Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

ADVANCES IN VIRUS RESEARCH, VOL. 44

MOLECULAR BIOLOGY OF RUBELLA VIRUS

Teryl K. Frey

Department of Biology
Georgia State University
Atlanta, Georgia 30303

I. Introduction

Rubella virus is a significant human pathogen. However, the characterization of the molecular biology of rubella virus has been slow in progressing. Recently, significant progress has been made in molecular characterization, most notably the determination of the sequence of the genomic RNA. The most recent comprehensive review of rubella virus molecular biology was by Horzinek (1981) in a monograph on the non-arthropod-borne togaviruses. A complete summary of rubella virion structure and virion morphogenesis was contributed by Murphy

(1980). Reviews by Bowden and Westaway (1989) and Wolinsky (1990a) covered progress since that time in a more brief fashion. Thus, an updated comprehensive review is due. Except for a summary of the current medical significance of rubella virus, this review will not cover medical aspects of rubella virus. There are a large number of excellent reviews on medical aspects of rubella virus. Four reviews on the subject to which the reader is referred are Plotkin (1988), Wolinsky (1990a), Alford and Preblud (1990), and Cherry (1992).

II. Current Medical Significance

Rubella virus is the etiological agent of a disease known as rubella, 3-day measles, or German measles (for a review of the disease, see any of the medical reviews referred to above). Rubella virus infection is systemic in nature and the accompanying symptoms are generally benign, the most pronounced being a mild rash of short duration. It is estimated that as many as one-half of the cases are asymptomatic. The most common complication of rubella virus infection is transient joint involvement such as polyarthralgia and arthritis. Joint involvement is more common in adults (especially women) than in children. Less frequent complications are thrombocytopenic purpura (approximately 1 in 3000 cases) and postinfectious encephalopathy (approximately 1 in 8000 cases). Deaths due to complications from acute rubella virus infection are rare [1 in 60,000 cases; S. Holmes, Centers for Disease Control (CDC), unpublished data].

Of the names given to the agent and the disease that it causes, "3-day measles" is descriptive of the length of duration of the rash, "German measles" is due to the fact that the disease was first described by German physicians who named it *Rötheln*, and "rubella" was a name proposed in 1866 as an alternative to *Rötheln* by the English physician H. Veale, who found the German name "harsh and foreign to our ears" (Veale, 1866; Forbes, 1969). The choice of these names was unfortunate because as a result rubella (both the virus and the disease) is almost invariably confused with measles. To make things worse, measles is also known as rubeola. For the record, measles virus is a negative-stranded RNA virus of the family Paramyxoviridae. Although the diseases of measles and rubella share similarities in symptomology, measles is a much more serious disease with a significant incidence of dangerous complications that can lead to death [measles was responsible for 132 deaths in the United States from 1989 through 1991 (CDC, 1993)].

The primary health impact of rubella virus is that it is a teratogenic

agent. The virus can cross the placenta and replicate in the fetus. Teratogenicity is confined to the first trimester of pregnancy. The incidence of birth defects in babies born to women infected during the first trimester approaches 80%. The birth defects induced, known collectively as congenital rubella syndrome (CRS), can be extremely serious in nature, most commonly involving heart defects, cataracts, deafness, and mental retardation. Although the connection between rubella and birth defects was first recognized in 1941 (Gregg, 1941), the virus was not isolated until 1962 (Parkman *et al.*, 1962; Weller and Neva, 1962). In part spurred by the rubella epidemic of 1964, in which 20,000 CRS cases occurred in the United States, there was a flurry of research activity on the virus in the 1960s resulting in the development of several live attenuated vaccines (for review see Plotkin, 1988). In the United States, a vaccination program was instituted in 1969 that has led to a dramatic reduction in the incidence of both rubella and CRS. From 57,686 and 81 cases of rubella and CRS, respectively, in 1969 (CDC, 1980), the incidence bottomed out at 225 cases and 2 cases, respectively, in 1988 (CDC, 1991a).

In the United States, the vaccination strategy is aimed at elimination of rubella and includes both universal vaccination of infants at 15 months of age with the trivalent measles, mumps, rubella (MMR) vaccine and specific targeting with the rubella vaccine of seronegative women planning pregnancy and seronegative adults who could come in contact with women of childbearing age (most particularly medical personnel), although it is recommended that any individual over the age of 12 months without evidence of natural infection or vaccination be vaccinated (CDC, 1990). The vaccine virus can cross the placenta; however, no case of CRS due to vaccination has ever been reported and the registry of women vaccinated with rubella vaccine during pregnancy was closed in 1989. Nevertheless, vaccination of pregnant women is contraindicated and pregnant women who are found to be seronegative are vaccinated postpartum. In both the United Kingdom and Japan, a strategy of vaccination of seronegative female teenagers was initially pursued in hopes of maintaining herd immunity against wild rubella and only vaccinating those at risk. However, this strategy did not result in elimination of CRS and subsequently the vaccination of teenage females was augmented with universal childhood vaccination with MMR in the United Kingdom in 1988 (Badenoch, 1988) and in Japan in 1989 (Sugiura and Yamada, 1991). Although vaccination against rubella virus is practiced in the United States, Europe, Japan, Australia, and some countries in the developing world (Galazka, 1991; Epidemiological Notes, 1989), in a larger number of developing countries vaccination is not pursued and the wild virus is still endemic (Miller,

1991). Rubella virus is not included in the World Health Organization's Expanded Programme for Immunization (EPI) against preventable childhood diseases (Bart and Lin, 1990).

The current medical significance of rubella virus is threefold. First is that despite the existence of a vaccination program for 20 years, a resurgence of rubella occurred in the United States between 1989 and 1991 (Lindegren *et al.*, 1991). The number of rubella and CRS cases rose from 225 and 2 cases in 1988 to 396 and 2 cases in 1989, 1125 and 25 cases in 1990, and 1401 and 31 cases in 1991 (CDC, 1992b; S. Holmes, National Congenital Rubella Syndrome Registry, unpublished data). The resurgence ended in 1991, and in 1992 totals of only 160 and 3 cases of rubella and CRS were reported. The resurgence, most pronounced among individuals less than 1 year of age and over 15 years of age, was concentrated in the West and mid-Atlantic regions (half of the reported cases in 1990 occurred in California and the majority of the reported cases in 1991 occurred in Pennsylvania and New York) and was notable for the large number of outbreaks in prisons, college campuses, workplaces, and religious communities in which vaccination is not practiced [the 1991 outbreak in Pennsylvania and New York was largely concentrated in Amish communities (CDC, 1991b, 1992a)]. The epidemiology of the resurgence indicated that it was primarily due to infection of those never exposed to natural infection or vaccinated and thus the challenge of preventing similar resurgences in the future lies in achieving complete vaccination coverage, particularly among children $\geq$ 15 months of age and women of childbearing age who lack evidence of vaccination or immunity. In conjunction with the recent measles epidemic, it has been recommended by both the Advisory Committee on Immunization Practices (ACIP) of the U.S. Public Health Service and the Committee on Infectious Diseases of the American Academy of Pediatrics that children receive a second dose of MMR vaccine at age 5 years (CDC, 1989) or at age 10–12 years (Plotkin *et al.*, 1989).

The second current medical significance of rubella virus is reports of a significant incidence of serious chronic complications following vaccination of adult women. As discussed above, natural rubella infection of adults is frequently associated with transient joint involvement. Intriguingly, joint involvement occurs at a much higher incidence in adult women than in adult men. In some studies the incidence of joint involvement in adult women with naturally acquired rubella is reported to be as high as 50% and some investigators consider joint involvement to be a symptom of the disease in adult women rather than a complication (Heggie and Robbins, 1969; Cooper *et al.*, 1969; Wolinsky, 1990a; CDC, 1990). Not surprisingly, all of the vaccine strains are also

associated with transient joint involvement following vaccination; however, the incidence of joint involvement is lower following vaccination with any of the vaccine strains than following naturally acquired rubella. The ACIP states that the expected incidence of transient joint involvement in adult women following vaccination is in the neighborhood of 25% (CDC, 1990). The vaccine originally employed in the United States was the HPV-77/DE5 strain. However, in 1979 this vaccine was replaced with the RA 27/3 strain, which was found to be associated with a lower incidence of joint involvement and to induce an immunity more comparable to that induced by natural rubella virus infection.

In addition to transient arthralgia and arthritis, the occurrence of chronic arthritis following rubella virus infection has been recognized (reviewed in Ford *et al.*, 1988). A number of laboratories have reported isolation of rubella virus from synovial fluids of patients suffering from transient or chronic rubella arthritis, indicating that localized virus replication and persistence plays a role in the pathogenesis of the disease (Hildebrant and Maasab, 1966; Weibel *et al.*, 1969; Ogra and Herd, 1971; Ford *et al.*, 1982; Fraser *et al.*, 1983; Chantler *et al.*, 1985b). There were reported cases of chronic joint complications following vaccination of both children and adults with the HPV-77/DE5 vaccine and in some cases chronic arthritis was accompanied by neurological involvement (the collection of symptoms was termed chronic arthropathy) (Thompson *et al.*, 1973; Schaffner *et al.*, 1974; Spruance *et al.*, 1977; Chantler *et al.*, 1982). Although the occurrence of such complications was a factor in the withdrawal of the HPV-77 vaccine, the incidence was seemingly rare. However, a group in British Columbia has presented evidence that chronic arthropathy following natural rubella or vaccination in adult women is relatively common. In their most recent reported study, these investigators found that the incidence of chronic arthropathy following natural rubella was 30% and following vaccination with the RA 27/3 vaccine was 5% (Tingle *et al.*, 1986). Rubella virus was isolated by this group from mitogenically stimulated peripheral blood lymphocytes from patients suffering from chronic arthopathy following both natural infection and vaccination (the isolated virus was not identified as being wild type or vaccine), indicating that virus persistence was a factor in the pathogenesis of the syndrome (Chantler *et al.*, 1981; Chantler *et al.*, 1982; Tingle *et al.*, 1985a,b). More recently, rubella virus RNA was detected by polymerase chain reaction in peripheral blood leukocytes taken 8 to 10 months postvaccination from two women suffering from chronic arthropathy following vaccination (Mitchell *et al.*, 1993). Interestingly, sera taken prior to vaccination of some patients who subsequently

developed chronic arthropathy was positive for anti-rubella antibodies, indicating previous exposure to rubella virus (Tingle *et al.*, 1983, 1985a,b, 1989). These reports have sparked controversy because the ACIP has stated its belief that this type of chronic complication, if it does exist following RA 27/3 vaccination, is extremely rare in incidence (CDC, 1990, 1991c). In a review of available evidence conducted by the Institute of Medicine, the conclusion was reached that the evidence was consistent with a causal relation between vaccination with the RA 27/3 vaccine and development of chronic arthritis, although the evidence did not provide for reliable estimates of the risk of occurrence (Howson and Fineberg, 1992).

Finally, rubella virus is of current medical significance due to its association with chronic disease, including autoimmune disease. The connection between virus infection and chronic disease is an issue of interest with a large number of viruses and the mechanism of pathogenesis of such diseases is being actively studied (Oldstone, 1989). The highest correlation between rubella virus infection and chronic disease is in the CRS population (for review see Wolinsky, 1990a). In addition to the nonprogressive symptoms apparent at birth, CRS patients suffer from an extremely high incidence of endocrine dysfunctions that manifest themselves later in life as insulin-dependent diabetes mellitus (IDDM) and thyroiditis (the incidence of IDDM in the CRS population is 50 times that in the general population) and a rare, fatal, neurodegenerative disorder, progressive rubella panencephalitis (PRP), was first recognized in the population. The primary issue in the pathogenesis of these diseases in this population is the role played by virus persistence (Rawls, 1974; Sever *et al.*, 1985). Rubella virus infection *in utero* is systemic and virus can be isolated from multiple organs from both aborted fetuses and children who die of CRS-related complications after birth. Children who survive excrete virus in the urine and nasopharyngeal fluid and virus can be isolated from the cerebrospinal fluid. Virus excretion generally ceases within 6 months to 1 year of age, although excretion has been documented to occur through 2 years of age in some patients. There have been a few reports of virus isolation from CRS patients later in life (Menser *et al.*, 1967; Wolinsky, 1985); however, whether lifelong persistent infection occurs uniformly or is an atypical event is not known.

With respect to endocrine dysfunction, it can be hypothesized that replication of virus (either *in utero* or as persisting virus late in life) in the pancreas or thyroid could lead to virus-induced cell destruction or immune-mediated call destruction aimed at virus clearance, or to generation of an autoimmune response triggered by either molecular mimicry between virus antigens and cell proteins or by release of cell

material from cells destroyed by virus replication. In terms of the ability of rubella virus to replicate in endocrine tissue, rubella virus has been shown to replicate in human pancreatic islet cells *in vitro* (Numazaki *et al.*, 1989), to infect pancreatic cells in fetuses infected *in utero* (Monif *et al.*, 1965; Cooper *et al.*, 1965; De Prins *et al.*, 1978), and to replicate in the pancreas of experimentally infected animals (Menser *et al.*, 1978; Rayfield *et al.*, 1986). Rubella virus antigen was demonstrated in the thyroid tissue of a CRS patient suffering from Hashimoto's thyroiditis (Ziring *et al.*, 1977). In terms of molecular mimicry, the rubella virus capsid protein shares reasonable amino acid homology with human thyroglobin (Wolinsky, 1990a). Finally, a significant number of CRS patients have anti-islet call and anti-thyroid antibodies (Ginsberg-Fellner *et al.*, 1984; Clarke *et al.*, 1984). However, the distribution of major histocompatibility complex (MHC) haplotypes in the population of CRS patients suffering from IDDM is the same as in the non-CRS IDDM population, and an alternative hypothesis is that congenital rubella infection somehow increases the penetrance of a genetic trait (Rubenstein *et al.*, 1982; Ginsberg-Fellner *et al.*, 1984, 1985).

In terms of rubella virus persistence and chronic disease in the non-CRS population, at least two cases of PRP have been described in individuals who were infected postnatally (Wolinsky, 1985). More generally, rubella virus has been linked to both arthritis and multiple sclerosis (MS). The link between rubella virus and arthritis was originally suggested by the arthritogenic nature of rubella virus infection, including the occasional occurrence of chronic rubella arthritis, as discussed above. Rare cases of development of rheumatoid arthritis following rubella arthritis have been reported (Martenis *et al.*, 1968; McCormick *et al.*, 1978). Other studies have reported the detection of rubella virus antigen in synovial biopsy material (Ogra *et al.*, 1975), the isolation of virus from peripheral blood lymphocytes (Chantler *et al.*, 1985a), and the presence of anti-rubella virus antibody-secreting cells in synovial fluid (Chattopadhyay *et al.*, 1979) from a significant fraction of rheumatoid and/or juvenile rheumatoid arthritis patients examined. However, a large number of studies have detected no connection between rubella virus and these forms of arthritis (reviewed in Phillips, 1989). In another study, using a sensitive polymerase chain reaction assay, rubella virus was detected in cells from synovial fluid in 3 of 11 rheumatoid arthritis patients analyzed and in a patient suffering from psoriatic arthritis; however, rubella virus was not detected in synovial fluid cells in any of 12 juvenile rheumatoid arthritis patients analyzed (Phillips *et al.*, 1993). The link between rubella virus and MS is on the basis of epidemiological studies that indicate that MS patients were exposed to a number of common childhood viruses (includ-

ing rubella virus) relatively late in childhood (Compston *et al.*, 1986; Alvord *et al.*, 1987; Alter *et al.*, 1987) and on the presence of antibodies to specific viruses in the cerebrospinal fluid (CSF) of MS patients (anti-measles virus antibodies are the most frequently encountered, whereas anti-rubella virus antibodies are second) (Felgenhauer *et al.*, 1985). It should be noted that MS is an autoimmune disease and numerous attempts have been made to demonstrate the presence of virus genetic material in the brains of MS patients, without success (Godec *et al.*, 1992). In this regard, both the E1 and E2 glycoproteins of rubella virus share short, but significant, stretches of amino acid homology with human proteolipid protein, a major structural component of myelin, again suggestive of a trigger of autoimmunity by molecular mimicry (Atkins *et al.*, 1990; Wolinsky, 1990a). It is of interest to note that in cases of PRP, rubella virus antigen and genetic material are difficult to demonstrate in brains at autopsy and thus the pathogenesis of this disease may also involve autoimmune mechanisms (Wolinsky, 1990b).

III. Togavirus Taxonomy Revisited

Under current classification (Francki *et al.*, 1991), rubella virus is the sole member of the genus *Rubivirus* within the family Togaviridae of animal viruses. The other genus within the family Togaviridae is the genus *Alphavirus* (Sindbis virus is the type species), which consists of 27 arboviruses known originally as group A arboviruses. Since its adoption by the International Committee on Taxonomy of Viruses in 1974, the family Togaviridae has contained a number of other viruses. The subsequent molecular characterization of most of these viruses led to their reclassification into other virus families. This reclassification will be discussed here because it has only recently occurred with a number of the original togaviruses and because the result demonstrates that rubella virus fills a unique taxonomic niche.

The original family Togaviridae was approved by the International Committee on Taxonomy of Viruses (ICTV) in 1974 and was based on the morphological criteria that the virus have an infectious, nonsegmented RNA genome of 3×10^6 to 4×10^6 Da in molecular mass and a virion consisting of an isometric nucleocapsid surrounded by a lipid envelope of cell origin in which was embedded one or more virus-specified glycoproteins (Fenner *et al.*, 1974). Within the original togavirus family was a second genus of arboviruses, the flaviviruses or group B arboviruses (60 members, with yellow fever virus the type species). Several other viruses that were not arboviruses had characteristics that qualified them as togaviruses: rubella virus, equine arteritis virus (EAV),

bovine viral diarrhea virus (BVDV), hog cholera virus (HCV), border disease virus of sheep (BDV), lactate dehydrogenase-elevating virus of mice (LDV), simian hemorrhagic fever virus (SHFV), cell-fusing agent (CFA) of *Aedes aegypti*, and carrot mottle virus. These viruses, known collectively as the non-arthropod-borne togaviruses, were all eventually included in the togavirus family (Matthews, 1982). From among these viruses were established the genera *Rubivirus* and *Pestivirus* (HCV, BVDV, and BDV) in 1975 (Fenner, 1975–1976) and the genus *Arterivirus* (EAV) in 1984 (Westaway *et al.*, 1985a).

Among the togaviruses, the most extensive molecular characterization has been done on members of the genus *Alphavirus* (reviewed in Strauss and Strauss, 1986; Schlesinger and Schlesinger, 1990). The key features of taxonomic concern of the alphaviruses are that their genomic RNAs have a 5′ terminal type 0 cap (no ribose methylation) and a 3′ terminal poly(A) sequence, and contain two long open reading frames (ORFs). The 5′ proximal ORF encodes nonstructural proteins and is translated from the genomic RNA in infected cells while the 3′ proximal ORF encodes the structural proteins found in the virion and is translated from a single subgenomic RNA species synthesized in infected cells. Within the structural protein ORF, the sequences encoding the capsid protein are 5′ terminal. Rubella virus is similar to the alphaviruses in these features, justifying its current retention within the togavirus family.

The flavivirus genomic RNA has a 5′ terminal type 0 cap but lacks the 3′ terminal poly(A). Determination of the sequence of the flavivirus genomic RNA revealed a single ORF that encodes both structural and nonstructural proteins (Rice *et al.*, 1985). Within the ORF, the structural protein sequences are 5′ proximal (the sequences encoding the capsid protein are at the 5′ end of the ORF) and the nonstructural protein-encoding sequences are 3′ proximal. Because of the significant difference in genomic coding strategy between the alphaviruses and the flaviviruses, the flaviviruses were classified as their own family in 1984 (Westaway *et al.*, 1985b). Subsequent determination of the genomic sequence of two pestiviruses, BVDV (Collett *et al.*, 1988) and HCV (Meyers *et al.*, 1989), revealed a coding strategy similar to those of the flaviviruses and the genus *Pestivirus* was reclassified into the family Flaviviridae (Francki *et al.*, 1991). CFA is currently included as a possible member of the flavivirus family (Francki *et al.*, 1991) and the completion of the sequence of the CFA genome ascertained that the CFA coding strategy is similar to those of flaviviruses (Cammisa-Parks *et al.*, 1992). Human hepatitis C virus also has a genomic coding strategy similar to those of the flaviviruses (Miller and Purcell, 1990).

Sequence determination of the genome of EAV revealed a coding

strategy different from those of both the alphaviruses and flaviviruses (den Boon *et al.*, 1991). The EAV genome contains multiple ORFs that are translated from multiple subgenomic RNAs synthesized in infected cells. The ORF encoding the capsid protein is the most 3′ proximal. In these regards, the genomic coding strategy is reminiscent of those of members of the families Coronaviridae and Toroviridae (genera *Coronavirus* and *Torovirus*), although EAV is distinct both in genome size and virion morphology. Therefore the genus *Arterivirus* has been reclassified as a free-standing genus (Francki *et al.*, 1991). Genomic sequencing of both LDV and SHFV indicate that these viruses have coding strategies similar to that of EAV (Godeny *et al.*, 1990, 1993; Kuo *et al.*, 1991; E. K. Godeny and M. A. Brinton, personal communication). LDV and SHFV are currently classified as members of the genus *Arterivirus* (Francki *et al.*, 1991). Carrot mottle virus is currently classified as a possible member of the family Togaviridae; however, not enough work has been done on it to resolve its proper classification.

The result of all of this taxonomic reclassification is that rubella virus is the sole surviving non-arthropod-borne togavirus and the only only known member of the genus *Rubivirus*. This is an unusual situation because most human viruses have animal relatives within the same genus. Whether no animal rubivirus exists or whether one exists but has been overlooked because of the benign nature of the disease is not known. This is a question of medical relevance because no animal model for rubella exists and rubella virus replicates only subclinically in most common laboratory animals, although *in utero* infection through placental transfer in ferrets and rabbits has been reported.

IV. Virion Structure

A. Virion Morphology and Physicochemical Characteristics

In the only reported determination of the chemical composition of rubella virions (Voiland and Bardeletti, 1980), it was found that of virion dry weight 2.4% was RNA, 74.8% was protein, 4.0% was carbohydrate, and 18.8% was lipid. This composition differs from the composition of alphavirions, which are 8.7% RNA, 60–64% protein, 7% carbohydrate, and 27–31% lipid (as given in Francki *et al.*, 1991). The rubella virion composition seems inaccurate, particularly with regard to the RNA composition. Rubella virions are exceedingly more difficult to purify in quantity than are alphavirions and thus chemical composition analysis is correspondingly more difficult.

Negatively stained rubella virions are generally round with a diam-

eter of approximately 60 nm and are decorated with a glycoprotein fringe 5–8 nm in diameter (reviewed in Murphy, 1980). Two micrographs of negatively stained rubella virions are shown in Fig. 1A and B. The glycoprotein fringe is more clearly defined on the virions in Fig. 1A; however, the definition of the glycoprotein fringe on the virions in the micrograph in Fig. 1B is more typical. Negatively stained rubella virions appear similar to negatively stained alphavirions, with a notable distinction being that anomalously shaped virions (elongated, multicored, aberrant cores), which are rare in preparations of alphavirions, are relatively common in preparations of rubella virions.

Within the exterior glycoprotein fringe of the rubella virion, individual glycoprotein spikes are rarely visualized. However, on occasional micrographs, regularly spaced 5- to 8-nm projections with enlarged distal ends have been resolved (Holmes *et al.*, 1969; Bardeletti *et al.*, 1975). Following degradation of virions by prolonged incubation at 37, 45, or 56°C, 5- to 6-nm spherical subunits that form hexagonal and pentagonal arrays are observed both free in solution and in association with membranes (Payment *et al.*, 1975a). It is proposed that these arrays represent an end-on view of the virion spike. The glycoproteins on the surface of alphavirions form a $T = 4$ icosahedral lattice (reviewed in Murphy, 1980; Fuller and Argos, 1987). It is not known if the glycoproteins on the rubella virion form a similar structure.

Thin-section electron microsopy of rubella virions reveals an electron-dense core of 30–35 nm surrounded by an electron-lucent zone between the core and the virion envelope (as shown in Fig. 1C). The core is considered to be smaller and the electron-lucent zone to be wider in rubella virions than in alphavirions (Murphy, 1980). Rotational analysis of thin sections of rubella virions suggested that the core has a $T = 3$ icosahedral symmetry and contains 32 capsomers (Matsumoto and Higashi, 1974). It has been shown that alphavirus capsids have a $T = 4$ symmetry (Paredes *et al.*, 1993). Cores present in negative-stained rubella virion preparations due to spontaneous degradation or in negatively stained preparations of virions purposefully degraded with saponin, urea, or heat, are 30 to 35 nm in diameter and have a fenestrated appearance (Horzinek *et al.*, 1971; Payment *et al.*, 1975a). The rubella virus cores in these preparations are similar in appearance to alphavirus cores (see, e.g., Horzinek and Mussgay, 1971; Söderlund *et al.*, 1979; Paredes *et al.*, 1992); however, alphavirus cores have a slightly larger diameter of between 38 and 40 nm. The fenestrated appearance of rubella virus cores has led to the suggestion that the core is composed of ringlike subunits (Horzinek *et al.*, 1971).

The buoyant density of rubella virions measured in sucrose gradients has been reported to be between 1.175 and 1.20 g/ml, with most of

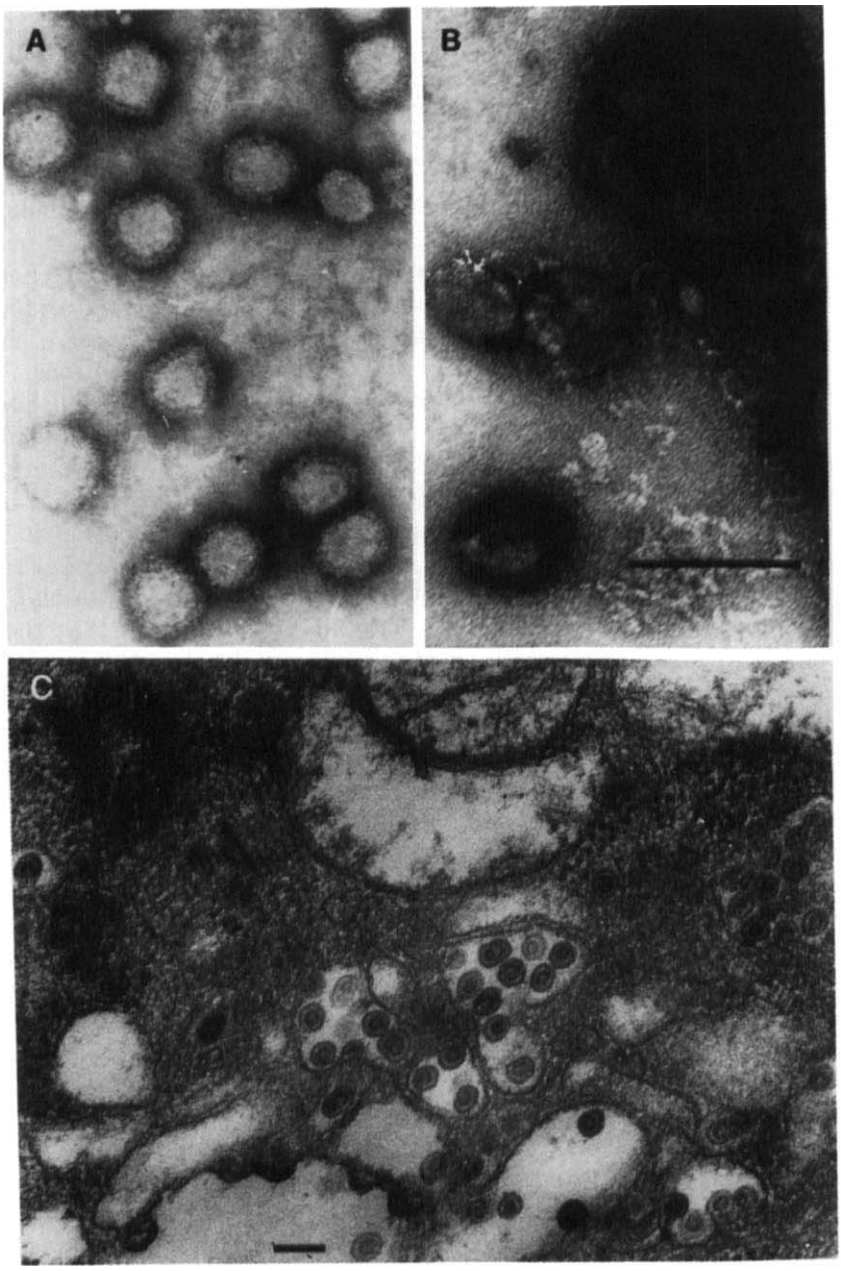

FIG. 1. Electron micrographs of rubella virions. (A and B) Rubella virions negatively stained with uranyl acetate. (B) Bar: 100 nm. [The micrograph in (A) was provided by C.-H. von Bonsdorff and the micrograph in (B) was provided by R. Simmons.] (C) Thin section of a rubella virus-infected BHK-21 cell. Cells were fixed in 1% glutaraldehyde and postfixed in 1% osmium tetroxide before embedding and thin sectioning. Bar: 100 nm. Note the presence of virions in the cytoplasmic matrix as well as in cytoplasmic vacuoles. [The micrograph was provided by L. Oshiro and is from Oshiro *et al.* (1969).]

the determinations yielding values between 1.18 and 1.19 g/ml (reviewed in Horzinek, 1981). In comparison, the buoyant density of alphavirions in sucrose is 1.2 g/ml (as given in Francki *et al.*, 1991). Widely varying determinations of the sedimenation coefficient of rubella virions have been reported: 342S (Russell *et al.*, 1967), 240S ± 25S (Thomssen *et al.*, 1968), and 350S ± 50S (Bardeletti *et al.*, 1975). The reason for the variation is not clear. In comparison, the sedimentation coefficient of alphavirions is 280S (as given in Francki *et al.*, 1991). Considering that rubella virions are morphologically similar to alphavirions, it seems likely that the sedimentation coefficients should be comparable and an experiment in which rubella virions and alphavirions were cosedimented in the same gradient would be appropriate.

Treatment of rubella virions with nonionic detergents in the presence or absence of ether leads to release of a moiety with a sedimentation coefficient of 150S (reviewed in Horzinek, 1981). This moiety contains the virion RNA and capsid protein and is devoid of virion glycoprotein and thus is the core or nucleocapsid. Because of instability and a propensity to aggregate, nucleocapsids isolated by gradient centrifugation following disruption of virions have never been visualized by electron microscopy. The only reliable determination of the density of the capsid (1.44 ± 0.04 g/ml) was made using isolated capsids that were fixed with glutaraldehyde prior to isopycnic centrifugation (Salmi, 1972a). In comparison, alphavirus capsids have a sedimentation coefficient of 150S to 160S and a density of 1.42–1.43 g/ml (Söderlund *et al.*, 1979). Treatment of the 150S rubella virus capsid with RNase leads to digestion of the virion RNA and disintegration of the 150S moiety, indicating that (1) either the virion RNA is exposed on or near the surface of the capsid or RNase can penetrate into the capsid and (2) it is likely that the virion RNA is an important structural component of the capsid [(Hovi, 1972); the virion RNA in the alphavirus capsid is also susceptible to digestion by RNase]. Also indicative of the importance of the virion RNA in the rubella virus capsid structure is that the 150S moiety can be dissociated with a number of polyanions, including mammalian rRNA and 5S RNA, but not by polycations (Hovi, 1972).

B. Stability

Rubella virus infectivity is stable within a pH range of 6.8 to 8.1 (reviewed in Horzinek, 1981). The virus is stable for years when frozen at temperatures below −20°C and for weeks at 4°C. At 37°C, infectivity decays, with a half-life of 1 to 2 hr, and at 56°C with a half-life of 5 to 20 min. The virus can be lyophilized and advantage is taken of this

property in that the vaccine is distributed in lyophilized form. The lyophilized vaccine is stable at 4°C for years, at room temperature for months, and at 37°C for weeks (McAleer *et al.*, 1980). Virus infectivity is susceptible to a number of common inactivating agents such as formaldehyde and ultraviolet (UV) light and to lipid solvents such as ether, chloroform, and detergents [sodium dodecyl sulfate (SDS), Nonidet P-40 (NP-40), Tween 80, deoxycholate (DOC), saponin]. When rubella virions are exposed to photodynamic dyes, virus infectivity is lost on exposure to light, indicating that the dyes can penetrate into the virion.

