Negatively Charged Residues in the IgM Stop-Transfer Effector Sequence Regulate Transmembrane Polypeptide Integration*

(Received for publication, July 23, 1999)

Domina Falcone, Hung Do‡, Arthur E. Johnson‡§, and David W. Andrews¶

From the Department of Biochemistry, McMaster University, Hamilton, Ontario L8N 3Z5, Canada and the ‡Departments of Medical Biochemistry and Genetics and the §Departments of Chemistry and of Biochemistry and Biophysics, Texas A & M University, College Station, Texas 77843-1114

A non-hydrophobic sequence that contributes to the biogenesis of a transmembrane protein is termed a stop-transfer effector (STE). To examine the mechanism of STE-mediated stop-transfer, a series of fusion proteins were constructed containing variants of a putative STE from murine IgM fused to an otherwise translocated hydrophobic sequence. Unexpectedly, the fraction of molecules adopting transmembrane topology was insensitive to many amino acid substitutions within the STE sequence but varied directly with the number of negative charges. Furthermore, when present at the amino terminus of a reporter, mutants were observed that adopted type I (amino terminus lumenal) and type II (amino terminus cytoplasmic) transmembrane topologies, demonstrating that the STE sequence can be located at either side of the endoplasmic reticulum membrane. Our results suggest that recognition of a broad structural feature formed primarily by negatively charged residues within the STE halts translocation and triggers membrane integration, even when the negative charges end up on the cytoplasmic side of the membrane. Since functional STE sequences photocross-link to two membrane proteins not previously identified at the translocon, these unique proteins are presumably involved in recognizing STE sequences and/or facilitating STE function.

Hydrophobicity is the primary driving force for polypeptide integration into membranes. Moreover, the precise sequence and/or length of a hydrophobic transmembrane sequence is a critical factor in regulating transmembrane topology (1). Yet, accumulating evidence also suggests a role for charged residues, primarily in determining the orientation of hydrophobic transmembrane segments in membranes (2, 3). In general, the more positively charged end of the transmembrane sequence is located on the cytoplasmic side of either the endoplasmic reticulum (ER)1 or bacterial plasma membrane (4, 5). In *Escherichia coli*, positively charged residues have been shown to play an active role in post-translational (mostly Sec-independent) integration into membranes (6, 7). Furthermore, it appears that positive charges regulate topology by blocking protein translocation (7). In yeast, it appears that both negatively and positively charged residues can contribute to the overall charge distribution to govern orientation (8). Recently, other non-hydrophobic effectors of membrane topology have been reported, including both negatively charged residues in *E. coli* (9) and the folded state of the amino terminus (3).

In eukaryotes, cell-free assay systems have revealed numerous details about the translocation process and also provided indications of considerable regulatory complexity (10). In addition to sequences that regulate topology, cell-free systems have been used to identify sequences that regulate translocation at the ER. A block in translocation has been described for plasminogen activator inhibitor-2 that alters the interaction of the nascent polypeptide with the signal recognition particle and with the translocation machinery in the ER membrane. The translocation block results in both secreted and cytoplasmic forms of the protein (11). There is also evidence that P-glycoprotein, a membrane protein predicted to span the membrane 12 times, adopting at least two different membrane topologies (12, 13).

Cell-free assays were also used to demonstrate that the protein ductin adopts two different topologies at the ER membrane (14). It was later determined that both forms are physiologically relevant. One form of ductin is a subunit of the vacuolar H^+-ATPase, and the other is found in gap junctions. Dual topology has also been demonstrated for the L protein of hepatitis B virus. In this case the molecule appears to be initially integrated into the ER membrane in a single orientation and then a subset of the polypeptides are post-translationally translocated across the ER membrane to generate the alternate orientation (15).

In contrast to these sequences, which appear to determine the final orientation of a transmembrane segment in the membrane, there is also evidence for sequence-specific regulation of polypeptide integration into the ER membrane (10). Integration into membranes may be regulated for some sequences that are not sufficiently hydrophobic to simply partition into the lipid bilayer. For example, correct membrane assembly of some integral membrane proteins involves the integration of multiple transmembrane segments that individually are unable to integrate into the lipid bilayer (10). Co- and post-translational translocation mechanisms involving the first two transmembrane sequences have been implicated in correct assembly of the cystic fibrosis transmembrane conductance regulator (16). Furthermore, an unusual transmembrane sequence that appears to be regulated for integration into the ER membrane was identified as part of transmembrane sequence 7 from P-glycoprotein (17).

Finally, an element termed a stop-transfer effector (STE) sequence has been identified that mediates membrane integra-
tion of an otherwise secreted hydrophobic domain (18). The two known STE sequences were identified in the transmembrane form of murine IgM (19) and in the prion-related protein, PrP (20, 21).

Many immunoglobulins are synthesized in both secreted and transmembrane forms. Over most of the length of the molecule, the two forms are identical, but the transmembrane forms contain an additional sequence at the carboxyl terminus of the polypeptide. This sequence encodes a hydrophilic sequence of approximately 20 amino acids, followed by the hydrophobic transmembrane domain of the protein. Differential splicing of the primary transcript that adds another exon to the mRNA accounts for the change in protein sequence (22, 23). In vitro, a sequence of amino acids including the hydrophilic sequence amino-terminal of the murine IgM transmembrane sequence was demonstrated to have STE activity when positioned at either the amino terminus or in the middle of a secreted protein (19).

Although the mechanism of STE-mediated membrane integration is poorly characterized, recent data suggest that it may be fundamentally different than membrane integration mediated by hydrophobicity alone. For example, unlike conventional topology-determining sequences, the function of the PrP STE sequence depends both on the composition of the cell-free translation extract (20) and involves a nascent polypeptide-encoded pause in translocation (24). Moreover, recent data demonstrate that when inappropriate membrane integration by the PrP STE is provoked in transgenic animals, neuron degeneration results that are very similar to that seen in Gerstmann-Strausssler Scheinker syndrome, a correlation that suggests that STE regulation is important in human disease (25). However, little is known about the specific properties of STE sequences that trigger integration into membranes because the two known STE sequences show no sequence similarity. Thus, important residues in these STE sequences must be identified experimentally.

