Deletion Mapping of the Encephalomyocarditis Virus Primary Cleavage Site

HARRY HAHN† AND ANN C. PALMENBERG*

*Institute for Molecular Virology and Department of Biochemistry,
University of Wisconsin, Madison, Wisconsin 53706*

Received 28 December 2000/Accepted 25 April 2001

The cotranslational, primary self-cleavage reaction of cardiovirus polyprotein relies on a highly conserved, short segment of amino acids at the 2A-2B protein boundary. The amino terminus of the required element for encephalomyocarditis virus has now been mapped to include Tyr₁₂₆ of the 2A protein, the 18th amino acid before the cleavage site.

Picornavirus polyproteins are processed in elaborate cascades of cotranslational and posttranslational cleavage events to produce functionally mature viral proteins and active precursors. Most polyprotein cleavages are mediated by viral proteinase, 3C^{pro}, encoded within the 3' portion of the genome, but the first or primary cleavage takes place before 3C^{pro} is synthesized (11). Picornaviruses of the *Cardiovirus* and *Aphthovirus* genera employ an unusual mechanism to achieve this scission. During translation, the C terminus of protein 2A, located in the middle region of the polyprotein, adopts an unstable conformation that affects a cotranslational break in the peptide bond between the 2A and 2B segments of the polyprotein (7, 13, 14). The core sequences adjacent this cleavage site are strongly conserved among all cardiovirus and aphthovirus strains and always contain the essential octomer, DvExNPGP. The primary cleavage is between the last two amino acids (G and P) of the motif. Site-directed mutagenesis and in vitro translation assays with cardiovirus sequences have clearly demonstrated that each of the core octomer residues contributes to optimal activity (7, 12).

The upstream and downstream sequence requirements for the primary cleavage reactions have been examined in detail for foot-and-mouth disease virus (FMDV), an aphthovirus (13, 14). The natural FMDV 2A segment is only 18 amino acids long, and the minimum primary cleavage maps to a 13-amino-acid element (12 N-terminal and 1 C-terminal to the G or P cleavage site) that includes most of this region and terminates with the essential octomer. This short segment will process cotranslationally in the absence of other FMDV sequences (6, 13, 16), although slightly longer cassettes (18 N-terminal and 1 C-terminal amino acid) may process more completely, depending upon the surrounding protein context (5).

Although cardioviruses have much larger natural 2A segments (133 to 143 amino acids) than aphthoviruses, work with *Encephalomyocarditis virus* (EMCV) and *Theiler's murine encephalitis virus* (TMEV) has likewise suggested that most of the additional 2A protein is also dispensable for primary cleavage

activity. EMCV and TMEV cleavage cassettes of 20 amino acids or longer will autoprocess in certain heterologous contexts (3, 14). However, in contrast to FMDV, shorter EMCV segments (<14 amino acids) that more closely mimic the minimal FMDV element are relatively inactive (5). To clarify the question of upstream sequences that contribute to the cardiovirus primary cleavage, we constructed a panel of nested EMCV cDNAs containing precise in-frame deletions (or substitutions) disrupting the wild-type 2A sequence near the primary cleavage site.

Engineered cDNAs. Standard recombinant methods were used (1). Nucleotide numbering was done according to EMCV-R (GenEMBL M81861). Plasmid pEC₉, containing infectious EMCV-R sequences, has been described, as has plasmid pSK+1D2A, containing cDNA from the EMCV internal ribosome entry site, linked to a viral segment encoding proteins 1C (partial), 1D, 2A, and 2B (partial) (7). To facilitate cloning, a silent C-to-A mutation, creating a *Bsr*GI site, was engineered at base 3894 according to the Transformer Site-Directed Mutagenesis System (Clontech). A panel of deletions 5' to the primary cleavage sequence was created by PCR within C3894A cDNA. The chosen primers (Fig. 1) contained the common *Bsr*GI site followed by EMCV sequences that omitted successive residues N-terminal to the cleavage site. The amplified fragments thus had deletions extending from the -23 amino acid position of 2A (the *Bsr*GI site) towards the primary cleavage site. After sequencing was done to confirm their content, the resultant cDNAs were designated Δ-21, Δ-19, Δ-17, Δ-16, Δ-15, Δ-14, Δ-12, Δ-10, and Δ-8, where "Δ" represents the deletion and "-21" (for example) refers to the wild-type amino acid within 2A that marks the C terminus of the deletion. Additionally, a variant of Δ-17, was created and called Δ-17*. Besides the directed deletion, this sequence contained a G-to-T substitution at base 3919, replacing the wild-type Gly residue at -16 with Val.