C. Virion Proteins

Owing to the appearance of a number of reports with conflicting data, the polypeptide composition of the rubella virion remained an enigma until 1982–1984, when six studies were published that reported consistent results (Ho-Terry and Cohen, 1982; Trudel *et al.*, 1982; Oker-Blom *et al.*, 1983; Waxham and Wolinsky, 1983; Toivonen *et al.*, 1983; Bowden and Westaway, 1984). Retrospectively, the earlier studies by Vaheri and Hovi (1972), Payment *et al.* (1975a), and Ho-Terry and Cohen (1980) had correctly determined the polypeptide composition. The rubella virion contains three virus-specific polypeptides: a capsid protein, C, and two envelope glycoproteins, E1 and E2. In polyacrylamide gels, C migrates as a doublet, the lower band always being of greater intensity than the upper band (see Fig. 2). The molecular weight of C has been reported to be between 33,000 and 38,000 (the range of reported molecular weights reflecting differences in the gel systems used; the mean molecular weight of these determinations is 34,500). E1 migrates as a discrete band with a molecular weight between 55,000 and 62,000 (mean, 59,000). In contrast, E2 migrates as a smear between 42,000 and 47,000 and between 47,000 and 54,000 (mean, 44,000 to 50,000). The intensity within the E2 smear is concentrated at the two ends of the smear. Using gradient gels, Oker-Blom *et al.* (1983) were able to separate E2 into two discrete bands that were termed E2a and E2b (upper and lower); however, no other group has been able to achieve a similar separation. The E2a and E2b ends of the virion E2 smear have similar amino acid compositions (Kalkkinen *et al.*, 1984), similar tryptic and chymotryptic peptide maps (Oker-Blom *et al.*, 1983; Ho-Terry and Cohen, 1982), the same amino-terminal sequence (Kalkkinen *et al.*, 1984), and are immunoprecipitated by antibodies raised against a synthetic peptide containing the COOH-terminal amino acids of E2 deduced from the nucleotide sequence (deduced amino acid sequence) (Baron and Forsell, 1991). Therefore, the

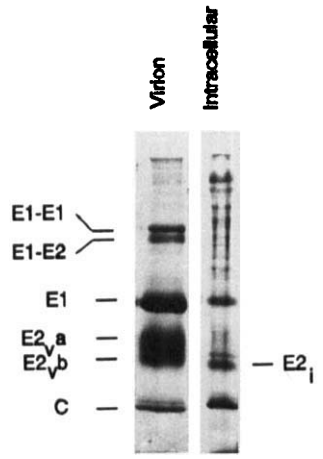

FIG. 2. Electropherogram of rubella virus structural proteins. Radiolabeled ([^{35}S]methionine) rubella virions and rubella virus structural proteins immunoprecipitated from infected Vero cells with human anti-rubella virus serum were electrophoresed in an SDS-polyacrylamide gel. The identity of each species is denoted: C, capsid protein; $E2_v$a and $E2_v$b, upper and lower ends of $E2_v$ glycoprotein smear; $E2_i$, intracellular form of E2 glycoprotein; E1, E1 glycoprotein; E1–E2, covalently linked E1–E2 heterodimers; E1–E1, covalently linked E1–E1 homodimers. [The electropherogram is from Marr *et al.* (1991).]

E2a and E2b ends of the virion smear share a common peptide backbone and the differences in molecular weight are due to differences in glycosylation. Strain differences in the relative amounts of E2a and E2b in virions have been noted in that whereas most strains studied have roughly equal amounts of E2a and E2b, the M33 strain contains primarily E2a (Lundström *et al.*, 1991). Whereas the initially detectable forms of both E1 and C in pulse-radiolabeled infected cells comigrate with their virion counterparts, the initially detectable intracellular form of E2 appears as a discrete band that migrates more rapidly than does the lower end of the virion E2 smear (M_r 41,000 to 43,000; mean, 42,000; Fig. 2) (Oker-Blom *et al.*, 1983; Bowden and Westaway, 1984) and intracellular bands comigrating with the virion E2 forms are not detectable until after a chase of 2 hr (Baron and Forsell, 1991). For convenience, the intracellular and virion forms of E2 will be referred to as $E2_i$ and $E2_v$, respectively, and the upper and lower ends of the $E2_v$ smear will be termed $E2_v$a and $E2_v$b.

A schematic diagram of the structural protein ORF with the relative location of landmarks on the structural proteins predicted from their deduced amino acid sequences is shown in Fig. 3. From the deduced

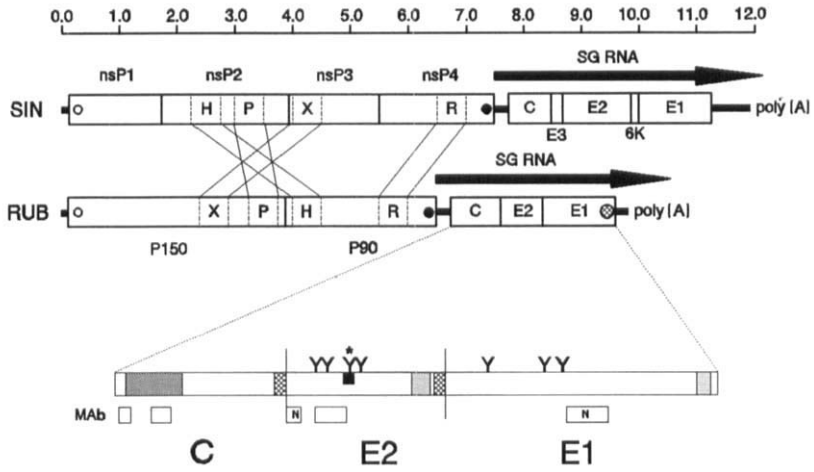

FIG. 3. Topography of the genome RNAs of rubella virus and Sindbis virus. The scale at the top of the diagram is in kilobases. Untranslated sequences are denoted by black lines and open reading frames (ORFs) by open boxes. In both viruses, the 5′ proximal ORF encodes nonstructural protein and the 3′ proximal ORF encodes structural proteins. The boundaries of the individual proteins processed from the precursor translated from each ORF are denoted (nsP = nonstructural protein). Within the nonstructural protein ORFs, the location of global amino acid motifs indicative of replicase (R), helicase (H), and cysteine protease activity (P) as well as the small region of homology between the deduced amino acid sequence of rubella virus and Sindbis virus (X motif) are shown. Also shown are the positions of regions of nucleotide homology between rubella virus and alphaviruses (○, 51/46 nucleotide conserved region; ●, subgenomic start site), the 3′ terminal stem-and-loop structure in the rubella virus genome (⊗), and the sequences included in the subgenomic RNAs (SG RNA) of both viruses. An expanded topography of the rubella virus structural protein ORF is shown at the bottom of the diagram. Within the ORF, the positioning of the following domains of the structural proteins is shown: ▨, the hydrophilic region of C, which contains a high concentration of basic amino acids and putatively interacts with the virion RNA; ▩, the hydrophobic signal sequences that precede the N termini of E2 and E1; ▩, the transmembrane sequences of E2 and E1; Y, potential N-linked glycosylation sites (the site marked with a Y is not present in the HPV-77 vaccine strain); □, a putative region for O-linked glycosylation. Below the diagram are shown the location of domains that contain epitopes recognized by mouse monoclonal antibodies (MAbs) (N denotes domains containing epitopes recognized by neutralizing MAbs).

amino acid sequence (Clarke *et al.*, 1987; Frey and Marr, 1988; Takkinen *et al.*, 1988), C is 293 or 300 amino acids in length, depending on which of two closely spaced, in-frame AUGs at the beginning of the coding sequence translation is initiated. The amino-terminal half of the C protein is highly hydrophilic and is particularly rich in arginine residues. This is presumably the domain of the protein that interacts with the viral RNA in the capsid. Unique among viruses with quasi-

spherical capsids or virions, the carboxy-terminal 23 amino acids of the rubella virus C protein are highly hydrophobic and function within the structural polyprotein precursor as the signal sequence for E2. The function of this sequence in processing of the structural protein precursor and the implications of its presence on capsid morphogenesis are discussed in Sections VI,E and VI,F,1, below. Additionally, it has been shown that C is phosphorylated, although the extent of phosphorylation was not determined (Marr *et al.*, 1991). On isoelectric focusing gels, C fractionates into two bands with p*I* values of 8.8 and 9.5 (Waxham and Wolinsky, 1985b), which could be due to differential phosphorylation. In the capsid, C is present as a disulfide-linked dimer (Waxham and Wolinsky, 1983; Baron and Forsell, 1991).

From the deduced amino acid sequence (Frey *et al.*, 1986; Clarke *et al.*, 1987; Vidgren *et al.*, 1987; Frey and Marr, 1988), E1 and E2 are 481 and 282 amino acids in length, respectively. Both E1 and E2 contain N-linked carbohydrate groups (Ho-Terry and Cohen, 1982; Oker-Blom *et al.*, 1983; Waxham and Wolinsky, 1983; Bowden and Westaway, 1984). The deduced amino acid sequence of E1 contains three potential N-linked glycosylation sites (Asn-X-Ser/Thr), all of which are utilized (Bowden and Westaway, 1985; Hobman *et al.*, 1991). Nonglycosylated E1 synthesized in tunicamycin-treated cells has a molecular weight of 53,000 (Oker-Blom *et al.*, 1983; Bowden and Westaway, 1984; Sanchez and Frey, 1991) and thus roughly 6 kDa of the molecular mass of E1 is contributed by carbohydrate moieties. In contrast, E2 is heavily glycosylated. The molecular weight of E2 in tunicamycin-treated cells is 29,000 to 30,000 (Oker-Blom *et al.*, 1983; Sanchez and Frey, 1991) and thus 30 to 40% of the molecular mass of E2 (15 to 20 kDa) is carbohydrate. The deduced amino acid sequence of E2 contains four potential N-linked glycosylation sites and all four appear to be utilized (Bowden and Westaway, 1985; Zheng *et al.*, 1989; Qiu *et al.*, 1992a). Interestingly, the genome of one of the attenuated rubella virus vaccine strains (HPV-77) was found to contain two nucleotide substitutions in the E2-coding region, in comparison to its wild-type parent strain (the M33 strain), which eliminate one of the potential N-linked sites (Zheng *et al.*, 1989). As predicted if this site is utilized in E2 of the M33 strain, both $E2_v$ and $E2_i$ of HPV-77 have molecular weights that are 3000 Da less than the molecular weights of their M33 counterparts. Whether this mutation is involved in the attenuated phenotype of this vaccine strain is not known. In addition to N-linked glycans, $E2_v$ is also O-glycosylated (Sanchez and Frey, 1991; Lundström *et al.*, 1991). *O*-Glycans maximally contribute from 4 to 9 kDa to the molecular mass of $E2_v$. Consensus sequences present at O-linked glycosylation sites have yet to be established; however, regions with concentrations

of hydroxy-amino acids in association with proline residues appear to be preferred (reviews in Will *et al.*, 1993). A 27-residue sequence in rubella virus E2 (amino acids 97–123) that contains 8 threonines, 1 serine, and 8 prolines would thus seem to be an excellent candidate site. On isoelectric focusing gels, E1 focuses as a single species of p*I* between 6.0 and 6.5 (Ho-Terry and Cohen, 1982; Waxham and Wolinsky, 1985b). Ho-Terry and Cohen (1982) found the p*I* values of $E2_v$a and $E2_v$b to be 8.8 and 6.0, respectively, whereas Waxham and Wolinsky (1985b) found $E2_v$ to resolve into at least 15 bands with a range of p*I* values from 5.0 to 8.6

The extent of processing of the glycans on both E1 and E2 in rubella virions is heterogeneous as determined both by resistance to endoglycosidase H (endo H) and binding of lectins specific to high mannose-type, complex-type, and hybrid-type glycans (Toivonen *et al.*, 1983; Sanchez and Frey, 1991; Lundström *et al.*, 1991). In the case of E2, the extent of processing is related to the molecular weight within the $E2_v$ smear as $E2_v$b is endo H sensitive, is preferentially labeled with [^{3}H]mannose, and binds lectins specific for high-mannose glycans whereas $E2_v$a is endo H resistant, is preferentially labeled with [^{3}H]galactose, and binds lectins specific for hybrid-type and complex-type glycans (Oker-Blom *et al.*, 1983; Bowden and Westaway, 1984; Sanchez and Frey, 1991). Lectin-binding assays also indicate that the *O*-glycans on $E2_v$b are less substituted than the *O*-glycans on $E2_v$a (Sanchez and Frey, 1991; Lundström *et al.*, 1991). [^{3}H]Sodium borohydride labeling of both E1 and E2 following reaction with galactose oxidase is equivalent with or without prior digestion with neuraminidase (Toivonen *et al.*, 1983) and mild acid hydrolysis has no effect on the migration of E1 and E2 glycopeptides (Bowden and Westaway, 1985), indicating that the majority of complex-type *N*-glycans on these glycoproteins terminate in galactose. However, both E1 and $E2_v$a bind lectins specific for terminal sialic acid (Sanchez and Frey, 1991; Lundström *et al.*, 1991) and thus some fraction of the complex-type carbohydrates contains sialic acid. Digestion of $E2_v$ with endoglycosidase F results in the production of two discrete bands with molecular weights of 33,000 and 37,500, digestion of $E2_v$b yielding only the 33,000 product whereas digestion of $E2_v$a yields both products (Sanchez and Frey, 1991). Digestion of M33 virions, which contain primarily $E2_v$a, with *N*-glycanase F yields a discrete 38-kDa product (Lundström *et al.*, 1991). These results indicate that the smeary nature of $E2_v$ is due to differential processing among the *N*-glycans and suggests that the difference in molecular weight between $E2_v$a and $E2_v$b is due to O-glycosylation, either the presence of additional *O*-glycans

on $E2_v$a in comparison to $E2_v$b or a greater extent of substitution of *O*-glycans on $E2_v$a than on $E2_v$b.

Information on interaction of E2 and E1 with the virion envelope comes entirely from the deduced amino acid sequences (Frey *et al.*, 1986; Clarke *et al.*, 1987; Vidgren *et al.*, 1987; Frey and Marr, 1988). As are most virus glycoproteins, rubella virus E1 and E2 are class I membrane proteins in that they contain a single potential transmembrane sequence [(Wiley, 1986); class II membrane proteins contain two or more membrane-spanning domains]. In both E1 and E2, the predicted transmembrane sequence is near the carboxy terminus and thus both proteins are type I glycoproteins (as opposed to type II glycoproteins, in which the transmembrane sequence is at the amino terminus). The putative E2 transmembrane sequence is 39 residues in length (which is bisected by an Asp-His dipeptide) and is followed by a heptapeptide of sequence Arg-Arg-Cys-Ala-Arg-Arg-Arg and a stretch of 20 hydrophobic residues that serves as the signal sequence for E1. The putative E1 transmembrane sequence is 22 amino acids in length and is followed by a stretch of 13 amino acids. This 13-amino acid stretch as well as the Arg-rich heptapeptide of E2 are probably located on the underside of the virion envelope and either one or both may play a role in mediating the interaction of the glycoproteins with the nucleocapsid. As is the case with the glycoproteins of a number of other enveloped viruses, [^{3}H]palmitic acid is incorporated into both rubella virus E2 and E1, indicating the covalent attachment of fatty acid to amino acid residues within the transmembrane region (Waxham and Wolinsky, 1985b). Although fatty acylation of proteins can occur through both oxy-ester bonds to threonine and serine residues and thio-ester bonds to cysteine residues, evidence on virus glycoproteins indicates that fatty acylation usually involves thioesterification (reviewed in Schultz *et al.*, 1988). The exact attachment site has been determined for three virus glycoproteins [G glycoprotein of vesicular stomatitis virus (VSV; Rose *et al.*, 1984), E1 of Semliki Forest virus (Schmidt *et al.*, 1988), and the fusion protein of respiratory syncytial virus (Arumugham *et al.*, 1989)] and in each case the cysteine is close to the COOH terminus of the transmembrane region (i.e., on the cytoplasmic side of the lipid bilayer). Significantly, in this regard, the rubella virus E2 transmembrane region contains a Cys residue at the exact carboxy terminus and the E1 transmembrane sequence contains Cys at both the second and third residues from the carboxy terminus. In contrast, Ivanova and Schlesinger (1993) showed that palymitoylation sites on E2 of Sindbis virus are three cysteine residues within the 31-amino acid cytoplasmic tail, a region that serves as the signal sequence for the 6K protein.

However, the homologous region of rubella virus E2 (the E1 signal sequence) lacks cysteine residues.

D. Conformation and Function of E1 and E2 on Virions

The conformation assumed by E1 and E2 on the virion is not clear and there is even disagreement as to the molar ratio of these species in the virion, with E1:E2 ratios of 1:1 (Bowden and Westaway, 1984; Baron and Forsell, 1991), 6:5 (Payment *et al.*, 1975a), and 5:1 (Vaheri and Hovi, 1972; Waxham and Wolinsky, 1985b) having been reported. In the virions of alphaviruses, E1 and E2 are noncovalently complexed in a heterodimer and three heterodimers form each surface spike (Rice and Strauss, 1982; Vogel *et al.*, 1986; Fuller and Argos, 1987; Anthony and Brown, 1991). Although the architecture of the rubella virus surface spikes has not been defined, there is evidence to indicate that E1 and E2 form a complex in the rubella virion. Antisera specific for either E1 or E2 immunoprecipitate both proteins from preparations of virions disrupted by nonionic detergents (Baron and Forsell, 1991). When nonionic detergent-disrupted virions are subjected to velocity centrifugation on sucrose gradients, both E1 and E2 sediment in two fractions, a monomer fraction and a fraction sedimenting more rapidly than monomers (Vaheri and Vesikari, 1971; Payment *et al.*, 1975a; Ho-Terry and Cohen, 1980; Trudel *et al.*, 1980; Dorsett *et al.*, 1985; Baron and Forsell, 1991). E1 and E2 in the monomer fraction are immunoprecipitable only with homologous antisera whereas E1 and E2 in the more rapidly sedimenting fraction are immunoprecipitable by antisera against either protein, indicating that E1 and E2 in this fraction are in a complex (Baron and Forsell, 1991). Depending on the virus, the noncovalent association of glycoproteins can be sensitive to disruption by the conditions used in such experiments (reviewed in Hurtley and Helenius, 1989) and thus it is likely that the monomer fraction observed in such experiments is due to breakdown of higher order complexes rather than to the existence of monomers in the rubella virion. As a point of comparison, the alphavirus E1–E2 heterodimer is generally resistant to disruption by nonionic detergents (Wahlberg *et al.*, 1989).

Disulfide-bonded glycoprotein complexes are also routinely observed when preparations of rubella virions are disrupted with SDS in the absence of reducing agent and subjected to polyacrylamide electrophoresis under nonreducing conditions (Ho-Terry and Cohen, 1980; Waxham and Wolinsky, 1983; Baron and Forsell, 1991). These complexes migrate as two bands with molecular weights of 105,000 and 95,000 and have been shown to contain E1 only and E1 and E2, respec-

tively (Waxham and Wolinsky, 1983; Dorsett *et al.*, 1985). The molecular weights of these bands indicate that they represent E1–E1 homodimers and E1–E2 heterodimers. Although it is apparent that only a small fraction of the glycoproteins in the rubella virion is present in disulfide-linked complexes, the E1–E1 homodimers are immunoprecipitated from virions disrupted with nonionic detergents by anti-E2 as well as anti-E1 serum, indicating that they are part of a larger complex (Baron and Forsell, 1991). It will be of interest to determine if these disulfide-linked complexes are a key element in the architecture of the virion surface spikes or whether they are aberrant intracellular forms that are fortuitously incorporated into some fraction of the virion population.

Following disruption of rubella virions with SDS, E1 migrates more rapidly in gels under nonreducing conditions than under reducing conditions, whereas E2 migrates similarly under both conditions, indicating that disulfide bonds are important in the intramolecular conformation of E1 but not of E2 (Waxham and Wolinsky, 1983; Dorsett *et al.*, 1985). Interestingly, all of the anti-E2 monoclonal antibodies thus far defined recognize contiguous epitopes whereas the majority of the anti-E1 monoclonal antibodies recognize noncontiguous epitopes (Green and Dorsett, 1986; Wolinsky *et al.*, 1991). The disulfide bonding of E1 is important in the formation of a significant fraction of these epitopes because a number of anti-E1 monoclonal antibodies bind to SDS-disrupted, non-reduced E1 but not to SDS-disrupted, mercaptoethanol-reduced E1 (Green and Dorsett, 1986). Disulfide bonding is also important to the maintenance of the structure and function of the glycoprotein complex because treatment of virions with mercaptoethanol markedly increases the susceptibility of both E1 and E2 to trypsin and destroys both the virion hemagglutinin activity and infectivity (Ho-Terry and Cohen, 1981; Katow and Sugiura, 1988b). Interestingly, mercaptoethanol treatment of virions sufficient to destroy hemagglutinin activity does not affect binding with antibodies present in human sera that exhibit hemagglutination inhibition (HI) as measured in an HI competition assay (Ho-Terry and Cohen, 1981). This finding suggests that HI antibodies bind at a site different from that interacting directly with erythrocytes.

Whatever the conformation of E1 and E2 in the virion, both are exposed on the virion surface as evidenced by the finding that they are both labeled with [^{3}H]borohydride following galactose oxidase treatment and that monoclonal antibodies specific to both bind to intact virions (Oker-Blom *et al.*, 1983; Toivonen *et al.*, 1983; Waxhan and Wolinsky, 1985a; Green and Dorsett, 1986). However, E2 is more resistant to digestion with trypsin and glycosidases than is E1 (Ho-Terry

and Cohen, 1980, 1984; Katow and Sugiura, 1988a), radioiodination of intact virions leads to the labeling of E1 but not E2 (Katow and Sugiura, 1988a), and the majority of anti-E2 monoclonal antibodies do not react with intact virions whereas the majority of anti-E1 monoclonals do (Waxham and Wolinsky, 1985a; Green and Dorsett, 1986). Thus, E2 appears to be less exposed on the virion surface than is E1. Although it is tempting to hypothesize that the relative inaccessibility of E2 is due to its being covered by E1 in a heterodimer, from available data it is also conceivable that the secondary structure of E2, particularly the heavy glycosylation, accounts for its resistance to enzymes and antibody binding.

Activities associated with rubella virions are hemagglutination (Stewart *et al.*, 1967; Halonen *et al.*, 1967; Schmidt *et al.*, 1971), cell fusion at low pH (Väänänen and Kääriäinen, 1980), hemolysis (Kobayashi, 1978; Väänänen and Kääriäinen, 1980; Kobayashi and Suzuki, 1982), and a weak neuraminidase (Bardeletti *et al.*, 1975). The hemagglutinin resides on E1, as (1) treatment of virions with concentrations of trypsin that digest E1 but not E2 destroy hemagglutinin activity (Ho-Terry and Cohen, 1981), (2) anti-E1 monoclonal antibodies, but not anti-E2 monoclonal antibodies, have been isolated that exhibit hemagglutination inhibition (HI) activity (Waxham and Wolinsky, 1983, 1985a; Green and Dorsett, 1986), and (3) purified E1 and a 13-kDa trypsin cleavage product of E1 adsorb to erythrocytes (Ho-Terry and Cohen, 1985). Because anti-E1 monoclonal antibodies have been isolated that exhibit HI but do not neutralize (Waxham and Wolinsky, 1985a; Green and Dorsett, 1986), hemagglutination (binding to erythrocytes) and binding to receptors on infectable cells are not strictly coincident.

Fusion of erythrocytes by rubella virions was observed when hemagglutination was allowed to occur at pH 5.8 at 0°C followed by incubation at 42°C at pH 5.8 or 6.3, but not at higher pH values (Väänänen and Kääriäinen, 1980). It was also observed that fusion of rubella virus-infected cells occurred following brief exposure to pH values of 6.0 and below, maximal fusion occurring at pH values of 5.0 and below, presumably mediated by virus glycoproteins present in the plasma membrane of infected cells (Katow and Sugiura, 1988a). Rubella virions exposed to pH 5 lost infectivity and hemagglutination activity, but gained the ability to bind liposomes (Katow and Sugiura, 1988a). Exposure of similar fusogenic activities on treatment with low pH has been demonstrated with a number of enveloped viruses (Wiley, 1986). The activity functions following endocytosis of virus particles that have attached to receptors on the surface of susceptible cells. In the reduced pH environment of endocytic vesicles, the fusogenic activity

becomes operative, causing fusion of viral membrane with the vesicular membrane and allowing release of the virus capsid into the cytoplasm. Low-pH treatment of rubella virions caused an irreversible conformational change of E1 and E2 exhibited by a reversal in relative sensitivity to trypsin (i.e., E2 became trypsin sensitive and E1 became trypsin resistant) and by the ability to radioiodinate E2 (Katow and Sugiura, 1988a). Trypsin-treated, low pH-exposed virus retained its ability to fuse with liposomes, indicating that the fusogenic activity resides on E1.

Exhibition of hemolysis activity requires damage of virions, which can be purposefully induced by repeated freeze-thawing (Kobayashi, 1978; Kobayashi and Suzuki, 1982) but often occurs during routine virion purification (Väänänen and Kääriäinen, 1980). Hemolysin requires a functional hemagglutinin. Unlike hemagglutination, which is exhibited by nonionic detergent-treated virions and isolated glycoprotein complexes, hemolysin activity, as well as fusion, requires complete virions. The mechanism of hemolysin activity is poorly understood but is thought to occur when membrane fusion is mediated by virus glycoproteins following erythrocyte binding and hemoglobin leaks out through damaged virion particles. Two groups reported that among their panels of anti-E1 monoclonal antibodies were antibodies that inhibited hemolysin (Waxham and Wolinsky, 1985b; Umino *et al.*, 1985). The hemolysin-inhibiting monoclonals isolated by Waxham and Wolinsky (1985b) also exhibited HI activity; however, the hemolysin-inhibiting monoclonals isolated by Umino *et al.* (1985) did not exhibit HI, indicating that hemolysin activity resides on E1.

E. Immunological Determinants on Virion Proteins

Only one serotype of rubella virus exists and thus naturally acquired rubella virus infection or vaccination generally confers immunity to recurrence of acute infection. As a result of natural infection or vaccination, a persisting circulating antibody titer is developed that can be measured by neutralization, HI, complement fixation, or a number of assays that measure binding of antibody to virus antigen [enzyme-linked immunosorbent assay (ELISA) or latex agglutination]. Despite the existence of a single serotype, strains of rubella virus (both wild and vaccine) can be differentiated from each other immunologically by using assays that measure the avidity with which serum raised against one strain reacts with other strains. Assays which discriminate between strains include neutralization (Gould and Butler, 1980), radioimmune precipitation (Ho-Terry *et al.*, 1982), and Western blotting (Dorsett *et al.*, 1985; Cusi *et al.*, 1989), but not HI (Best and Banatvala, 1970). In one

Western blotting study (Dorsett *et al.*, 1985), the determinants of strain specificity resided on E2, whereas in another study (Cusi *et al.*, 1989) strain differences were apparent with both E2 and C. This finding correlates with the lack of strain specificity distinguishable by HI, because the hemagglutinin resides on E1, and suggests that at least some neutralization epitopes recognized by polyclonal sera reside on E2. Best *et al.* (1992) reported that 9 strains of rubella virus reacted similarly with a panel of 28 monoclonal antibodies in neutralization, HI, immunofluorescence, and enzyme immunoassays. Twenty-four of these monoclonal antibodies were specific for E1, thus confirming the lack of strain-specific determinants on this protein. Only two monoclonal antibodies each were specific for E2 and C and thus conclusions as to determinants of strain specificity on these proteins are limited to the regions of the proteins recognized by these antibodies (neither of the anti-E2 monoclonal antibodies exhibited neutralizing activity).

Whether the detectable immunological differences between strains are of biological consequence is unknown. One phenomenon in which such differences could conceivably play a role is reinfection of individuals who are demonstrably seropositive if the reinfecting virus differed antigenically from the original virus to which the individual was exposed. During rubella outbreaks, reinfections have been found to occur in 5% of seropositive individuals with immunity due to natural infection and in over 50% of seropositive individuals who were vaccinated (reviewed in Miller, 1990; Best, 1993). Reinfection usually occurs in individuals who have low antibody titers. Although reinfection is usually asymptomatic and virus is confined to the pharynx, occasionally viremia develops and cases of CRS resulting from reinfection have been reported. The incidence of CRS in babies born to women who were reinfected during the first trimester of pregnancy is 8% or less.

In response to rubella virus infection, anti-rubella virus antibodies in all of the immunoglobulin classes are induced (Salonen *et al.*, 1985). The transient IgM response is directed primarily against E1 (Partanen *et al.*, 1985; Pettersson *et al.*, 1985) whereas the persisting IgG response is directed against all three virion proteins, although predominant reactivity is against E1 (Partanen *et al.*, 1985; Pettersson *et al.*, 1985; Dorsett *et al.*, 1985; Katow and Suguira, 1985; de Mazancourt *et al.*, 1986; Cusi *et al.*, 1988, 1989). Persisting IgG antibodies are primarily, if not solely, of the IgG_1 subtype (Sarnesto *et al.*, 1985; Stokes *et al.*, 1986). Serum IgA response is directed solely against the C protein (Ho-Terry and Cohen, 1979; Partanen *et al.*, 1985), a seemingly anomalous finding considering the importance of this immunoglobulin class in preventing reinfection in the nasopharynx (Ogra *et al.*, 1971; Al-Nakib *et al.*, 1975). Presumably, the targeting of serum IgA differs from that

of IgA in nasopharyngeal secretions. Interestingly, MS patients exhibit a comparatively reduced humoral reactivity to E1 and increased reactivity to E2 in comparison with the control population, although the significance of this observation is unclear (Nath and Wolinsky, 1990).

Despite the fact that the conditions of congenital infection with rubella virus could lead to immunological tolerance, CRS patients exhibit a humoral response against rubella virus and the presence of antibodies (both maternal and fetal) that exhibit neutralizing activity is detectable at birth (reviewed in Rawls, 1974). In a study of CRS sera obtained within the first 16 months of life, the humoral response was directly primarily, and in some cases solely, against E1 and reactivity against C was rarely detected (de Mazancourt *et al.*, 1986) whereas in a study of sera from older CRS patients, an accentuated antibody response to E2 was noted (Katow and Sugiura, 1985). Frank virus persistence in CRS patients thus occurs in the presence of a neutralizing humoral response and it has been found that rubella virus persistence can be maintained in cell culture in the presence of neutralizing antibodies (reviewed in Abernathy *et al.*, 1990). In such cell culture systems, the ability of the virus to mature at intracellular locations is the key factor in the maintenance of persistence and this feature is certainly important in maintenance of persistence *in vivo*. However, the mechanism by which infected cells avoid destruction by complement, the antibody-dependent cytotoxic cell system, or the cell-mediated immune system *in vivo* is not clear. In the latter regard, several studies have detected partial or total impairment of the cell-mediated immune response against rubella virus in CRS patients (reviewed in Buimovici-Klein and Cooper, 1985). A final point of interest on the humoral anti-rubella virus response in CRS patients is the finding that serum antibody levels generally decline during childhood and become undetectable in a fraction of patients [3 to 35%, depending on the study (Kenrick *et al.*, 1968; Hardy *et al.*, 1969; Cooper *et al.*, 1971; Dudgeon *et al.*, 1972; Ueda *et al.*, 1975)]. Vaccination of such patients does not result in systemic infection as measured by viremia or virus shedding; however, reappearance of a titer following vaccination is rarely observed (Cooper *et al.*, 1971).

As alluded to several times in preceding discussions, several panels of monoclonal antibodies have been raised against rubella virus (Tedder *et al.*, 1982; Waxham and Wolinsky, 1985a; Umino *et al.*, 1985; Green and Dorsett, 1986; Ho-Terry *et al.*, 1986; Gerna *et al.*, 1987). In these panels, the greatest number of monoclonal antibodies is specific for E1 and, depending on the panel, three to six epitopes on E1 were defined. Anti-E1 monoclonal antibodies exhibiting both neutralization

and HI, neutralization but not HI, HI but not neutralization, and neither activity have been characterized. Within these panels, monoclonal antibodies with anti-E2 specificity are rare. Wolinsky *et al.* (1991) have isolated six anti-E2 monoclonals that recognize at least three epitopes, none of which neutralize. However, the single anti-E2 monoclonal antibody isolated by Green and Dorsett (1986) does neutralize. Finally, monoclonals defining at least four epitopes of C have been isolated (Waxham and Wolinsky, 1985a; Wolinsky *et al.*, 1991). As expected from the internal location of C in the virion, anti-C monoclonal antibodies exhibit neither neutralization nor HI activity.