Here we have used a mutagenesis strategy along with photocross-linking to decipher the signal within the IgM STE that leads to membrane integration. Membrane integration is shown to depend on the unusually large number of negatively charged residues within the STE. Surprisingly, the precise position of the charges is unimportant, and the contribution of each of the negatively charged residues to integration of the polypeptide into the membrane is additive. It appears that STE recognition occurs within the translocon, since functional and non-functional STE sequences both photocross-link the translocon-associated protein TRAM. More importantly, functional STE sequences photocross-link to two proteins, not previously identified as translocon components that may recognize STE elements and effect their function.

**EXPERIMENTAL PROCEDURES**

*Plasmid Constructs—* Plasmids were constructed by introducing restriction sites into the wild-type sequence by site-directed mutagenesis, and then all further substitutions were made by inserting oligonucleotides into these sites within the plasmids. The positions of the relevant restriction sites relative to the coding sequence is indicated below the sequence in Fig. 1A. All of the plasmids were sequenced using standard techniques. All plasmids except C1 were inserted following the SP6 untranslated leader in pSPUTK. In C1, the plasmid is identical to pSPUTK except that an Ala at the −3 position of the UTK leader is substituted with a Cys. The coding sequences of the different plasmids differ only in the regions indicated as the hydrophobic or the STE region in Fig. 1. Unless specified otherwise, the hydrophobic region was derived from the hydrophobic core of the preprolactin signal sequence by deleting the last seven amino acids and replacing an amino-terminal arginine with a cysteine (19). In WT, the hydrophobic region contains the hydrophobic core sequence from IgM (19). An additional set of plasmids was generated by replacing the hydrophobic region with the regulated portion of transmembrane sequence 7 from human P-glycoprotein (17). The signal cleavage region is common to all of the constructs and is arbitrarily defined as the 9 amino acids after the hydrophobic region. This sequence contains a cryptic signal peptide cleavage site. Thus the signal cleavage region is the first 9 amino acids of the Pt domain (amino acids 58–199 of prolactin) that was used as a reporter.

The control plasmids encode related non-hydrophobic residues and positive charges (C1 and C2) or a single Met (C3) in place of the STE region in the other constructs. Five plasmids were constructed in which the STE and hydrophobic region are positioned internally in the preprolactin signal sequence. The two constructs included STEs A2, A7, A16, C1, and C2. Another three plasmids were constructed in which the hydrophobic region was replaced by a segment (NGGLQPFAAFISFKTTG弗DIDPETKQR) from transmembrane sequence 7 from P-glycoprotein (17). In these plasmids, the first reporter sequence (globin amino acids 1–107) is preceded with a secretory signal peptide (from preprolactin), and thus the entire polypeptide is translocated into the interior of the microsome unless the STE stops translocation. The construction of these plasmids was via cassette mutagenesis using the restriction sites in Fig. 1 but starting with a plasmid encoding the globin and Pt domains reported previously (19). Construction and characterization of the plasmid encoding the WT STE and hydrophobic transmembrane region of murine membrane IgM fused to two non-hydrophobic residues of bovine preprolactin was reported previously (19). The sequence MVTER overlined in Fig. 1 indicates the extra amino acids added to the WT STE during cloning. Full construction details for each plasmid are available from the authors.

**Analysis of Topology and Membrane Integration—** Plasmids were transcribed in vitro using SP6 polymerase, and then polypeptides were synthesized from unpurified transcription products using either a rabbit reticulocyte lysate translation system or a wheat germ extract supplemented with canine pancreatic microsomes. Prior to proteolysis, the canine pancreatic microsomes were separated from untagged molecules in the reticulocyte lysate translation reaction by gel filtration chromatography on Sepharose CL-2B resin. The excluded fractions containing microsomes and microsome-associated proteins (150 µl) were pooled, divided into 3 aliquots of 45 µl, and proteinase K was added to two of the samples at 0.03 µg/ml for 15 min at 0 °C. Proteolysis was terminated by adding trichloroacetic acid to a final concentration of 16% (w/v). Precipitated proteins were washed with ethanol/ether (1:1, v/v), and resuspended in SDS-PAGE loading buffer. After electrophoresis, radioactivity was recorded from the dried polyacrylamide gel using a PhosphorImager (Molecular Dynamics).

To create a reagent that would modify cysteine residues in the fusion protein, but would not cross the ER membrane, we reacted N-(γ-maleimidobutyryl) succinimide ester (GMBS, Pierce) with spermine to synthesize N-(γ-maleimidobutyryl) spermine amide (GMBSA). The resulting compound contains a reactive maleimide attached to one of the four amino groups of spermine. For this reaction, spermine was dissolved in 50% (v/v) acetonitrile/water to a final concentration of 200 µM. A pH 7.5 solution of 0.5 µl spermine was added to the spermine solution, and the resulting solution was covered with paraffin to prevent evaporation. GMBS was dissolved in 100% acetonitrile to a final concentration of 180 mM, and 200 µl was added to the spermine solution in 20-µl aliquots at 1-min intervals through a small hole in the parafilm cover. The reaction was allowed to proceed for an additional 40 min at room temperature and then the reaction mixture containing GMBSA was reduced to approximately 200 µl in a speedvac.

