Systematic Introduction of Proline in a Eukaryotic Signal Sequence Suggests Asymmetry within the Hydrophobic Core*

Patrick Ryan‡ and Chad O. Edwards

From the Department of Microbiology and Immunology, University of Tennessee, Memphis, Tennessee 38163

The hydrophobic core or h region of both prokaryotic and eukaryotic signal sequences is the predominant structural domain that controls the efficiency of protein translocation across membranes. Characteristically, hydrophobic cores appear to assume α-helical conformations, and studies in prokaryotes have indicated that this conformation is necessary for efficient signal sequence function. To address the conformational constraints of an eukaryotic signal sequence, we have introduced a single proline in almost each position of the signal sequence hydrophobic core of glycoprotein C (gC) of the swine herpesvirus, pseudorabies virus. When the resulting mutant virus strains were used to infect cells, we found that substitution of proline at certain positions affected gC translocation greater than its introduction at other sites within the hydrophobic core. The observed positional effects did not completely correlate with reductions in overall hydrophobicity or linear position within the hydrophobic core. Rather, it appeared that one face of the gC signal sequence α-helix is far more sensitive to proline disruption than the other, potentially indicating a functional asymmetry.

The N-terminal signal sequences of prokaryotic and eukaryotic polypeptides, required for protein export to extracytoplasmic sites, are characterized not by a conservation of primary amino acid sequence but rather by the maintenance of three structural domains (1). At the N terminus, each signal sequence contains a hydrophilic or n region composed of 1–5 residues, several of which are positively charged. This is followed by the hydrophobic core or h region that is typically 7–15 amino acids in length and highly hydrophobic. At the C terminus, a more polar region (c region) of 3–7 residues defines the signal peptidase cleavage site that is used once the signal sequence has traversed the membrane. Numerous genetic and biochemical studies in prokaryotes and eukaryotes have demonstrated that the predominant feature required for signal sequence function is the hydrophobicity of the core structure (2–4).

Another conserved structural feature of signal sequences is the α-helical nature of the hydrophobic core (1). This conservation has been interpreted to indicate an important role for secondary structure in the translocation process, a conclusion supported by studies of prokaryotic signal sequences. In particular, genetic analysis of suppressors of Escherichia coli LamB signal sequence mutants and subsequent biophysical assessments of mutant LamB signal peptides strongly suggest that an α-helical conformation is necessary for efficient LamB export (5, 6).

We have been conducting a genetic analysis of the gC1 signal sequence encoded by PRV, a swine herpesvirus. A 479-amino acid protein, gC contains 8 N-linked glycosylation sites and is localized as a mature species on the infected cell surface and in the virus envelope (7, 8). Importantly, gC is nonessential for virus growth (9), and it has been previously shown that a complete deletion of the 22-amino acid signal sequence abolishes gC export and renders it a cytoplasmic, nonglycosylated protein (10). We have also introduced more discrete alterations into the hydrophilic and hydrophobic regions of the gC signal sequence and have determined that it is typical of other characterized N-terminal signal sequences (4, 11). Here, we report on the systematic introduction of proline, an accepted α-helix breaker (12), throughout most of the hydrophobic core as a way to evaluate the importance of secondary structure to gC signal sequence function. Our results indicate that some positions within the hydrophobic core are more sensitive to proline substitution than others.

MATERIALS AND METHODS

Cells and Virus Strains—Our wild-type strain, PRV-Be, and the porcine kidney cell line PK15 have been previously described (8).

Construction and Identification of gC Signal Sequence Mutants—Site-specific mutagenesis using the vector pAXS (a derivative of pALTER-1 of Promega Corp.) was performed as described previously (11). Plasmid-borne gC mutants were confirmed by DNA sequencing and introduced into the PRV genome by calcium phosphate-mediated cotransfection of plasmid and viral DNA (11, 13). Recombinant viruses were identified using gC-specific antibodies in a peroxidase-linked immunoassay, the black plaque assay (8, 14). Recombinants were plaque-purified and confirmed by Southern analysis of viral DNA (9).

