JOURNAL OF VIROLOGY, June 1984, p. 957–959
0022-538X/84/060957-03$02.00/0

Vol. 50, No. 3

Systematic Nomenclature of Picornavirus Proteins

ROLAND R. RUECKERT[1]* AND ECKARD WIMMER[2]

Biophysics Laboratory, The Graduate School, and Biochemistry Department, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706,[1] and Department of Microbiology, School of Medicine, State University of New York, Stony Brook, New York 11794[2]

Received 20 January 1984/Accepted 1 March 1984

An easily learned convention for systematizing the nomenclature of picornavirus proteins is described. The convention is based upon an idealized map, called the L434 diagram, of the picornavirus polyprotein.

The cleavage pathway of picornavirus proteins is complex and varies somewhat with the virus species. Moreover, electrophoretic mobility on sodium dodecyl sulfate-containing polyacrylamide gels, the criterion originally used to name the proteins (12), varies not only with virus species (1) but also with variations in technique required to resolve the proteins (6, 9, 13). These circumstances, together with the existence of a large number of laboratories working independently on different picornaviruses, have led to a variety of nomenclatures, which hampers communication between research workers and unnecessarily complicates the teaching of picornavirology.

This problem was recently addressed at the third meeting of the European Study Group on the Molecular Biology of Picornaviruses, held 5 to 10 September 1983 at Urbino, Italy. It was agreed at this meeting that a systematic nomenclature is needed. It was further agreed, by unanimous vote, to adopt a convention, hereafter called the L434 convention, based upon an idealized map of the picornavirus polyprotein (Fig. 1). The purpose of this communication is to describe the L434 convention and to summarize recommendations adopted at the Urbino meeting.

L434 diagram. The polyproteins of poliovirus, encephalomyocarditis (EMC) virus, foot-and-mouth disease virus (FMDV), and probably also human rhinovirus 1A can be accommodated into the pattern L-ABCD-ABC-ABCD (L434 diagram), in which L is a leader protein and A,B,C, . . . represent end products of a capsid piece (P1), a midpiece (P2), and a right piece (P3) (Fig. 1). The poliovirus polyprotein appears to lack a leader protein whereas that of FMDV appears to lack a protein 2B in the midpiece; in addition, FMDV encodes three VPg proteins, whereas other picornaviruses appear to encode only one. The L434 diagram provides a basis for systematically classifying and naming virtually all of the known proteins (Table 1). A few exceptions will be discussed below.

Leader region. The leader is defined as that region of the polyprotein preceding the capsid precursor protein. Knowledge of the synthesis and processing of leader proteins is still too scant to discern homologies; thus L protein nomenclature remains to be defined. The symbol L was adopted, in preference to P0, to minimize possible confusion with coat protein, VP0.

P1 region. The P1 region is defined as that part of the polyprotein which generates coat proteins VP4, VP2, VP3, and VP1. Thus FMDV P91 (Plum Island nomenclature), called P88 by the Pirbright group (actual molecular weight determined by nucleotide sequencing is about 80,000), would be called P1 or 1ABCD or 1, whereas FMDV P72 would be called 1ABC. Once the basic pattern is committed to memory one hardly needs a figure to imagine the map position of the protein. Given that the L434 name of poliovirus protein 3c is 1BCD, for example, its coordinates on the L434 diagram are easily visualized.

P2 region. The midpiece, P2, is defined as the region between P1 and P3. The P2 regions of poliovirus and EMC virus are cleaved into three proteins, whereas that of FMDV appears to generate only two (4), due apparently to deletion of protein 2B. If so, FMDV P56, also known as P52, would be called 2AC. This makes the deletion immediately evident. Should the deletion ultimately prove artifactual there would be no problem in reverting to standard description. Protein 3b of poliovirus is described as P2 or 2ABC, whereas its protein 5b is 2BC and X becomes 2C.

P3 region. The P3 region is the entire right piece of the polyprotein beginning with the protein 3A flanking VPg (3B) on the left. Thus, the L434 name of poliovirus protein 9 and EMC virus protein H is 3AB. The multiple VPg's of FMDV are also readily accommodated in this system. Thus P81 (also called P72), carrying zero, one, two or three VPg's, becomes 3CD, 3BCD, 3BBCD, or 3BBBCD, respectively.

Overlap proteins. Certain proteins overlap the borders defined by the L, P1, P2, and P3 regions. Thus the L434 name of EMC virus protein pre-A, also known as A1, is L-1-2A, whereas that of protein A is 1-2A. Similarly, unusual proteins such as 3b/9 and X/9 of poliovirus, which overlap the P2 and P3 regions, become 2BC-3AB and 2C-3AB, respectively.