For labeling membrane proteins from translation reactions, microsomes were isolated from a 10-µl translation reaction using a Sepharose CL-2B gel exclusion column equilibrated in 10 mM Tris-HCl (pH 7.5), 50 mM KCl, 2 mM MgCl2. Protease inhibitors were added to the excluded volume (40 µl) to final concentrations of 1 µg/ml chymostatin, antipain, leupeptin, pepstatin, and 2 µg/ml aprotinin. The excluded fraction containing the membranes was divided into two equal fractions, adjusted to 30 mM Tris-HCl (pH 7.5), and one was labeled by adding 4.5 µl GMBSA and incubating at 24 °C for 1 h. The reactions were quenched by adding dithiothreitol to 20 mM and incubating for 5 min at 24 °C. The membranes were then pelleted from the reactions by centrifugation at 100,000g for 15 min at 4 °C through a 100 µl 0.5 µl sucrose cushion in a 200-µl Airfuge tube. To assess the GMBSA reactivity (topology) of integrated proteins, the reactions were processed as above except that they were adjusted to 1 M urea, 0.2 M sodium carbonate (pH 11.5) before centrifugation. Membrane pellets were resuspended directly in SDS-PAGE loading buffer and separated on a 15%–20% Laemmli polyacrylamide gel.

**Photocross-linking—** Photoreactive probes were incorporated into...
Negative Charges Regulate Polypeptide Integration

**A The Topogenic Element**

- STE region: MGERGEAEGGFEYW
- H region: TCCLLAAGNL
- C region: CTISLPL

**B Properties of the Mutant STE Sequences**

<table>
<thead>
<tr>
<th>Name</th>
<th>STE sequence</th>
<th>STE charges</th>
<th>H length</th>
<th>%TM</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>MVTERVDKS(T)</td>
<td>negative: 2 positive: 26</td>
<td>26</td>
<td>89 (6.0)</td>
</tr>
<tr>
<td>A14</td>
<td>MDSKGSQK</td>
<td>8</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>A10</td>
<td>M . . . . . .</td>
<td>6</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>A11</td>
<td>M . L . . . .</td>
<td>6</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>A12</td>
<td>M . P . . . .</td>
<td>5</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>A11</td>
<td>M . . . . . .</td>
<td>5</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>A10</td>
<td>M . . . . . .</td>
<td>4</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>A12</td>
<td>M . . . . . .</td>
<td>4</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>A11</td>
<td>M . . . . . .</td>
<td>4</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>A6</td>
<td>MAXVNLMS</td>
<td>4</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>A5</td>
<td>M . . . . . .</td>
<td>4</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>A4</td>
<td>M . . . . . .</td>
<td>3</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>A3</td>
<td>M . . . . . .</td>
<td>2</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>A2</td>
<td>M . . . . . .</td>
<td>1</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>A1</td>
<td>MAXVNLMS</td>
<td>1</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>A10</td>
<td>MS . T . KL . .</td>
<td>4</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>A15</td>
<td>MGDLDDVCAWGQDSELGQLM</td>
<td>7</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>C1</td>
<td>MDSKGSQK</td>
<td>1</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>C2</td>
<td>MDSKGSQK</td>
<td>1</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>C3</td>
<td>M . . . . . .</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
</tbody>
</table>

**FIG. 1. A**, amino acid sequences (in single letter code) of the STE, H (hydrophobic), and C (signal cleavage) regions of the constructs analyzed. The approximate positions of the restriction endonuclease sites in the corresponding cDNA sequence used to construct the mutants are indicated below the sequence. **B**, properties of the mutant STE sequences. In WT the bar above the sequence MVTER indicates extra amino acids added to the STE sequence during cloning (19). The residue indicated with the 1 is the first amino acid of the IgM STE. The M at the amino terminus of each sequence is the initiator methionine. Dots indicate residues unchanged from the WT sequence. An extra Ala residue inserted in A1–A5 is indicated by a superscript. The length of each hydrophobic region is the number of uncharged amino acids between the last charge in the STE and the first residue of the signal cleavage region. The fraction of polypeptides inserted into the ER membrane when the STE was positioned at the amino terminus of the Pt reporter is indicated by %TM; the standard error for that determination is given by the number in brackets to the right. * indicates the hydrophobic region in the control molecule with the WT STE is the authentic IgM hydrophobic sequence TTASTFIVLFLSLYPYTVTLF. ***, molecules for which %TM was not determined when located at the amino terminus of the reporter.

The approximate positions of the restriction endonuclease sites in the corresponding cDNA sequence used to construct the mutants are indicated in the STE region. The distribution of charged residues in the STE and the uncharged portion of glutamic and aspartic acid residues vary. In mouse, the transmembrane form of IgM contains one aspartic and seven glutamic residues, whereas immunoglobulin 2a (A15) contains five aspartic and two glutamic acid residues (Fig. 1).

There is no precedent for negatively charged amino acids controlling polypeptide integration into membranes (28). However, our preliminary deletion analysis of the IgM STE suggested that the Glu residues in the sequence might contribute to STE function (data not shown). Therefore, plasmids encoding a series of fusion proteins were constructed to determine if the negatively charged residues within the IgM STE sequence can contribute to polypeptide integration into ER membranes (Fig. 1). Each fusion protein contains the topogenic element being tested for STE activity and one or more protein domains used as reporter sequences.

**Characterization of the STE Sequence**—The additional exon expressed in the transmembrane form of murine IgM includes both the hydrophobic transmembrane sequence and a 21-residue putative STE sequence (19). Similar hydrophilic sequences, containing 5–7 negatively charged residues, are found directly amino-terminal of the transmembrane sequence of many immunoglobulins from species as diverse as human and teleost fish. However, both the positions and the relative proportion of glutamic and aspartic acid residues vary. In mouse, the transmembrane form of IgM contains one aspartic and seven glutamic residues, whereas immunoglobulin 2a (A15) contains five aspartic and two glutamic acid residues (Fig. 1).

There is no precedent for negatively charged amino acids controlling polypeptide integration into membranes (28). However, our preliminary deletion analysis of the IgM STE suggested that the Glu residues in the sequence might contribute to STE function (data not shown). Therefore, plasmids encoding a series of fusion proteins were constructed to determine if the negatively charged residues within the IgM STE sequence can contribute to polypeptide integration into ER membranes (Fig. 1). Each fusion protein contains the topogenic element being tested for STE activity and one or more protein domains used as reporter sequences.