Radiolabeling and Quantitation of Infected Cell Proteins—Steady-state or pulse-chase radiolabeling of infected cell proteins was performed as described previously (8). gC species were immunoprecipitated with a polyclonal antiserum and resolved on SDS-10% polyacrylamide gels followed by fluorography and autoradiography (8). Appropriately exposed autoradiograms of each pulse-chase experiment were scanned using a PDI densitometer, and the relative levels of the cytoplasmic and ER forms of gC were determined for the 15-min chase point as described (4).

Computer Analyses—The predicted secondary structure of wild-type and mutant gC signal sequences was determined using the Peptide-structure program of the University of Wisconsin Genetics Computer Group (15). The Helicalwheel program of the same package was used to generate Fig. 4.

* This work was supported by Grant AI28520 from the NIAID, National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) U29121-U29129.

‡ To whom correspondence should be addressed: Dept. of Microbiology and Immunology, University of Tennessee at Memphis, 858 Madison Ave., Rm. 201, Memphis, TN 38163. Tel.: 901-448-8764; Fax: 901-448-8462.

1 The abbreviations used are: gC, glycoprotein C; PRV, pseudorabies virus; ER, endoplasmic reticulum; SRP, signal recognition particle; MOI, multiplicity of infection.

27876
Proline Mutations of the PRV gC Signal Sequence

Proline substitution and single codon deletion mutants of the gC signal sequence hydrophobic core. The nucleotide sequence of the sense strand and the predicted amino acid sequence of the wild-type gC signal sequence are shown. The arrows below the sequence indicate the alterations made to substitute a single proline into positions 9–14. The shaded boxes demarcate the deletion of residue 10 and 11, whereas the double mutant gC A10PL11A is shown at the bottom. In all mutants, the strain number is listed below the substituted or deleted amino acid. The downward pointing arrow above the wild-type sequence denotes the predicted signal peptidase cleavage site.

Fig. 1. Proline substitution and single codon deletion mutants of the gC signal sequence hydrophobic core.

RESULTS

Systematic Introduction of Proline Residues into the Hydrophobic Core of the gC Signal Sequence—The hydrophobic core of the gC signal sequence is composed of residues 7–15 and is predicted to exist as an α-helix (12). Initially, six mutant alleles were produced in plasmid-borne copies of gC such that a single codon lying within positions 9–14 was replaced with a proline codon (Fig. 1). Each of the mutant alleles was then independently recombined into the wild-type PRV genome; this gave rise to six different gC signal sequence mutants designated strains PRV543–548 (Fig. 1). Collectively, the mutants introduced a proline residue in each position of one complete turn of the α-helix of the hydrophobic core.

The export competency of each mutant form of gC was evaluated by steady-state radiolabeling infected cell proteins as described previously (8). Wild-type gC was found in the infected cell in two well-characterized forms (8); as a 74-kDa ER-localized glycoprotein and as a 92-kDa Golgi-modified species that resides in the plasma membrane and the mature virus envelope (Fig. 2). An additional form of gC was evident in each of the mutant-infected cells; this protein has been previously demonstrated to be the nonglycosylated, 58-kDa precursor species localized in the cytoplasm (10). The extent of each precursor’s accumulation reflected the degree of the translocation defect, ranging from almost no impairment for a proline substitution at position 10 (strain PRV544) to a near complete block in translocation across the ER membrane as a result of a proline substitution at position 14 (strain PRV548). Interestingly, a vaccine strain of PRV (Bartha) contains greatly reduced levels of gC and has been found to harbor a proline substitution at position 14 (16). In general, mutants in which proline replaced a leucine residue exhibited a more pronounced defect than those exchanging proline and alanine. This was not unexpected because leucine is more hydrophobic than alanine, and it has been established that the principal determinant of gC signal sequence function is its overall hydrophobicity (4). Still, the correspondence of loss of hydrophobicity and loss of translocation proficiency was imperfect. For example, the gC signal sequences of strains PRV543, 545, 546, and 548 were identical in hydrophobic content, yet PRV548-encoded gC was far more defective than the others for translocation across the ER membrane.