Processing of mutant cleavage sequences. The cleavage activities of the recombinant proteins were assayed in the absence of other viral processing events by translation of transcript RNAs in reticulocyte lysates (2). Conversion of precursor 1C*1D2A2B* (59 kDa) into product 1C*1D2A (49 kDa) was the benchmark for primary cleavage (Fig. 2A). The small 2B* product (10 kDa) was not retained by the gels. Typical reactions were for 1 h at 30°C, but longer incubation (up to 4 h) or

* Corresponding author. Mailing address: Department of Biochemistry, 433 Babcock Dr., Madison, WI 53706. Phone: (608) 262-7519. Fax: (608) 262-6690. E-mail: acpalmen@facstaff.wisc.edu.

† Present address: 2440 Barrington Ave. #309, Los Angeles, CA 90064.

FIG. 1. Deletions N-terminal to the primary cleavage. PCR reactions with a series of nested 5' primers (33 bases each) created defined cDNA deletions upstream of the EMCV primary cleavage region. The 5' primer for each reaction, Δ-8 to Δ-21, is illustrated. A common 3' primer (not shown) complemented viral bases 4354 to 4385. The resultant amplicons were digested with *Bsr*GI (encoded in the 5' primer) and *Nsi*I and then used to replace the *Bsr*GI to *Nsi*I fragment of pSK+1D2A(C3894A).

shorter incubation (30 min) did not alter the relative cleavage patterns (not shown), as is consistent with other reports for cotranslational primary reactions in cardioviruses (7, 14).

The silent C3984A mutation in the parental cDNA had no effect on the extent of cleavage (lane C3984A) relative to a wild-type sequence. The two smallest deletions (Δ-21 and Δ-19) also cleaved as well as the wild type, with the predominant band being that of the cleaved form. Conversely, the largest deletions (Δ-14, Δ-12, Δ-10, and Δ-8) produced proteins that were poorly cleaved, and most of the product was in the uncleaved form.

The proteins between the longest and shortest extremes had various activities. The Δ-15 and Δ-17 proteins were only partially active, while the intermediate deletion, Δ-16, cleaved nearly as efficiently as did the wild type. These results were highly reproducible (not shown) and indicated that the sequences upstream of these deletions somehow contributed to a more effective cleavage environment when in a Δ-16, rather than in a Δ-17, context. Indeed, the Δ-16 sequence (but not Δ-17) shared a fortuitous Tyr residue at -18, that was displaced from the -25 position of the wild-type sequence. Tyr at -18 is common also to the highly active Δ-19, Δ-21, and wild-type proteins. The patterns suggest the boundary of the EMCV cleavage element must be at or near this location.

Primary cleavage model. The primary cleavage reaction in aphthoviruses and cardioviruses is cotranslational and autocatalytic. It is not mediated by exogenous host proteases or other viral proteases, including 3C^{Pro} (5, 7, 12, 14), which instead has a preferred specificity for Gln-Gly, Gln-Ser, or Glu-Gly substrates that can center within its 5- to 6-amino-acid binding pocket (4, 8, 10). The minimum aphthovirus primary cleavage element is encoded by a peptide segment 12 to 13 amino acids in length (6, 13, 14, 16). Our results support a similar situation in the cardiovirus system, in that a short, contiguous peptide sequence at the carboxyl end of 2A was necessary to produce

this cleavage, and the segment functioned regardless of deletions in the upstream protein context. Specifically, the cleavage occurred to completion when only when 18 amino acids N-terminal to the primary cleavage site were maintained as wild-type EMCV. The C-terminal site requirements, as previously mapped (7), are known to include the singlet Pro at the end of the NPG/P segment, which begins the 2B sequence. Thus, the cardiovirus primary cleavage element probably extends for a total of 19 amino acids beginning with 2A Tyr₁₂₆ and, if left intact, can operate during translation as a self-contained, autocatalytic cleavage cassette.

Deletions which impinged upon this segment provide additional information about the required sequences. From size alone, for example, we expected Δ-16 to cleave more poorly than Δ-17. However, Δ-16 cleaved nearly to completion, while Δ-17 cleaved only partially, despite having an additional residue of wild-type protein (Fig. 2B). The wild-type 2A segment has a Tyr at the -18 position and, due to the deletion method employed, Δ-16 had a similar Tyr at position -18, displaced from its wild-type position at -25. The Δ-17 construct, on the other hand, had Arg at its -18 location. Moreover, we have previously reported that another EMCV deletion mutant, Δ2A* which is missing 2A residues 6 to 125, undergoes the primary cleavage reaction with an efficiency similar to that of the wild type (15). Effectively, Δ2A* is an authentic Δ-18 deletion (Fig. 2C), albeit with more distant upstream sequences linked to the cleavage element. The activity of this mutant was consistent with our panel, in that every tested construction with a substitution at -18 had a deleterious effect on the cleavage efficiency. Therefore, Tyr at -18 correlates most closely with the activity of the primary mechanism and probably marks the true N-terminal boundary for the EMCV primary cleavage element.