Alternative cleavage proteins. One class of proteins not accommodated by the simple L434 diagram is the set represented by poliovirus P3 proteins 6a and 6b. They are products of a tyrosine-glycine cleavage site located in P3 protein 2 (see Fig. 7 in reference 9). Homologous proteins are known for HRV-1A and FMDV. It is proposed that these alternative cleavage products of region 3CD be named 3C′ (6a of poliovirus, P56c of FMDV, r39a of HRV-1A) and 3D′ (6b of poliovirus, P56a of FMDV, and r39b of HRV-1A), respectively. Other proteins which do not fit the L434 pattern will retain their trivial names. Examples are poliovirus P3 proteins 4a, 5a, and 7d (see Fig. 7 of reference 9).

Trimmed or modified proteins. Trimming, such as removal of amino acids from an end of a protein, or modification, such as removal of an amino group, phosphorylation, or methylation, would be described as a modified form of a standard protein. For example EMC protein epsilon-1 (11) is a modified form of protein 1AB.

* Corresponding author.

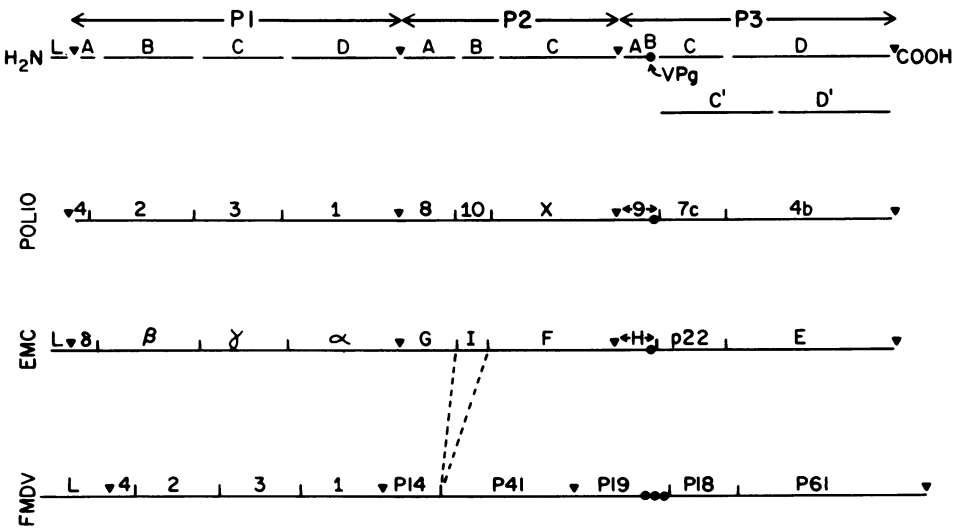

FIG. 1. L434 diagram (top) illustrating organization of the idealized picornavirus polyprotein. Filled triangles mark junctions of the leader (L), capsid piece (P1), midpiece (P2), and right piece (P3). Pieces P1, P2, and P3 contain four, three, and four (counting A and B as separate segments) segments, respectively. C' and D' refer to alternative cleavage proteins (see text). The polyprotein of poliovirus lacks a leader piece (7), whereas the polyprotein of FMDV appears to lack a segment in the P2 piece; it also contains three, rather than one, VPg sequences (4).

Gene functions. The terms *vpg*, *pro*, *pol*, and *ncp* identify genetic loci responsible for synthesis of VPg, protease, polymerase, and nucleocapsid proteins, respectively. These terms are readily incorporated into L434 nomenclature; thus: EMC p22 = $3C^{pro}$; polio 4b = $3D^{pol}$; VPg = $3B^{vpg}$; P1 = 1^{ncp}.

This feature is useful for indicating functions associated with specific proteins and would permit easy distinction between a hypothetical polymerase associated with 2C, for example, from the polymerase now known to be associated with 3D.

Recommendations. It is recognized that the L434 nomenclature cannot be applied in cases where the mapping position of a protein has not yet been established. There are also cases when authors prefer older names to describe special proteins, such as ambiguous forms of poliovirus coat proteins. For these reasons, it was agreed at Urbino that abandonment of old nomenclatures would not be mandatory. That is, an author may choose to retain an old nomenclature provided that the L434 names, when known, are identified within the publication. It was also decided that the widely used VP0, 1, 2, 3, and 4 nomenclature for coat proteins be retained. In addition, the L434 convention does not address proteins which may be encoded by cistrons outside the polyprotein cistron.