The epitopes recognized by some of the monoclonal antibodies specific for all three proteins have been localized (see Fig. 3) and in some cases the epitopes recognized have been correlated with recognition by antibodies present in human sera. Epitopes recognized by anti-C monoclonal antibodies have been localized between amino acids 9 and 29 and between amino acids 64 and 97 of protein C (Wolinsky *et al.*, 1991). Human sera reacted with peptides containing amino acids 1 to 30 and 96 to 123 of protein C (Ou *et al.*, 1992a). The C_{96}–C_{123} peptide was also recognized by rabbit polyclonal anti-rubella virus serum. The epitopes of all of the anti-E2 monoclonals isolated by Wolinsky *et al.* (1991) are located between amino acids 51 and 105 of the E2 protein (Wolinsky *et al.*, 1991; J. Wolinsky, personal communication) whereas the epitope recognized by the neutralizing anti-E2 monoclonal isolated by Green and Dorsett (1986) is located between amino acids 1 and 26 of E2 (O'Brien, 1989). Although most of the anti-E1 monoclonal antibodies recognize nonlinear epitopes, the linear epitopes recognized by some anti-E1 monoclonal antibodies have been finely mapped. Terry *et al.* (1988) defined three nonoverlapping epitopes between amino acids 245 and 285 of E1 that are recognized by monoclonals that exhibit neutralization and HI or neutralization alone. Chaye *et al.* (1992) and Wolinsky *et al.* (1991, 1993) each defined overlapping epitopes between amino acids 214 and 240 of E1 that were recognized by monoclonal antibodies that exhibited HI, neutralization, or neutralization and weak HI.

Peptides corresponding to the region of E1 containing all of the defined monoclonal antibody-reactive epitopes (amino acids 214 through 285) were recognized by human sera in enzyme-linked immunoassays (Lozzi *et al.*, 1990; Neri *et al.*, 1991; Mitchell *et al.*, 1992; Ilonen *et al.*, 1992). Thus, this region of E1 may be a major neutralizing epitope recognized by human antibodies, a finding of considerable significance in the development of peptide vaccines. Inoculation of rabbits with a peptide corresponding to amino acids 208 to 239 of E1 resulted in induction of neutralizing (but interestingly not HI) antibodies

(Wolinsky *et al.*, 1993). However, as of yet the neutralizing capacity of human antibodies recognizing these epitopes has not been confirmed. Studies in which the reactivity of human sera to peptides representing the entirety of E1 revealed the presence of major antibody-binding domains outside of this one region; however, whether any of these regions contain neutralizing epitopes is not known (Ilonen *et al.*, 1992; Newcombe *et al.*, 1993). As discussed above, evidence generated on the strain specificity of neutralizing antibodies indicated that epitopes recognized by human neutralizing antibodies are also present on E2. Other evidence supporting this conclusion is that rubella virions digested with concentrations of trypsin that destroy E1, but not E2, retain the capacity to combine with human neutralizing antibodies as measured in a neutralization competition assay (Ho-Terry and Cohen, 1980). Thus, it would also be of interest to measure the reactivity of human sera with peptides corresponding to the sequence of E2, particularly the epitope recognized by the neutralizing anti-E2 monoclonal antibody ($E2_1$–$E2_{26}$), and to determine if E2 peptides are capable of inducing neutralizing antibodies in rabbits.

Progress has also been made on definition of epitopes on the rubella virion proteins recognized by the human cell-mediated immune system. On the basis of both immunoproliferative assays and class II-restricted cytotoxic assays, immunodominant epitopes have been recognized that are located between amino acids 14–29 and 255–280 of C, 54–74 of E2, and 273–284 and 402–422 of E1 (McCarthy *et al.*, 1993; Ou *et al.*, 1992a–c, 1993; Ilonen *et al.*, 1992). It is worth noting that the $E1_{273-284}$ peptide contains a neutralization and HI epitope defined by monoclonal antibodies. However, none of these epitopes was universally recognized by all donors and peptides containing other regions of the virion proteins were recognized by some donors. It has been shown that cells expressing C protein are recognized by $CD8^+$ cytotoxic cells from a large percentage of donors in a class I-restricted assay and an immunodominant epitope at the amino terminus of C (amino acids 9 to 29) was defined (Lovett *et al.*, 1993).

V. Genomic RNA: Structure, Sequence, and Coding Strategy

The genomic RNA of rubella virus, which has a sedimentation coefficient of 38S to 40S (Hovi and Vaheri, 1970a; Sedwick and Sokol, 1970), is 9756 nucleotides in length excluding the 3′ terminal poly(A) tract (Dominguez *et al.*, 1990). The RNA was shown by Hovi and Vaheri (1970a) and Sedwick and Sokol (1970) to be infectious in BHK-21 cells, using DEAE-dextran-mediated transfection; both studies noted that

the relative infectivity was low in comparison to the RNAs of other viruses, although comparative specific infectivities were not reported. In contrast, we find that the infectivity of the rubella virion RNA is roughly equal to that of Sindbis virion RNA, using lipofectin-mediated transfection of Vero cells (Wang *et al.*, 1994).

Oker-Blom *et al.* (1984) reported the presence of two products in complete RNase digests of ^{32}P-labeled RNA extracted from purified rubella virions that had properties expected for 7-methyl guanosine cap structures [m^{7}G(5′)ppp(5′)Np] found on the 5′ terminus of most eukaryotic messenger RNAs. The penultimate nucleotide (N) of these products was not determined and neither product comigrated in high-voltage paper electrophoresis with the m^{7}G(5′)ppp(5′)Ap cap structure isolated from Semliki Forest virus that was used as a control. The presence of two products could have been due to variable ribose methylation on the penultimate or subsequent nucleotides, a feature found on many cell mRNAs and negative-polarity RNA virus mRNAs, but not on alphavirus RNAs (Banerjee, 1980). Alternatively, one of the two products could have been from the rubella virus subgenomic RNA that frequently is present in virion RNA preparations (Sedwick and Sokol, 1970) and was observable in the gradients of virion RNA used by these authors. Consistent with this latter possibility, the penultimate nucleotides of the genomic and subgenomic RNAs are C and U, respectively (Dominguez, 1991; Frey *et al.*, 1989). Recently, it was shown that the subgenomic RNA is packaged into virus particles (Wang *et al.*, 1994). The poly(A) tract at the 3′ end of the rubella virus genomic RNA is heterogeneous in length, the mean length being 53 nucleotides (Wang *et al.*, 1994). The mean length of the poly(A) tract on alphavirus genomic RNAs is 68 nucleotides (Frey and Strauss, 1978).

The complete sequence of the genomic RNA of rubella virus has been determined for two strains: (1) the Therien strain [Frey *et al.* (1986); Frey and Marr (1988); Dominguez *et al.* (1990) (accession number in the GenBank Library of Nucleotide Sequences is M15240); the sequence of the structural protein-coding region of the Therien strain was also determined by Vidgren *et al.* (1987) and Takkinen *et al.* (1988) (D00242 and D00156)] and (2) the M33 strain [Clarke *et al.* (1987) (X05259; S. Gillam (personal communication) (X72393)]. Additionally, the structural protein-coding region sequence was determined for the HPV-77 vaccine strain (which was derived from the M33 strain) [Zheng *et al.* (1989) (M30776)], and the RA 27/3 vaccine strain (the parent strain of which no longer exists) [Nakhasi *et al.* (1989) (X14871)]. The base composition is 14.9% A, 15.4% U, 30.8% G, and 38.7% C, the G + C content of 69.5% being the highest of any RNA virus thus far sequenced. The virus RNA with the next highest G + C

content in the GenBank and European Molecular Biology Organization (EMBO) repositories is that of Semliki Forest virus, which is 53% G + C.

Needless to say, the high G + C content made sequence determination of the rubella genome extremely difficult and sequencing errors are present in the original sequences reported for the Therien (Frey *et al.*, 1986) and M33 strains (Nakhasi *et al.*, 1986; Clarke *et al.*, 1987). Most of these errors have since been corrected in the literature (Frey and Marr, 1988; Zheng *et al.*, 1989; Clarke *et al.*, 1988); however, at the time of writing the J02620 and X05259 GenBank entries had not been completely corrected. A sequencing error of note is the presence of a CG dinucleotide at nucleotide 6292, which was omitted from the Therien strain sequence (Dominguez *et al.*, 1990). In the presence of these additional nucleotides, the 5′ proximal ORF uses different termination codons than originally deduced. Consequently, the ORF is 90 amino acids shorter than originally reported and the 5′ and 3′ proximal ORFs are separated by 123 nucleotides and do not overlap, as was originally reported.

A diagram of the positive-polarity orientation of the rubella virus genome, including the position of the two long ORFs and the coding arrangement of both proteins and amino acid motifs within these ORFs, is shown in Fig. 3. The 5′ proximal ORF is 6345 nucleotides in length, extending from nucleotide 41 to nucleotide 6385, whereas the 3′ proximal ORF is 3189 nucleotides in length, extending from nucleotide 6506 to nucleotide 9694. These ORFs are in the same translational frame and are separated by 123 nucleotides. As discussed above, the 3′ proximal ORF encodes the structural proteins; the order of the coding sequences for the structural proteins within this ORF is NH_2-C-E2-E1-COOH [this order was determined by Oker-Blom (1984a) before the sequence was available]. The start site for the subgenomic RNA from which the structural protein ORF is translated is at nucleotide 6430 of the genomic RNA (Frey *et al.*, 1989). The 5′ proximal ORF codes for nonstructural proteins, including proteins involved in RNA replication, as indicated by the presence of global amino acid motifs in the deduced amino acid sequences indicative of RNA-dependent RNA polymerase and helicase activity that are found in a large number of positive-polarity RNA viruses, including the 5′ proximal ORF of the alphavirus genome (Kamer and Argos, 1984; Gorbalenya *et al.*, 1988). The polyprotein product of the alphavirus 5′ proximal ORF is cleaved into four nonstructural proteins (nsPs) by a cysteine protease activity present within the polyprotein (reviewed in Strauss and Strauss, 1990). Gorbalenya *et al.* (1991) located a putative cysteine protease within the deduced amino acid sequence of the

rubella virus 5′ proximal ORF that shares homology with the alphavirus nonstructural protease and several cell papain-like proteases. As is discussed in Section VI,C, below, this protease is active and catalyzes at least one proteolytic cleavage of the polyprotein product of the rubella virus 5′ proximal ORF. Interestingly, a direct homology comparison of the deduced amino acid sequences of the 5′ proximal ORFs of rubella virus and alphaviruses reveals only one short (50 amino acids) region of homology (Dominguez *et al.*, 1990). This region, termed the X motif, is located at the amino terminus of alphavirus nsP3 and thus far has been found only in the genomes of the alphaviruses, rubella virus, hepatitis E virus, and possibly the coronaviruses (Gorbalenya *et al.*, 1991; Koonin *et al.*, 1992).

There are ORFs of significant length in the negative-polarity orientation of the rubella virus genome (Dominguez *et al.*, 1990). The longest of these occupies the nucleotides complementary to those that encode the structural protein ORF in the positive-polarity orientation and is potentially translated into a polypeptide 924 amino acids in length. In addition, there are five ORFs that potentially encode polypeptides 200 amino acids or greater in length. The unusual base composition of the rubella virus RNA would make the random occurrence of termination codons less frequent than in an RNA with a lower G + C content. However, the expected random frequency of occurrence of termination codons in an RNA with the base composition of the rubella virus negative-polarity RNA is 1 in 47 (as compared with 1 in 21 in an RNA in which the frequency of each nucleotide was 0.25) and thus these ORFs are much longer than would be expected if the presence of termination codons in the negative-polarity orientation were occurring by random chance. There is no evidence that any of the negative-polarity ORFs are translated.

The high G + C content of the rubella virus genomic RNA has a pronounced impact on the codon usage in the two positive-polarity ORFs. In Table I is shown a comparison of codon usage in the rubella virus ORFs (codon usage in the 5′ proximal and 3′ proximal ORFs is similar and therefore these ORFs have been combined) with the homologous ORFs from the prototype alphavirus, Sindbis virus (Strauss and Strauss, 1986), and a compilation of human genes (Maruyama *et al.*, 1986). As expected, use of G + C-rich codons is favored in the rubella virus ORFs, particularly codons that end in G or C. The nucleotide frequencies at each codon position in the rubella virus and Sindbis virus ORFs and the human genes are given in Table II. Selection of C and G residues at third-codon positions is a recognized tendency in eukaryotic genes (reviewed in Ikemura, 1985) and in the coding regions from all three sources, the percentage of C residues at third-

TABLE I

CODON USAGE IN RUBELLA VIRUS AND SINDBIS VIRUS OPEN READING FRAMES AND IN HUMAN GENES

Amino acid	Codon	RUB[a]	SIN[b]	HUM[c]	Amino acid	Codon	RUB	SIN	HUM
Arg	CGA	0.44	0.37	0.52	Val	GUA	0.25	1.94	0.62
	CGG	1.50	0.45	0.77		GUG	2.48	1.99	3.09
	CGC	6.03	1.22	1.11		GUC	3.18	2.29	1.62
	CGU	0.63	0.66	0.36		GUU	0.82	1.28	1.06
	AGA	0.16	1.65	1.08					
	AGG	0.35	0.96	1.14	Lys	AAA	0.41	3.00	2.20
						AAG	1.01	3.46	3.58
Leu	CUA	0.16	1.09	0.61					
	CUG	2.52	2.31	4.37	Asn	AAC	1.16	2.39	2.36
	CUC	3.96	1.06	2.20		AAU	0.57	1.25	1.69
	CUU	0.79	1.14	1.07					
	UUA	0.15	0.48	0.59	Gln	CAA	0.82	1.43	1.15
	UUG	0.69	1.62	1.15		CAG	2.14	2.07	3.27
Ser	UCA	0.19	1.30	0.92	His	CAC	2.58	1.38	1.43
	UCG	0.72	1.41	0.43		CAU	1.00	1.30	0.98
	UCC	0.89	0.98	1.87					
	UCU	0.41	0.69	1.43	Glu	GAA	1.04	3.16	2.68
	AGC	1.64	1.43	2.05		GAG	3.99	2.60	4.16
	AGU	0.25	0.96	0.87					
					Asp	GAC	4.15	3.06	3.02
Thr	ACA	0.35	2.07	1.41		GAU	0.91	1.70	2.14
	ACG	1.10	1.25	0.55					
	ACC	3.74	2.55	2.51	Tyr	UAC	2.30	2.07	1.84
	ACU	0.88	1.38	1.34		UAU	0.22	1.25	1.24
Pro	CCA	1.16	2.13	1.18	Cys	UGC	3.11	2.10	1.51
	CCG	2.58	1.94	0.60		UGU	0.28	0.59	1.00
	CCC	4.25	1.22	1.85					
	CCU	1.48	1.28	1.43	Phe	UUC	1.79	1.86	2.57
						UUU	0.35	1.62	1.73
Ala	GCA	0.97	2.55	1.27					
	GCG	4.09	1.73	0.59	Ile	AUA	0.19	1.17	0.51
	GCC	7.20	3.11	2.98		AUC	1.57	2.07	2.40
	GCU	1.64	1.28	1.43		AUU	0.57	1.62	1.35
Gly	GGA	0.50	2.37	1.46	Met	AUG	1.48	2.29	2.51
	GGG	1.82	1.01	1.37					
	GGC	5.28	1.36	2.43	Trp	UGG	2.48	0.98	1.42
	GGU	0.57	0.85	1.03					
					Ter	UAA	0.03	0.00	0.11
						UAG	0.03	0.03	0.05
						UGA	0.00	0.05	0.17

[a] Codon usage in the rubella virus nonstructural protein and structural protein ORFs combined. Codon usage in the two ORFs is similar. Usage of each codon is given as the percentage of total codon usage. Data from Dominguez (1991).

[b] Codon usage in the Sindbis virus nonstructural protein and structural protein ORFs combined. Codon usage in the two ORFs is similar. Data from Strauss and Strauss (1986).

[c] Codon usage in 135 human genes as compiled by Maruyama *et al.* (1986).

TABLE II
NUCLEOTIDE FREQUENCIES AT DIFFERENT CODON POSITIONS IN RUBELLA VIRUS AND SINDBIS VIRUS OPEN READING FRAMES AND IN HUMAN GENES

Frequency[a]	A	C	G	U
Rubella virus				
Overall[b]	0.15	0.39	0.31	0.15
First position	0.15	0.32	0.39	0.14
Second position	0.22	0.32	0.25	0.21
Third position	0.07	0.53	0.29	0.11
Sindbis virus				
Overall[b]	0.28	0.26	0.25	0.21
First position	0.30	0.21	0.32	0.17
Second position	0.30	0.27	0.17	0.26
Third position	0.25	0.30	0.26	0.19
Human genes				
Overall	0.25	0.26	0.26	0.22
First position	0.28	0.23	0.31	0.18
Second position	0.32	0.22	0.18	0.28
Third position	0.17	0.34	0.29	0.20

[a] Compiled from data in Table I.
[b] Base composition within ORFs (excluding nontranslated sequences).

codon positions is elevated in comparison with the overall percentage of C residues (the selection is proportionately greater in the rubella virus ORFs and the human genes than in the Sindbis virus ORFs). Interestingly, only in the human genes is a selection of G residues at third-codon positions apparent. The rubella virus and Sindbis virus ORFs and the human genes also share an excess of Gs at the first codon position and of Us at the second codon position and a deficiency of Gs at the second codon position. The rubella virus ORFs and the human genes share a deficiency of As at third codon positions, a tendency not present in the Sindbis virus ORFs.

In the rubella virus ORFs, there is also a selection among isofunctional amino acids toward those that are encoded by G + C-rich codons. Proportional usage among five classes of isofunctional amino acids in the rubella virus and Sindbis virus ORFs and the human genes is shown in Table III. The most dramatic example of selection for amino acids with G + C-rich codons in the rubella virus ORFs is in the basic amino acids, 87% of which are arginine (codons CGN, AGA, and AGG)

TABLE III

USAGE OF ISOFUNCTIONAL AMINO ACIDS IN RUBELLA VIRUS AND SINDBIS VIRUS OPEN READING FRAMES AND HUMAN GENES

Virus/gene	Isofunctional amino acids for which codons exhibit GC-rich vs AU-rich distribution[a]				
	Basic (Arg/Lys)	Arg (CGN, AGPu)	Lys (AAPu)		
Rubella	0.105[b]	87%[c]	13%[c]		
Sindbis	0.118	45%	55%		
Human	0.108	46%	54%		
	Polar (Gln/Asn)	Gln (CAPu)	Asn (AAPy)		
Rubella	0.047	63%	37%		
Sindbis	0.071	49%	51%		
Human	0.085	52%	48%		
	Hydrophobic (Ala/Leu/ Val/Ile)	Ala (GCN)	Leu (CUN, UUPu)	Val (GUN)	Ile (AUA/C/U)
Rubella	0.312	44%	27%	22%	7%
Sindbis	0.287	30%	27%	26%	17%
Human	0.269	23%	37%	24%	16%
Virus/gene	Isofunctional amino acids for which codons exhibit no GC-rich vs AU-rich distribution				
	Acidic (Asp/Glu)	Asp (GAPy)	Glu (GAPu)		
Rubella	0.101	50%	50%		
Sindbis	0.105	45%	55%		
Human	0.120	43%	57%		
	Polar (Ser/Thr)	Ser (UCN, AGPy)	Thr (ACN)		
Rubella	0.102	40%	60%		
Sindbis	0.140	48%	52%		
Human	0.134	57%	43%		

[a] Amino acid encoded by GC-rich codons listed first.

[b] Fraction of total amino acid composition composed by isofunctional group. Compiled from data in Table II.

[c] Percentage usage within isofunctional group.

and only 13% of which are lysine (codons AAA and AAG). In both the Sindbis virus genome and the human genes, roughly 45% of the basic residues are arginine and 55% are lysine. In the rubella virus ORFs, there is also a selection of glutamine (CAG or CAA) over asparagine (AAC or AAU) and, among hydrophobic amino acids, alanine (GCN) is selected for whereas isoleucine (AUA, AUC, or AUU) is selected against. That these selections are driven by nucleotide content is apparent when the acidic amino acids are considered. The codons for glutamic acid are GAA and GAG and for aspartic acid are GAC and GAU and thus are of similar G + C content. Within the rubella virus ORFs, 50% of the acidic residues are glutamic acid and 50% are aspartic acid.

A tendency in nucleotide frequency in the genomes of higher vertebrates is a deficit in the frequency of the dinucleotide CG; this tendency is not found in the genomes of invertebrates or plants (reviewed in Strauss *et al.*, 1990). The CG deficit may be related to the function of CG as a signal for methylation of C, leading to chromatin inactivation. The deficiency of CG dinucleotides affects codon usage in the mRNAs of higher vertebrates. Interestingly, a CG deficit is also present in the genomes of some RNA viruses of higher vertebrates such as poliovirus and all of the negative-polarity RNA viruses, whereas the genomes of plant RNA viruses rarely exhibit a CG deficit. Because a CG deficit in the genome of an RNA virus could not involve methylation, it is proposed that it is due to codon usage. However, other viruses of higher vertebrates, such as the aphthoviruses (foot-and-mouth disease viruses) and the coronaviruses, do not exhibit a CG deficiency and occasional plant viruses, such as cowpea mosaic virus, do exhibit a CG deficiency. Among viruses that replicate in both invertebrates and higher vertebrates, the flaviviruses exhibit a pronounced genomic CG deficiency whereas the alphaviruses exhibit a modest genomic CG deficiency (Weaver *et al.*, 1993). In the rubella virus genome, the frequency of CG dinucleotides is as expected from the base composition.

VI. Virus Replication Cycle

A. Permissive Cell Lines and Replication Curve Characteristics

Rubella virus replicates in a number of primary cell cultures and continuous cell lines of vertebrates. Of historic interest, when rubella virus was initially shown to have a togavirus-like morphology, it was speculated that rubella virus might be an arbovirus (Holmes and Warburton, 1967; Carver and Marcus, 1968). However, rubella virus fails

to replicate in mosquitoes (Tesh and Rosen, 1975). In most types of vertebrate culture cells, rubella virus replicates to low titers and induces little or no cytopathic effect (CPE). The molecular characterization of rubella virus has been undertaken almost exclusively in two continuous cell lines, BHK-21 cells (hamster kidney) and Vero cells (African Green monkey kidney). BHK-21 cells, in which titers of 10^7 plaque-forming units (PFU) or 50% tissue culture infectious doses ($TCID_{50}$)/ml can be produced (Maes *et al.*, 1966; Vaheri *et al.*, 1967), were used almost exclusively in studies done through 1979. However, the demonstration that titers of 10^8 PFU or $TCID_{50}$/ml can be produced in Vero cells (Bardeletti *et al.*, 1979; Oker-Blom *et al.*, 1983) led to conversion to that line. Plaque assays can be done using BHK-21, Vero, and RK-13 (rabbit kidney) cells (Vaheri *et al.*, 1967; Rhim and Schell, 1967; Taylor-Robinson *et al.*, 1964). Production of low virus titers in RK-13 cells precludes their use in molecular characterization experiments (Maes *et al.*, 1966).

The two cell lines in which rubella virus replication is most productive, BHK-21 and Vero, both lack a functional interferon system (Mifune *et al.*, 1970). This is an important factor in the ability of these cell lines to support rubella virus replication to high titers because rubella virus replication is substantially reduced in cells pretreated with interferon (Wong *et al.*, 1967; Stanwick and Hallum, 1974; Nakhasi *et al.*, 1988). The former two of these studies used undifferentiated interferon from induced cell cultures; in the latter study, using recombinant interferons, it was shown that rubella virus replication was sensitive to the action of both interferon α and γ; however, the effect of interferon β was not analyzed. In interferon-competent primary African Green monkey kidney (AGMK) cells infected with rubella virus at a multiplicity of infection (MOI) of 0.1 infectious doses per cell or less, interferon is induced leading to reduced yields of virus because multiple rounds of virus replication are inhibited (the yields of virus are directly proportional to the MOI) (Wong *et al.*, 1967). Interestingly, interferon is also induced in primary AGMK cells infected at an MOI of 10 infectious doses per cell and the culture becomes refractory to infection by heterologous viruses. Thus, an antiviral state can be induced in cells already infected with rubella virus. Presumably, the antiviral state induced after rubella virus infection subsequently interferes with ongoing rubella virus replication, because it has been shown that when Vero cells infected with rubella virus are treated with exogenous interferon, virus replication is reduced (Wong *et al.*, 1967; Mifune *et al.*, 1970; Stanwick and Hallum, 1974). Despite the sensitivity of rubella virus replication to interferon, the virus is adapted to survival in the presence of interferon because persistent infec-

tions have been established in a number of interferon-competent cell lines in which detectable interferon is present in the persistently infected culture fluid and a majority of the cells in the persistently infected culture are infected (Mifune *et al.*, 1970; Stanwick and Hallum, 1974).

The induction of interferon by rubella virus is the most likely mechanism behind the phenomenon of rubella virus interference with the replication of a number of heterologous viruses, a phenomenon that was recognized in the study of Parkman *et al.* (1962), in which the isolation of rubella virus was originally reported. A number of studies on rubella virus interference in interferon-competent cell lines led to an overall connection between interferon induction and interference (reviewed in Desymter *et al.*, 1969; Horzinek, 1981). However, the connection was not completely conclusive and some investigators concluded that rubella virus expressed an "intrinsic interference" against some viruses. However, the finding that interference is not exhibited in interferon-incompetent Vero cells seemingly confirms the connection (Desmyter *et al.*, 1969; Stanwick and Hallum, 1974). Of clinical relevance, the standard initial laboratory assay to detect the presence of rubella virus consists of inoculation of primary African Green monkey kidney cell cultures with a specimen followed after incubation for 1 week to 10 days by challenge with an enterovirus such as coxsackievirus A9, echovirus 11, or echovirus 40 (Herrmann, 1979). Sometimes challenge follows an intermediate passage of the fluid from the inoculated culture. Interference with enterovirus-induced CPE provides presumptive evidence for the presence of rubella virus. Definitive identification of rubella virus requires neutralization of interference with anti-rubella virus serum or analysis of infected cultures using a specific immunodiagnostic assay (immunofluorescence or immunoperoxidase staining).

A fascinating (and, to the molecular virologist, frustrating) feature of rubella virus replication in cell culture is that all of the cells in a cell culture are not infectable at any given time, even in the most permissive cell lines. This is best exemplified by infectious center experiments done at 2 to 4 hr postinfection on cultures infected with an MOI of 5 to 10 PFU/cell in which less than 100% of the cells are found to give rise to infectious centers. The lack of uniform infection is also apparent in immunofluorescence studies because at the time postinfection of initial detectability of virus proteins, only a fraction of the cells are positive. In BHK-21 cells, the percentage of initially infectable cells is in the range of 10% (Sedwick and Sokol, 1970; Wong *et al.*, 1969) whereas in Vero cells the percentage is on the order of 50% (Hemphill *et al.*, 1988). By 24 to 96 hr postinfection, a much higher

percentage of the cells is infected as detected by either infectious center assay or immunofluorescence (in BHK-21 cells, the percentage of infected cells approaches 50%, whereas in Vero cells the percentage approaches 100%). Thus, cells not initially infectable become infectable during the course of the experiment. Experimentally, the limited infectability means that, even at high input MOIs, a synchronous infection is not achieved. The basis of the phenomenon is unknown but presumably involves a cellular component that is present in limiting but fluctuating amounts in cells in a culture.

The replication rate of rubella virus is slow in comparison to rates of other viruses. In BHK-21 cells, Vero cells, and RK-13 cells infected with MOIs of 5 PFU or $TCID_{50}$/cell or greater, the eclipse period is at least 12 hr followed by a gradual rise in virus titers through 36 to 48 hr postinfection, when peak titers are obtained (Vaheri *et al.*, 1965; Maes *et al.*, 1966; Wong *et al.*, 1969; Oker-Blom, 1984b). Peak titers are maximally 30 PFU or $TCID_{50}$/cell in BHK-21 cells and 300 PFU or $TCID_{50}$/cell in Vero cells (Vaheri *et al.*, 1967; Bardeletti *et al.*, 1979). In addition to extracellular virus, intracellular virus is also recoverable. In infected RK-13 and BHK-21 cells, the titer of intracellular virus is equal to or greater than the titer of extracellular virus, whereas in infected Vero cells the titer of intracellular virus is only 1–2% of the extracellular titer (Maes *et al.*, 1966; Bardeletti *et al.*, 1979). If infected cultures are maintained, the peak virus titers are maintained for several days or longer (Maassab and Veronelli, 1966; Vaheri *et al.*, 1967; Wong *et al.*, 1967). Cytopathic effect is initially detectable between 36 and 48 hr postinfection (Vaheri *et al.*, 1967; Hemphill *et al.*, 1988). The CPE generally consists of an increase in refractile index, cell rounding and detachment, and the presence of debris adhering to the monolayer. In infected monolayers, CPE is often concentrated in foci, the number of which is generally related to the amount of input virus. In no cell line is cell destruction complete and in all cell lines tested a persistent infection is established (Mifune *et al.*, 1970; Stanwick and Hallum, 1974).

B. Attachment and Penetration

Attachment of rubella virus to susceptible cells is relatively rapid. In studies using BHK-21 cells, Bardeletti *et al.* (1972) found that 99% of input virus disappeared from the inoculum during a 30-min adsorption period whereas Vaheri *et al.* (1967) found that in a plaque assay with a 1-hr adsorption the number of plaques formed was 80% of the number of plaques formed with a 3-hr adsorption. The cellular receptor for rubella virus has not been identified. In one study, anti-idiotypic anti-

bodies raised against anti-E1 monoclonal antibodies that exhibit HI and/or neutralization failed to bind to either 1-day-old chick erythrocytes or Vero cells or to react with protein extracts of either cell type on Western blots (Nath *et al.*, 1989). In another study, it was shown that the lipid fraction, but not the protein fraction, of solubilized goose erythrocyte membranes inhibited hemagglutination (Mastromarino *et al.*, 1989). Hemagglutination inhibition activity was exhibited specifically by glycolipid and phospholipid extracted from the erythrocyte membranes, and it was found that purified phosphatidylserine and cerebroside sulfate, but not a variety of other phospho- and glycolipids, exhibited HI. In an extension of this study, it was found that treatment of Vero cells with high concentrations of phospholipases A_2 and C completely inhibited rubella virus infection whereas treatment with a variety of proteases and glycosidases only modestly inhibited rubella virus infection (Mastromarino *et al.*, 1990). Incubation of rubella virions with a number of phospho- and glycolipids (most effectively phosphatidylserine and phosphatidylinositol) also inhibited infection. The implication of these experiments is that the cellular receptor for rubella virus could be a lipid. However, it is not known if treatment of virions with lipids could nonspecifically disrupt virion structure and whether treatment of cells with phospholipases could cause rearrangements to plasma membrane structure deleterious to virion binding. In this regard, it would be of great interest to determine the effect of these procedures using a virus known to have a proteinaceous cell receptor (e.g., Sindbis virus).

The majority of enveloped animal viruses enter the cell by receptor-mediated endocytosis (reviewed in Wiley, 1986). Following fusion of the endocytosed vesicle containing the virus with an endosome, the low-pH environment of the endosome triggers exposure of the fusogenic activity of the viral glycoproteins, causing fusion of the virus membrane with the endosomal membrane and resulting in the release of the viral nucleocapsid into the cytoplasm. The requisite experiments to prove that rubella virus enters the cell by receptor-mediated endocytosis have not been done. However, exposure of the rubella virus glycoproteins to pH 6.0 or less exposes a fusogenic activity, which is consistent with entry by receptor-mediated endocytosis (Katow and Suguira, 1988b). Interestingly, nonionic detergent extraction of rubella virions at a pH of 5.0 or less results in partitioning of the C protein into the detergent phase (Mauracher *et al.*, 1991). Presumably the C protein undergoes a conformational change at low pH that renders it hydrophobic in nature (it is tempting to speculate that this is due to exposure of the COOH-terminal hydrophobic sequence). The genomic RNA is released into the aqueous phase of the extraction,

indicating that nucleocapsid disassembly occurred. Thus, the low-pH environment of the endosome possibly triggers uncoating of the genome as well as membrane fusion. Presumably, following exposure to the low pH of the endosome, the C protein adheres to the interior surface of the viral envelope, which becomes the exterior surface of the endosome following membrane fusion.