**Regulation of polypeptide integration into membranes and the topologies of the fusion proteins containing each of the various STE sequences were assayed by proteolysis, carbonate extraction, and chemical labeling, after synthesis in reticulocyte lysate or wheat germ extract containing canine pancreatic microsomes. Finally, photocross-linking was used to compare the environments occupied by functional and non-functional STE sequences at the ER membrane.**

**RESULTS**

**Characterization of the STE Sequence**—The additional exon expressed in the transmembrane form of murine IgM includes both the hydrophobic transmembrane sequence and a 21-residue putative STE sequence (19). Similar hydrophilic sequences, containing 5–7 negatively charged residues, are found directly amino-terminal of the transmembrane sequence of many immunoglobulins from species as diverse as human and teleost fish. However, both the positions and the relative proportion of glutamic and aspartic acid residues vary. In mouse, the transmembrane form of IgM contains one aspartic and seven glutamic residues, whereas immunoglobulin 2a (A15) contains five aspartic and two glutamic acid residues (Fig. 1).

There is no precedent for negatively charged amino acids controlling polypeptide integration into membranes (28). However, our preliminary deletion analysis of the IgM STE suggested that the Glu residues in the sequence might contribute to STE function (data not shown). Therefore, plasmids encoding a series of fusion proteins were constructed to determine if the negatively charged residues within the IgM STE sequence can contribute to polypeptide integration into ER membranes (Fig. 1). Each fusion protein contains the topogenic element being tested for STE activity and one or more protein domains used as reporter sequences.

**Characterization of the STE Sequence**—The additional exon expressed in the transmembrane form of murine IgM includes both the hydrophobic transmembrane sequence and a 21-residue putative STE sequence (19). Similar hydrophilic sequences, containing 5–7 negatively charged residues, are found directly amino-terminal of the transmembrane sequence of many immunoglobulins from species as diverse as human and teleost fish. However, both the positions and the relative proportion of glutamic and aspartic acid residues vary. In mouse, the transmembrane form of IgM contains one aspartic and seven glutamic residues, whereas immunoglobulin 2a (A15) contains five aspartic and two glutamic acid residues (Fig. 1).

There is no precedent for negatively charged amino acids controlling polypeptide integration into membranes (28). However, our preliminary deletion analysis of the IgM STE suggested that the Glu residues in the sequence might contribute to STE function (data not shown). Therefore, plasmids encoding a series of fusion proteins were constructed to determine if the negatively charged residues within the IgM STE sequence can contribute to polypeptide integration into ER membranes (Fig. 1). Each fusion protein contains the topogenic element being tested for STE activity and one or more protein domains used as reporter sequences.

**Regulation of polypeptide integration into membranes and the topologies of the fusion proteins containing each of the various STE sequences were assayed by proteolysis, carbonate extraction, and chemical labeling, after synthesis in reticulocyte lysate or wheat germ extract containing canine pancreatic microsomes. Finally, photocross-linking was used to compare the environments occupied by functional and non-functional STE sequences at the ER membrane.**

The hydrophobic region used as a potential transmembrane domain was derived from the preprolactin signal sequence because it is unlikely that a hydrophobic sequence from a cleaved secretory signal sequence would contain information other than hydrophobicity that specifies membrane topology or integration. To generate the hydrophobic region from the preprolactin signal sequence, the last seven amino acids of the sequence were deleted (including the signal cleavage site), and an amino-terminal Arg was replaced with Cys. Deletion of the signal cleavage site greatly reduced signal peptide cleavage when the sequence is positioned internally within a secreted protein (less than 5% of molecules are cleaved, data not shown).

Although all of the polypeptides contained the same hydrophobic region, the overall length of the hydrophobic portion of the potential transmembrane segment is also determined by the distribution of charged residues in the STE and the uncharged residues in the reporter that immediately precedes the putative STE sequence. Substitution of the negative charges within the STE with uncharged residues can also change the length of the potential transmembrane segment (Fig. 1).

**Stop-Transfer Activity at an Internal Location**—To examine assembly of a conventional type I transmembrane protein, chimeric proteins were constructed containing a secretory signal sequence derived from globin (Gt, amino acids 1–107 of chimpanze globin) and another reporter sequence from preprolactin (Pt, amino acids 58–199 of bovine preprolactin). The sequences used as reporters have been shown previously to contain no cryptic topogenic elements (29). To test the stop-transfer activity of the various STE sequences and hydrophobic region combinations, they were positioned between the Gt and Pt reporter sequences (Fig. 2).
Negative Charges Regulate Polypeptide Integration

Fig. 2. Negative charges are required for STE-mediated integration of marginally hydrophobic sequences into the ER membrane. A, sequences were tested for STE activity with hydrophobic sequences of 15, 17, or 23 residues (hatched, black, and open bars, respectively). Nucleotides encoding the test sequences (containing 1 or 4 Glu residues and the indicated hydrophobic sequence) were positioned between sequences coding for the Gt and Pt reporters. Proteolysis of the Pt reporter was used to identify molecules with transmembrane topology after transcription of the plasmid in vitro and translation of the transcription products in reticulocyte lysate with added canine pancreatic microsomes. The STE sequences analyzed were C1, C2, A2 (one negative charge), and A16 and A7 (4 negative charges), from left to right. B, when positioned within a secreted reporter protein, molecules containing a functional STE (4 Glu residues, open box labeled 4) and a marginally hydrophobic region (15–17 residues) integrate as type I (amino terminus luminal) proteins with the Gt domain inside and the Pt domain outside of the microsome. Non-functional STE mutants (1 Glu residue, open box labeled 1) are unable to stop translocation and are released into the ER lumen. The secretory signal sequence at the amino terminus of the Gt domain is cleaved from both molecules and is not shown. C, the negative charges in the STE sequence are required for integration into membranes of a segment of the transmembrane sequence 7 segment from P-glycoprotein. Upward arrowheads indicate the protease-protected globin fragment diagnostic of transmembrane molecules. The intensity of this band is much less than for the full-length protein because it contains only 4 of the 8 methionine residues in the full-length protein. Correction for the number of methionines revealed that 11, 18, and 23% of the molecules were transmembrane for STEs A3, A7, and A13, respectively. Downward arrowheads indicate secreted signal-cleaved molecules. Comparison of total reaction products (lanes 1, 5, and 9) with membrane-bound molecules indicates that the amino-terminal signal peptide is cleaved when the molecules are translocated. The number of negative charges and the name of each STE are indicated above the lanes. The STE sequences are as shown in Fig. 1 except that the last three amino acids (NLW) were replaced with a single amino acid (Val). The migration positions of molecular mass markers (in kDa) are shown to the right of the panel.
Negative Charges Regulate Polypeptide Integration