Proline Substitutions Impact on gC Translocation Efficiency Beyond Their Diminution of Overall Hydrophobicity—Proline substitutions reduce the hydrophobicity of signal sequences owing to a net hydrophilicity value of 0 as assigned by Hopp and Woods (17). We therefore reasoned that a simple deletion of one of the 9–14 codons should mimic an introduction of proline at the same position if loss of hydrophobicity was the only detrimental consequence of the substitution. Accordingly, two additional mutant strains were generated; PRV549 contains a deletion of the tenth gC codon, whereas PRV550 is deleted for the eleventh codon. A final mutant, PRV551, was also produced. This mutant encoded gC A10PL11A, an allele similar to that harbored by PRV544 but in which the additional replacement of leucine at position 11 with alanine resulted in an overall hydrophobicity identical to the gC signal sequences encoded by strains PRV543, 545, 546, and 548 (Fig. 1).

All of the mutant strains and wild type were used in pulse-chase experiments (8, 11) to accurately determine the translocation defect of each gC signal sequence mutant. The results are shown in Fig. 3, and a quantitation of translocation efficiency, determined from the 15-min chase point, is provided in Table I. Wild-type gC was rapidly and efficiently exported; the cytoplasmic 58-kDa form was not detected even at the earliest chase times. Each of the mutants also exported a fraction of the radiolabeled gC with a wild-type rate of conversion from the 74–92-kDa form. However, virtually all of the mutants produced a precursor form of gC that remained throughout the 2-h chase period (i.e. no post-translational translocation or, in general, substantial degradation of the cytoplasmic species was observed). In some panels, a species just smaller than 74 kDa
Proline Mutants of the PRV gC Signal Sequence

Fig. 3. Pulse-chase analysis of wild-type and mutant forms of gC. PK15 cells were infected at a MOI of 10 with wild-type or mutant virus. At 4 h postinfection, cells were radiolabeled for 2 min with 100 μCi/ml [35S]cysteine in Dulbecco's modified Eagle's medium lacking fetal bovine serum. Following the addition of medium supplemented with 20-fold the normal amounts of cystine and methionine, monolayers were lysed at specific times (indicated in min across the top), and the gC species were immunoprecipitated. The immunoprecipitates were resolved by electrophoresis and visualized by autoradiography. The virus strain used to infect cells is indicated to the side of each panel, and molecular mass markers (in kDa) are indicated between panels.

TABLE I

| Strain | Genotype   | % Translocation |
|--------|------------|-----------------|
| Be     | gC L9P     | 100             |
| 543    | gC A10P    | 98              |
| 544    | gC L11P    | 90              |
| 545    | gC L12P    | 33              |
| 546    | gC L14P    | 19              |
| 547    | gC L13P    | 5               |
| 548    | gC L14P    | 5               |
| 549    | gC L13P    | 5               |
| 550    | gC L13P    | 99              |
| 551    | gC L13P    | 5               |
| 552    | gC L13P    | 90              |
| 555    | gC A10PL11A| 74              |

Table I: Summary of export efficiencies

could be seen and may have represented an under-glycosylated ER form of gC. A protein just smaller than the 58-kDa precursor accumulated throughout the 2-h chase period in PRV547-infected cells; the kinetics suggested that it was a breakdown product of the cytoplasmic gC species. Additional minor bands were also evident, but their relationship, if any, to gC was not clear. Overall, the pulse-chase results were in good agreement with the outcome of the steady-state analysis of each of the original proline substitutions, with quantitated translocation efficiencies ranging from 5–98%. However, the inclusion of strains PRV549–551 yielded additional results indicating that proline’s negative effects on signal sequence function were the result of more than just loss of hydrophobicity. Deletion of codons 10 or 11 had little effect on the translocation efficiency of gC, even though a proline substitution at position 11 resulted in nearly a 70% reduction in the process. Furthermore, the double mutant gC A10PL11A was somewhat more defective than the gC A10P mutant alone (presumably due to the reduced hydrophobicity at position 11) but remained far more translocation competent than the similarly hydrophobic gC products encoded by strains PRV545, 546, and 548.

DISCUSSION

Proline residues are rarely found in the hydrophobic cores of prokaryotic or eukaryotic signal sequences (18). This is likely due to proline’s nonhydrophobic character and its propensity to disrupt α-helices (12). Thus, it is not surprising that the introduction of proline residues into wild-type hydrophobic cores or near truncated cores results in translocation defects (Refs. 5, 16, and 19 and this study). However, it has not been clear for all such mutants whether the majority of signal sequence dysfunction should be attributed to proline’s reduction of hydrophobicity or to its disturbance of secondary structure. Our results indicate that proline affects signal sequence function via both properties, but certain positions within the gC signal sequence are more susceptible to secondary structure disruptions than others.

