperature-sensitive intracellular transport and are subject to aberrant intermolecular disulfide bonding. *J. Biol. Chem.* 263:5955-5960.
  37. Machamer, C. E., R. W. Doms, D. G. Bole, A. Helenius, and J. K. Rose. 1990. Heavy chain binding protein recognizes incompletely disulfide-bonded forms of Vesicular Stomatitis Virus G protein. *J. Biol. Chem.* 265:6879-6883.
  38. Matsumura, M., G. Signor, and B. W. Matthews. 1989. Substantial increase of protein stability by multiple disulphide bonds. *Nature (Lond.)*. 342:291-293.
  39. Morgan, E. H., and T. Peters, Jr. 1985. The biosynthesis of rat transferrin: evidence for rapid glycosylation, disulfide bond formation, and tertiary folding. *J. Biol. Chem.* 260:14793-14801.
  40. Munro, S., and H. R. B. Pelham. 1986. An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell*. 46:291-300.
  41. Ng, D. T. W., R. E. Randall, and R. A. Lamb. 1989. Intracellular maturation and transport of the SV5 type II glycoprotein hemagglutinin-neuraminidase: specific and transient association with GRP78-BiP in the endoplasmic reticulum and extensive internalization from the cell surface. *J. Cell Biol.* 109:3273-3289.
  42. Owen, M. J., A.-M. Kissenerghis, and H. F. Lodish. 1980. Biosynthesis of HLA-A and HLA-B antigens *in vivo*. *J. Biol. Chem.* 255:9678-9684.
  43. Pace, C. N., G. R. Grimsley, J. A. Thompson, and B. J. Barnett. 1988. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem.* 263:11820-11825.
  44. Parmley, S. F., and G. P. Smith. 1988. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene*. 73:305-318.
  45. Perona, R., and R. Serrano. 1988. Increased pH and tumorigenicity of fibroblasts expressing a yeast proton pump. *Nature (Lond.)*. 324:438-440.
  46. Peters, T., Jr., and L. K. Davidson. 1982. The biosynthesis of rat serum albumin: *in vivo* studies on the formation of the disulfide bonds. *J. Biol. Chem.* 257:8847-8853.
  47. Rothman, J. E. 1989. Polypeptide chain binding proteins: catalysts of protein folding and related processes in cells. *Cell*. 59:591-601.
  48. Sambrook, J., T. Maniatis, and E. F. Fritsch. 1989. *Molecular Cloning: A Laboratory Manual*. 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. 545 pp.
  49. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*. 74:5463-5467.
  50. Skehel, J. J., P. M. Bayley, E. B. Brown, S. R. Martin, M. D. Waterfield, J. M. White, I. A. Wilson, and D. C. Wiley. 1982. 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.
  51. Thornton, J. M. 1981. Disulphide bridges in globular proteins. *J. Mol. Biol.* 151:261-287.
  52. Vuust, J., and K. A. Piez. 1972. A kinetic study of collagen biosynthesis. *J. Biol. Chem.* 247:856-862.
  53. Waterfield, M. D., M. J. Gething, G. Scrace, and J. J. Skehel. 1980. The carbohydrate side chains and disulphide bonds of the haemagglutinin of the influenza virus A/Japan 305/57 (H2N1). Elsevier, North Holland, Australia. 11-20.
  54. Weissman, J. S., and P. S. Kim. 1991. Reexamination of the folding of BPTI: predominance of native intermediates. *Science (Wash. DC)*. 253:1386-1393.
  55. Wiley, D. C., and J. J. Skehel. 1987. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annu. Rev. Biochem.* 56:365-394.
  56. Wilson, I. A., J. J. Skehel, and D. C. Wiley. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature (Lond.)*. 289:366-373.
  57. Zoller, M. J., and M. Smith. 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 derived vectors. *Methods Enzymol.* 100:468-500.
  58. Zoller, M. J., and M. Smith. 1985. Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. *DNA*. 3:479-488.