Viral sequence comparisons and the data presented here converge in a consistent model for EMCV primary cleavage

FIG. 2. Primary cleavage activity of deletion sequences. (A) Plasmid pSK+1D2A (0.5 to 1 μ g) containing wild-type or mutant cDNA sequences was used to program coupled transcription-translation reactions (12.5 μ l; Novagen Single Tube Protein System with [35 S]methionine) (2). After 1 h at 30°C, loading buffer (12.5 μ l) was added, and the samples were boiled and then fractionated through a 5 to 20% sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel. A molecular weight marker is in the leftmost lane, followed by samples programmed with the indicated cDNAs. (B) Protein sequence of the deletion constructs near the primary cleavage site are shown. Upper-case letters highlight amino acids that share the same position and identity as the wild-type sequence. The conversion of precursor (1C*1D2A2B) to product (1C*1D2A) was quantitated from the autoradiograph in panel A using ImageQuant software. The cleavage column gives the percentage of total integrated intensity in the product band relative to the sum of the precursor and product bands. The relative values were highly repeatable over a range of experiments. (C) Plasmid Δ-2A*, missing 120 codons from the EMCV 2A region, has been described elsewhere (15). It has a very active primary cleavage function (94%) and is included here for comparison with the current deletion panel.

requirements (Fig. 3). Surprisingly, the 2A proteins of natural EMCV strains share the lowest degree of amino acid identity (77.6% average) of any region in the viral polyproteins (93.7% average), value lower even than that of the combined capsid immunogenic sites (88% average). Yet sequence variations within the 2A primary cleavage elements are exceedingly rare (98.2% identity) and were observed only at the -6 and -13 positions (I₁₃₈V and A₁₃₂S, respectively) of mengovirus. Deliberate substitution at the -17 position (G₁₂₉V) in the Δ-17* EMCV cDNA produced a less active fragment than the Δ-17 mutation (Fig. 2A), again implicating the wild-type sequence as the preferred format in this position.

The striking periodicity of aliphatic residues within this region of the EMCV, TMEV, and FMDV 2A proteins has been previously observed (5) and suggests the segments should be modeled in helical conformations (Fig. 3). For EMCV, a four-turn helix initiating with Tyr₁₂₆ (at -18) and ending with T₁₄₀ (-4) would have one face composed of large, hydrophobic residues (Tyr, -18; Tyr, -15; Phe, -14; Leu, -11; Leu, -10; Ile, -6) with strong stacking potential and another face of

highly charged residues (D, -12; H, -8; D, -7; E, -5). The 2A proteins of the related theiloviruses can likewise be modeled with similar conserved, four-turn helices (beginning at a His -18) that share many identities (11 of 18) with the EMCV residues and also their relative distributions. Outside of this primary cleavage region, the theilovirus and EMCV share less than 28% identity in their 2A proteins (50.3% overall for their polyproteins). Since hydrophobic faces usually indicate an affinity for other hydrophobic surfaces, perhaps in this case on the translating ribosome, the cleavage cassette structure must help provide a framework or correct environment for the unusual PGP triplet to become susceptible to proteolysis. The electrophilic attack on the cleaved bond probably comes from an internal source, such as the amide group of the obligate Asn residue at -3 (13).

FIG. 3. Primary cleavage models. Helical net representations (DNASTAR Protean software package) illustrate the character of residues near the NPGP cleavage sites of viruses from the four taxonomic clades known to undergo autocatalytic, cotranslational reactions. The amino acids are arranged as they might appear in an alpha helix (110°). The dark circles highlight large aliphatic residues. The gray circles highlight charged residues. Numbering is relative to the cleaved NPG/P site (open-faced characters). Sequence variations in other known viral strains are indicated (small letters), as are the minimum primary cleavage elements for FMDV (12 amino acids of 2A [13]) and EMCV (18 amino acids of 2A). The complete sequence alignments for EMCV (6 strains), TMEV (4 strains), and FMDV (32 strains) picornaviruses are available on the internet (www.bocklabs.wisc.edu/acp). Human, bovine, and porcine rotavirus sequences are from accession numbers AJ132203, P34717, and P27586, respectively.