A good example of where proline, exclusively through its nonhydrophobic character, has a very slight effect on translocation efficiency is position 10 of the gC signal sequence. Regardless of whether the alanine at this position was replaced with proline (hydrophilicity value of 0) or simply deleted, the encoded gC polypeptide was translocated with near wild-type efficiency. A similar pattern of highly efficient translocation was found for proline substitutions at residues 9 or 13, suggesting that proline’s effect at these positions too was due to a reduction in hydrophobicity alone. However, two observations in particular ruled out the possibility that proline always exerts its effects solely by a reduction in overall signal sequence hydrophobicity. First, the deletion of codon 11 resulted in only a 10% loss of translocation efficiency for gC. In contrast, the introduction of proline at this position produced a gC species that was seven times more export defective. Second, several of our mutants had identical hydrophobicity values for their gC core structures, yet the translocation efficiencies of these mutant gCs ranged from 5 to 89%. In addition, proline substitutions at position 12, 13, or 14 resulted in similarly sized, interrupted α-helices, but gC A13P was translocated 4.5- and 18-fold more efficiently than gC L12P or gC L14P, respectively. Thus, the positional effects that we observed did not appear to be directly related to the lengths of the α-helices that resided on either side of the inserted proline. This was a consideration because Emr and Silhavy (5) isolated suppressors of an E. coli LamB signal sequence deletion mutant in which proline or glycine residues flanking the deletion were replaced with hydrophobic amino acids. They concluded that suppression was accomplished through the restoration of a suitably long α-helix to the truncated hydrophobic core. In follow up studies, Bruch and Gierasch (6) placed peptides corresponding to the LamB signal sequence of the suppressor strains in membrane mimetic environments and found that suppression efficiency correlated with the stability and not necessarily the length of the restored α-helix. In the absence of biophysical data, we cannot address the stability of the helices formed by our gC mutants.

What then may explain the positional effects on translocation that we observe for the proline insertions in the gC signal sequence hydrophobic core? Fig. 4 shows a helical wheel plot of the wild-type gC signal sequence. The α-helix has been empirically divided into two halves that reflect the severity of the translocation defect imposed by proline substitutions throughout the hydrophobic core. Coincidentally, the same axis divides the helix into its most hydrophobic (Fig. 4, left side) and least hydrophobic (Fig. 4, right side) faces. As depicted in Fig. 4, residues on the left side of the helix are much more sensitive to proline substitutions than positions on the right side. In fact, a
signal sequence with respect to SRP, thus promoting a sidedness (21), the von Heijne proposal for kinked helix-protein interaction, we would caution that if the gC signal sequence is uniformly facing away from any nearby protein sequences. However, we would be true for residues 7, 8, 11, 12, 14, and 15, whose placed them on the convex side of a kinked helix. The opposite consequence, positions 9, 10, and 13 of the hydrophobic core 4 interact with another protein, possibly the 54-kDa subunit tweenthehydrophobiccoreandSRP. ButvonHeijne addition-
certain kinks may sterically interfere with the interaction be-
 tween the hydrophobic core and SRP. Although, von Heijne previously been shown to result in a 40\% translocation efficiency (4). von Heijne (20) has noted that prolines create kinks unpaired amide and carbonyl groups of the introduced kink. Extending this proposal to the gC signal sequence, we would suggest that residues lying on the right side of the helix in Fig. 4 interact with another protein, possibly the 54-kDa subunit (SRP54) of SRP, whereas those on the left side may not. As a consequence, positions 9, 10, and 13 of the hydrophobic core would be reasonably tolerant of a proline substitution that placed them on the convex side of a kinked helix. The opposite would be true for residues 7, 8, 11, 12, 14, and 15, whose exposed amide and carbonyl groups of the convex side would face away from any nearby protein sequences. However, we should caution that if the gC signal sequence is uniformly surrounded by protein, which may very well be the case given the number of proteins implicated in ER membrane translocation (21), the von Heijne proposal for kinked helix-protein interactions is not valid (20).