C. Translation of 5′ Proximal Open Reading Frame

Following dissociation of the capsid and release of the genomic RNA into the cytoplasm, the virion RNA is translated to produce the 2115-amino acid polypeptide encoded by the 5′ proximal ORF. Although proteolytic cleavage of this polypeptide is predicted by the alphavirus model and a cysteine protease motif is present within the deduced amino acid sequence of the polypeptide, evidence that this polypeptide is cleaved was difficult to obtain. Because of the limited amounts of virus-specific proteins synthesized in infected cells and the lack of inhibition of cell protein synthesis, neither virus structural nor nonstructural proteins can routinely be visualized over the cell background in polyacrylamide gels of radiolabeled infected cell lysates (Hemphill *et al.*, 1988). Using human convalescent serum to immunoprecipitate proteins from infected cell lysates radiolabeled under hypertonic salt conditions to selectively favor translation of virus proteins, Bowden and Westaway (1984) were able to detect proteins with molecular weights of 200,000, 150,000, 87,000, 75,000, and 27,000, which were present in addition to the structural proteins.

The rubella virus 5′ proximal ORF was successfully expressed in transfected cells by using a vector [pTM3 (Moss *et al.*, 1990)] in which the 5′ proximal ORF sequences were placed downstream from a T7 RNA polymerase promoter and the encephalomyocarditis virus cap-independent translation initiation sequences (Marr *et al.*, 1994). T7 RNA polymerase is provided by infection of transfected cells with a vaccinia virus recombinant that expresses T7 RNA polymerase. Rubella virus-specific products with molecular weights of 200,000, 150,000, and 90,000 were clearly resolved (products with lower molecular weights were obscured by the presence of vaccinia virus proteins). The size of these products is close to those of the largest three products observed in rubella virus-infected cells by Bowden and Westaway (1984). The catalytic residue of the predicted protease within the deduced amino acid sequence of the 5′ proximal ORF was a cysteine at residue 1151 (Gorbalenya *et al.*, 1991). Mutagenesis to change this cysteine to a glycine resulted in the production of only the 200-kDa species in transfected cells. This result confirmed that the 200-kDa

species was the polyprotein precursor translated from the 5′ proximal ORF, that processing of the precursor occurred, that the Cys at residue 1151 was important in protease activity, and that the 150- and 90-kDa species were processing products. Deletion mutagenesis showed that the 150-kDa product was derived from the amino-terminal region of the polyprotein precursor.

More recently, antibodies raised against bacterial fusion proteins containing regions encoded by the 5′ proximal ORF were used for immunoprecipitation experiments with lysates from cells in which the 5′ proximal ORF was expressed from the pTM3 vector and from rubella virus-infected cells (R.-Y. Forng, unpublished observations). In both systems, the 200-kDa species was immunoprecipitated by all of the antibodies, the 150-kDa species was immunoprecipitated by antibodies against fusion proteins containing sequences from the amino-terminal region of the ORF, and the 90-kDa species was immunoprecipitated by antibodies against fusion proteins containing sequences from the COOH-terminal region of the ORF. Thus, the order of these species within the 200-kDa precursor is NH_2-P150-P90-COOH. The combined molecular weight of these two species (240 kDa) is similar to that predicted from the deduced amino acid sequence encoded by the ORF (237 kDa) and thus these are probably the only species processed from the precursor. (Parenthetically, proteins greater than 200,000 in molecular weight often migrate more rapidly than expected in gels and thus the molecular weight of the 200,000 of the precursor in relation to the calculated molecular weight of 237,000 is not unexpected). From the relative sizes of these two products, the cleavage site within the precursor is COOH terminal to the protease domain (Fig. 3). Thus, the X motif and the protease reside on the 150-kDa product and the helicase and replicase motifs reside on the 90-kDa product. Thus far, even with the use of prolonged radiolabeling periods, further processing products have not been convincingly demonstrated. This is in contrast to the findings of Bowden and Westaway (1984); however, it is possible that the 75- and 27-kDa products these researchers observed were not from the 5′ proximal ORF.

Interestingly, the AUG initiating the 5′ proximal ORF [nucleotides 41 to 43 of the genomic RNA (AUG_{41-43})] is not the 5′-most AUG on the genomic RNA. An AUG at nucleotides 3–5 (AUG_{3-5}) that is in a different translation frame potentially encodes a 17-amino acid product (see Fig. 4). Because initiation of translation on eukaryotic mRNAs generally occurs at the AUG nearest the 5′ cap (Kozak, 1989), the presence of AUG_{3-5} could limit translation of the 3′ proximal ORF. However, it has been found that initiation of translation of eukaryotic

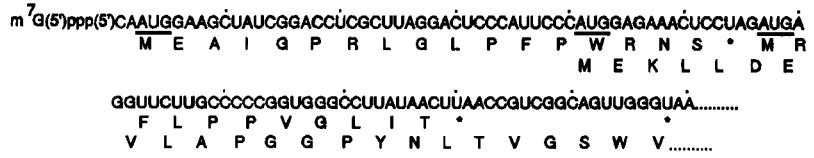

FIG. 4. 5′ Terminal sequence of the rubella virus genome RNA. The deduced amino acid sequences (one-letter code) encoded by potential ORFs are given. The ORF initiated by the second AUG from the 5′ end of the genome (nucleotides 41–43) is the long 5′ proximal ORF shown in the genome diagram in Fig. 3 and encodes the nonstructural proteins. *, termination codon.

mRNAs at AUGs fewer than 15 nucleotides from the 5′ end is extremely inefficient, resulting in relatively efficient usage of downstream AUGs (Sedman *et al.*, 1990). It should be noted, however, that even though initiation at AUG_{3-5} is probably an infrequent event, potential translation from AUG_{3-5} is tightly controlled due to the presence of in-frame termination codons 18, 30, and 36 codons downstream from AUG_{3-5}. The in-frame codon following the first of these termination codons is an AUG [nucleotides 57–59 (AUG_{57-59})] resulting in three AUGs within the 5′ terminal 60 nucleotides of the rubella virus genome.

The initiation of translation at the AUGs present at the 5′ end of the rubella virus genomic RNA was studied by insertion of the 5′ terminal 65 nucleotides into vectors containing the coding sequences for bacterial chloramphenicol acetyltransferase (CAT) (Pogue *et al.*, 1993). *In vitro* translation of *CAT* from RNA transcripts synthesized from a construct in which the *CAT* ORF was in frame with AUG_{41-43} was preferentially initiated at AUG_{41-43} rather than at the *CAT* AUG (the products could be distinguished on the basis of size), whereas translation of *CAT* from RNA transcripts from a construct in which the *CAT* ORF was in frame with AUG_{57-59} was preferentially initiated at the *CAT* AUG (AUG_{41-43} was deleted in this latter construct). In Vero cells transfected with a vector that expresses *CAT* under control of the adenovirus major late promoter, the presence of the rubella virus 5′ terminal sequences upstream from the *CAT* ORF such that they were in frame with AUG_{41-43} increased expression of CAT twofold in comparison to the native *CAT* upstream sequences. Thus, initiation of translation at AUG_{41-43} is relatively efficient both *in vivo* and *in vitro*, whereas initiation of translation at one of the alternate AUGs was not. Interestingly, insertion of the 3′ terminal 165 nucleotides of the rubella virus genome [including 27 residues of the poly(A) tail] down-

stream from the *CAT* gene was necessary for successful translation of transcripts *in vitro* (irregardless of whether rubella virus 5′ terminal sequences were present in the construct) but not for expression *in vivo*.

D. RNA Synthesis

1. RNA Species Produced

The proteins translated from the 5′ proximal ORF putatively use the infecting virion RNA as a template for the transcription of a complementary, negative-polarity RNA of genome length. A genome-length, negative-polarity RNA has been detected in infected cells and is present only in double-stranded form, indicating its function solely as a template for positive-polarity RNA synthesis (Hemphill *et al.*, 1988). Analysis of double-stranded RNA species in infected cells demonstrated the presence of both completely double-stranded forms and forms that contain both single-stranded and double-stranded RNA [termed replicative forms (RFs) and replicative intermediates (RIs), respectively] (Sedwick and Sokol, 1970). Replicative intermediates are presumably derived from double-stranded replicative complexes undergoing active transcription, because this species is preferentially labeled during short radiolabeling periods. The double-stranded RNA species of rubella virus are infectious (Sedwick and Sokol, 1970; C.-Y. Wang, unpublished observations). Data on the infectivity of alphavirus double-stranded RNA species are conflicting. In two studies it was found that the alphavirus double-stranded RNA species were infectious (Yoshinaka and Hotta, 1971; Segal and Sreevalsan, 1974) whereas in two other studies it was found that the alphavirus double-stranded RNA species were not infectious but became infectious following denaturation of the double-stranded structure (Friedman, 1968; Wengler *et al.*, 1976). Because the genome-length species is the only negative-polarity RNA routinely detectable in infected cells (Hemphill *et al.*, 1988), the subgenomic RNA is also transcribed from this template. The sequence of the subgenomic RNA is identical to the 3′ terminal 3326 nucleotides of the genomic RNA, indicating that initiation of the sub-genomic RNA occurs at an internal site on the genome-length, negative-polarity RNA (Frey *et al.*, 1989). The subgenomic and genomic RNAs are synthesized in infected cells at a molar ratio of 1.6:1 (Hemphill *et al.*, 1988).

2. Potential Regulatory Sequences

Four regions in the genomic RNAs of alphaviruses are highly conserved among alphaviruses and thus are thought to be regulatory sig-

nals for virus RNA replication (Strauss and Strauss, 1986). Significantly, stretches of nucleotides sharing homology with three of these four conserved regions are present in the rubella virus genome (Dominguez *et al.*, 1990). The presence of these regions of homology in the rubella virus genome strengthens the hypothesis that these regions are important in viral RNA replication.

The first of these regions of shared homology/structure is a predicted secondary structure occurring at the exact 5′ end of the genome, the negative-polarity complement of which is thought to serve as a recognition site for initiation of genomic RNA synthesis (possibly in conjunction with the second conserved region, the 51-nucleotide sequence, as discussed below). Site-directed mutagenesis of the Sindbis virus infectious clone showed that mutations that alter the structure were lethal or resulted in a virus that replicated poorly (Niesters and Strauss, 1990b). The comparable rubella virus and alphavirus secondary structures are shown in Fig. 5A and C, respectively. Experimental evidence suggests that the predicted rubella virus 5′ structure is formed. When primer extension is performed using genome RNA as a template and an oligonucleotide primer complementary to a nucleotide stretch downstream from the 5′ structure, two strong stop bands are produced, one corresponding to the 5′ end of the genome and one corresponding to the exact base of the stem of the 5′ structure (Dominguez, 1991). Considering the relatively low theoretical thermal stability of this structure (−20.6 kcal/mol), it is surprising that reverse transcription is impeded. As shown in Fig. 5B, it is possible to draw a pseudoknot structure pairing nucleotides in the side stem–loop of the 5′ structure with nucleotides in the 5′ single-stranded region. Formation of a pseudoknot greatly increases the stability of a stem-and-loop structure (Pleij and Bosch, 1989), possibly explaining the occurrence of the strong stop. However, there is no experimental evidence that such a pseudoknot is formed.

The second alphavirus conserved region is a stretch of 51 nucleotides beginning about 150 nucleotides from the 5′ end of the alphavirus genome. This region can be configured to form a double stem-and-loop structure and mutations in the infectious clone that disrupt this structure lead to virus with a reduced growth rate (Niesters and Strauss, 1990a). A computer-derived configuration of the secondary structure of the first 500 nucleotides of the alphavirus genome places the 51-nucleotide conserved region and the 5′ terminal structure in close proximity and thus it is thought that these two regions form a binding site for factors involved in replication. The region of the rubella virus genome sharing homology with the 51-nucleotide region contains 46 nucleotides and is located 224 nucleotides from the 5′ end of the ge-

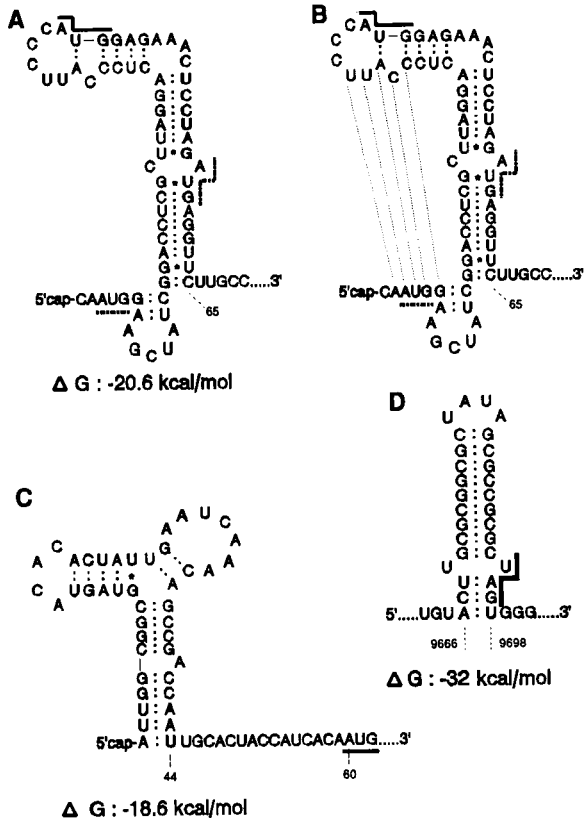

FIG. 5. Potential stem-and-loop structures in the rubella virus genome RNA. (A) Stem-and-loop structure formed by nucleotides at the 5′ end of the rubella virus genome predicted by an RNA secondary structure computer program (C. W. A. Pleij, personal communication). The AUG at the beginning of the long 5′ proximal ORF (nonstructural protein ORF) is underlined with a solid line whereas AUGs in an alternate translation frame (see Fig. 4) are underlined with dashed lines. (B) The 5′ terminal structure showing potential pairing between nucleotides in the extreme 5′ terminal single-stranded stretch and nucleotides in the side stem-and-loop shown. Such pairing would result in formation of a pseudoknot structure. (C) Stem-and-loop structure predicted to be formed at the 5′ end of the genome RNAs of alphaviruses (the prototype Sindbis virus structure is shown) (from Niesters and Strauss, 1990b). The AUG at the beginning of the nonstructural protein ORF is underlined with a solid line. (D) Predicted stem-and-loop structure formed by nucleotides near the 3′ end of the rubella virus genome RNA. The termination codon at which the structural protein ORF ends is underlined with a solid line. In (A)–(D), the numbers indicate nucleotides from the 5′ end of the genome. The stabilities (ΔG) for each structure were calculated using the method of Tinoco *et al.* (1973) [formulas have not been developed for calculating the ΔG of a pseudoknot as shown in (B)].

nome (Fig. 6A). Interestingly, the characteristics of the rubella virus 46-nucleotide region are dissimilar from those of the alphavirus 51-nucleotide region. The rubella virus 46-nucleotide region does not form a distinctive secondary structure and a computer-derived configuration of the secondary structure of the first 500 nucleotides of the rubella virus genome does not place the 46-nucleotide region and 5′ terminal structure in close proximity (Dominguez, 1991). Surprisingly, the alphavirus 51-nucleotide region and the rubella virus 46-nucleotide region are in frame translationally and encode a pocket of amino acid homology, part of which has been shown by computer-assisted alignment to be conserved in a large number of positive-polarity RNA viruses and is postulated to be functional in methyltransferase activity (Rozanov *et al.*, 1992). However, it seems doubtful that conservation of these regions of nucleotide sequence in the

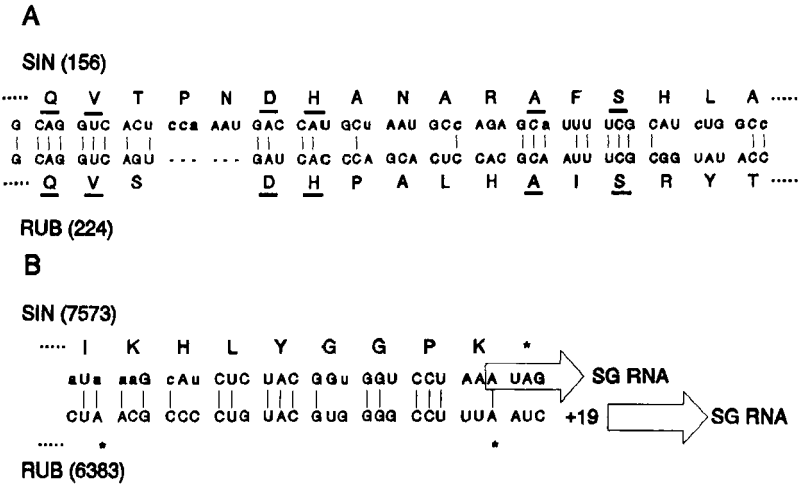

FIG. 6. Alignment of sequences in the rubella virus genome with two sequences conserved in the genomes of all alphaviruses. (A) The 51-nucleotide conserved sequence; (B) the subgenomic promoter conserved sequence. For simplicity, the prototype Sindbis virus sequence is shown, the nucleotides found in all alphavirus genome RNAs being capitalized. The nucleotides within the genome at which these sequences begin are denoted. The amino acids in the nonstructural protein ORF of each virus encoded by these sequences are shown; amino acids that are identical in the Sindbis virus and rubella virus ORFs are underlined. It is to be noted that the Sindbis virus ORF terminates within the subgenomic start site whereas the rubella virus ORF terminates at the beginning of the putative subgenomic promoter. On the diagram of the sequences at the subgenomic promoter, the 5′ end of the subgenomic RNA is denoted by an arrow (the 5′ end of the rubella virus subgenomic RNA is 19 nucleotides downstream from the putative promoter).

genomes of rubella virus and the alphaviruses due to selection for amino acid sequence has occurred, because some of the conserved nucleotides are at third-codon positions.

The third alphavirus conserved region is a stretch of 20 nucleotides immediately upstream from the subgenomic RNA start site. This region has been shown to contain the minimal region necessary for subgenomic RNA synthesis and has been termed the subgenomic promoter (Levis *et al.*, 1990). The stretch of nucleotides in the rubella virus genome sharing homology with the alphavirus subgenomic promoter is located 23 nucleotides upstream from the subgenomic start site (Figs. 6B and 7). Thus, if this region of the rubella virus genome serves as a promoter for subgenomic RNA synthesis, either the promoter is larger than in the alphavirus genome or a spacer between the promoter and the start site has evolved. When the rubella virus sequence sharing homology with the alphavirus subgenomic promoter was used to replace the Sindbis virus subgenomic promoter in a Sindbis virus expression vector, no subgenomic RNA synthesis was observed, whereas when the same procedure was done using the subgenomic promoters from several alphaviruses, subgenomic RNA synthesis was detected (Hertz and Huang, 1992).

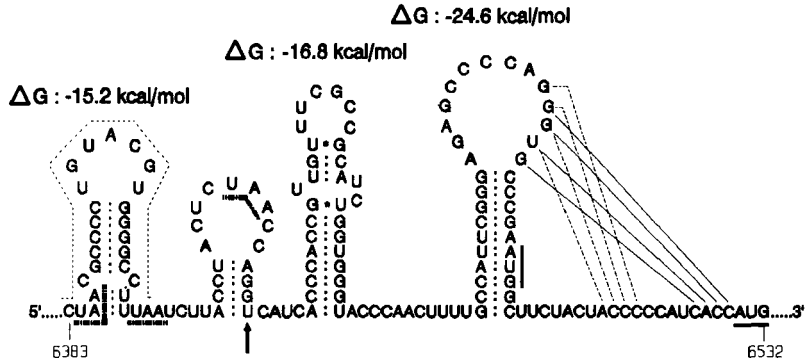

FIG. 7. Potential secondary structures at the rubella virus subgenomic RNA start site. Nucleotides are numbered from the 5′ end of the genome. The nucleotide at which the subgenomic RNA begins is marked with an arrow. The region sharing homology with the alphavirus subgenomic promoter is overlined with a dashed line. The in-frame UAA codons that terminate the nonstructural ORF are underlined with a dashed line and the in-frame AUGs at the 5′ end of the structural protein ORF (SP-ORF) are denoted with solid lines. Potential pairing sites between nucleotides in the loop of the downstream stem-and-loop and nucleotides in the single-stranded region between the SP-ORF AUGs to form a pseudoknot structure are shown. The stabilities (ΔG) of the stem-and-loops were calculated using the method of Tinoco *et al.* (1973).

The fourth alphavirus conserved region is the 20 nucleotides immediately preceeding the 3′ terminal poly(A) tract. These nucleotides share no homology with nucleotides at the 3′ end of the rubella virus genome. Independent of alphavirus homology, a stable stem-and-loop structure 58 nucleotides from the 3′ terminal poly(A) tract of the rubella virus genome has been pointed out by several groups (Frey *et al.*, 1986; Vidgren *et al.*, 1987) (Fig. 5D). Interestingly, both the positive-polarity sequence of this structure and its negative-polarity complement resemble a eukaryotic TATA promoter sequence. When placed in a CAT expression plasmid downstream from a simian virus (SV40) enhancer, a DNA equivalent of the stem-and-loop sequence exhibits promoter activity in the negative-polarity, but not in the positive-polarity, orientation in transfected COS cells (Cao *et al.*, 1992). Whether this finding has any relevance toward the rubella virus replication strategy is unknown, but the implications are intriguing.

3. Involvement of Cell Proteins in Virus RNA Synthesis

Cell proteins are essential components of the RNA-dependent RNA polymerases of bacteriophage QB (Blumenthal and Carmichael, 1979) and cucumber mosaic virus (Hayes and Buck, 1990) and it is assumed that this will be the case for the RNA polymerases of other RNA viruses. One line of evidence for involvement of cell factors in alphavirus RNA replication is that treatment of cells before infection with actinomycin D or α-amanitin (drugs that block transcription of cell mRNAs) results in inhibition of virus replication due to a specific block in virus RNA synthesis, implying that a labile cellular component is necessary for virus RNA synthesis (Baric *et al.*, 1983). With rubella virus, a number of investigators reported that treatment of cells with actinomycin D either before infection or during the eclipse period of the infection, but not during later periods, decreased the final yield of virus (Woods and Robbins, 1968; Wong *et al.*, 1969; Hovi and Vaheri, 1970b; Sedwick and Sokol, 1970; Payment *et al.*, 1975b). Nakhasi *et al.* (1988) showed that continuous treatment of Vero cells with either actinomycin D or α-amanitin beginning at the time of infection or at any time during the first 8 hr of infection (but not after 8 hr of infection) resulted in a reduction in the amount of virus-specific RNA accumulated through 48 hr postinfection and an absence of detectable virus-specific protein synthesis at 48 hr postinfection. Similar treatment had no effect on virus-specific protein synthesis in vesicular stomatitis virus-infected Vero cells, and therefore the results were due to specific interference with rubella virus replication and not to generalized cell deterioration caused by long-term treatment with these drugs. These results are similar to those obtained with alphaviruses; however,

virus-specific macromolecular synthesis in actinomycin D- or α-amanitin-treated cells has not been analyzed at early times during the replication cycle to determine which specific step is inhibited.

It has been shown that host proteins potentially involved in viral RNA transcription can be identified in binding assays in which specific viral RNA sequences are incubated with cellular lysates (Andino *et al.*, 1990, 1993). Using UV light to cross-link RNA–protein complexes, specific binding was demonstrated between proteins in lysates of both uninfected and rubella virus-infected Vero cells and RNAs consisting of the sequences from the distinctive regions of the rubella virus genome described above. The 3′-terminal stem-and-loop structure bound cellular proteins with molecular weights of 61,000, 63,000, and 68,000 (Nakhasi *et al.*, 1990). Binding was greater in extracts from infected cells than in extracts of uninfected cells and binding was completely abrogated when extracts were treated with alkaline phosphatase, indicating that protein phosphorylation was important in binding activity.

In a subsequent study, the negative-polarity complements of both the 5′ terminal stem-and-loop structure and the 46-nucleotide conserved sequence were found to bind to cellular proteins with molecular weights of 56,000, 79,000, and 97,000 (Nakhasi *et al.*, 1991). The amount of the 97-kDa protein increased substantially in lysates from infected cells in comparison to lysates from uninfected cells and, in some of the experiments shown, it was not clear that the 97-kDa protein was present in uninfected cell lysates (the size of this protein is tantalizingly close to the size of one of the cleavage products detected by expression of the rubella virus 5′ proximal ORF). The 56-kDa protein that bound to these RNAs was found to be identical to the 61-kDa protein identified in the 3′ stem-and-loop binding study. Significantly, it was found that the negative-polarity equivalents of the 5′ stem-and-loop structures on the genomes of two alphaviruses, Sindbis virus and eastern equine encephalitis virus, bound the same proteins and the structures from these viruses and rubella virus could compete with each other for binding. This finding indicates equivalence of function of this secondary structure in the genomic RNAs of these distantly related viruses. However, binding of these proteins to the negative-polarity equivalent of the 46-nucleotide region of the rubella virus genome was not competed for by the negative-polarity equivalent of the homologous 51-nucleotide region of the Sindbis virus genome. Finally, it was found that the 5′ stem-and-loop structure on the rubella virus genome RNA bound proteins with molecular weights of 52,000 and 59,000, which were present in extracts of both infected and uninfected cells (Nakhasi *et al.*, 1994; Pogue *et al.*, 1993). Analysis of the effect of site-directed mutagenesis of the stem-and-loop structure on binding to

these proteins indicated that the unpaired nucleotide bulge in the main stem of the structure was a primary determinant in binding.

Substantial progress has been made in determining the identity of the cellular proteins that bind to specific regions of the rubella virus RNA (Nakhasi *et al.*, 1994; Pogue *et al.*, 1993). The amino acid sequence of purified 61/56-kDa protein from monkey kidney cells that bound to both the 3′ terminal stem-and-loop structure and the negative-polarity equivalent of the 5′ terminal stem-and-loop structure was identical to that of human calreticulin, a protein associated with the endoplasmic reticulum that is also found in other cellular compartments and the cytoplasm. The native function of the protein has not been determined. Significantly, simian calreticulin is an autokinase and the phosphorylated form exhibits increased RNA-binding activity, which correlates with the results of alkaline phosphatase treatment of cellular extracts, which abrogated binding of the cellular proteins to the rubella virus stem-and-loop structure from the 3′ end of the genome. Autophosphorylation of calreticulin is stimulated by conditions of cell stress, such as serum deprivation, and the findings of increased binding activity in rubella virus-infected cells indicates that autokinase activity is also stimulated by rubella virus infection.

Calreticulin is also associated with cytoplasmic ribonucleoprotein complexes that contain a group of small RNA species known as hYRNAs and components of the Ro/SS-A autoantigen complex (Sontheimer and Capra, 1993). These autoantigens are the target of autoantibodies in patients with autoimmune diseases such as systemic lupus erythematosus and Sjögren's syndrome. The tentative binding site of the Ro/SS-A autoantigens on the hYRNAs is a bulged stem similar to the major stem of the stem-and-loop structure at the 5′ end of the rubella virus genomic RNA. It was found that hY3RNA competes with the rubella virus stem-and-loop structure for binding of the 59-kDa, but not the 52-kDa, protein (Pogue *et al.*, 1993). The rubella virus 5′ stem-and-loop structure–52/59-kDa protein complex was specifically immunoprecipitated by autoimmune patient serum that has specificity for Ro/SS-A autoantigen.

The precise function of binding of cellular proteins to specific regions of the rubella virus RNA is not known. However, it is thought that binding, particularly to structures or sequences near the 3′ end of an RNA such as the stem-and-loop structure near the 3′ terminus of the rubella virus genomic RNA or the negative-polarity equivalent of the 5′ terminal stem-and-loop structure, could serve as an initial recognition event in the formation of the virus-specific replication complex that subsequently functions to synthesize the complementary strand. The function in the virus RNA replication process of binding of

cell proteins to the 5′ end of an RNA species is not so clear. For this reason, it was proposed that binding of the Ro/SS-A autoantigen to the 5′ end of the rubella virus genomic RNA could facilitate initiation of translation by destabilizing and unwinding the stem-and-loop structure (Nakhasi *et al.*, 1994; Pogue *et al.*, 1993). Alternatively, the negative-polarity genome-length RNA species is found exclusively in double-stranded complexes in infected cells, and initiation of positive-polarity genome RNA synthesis could involve recognition of both the 3′ end of the negative-polarity RNA and the 5′ end of the genome RNA to which it is hybridized. Interestingly, the propensity of Ro/SS-A antigen and calreticulin to coexist in complexes means that these proteins could function to cyclize the rubella virus genomic RNA (Nakhasi *et al.*, 1994).

4. Generation of Defective-Interfering RNAs

A feature of RNA replication that has been recognized with most RNA viruses is the generation of deletion mutants known as defective-interfering (DI) RNAs (reviewed in Holland, 1990). Optimal conditions for generation of DI RNAs are serial undiluted passage of virus stocks and persistent infection in cell culture. Defective-interfering RNAs are dependent on the presence of the genome RNA for replication but, at the same time, often interfere with genome RNA replication. By so doing, DI RNAs tend to attenuate the cytopathogenicity of virus infection. The analysis of DI RNAs is of interest because these RNAs retain the genomic sequences critical for replication and encapsidation. Generation of DI RNAs by rubella virus was reported by four groups. In two of the reports, DI RNAs were present in the infecting stocks (Bohn and Van Alstyne, 1981; Terry *et al.*, 1985) and it was proposed that their presence could account in part for the relatively noncytopathic replication exhibited by rubella virus (Bohn and Van Alstyne, 1981). However, it was subsequently shown that the replication of plaque-purified rubella virus free of detectable DI RNAs was also relatively noncytopathic and thus similar to that reported for rubella virus that had not been plaque purified (Hemphill *et al.*, 1988). On undiluted serial passage of plaque-purified rubella virus, DI RNAs were detectable after four passages (Frey and Hemphill, 1988), a fairly typical observation with respect to generation of DI RNAs by other viruses. Concomitant with the appearance of DI RNAs in the serially passaged stocks, the amount of genomic RNA present in cells infected with these stocks decreased dramatically and the titer of the stocks declined, showing that interference was occurring. Defective-interfering RNAs were also detected by two groups in long-term persistently infected cell cultures (Norval, 1979; Abernathy *et al.*, 1990). When the pres-

ence of DI RNAs at various times postinitiation of persistence was analyzed, it was found that DI RNAs were not detectable until 19 days, indicating that the presence of DI RNAs is not necessary for rubella virus to initiate a persistent infection (Frey and Hemphill, 1988). This is in contrast to a number of more lytic viruses, with which the presence of DI particles in the infecting stocks is essential for initiation of persistence to occur. Following the appearance of DI RNAs in the persistently infected cells, the amount of genomic RNA present gradually decreased to undetectable levels. The size of the rubella virus DI RNAs detected during both serial undiluted passage and persistent infection ranged from 1000 nucleotides to 7000–8000 nucleotides. It has been shown that large DI RNAs 6700 to 7200 nucleotides in length, which are generated during undiluted serial passage of rubella virus in Vero cells, contain the intact nonstructural protein open reading frame and a large deletion within the structural protein open reading frame (C. A. Derdeyn, unpublished observations). Short DI RNAs 600 to 800 nucleotides in length, which are observed in the same cells, are actually subgenomic RNAs synthesized from the large DI RNA template. The large DI RNA species putatively have the capacity for self-replication, but are of course dependent on standard genomic RNA for structural protein synthesis.

E. Translation and Processing of Structural Proteins

At the 5′ end of the structural protein ORF are two in-frame AUGs separated by seven codons that are the initial AUGs on the subgenomic RNA, the first being 78 nucleotides from the 5′ end (Fig. 7). The 5′ nontranslated region of the subgenomic RNA contains two stem-and-loop structures that are predicted to be relatively stable (Fig. 7). The amino terminus of the C protein is blocked, preventing determination of the amino-terminal sequence of the capsid protein (Kalkkinen *et al.*, 1984), and thus it is not known if one or both of these AUGs is used. In terms of favorable neighboring nucleotide context for initiation of translation as compiled by Kozak (1987) (CCACCAUGG being most favorable, with the boldface purines being of most importance), the downstream AUG is in an excellent context (UCACCAUGG) whereas the upstream AUG is in a less favorable context (CCCGAAUGG). Translation *in vitro* of RNAs transcribed from constructs containing both AUGs or mutagenized to remove one of the two AUGs demonstrated that initiation of translation can occur at either AUG; however, the AUG at which initiation occurred preferentially was not convincingly demonstrated (Clarke *et al.*, 1988; Marr *et al.*, 1991). The C proteins translated *in vitro* from RNAs transcribed from these constructs

all migrate as a doublet (or occasionally as a triplet) in polyacrylamide gel electrophoresis (PAGE) (Suomalainen *et al.*, 1990; Marr *et al.*, 1991) and thus, the C doublet synthesized in virus-infected cells and present in virions does not appear to be due to initiation at both of the AUGs.

