Assessment of determinants affecting the dual topology of hepadnaviral large envelope proteins

Carsten Lambert, Sylvia Mann and Reinhild Prange

Department of Medical Microbiology and Hygiene, University of Mainz, Augustusplatz, D-55101 Mainz, Germany

For functional diversity, the large (L) envelope protein of hepatitis B virus (HBV) acquires a dual transmembrane topology via co-translational membrane integration of the S region and partial post-translational translocation of the preS subdomain. Because each process requires the second transmembrane segment (TM2), we explored the action of this determinant by using protease protection analysis of mutant L proteins. We demonstrated that neither the disruption of a leucine zipper-like motif by multiple alanine substitutions nor the flanking charges of TM2 affected the topological reorientation of L. The dispensability of both putative subunit interaction modules argues against a link between preS post-translocation and envelope assembly. Phenotypic mixing experiments revealed that the preS and S protein domains of the related duck HBV L polypeptide failed to substitute functionally for the topogenic elements of HBV in directing the correct L topography, implicating different translocation mechanisms used by the two hepadnavirus genera.

Owing to the small size of their genome, viruses have evolved by retaining a maximum amount of information in a minimum of polypeptide sequences. As a consequence, many of the proteins or protein domains encoded by viruses are multifunctional. The large (L) envelope protein of hepatitis B virus (HBV), the prototype member of the Hepadnaviridae family, is an example of such multifunctionality. This protein serves in virus entry as a receptor protein and in virus biogenesis, the L protein, together with the related middle (M) and small (S) envelope proteins, are expressed from a single open reading frame by differential translation initiation. As a result, the sequence of S is repeated at the C termini of M and L, which contain the additional preS2 domain or preS2 and preS1 domains, respectively (reviewed by Heermann & Gerlich, 1991). All three proteins are post-translationally integrated into the endoplasmic reticulum (ER) membrane, most likely directed by a signal-anchor and stop-transfer sequence encoded within the first and second transmembrane (TM) segments (TM1 and TM2) of their S domains (Eble et al., 1987). These signals also direct co-translational translocation of the upstream preS2 region of M into the ER lumen (Eble et al., 1990). In contrast, the preS2 plus preS1 (preS) domain of L initially remains cytosolic. During maturation, about half of the L molecules post-translationally translocate their preS region into the ER, thereby generating a dual topology that is maintained in the virion envelope (Bruss et al., 1994) (Fig. 1). By orientating the preS domain at both the cytosolic (inside the virus) and luminal (outside the virus) location, L serves its topological opposing functions in the virus life cycle, capsid envelopment and receptor binding, respectively.

We previously demonstrated that the partial preS post-translocation is seemingly not established by an HBV-specific channel generated from lateral interactions between the TM regions of the envelope proteins during virus assembly, as this process requires neither S and M proteins nor any of the amphipathic TM segments implicated in channel formation (Lambert & Prange, 2001). Rather, we found that the topological reorientation of L is physically linked to the ER membrane, suggesting the involvement of a host-cell transmembrane transport machinery. These studies also led to identification of TM2 as a critical determinant of preS post-translocation. A second topogenic element was mapped to a preS1-specific subdomain of L, termed the cytosolic anchorage determinant (CAD), which interacts with the cognate heat-shock protein Hsc70, thereby preventing co-translational preS translocation (Löffler-Mary et al., 1997; Lambert & Prange, 2003). Both topogenic determinants
have been shown to act in cis, i.e. within the same L molecule. This is in contrast to the L homologue of the phylogenetically distant duck HBV (DHBV). This protein adopts the same dual topology but engages preS-specific determinants in cis, such as a cluster of lysine residues, together with Hsc70-binding elements and S-specific determinants in trans, like the TM1 domain, with the latter provided by co-expressed S chains (Swameye & Schaller, 1997; Grgacic, 2002). As a further difference, the topological switch of DHBV L has been shown to coincide with envelope particle morphogenesis and export (Grgacic, 2002). Together, these observations implicate that the folding pathways of the two hepadnaviral L proteins appear to diverge on a common final phenotype. Addressing this issue, here we focused on the topogenic determinants and investigated their molecular nature, function and contribution to the reorientation of the HBV L protein.

A series of mutant and chimeric proteins was constructed and their topological features were analysed by transient expression in COS-7 cells. As a reliable assay, trypsin protection experiments of microsomes were used to monitor the partial preS post-translocation of L across membranes (Lambert & Prange, 2001). Upon expression of an HA-tagged version of the L gene, driven by plasmid pN12.LHA, L appeared in its characteristic doublet of a 39 kDa non-glycosylated (p39) and a 42 kDa single-glycosylated (gp42) form as a consequence of partial N-glycosylation at Asn-309 in its S domain. While the preS domain of newly synthesized L chains is almost fully sensitive to cleavage with trypsin, over time it becomes increasingly protected due to its partial post-translocation into the ER, yielding up to 50–60 % resistant chains at steady state, as summarized in Fig. 1. Trypsin treatment in the absence of detergent (NP-40) yielded two fractions of L: trypsin-resistant full-length molecules with preS post-translationally translocated into the ER and trypsin-sensitive chains with preS orientated to the cytosol where they are cleaved to a 25 kDa non-glycosylated (T) and a 28 kDa single-glycosylated (gT) fragment, corresponding to the C-terminal S parts. Upon disruption of microsomes with NP-40, trypsin completely converted L to these fragments. The post-translational mode of preS translocation has been further evidenced by the absence of N-linked glycosylation of preS carrying two modification-competent glycosylation sites (Asn-4 and Asn-123) (Löffler-Mary et al., 1997).