converts the mechanism of integration from one dependent on the STE to one entirely due to hydrophobicity. It was not possible to avoid this complication by substitution of the negatively charged amino acids with sequences containing positively charged residues because the positively charged residues abolished STE function (data not shown). When the hydrophobic region was positioned at the amino terminus of a reporter protein, it was cleaved by signal peptidase and removed from the reporter protein if it failed to be recognized as a bona fide transmembrane domain. Cleavage occurred even when the hydrophobicity was increased by the various charged residue substitutions. For this reason we examined the function of the hydrophobic region alone and fused behind the various STE sequences when positioned at the amino terminus of the Pt reporter domain. Stop-Transfer Activity at an Amino-terminal Location— When the hydrophobic region was positioned at the amino terminus of a reporter domain, it functioned primarily as a secretory signal sequence, and approximately 80% of the polypeptides were cleaved by signal peptidase. The remaining molecules were not cleaved by signal peptidase, presumably because the authentic signal cleavage site was deleted, and cleavage at the cryptic site is less efficient. Since relatively long uncharged sequences may still function as cleaved secretory signal sequences instead of integrated signal-anchor sequences (28), it was possible to substitute the charged residues in the STE with uncharged amino acids and assay STE activity by the extent of conversion of the signal sequence (translocation) to a signal-anchor sequence (integration). Thus, at either an amino-terminal or internal location, an STE and hydrophobic region become transmembrane if the STE is functional but pass through the ER membrane if the STE is nonfunctional or absent.

Mutants were analyzed to evaluate the importance of various aspects of the STE sequence as follows: total length (e.g. A14 versus A13); addition of a positively charged residue (e.g. A13 versus A12); length of hydrophobic sequence (substitution of Glu with non-charged residues at the carboxyl end of the STE sequence increases the length of the hydrophobic region, e.g. A11 versus A13); and/or replacement of single or multiple residues as indicated in Fig. 1B. To compare the Glu-rich IgM STE with a similar sequence containing Asp residues, we also examined the putative STE from murine immunoglobulin 2a (A15). As secreted controls we used a sequence containing primarily hydrophilic residues or a single methionine (STE C1 and C3, respectively). As a transmembrane control we used a previously characterized construct containing both the wild-type STE and the hydrophobic transmembrane sequence of IgM fused to the Pt reporter (19). This fusion protein also contains the additional amino acids MVTER at the extreme amino terminus (Fig. 1B).

In the constructs assayed here, if a functional STE sequence is fused to the amino terminus of the hydrophobic region, the chimera is converted from a signal peptide into a signal-anchor sequence, and the Pt domain is anchored as a transmembrane protein with a mixture of type I and type II topologies (Fig. 3). Thus, depending on the extent of signal peptide cleavage, any particular fusion protein can have up to four fates when it interacts with membranes in vitro (Fig. 3B). If the STE does not halt translocation (i.e. is non-functional), the molecule is secreted and the signal peptide is or is not cleaved (Fig. 3A, upward and downward arrowheads, respectively). If translocation is halted by the STE, the transmembrane proteins can adopt type I (protease-sensitive) or type II (partial protease protection, Fig. 3A, dots) orientations. All of the fusion proteins were examined for translocation by translation in reticulocyte lysate containing canine pancreatic microsomes, followed by separation of targeted from non-targeted molecules by gel filtration chromatography and proteolysis of the former. Translocated molecules (containing a functional STE) are protected from proteolysis because they are inside the microsome. Most of the translocated molecules were also cleaved by signal peptidase. In contrast, transmembrane molecules (containing a non-functional STE) are partially or com-

![Fig. 3. Proteolysis assays used to determine membrane topology. A, secreted molecules (arrowheads) and the lumenal portion of the type II transmembrane molecules (dots) are protected from added protease (Ptk) unless the membrane is solubilized with detergent (Det). For type I (amino terminus, luminal) transmembrane proteins, the Pt reporter is protease-sensitive. In type II proteins, only the cytoplasmically localized STE sequence at the amino terminus is digested by the protease. Fully secreted molecules, both signal sequence-uncleaved (downward pointing arrowheads) and signal sequence-cleaved (upward pointing arrowheads), are protected from protease. Since membrane fractions were separated from the reactions by gel filtration chromatography prior to analysis by proteolysis, non-targeted molecules are not seen. The number of negative charges and the name of each STE are indicated above the lanes. The migration positions of molecular mass markers are indicated in kDa to the right of the panels. B, when positioned at the amino terminus of the reporter protein, molecules containing a functional STE integrate as type I (amino terminus luminal) or type II (amino terminus cytoplasmic) proteins. Non-functional STE mutants are unable to stop translocation and may or may not be signal-cleaved when they are released into the ER lumen (arrowhead labeled S).]
pletely digested by the added protease.

When polypeptide disposition was assessed as detailed below, a direct correlation (slope = 9.1, r = 0.96) was observed between the number of glutamic residues and the stop-transfer efficiency of the STE sequence (Fig. 4A). The numerical data used to generate Fig. 4A and the error associated with each data point are presented in Fig. 1B. This direct correlation was particularly surprising given the variety of differences between the STE sequences tested. For example, polypeptides with the A6 and A7 STEs both contain four negative charges but differ at eight other positions, and yet both polypeptides are about 50% transmembrane. In contrast, the polypeptides with STEs A13 and A7 (six and four negative charges, respectively) vary at only these two positions, yet the stop-transfer efficiencies of these STEs (approximately 75 and 50%, respectively) are significantly different.