In the absence of microsomes *in vitro*, translation of the structural protein ORF gives rise to a 110-kDa precursor (Oker-Blom *et al.*, 1984; Clarke *et al.*, 1987). Thus, rubella virus lacks the capsid protein autoprotease found in the alphaviruses. In the deduced amino acid sequence of the structural protein ORF, the amino termini of E2 and E1 are preceded by stretches of 23 and 20 hydrophobic amino acids, respectively, which have the characteristics of consensus signal sequences that mediate the translocation of membrane-bound and secreted proteins into the lumen of the endoplasmic reticulum (ER) (reviewed in Wiley, 1986). Thus, it is predicted that these sequences mediate translocation of E2 and E1 into the lumen of the ER and that the C–E2 and E2–E1 cleavages are catalyzed by signalase, a lumenal enzyme that cleaves signal sequences from proteins following translocation. As predicted, when the structural protein ORF is translated *in vitro* in the presence of microsomes, accurate processing of the structural proteins occurs (Marr *et al.*, 1991). Deletion mutagenesis studies have shown that the hydrophobic sequences that precede E2 and E1 serve as the signal sequences for E2 and E1 in terms of being required for translocation into the lumen of the ER and glycosylation (Hobman *et al.*, 1988; Hobman and Gillam, 1989; Oker-Blom *et al.*, 1990). An E2 construct containing the precise 23-amino acid signals sequence at its NH_2 terminus is accurately processed both *in vitro* in the presence of microsomes and *in vivo*, showing that this sequence can function externally as well as in its native internal context (Marr *et al.*, 1991; Sanchez and Frey, 1991). In terms of evidence that definitively shows that signalase catalyzes the structural protein cleavages, site-directed mutagenesis was performed on a structural protein ORF construct to change the COOH-terminal alanine residue of the E2 signal to a proline, a residue never found within signalase cleavage sites (McDonald *et al.*, 1991; von Heijne, 1984). Translation *in vitro* in the presence of microsomes and expression *in vivo* of the mutagenized construct resulted in production of an uncleaved C–E2 product (which, interestingly, was properly glycosylated). However, cleavage still occurred in that C and E2 were readily detectable. The production of C and E2 was presumably due to cleavage at an alternate site within the E2 signal sequence (von Heijne, 1984); however, this was not confirmed. In the same study, it was shown that translation of at least 26 amino acids of E2 was necessary for the C–E2 cleavage to occur.

It has been shown that the E2 and E1 signal sequences remain attached to mature C and E2, respectively (Suomalainen *et al.*, 1990; Marr *et al.*, 1991; Baron *et al.*, 1992). The most definitive evidence in this regard is that antibodies raised to a peptide consisting of the COOH-terminal 10 amino acids of the E2 signal sequence and antibodies raised to a peptide consisting of the COOH-terminal 12 amino acids of the E1 signal sequence immunoprecipitate C and E2, respectively, from both virions and infected cells (Suomalainen *et al.*, 1990; Baron *et al.*, 1992). Thus, the only proteolytic cleavages that have been shown to occur in the processing of the rubella virus structural protein precursor are the two mediated by signalase. However, the amino terminus of C has not been determined and carboxy-terminal sequencing has not been successfully done with C, E2, or E1 (Kalkkinen *et al.*, 1984) and thus it is possible that a limited amount of proteolytic tailoring at these locations could occur.

The maintenance of the E2 signal sequence as part of the capsid protein is, as far as has been determined, unique to rubella virus. In the structural protein precursor of three other virus genera (*Alphavirus*, *Flavivirus*, and *Pestivirus*), the COOH-terminus of the capsid protein is also immediately adjacent to the signal sequence for a membrane protein. In both the alphaviruses and the flaviviruses, the signal sequence is removed from the mature capsid protein: by the capsid autoprotease in the case of the alphaviruses, and putatively by the viral-encoded nonstructural protein protease (an NS2b–NS3 complex) in the case of the flaviviruses (Chambers *et al.*, 1990, 1991). It has not yet been determined if the signal sequence is removed from the pestivirus capsid protein (Rümenapf *et al.*, 1993). Retention of the signal sequence appears to have a profound impact on the capsid morphogenesis of these viruses. Autoproteolysis of the alphavirus capsid protein is so rapid that the signal for the adjacent membrane protein (PE2) is not even transiently associated with the C protein and alphavirus capsids form in the cytosol. In contrast, following cleavage of the flavivirus C protein and the adjacent membrane protein (pre-M) by signalase, the pre-M signal sequence remains attached to C and is thought to be removed only at the time of capsid formation (reviewed in Nowak *et al.*, 1989; Chambers *et al.*, 1990). Capsid morphogenesis of both the flaviviruses and rubella virus occurs in association with membranes and it is hypothesized that the presence of a signal sequence at the COOH-terminus of the capsid proteins of these viruses is important in mediating the interaction between the capsid protein and membranes. As is discussed below (Section VI,F,1), the E2 signal sequence mediates an association between the rubella virus C protein and membranes.

F. Posttranslational Maturation and Transport

1. C Protein

The rubella virus C protein forms a noncovalently bonded dimer soon after translation in infected cells (Baron and Forsell, 1991). The use of a reducing agent such as iodoacetamide in the lysis buffer is necessary to recognize the nonconvalent nature of the initial dimer, as disulfide bonds are formed by the oxidizing conditions of the lysis buffer. Within 2 hr after translation, covalently linked C dimers appear in infected cells, even in the presence of reducing agent in the lysis buffers. When the C protein was expressed using vaccinia virus, noncovalently bonded dimers also formed soon after translation; however, covalently linked C dimers were not detected, indicating that disulfide bond formation in infected cells occurs within the virion following budding.

An association of C with membranes mediated by the presence of the E2 signal sequence has been demonstrated both *in vitro* and *in vivo*. Following translation *in vitro* in the presence of microsomes or expression *in vivo* of a C construct produced by introduction of termination codons at the end of the E2 signal sequence (C + S) or a C–E2 construct, C is associated with the microsome or membrane fraction (Suomalainen *et al.*, 1990; Baron *et al.*, 1992). Under similar conditions, C produced from a construct in which termination codons were placed before the E2 signal sequence (C − S) segregates into the soluble fraction and thus the association of C with membranes is mediated by the E2 signal sequence. Association with the microsome or membrane fraction was tighter with the C processed from the C–E2 construct than from the C + S construct; in neither case was membrane association of C as tight as the membrane association of an integral membrane protein such as E2. As referred to above (Section VI,E), rubella virus capsid formation occurs in association with membranes and, presumably, the E2–signal sequence-mediated association of C with membranes is important in this event.

By immunofluorescence, rubella virus C protein is present in both the ER and the Golgi in infected cells (Hobman *et al.*, 1990; T. K. Frey and E. S. Abernathy, unpublished observations). In transfected cells, the C − S product is distributed in a punctate pattern throughout the cytoplasm, the C + S product as well as C processed from a construct containing C plus one-third of E2 is localized in the ER, whereas C processed from a C–E2 or structural protein ORF construct is localized in the Golgi (Hobman *et al.*, 1990; Baron *et al.*, 1992). Thus, the E2–signal-mediated association of C to membranes is important in localization of C to membranous structures in cells and transport of C from

the ER to the Golgi region occurs in conjunction with the viral glycoproteins, minimally E2. As is discussed below (Section VI,F,2,b), the rubella virus glycoproteins are targeted to the Golgi apparatus and presumably the association between C and the glycoproteins allows C to be transported to the Golgi region as a passenger on the outside of transitory vesicles containing the glycoproteins. In transfected cells in which C and E2 are coexpressed from independent constructs rather than coordinately expressed, C remains in the ER, indicating that transport of C to the Golgi requires a tight association of C and E2 that is formed only during coordinate expression.

2. E2 and E1

a. Maturation. The most striking feature of the posttranslational maturation of the rubella virus glycoproteins is the conversion of $E2_i$ to $E2_v$. The $E2_v$ forms are not immunoprecipitable by polyclonal anti-rubella virus serum from a number of species (including sera from some humans) and most of the anti-E2 monoclonal antibodies and thus the presence of intracellular $E2_v$ forms was not detected in a number of studies that used these antisera (Oker-Blom *et al.*, 1983; Oker-Blom, 1984; Bowden and Westaway, 1984; Sanchez and Frey, 1991). Detection of the intracellular $E2_v$ forms by immunoprecipitation was first accomplished by use of polyclonal serum raised against a peptide containing the COOH-terminal 12 amino acids of the E1 signal sequence (Baron and Forsell, 1991). As shown in Fig. 8, in pulse-chase radiolabeling experiments using infected Vero cells the intracellular $E2_v$ forms become apparent after chases of from 2 to 4 hr and, concomitantly, the $E2_i$ form disappears [when immunoprecipitation is done using an antibody that fails to react with $E2_v$, the $E2_i$ species simply disappears following chases of from 2 to 4 hr in this type of experiment (Fig. 8, α-E2-1 lanes)]. The intracellular $E2_v$ forms are also detectable on Western blots of infected cell lysates probed with anti-E2 monoclonal antibodies (Sanchez and Frey, 1991). Interestingly, although the migration of intracellular $E2_v$ is smeary in nature, three to four discrete bands are distinguishable within the smear, unlike $E2_v$ from virions (Fig. 8) (Sanchez and Frey, 1991). The presence of these bands indicates that the posttranslational processing of E2 occurs in discrete steps.

Presumably, a major event in the conversion of $E2_i$ to $E2_v$ is the addition of *O*-glycans. The molecular weight of both endoglycosidase F (endo F)-digested $E2_i$ and E2 synthesized in tunicamycin-treated cells is roughly 30,000 (Oker-Blom *et al.*, 1983; Sanchez and Frey, 1991), a figure close to the molecular weight for the amino acid backbone of E2 of 30,100 calculated from the deduced amino acid sequence (Frey and

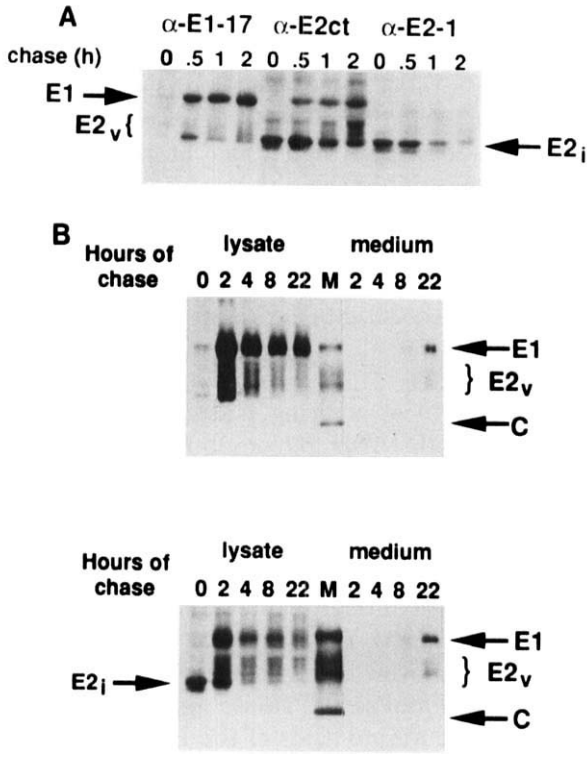

FIG. 8. Pulse–chase radiolabeling of the rubella virus structural proteins. Rubella virus-infected Vero cells were radiolabeled with a 15-min pulse of [^{35}S]methionine, chased with nonradioactive medium for the indicated times, and lysed with buffer containing nonionic detergent. Immunoprecipitations were done with anti-E1 or -E2 monoclonal antibodies (α-E1–17 or α-E2-1) or a polyclonal antiserum to a peptide containing the E1 signal sequence (α-E2ct). Coprecipitation of E1 and E2 by serum specific for either protein is indicative of heterodimer formation (note that α-E2-1 does not coprecipitate E1 whereas α-E2ct does; α-E1-17 coprecipitates E2). In (A), coprecipitation is apparent in the 0.5-hr chase and the presence of $E2_v$ forms is apparent in the 2-hr-chase (α-E2ct precipitates the $E2_v$ forms whereas α-E2-1 does not). In the top half of (B), immunoprecipitation was done with α-E1-17 and in the bottom half, immunoprecipitation was done with α-E2ct. Note the presence of the $E2_v$ forms at the 2-hr chase and the subsequent decrease in intensity during chases of 2, 8, and 24 hr. In the top half of (B), immunoprecipitation is with α-E1–17 and thus the amount of E1 remains constant while the amount of the $E2_v$ forms declines dramatically. In the bottom half of (B), immunoprecipitation is with α-E2ct and thus both the amount of immunoprecipitable E1 and E2 decline during the chase periods. The amount of E1 and E2 in the medium did not become detectable until a 24-hr chase and thus the decline in the amount of E2 was not due to export of virions into the medium. [The electropherograms were provided by M. Baron and are from Baron and Forsell (1991).]

Marr, 1988). The molecular weights of the endo F digestion products of $E2_v$ are 33,000 and 37,500. Thus, $E2_i$ most likely contains no *O*-glycans. Because the intracellular $E2_v$ species comigrate with $E2_v$ from virions, they presumably contain *O*-glycans, although studies to detect the presence of *O*-glycans on the intracellular $E2_v$ species have not actually been done. The site of O-glycosylation of E2 is not known. In studies on the O-glycosylation of other glycoproteins, the addition of the initial *N*-acetylgalatosamine has been reported to occur in the ER, transitional elements between the ER and Golgi, and various compartments of the Golgi (reviewed in Tooze *et al.*, 1988; Locker *et al.*, 1992). There is general agreement that subsequent addition of the galactose and sialic acid residues occurs in the *trans*-Golgi.

As detected by coimmunoprecipitation with anti-E1- or anti-E2-specific serum, heterooligomeric complexes containing E1 and E2 can be detected within 30 min of synthesis (Baron and Forsell, 1991) (Fig. 8). The initially detectable form of E2 coimmunoprecipitated by anti-E1 serum is the $E2_i$ species; however, after chases of 2 hr or longer, the $E2_v$ forms are also coprecipitated. Sucrose gradient centrifugation of infected cell lysates reveals the presence of E1–E2 complexes cosedimenting with Semliki Forest virus E1–E2 heterodimers, suggesting that the heterooligomeric complex formed by the rubella virus glycoproteins is a heterodimer. As with similar experiments on disrupted virions, roughly half of each glycoprotein segregates into the oligomer and monomer fractions in the gradients, presumably indicative of the inherent instability of the heterodimer. Studies on the glycoproteins of a number of other enveloped viruses have shown that oligomerization occurs in the ER and is a prerequisite for export from the ER (reviewed in Hurtley and Helenius, 1989). In the case of the alphaviruses, formation of the pE2–E1 heterodimer occurs in the ER soon after translation (Wahlberg *et al.*, 1989) and it is assumed that the hexamer unit consisting of three heterodimers is formed before transport from the ER occurs (Doms *et al.*, 1993). The finding that the rubella virus glycoproteins heterodimerize is consistent with these observations and it is likely that most of the posttranslational maturation and modification undergone by the rubella virus glycoproteins occur in the context of the heterodimer. Because it has not been established that the rubella virus glycoproteins form a hexameric complex, it is not known if such a complex is formed in the ER.

The rubella virus glycoproteins have been expressed using a number of vectors: vaccinia virus (Baron and Forsell, 1991; Sanchez and Frey, 1991; Baron *et al.*, 1992), baculovirus (Oker-Blom *et al.*, 1989), and plasmids in which expression is driven by the SV40 early or late promoter or the human cytomegalovirus immediate early promoter used

for both transient and stable transfection (Hobman and Gillam, 1989; Hobman *et al.*, 1988, 1990, 1992, 1993; Baron *et al.*, 1992). In the case of both vaccinia virus and plasmid-driven expression, as far as was analyzed in each study in which the entire structural protein ORF was expressed or the glycoproteins were expressed coordinately (i.e., without the C sequences), the processing of the glycoproteins was similar to that in infected cells. However, when the rubella virus glycoproteins were expressed in baculovirus in *Spodoptera frugiperda* cells both E2 and E1 were smaller than their authentic counterparts and a 90- to 95-kDa uncleaved but glycosylated E2–E1 product was a major species produced.

When E2 and E1 are expressed coordinately using transient transfection, conversion of $E2_1$ to $E2_v$ was detectable but incomplete (Hobman *et al.*, 1990), whereas using vaccinia virus expression conversion was noticeably slower than in rubella virus-infected cells (Baron and Forsell, 1991; Sanchez and Frey, 1991). No detectable conversion occurred when these proteins were expressed in baculovirus (Oker-Blom *et al.*, 1989). Interestingly, in stably transfected cells quantitative conversion occurred between 60 and 180 min postsynthesis (Hobman *et al.*, 1993). Expressed E2 and E1 form heterooligomers that are detectable within 5 min of synthesis (Hobman *et al.*, 1993; Baron and Forsell, 1991). Thus far, the only detectable form of expressed E2 reported in heterooligomers is $E2_i$, although this is most likely a function of the chase periods analyzed in each study. When E2 is expressed alone, the $E2_i$ form is stable; however, limited conversion to $E2_v$ does occur (Hobman and Gillam, 1989; Hobman *et al.*, 1990; Sanchez and Frey, 1991). In studies in which expression of E2–E1 and E2 alone is analyzed using the same vector system, conversion of $E2_i$ to $E2_v$ is more efficient when E2 and E1 are coexpressed (Hobman *et al.*, 1990; Sanchez and Frey, 1991). These results indicate that E2 itself contains the signals necessary for processing to mature form and does not require conformation acquired in the heterooligomeric form. Efficient conversion in the heterooligomeric form is thus likely to be a function of efficient transport to the site of conversion.

As discussed above, the E1 in virions has a complicated intramolecular tertiary structure that involves disulfide bonding. It was shown using pulse–chase radiolabeling of stably transfected cells that coexpress E1 and E2 and gel electrophoresis of radiolabeled proteins under nonreducing conditions that E1 requires 30 to 60 min to acquire its mature tertiary structure (Hobman *et al.*, 1993). This finding was consistent with earlier studies in which it was found that immediately following a pulse-radiolabeling period in rubella virus-infected cells, E1 was not immunoprecipitable by anti-E1 monoclonal antibodies that

recognized nonlinear epitopes but, following a 30-min chase, became immunoprecipitable (Baron and Forsell, 1991). When E1 is expressed alone it is recognized by most anti-E1 monoclonal antibodies by either immunofluorescence or immunoprecipitation and thus the presence of E2 is not necessary for acquisition of tertiary structure recognizable by these monoclonal antibodies (Hobman *et al.*, 1990, 1993; Baron and Forsell, 1991; Baron *et al.*, 1992). Interestingly, anti-E1 monoclonal antibodies that recognize nonlinear epitopes do not stain the ER of rubella virus-infected cells when fixation is with formaldehyde but do stain the ER when fixation is with ethanol (T. K. Frey and E. S. Abernathy, unpublished observations). This observation implies that E1 is complexed with another protein in the ER and that this complex is not disrupted by formaldehyde. Although this complex may be between E1 and E2, a class of proteins in the ER (known as molecular chaperones) forms complexes with glycoproteins that may, among other things, assist in folding (reviewed in Doms *et al.*, 1993). In this regard it has been found that gp78–BiP, a well-characterized molecular chaperone, is coprecipitated with E1 using both anti-rubella virus serum and anti-E1 monoclonal antibodies (Hobman *et al.*, 1990, 1991).

A final point on the posttranslational processing of the rubella virus glycoproteins is that, on the basis of the observation that in Western blots of infected cells probed with a cocktail of anti-E2 monoclonal antibodies the labile $E2_i$ species was the predominant species detected, we proposed that E2 is selectively turned over in infected cells (Sanchez and Frey, 1991). In some pulse–chase radiolabeling experiments in infected cells, the intensity of the E2 bands appears to decline significantly during the course of the chase in comparison to E1 (see Fig. 8). However, considering the differential affinity for antibodies for the $E2_i$ and $E2_v$ forms, this hypothesis needs more extensive experimental analysis to be confirmed.

b. Targeting. In rubella virus-infected cells, the glycoproteins are detectable by immunofluorescence in the ER, Golgi, intracellular vacuoles, and on the cell surface (Waxham and Wolinsky, 1983; Bowden and Westaway, 1989; Hobman *et al.*, 1990), all of which are reported to be sites of virus budding. That surface fluorescence is due to incorporation of the virus glycoproteins into the plasma membrane and not due to aggregation of released virions on the cell surface was shown by binding of both erythrocytes and ferritin-labeled antibodies to regions of plasma membrane devoid of virions (Oshiro *et al.*, 1969; Matsumoto and Higashi, 1974). Little evidence has been obtained on the targeting of the glycoproteins or the kinetics of transport between the sites of accumulation in infected cells. In a time course study, by immunofluorescence using anti-E1 and -E2 monoclonal antibodies, E1 and

E2 were initially detectable in the ER and Golgi and appeared on the cell surface only at later times postinfection (T. K. Frey and E. S. Abernathy, unpublished observations). Both E1 and E2 were retained in the Golgi during prolonged treatment with cycloheximide, a drug that inhibits translation, but not posttranslational processing or transport. Taken together, these observations indicate that E1 and E2 are only slowly transported from the Golgi to the cell surface and that such transport may not be quantitative.

These indications have been confirmed by the findings that in cells transiently or stably transfected with plasmid constructs or infected with vaccinia virus constructs containing the structural protein ORF or E2–E1 the primary site of glycoprotein accumulation is the Golgi (Hobman *et al.*, 1990, 1993; Baron *et al.*, 1992). Immunoelectron microscopy of stably transformed cells showed that the glycoproteins were distributed across all cisternae of the Golgi stack as well as in Golgi-associated vesicles (Hobman *et al.*, 1993). In all of the expression studies, the presence of both glycoproteins on the cell surface could be detected; however, only a small fraction of the expressed glycoproteins was transported to the cell surface. These studies encompassed five different cell lines from three different species and no difference in amount of surface expression was apparent among these cell lines. When E2 was expressed alone, it was localized in both the ER and the Golgi, although a fraction was transported to the cell surface (Hobman and Gillam, 1989; Hobman *et al.*, 1990, 1993; Baron *et al.*, 1992). When E1 was expressed alone, it was found to localize in a novel ER–Golgi intermediate compartment (Baron *et al.*, 1992; Hobman *et al.*, 1993). The significance of this compartment in terms of rubella virus replication is unclear. Thus, E2 seems to contain the signals for transport out of the ER to the Golgi and eventually the cell surface. However, out of the context of the heterodimer, such transport is inefficient. Although the preferred site of localization of the rubella virus glycoproteins is the Golgi, there is no absolute retention signal as evidenced by the fact that migration to the cell surface does eventually occur.

The only data on the rate of transport of the rubella virus glycoproteins between compartments in infected cells are those that show that the $E2_v$ forms are initially detectable within a 2-hr chase and that conversion of $E2_i$ to $E2_v$ is essentially complete within 4 hr (Baron and Forsell, 1991) (Fig. 8). Because the largest of these forms in virions contains complex carbohydrates, this implies that transport to the Golgi occurs within this time frame. In transiently transfected cells, endo H-resistant forms of E2 and E1 could be detected within 30 and 60–120 min of synthesis, respectively; although a substantial amount of E2 was converted to an endo H-resistant form, only a fraction of E1

was converted to an endo H-resistant form (Hobman *et al.*, 1990). In vaccinia virus-infected cells, binding of E2 and E1 to a lectin specific for β-D-galactose was apparent within a 2-hr chase (Baron *et al.*, 1992). In stably transfected CHO cells complete conversion of both glycoproteins to an endo H-resistant form occurred with a half-life of 60–90 min and sialylation of both glycoproteins could be detected by 120 min after synthesis (Hobman *et al.*, 1993). Despite the different protocols employed and the discrepancy in completeness of some of the results, the time of transport of the rubella virus glycoproteins from the ER to the Golgi seems to be in the range of 1 to 2 hr, both in infected cells and when the glycoproteins are expressed. Compared to other viral glycoproteins, this is a relatively slow rate of transport (e.g., the time required for transport of the alphavirus glycoproteins to the Golgi is 25 min; reviewed in Doms *et al.*, 1993).

The factors controlling the rate of glycoprotein transport are not completely understood; however, it has been recognized that essentially complete acquisition of tertiary structure by glycoproteins occurs before exit from the ER and that incompletely folded or misfolded proteins are retained in the ER (reviewed in Hurtley and Helenius, 1989). Because extensive time is required for complete folding of rubella virus E1, it has been proposed that this is the rate-limiting step in the transport of the rubella virus glycoproteins out of the ER (Hobman *et al.*, 1993).

In a different vein, site-directed mutagenesis studies were undertaken to determine if glycosylation played a role in rubella virus glycoprotein transport. When expressed in the absence of E1 in transiently transfected COS cells, a series of mutants of E2, each lacking one or more of the N-linked glycosylation sites, was found not to be processed to the $E2_v$ form (in these experiments, roughly 50% of wild-type $E2_1$ was converted to $E2_v$) and to be relatively unstable in comparison to wild-type E2 (Qiu *et al.*, 1992a). The mutant completely lacking glycosylation sites was the most unstable, being degraded completely within 4 hr. Whereas wild-type E2 was primarily localized in the Golgi in transfected cells as determined by immunofluorescence, the major site of accumulation of all of the glycosylation mutants was in the ER, although some Golgi localization was apparent. Unlike wild-type E2, none of the E2 produced by any of the glycosylation mutants was transported to the cell surface. Thus, mutagenesis of any of the E2 glycosylation sites has an effect on the efficiency of transport and processing. In corollary studies on E1, abrogation of each glycosylation site either singly or in combination had no effect on the distribution within the cell of E1 expressed alone or with E2; however, mutants lacking either of the COOH-terminal two glycosylation sites were not

transported to the cell surface when coexpressed with E2 (Hobman *et al.*, 1991; Qiu *et al.*, 1992b). It is thought that the deleterious effects on glycoprotein transport often elicited by interference with glycosylation is due generally to alteration of tertiary structure rather than to glycans being specific transport signals (reviewed in Doms *et al.*, 1993). In this regard, all of the E1 glycosylation mutations had a dramatic effect on the conformation of E1 as evidenced both by the binding pattern of these proteins to both human sera and monoclonal antibodies in comparison to wild-type E1 and the relative inability of vaccinia recombinants expressing the mutated proteins to stimulate production of both neutralizing and HI antibodies in rabbits.

There is evidence that the rubella virus glycoproteins can be secreted from the infected cell in that a number of studies reported the release of rubella virus-specific "soluble antigens" into the culture medium of infected cells (Furukawa *et al.*, 1967; Schmidt and Lennette, 1969; Le Bouvier, 1969a,b; Vaheri and Vesikari, 1971). These antigens exhibited neither infectivity nor hemagglutinin activity but were able to react with antibodies to fix complement ("complement-fixing antigens") and to aggregate platelets. The size of these particles was 3–7S and the density was 1.08–1.11 g/ml as measured in sucrose gradients (the density of virions in such experiments is 1.19–1.21 g/ml). Immunologically, the soluble antigens reacted similarly to virions in immunodiffusion tests utilizing either convalescent human antiserum or anti-rubella virus rabbit serum, and when soluble antigen preparations were used to immunize rabbits, antibodies that exhibit neutralization, HI, and complement fixation were induced. The release of soluble antigens was initially detectable relatively late in infection (after the initial detectable release of virions) and production was particularly abundant in persistently infected cultures. Soluble antigens with similar properties could be released from infected cells by extraction with nonionic detergent, freeze-thawing, and sonication (Furukawa *et al.*, 1967; Vaheri and Vesikari, 1971; Salmi, 1972a,b). Although the protein content of the soluble antigens was never determined, they apparently contain at least the virus glycoproteins. The low density is indicative of the absence of nucleic acid and the size is similar to that of heterooligomeric complexes released from virions by treatment with nonionic detergents. However, if the soluble antigens consist of glycoprotein heterooligomers, the lack of hemagglutinin activity is curious because nonionic detergent disruption of virions enhances hemagglutinin activity. More work is necessary to determine the precise content and structure of these moieties.

G. Virion Morphogenesis

A large number of electron microscopy studies has examined the budding of rubella virus (reviewed in Murphy, 1980; Horzinek, 1981). Budding can be observed to commence with a thickening of a patch of membrane modified by the insertion of viral glycoprotein as detected by immunoferritin labeling (Oshiro *et al.*, 1969) (Fig. 9). As budding progresses, the membrane within the bud also thickens and the envelope of the complete virion retains the thickened morphology. Capsids are evident in association with partially budded membranes. The property of capsid formation occurring coincident with budding is shared by rubella virus with the type C retroviruses. In contrast, alphavirus capsids coalesce independently of both membranes and the budding process and form crystalline arrays in the cytoplasm of infected cells.

Apparent discrepancies exist in the reported sites of rubella virus budding. Most studies were done with BHK-21 cells and although the majority reported the budding of virions at both intracellular membranes and at the plasma membrane (von Bonsdorff and Vaheri, 1969; Higashi *et al.*, 1969; Holmes *et al.*, 1969; Oshiro *et al.*, 1969; Matsumoto and Higashi, 1974; Bardeletti *et al.*, 1979), budding solely at intracellular membranes (Tuchinda *et al.*, 1969; Edwards *et al.*, 1969) and, in contrast, budding predominantly at the plasma membrane (Murphy *et al.*, 1968) were also reported. Holmes *et al.* (1968) and Hamvas *et al.* (1969) observed virions in intracellular vacuoles, but did not detect budding forms. Of the three studies done in Vero cells, Tuchinda *et al.* (1969) reported budding only at intracellular membranes, Bowden and Westaway (1989) reported budding at both intracellular membranes and the cell surface, whereas Payment *et al.* (1975b) reported budding solely at the cell surface. In RK-13 cells, Matsumoto and Higashi (1974) reported budding at both intracellular membranes and the cell surface, Holmes *et al.* (1968), Hamvas *et al.* (1969), and Kouri *et al.* (1974) reported accumulation of virions in intracellular vacuoles but did not observe budding forms, and Chatterji *et al.* (1969) detected the presence only of extracellular virions and did not observe budding forms. In SIRC (rabbit cornea) cells, McCombs *et al.* (1968) observed budding only at intracellular membranes whereas Hamvas *et al.* (1969) reported accumulation of virions in intracellular vacuoles but did not observe budding forms. It is clear that the differences observed in these studies were not due to cell type and the available data do not warrant the conclusion that differential patterns of budding occur in different cell types.