Comparison of the cardioviruses with the only other viral sequences known to cleave in a similar manner, the FMDV and type-C rotavirus (9), clearly indicates a similar strong proclivity towards clustering hydrophobic residues on the face of a putative helix (Fig. 3). For the shorter primary cleavage cassettes of FMDV, two or three helical turns (−12 to −4) appear to suffice for autoproteolytic activity (13), although four full turns, initiating at −18, may still be the preferred conformation (5). In contrast to the cardioviruses, the naturally truncated 2A region within FMDV may have contributed to the evolution of partial reactivity within shorter segments. The carboxyl border of the type-C rotavirus cleavage element has yet to be mapped genetically (9). However, similarities with the other sequences again suggest a likely boundary at the −17 (Phe) or −18 (Lys) position.

We thank J.-Y. Sgro for assistance with protein modeling.

This work was supported by National Institutes of Health grants AI-17331 to A.C.P. and training grant GM-07215 to H.H.

REFERENCES

1. Ausubel, F., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1993. Current protocols in molecular biology. John Wiley & Sons, Inc., New York, N.Y.
2. Barber, G. N., M. Wambach, S. Thompson, R. Jagus, and M. G. Katze. 1995. Mutants of the RNA-dependent protein kinase (PKR) lacking double-stranded RNA binding domain I can act as transdominant inhibitors and induce malignant transformation. *Mol. Cell. Biol.* **15**:3138–3146.
3. Batson, S., and K. Rundell. 1991. Proteolysis at the 2A/2B junction in Theiler's murine encephalomyelitis virus. *Virology* **181**:764–767.
4. Bergmann, E. M., S. C. Mosimann, M. M. Chernaia, B. A. Malcolm, and M. N. G. James. 1997. The refined crystal structure of the 3C gene product from hepatitis A virus: specific proteinase activity and RNA recognition. *J. Virol.* **71**:2436–2448.
5. Donnelly, M. L., D. Gani, M. Flint, S. Monaghan, and M. D. Ryan. 1997. The cleavage activities of aphthovirus and cardiovirus 2A proteins. *J. Gen. Virol.* **78**:13–21.
6. Haghighat, A., Y. Svitkin, I. Novoa, E. Kuechler, T. Skern, and N. Sonenberg. 1996. The eIF4G-eIF4E complex is the target for direct cleavage by the rhinovirus 2A proteinase. *J. Virol.* **70**:8444–8450.
7. Hahn, H., and A. C. Palmenberg. 1996. Mutational analysis of the encephalomyocarditis virus primary cleavage. *J. Virol.* **70**:6870–6875.
8. Hall, D. J., and A. C. Palmenberg. 1996. Mengo virus 3C proteinase: recombinant expression, intergenus substrate cleavage and localization in vivo. *Virus Genes* **13**:99–110.
9. Langland, J. O., S. Pettiford, B. Jiang, and B. L. Jacobs. 1994. Products of the porcine group c rotavirus NSP3 gene bind specifically to double-stranded RNA and inhibit activation of the interferon-induced protein kinase PKR. *J. Virol.* **68**:3821–3829.
10. Matthews, D. A., W. W. Smith, R. A. Ferre, B. Condon, G. Budahazi, W. Sisson, J. E. Villafranca, C. A. Janson, H. E. McElroy, C. L. Gribskov, and S. Worland. 1994. Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. *Cell* **77**:761–771.
11. Palmenberg, A. C. 1990. Proteolytic processing of picornaviral polyprotein. *Annu. Rev. Microbiol.* **44**:603–623.
12. Palmenberg, A. C., G. D. Parks, D. J. Hall, R. H. Ingraham, T. W. Seng, and P. V. Pallai. 1992. Proteolytic processing of the cardioviral P2 region: 2A/2B cleavage in clone-derived precursors. *Virology* **190**:754–762.
13. Ryan, M. D., and J. Drew. 1994. Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. *EMBO J.* **13**:928–933.
14. Ryan, M. D., A. M. Q. King, and G. P. Thomas. 1991. Cleavage of FMDV polyprotein is mediated by residues located within a 19 amino acid sequence. *J. Gen. Virol.* **72**:2727–2732.
15. Svitkin, Y. V., H. Hahn, A. C. Gingras, A. C. Palmenberg, and N. Sonenberg. 1998. Rapamycin and wortmannin enhance replication of a defective encephalomyocarditis virus. *J. Virol.* **72**:5811–5819.
16. van Kuppeveld, F. J. M., J. G. J. Henderop, R. L. L. Smeets, P. H. G. M. Willems, B. P. M. Dijkman, J. M. D. Galama, and W. J. G. Melchers. 1997. Cocksakievirus protein 2B modifies endoplasmic reticulum membrane and plasma membrane permeability and facilitates virus release. *EMBO J.* **16**:3519–3532.