Remarkably, the putative STE from Igγ2a (A15) containing 5 Asp and only 2 Glu residues resulted in a quantitatively similar number of transmembrane molecules as that obtained for polypeptides with the A13 STE that contains 6 Glu residues (78 and 75%, respectively). Thus, we conclude that STE function is effected by the negative charges, and it is not important whether the charges result from Asp or Glu residues. Most of the translocated molecules are signal-cleaved. Hence, it is not surprising that the extent of signal peptide cleavage clearly correlated inversely (slope = -9.5, r = 0.96) with the number of glutamic residues in the STE sequence (Fig. 4B). However, lack of signal cleavage was not in itself sufficient to halt translocation, since some uncleaved molecules were secreted (Fig. 3A, downward pointing arrowheads). Furthermore, lack of signal peptide cleavage for the 15-residue hydrophobic region used here did not prevent translocation when located in the interior of a secreted protein (Fig. 2A).

Transmembrane Orientation Determined by Proteolysis—In addition to determining the fraction of the molecules that are transmembrane, proteolysis data were also used to determine the transmembrane orientation of the molecules. Molecules with type II topology are easily measured since these molecules are partially protected from added protease. As expected there is a strong correlation between the fraction of molecules with type II transmembrane topology and the number of negative charges in the STE sequences (slope 6.8, r = 0.87, Fig. 4C). This correlation can be appreciated by visual inspection of the PhosphorImager data since the band that results from proteolytic removal of the short cytoplasmic amino terminus increases in intensity (dots, Fig. 3).

Unlike type II molecules, polypeptides with type I topology are almost completely digested by added protease (the lumenal and transmembrane domain together are predicted to be less than 50 amino acids long and contain only a single methionine).
The fraction of type I molecules was calculated from the PhosphorImager data after correcting for the number of methionine residues in each species by first adding together the signals from the bands resulting from uncleaved but secreted polypeptides and the type II transmembrane polypeptides in the plus protease lane (downward pointing arrowheads and dots, respectively) and subtracting this sum from the signal from the upper band (uncleaved molecules) in the minus protease lane (Fig. 4D). The fraction of molecules adopting a type I topology varied between just less than 20 and approximately 50%. With the exception of molecules with STE A15, the correlation between the fraction of the polypeptides that adopt type I topology and the number of negative charges in the STE (slope = 2.3, \( r = 0.35 \)) is marginal. If the data point due to STE A15 (7 negative charges) is excluded, the calculated slope of is 1.7 with \( r = 0.30 \).

Because of the uncertainty involved in determining the fraction of molecules with type I topology by this calculation, we confirmed these results using chemical labeling. Very similar results were obtained when type I molecules were quantified after labeling with GMBSA, a membrane impermanent modifier of cysteine (see below).

Although most of the STE-dependent increase in transmembrane molecules was due to an increase in the number of molecules with type II topology, STE-dependent triggering of transmembrane topology is not dependent on which side of the ER membrane the STE ultimately resides. For example, approximately half of the roughly 75% transmembrane molecules with STE A12 or A13 adopted the type I topology.

Examination of the topology adopted by each of the mutants revealed that individual glutamic residues, the single lysine and the length of the hydrophobic sequence each, by itself, has only limited influence on the topology adopted. This result contrasts directly with data obtained using conventional signal-anchor sequences where topology is directly related to the distribution of charged residues surrounding the hydrophobic segment.

**Hydrophobicity Does Not Correlate with Stop-Transfer Activity**—Unexpectedly, no correlation was observed between the number of uninterrupted uncharged residues in the hydrophobic sequence and stop-transfer activity (Fig. 4E). Moreover, the A2 STE and hydrophobic region functioned as a relatively efficient secretory signal peptide (71% fully translocated) even though it contains a 23 hydrophobic residue sequence followed by a sequence (CHTSSLPTP) that is not particularly hydrophilic. Only the His and Pro in this latter sequence are rare in transmembrane sequences (30). There was also no correlation when stop-transfer activity was compared with total calculated hydrophobicity of the hydrophobic sequence in each molecule (data not shown). Thus, changes in hydrophobicity cannot account for the differences in stop-transfer activity observed for the various STE sequences.

**Integration of Transmembrane Proteins**—To determine if the transmembrane molecules actually integrated in the ER membrane or if they are translocation intermediates that are “stuck” in the translocon, membrane integration was assayed by the rigorous criterion of resistance to extraction from membranes by incubation in 1 M urea, 0.2 M sodium carbonate (pH 11.5) (19) (Fig. 4F). As expected when using such harsh conditions, the fraction of molecules scored as integrated is in general slightly less than the number determined to be transmembrane by proteolysis. Nevertheless, the correlation between the number of negative charges and membrane integration (slope = 8.3; \( r = 0.79 \)) is still striking. Since, non-targeted molecules would complicate quantification of carbonate extraction data, gel filtration chromatography was used to remove non-targeted molecules prior to incubation in 1 M urea, 0.2 M sodium carbonate (pH 11.5).

In contrast to proteolysis assays, type I molecules cannot be distinguished from type II molecules by carbonate extraction. To confirm that both type I and type II transmembrane molecules were integrated into the ER membrane, we labeled the Cys residues in type I molecules with GMBSA prior to carbonate extraction. Unlike other Cys-modifying reagents, GMBSA does not cross ER membranes.2 Reaction with GMBSA results in covalent attachment of the reagent to Cys residues, thereby adding 3 positive charges (but only approximately 300 in molecular weight) to the polypeptide for each modified Cys. Although not all Cys residues are equally accessible to the labeling reagent, sufficient labeling occurs at the 4 Cys residues in the Pt domain, to permit labeled type I molecules to be differentiated from unlabeled type II molecules. As controls we labeled the \( \beta \)-subunit of signal recognition particle receptor (a type I transmembrane protein containing three Cys residues) and showed that the 6 Cys residues in prolactin were protected from modification with GMBSA after translocation into the ER (data not shown).