As discussed above (Section VI,F,2,b), expression studies have dem-

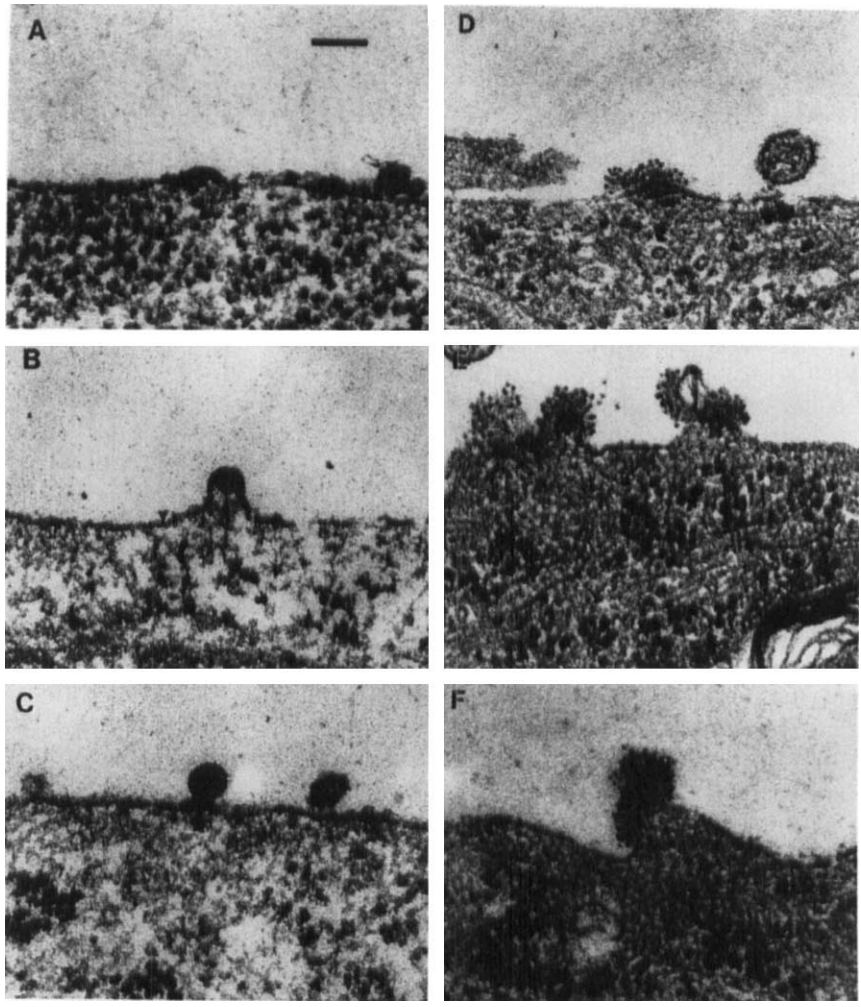

FIG. 9. Budding of rubella virus. Budding forms of rubella virus from the cytoplasmic membrane observed in infected BHK-21 cells. Cells were fixed with 1% glutaraldehyde and postfixed with 1% osmium tetroxide prior to embedding and thin sectioning. The initiation of the budding process is indicated by the thickening of a patch of membrane (A and D). Cores are seen to form in association with partially deformed membranes (B and E). Freshly budded virions are shown in (C) and (F). In (D) through (F), the thin sections were reacted with ferritin-conjugated human anti-rubella virus serum to distinguish regions of the membrane modified by the presence of rubella virus antigens. [The micrographs were provided by L. Oshiro and are from Oshiro *et al.* (1969).]

onstrated that the rubella virus glycoproteins are targeted primarily to intracellular locations, although inefficient transport to the cell surface occurs (Hobman *et al.*, 1990, 1993; Baron *et al.*, 1992). Consistent with these findings, in infected cells the glycoproteins are initially detectable at intracellular locations and appear on the cell surface only at later times in the replication cycle (T. K. Frey and E. S. Abernathy, unpublished observations). It would thus be expected that the primary site of virus budding is at intracellular locations, with budding at the cell surface occurring at later times postinfection when sufficient concentrations of glycoproteins had accumulated in the plasma membrane. Another indication that the primary site of virus budding is at intracellular locations is that the lipid content of the virion envelope is similar to that of intracellular membranes rather than the plasma membrane (intriguingly, the virion membrane was also found to contain cardiolipin, a compound specific for the inner mitochondrial membrane) (Bardeletti and Gautheron, 1976).

Most of the electron microscopic studies were done on cultures infected with an MOI of 1 infectious dose per cell or less and, as expected when a low MOI is used, virions were not observable in these studies until 48 hr postinfection. Indeed, in several studies, the cells were not processed for electron microscopy until 7 to 10 days postinfection. Thus, in these studies the degree of asynchrony of infection is greater than in cells infected with a high MOI and at each time point the specimen analyzed would contain cells at all stages of the virus replication cycle. In one study (von Bonsdorff and Vaheri, 1969), however, a high MOI was used and virus could be observed in up to 50% of the cells as early as 15 hr postinfection. Thus, in this study conditions of synchronous infection were approached as closely as is possible with rubella virus. In this study, the accumulation of virus and virus budding was initially observed in the Golgi. At later times postinfection, budding at the plasma membrane was observed but with much less frequency than at intracellular membranes. These observations are entirely consistent with the findings on the intracellular sites of accumulation of the virus glycoproteins and it is thus likely that this is the normal course of progression of rubella virion morphogenesis. The studies that describe budding solely at intracellular membranes or at both intracellular membranes and the plasma membrane are consistent with this progression; however, the fact that a progression exists was missed due to the asynchronous infection conditions or that the infected cells were analyzed very late in infection when budding at both sites occurs. Interestingly, despite using an MOI of 1 $TCID_{50}$/cell, Bardeletti *et al.* (1979) also observed that budding and virus accumulation were initially detectable at intracellular sites and that budding

from the plasma membrane occurred only at later time points, at which time budding was predominant at intracellular sites. The only reports inconsistent with this progression are those in which budding is observed primarily at the plasma membrane. However, in only one study was budding solely at the plasma membrane reported (Payment *et al.*, 1975b), and in this study only 10% of the cells contained virus and thus a representative sample might not have been observed.

The intracellular budding of virions is observed both into the Golgi apparatus (von Bonsdorff and Vaheri, 1969; Matsumoto and Higashi, 1974; Bardeletti *et al.*, 1979) and into cytoplasmic vacuoles (Higashi *et al.*, 1969; McCombs *et al.*, 1968; Tuchinda *et al.*, 1969; Holmes *et al.*, 1969; Oshiro *et al.*, 1969; Edwards *et al.*, 1969; Bardeletti *et al.*, 1979). In these reports of vacuolar budding, Higashi *et al.* (1969), Holmes *et al.* (1969), and Edwards *et al.* (1969) ascribed the vacuoles as being derived from the Golgi. In only one report (Murphy *et al.*, 1968) was budding into the ER observed. Therefore, although it has become a tenet that rubella virus buds into the ER (Pettersson, 1991), this has not been definitely determined. Whether there is a preferred compartment for rubella virus budding is not known and it would be of great interest to do immunoelectron microscopy using marker antibodies specific for ER, transitional, and Golgi compartments on rubella virus-infected cells at relatively early times in the replication cycle. Later in the replication cycle, glycoproteins are transported to apparently non-predominant sites (e.g., the cell surface) and budding can occur at these sites. This is the case with the coronavirus, murine hepatitis virus-A59, which buds initially exclusively into transitional elements between the ER and Golgi (Tooze *et al.*, 1984). However, later in infection the virus glycoproteins accumulate in the smooth ER and budding occurs at this site as well. Another reason to avoid analysis at late times in infection in rubella virus-infected cells is that the cell cytoplasm becomes highly vacuolated and rearrangements of the ER occur, including association with vacuoles containing virions (Holmes *et al.*, 1968; Tuchinda *et al.*, 1969; Edwards *et al.*, 1969; von Bonsdorff and Vaheri, 1969). This disorganization and degeneration of the cytoplasm could have an effect on the specificity of budding.

H. Time Course of Virus Macromolecular Synthesis

In rubella virus-infected cells, virus-specific RNA and proteins are initially detectable at 10 to 12 hr postinfection, a finding that parallels the eclipse period observed in virus production (Sedwick and Sokol, 1970; Vaheri and Vesikari, 1971; Hemphill *et al.*, 1988). Rates of virus-specific RNA synthesis peak between 26 and 30 hr postinfection whereas rates of virus-specific protein synthesis are at peak levels by

16 hr postinfection. This is in contrast to alphavirus-infected cells, in which virus-specific macromolecules can be detected within 2 hr postinfection. Although quantitative comparisons of macromolecular synthesis have not been done, the relative level of incorporation of radiolabeled precursor into virus RNA and proteins in infected cells is qualitatively much greater in alphavirus-infected cells than in rubella virus-infected cells. The number of virions produced in alphavirus-infected cells in comparison to rubella virus-infected cells is greater by 100-fold or more. Thus, in some process or processes in the replication cycle, rubella virus is slow and/or inefficient in comparison with the alphaviruses.

The molecular basis of the relatively inefficient rate of rubella virus replication is unknown and is probably due to a complex set of factors. However, the high G + C content of the rubella virus RNA possibly contributes to the relative inefficiency of rubella virus replication. First, as discussed above, the high G + C content leads to a pattern of codon usage that is different than that found in human genes and thus the rate of translation of the rubella virus RNAs could be limited by the availability in infected cells of the tRNA isoacceptor species that their coding sequence specify. However, the rate of translation of mRNAs of herpes simplex viruses I and II, the genomic DNAs of which are 67 and 69% G + C, respectively, has been shown to be equivalent to the rate of translation of cellular mRNAs (Honess, 1984; Honess *et al.*, 1989). Second, the high G + C content of the rubella virus RNA imparts greater stability to both the secondary structure formed by the single-stranded RNA species and the double-stranded replicative complexes in comparison to viruses whose RNAs are of lower G + C content [the T_m of the rubella virus double-stranded RNA is 12°C higher than that of the alphaviruses, using the formula that T_m varies by 0.584(%G + C)]. Because of the greater stability of the rubella virus single- and double-stranded RNA species, a greater expenditure of energy is required to denature them during transcription (and translation in the case of the single-stranded RNAs). In this regard, it is of interest that the only two regions of the rubella virus genome in which the G + C content is near 50% are at the 5′ end of the RNA and at the subgenomic RNA start site. Thus, these regions of the double-stranded RNA replicative complex may be relatively easy to denature such that initiation of genomic and subgenomic RNA can be facilitated.

I. Effect of Virus Replication on Host Cell

Rubella virus infection appears to stimulate the rate of metabolism. Bardeletti *et al.* (1972) and Bardeletti (1977) noted an immediate effect in that oxygen uptake and lactic acid production were stimulated

during the first hour of infection of BHK-21 cells. Concomitantly, intracellular ATP levels dropped. After the first hour of infection, no differences between metabolism levels in infected and uninfected cells were detected during the early eclipse phase (2 to 5 hr postinfection). However, as infection proceeded oxygen uptake was stimulated such that by the time of peak virus production (20–22 hr post infection) oxygen uptake was 20% greater in infected cells than in control uninfected cells. At this time in the infection cycle, ATP levels were also higher in infected cells than in uninfected cells. Vaheri and Cristofalo (1967) also noted a stimlation of metabolism in infected BHK-21 cells as measured by an increase in glucose utilization and lactate production in comparison to uninfected controls at time points after 24 hr postinfection. An interesting finding in the studies of Bardeletti *et al.* (1972) and Bardeletti (1977) was that within the first hour of rubella virus infection, mitochondria were protected from the uncoupling effect of 2,4-dinitrophenol. These authors interpreted this finding as indicative of mitochondrial involvement of rubella virus replication.

Stimulation of membrane biosynthesis has also been reported in rubella virus-infected cells. Bardeletti and Gautheron (1976) found an increased lipid content in infected BHK-21 cells in comparison to uninfected cells. In addition to an increase in the amount of lipid present, the overall lipid composition in infected cells differed from that of uninfected cells (Bardeletti and Gautheron, 1976; Voiland and Bardeletti, 1980; Bardeletti and Voiland, 1981). Specifically, in comparison to uninfected cells the overall percentage of phospholipid decreased, the relative amounts of phosphatidylcholine and linoleic acid increased, and a novel glycolipid was discovered (this glycolipid was not present in virions). Payment *et al.* (1975c) isolated a novel membrane fraction from infected BHK-21 cells. Electron microscopy of specimens taken early in infection of BHK-21 cells revealed a proliferation of smooth membranes in the Golgi, particularly those budding virus (von Bonsdorff and Vaheri, 1969). During prolonged infection (2 days and beyond) of BHK-21, Vero, RK-13, and SIRC cells, membrane alterations observed include vacuolarization of the cytoplasm (McCombs *et al.*, 1968; Higashi *et al.*, 1969; Tuchinda *et al.*, 1969; Hamvas *et al.*, 1969; Holmes *et al.*, 1969; von Bonsdorff and Vaheri, 1969), proliferation and distention of the membranes of both the ER and the Golgi (Higashi *et al.*, 1969; Tuchinda *et al.*, 1969; Holmes *et al.*, 1969; Edwards *et al.*, 1969; Chatterji *et al.*, 1969), and occasionally the appearance of structures with unusual morphologies such as crystalline inclusions (Higashi *et al.*, 1969; Kim and Boatman, 1967; Holmes *et al.*, 1968, 1969) and annulate lamellae (Kim and Boatman, 1967; Patrizi

and Middelkamp, 1970). Expansion, alteration, and rearrangement of cytoplasmic membranous structures is commonly observed in cells infected with other viruses that bud from intracellular membranes (Kuismanen *et al.*, 1984; Tooze *et al.*, 1985). In the case of the bunyavirus, Uukuniemi virus, alteration of the Golgi apparatus is triggered by the presence of the virus glycoproteins in the absence of other virus gene products or virus maturation (Gahmberg *et al.*, 1986). The only evidence to indicate an effect of the rubella virus glycoproteins on the structure of intracellular membranes is the appearance of the novel ER–Golgi intermediate compartment in cells in which E1 is expressed alone (Baron *et al.*, 1992; Hobman *et al.*, 1993).

Studies on the effect of rubella virus replication on host cell macromolecular synthesis have led to seemingly contradictory results. Maes *et al.* (1966) and Hemphill *et al.* (1988) found no inhibition of cell RNA synthesis in BHK-21 and Vero cells, respectively, through 72 hr postinfection. However, although Vaheri and Cristafalo (1967) also found no inhibition of cell RNA synthesis in BHK-21 cells at 24 hr postinfection, after this time inhibition of cell RNA synthesis was detectable and progressed to complete inhibition by 56 hr postinfection. Similar to the results with cell RNA synthesis inhibition, no inhibition of cell protein synthesis could be detected at 24 hr postinfection in BHK-21 (Maes *et al.*, 1966; Vaheri and Cristofalo, 1967), Vero (Hemphill *et al.*, 1988), and RK-13 cells as well as in human peripheral blood lymphocytes (PBLs) (Chantler and Tingle, 1980). However, late in infection (48 to 72 hr postinfection), results varied. Maes *et al.* (1966) and Chantler and Tingle (1980) found no inhibition of cell protein synthesis in BHK-21 and RK-13 cells, respectively. Payment *et al.* (1975b) and Hemphill *et al.* (1988) found in infected Vero cells that cell protein synthesis was inhibited by 50%. In contrast, Vaheri and Cristofalo (1967) and Chantler and Tingle (1980) found complete inhibition of cell protein synthesis in BHK-21 and human PBLs, respectively. With the exception of Payment *et al.* (1975b), all of these studies were conducted using an input MOI of 5 infectious doses per cell or greater and in all of these studies it is clear that the inhibition of cell macromolecular synthesis observed late in infection was not simply due to cell death. Overall, no effect on cell macromolecular synthesis is exhibited early in infection, although the lack of synchronous infection would delay the time postinfection at which inhibition could be detected. Varying effects on cell macromolecular synthesis are observable late in infection and the variability is not completely due to differences in cell type.

A number of studies have shown that cells infected with rubella virus grow and divide more slowly than do uninfected cells. The range

of cell types in which this effect has been reported includes standard cell culture lines (both following initial infection and during persistent infection) (Maassab and Veronelli, 1966; Vaheri and Cristofalo, 1967; Gerna *et al.*, 1973; Stanwick and Hallum, 1974; Norval, 1979), cell strains and organ cultures derived from several fetal human organs (Plotkin *et al.*, 1965; Hoskins and Plotkin, 1967; Boue and Boue, 1969; Heggie, 1976; Yoneda *et al.*, 1986), and strains derived from congenitally infected fetuses that were found to be persistently infected with rubella virus (Rawls and Melnick, 1966). The most striking effect of rubella virus infection on cell growth and division is observed in diploid cell strains derived from human fetal lungs (Plotkin *et al.*, 1965; Hoskins and Plotkin, 1967; Boue and Boue, 1969). Rubella virus infection of such cell strains results in cessation of growth within a few passages. Rubella virus infection of cell strains from other human fetal organs (skin, pharyngeal mucosa, hypophysis, thymus, pericardium, brain, kidney, and pituitary) results in effects on cell growth and division ranging from modest to inapparent and differences in effect on cell growth and division of these cell strains were observed in different studies. Interestingly, other than the effect on cell growth and division, rubella virus infection of cell strains derived from human fetal organs results in no observable cytopathic effect.

The cessation of growth of human diploid lung strains following rubella virus infection was attributed to an inhibition of mitosis because the mitotic indices of infected cells were significantly below those of uninfected cells (Plotkin *et al.*, 1965). Gerna *et al.* (1972a,b) reported a significant decrease in the mitotic indices of Vero, primary rabbit kidney and SIRC (rabbit cornea) cells infected with virulent, but not attenuated vaccine, strains of rubella virus. The mechanism by which rubella virus inhibits mitosis is not known. Vaheri and Cristafalo (1967) and Gerna *et al.* (1973) reported an inhibition of DNA synthesis in rubella virus-infected cells. Plotkin and Vaheri (1967) described a protein present in the culture fluid of rubella virus-infected WI-38 (diploid human lung) cells that inhibited mitosis in uninfected WI-38 and human skin cells. Anti-rubella virus antiserum had no effect on the activity of this protein and the protein was not an interferon. Bowden and Westaway (1987) reported disaggregation of the microfilament system in infected Vero cells and proposed that this could lead to inhibition of mitosis by interfering with construction of the mitotic spindle.

The effect of rubella virus on cell division is of clinical significance because one of the manifestations of CRS is a reduction of the number of cells in affected organs and it is thought that this could be caused by reduced growth rate of lack of division of rubella virus-infected cells (reviewed in Rawls, 1974). Interestingly, only one in 10^3 to 10^5 of the

cells from infected fetal organs can be shown to harbor virus. However, rubella virus infection of precursor or stem cells early in embryogenesis with resultant inhibition of cell growth and division could certainly lead to a reduction in the number of cells in the organ that eventually develops. Elaboration of an antimitotic substance by a small number of infected cells would also lead to a reduction of cell number. Another mechanism by which rubella virus infection *in utero* could stunt organ development was suggested by the studies of Yoneda *et al.* (1986) on human palate-derived embryonic mesenchymal cells that maintained differentiated function *in vitro*. Persistent infection of these cells *in vitro* by rubella virus led to a twofold decrease in responsiveness to human epidermal growth factor in comparison to uninfected cells.

In association with the studies on the effect of rubella virus infection on mitosis, it was found that the number of chromosomal breakages was increased in human diploid cell strains infected *in vitro* with rubella virus in comparison to uninfected cell strains (Plotkin *et al.*, 1965; Chang *et al.*, 1966; Hoskins and Plotkin, 1967; Boue and Boue, 1969). It was also found that cell strains derived from a percentage of congenitally infected fetuses contained an increased number of chromosome breakages in comparison to strains derived from fetuses aborted for reasons other than congenital rubella virus infection (Chang *et al.*, 1966). However, because the occurrence of significantly higher rates of chromosomal breakage in cell strains from congenitally infected fetuses was not uniform, induction of chromosome breakage is not considered to be a primary mechanism of teratogenesis (Chang *et al.*, 1966).

Despite the lack of recognizable cytopathology in rubella virus-infected cell strains derived from human fetuses, in tissues from congenitally infected human fetuses noninflammatory necrotic lesions are present that are apparently due to direct virus-mediated cell destruction (reviewed in Wolinsky, 1990a). Electron microscopic examination of tissues from a large number of congenitally infected fetuses revealed neither the presence of virions nor the type of cytopathic changes encountered in culture cells as described above (Kistler, 1975). However, the presence of tubular aggregates enclosed in cisternae of the ER and an increased number of nuclear bodies in a proportion of cells from a number of organs were observed. The proportion of cells containing tubuloreticular complexes was highest in the vicinity of lesions. Similar tubular aggregates are routinely observed in cells from patients infected with other viruses, in tumors, and in tissues from patients suffering from autoimmune diseases such as systemic lupus erythematosus and thus are not specific to rubella virus.

VII. Rubella Virus and Togavirus Evolution

A. Microevolution

The mutability of RNA virus genomes due to the lack of proofreading enzymes associated with RNA-dependent RNA replication was first recognized in 1982 (Holland *et al.*, 1982). The initial predictions to be made following recognition of the phenomenon were that RNA viruses should evolve with extreme rapidity and exhibit considerable sequence diversity. Both of these predictions were borne out with a number of viruses, most nefariously HIV-1, isolates of which can vary from each other in nucleotide sequence by 13% (30% in the envelope protein-coding region) and with which antigenic variation can be documented within a single infected individual (Wong-Staal, 1990). Within a single serotype of other viruses [e.g., poliovirus serotype 1 and both the Indiana and New Jersey serotypes of vesicular stomatitis virus (VSV)], nucleotide variabilities between strains of up to 18 to 23% have been reported (Rico-Hesse *et al.*, 1987; Nichol *et al.*, 1989; Bilsel *et al.*, 1990; Bilsel and Nichol, 1990). However, in a number of RNA viruses, nucleotide variability among strains is not of this magnitude and rubella virus falls into this group of viruses. From the reported nucleotide sequences, sequence variability in the structural protein-coding region is between 2.5 and 3.5% among independent strains of rubella virus (Table IV). In a study that compared the E1 coding sequence of 11 independent strains and isolates of rubella virus from both North America and Europe, sequence variability was between 0.7 and 3.6% (Frey and Abernathy, 1993). The recently completed nucleotide sequence of the nonstructural protein ORF of the M33 strain of rubella virus (S. Gillam, personal communication) varies from the Therien nonstructural protein ORF sequence by 2.5%. Other viruses that exhibit limited interstrain variability in structural protein-coding regions are measles virus [up to 7.2% in one study; (Taylor *et al.*, 1991), and from 0.5–4.8% in a second study; (Rota *et al.*, 1994)] human type 3 parainfluenza virus (1.2 to 5.8%; van Wyke Coelingh *et al.*, 1988); influenza C virus (0.1 to 6.6%; Buonagurio *et al.*, 1985); and mumps virus (0.1 to 12%; Yamada *et al.*, 1989). Among alphaviruses, Sindbis virus strains vary by 5.0 to 5.7% in nucleotide sequence of the total genome or the glycoprotein-coding region, but can vary by up to 20% in the 3′ noncoding region (Russell *et al.*, 1989; Shirako *et al.*, 1991). North American isolates of Eastern equine encephalitis were found to vary by only 0.7% in the structural protein-coding region (Weaver *et al.*, 1991).

With study of the mutability of RNA viruses, it has become apparent

TABLE IV
PERCENTAGE DIFFERENCES BETWEEN REPORTED NUCLEOTIDE AND DEDUCED AMINO ACID SEQUENCES OF RUBELLA VIRUS

Strain	Therien (F)	RA 27/3	M33	HPV-77
		Nucleotide sequence[a]		
Therien (G)	0.6	2.4	3.5	3.3
Therien (F)		2.5	3.4	3.3
RA 27/3			2.9	2.8
M33				0.3
		Amino acid sequence[b]		
Therien (G)	0.8	3.0	2.6	2.4
Therien (F)		3.0	2.5	2.2
RA 27/3			2.5	2.2
M33				0.6

[a]Percentage difference in reported nucleotide sequence of the subgenomic RNA of three independent strains of rubella virus. The sequence of the Therien strain was determined independently in two labs (G = Georgia State University; F = University of Helsinki, Finland) and the HPV-77 is a vaccine strain derived by serial passage of the M33 strain in cell culture. Sequence citations: Therien (G): Frey *et al.* (1986), Frey and Marr (1988); Therien (F): Vidgren *et al.* (1987), Takkinen *et al.* (1988); RA 27/3: Nakhasi *et al.* (1989); M33: Clarke *et al.* (1987); HPV-77: Zheng *et al.* (1989).

[b]Difference in deduced amino acid sequence of the structural protein ORF.

that the high mutability of RNA viruses leads to their existence both in the laboratory and in nature as complex populations of "quasispecies" (Holland *et al.*, 1992). Although the rate of evolution of a particular virus can be rapid and intraclonal variability can exist, both population equilibrium dynamics as well as selection can lead to maintenance of a consensus sequence within the virus population in a microenvironment. Thus, even viruses that are capable of rapid evolution under laboratory conditions or that exhibit a wide degree of nucleotide sequence divergence on a worldwide basis can exhibit local or regional stasis. Consistent with the quasispecies population concept, with most viruses that cause human disease, it is found that several distinct genetic lineages cocirculate [the exception to this rule is the influenza A viruses, which show a linear evolutionary progression on a worldwide basis (Buonagurio *et al.*, 1985)]. Sequence analysis of rubella virus isolates provided evidence for the existence of cocirculating lineages that have a worldwide distribution (Frey and Abernathy, 1993).

At present, there is no answer to the question concerning the reason for relative lack of interstrain sequence divergence exhibited by

rubella virus and discussion of the question must consider both the inherent mutation rate exhibited by the rubella virus replicase and the selective constraints faced by rubella virus in its replication and transmission in nature. The generally accepted range of the rate of RNA virus mutation is between 10^{-3} and 10^{-6} mutations per nucleotide per round of replication and rates of evolution in nature of between 10^{-2} and 10^{-4} substitutions per nucleotide per year have been determined. There is no evidence as to the rate of mutation during rubella virus replication and, because of the phylogenetic grouping of the genetic lineages of rubella virus, calculation of the rate of rubella virus evolution in nature has thus far not been possible (Frey and Abernathy, 1993), although considering the low degree of nucleotide variability among rubella virus strains the rate of rubella virus evolution is probably on the low end of the spectrum. It is conceivable that the low degree of variability exhibited among rubella virus strains is due to a lower rate of mutation on the part of the rubella virus replicase, possibly due its being less error prone than the replicases of other RNA viruses or because rubella virus expresses a replicase with proofreading capacity. There is no experimental evidence relating to either of these possibilities. It should also be considered that, due to its slow replication cycle and relatively small number of progeny virus produced per infected cell, rubella virus undergoes fewer rounds of RNA replication per given amount of time in culture or in nature than do viruses that replicate more rapidly and to higher titers, thus providing less opportunity for mutation to occur. Interestingly, however, the sequence diversity of the Therien strain of rubella virus maintained in different laboratories for over 13 years is 0.6% whereas the M33 progenitor wild-type strain and the HPV-77 vaccine strain attenuated by 77 passages in AGMK cells vary by 0.3% (Table IV). These differences in sequence are similar to the sequence diversity of strains of other viruses maintained in different laboratories or passaged in cell culture to produce attenuated vaccines. For example, the same Sindbis virus strains maintained in different laboratories differ in sequence by 0.2 to 0.3% (Lustig *et al.*, 1988), attenuation of yellow fever vaccine by 240 passages in culture cells and animals led to a sequence difference of 0.6% (Hahn *et al.*, 1987), and the Sabin type 1 and type 3 poliovirus vaccines differ in sequence by 0.8 and 0.1%, respectively, from their wild-type parents despite over 70 passages in culture cells and animals in each case (Nomoto *et al.*, 1982; Stanway *et al.*, 1984).

In terms of the effect of selection on the interstrain sequence diversity exhibited by rubella virus, an obvious selective force is the interaction of the virus structural proteins with the human immune system. Considering the tendency of other viruses to exhibit antigenic

drift or to diverge into multiple serotypes, it is amazing that viruses such as rubella virus stubbornly maintain a single serotype. Presumably, the structure of the virus structural proteins is sufficiently rigid such that changes in amino acid sequence are not tolerated.

Analysis of the distribution of the nucleotides at which sequence variation between rubella virus strains occurs is also instructive in attempting to elucidate the selective pressures operative in rubella virus evolution. Among the five sequence determinations of the structural protein-coding region (Table IV), 148 of the 3189 nucleotides (4.6%) show variation. Of these, 20% are at first-codon positions, 15% are at second-codon positions, and 64% are at third-codon positions. In a collection of sequences of the E1 coding regions of 11 independent strains (Frey and Abernathy, 1993), 123 of 1300 nucleotides sequenced (9.5%) showed variation (the greater number of nucleotides exhibiting variation in this study in comparison to the structural protein ORF sequences is undoubtedly due to the inclusion of sequences from a larger number of strains). Of these, 18% were at first-codon positions, 11% were at second-codon positions, and 71% were at third-codon positions. Thus, although as expected the majority of nucleotides at which differences were observed are at third-codon positions, only 9% of the third-codon positions in the structural protein ORF sequence collection and 20% of the third-codon positions in the E1 sequence collection exhibited variability. If selective pressure were primarily exerted by maintenance of amino acid sequence, then variation would be expected at a much higher percentage of third-codon positions. Indeed, in a comparative sequence study of poliovirus serotype I strains across a region of nucleotides encoding parts of the VP1 capsid protein and NVP2A nonstructural protein, variability was encountered at every third-codon position (Rico-Hesse *et al.*, 1987) and in a collection of sequences of the G protein gene from multiple strains of the VSV New Jersey serotype, variability was present at 89% of the third-codon positions (Nichol *et al.*, 1989). In actuality, of the nucleotides that exhibit interstrain variability, in rubella virus a lower percentage are at third-codon positions (64 to 71%) than in either poliovirus (98%) or VSV (84%). This results in the range of variability in deduced amino acid sequence among strains of rubella virus (2.2 to 3.0% in the complete structural protein ORF, Table IV; 0.2 to 2.9% in the E1 comparison study) being closer to the range of variability in the poliovirus and VSV amino acid sequences (up to 4% and 3 to 9%, respectively) than is the range of variability in nucleotide sequences.

Considering the interstrain invariability of nucleotide sequence at over 90% of the total residues and 80% of the third-codon positions in the rubella virus structural protein-coding sequences, selective pres-

sure other than maintenance of amino acid sequence must be operative in the evolution of rubella virus. The selective pressures are unknown but possibilities include the distribution of isoacceptor species in cells in which the virus replicates, maintenance of the negative-polarity ORFs, the maintenance of high G and C content, (G → C and C → G substitutions are transversions, which are much less likely to occur than are transitions), or conformational requirements of the virion RNA. It should be pointed out that the conservation of the majority of nucleotides among strains is not unique to rubella virus and is evident in the sequence collections cited above for a number of other viruses (measles, mumps, influenza C, and human parainfluenza type 3).

B. Macroevolution

From the comparative schematic genetic map of the genomes of rubella virus and Sindbis virus shown in Fig. 3, the most straightforward conclusion as to the evolution of the togaviruses is that rubella virus and the alphaviruses diverged from a common ancestor. However, extensive computer-assisted homology analysis has shown that the evolution of these genera was probably much more complicated.

To recapitulate the similarities in genomic coding and replication strategy between the genera *Rubivirus* and *Alphavirus*, the following features are shared: the presence of two long ORFs, the 5′ terminal of which encodes nonstructural proteins and the 3′ terminal of which encodes structural proteins; the gene order within the structural protein ORF of a soluble capsid protein followed by two envelope glycoproteins; and the translation of the structural protein ORF from a single species of subgenomic RNA that is synthesized by internal initiation on a genome-length minus-polarity RNA template. An interesting comparative feature of the genomes of the two genera is that the rubella virus genome is 20% shorter than is the Sindbis virus genome and that this compression is apparent in all regions of the genome: the nonstructural protein ORF is 2115 amino acids in rubella virus versus 2513 amino acids in Sindbis virus; the structural protein ORF is 1063 amino acids in rubella virus versus 1245 amino acids in Sindbis virus; and the 3′ terminal nontranslated region is 62 nucleotides in rubella virus versus 319 nucleotides in Sindbis virus.

There are two clear differences in the genomic coding strategies of rubella virus and the alphaviruses. The first comprises several variations in the structural protein expression strategy: the retention of the E2 signal sequence on the rubella virus capsid protein and its absence on the alphavirus capsid protein, the alphavirus PE2 → E2 + E3 cleav-

age in contrast to the absence of cleavages in the maturation of the rubella virus glycoproteins; the presence of *O*-glycans on rubella virus E2 that are absent from both alphavirus glycoproteins, and the existence of the 6K protein in the Sindbis virus structural protein ORF between E2 and E1 that is not present in the rubella virus structural protein ORF. Whether these differences could be the result of simple evolution from a common ancestor is not clear, particularly considering that the first of these differences, the retention or absence of the E2 signal sequence on the capsid protein, underscores a significant difference between the two genera in both the mechanism of capsid morphogenesis and site of budding. The second difference between the genomes of rubella virus and the alphaviruses is in the order of motifs within the nonstructural protein ORF. It was this difference that gave the initial indication of the complicated nature of togavirus evolution.