These proposals raise the question of what orients the gC signal sequence with respect to SRP, thus promoting a sidedness to the hydrophobic core. This is a particularly pertinent question given the inherent flexibility of SRP that allows it to recognize a myriad of signal sequences differing in their primary amino acid sequences and in the lengths of their three structural domains. One possibility would be the hydrophilic N terminus whose positively charged residues could provide an electrostatic anchor. We conducted a preliminary experiment to test this by replacing arginine with glycine at position 6 of the gC L12P signal sequence. We have previously shown that a gC R6G mutant is exported with 94\% efficiency (11); the translocation efficiency of the gC R6GL12P double mutant was reduced by an amount equal to the combined defects of each mutant alone.\(^2\) Thus, reducing the potential electrostatic interaction of the N terminus of the signal sequence did not appear to alter the orientation of the gC hydrophobic core. In fact, if anything, the additive effect of the two mutations suggested that the hydrophilic and hydrophobic domains of the gC signal sequence function independently.

Previous analyses of sporadic disruptions of the \( \alpha \)-helices of hydrophobic cores have strongly suggested a role for secondary structure in signal sequence function. Our work represents the first systematic, genetic analysis of proline insertions in a signal sequence hydrophobic core, prokaryotic or eukaryotic. Our results indicate that the \( \alpha \)-helix of the gC signal sequence, although not strongly amphipathic, may be functionally asymmetric.

Acknowledgments—We thank Shannon Flynn for critical reading of the manuscript prior to its submission. All oligonucleotides were synthesized by the Molecular Resource Center of the University of Tennessee at Memphis.

REFERENCES

1. von Heijne, G. (1990) J. Membr. Biol. 115, 195–201
2. Gennity, J., Goldstein, J., and Inouye, M. (1990) J. Bioenerg. Biomembr. 22, 233–269
3. Bird, P., Getting, M.-J., and Sambrook, J. (1990) J. Biol. Chem. 265, 8420–8425
4. Ryan, P., Robbins, A., Whealy, M., and Enquist, L. W. (1993) Virus Genes 7, 5–21
5. Emr, S. D., and Silhavy, T. J. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 4599–4603
6. Bruch, M. D., and Gierasch, L. M. (1990) J. Biol. Chem. 265, 3851–3858
7. Robbins, A. K., Watson, R. J., Whealy, M. E., Hays, W. W., and Enquist, L. W. (1988) J. Virol. 58, 339–347
8. Ryan, J. P., Whealy, M. E., Robbins, A. K., and Enquist, L. W. (1987) J. Virol. 61, 2962–2972
9. Robbins, A. K., Whealy, M. E., Watson, R. J., and Enquist, L. W. (1986) J. Virol. 59, 635–645
10. Enquist, L. W., Keker, C. L. Jr., Robbins, A. K., Ryan, J. P., and Whealy, M. E. (1988) J. Virol. 62, 3565–3573
11. Tomlino, M., Wilkinson, K. S., and Ryan, P. (1994) J. Biol. Chem. 269, 32016–32021
12. Chou, P. Y., and Fasman, G. D. (1978) Adv. Enzymol. Relat. Areas Mol. Biol. 47, 45–148
13. Graham, F. L., and Van der Eb, A. S. (1973) Virology 52, 456–467
14. Holland, T. C., Sandri-Goldin, R. M., Holland, L. E., Marlin, S. D., Levine, M., and Gluzo, O. C. (1983) J. Virol. 46, 649–652
15. Devereux, J., Haederli, P., and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395
16. Robbins, A. K., Ryan, J. P., Whealy, M. E., and Enquist, L. W. (1989) J. Virol. 63, 250–259
17. Hopp, T. P., and Woods, K. R. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 3824–3828
18. von Heijne, G. (1985) J. Mol. Biol. 184, 59–105
19. Bedouelle, H., Bassford, P. J., Jr., Fowler, A. V., Zabin, I., Beckwith, J., and Hofnung, M. (1980) Nature 285, 78–81
20. von Heijne, G. (1991) J. Mol. Biol. 218, 499–503
21. Walter, P., and Johnson, A. E. (1994) Annu. Rev. Cell Biol. 10, 87–119

\(^2\) P. Ryan, unpublished observations.