Previously, we showed that among the four membrane-spanning segments of L, only TM2 is needed for preS post-translocation (Lambert & Prange, 2001). While these results largely excluded an HBV-specific envelope structure serving as a preS conduit, there still remained the possibility that lateral interactions of TM2 domains by helical packing might form an autonomous translocation channel, as, for example, exemplified by the M2 transmembrane proton channel protein of influenza A (Sansom et al., 1998). In favour of this assumption, database searching identified a leucine-rich heptad motif within TM2 (Fig. 2A), a candidate self-interaction motif (Gurecka et al., 1999; Scholze et al., 2002), which might form a pore structure. To assess whether the leucine zipper-type motif comprised within Leu-247, Leu-254 and Leu-261 might contribute to preS reorientation, double and triple substitutions by alanine were made by

Fig. 1. Domain structure and transmembrane topology of the HBV L envelope protein. (A) Schematic representation of L consisting of the preS1, preS2 and S domains. Numbers above the domains refer to the corresponding amino acid positions. The four transmembrane TM segments are indicated by open boxes, with the first two designated 1 and 2; partial N-glycosylation occurring at Asn-309 is indicated by ¥. (B) Model of the mixed topology of L at the ER membrane. Upon co-translational membrane insertion, the preS domain of L is initially located on the cytosolic surface of the ER (left). During maturation, ~50 % of the L molecules post-translationally translocate preS into the ER lumen, thereby leading to the dual topology (right). (C) Trypsin protection assay of HA-tagged L protein synthesized in transiently transfected COS-7 cells. Two days after transfection, microsomes were prepared and either left untreated or digested with trypsin in the absence (−) or presence (+) of NP-40. Samples were analysed by HA-specific immunoblotting. The non-glycosylated (p39) and single-glycosylated (gp42) forms of L and its non-glycosylated (T) and single-glycosylated (gT) tryptic fragments are indicated. Of note, this assay was performed within the so-called Lo-only-background in which synthesis of the S and M proteins is prevented by missense mutations of their initiator codons. The topological features of L o are virtually identical to wt L (Lambert & Prange, 2001).
PCR-directed mutagenesis. Upon synthesis in COS-7 cells, each of the L247/254A, L247/261A, L254/261A and L247/254/261A mutants were obtained in a doublet that is characteristic for the non- and single-glycosylated chains of wild-type (wt) L and hence for co-translational insertion into the ER membrane (Fig. 2A). The analysis of their topological properties, however, revealed that all double mutants displayed the wt translocational phenotype, occurring in a fraction of ~40–50% of polypeptides that were protected from trypsin by the microsomal membrane unless NP-40 was present. The triple mutant supported preS post-translocation even to a greater extent. Accordingly, the leucine zipper-like motif is not required for preS post-translocation but appears to affect its degree.

Another prominent feature of TM2 is the asymmetric distribution of adjacent charges contained within three preceding arginines and one succeeding aspartic acid. Consistent with the well-established function of polar residues flanking membrane-spanning domains (Boyd & Beckwith, 1990), these charges are likely to be important for the stop-transfer function of TM2. In addition, however, their polar side-chains might be also involved in interhelical hydrogen bonding and hence in transmembrane helix interactions (reviewed by Ubarretxena-Belandia & Engelman, 2001), thereby partnering L chains to create a protein–protein interface, which could allow preS post-translocation. To examine the impacts of Arg-236, Arg-241, Arg-242 and Asp-262 on L refolding, they were individually mutated to uncharged residues, such as alanine or leucine. Protease protection analysis demonstrated that R236L, R242A and D262A formed the correct L topology, as shown by their partial resistance against trypsin in intact microsomes and their lack of preS-linked N-glycans (Fig. 2B). Similarly, R241A post-translationally translocated preS but to a higher degree. Possibly, this substitution might lead to a structurally imposed loss of translocational control, thereby altering the equilibrium of luminal- and cytosolic-oriented preS domains. Together, these data indicated that the putative transmembrane helix–helix packing motifs of TM2, the leucine-rich heptad repeat and the polar residues are dispensable for the post-translational reorientation of L, thereby further implicating that L refolding is not coupled to envelope assembly.