Both rubella virus and the alphaviruses belong to a large superfamily that consists of the togavirus and hepatitis E virus families of animal viruses and several families of plant viruses (reviewed in Goldbach *et al.*, 1991). The existence of this superfamily was originally recognized as a result of computer-assisted homology analysis, which detected the presence of homology between the deduced amino acid sequence of nsP1, nsP2, and nsP4 of Sindbis virus and members of three plant virus families, the family Tobamoviridae (tobacco mosaic virus), the family Bromoviridae (brome mosaic virus), and the family Ilaviridae (alfalfa mosaic virus) (Haselhoff *et al.*, 1984; Ahlquist *et al.*, 1985). Later analysis demonstrated that these regions of homology were regions associated with methyltransferase, helicase, and replicase activity, respectively (Kamer and Argos, 1984; Gorbalenya *et al.*, 1988; Rozanov *et al.*, 1992). The families in the alphavirus-like superfamily are diverse in genomic structure and coding strategy, sharing only the presence of these three regions of homology and the translation of the structural protein(s) from a subgenomic RNA. As a demonstration of the diversity within the superfamily, it contains viruses with mono-, bi-, and tripartite genomes, viruses with naked helical, naked icosahedral, and enveloped icosahedral virions, and viruses with 3′ poly(A) tracts and 3′ tRNA-like structures. Of interest, only the animal virus families within the superfamily contain the X domain.

Because the order of the methyltransferase, helicase, and replicase domains is similar to all of the families within the alphavirus-like superfamily, it is assumed that these families diverged from a common ancestor (reviewed in Goldbach, 1990). The differences in capsid structure and presence of domains in subsets of families within the superfamily are assumed to be due to interviral recombination. Recombina-

tion between nonsegmented, positive-polarity RNA viruses has been demonstrated in nature and in the laboratory in the families Picornaviridae, Togaviridae (alphavirus), and Coronaviridae families (reviewed in Lai, 1992). In this regard, interviral recombination between rubella virus and a latent retrovirus (R-virus) of BHK-21 cells has been reported (Sato *et al.*, 1976, 1978; Yamamoto and Urade, 1989; Urade and Yamamoto, 1990; Urade *et al.*, 1993). The hybrid viruses have a virion morphology similar to that of the R-virus and contain a DNA polymerase activity. The hybrids reacted with anti-rubella virus complement-fixing antibodies but were not neutralized by anti-rubella virus serum and were recognized by anti-C and anti-E2, but not anti-E1, monoclonal antibodies. It was hypothesized that the hybrids were formed by a recombination event in which the E1-coding region of rubella virus was replaced with the reverse transcriptase and envelope-coding region of the R-virus. However, this has not been confirmed by sequence determination.

In the picornaviruses, recombination has been shown to be achieved by a "copy choice" mechanism by which an RNA replicase and associated RNA transcript leave one template and continue transcription on another, resulting in a chimeric transcript. With respect to the difference in order of global motifs in the nonstructural protein ORFs of rubella virus and the alphaviruses, rearrangement by copy choice mechanism would involve minimally four copy choice events. The other alternative would be independent donation of the protease and X domains to a common ancestor.

Surprisingly, when the deduced amino acid sequences surrounding the helicase and replicase motifs of members of the alphavirus-like superfamily were used to create phylogenetic trees (Koonin, 1991; Weaver *et al.*, 1993) (Fig. 10), the alphaviruses segregated onto one branch with the hordei-, tobra-, tobamo-, ilar-, bromo-, cucumo-, and closteroviruses whereas rubella virus segregated onto a different branch with hepatitis E virus and the furoviruses [beet yellow necrotic vein virus (BYNVV)]. A third branch contained the carla-, potex-, and tymoviruses along with apple chlorotic leafspot virus (previously classified as a closterovirus). More detailed analysis of the complete nonstructural protein ORFs of BYNVV, hepatitis E virus, and rubella virus revealed additionally a "Y domain" of unknown function in all three viruses and a proline-rich hinge region in hepatitis E virus and rubella virus, neither of which were present in the alphavirus nonstructural protein ORF (Koonin *et al.*, 1992; Koonin and Dolja, 1993). The order of the motifs was methyltransferase–Y–helicase–protease–replicase in BYNVV, methyltransferase–Y–protease–proline hinge–X–helicase–replicase in hepatitis E virus, and methyltransferase–Y–proline hinge–X–protease–helicase–replicase in rubella virus. Thus,

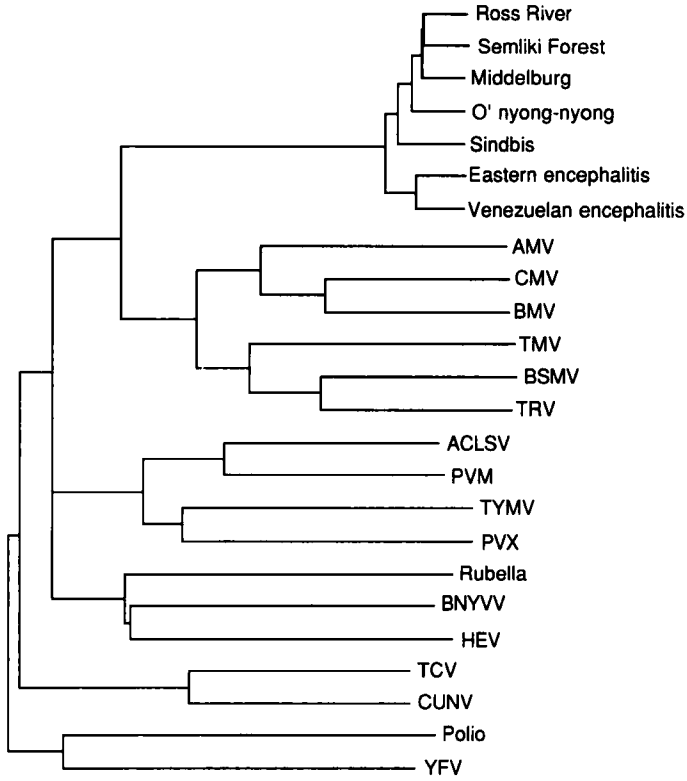

FIG. 10. Phylogenetic tree of replicases of members of the alphavirus-like superfamily. Relatedness is proportional to horizontal distance between viruses and branch points on the tree; vertical spacing is irrelevant to relatedness. The tree is rooted by using members of other virus superfamilies (YFV, yellow fever virus). Viruses at the top of the tree with unabbreviated names are all alphaviruses. Abbreviations: AMV, alfalfa mosaic virus (family Ilarviridae); CMV, cucumber mosaic virus (family Cucumoviridae); BMV, brome mosaic virus (family Bromoviridae); TMV, tobacco mosaic virus (family Tobamoviridae); BSMV, barley stripe mosaic virus (family Hordeiviridae); TRV, tobacco rattle virus (family Tobraviridae); ACLSV, apple chlorotic leafspot virus (family Closteroviridae); PVM, potato virus M (family Carlaviridae); TYMV, turnip yellow mosaic virus (family Tymoviridae); PVX, potato virus X (family Potexviridae); HEV, hepatitis E virus; BYNVV, beet yellow necrotic vein virus (family Furoviridae); TCV, turnip crinkle virus (family Carmoviridae); CUNV, cucumber necrosis virus (family Tombusviridae). [The tree was provided by S. Weaver and is from Weaver *et al.* (1993).]

with the exception of the protease domain, the order of these domains is similar among these viruses.

Overall, the conclusion of these phylogenetic analyses is that the nonstructural protein ORFs of rubella virus and the alphaviruses are more closely related to the nonstructural protein-coding regions of other virus families than to each other. This makes it unlikely that

rubella virus and the alphaviruses simply diverged from a common ancestor unless a virus with an alphavirus/rubella virus-like genome organization was the ancestral virus of most or all of the alphavirus-like superfamily (Weaver *et al.*, 1993). In this scenario, the alphavirus and rubella virus branches diverged from this ancestor, accumulating the differences in the helicase and replicase motifs in the process. The X, Y, proline hinge, and protease motifs were donated by interviral recombination after divergence of these branches. The other viruses in the superfamily evolved from these branches by interviral recombination (e.g., hepatitis E virus could have been generated by truncation of the rubella virus subgenomic ORF or by recombination between a rubella-like virus and a calicivirus). An alternative to the ancestral alphavirus theory is that the genome organization of alphaviruses and rubella virus is an example of convergent evolution. A final possibility is that a number of recombinational events among members of different branches of the alphavirus-like superfamily could have led to the evolution of the alphaviruses and rubella virus. For example, rubella virus could have arisen from a recombinational event between an alphavirus-like and a BYNVV-like or hepatitis E-like virus. Conversely, the alphaviruses could have arisen by recombination between a rubella-like virus and another alphavirus-like superfamily member such as TMV. The upcoming studies of the nonstructural proteins of rubella virus will be fascinating in terms of resolving the evolution of these viruses. If interviral recombination among members of the alphavirus-like superfamily led to the generation of the alphaviruses and rubella virus, the most closely related regions of the rubella virus and alphavirus genomes would be the structural protein ORF. Thus, studies of the structural proteins of rubella virus and alphaviruses in terms of basic conformation may also be of great evolutionary interest. Taxonomically, because of the differences in the nonstructural protein-coding region, it has been suggested that the family Togaviridae be disbanded or elevated to the Order level and that the genera *Alphavirus* and *Rubivirus* be elevated to family status (Koonin and Dolja, 1993; Ward and Shukla, 1993).

VIII. Future Directions

Medically, the current challenge posed by rubella virus is to achieve complete vaccination coverage to prevent resurgences such as occurred in the United States between 1989 to 1991 and to come to grips with the problem of vaccine-associated complications in adult women. Large-scale studies done at independent facilities are needed to determine an accurate incidence of occurrence. The pathogenesis of vaccine-

associated complications also needs examination to determine what role such mechanisms as previous immunological exposure, persistent virus infection, and autoimmunity play. With respect to pathogenesis, animal models for rubella are urgently required. Molecular analysis and manipulation of the rubella virus genome has provided a host of alternative vaccine strategies such as bioengineered antigens, noninfectious particles, synthetic peptides, and infectious clones. However, the necessity for use of such alternate vaccines and their intelligent employment requires the basic epidemiological and pathogenesis studies cited above.

Worldwide, the medical emphasis differs because rubella is not considered a serious disease in comparison to other viral scourges of the developing world such as HIV, measles, influenza, hepatitis, and viral diarrhea. This is the main reason that rubella eradication is not currently being emphasized. However, rubella imposes load on any society and development of alternative vaccines could be of use in a global eradication program.

At the molecular level, the characterization of rubella virus has lagged behind other viruses, a situation that is particularly apparent in comparison with the related alphaviruses. Considering the ease of the alphavirus system and the intractibility of the rubella virus system, this will probably always be the case. However, characterization of the alphaviruses is always of use to rubella virus researchers as a comparative example. As emphasized above, the evolution of rubella virus and alphaviruses turned out to be more complicated than expected and this should stimulate interest in characterization of rubella virus. Of particular current interest with regard to the evolution of these viruses are the ongoing studies to determine the coding capacity of the rubella virus nonstructural protein open reading frame and identify the functions of RNA control sequences conserved in the two genera. An experimental development that will also help to spur molecular characterization of rubella virus is the development of a rubella virus infectious clone (Wang *et al.*, 1994), a cDNA copy of the virus genome from which infectious RNA transcripts can be synthesized *in vitro*. Thus, owing both to an increase in interest in rubella virus (both medically and molecularly) and to facilitation of laboratory technology, the pace of work on rubella virus should speed up in the near future.

ACKNOWLEDGMENTS

I thank Shirley Gillam, Eugene Koonin, Hira Nakhasi, and Scott Weaver for releasing data prior to publication. I thank Sandra Holmes for providing statistics and discussion concerning medical aspects of rubella virus. I thank Robert B. Simmons for photographic

work. I thank Lee Marr for critical reading of the manuscript and translation of articles written in French and Rosanna Petruziello for translation of articles written in Italian. I am especially indebted to James Strauss and Jerry Wolinsky for reviewing the manuscript. Research in my laboratory is supported by a grant (AI-21389) from the NIH. This manuscript was prepared while I was the recipient of a Research Career Development Award (AI-00923) from the NIH.

References

Abernathy, E. S., Wang, C.-Y., and Frey, T. K. (1990). *J. Virol.* **64,** 5183–5187.

Ahlquist, P., Strauss, E. G., Rice, C. M., Strauss, J. H., Haseloff, J., and Zimmern, D. (1985). *Proc. Natl. Acad. Sci. U.S.A.* **82,** 536–542.

Alford, C. A., and Preblud, S. R. (1990). *In* "Infectious Diseases of the Fetus and Newborn Infant" (J. S. Remington and J. O. Klein, eds.), 3rd Ed., pp. 196–240. Saunders, Philadelphia.

Al-Nakib, W., Best, J. M., and Banatvala, J. E. (1975). *Lancet* **i,** 182–185.

Alter, M., Zhen-Xin, A., Davanipour, A., Sobel, E., Min-Lai, S., and LaRue, L. (1987). *Ital. J. Neurol. Sci., Suppl.* **6,** 11.

Alvord, E. C., Jahnke, U., Fischer, E. H., Kies, K. W., Driscoll, B. F., and Compston, D.A.S. (1987). *J. Child Neurol.* **2,** 313–321.

Andino, R., Rieckhof, G. E., and Baltimore, D. (1990). *Cell* **63,** 369–380.

Andino, R., Rieckhof, G. E., Achacoso, P. L., and Baltimore, D. (1993). *EMBO J.* **12,** 3587–3598.

Anthony, R. P., and Brown, D. T. (1991). *J. Virol.* **65,** 1187–1194.

Arumugham, R. C., Seid, R. C., Jr., Doyle, S., Hildreth, S. W., and Paradiso, P. R. (1989). *J. Biol. Chem.* **264,** 10339–10342.

Atkins, G. J., Daly, E. A., Sheahan, B. J., Higgins, D. G., and Sharp, P. M. (1990). *Neuropathol. Appl. Neurobiol.* **16,** 179–180.

Badenoch, J. (1988). *Br. Med. J.* No. **297,** 750–751.

Banerjee, A. K. (1980). *Microbiol. Rev.* **44,** 175–205.

Bardeletti, G. (1977). *Intervirology* **8,** 100–109.

Bardeletti, G., and Gautheron, D. (1976). *Arch. Virol.* **52,** 19–27.

Bardeletti, G., and Voiland, A. (1981). *Arch. Virol.* **68,** 285–289.

Bardeletti, G., Henry, M., Sohier, R., and Gautheron, D. C. (1972). *Arch. Gesamte Virusforsch.* **39,** 26–34.

Bardeletti, G., Kessler, N., and Aymard-Henry, M. (1975). *Arch. Virol.* **49,** 175–186.

Bardeletti, G., Tektoff, J., and Gautheron, D. (1979). *Intervirology* **11,** 97–103.

Baric, R. S., Carlin, L. J., and Johnston, R. E. (1983). *J. Virol.* **45,** 200–205.

Baron, M. D., and Forsell, K. (1991). *Virology* **185,** 811–819.

Baron, M. D., Ebel, T., and Suomalainen, M. (1992). *J. Gen. Virol.* **73,** 1073–1086.

Bart, K. J., and Lin, K. F., -Y. C. (1990). *Pediatr. Clin. North Am.* **37,** 735–756.

Best, J. M. (1993). *Curr. Med. Lit.—Virol.* **2,** 35–40.

Best, J. M., and Banatvala, J. E. (1970). *J. Gen. Virol.* **9,** 215–223.

Best, J. M., Thomson, A., Nores, J. R., O'Shea, S., and Banatvala, J. E. (1992). *Intervirology* **34,** 164–168.

Bilsel, P. A., and Nichol, S. T. (1990). *J. Virol.* **64,** 4873–4883.

Bilsel, P. A., Rowe, J. E., Fitch, W. M., and Nichol, S. T. (1990). *J. Virol.* **64,** 2498–2504.

Blumenthal, T., and Carmichael, G. G. (1979). *Annu. Rev. Biochem.* **48,** 525–548.

Bohn, E. M., and Van Alstyne, D. (1981). *Virology* **111,** 549–554.

Boue, A., and Boue, J. (1969). *Am. J. Dis. Child.* **118,** 45–48.

Bowden, D. S., and Westaway, E. G. (1984). *J. Gen. Virol.* **65,** 933–943.

Bowden, D. S., and Westaway, E. G. (1985). *J. Gen. Virol.* **66,** 201–206.
Bowden, D. S., and Westaway, E. G. (1989). *Subcell. Biochem.* **15,** 203–231.
Buimovici-Klein, E., and Cooper, L. Z. (1985). *Rev. Infect. Dis.* **7,** S123–S128.
Buonagurio, D. A., Nakada, S., Desselberger, U., Krystal, M., and Palese, P. (1985). *Virology* **146,** 221–232.
Cammisa-Parks, H., Cisar, L. A., Kane, A., and Stollar, V. (1992). *Virology* **189,** 511–524.
Cao, X.-Q., Liu, T.-Y., and Nakhasi, H. L. (1992). *Gene* **114,** 251–256.
Carver, D. H., and Marcus, P. I. (1968). *Bacteriol. Proc.* p. 181.
CDC (1980). "Rubella Surveillance, Jan. 1976–Dec. 1978," May 1980. p. 2.
CDC (1989). *Morbid. Mortal. Wkly. Rep.* **38,** S-9, p. 3.
CDC (1990). *Morbid. Mortal. Wkly. Rep.* **39,** Recommendations and Reports (RR-15), 8.
CDC (1991a). *Morbid. Mortal. Wkly. Rep.* **40,** 93–99.
CDC (1991b). *Morbid. Mortal. Wkly. Rep.* **40,** 264–265.
CDC (1991c). *Morbid. Mortal. Wkly. Rep.* **40,** Recommendations and Reports (RR-12), 24–26
CDC (1992a). *Morbid. Mortal. Wkly. Rep.* **41,** 468–476.
CDC (1992b). *Morbid. Mortal. Wkly. Rep.* **41,** 67.
CDC (1993). *Morbid. Mortal. Wkly. Rep.* **42,** 378–381.
Chambers, T. J., Hahn, C. S., Galler, R., and Rice, C. M. (1990). *Annu. Rev. Microbiol.* **44,** 649–688.
Chambers, T. J., Grakoui, A. and Rice, C. M. (1991). *J. Virol.* **65,** 6042–6050.
Chang, T. H., Moorhead, P. S., Boue, J. G., Plotkin, S. A., and Hoskins, J. M. (1966). *Proc. Soc. Exp. Biol. Med.* **122,** 236–243.
Chantler, J. K. (1981). *Infect. Immun.* **32,** 1274–1280.
Chantler, J. K., and Tingle, A. J. (1980). *J. Gen. Virol.* **50,** 317–328.
Chantler, J. K., Ford, D. K., and Tingle, A. J. (1982). *Lancet* **1,** 1323–1325.
Chantler, J. K., Tingle, A. J., and Petty, R. E. (1985a). *N. Engl. J. Med.* **313,** 1117–1123.
Chantler, J. K., da Roza, D. M., Bonnie, M. E., Reid, G. D., and Ford, D. K. (1985b). *Ann. Rheum. Dis.* **44,** 564–568.
Chatterji, J., Beswick, T. S. L., and Chapman, J. A. (1969). *J. Gen. Virol.* **4,** 371–377.
Chattopadhyay, H., Chattopadhyay, C., Natvig, J. B., Wiger, D., and Mellbye, O. J. (1979). *Scand. J. Immunol.* **10,** 47–54.
Chaye, H., Chong, P., Tripet, B., Brush, B., and Gillam, S. (1992). *Virology* **189,** 483–492.
Cherry, J. D. (1992). *In* "Textbook of Pediatric Infectious Diseases" (R. D. Feigin and J. D. Cherry, eds.), 3rd Ed., pp. 1792–1817. Saunders, Philadelphia.
Clarke, D. M., Loo, T. W., Hui, I., Chong, P., and Gillam, S. (1987). *Nucleic Acids Res.* **15,** 3041–3057.
Clarke, D. M., Loo, T. W., McDonald, H., and Gillam, S. (1988). *Gene* **65,** 23–30.
Clarke, W. L., Shaver, K. A., Bright, G. M., Rogol, A. D., and Nance, W. E. (1984). *J. Pediatr.* **104,** 370–373.
Collett, M. S., Larson, R., Gold, C., Strick, D., Anderson, D. K., and Purchio, A. F. (1988). *Virology* **165,** 191–199.
Compston, D. A. S., Vadarelis, B. N., Paul, E., McDonald, W. I., Batchelor, J. R., and Mims, C. A. (1986). *Brain* **109,** 325–344.
Cooper, L. Z., Green, R. H., Krugman, S., Giles, J. P., and Mirick, G. S. (1965). *Am. J. Dis. Child.* **110,** 416–427.
Cooper, L. Z., Ziring, P. R., Weiss, H. J., Matters, B. A., and Krugman, S. (1969). *Am. J. Dis. Child.* **118,** 218–225.
Cooper, L. Z., Florman, A. L., Ziring, P. R., and Krugman, S. (1971). *Am. J. Dis. Child.* **122,** 397–403.

Cusi, M. G., Rossolini, G. M., Cellesi, C., and Valensin, P. E. (1988). *Arch. Virol.* **101,** 25–33.
Cusi, M. G., Metelli, R., and Valensin, P. E. (1989). *Arch. Virol.* **106,** 63–72.
de Mazancourt, A., Waxham, M. N., Nicolas, J. C., and Wolinsky, J. S. (1986). *J. Med. Virol.* **19,** 111–122.
den Boon, J. A., Snijder, E. J., Chirnsie, E. D., de Vries, A. A. F., Horzinek, M. C., and Spaan, W. J. M. (1991). *J. Virol.* **65,** 2910–2920.
De Prins, F., Van Assche, F. A., Desmyter, J., DeGroote, G., and Gepts, W. (1978). *Lancet* **i,** 439–440.
Desmyter, J., DeSomer, P., Rawls, W. E., and Melnick, J. L. (1969). *Symp. Ser. Immunobiol. Stand.* **11,** 139–148.
Dominguez, G. (1991). Ph.D. Thesis, Georgia State Univ., Atlanta.
Dominguez, G., Wang, C.-Y., and Frey, T. K. (1990). *Virology* **177,** 225–238.
Doms, R. W., Lamb, R. A., Rose, J. K., and Helenius, A. (1993). *Virology* **193,** 545–562.
Dorsett, P. H., Miller, D. C., Green, K. Y., and Byrd, F. I. (1985). *J. Infect. Dis.* **7,** S150–S156.
Dudgeon, J. A., Marshall, W. C., and Peckham, C. S. (1972). *Lancet* **ii,** 480–481.
Edwards, M. R., Cohen, S. M., Bruno, M., and Deibel, R. (1969). *J. Virol.* **3,** 439–444.
Epidemiological Notes (1989). *Epidemiol. News Bull. (Singapore)* **15,** 66–68.
Felgenhauer, K., Schaklich, H. J., Nedic, M., and Ackermann, R. (1985). *J. Neurol. Sci.* **71,** 291–299.
Fenner, F. (1975-1976). *Intervirology* **6,** 1–12.
Fenner, F., Pereira, H. G., Porterfield, J. S., Joklik, W. K., and Downie, A. W. (1974). *Intervirology* **3,** 193–198.
Forbes, J. A. (1969). *Am. J. Dis. Child.* **118,** 5–11.
Ford, D. K., da Roza, D. M., Reid, G. D., Chantler, J. K., and Tingle, A. J. (1982). *J. Rheumatol.* **9,** 420–423.
Ford, D. K., Tingle, A. J., and Chantler, J. K. (1988). *In* "Infections in the Rheumatic Diseases" (L. Espinoza, ed.), pp. 103–106. Grune & Stratton, Orlando, Florida.
Francki, R. I. B., Fauquet, C. M., Knudson, D. L., and Brown, F., eds. (1991). "Classification and Nonmenclature of Viruses. Fifth Report of the International Committee on Taxonomy of Viruses," Archives of Virology, Suppl. 2. Springer-Verlag, Vienna.
Fraser, J. R. E., Cunningham, A. L., Hayes, K., Leach, R., and Lunt, R. (1983). *Clin. Exp. Rheumatol.* **1,** 287–293.
Frey, T. K., and Abernathy, E. S. (1993). *J. Infect. Dis.* **168,** 854–864.
Frey, T. K., and Hemphill, M. L. (1988). *Virology* **164,** 22–29.
Frey, T. K., and Marr, L. D. (1988). *Gene* **62,** 85–99.
Frey, T. K., and Strauss, J. H. (1978). *Virology* **86,** 494–506.
Frey, T. K., Marr, L. D., Hemphill, M. L., and Dominguez, G. (1986). *Virology* **154,** 228–232.
Frey, T. K., Marr, L. D., Sanchez, A., and Simmons, R. B. (1989). *Virology* **168,** 191–194.
Friedman, R. M., (1968). *J. Virol.* **2,** 547–552.
Fuller, S. D., and Argos, P. (1987). *EMBO J.* **6,** 1099–1105.
Furukawa, T., Vaheri, A., and Plotkin, S. A. (1967). *Proc. Soc. Exp. Biol. Med.* **125,** 1098–1102.
Gahmberg, N., Kuismanen, E., Keränen, S., and Pettersson, R. F. (1986). *J. Virol.* **57,** 899–906.
Galazka, A. (1991). *Epidemiol. Infect.* **107,** 43–54.
Gerna, G., Franchi, P., and Poletti, L. (1972a). *Boll. Soc. Ital. Biol. Sper.* **48,** 419–423.
Gerna, G., Franchi, P., and Poletti, L. (1972b). *Boll. Soc. Ital. Biol. Sper.* **48,** 822–824.

Gerna, G., Favino, A., and Franchi, P. (1973). *Boll. Soc. Ital. Biol. Sper.* **49,** 362–367.
Gerna, G., Revello, M. G., Dovis, M., Petruzzelli, E., Achilli, G., Percivalle, E., and Torsellini, M. (1987). *J. Gen. Virol.* **68,** 2007–2012.
Ginsberg-Felner, F., Witt, M. E., Yagihashi, S., Dobersen, M. J., Taub, F., Fedun, B., McEvoy, R. C., Roman, S. H., Davies, T. F., Cooper, L. Z., Rubenstein, P., and Notkins, A. L. (1984). *Diabetologia* **27,** 87–89.
Ginsberg-Felner, F., Witt, M. E., Fedun, B., Taub, F., Dobersen, M. J., McEvoy, R. C., Cooper, L. Z., Notkins, A. L., and Rubenstein, P. (1985). *Rev. Infect. Dis.* **7,** S170–S176.
Godec, M. S., Asher, D. M., Murray, R. S., Shin, M. L., Greenham, L. W., Gibbs, C. J., and Gajdusek, D. C. (1992). *Ann. Neurol.* **32,** 401–404.
Godeny, E. K., Speicher, D. W., and Brinton, M. A. (1990). *Virology* **177,** 768–771.
Godeny, E. K., Chen, L., Kumar, S. N., Methven, S. L., Koonin, E. V., and Brinton, M. A. (1993). *Virology* **194,** 585–596.
Goldbach, R. W. (1990). *In* "New Aspects of Positive Strand RNA Viruses" (M. A. Brinton and F. X. Heinz, eds.), pp. 1–12. Am. Soc. Microbiol., Washington, D.C.
Goldbach, R., Le Gall, O., and Wellink, J. (1991). *Semin. Virol.* **2,** 19–25.
Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1988). *FEBS Lett.* **235,** 16–24.
Gorbalenya, A. E., Koonin, E. V., and Lai, M. M.-C. (1991). *FEBS Lett.* **288,** 201–205.
Gould, J. J., and Butler, M. (1980). *J. Gen. Virol.* **49,** 423–426.
Green, K. Y., and Dorsett, P. H. (1986). *J. Virol.* **57,** 893–898.
Gregg, N. M. (1941). *Trans. Ophthalmol. Soc. Aust.* **3,** 35–46.
Hahn, C. S., Dalrymple, J. M., Strauss, J. H., and Rice, C. M. (1987). *Proc. Natl. Acad. Sci. U.S.A.* **84,** 2019–2023.
Halonen, P. E., Ryan, J. M., and Stewart, J. A. (1967). *Proc. Soc. Exp. Biol. Med.* **125,** 162–167.
Hamvas, J. J., Ugovsek, S., Iwakata, S., and Labzoffsky, N. A. (1969). *Arch. Gesamte Virusforsch.* **26,** 287–294.
Hardy, J. B., Sever, J. L., and Gilkeson, M. R. (1969). *J. Pediatr.* **75,** 213–220.
Haseloff, J., Goelet, P., Zimmern, D., Ahlquist, P., Dasgupta, R., and Kaesberg, P. (1984). *Proc. Natl. Acad. Sci. U.S.A.* **81,** 4358–4362.
Hayes, R. J., and Buck, K. W. (1990). *Cell* **63,** 363–368.
Heggie, A. D. (1976). *Teratology* **15,** 47–56.
Heggie, A. D., and Robbins, F. C. (1969). *Am. J. Dis. Child.* **118,** 12–17.
Hemphill, M. L., Forng, R.-Y., Abernathy, E. S., and Frey, T. K. (1988). *Virology* **162,** 65–75.
Herrmann, K. L. (1979). *In* "Diagnostic Procedures for Viral, Rickettsial and Chlamydial Infections," 5th Ed., pp. 725–766. Public Health Serv. U.S. Dep. Health Hum. Serv., Washington, D.C.
Hertz, J. M., and Huang, H. V. (1992). *J. Virol.* **66,** 857–864.
Higashi, N., Arimura, H., and Fujiwara, E. (1969). *Ann. Rep. Inst. Virus Res., Kyoto Univ.* **12,** 100–104.
Hildebrant, H. M., and Maasab, H. F. (1966). *N. Engl. J. Med.* **274,** 1428–1430.
Hobman, T. C., and Gillam, S. (1989). *Virology* **173,** 241–250.
Hobman, T. C., Shukin, R., and Gillam, S. (1988). *J. Virol.* **62,** 4259–4264.
Hobman, T. C., Lundström, M., and Gillam, S. (1990). *Virology* **178,** 122–133.
Hobman, T. C., Qiu, Z., Chaye, H., and Gillam, S. (1991). *Virology* **181,** 768–772.
Hobman, T. C.., Woodward, L., and Farquhar, M. G. (1992). *J. Cell Biol.* **118,** 795–811.
Hobman, T. C., Woodward, L., and Farquhar, M. G. (1993). *J. Cell Biol.* **121,** 269–281.
Holland, J. J. (1990). *In* "Virology" (B. N. Fields *et al.*, eds.), pp. 151–165. Raven, New York.