By using this technique, even type I molecules that make up only 10% of the total can be identified and shown to be resistant to extraction with sodium carbonate (Fig. 5A, compare lanes 3 and 6, downward-pointing arrowheads, with lanes 2 and 5, respectively). The upward pointing arrowheads indicate molecules integrated in the type II orientation that were not modified with GMBSA. Signal cleaved molecules (upward arrows) are protected from modification with GMBSA but are not resistant to extraction with sodium carbonate because they are located in the lumen of the ER.

Only a small fraction of the molecules with the A2 STE pellet after carbonate extraction (Fig. 5A, compare the signal cleaved molecules indicated by the upward arrow in lane 2 with the corresponding band in lane 3). However, more than half of the pelleted molecules are integrated with type I topology because they are modified with GMBSA (downward pointing arrowhead in lane 3). In contrast, the A8 STE efficiently stops translocation, and molecules with either type I (downward pointing arrowhead, lane 6) or type II topology (upward pointing arrowhead, lane 6) were integrated in the ER membrane. Extending this analysis to other STE sequences revealed that in all cases the majority of type I membrane-associated molecules were carbonate-resistant.

GMBSA labeling also permitted us to measure the fraction of each molecule that adopted the type I topology. After GMBSA labeling the fraction of molecules adopting type I topology can be measured directly from the PhosphorImager plate rather than by calculation from proteolysis data. Direct comparison of the results obtained using both methods revealed excellent agreement. Thus, our estimates for the fraction of molecules that adopted type I topology are reliable (Fig. 5B).

**Probing the Environment of the STE Sequence during Integration**—To examine the mechanism of STE-mediated membrane integration in more detail, we used photocross-linking (26, 31). For these experiments, nascent chains containing the STE, hydrophobic region, and 51 amino acids of the Pt passenger were generated by \textit{in vitro} translation of truncated mRNAs. Truncation of the nascent chain at amino acid 51 of Pt (by restriction digestion of the plasmid with \textit{Taq}I) positions the amino-terminal end within the translocon but not far enough to permit signal peptide cleavage. This allowed us to compare the immediate environment of both functional and non-functional

---

STE sequences within the translocon by photocross-linking. To identify cross-links dependent on proximity to a functional STE, we compared the photoadducts obtained for nascent chains with the A13, A12, and A2 STEs. The only difference between the A12 and A13 STEs is the presence of a lysine in place of asparagine at amino acid 11 of STE A12. Since these STEs function similarly (70–75% transmembrane molecules), any photocross-links observed for STE A12, but not with STE A13, are due to the additional probe inserted in the center of the A12 STE. Photoadducts formed in both A12 and A13 nascent polypeptide samples result from probes positioned within the coding region of the mature portion of the Pt passenger (amino acids 12 and 49 of Pt). Since the truncation point is at amino acid 51 of Pt, residue 12 is predicted to have exited from the ribosome while residue 49 will be within the ribosome.

The centrally located lysine in STE A12 is also present in STE A2, but 5 of the 6 glutamic acids in A12 were replaced with uncharged residues (Fig. 1). Since less than 30% of the nascent polypeptides with STE A2 end up transmembrane, any photocross-links observed for STE A12, but not with STE A13, are due to the additional probe inserted in the center of the A12 STE. Photoadducts formed in both A12 and A13 nascent polypeptide samples result from probes positioned within the coding region of the mature portion of the Pt passenger (amino acids 12 and 49 of Pt). Since the truncation point is at amino acid 51 of Pt, residue 12 is predicted to have exited from the ribosome while residue 49 will be within the ribosome.

After photolysis, the translation reactions were fractionated by carbonate extraction or by affinity chromatography and analyzed by SDS-PAGE. When samples were pelleted after incubation in sodium carbonate (pH 11.5), and integral membrane proteins (lanes 3 and 6) were pelleted by centrifugation (P), whereas secreted or peripheral proteins (lanes 2 and 5) remained in the supernatant (S). Upward arrows, signal-cleaved secreted molecules; upward arrowheads, type II integral membrane proteins; downward arrowheads, type I integral membrane proteins. B, the fraction of transmembrane molecules with type I topology as determined by GMBSA labeling (squares) is similar to that calculated from proteolysis data (triangles).

When binding to concanavalin A-Sepharose was used to concentrate and detect glycoprotein-containing photoadducts generated with the non-glycosylated nascent polypeptides, no photoadduct with a 69-kDa photocross-linking target was observed, suggesting that this photocross-linking target is not glycosylated. Instead, a new cross-link was visualized with STE A12 (Fig. 6, lane 5, downward arrow). In addition, a band observed just below the 43-kDa marker is darker in lane 5 than in lanes 4 or 6 (upward arrow). Subtracting the molecular mass of the nascent polypeptide from that of this photoadduct yields a predicted molecular mass of 69 ± 6 kDa for the photocross-linking target.

When binding to concanavalin A-Sepharose was used to concentrate and detect glycoprotein-containing photoadducts generated with the non-glycosylated nascent polypeptides, no photoadduct with a 69-kDa photocross-linking target was observed, suggesting that this photocross-linking target is not glycosylated. Instead, a new cross-link was visualized with STE A12 (Fig. 6, lane 5, downward arrow). In addition, a band observed just below the 43-kDa marker is darker in lane 5 than in lanes 4 or 6 (upward arrow). Subtracting the molecular mass of the nascent polypeptide from that of this photoadduct yields a predicted molecular mass of 69 ± 6 kDa for the photocross-linking target.
the STE A12 photoadduct. Finally, no STE-dependent photocross-links to Sec61α were observed using antisera to Sec61α, and photocross-linking to Sec61α was very inefficient for all three nascent polypeptides (data not shown). The extent of cross-linking to Sec61α differs for probes at different positions within a nascent polypeptide (32, 33). Thus, It seems likely that the lack of cross-linking to Sec61α observed here is due to the position of the probes within the STEs.