To gain additional insight into the action of the topogenic signals of L, we performed domain-exchange experiments by shuffling mammalian and avian HBV envelope subdomains. First, a chimera was generated by substituting the S region of L with the corresponding domain of DHBV. For construction of preS(h) : : S(d), the DHBV S region was amplified from plasmid pDHBV (Wildner et al., 1991) prior to cloning into pNI2.LHA. In accordance with the calculated molecular mass, preS(h) : : S(d) was expressed as a membrane-bound 37 kDa polypeptide (Fig. 3A). Surprisingly, two higher molecular mass forms were found in addition. These bands disappeared on deglycosylation with peptide- N-glycosidase F (PNGase F), indicating that they represented N-glycosylated versions. Because the DHBV S domain

---

**Fig. 2.** Role for TM2 in partial post-translational preS translocation of L. (A) A heptad motif of leucine residues in TM2 is dispensable for preS post-translocation. The amino acid sequence of TM2 is shown with the leucine residues at each seventh position of the potential zipper in bold and their amino acid positions indicated. COS-7 cells were transfected with HA-tagged versions of L mutants carrying double or triple substitutions of the designated leucine residues with alanine. Trypsin protection experiments were done as described in the legend to Fig. 1(C). The positions of non-glycosylated (p39) and single-glycosylated (gp42) mutant L polypeptides are indicated. (B) Charged residues flanking TM2 are dispensable for preS post-translocation. The polar R-236, R-241, R-242 and D-262 residues at the borders of TM2 are shown in bold on the left. Analysis of the topology of the mutants designated by their substitutions using trypsin is shown on the right. Their p39/gp42 forms are indicated.
is known to lack N-glycans, the modification of preS(h)::S(d) must have occurred at Asn-4 and Asn-123 within preS. Given that both glycosylation sites of preS(h)::S(d) were on the luminal side of the ER membrane, this observation implicated an unexpected co-translational mode of preS translocation. Consistent with this, the majority of preS(h)::S(d) glycosylated chains (~95%) were protected from trypsin in intact microsomes. Non-glycosylated p37, however, was fully accessible to trypsin. This might be due to inefficient co-translational preS import in the ER and/or retrograde preS dislocation out of the ER.

A similar outcome but different translocational phenotype was obtained with a reciprocal chimera consisting of the DHBV preS domain fused to the HBV S domain. On synthesis, preS(d)::S(h) appeared as a non- and single-glycosylated species, as confirmed by digestion with PNGase F (Fig. 3B). The N-glycosylation, most likely occurring at Asn-309 in the HBV S domain, indicated that preS(d)::S(h) was active in proper co-translational membrane integration. However, when probed with trypsin, preS(d)::S(h) was fully sensitive to cleavage independent of whether or not membranes were disrupted. Under both conditions, two HA-reactive fragments were generated, resembling the proteolytic fragments of wt L. Hence, preS(d)::S(h) completely failed to translocate the DHBV preS region, either co- or post-translationally.

As evident from Fig. 3, faster-migrating polypeptides appeared in the case of preS(h)::S(d), which probably represented preS2(h)::S(d) and S(d) forms derived from internal translational initiation. Such forms were missing in preS(d)::S(h) because the S promoter of DHBV located within preS is inactive in non-hepatic cell lines (Welsheimer & Newbold, 1996). This difference, however, is unlikely to account for the divergent behaviour of the two chimeras because co-translational polypeptide translocation into the ER is known to proceed independently from interprotein subunit interaction.

Given that the mixing of the two hepadnaviral L subdomains in either order blocked the correct (re)folding of L, we conclude that preS- and S-specific topogenic determinants of these hepadnavirus genera cannot act synergistically and in concert. To account for the co-translational preS phenotype of preS(h)::S(d), the activity of the HBV-specific CAD element in preventing co-translational preS translocation through interaction with Hsc70 appeared to be abrogated by the DHBV S domain. This might imply that the DHBV S domain has a high preference to orient its N terminus towards the ER lumen, thereby enforcing co-translational N-tail translocation even in the presence of CAD. Such a strong orientation effector function might explain why the DHBV L protein – unlike the corresponding protein of
HBV – engages a stretch of four lysines located within its preS domain in addition to Hsc70-binding elements to block cotranslational preS translocation (Swamey & Schaller, 1997). Reversibly, these binary DHBV preS-specific determinants might dominantly interfere with the topogenic function of the HBV S-specific signals in such a way that the TM2-directed post-translational N-tail reorientation is inhibited in the preS(d) : : S(h) chimera. Taken together, our finding that the specialized topogenic determinants of the DHBV and HBV L proteins were functionally non-interchangeable and failed to cross-operate provides further evidence for distinct folding pathways used by the two hepadnaviral proteins to acquire their dual topology.

ACKNOWLEDGEMENTS

We thank V. Bruss and H. Will for generously providing plasmids carrying cloned DHBV genome sequences. We acknowledge the assistance of T. Döring and thank M. Sapp, H.-C. Selinka and R. E. Streek for helpful discussions. This work was supported by grants to R.P. from the Deutsche Forschungsgemeinschaft (SFB 490-A1).

REFERENCES