Holland, J. J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and VandePol, S. (1982). *Science* **215,** 1577–1585.
Holland, J. J., De La Torre, J. C., and Steinhauer, D. A. (1992). *Curr. Top. Microbiol. Immunol.* **176,** 1–20.
Holmes, I. H., and Warburton, M. F. (1967). *Lancet* **ii,** 1233–1236.
Holmes, I. H., Wark, M. C., Jack, I., and Grutzner, J. (1968). *J. Gen. Virol.* **2,** 37–42.
Holmes, I. H., Wark, M. C., and Warburton, M. F. (1969). *Virology* **37,** 15–25.
Honess, R. W. (1984). *J. Gen. Virol.* **65,** 2077–2107.
Honess, R. W., Gomples, U. A., Barrell, B. G., Craxton, M., Cameron, K. R., Staden, R., Chang, Y.-N., and Hayward, G. S. (1989). *J. Gen. Virol.* **70,** 837–855.
Horzinek, M. C. (1981). "Non-Arthropod-Borne Togaviruses." Academic Press, London.
Horzinek, M. C., and Mussgay, M. (1971). *Arch. Gesamte Virusforsch.* **33,** 296–305.
Horzinek, M. C., Maess, J., and Laufs, R. (1971). *Arch. Gesamte Virusforsch.* **33,** 306–318.
Hoskins, J. M., and Plotkin, S. A. (1967). *Arch. Gesamte Virusforsch.* **21,** 296–308.
Ho-Terry, L., and Cohen, A. (1979). *J. Med. Microbiol.* **12,** 441–448.
Ho-Terry, L., and Cohen, A. (1980). *Arch. Virol.* **65,** 1–13.
Ho-Terry, L., and Cohen, A. (1981). *Arch. Virol.* **70,** 199–206.
Ho-Terry, L., and Cohen, A. (1982). *Arch. Virol.* **72,** 47–54.
Ho-Terry, L., and Cohen, A. (1984). *Arch. Virol.* **79,** 139–146.
Ho-Terry, L., and Cohen, A. (1985). *Arch. Virol.* **84,** 207–215.
Ho-Terry, L., Cohen, A., and Londesborough, P. (1982). *J. Med. Microbiol.* **15,** 393–398.
Ho-Terry, L., Terry, G. M., Cohen, A., and Londesborough, P. (1986). *Arch. Virol.* **90,** 145–152.
Hovi, T. (1972). *J. Virol.* **9,** 879–882.
Hovi, T., and Vaheri, A. (1970a). *Virology* **42,** 1–8.
Hovi, T., and Vaheri, A. (1970b). *J. Gen. Virol.* **6,** 77–83.
Howson, C. P., and Fineberg, H. V. (1992). *J. Am. Med. Assoc.* **267,** 392–396.
Hurtley, S. M., and Helenius, A. (1989). *Annu. Rev. Cell Biol.* **5,** 277–307.
Ikemura, T. (1985). *Mol. Biol. Evol.* **2,** 13–34.
Ilonen, J., Seppänen, H., Närvänen, A., Korkolainen, M., and Salmi, A. A. (1992). *Viral Immunol.* **5,** 221–228.
Ivanova, L., and Schlesinger, M. J. (1993). *J. Virol.* **67,** 2546–2551.
Kalkkinen, N., Oker-Blom, C., and Pettersson, R. F. (1984). *J. Gen. Virol.* **65,** 1549–1557.
Kamer, G., and Argos, P. (1984). *Nucleic Acids Res.* **12,** 7269–7282.
Katow, S., and Sugiura, A. (1985). *J. Clin. Microbiol.* **21,** 449–451.
Katow, S., and Sugiura, A. (1988a). *J. Gen. Virol.* **69,** 2797–2807.
Katow, S., and Sugiura, A. (1988b). *Jpn. J. Med. Sci. Biol.* **41,** 109–115.
Kenrick, K. G., Slinn, R. F., Dorman, D. C., and Menser, M. A. (1968). *Lancet* **i,** 548–551.
Kim, K. S., and Boatman, E. S. (1967). *J. Virol.* **1,** 205–214.
Kistler, G. S. (1975). *Beitr. Pathol. Bd.* **155,** 101–138.
Kobayashi, N. (1978). *Virology* **89,** 610–612.
Kobayashi, N., and Suzuki, M. (1982). *Arch. Virol.* **71,** 149–154.
Koonin, E. V. (1991). *J. Gen. Virol.* **72,** 2197–2206.
Koonin, E. V., and Dolja, V. V. (1993). *Crit. Rev. Biochem. Mol. Biol.* **28,** 375–430.
Koonin, E. V., Gorbalenya, A. E., Purdy, M. A., Rozanov, M. N., Reyes, G. R., and Bradley, D. W. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89,** 8259–8263.
Kouri, G. K., Aguilera, A., Rodriguez, P., and Korolev, M. (1974). *J. Gen. Virol.* **22,** 71–80.
Kozak, M. (1987). *Nucleic Acids Res.* **15,** 8125–8148.
Kozak, M. (1989). *J. Cell Biol.* **108,** 229–241.
Kuismanen, E., Bang, B., Hurme, M., and Pettersson, R. F. (1984). *J. Virol.* **51,** 137–146.

Kuo, L., Harty, J. T., Erickson, L., Palmer, G. A., and Plagemann, P. G. W. (1991). *J. Virol.* **65,** 5118–5123.

Lai, M.M.C. (1992). *Curr. Top. Microbiol. Immunol.* **176,** 21–32.

Le Bouvier, G. (1969a). *Nature (London)* **221,** 78–79.

Le Bouvier, G. (1969b). *Proc. Soc. Exp. Biol. Med.* **130,** 51–54.

Levis, R., Schlesinger, S., and Huang, H. V. (1990). *J. Virol.* **64,** 1726–1733.

Lindegren, M. L., Fehrs, L. J., Hadler, S. C., and Hinman, A. R. (1991). *Epidemiol. Rev.* **13,** 341–348.

Locker, J. K., Griffiths, G., Horzinek, M. C., and Rottier, P. J. M. (1992). *J. Biol. Chem.* **267,** 14094–14101.

Lovett, A. E., Hahn, C. S., Rice, C. M., Frey, T. K., and Wolinsky, J. S. (1993). *J. Virol.* **67,** 5849–5858.

Lozzi, L., Rustici, M., Corti, M., Cusi, M. G., Valensin, P. E., Bracci, L., Santucci, A., Soldani, P., Spreafico, A., and Neri, P. (1990). *Arch. Virol.* **110,** 271–276.

Lundström, M. L., Mauracher, C. A., and Tingle, A. J. (1991). *J. Gen. Virol.* **72,** 843–850.

Lustig, S., Jackson, A. C., Hahn, C. S., Griffin, D. E., Strauss, E. G., and Strauss, J. H. (1988). *J. Virol.* **62,** 2329–2336.

Maassab, H. F., and Veronelli, J. A. (1966). *J. Bacteriol.* **91,** 436–441.

Maes, R., Vaheri, A., Sewick, D., and Plotkin, S. (1966). *Nature (London)* **210,** 384–385.

Marr, L. D., Sanchez, A., and Frey, T. K. (1991). *Virology* **180,** 400–405.

Marr, L. D., Wang, C.-Y., and Frey, T. K. (1994). *Virology* **198,** 586–592.

Martenis, T. W., Bland, J. H., and Phillips, C. A. (1968). *Arthritis Rheum.* **11,** 683–687.

Maruyama, T., Gojobori, T., Aota, S., and Ikemura, T. (1986). *Nucleic Acids Res.* **14,** r151–r197.

Mastromarino, P., Rieti, S., Cioe, L., and Orsi, N. (1989). *Arch. Virol.* **107,** 15–26.

Mastromarino, P., Cioe, L., Rieti, S., and Orsi, N. (1990). *Med. Microbiol. Immunol.* **179,** 105–114.

Matsumoto, A., and Higashi, M. (1974). *Ann. Rep. Inst. Virus Res., Kyoto Univ.* **17,** 11–22.

Matthews, R. E. F. (1982). *Intervirology* **17,** 1–199.

Mauracher, C. A., Gillam, S., Shukin, R., and Tingle, A. J. (1991). *Virology* **181,** 773–777.

McAleer, W. J., Markus, H. Z., McLean, A. A., Buynak, E. B., and Hilleman, M. R. (1980). *J. Biol. Stand.* **8,** 281–287.

McCarthy, M., Lovett, A., Kerman, R. H., Overstreet, A., and Wolinsky, J. S. (1993). *J. Virol.* **67,** 673–681.

McCombs, R. M., Brunschwig, J. P., and Rawls, W. E. (1968). *Exp. Mol. Pathol.* **9,** 27–33.

McCormick, J. N., Duthie, J. J. R., Gerber, H., Hart, H., Baker, S., and Marmion, B. P. (1978). *Ann. Rheum. Dis.* **39,** 266–272.

McDonald, H., Hobman, T. C., and Gillam, S. (1991). *Virology* **183,** 52–60.

Menser, M. A., Harley, J. D., Hertzberg, R., Dorman, D. C., Murphy, A. M. (1967). *Lancet* **ii,** 387–388.

Menser, M. A., Forrest, J. M., and Bransby, R. D. (1978). *Lancet* **i,** 57–60.

Meyers, G., Rümenapf, T., and Thiel, H-J. (1989). *Virology* **171,** 555–567.

Mifune, K., Desmyter, J., and Rawls, W. E. (1970). *Infect. Immun.* **2,** 132–138.

Miller, C. L. (1991). *Epidemiol. Infect.* **107,** 63–68.

Miller, E. (1990). *Arch. Dis. Child.* **65,** 820–821.

Miller, R. H., and Purcell, R. H. (1990). *Proc. Natl. Acad. Sci. U.S.A.* **87,** 2057–2061.

Mitchell, L. A., Zhang, T., Ho, M., Decarie, D., Tingle, A. J., Zrein, M., and Lacroix, M. (1992). *J. Clin. Microbiol.* **30,** 1841–1847.

Mitchell, L. A., Tingle, A. J., Shukin, R., Sangeorzan, J. A., McCune, J., and Braun, D. K. (1993). *Arch. Intern. Med.* **153,** 2268–2274.

Monif, G.R.G., Avery, G. B., Korones, S. B., and Sever, J. L. (1965). *Lancet* **i,** 723–724.
Moss, B., Elroy-Stein, O., Mizukami, T., Alexander, W. A., and Fuerst, T. R. (1990). *Nature (London)* **348,** 91–92.
Murphy, F. A. (1980). *In* "The Togaviruses" (R. W. Schlesinger, ed.), pp. 241–316. Academic Press, New York.
Murphy, F. A., Halonen, P. E., and Harrison, A. K. (1968). *J. Virol.* **2,** 1223–1227.
Nakhasi, H. L., Meyer, B. C., and Liu, T.-Y. (1986). *J. Biol. Chem.* **261,** 16616–16621.
Nakhasi, H. L., Zheng, D., Hewlett, I. K., and Liu, T.-Y. (1988). *Virus Res.* **10,** 1–15.
Nakhasi, H. L., Thomas, D., Zheng, D., and Liu, T.-Y. (1989). *Nucleic Acids Res.* **17,** 4393–4394.
Nakhasi, H. L., Rouault, T. A., Haile, D. J., Liu, T.-Y., and Klausner, R. D. (1990). *New Biol.* **2,** 255–264.
Nakhasi, H. L., Cao, X-Q., Rouault, T. A., and Liu, T.-Y. (1991). *J. Virol.* **65,** 5961–5967.
Nakhasi, H. L., Singh, N. K., Pogue, G. P., Cao, X.-Q., and Rouault, T. A. (1994). *In* "Positive Strand RNA Viruses" (M. A, Brinton, C. H. Calisher, and R. Rueckert, eds.), *Arch Virol. Suppl.* **9,** pp. 255–267. Springer-Verlag, Vienna.
Nath, A., and Wolinsky, J. S. (1990). *Ann. Neurol.* **27,** 533–536.
Nath, A., Slagle, B., and Wolinsky, J. S. (1989). *Arch. Virol.* **107,** 159–167.
Neri, P., Corti, M., Lozzi, L., and Valensin, P. E. (1991). *Biopolymers* **31,** 631–635.
Newcombe, J., Best, J. M., and Sanders, P. G. (1993). *Abstr., Int. Cong. Virol., 9th* p. 129.
Nichol, S. T., Rowe, J. E., and Fitch, W. M. (1989). *Virology* **168,** 281–291.
Niesters, H. G. M., and Strauss, J. H. (1990a). *J. Virol.* **64,** 1639–1647.
Niesters, H. G. M., and Strauss, J. H. (1990b). *J. Virol.* **64,** 4162–4168.
Nomoto, A., Omata, T., Toyoda, H., Kuge, S., Horie, H., Kataoka, Y., Genba, Y. Nakano, Y., and Imura, N. (1982). *Proc. Natl. Acad. Sci. U.S.A.* **79,** 5793–5797.
Norval, M. (1979). *J. Gen. Virol.* **43,** 289–298.
Nowak, T., Farber, P. M., Wengler, G., and Wengler, G. (1989). *Virology* **169,** 365–376.
Numazaki, K., Goldman, H., Wong, I., and Wainberg, M. A. (1989). *Am. J. Clin. Pathol.* **91,** 446–451.
O'Brien, J. (1989). Ph.D. Thesis, Univ. of Tennessee Med. Sch., Memphis.
Ogra, P. L., and Herd, J. K. (1971). *J. Immunol.* **107,** 810–813.
Ogra, P. L., Kerr-Grant, D., Umana, G., Dzeirba, J., and Weintraub, D. (1971). *N. Engl. J. Med.* **285,** 1333–1339.
Ogra, P. L., Ogra, S. S., Chiba, Y., Dzierba, J. L., and Herd, J. K. (1975). *Lancet* **1,** 1157–1161.
Oker-Blom, C. (1984a). *J. Virol.* **51,** 354–358.
Oker-Blom, C. (1984b). Ph.D. Thesis, Univ. of Helsinki.
Oker-Blom, C., Kalkkinen, N., Kääriäinen, L., and Pettersson, R. F. (1983). *J. Virol.* **46,** 964–973.
Oker-Blom, C., Ulmanen, I., Kääriäinen, L., and Pettersson, L. (1984). *J. Virol.* **49,** 403–408.
Oker-Blom, C., Pettersson, R. F., and Summers, M. D. (1989). *Virology* **172,** 82–91.
Oker-Blom, C., Jarvis, D. L., and Summers, M. D. (1990). *J. Gen. Virol.* **71,** 3047–3053.
Oldstone, M. B. A. (1989). *Cell* **56,** 517–520.
Oshiro, L. S., Schmidt, N. J., and Lennette, E. H. (1969). *J. Gen. Virol.* **5,** 205–210.
Ou, D., Chong, P., Tripet, B., and Gillam, S. (1992a). *J. Virol.* **66,** 1674–1681.
Ou, D., Chong, P., McVeish, P., Jeffries, W. A., and Gillam, S. (1992b). *Virology* **191,** 680–686.
Ou, D., Chong, P., McVeish, P., Jeffries, W. A., Koloitis, G., Tingle, A. J., and Gillam, S. (1992c). *J. Virol.* **66,** 6788–6793.
Ou, D., Mitchell, L. A., Tingle, A. J., Nepom, G. T., Zrein, M., LaCroix, M., *et al.* (1993). *Abstr., Int. Congr. Virol., 9th* p. 129.

Paredes, A. M., Simon, M. N., and Brown, D. T. (1992). *Virology* **187,** 329–332.
Paredes, A. M., Brown, D. T., Rothnagel, R., Chiu, W., Schoepp, R. J., Johnston, R. E., and Prasad, B. V. V. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90,** 9095–9099.
Parkman, P. D., Buescher, R. L., and Artenstein, M. S. (1962). *Proc. Soc. Exp. Biol. Med.* **111,** 225–230.
Partanen, P., Seppänen, H., Suni, J., and Vaheri, A. (1985). *J. Clin. Microbiol.* **21,** 800–802.
Patrizi, G., and Middelkamp, J. N. (1970). *J. Ultrastruct, Res.* **31,** 407–423.
Payment, P., Ajdukovic, D., and Pavilanis, V. (1975a). *Can. J. Microbiol.* **21,** 703–709.
Payment, P., Ajdukovic, D., and Pavilanis, V. (1975b). *Can. J. Microbiol.* **21,** 710–717.
Payment, P., Ajdukovic, D., and Pavilanis, V. (1975c). *Can. J. Microbiol.* **21,** 718–722.
Pettersson, R. F. (1991). *Curr. Top. Microbiol. Immunol.* **170,** 67–106.
Pettersson, R. F., Oker-Blom, C., Kalkkinen, N., Kallio, A., Ulmanen, I., Kääriäinen, L., Partanen, P., and Vaheri, A. (1985). *J. Infect. Dis.* **7,** S140–S149.
Phillips, P. E. (1989). *In* "Arthritis" (D. J. McCarty, ed.), 11th Ed., pp. 482–504. Lea & Febiger, Philadelphia.
Phillips, P. E., Fraser, A., Shore, M., and Dougherty, R. (1993). *Arthritis Rheum.* **31,** R41. (Abstr.)
Pleij, C. W. A., and Bosch, L. (1989). *In* "RNA Processing, Part A: General Methods" (J. Dahlberg and J. Abelson, eds.), Methods in Enzymology, Vol. 180, pp. 289–303. Academic Press, San Diego.
Plotkin, S. A. (1988). *In* "Vaccines" (S. A. Plotkin and E. A. Mortimer, eds.) pp. 235–262. Saunders, Philadelphia.
Plotkin, S. A., and Vaheri, A. (1967). *Science* **156,** 659–661.
Plotkin, S. A., Boue, A., and Boue, J. G. (1965). *Am. J. Epidemiol.* **81,** 71–85.
Plotkin, S. A., Daum, R. S., Giebink, G. S., Hall, C. D., Lepow, M., Marcuse, E. K., McCracken, G. H., Jr., Phillips, C. F., Scott, G. B., and Wright, H. T., Jr. (1989). *Am. Acad. Pediatr. News* July, p. 6.
Pogue, G. P., Cao, W.-Q., Singh, N. K., and Nakhasi, H. L. (1993). *J. Virol.* **67,** 7106–7117.
Qiu, Z., Hobman, T. C., McDonald, H. L., Seto, N. O. L., and Gillam, S. (1992a). *J. Virol.* **66,** 3514–3521.
Qiu, Z., Tufaro, F., and Gillam, S. (1992b). *Virology* **190,** 876–881.
Rawls, W. E. (1974). *Prog. Med. Virol.* **18,** 273–288.
Rawls, W. E., and Melnick, J. L. (1966). *J. Exp. Med.* **123,** 795–816.
Rayfield, E. J., Kelly, K. J., and Yoon, J. W. (1986). *Diabetes* **35,** 1278–1281.
Rhim, J. S., and Schell, K. (1967). *Proc. Soc. Exp. Biol. Med.* **125,** 602–606.
Rice, C. M., and Strauss, J. H. (1982). *J. Mol. Biol.* **154,** 325–348.
Rice, C. M., Lenches, E. M., Eddy, R. R., Shin, S. J., Sheets, R. L., and Strauss, J. H. (1985). *Science* **229,** 726–733.
Rico-Hesse, R., Pallansch, M. A., Nottay, B. K., and Kew, O. M. (1987). *Virology* **160,** 311–322.
Rose, J. K., Adams, G. A., and Gallione, C. J. (1984). *Proc. Natl. Acad. Sci. U.S.A.* **81,** 2050–2054.
Rota, P. A., Bloom, A. E., Vanchieve, J. A., and Bellini, W. J. (1994). *Virology* **198,** 724–730.
Rozanov, M. N., Koonin, E. V., and Gorbalenya, A. E. (1992). *J. Gen. Virol.* **73,** 2129–2134.
Rubenstein, P., Walker, M. E., Fedun, B., Witt, M. E., Cooper, L. Z., and Ginsberg-Fellner, F. (1982). *Diabetes* **31,** 1088–1091.
Rümenapf, T., Unger, G., Strauss, J. H., and Thiel, H.-J. (1993). *J. Virol.* **67,** 3288–3294.
Russell, B., Selzer, G., and Goeze, H. (1967). *J. Gen. Virol.* **1,** 305–310.

Russell, D. L., Dalrymple, J. M., and Johnston, R. E. (1989). *J. Virol.* **63,** 1619–1629.
Salmi, A. A. (1972a). *Acta Pathol. Microbiol. Scand.. Sect. B* **80,** 534–544.
Salmi, A. A. (1972b). *Acta Pathol. Microbiol. Scand., Sect. B* **80,** 545–558.
Salonen, E.-M., Hovi, T., Meurman, O., Vesikari, T., and Vaheri, A. (1985). *J. Med. Virol.* **16,** 1–9.
Sanchez, A., and Frey, T. K. (1991). *Virology* **183,** 636–646.
Sarnesto, A., Ranta, S., Väänänen, P., and Mäkelä, O. (1985). *Scand. J. Immunol.* **21,** 275–282.
Sato, M., Yamada, T., Yamamoto, K., and Yamamoto, N. (1976). *Virology* **69,** 691–699.
Sato, M., Urade, M., Maeda, N., Miyazaki, T., Watanabe, M., Shibata, T., and Yamamoto, N. (1978). *Arch. Virol.* **56,** 89–103.
Schaffner, W., Fleet, W. F., Kilroy, A. W., Lefkowitz, L. B., Herrmann, K. L., Thompson, J., and Karzon, D. T. (1974). *Am. J. Dis. Child.* **127,** 684–688.
Schlesinger, S., and Schlesinger, M. J. (1990). *In* "Virology" (B. N. Fields, D. M. Knipe, *et al.*, eds.), 2nd Ed., pp. 121–148. Raven, New York.
Schmidt, M., Schmidt, M. F. G., and Rott, R. (1988). *J. Biol. Chem.* **263,** 18635–18639.
Schmidt, N. J., and Lennette, E. H. (1969). *Am. J. Dis. Child.* **118,** 89–93.
Schmidt, N. J., Dennis, J., and Lennette, E. H. (1971). *Appl. Microbiol.* **22,** 469–470.
Schultz, A. M., Henderson, L. E., and Oroszlan, S. (1988). *Annu. Rev. Cell Biol.* **4,** 611–647.
Sedman, S. A., Gelembiuk, G. W., and Mertz, J. E. (1990). *J. Virol.* **64,** 453–457.
Sedwick, W. D., and Sokol, F. (1970). *J. Virol.* **5,** 478–489.
Segal, S., and Sreevalsan, T. (1974). *Virology* **59,** 428–442.
Sever, J. L., South, M. A., and Shaver, K. A. (1985). *Rev. Infect. Dis.* **7,** S164–S169.
Shirako, Y., Niklasson, B., Dalrymple, J. M., Strauss, E. G., and Strauss, J. H. (1991). *Virology* **182,** 753–764.
Söderlund, H., von Bonsdorff, C.-H., and Ulmanen, I. (1979). *J. Gen. Virol.* **45,** 15–26.
Sontheimer, R. D., Lieu, T. S., and Capra, J. D. (1993). *Immunologist* **1,** 155–160.
Spruance, S. L., Metcalf, R., Smith, C. B., Griffiths, M. M., and Ward, J. R. (1977). *Arthritis Rheum.* **20,** 741–747.
Stanway, G., Hughes, P. J., Mountford, R. C., Reeve, P., and Minor, P. D. (1984). *Proc. Natl. Acad. Sci. U.S.A.* **81,** 1539–1543.
Stanwick, T. L., and Hallum, J. V. (1974). *Infect. Immun.* **10,** 810–815.
Stewart, G. L., Parkman, P. D., Hopps, H. E., Douglas, R. D., Hamilton, J. P., and Meyer, H. M., Jr. (1967). *N. Engl. J. Med.* **276,** 554–557.
Stokes, A., Mims, C. A., and Grahame, R. (1986). *J. Med. Microbiol.* **21,** 283–285.
Strauss, E. G., and Strauss, J. H. (1986). *In* "The Togaviridae and the Flaviviridae" (S. Schlesinger and M. J. Schlesinger, eds.), pp. 35–90. Plenum, New York.
Strauss, E. G., and Strauss, J. H. (1990). *Semin. Virol.* **1,** 347–356.
Strauss, E. G., Strauss, J. H., and Levine, A. J. (1990). *In* "Virology" (B. N. Fields, D. M. Knipe, R. M. Channo, eds.), 2nd Ed., pp. 167–190. Raven, New York.
Sugiura, A., and Yamada, A. (1991). *Pediatr. Infect. Dis. J.* **10,** 209–213.
Suomalainen, M., Garoff, H., and Baron, M. D. (1990). *J. Virol.* **64,** 5500–5509.
Takkinen, K., Vidgren, G., Ekstrand, J., Hellman, U., Kalkkinen, N., Wernstedt, C., and Pettersson, R. F. (1988). *J. Gen. Virol.* **69,** 603–612.
Taylor, M. J., Godfrey, E., Baczko, K., ter Meulen, V., Wild, T. F., and Rima, B. K. (1991). *J. Gen. Virol.* **72,** 83–88.
Taylor-Robinson, C. H., McCarthy, K., Grylls, S. G., and O'Ryan, E. M. (1964). *Lancet* No. 286, 1364–1365.
Tedder, R. S., Yao, J. L., and Anderson, J. J. (1982). *J. Hyg.* **88,** 335–350.
Terry, G. M., Ho-Terry, L., Cohen, A., and Londesborough, P. (1985). *Arch. Virol.* **86,** 29–36.

Terry, G. M., Ho-Terry, L., Londesborough, P., and Rees, K. R. (1988). *Arch. Virol.* **98,** 189–197.

Tesh, R. B., and Rosen, L. (1975). *Intervirology* **5,** 216–219.

Thompson, G. R., Weiss, J. J., Eloise, M. I., Shillis, J. L., and Brackett, R. G. (1973). *Am. J. Dis. Child.* **125,** 526–530.

Thomssen, R., Laufs, R., and Müller, J. (1968). *Arch. Gesamte Virusforsch.* **23,** 332–345.

Tingle, A. J., Yang, T., Allen, M., Kettyls, G. D., Larke, R. P. B., and Schulzer, M. (1983). *Infect. Immun.* **40,** 22–28.

Tingle, A. J., Chantler, J. K., Kettyls, G. D., Larke, R. P. B., and Schulzer, M. (1985a). *J. Infect. Dis.* **151,** 330–336.

Tingle, A. J., Chantler, J. K., Pot, K. H., Paty, D. W., and Ford, D. K. (1985b). *J. Infect. Dis.* **152,** 606–612.

Tingle, A. J., Allen, M., Petty, R. E., Kettyls, G. D., and Chantler, J. K. (1986). *Ann. Rheum. Dis.* **45,** 110–114.

Tingle, A. J., Pot, K. H., Puterman, M. L., and Hancock, E. J. (1989). *Clin. Immunol. Immunopathol.* **53,** S99–S106.

Tinoco, I., Borer, P. N., Dengler, B., Levine, M. C., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1973). *Nature (London)* **246,** 40–41.

Toivonen, V., Vainionpää, R., Salmi, A., and Hyypiä, T. (1983). *Arch. Virol.* **77,** 91–95.

Tooze, J., Tooze, S. A., and Warren, G. (1984). *Eur. J. Biochem.* **33,** 281–293.

Tooze, J., Tooze, S. A., and Warren, G. (1985). *Eur. J. Cell Biol.* **36,** 108–115.

Tooze, S. A., Tooze, J., and Warren, G. (1988). *J. Cell Biol.* **106,** 1475–1487.

Trudel, M., Ravaoarinoro, M., and Payment, P. (1980). *Can. J. Microbiol.* **26,** 899–904.

Trudel, M., Nadon, F., Comtois, R., Payment, P., Bonneau, A., and Lecomte, J. (1982). *J. Virol. Methods* **5,** 191–197.

Tuchinda, P., Nii, S., Sasada, T., Naito, T., Ono, N., and Chatiyanon, K. (1969). *Biken J.* **12,** 201–219.

Ueda, K., Nishida, Y., Oshima, K., Yoshikawa, H., Ohashi, K., and Nonaka, S. (1975). *Jpn. J. Microbiol.* **19,** 181–185.

Umino, Y., Sato, T. A., Katow, S., Matsuno, T., and Sugiura, A. (1985). *Arch. Virol.* **83,** 33–42.

Urade, M., and Yamamoto, N. (1990). *Res. Virol.* **141,** 285–303.

Urade, M., Mima, T., and Yamamoto, N. (1993). *Res. Virol.* **144,** 129–139.

Väänänen, P., and Kääriäinen, L. (1980). *J. Gen. Virol.* **46,** 467–475.

Vaheri, A., and Cristofalo, V. J. (1967). *Arch. Gesamte Virusforsch.* **21,** 425–436.

Vaheri, A., and Hovi, T. (1972). *J. Virol.* **9,** 10–16.

Vaheri, A., and Vesikari, T. (1971). *Arch. Gesamte Virusforch.* **35,** 10–24.

Vaheri, A., Sedwick, W. D., Plotkin, S. A., and Maes, R. (1965). *Virology* **27,** 239–241.

Vaheri, A., Sewick, W. D., and Plotkin, S. A. (1967). *Proc. Soc. Exp. Biol. Med.* **125,** 1086–1092.

van Wyke Coelingh, K. L., Winter, C. C., and Murphy, B. R. (1988). *Virology* **162,** 137–143.

Veale, H. (1866). *Edinburgh Med. J.* **12,** 404–414.

Vidgren, G., Takkinen, K., Kalkkinen, N., Kääriäinen, L., and Pettersson, R. F. (1987). *J. Gen. Virol.* **68,** 2347–2357.

Vogel, R. H., Provencher, S. W., von Bonsdorff, C.-H., Adrian, M., and Dubochet, J. (1986). *Nature (London)* **320,** 533–535.

Voiland, A., and Bardeletti, G. (1980). *Arch. Virol.* **64,** 319–328.

von Bonsdorff, C.-H., and Vaheri, A. (1969). *J. Gen. Virol.* **5,** 47–51.

von Heijne, G. (1984). *J. Mol. Biol.* **173,** 243–251.

Wahlberg, J. M., Boere, W. A., and Garoff, H. (1989). *J. Virol.* **63,** 4991–4997.

Wang, C.-Y., Dominguez, G., and Frey, T. K. (1994). *J. Virol.* **68.** In Press.

Ward, C. W., and Shukla, D. D. (1993). *Abstr., Int. Congr. Virol., 9th* p. 116.
Waxham, M. N., and Wolinsky, J. S. (1983). *Virology* **126,** 194–203.
Waxham, M. N., and Wolinsky, J. S. (1985a). *Virology* **143,** 153–165.
Waxham, M. N., and Wolinsky, J. S. (1985b). *Rev. Infect. Dis.* **7,** S133–S139.
Weaver, S. C., Scott, T. W., and Rico-Hesse, R. (1991). *Virology* **182,** 774–784.
Weaver, S. C., Hagenbaugh, A., Bellew, L. A., Netesov, S. V., Volchkov, V. E., Chang, G.-J. J., Clarke, D. K., Gousset, L., Scott, T. W., Trent, D. T., and Holland, J. J. (1993). *Virology* **197,** 375–390.
Weibel, R. E., Stokes, J., Buynak, E. B., and Hillemann, M. R. (1969). *Am. J. Dis. Child.* **118,** 226–229.
Weller, T. H., and Neva, F. A. (1962). *Proc. Soc. Exp. Biol. Med.* **111,** 215–225.
Wengler, G., Wengler, G., and Wahn, K. (1976). *Arch. Virol.* **50,** 45–53.
Westaway, E. G., Brinton, M. A., Gaidamovich, S. Y., Horzinek, M. C., Igarashi, A., Kääriäinen, L., Lvov, D. K., Porterfield, J. S., Russell, P. K., and Trent, D. W. (1985a). *Intervirology* **24,** 125–139.
Westaway, E. G., Brinton, M. A., Gaidamovich, S. Y., Horzinek, M. C., Igarashi, A., Kääriäinen, L., Lvov, D. K., Porterfield, J. S., Russell, P. K., and Trent, D. W. (1985b). *Intervirology* **24,** 183–192.
Wiley, D. C. (1986). *In* "Fundamental Virology" (B. N. Fields and D. M. Knipe, eds.) pp. 45–68. Raven, New York.
Will, C., Mühlberger, E., Linder, D., Slenczka, W., Klenk, H.-D., and Feldmann, H. (1993). *J. Virol.* **67,** 1203–1210.
Wolinsky, J. S. (1985). *In* "Handbook of Clinical Neurology. Vol. 34: Infections of the Central Nervous System" (P. J. Vinker and C. W. Bruyn, eds.), pp. 331–342. North-Holland Publ., Amsterdam.
Wolinsky, J. S. (1990a). *In* "Virology" (B. N. Fields, D. M. Knipe *et al.*, eds.), 2nd Ed., pp. 815–838. Raven, New York.
Wolinsky, J. S. (1990b). *In* "Immunologic Mechanisms in Neurologic and Psychiatric Disease" (B. H. Waksman, ed.), pp. 259–268. Raven, New York.
Wolinsky, J. S., McCarthy, M., Allen-Cannady, O., Moore, W. T., Jin, R., Cao, S.-N., Lovett, A., and Simmons, D. (1991). *J. Virol.* **65,** 3986–3994.
Wolinsky, J. S., Sukholutsky, E., Moore, W. T., Lovett, A., McCarthy, M., and Adame, B. (1993). *J. Virol.* **67,** 961–968.
Wong, K. T., Baron, S., and Ward, T. G. (1967). *J. Immunol.* **99,** 1140–1149.
Wong, K. T., Robinson, W. S., and Merigan, T. C. (1969). *J. Virol.* **4,** 901–903.
Wong-Staal, F. (1990). *In* "Virology" (B. N. Fields, D. M. Knipe *et al.*, eds.), 2nd Ed., pp. 1529–1543. Raven, New York.
Woods, W. A., and Robbins, F. C. (1968). *J. Gen. Virol.* **3,** 43–49.
Yamada, A., Takeuchi, K., Tanabayashi, K., Hishiyama, M., and Suguira, A. (1989). *Virology* **172,** 374–376.
Yamamoto, N., and Urade, M. (1989). *Photochem. Photobiol.* **50,** 351–358.
Yoneda, T., Urade, M., Sakuda, M., and Miyazaki, T. (1986). *J. Clin. Invest.* **77,** 1613–1621.
Yoshinaka, Y., and Hotta, S. (1971). *Virology* **45,** 524–526.
Zheng, D., Dickens, L., Liu, T.-Y., and Nakhasi, H. L. (1989). *Gene* **82,** 343–349.
Ziring, P. R., Gallo, G., Finegold, M., Buimovici-Klein, E., and Ogra, P. (1977). *J. Pediatr.* **90,** 419–420.