DISCUSSION

The behavior of the STE sequences examined here is unprecedented. Current models presume that hydrophobicity is the sole or primary determinant of integration into the ER membrane, yet the negatively charged Glu residues in the IgM STE sequence clearly affect the extent of protein integration for hydrophobic sequences of as few as 15 residues. How does a functional STE cause the integration of marginally hydrophobic sequences that would otherwise be translocated across the ER membrane? Do specific proteins mediate the STE effect?

Although we cannot determine where in the ER membrane the STE is recognized, our cross-linking data suggest that it occurs within the translocon.

Our photocross-linking experiments have identified three proteins that are adjacent to the STE sequence in the translocon. One of these proteins is TRAM, a previously identified component of the translocon. But the other two proteins, a 69-kDa membrane protein and a 54-kDa glycoprotein, do not resemble any proteins identified previously in studies with nascent secretory or membrane proteins (26, 27, 31–33). These results therefore suggest that STE-mediated integration into membranes may be affected by specific proteins, a conclusion that is consistent with the relative insensitivity of this effect to the hydrophobicity of the nascent chain.

Surprisingly, the effect of the Glu residues on conversion of a signal sequence to a signal-anchor sequence was additive. The additive effect of the negatively charged residues on membrane integration suggests that, similar to the effects of positively charged amino acids on membrane topology (34), the negative charges in the STE contribute to a feature required to trigger membrane integration. This result, obtained in two different eukaryotic cell-free systems and for two different hydrophobic sequences, is in direct contrast with results obtained using E. coli in which there was no effect of amino-terminal negative charges on signal sequences integration into membranes (28). Furthermore, although negative charges at the amino terminus of a polypeptide have been predicted to cause the polypeptide to adopt a type I topology (35–37), the cumulative effect of the negative charges was primarily to increase the fraction of molecules that integrated into the bilayer and adopted the type II topology. The integration of these proteins with a topology opposite to that predicted by rules based on the distribution of charges surrounding a transmembrane sequence (2, 5) is consistent with the possibility that STE-mediated integration into membranes occurs via a mechanism distinct from integration of conventional signal-anchor sequences.

As expected for true integral membrane proteins, polypeptides that adopted a transmembrane topology due to a functional STE sequence were not only resistant to extraction with sodium carbonate (pH 11.5) (Fig. 5B) but were also resistant to extraction with 2 M urea, 1 M NaSCN, or 0.08% (w/v) sodium deoxycholate (data not shown). Therefore, transmembrane orientation does not result from a simple STE-dependent steric block of translocation. Consistent with this interpretation, the various STE sequences also did not alter the targeting efficiency of the molecules (data not shown). Significantly, the only other STE sequence characterized thus far (that of PrP) was recently demonstrated to integrate PrP molecules in the ER...
membrane with both type I and type II topology (25). Thus, one characteristic of STE sequences may be that they can eventually reside on either side of the ER membrane. Furthermore, the IgM STE sequences that function at the amino terminus also stop translocation of 40 and 70% of hydrophobic sequences containing 15 or 17 uncharged residues, respectively, when located in the middle of a translocating polypeptide (Fig. 2). We conclude therefore, that the STE and hydrophobic sequences elicit membrane integration no matter where they are located in the translocating polypeptide.

To analyze the IgM STE by proteolysis and chemical labeling, we used a reticulocyte lysate cell-free system because it has been shown to reflect more accurately the topology of transmembrane molecules in transfected cells than does wheat germ extract (38). Yet while the integration of most polypeptides into the ER membrane is qualitatively similar in wheat germ extract, reticulocyte lysate, and in transfected mammalian cells, the only other STE sequence identified to date behaves differently in wheat germ extract and reticulocyte lysate. In wheat germ extract, the PrP STE exhibits dramatically increased membrane integration (20). Therefore, we also assayed all of the molecules in Fig. 1 using this system. Unlike PrP, membrane integration of molecules with the IgM STE and the mutant STEs examined here was quantitatively similar in both systems (data not shown). Thus, it appears that for the IgM STE, a membrane-dependent recognition event leads to STE-mediated integration of the nascent polypeptide into the membrane. The nature of this recognition event is unknown, but our photocross-linking experiments showed that when a functional STE enters the translocon, it is proximal to different proteins than either the mature portion of the nascent chain or a non-functional STE (Fig. 6). The simplest explanation for these data is that recognition of a functional STE moves the nascent chain to a distinct region within the translocon.

However, this transmembrane domain has been shown to interact with several different proteins in B lymphocytes and immature B cells (39, 40). Although one of these has a molecular mass of 32 kDa (prohibitin), it is not known if there is a counterpart in canine pancreatic microsomes. Furthermore, prohibitin is not a glycoprotein. Thus, it seems likely that two of the cross-linking partners identified here are novel ER proteins. It remains to be determined whether the same polypeptide will be found adjacent to other STE sequences.

Taken together, our results demonstrate that co-translational integration of proteins into the ER membrane is not always regulated strictly by hydrophobicity, and suggest that current models for membrane integration should be modified to include a mechanism for regulating integration. Indeed our demonstration that a hydrophobic sequence of 23 uncharged amino acids functioned as an efficient signal peptide, while many constructs containing only 17 uncharged amino acids were integrated into the ER membrane, suggests that predictive algorithms based only on hydrophobicity provide an incomplete view of membrane topology.

REFERENCES


Downloaded from http://www.jbc.org/ by guest on June 25, 2017
Negatively Charged Residues in the IgM Stop-Transfer Effector Sequence Regulate Transmembrane Polypeptide Integration
Domina Falcone, Hung Do, Arthur E. Johnson and David W. Andrews

doi: 10.1074/jbc.274.47.33661

Access the most updated version of this article at http://www.jbc.org/content/274/47/33661

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 41 references, 21 of which can be accessed free at http://www.jbc.org/content/274/47/33661.full.html#ref-list-1