TABLE 1. Relationship of picornavirus proteins

Previous name					L434 name
Poliovirus		EMC virus[c]	FMDV		
Reference 12[a]	Reference 7[b]		Reference 4[d]	Reference 3[e]	
		p12/p14	P16	p16/20a	L
		pre-A or A1			L-1-2A
		A			1-2A
1a	P1-la	B	P91	P88	1 or P1
			P56	P52	2AC
3b	P2-3b				2 or P2
5b	P2-5b				2BC
8	P2-8	G	P14	P20c	2A
10	P2-10	I	Deleted?		2B
X	P2-X	F	P41	P34	2C
7a	P2-7a				2AB
	X/9				2C-3AB
1b	P3-1b	C	P102	P100	3 or P3
9	P3-9	H	P19	P14	3AB
7c	P3-7c	P22	P18	P20b	3C
4	P3-4b[f]	E	P61	P56a	3D
2	P3-2	D	P81	P72	3CD
			VPgP81		3BCD
VP0	VP0	ε	VP0	VP0	1AB
VP1	VP1	α	VP1[g]	VP1	1D
VP2	VP2	β	VP2	VP2	1B
VP3	VP3	γ	VP3	VP3	1C
VP4	VP4	δ	VP4	VP4	1A

[a] Summers et al. (12) as modified by Butterworth (1).
[b] Kitamura et al. (7).
[c] See references 2, 5, 10, and 11.
[d] Plum Island nomenclature (4).
[e] Pirbright nomenclature (3).
[f] Also known as NCVP4, p63, poly(U) polymerase, and RNA polymerase.
[g] Formerly called VP3.

LITERATURE CITED

1. **Butterworth, B. E.** 1973. Comparison of the virus-specific polypeptides of encephalomyocarditis virus, human rhinovirus 1A and poliovirus. Virology **56**:439–453.
2. **Campbell, E. A., and R. J. Jackson.** 1983. Processing of the encephalomyocarditis virus capsid precursor protein studied in rabbit reticulocyte lysates incubated with *N*-formyl-[^{35}S]methionine-tRNA$_f^{Met}$. J. Virol. **45**:439–441.
3. **Doel, T. R., D. V. Sangar, D. J. Rowlands, and F. Brown.** 1978. A reappraisal of the biochemical map of foot-and-mouth disease virus RNA. J. Gen. Virol. **41**:395–404.
4. **Grubman, M. J., B. H. Robertson, D. O. Morgan, D. M. Moore, and D. Dowbenko.** 1984. Biochemical map of polypeptides specified by foot-and-mouth disease virus. J. Virol. **50**:579–586.
5. **Kazachkov, Y. A., T. V. Chernovskaya, E. Y. Siganova, Y. V. Svitkin, T. Y. Ugarova, and V. I. Agol.** 1982. Leader polypeptides encoded in the 5'-region of the encephalomyocarditis virus. FEBS Lett. **141**:153–156.
6. **Kew, O. M., M. A. Pallansch, D. R. Omilianowski, and R. R. Rueckert.** 1980. Changes in three of the four coat proteins of oral

poliovaccine strain derived from type 1 poliovirus. J. Virol. **33**:256–263.

7. **Kitamura, N., B. L. Semler, P. G. Rothberg, G. R. Larsen, C. J. Adler, A. J. Dorner, E. A. Emini, R. Hanecak, J. J. Lee, S. Van der Werf, C. W. Anderson, and E. Wimmer.** 1981. Primary structure, gene organization and polypeptide expression of poliovirus RNA. Nature (London) **291**:547–553.
8. **McLean, C., T. J. Matthews, and R. R. Rueckert.** 1976. Evidence of ambiguous processing and selective degradation in the noncapsid proteins of rhinovirus 1A. J. Virol. **19**:903–914.
9. **Pallansch, M. A., O. M. Kew, B. L. Semler, D. R. Omilianowski, C. W. Anderson, E. Wimmer, and R. R. Rueckert.** 1984. Protein processing map of poliovirus. J. Virol. **49**:873–880.
10. **Rueckert, R. R., T. J. Matthews, O. M. Kew, M. Pallansch, C. McLean, and D. Omilianowski.** 1979. Synthesis and processing of picornaviral polyprotein, p. 113–125. *In* R. Perez-Bercoff (ed.), Molecular biology of picornaviruses. Plenum Publishing Corp., New York.
11. **Rueckert, R. R., A. C. Palmenberg, and M. A. Pallansch.** 1980. Evidence for a self-cleaving precursor of virus-coded protease, RNA-replicase and VPg, p. 263–275. *In* G. Koch and D. Richter (ed.), Biosynthesis, modification and processing of cellular and viral polyproteins. Academic Press, Inc., New York.
12. **Summers, D. F., J. V. Maizel, and J. E. Darnell.** 1965. Evidence for virus-specific noncapsid proteins in poliovirus-infected cells. Proc. Natl. Acad. Sci. U.S.A. **54**:505–513.
13. **Swaney, J. B., G. F. Van de Woude, and H. L Bachrach.** 1971. Sodium dodecyl sulfate-dependent anomalies in gel electrophoresis: alterations in the binding patterns of foot-and-mouth disease virus polypeptides. Anal. Biochem. **58**:337–346.