Journal of Cellular Biochemistry 33:191–198 (1987)
Proteases in Biological Control and Biotechnology 175–182

Picornaviral Processing: Some New Ideas

Ann C. Palmenberg

Biophysics Laboratory of the Graduate School, and Biochemistry Department of the College of Agriculture and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706

Mature picornaviral proteins are derived by progressive, post-translational cleavage of a giant precursor polyprotein. At least three viral-encoded proteolytic activities are involved in the processing. The first cleavage takes place while the polyprotein is still nascent on a ribosome. In poliovirus, this event is probably catalyzed by peptide 2A, a protein from the middle portion of the genome. Most subsequent processing is effected by viral protease 3C, a thiol-type enzyme, responsible for eight to ten self-cleaving and autocatalytic reactions within the polyprotein. The final proteolytic processing event, maturation of the VPO peptide, may occur by a novel, autocatalytic, serine-type mechanism, where viral RNA serves as proton-acceptor during the cleavage reaction.

Key words: processing cascade, VPO maturation processing, RNA-induction, hydrazine induction, picornaviral protease 3C, picornaviral protease 2A, picornaviral maturation processing

Although physically among the smallest of positive-strand RNA viruses, the picornaviruses are of major historic, economic, and medical importance. The family contains a diverse variety of highly virulent human and animal pathogens, which traditionally are subdivided into four groups (or genera) on the basis of physical properties of the virions (sedimentation coefficient, pH stability, etc). The subgroups include the enteroviruses (eg, polio, hepatitis A, and coxsackie viruses), cardioviruses (eg, encephalomyocarditis [EMC] and Mengo virus), aphthoviruses (foot-and-mouth disease [FMD] virus), and rhinoviruses (human and bovine).

In spite of the disparate afflictions caused by these agents, recent advances in molecular virology, nucleotide sequencing, and X-ray crystallography have shown that all picornaviruses share remarkable similarity in their particle structure and genome organization. The virions contain a single-stranded RNA genome enclosed in a protein capsid shell. The capsids are composed of 60 subunits, each of which contains four nonidentical polypeptide chains [1,2]. The RNAs vary in length from 7,102 bases (rhinovirus-2) to 8,450 bases (approximate size of FMDV-01K), not including the poly(A) tail (50–150 bases) [3,4]. Cardio- and aphthoviruses are distin-

Received May 5, 1986, revised and accepted October 14, 1986.

guished by the presence of a 5′-proximal poly(C) tract, whose length (50–150 bases) and exact location relative to the 5′ end of the RNA (150–400 bases) vary with different isolates of virus [5–7]. (For reviews of comparative genome structure, see refs. [8,9].)

Translation of viral RNA proceeds primarily from a single strong initiation site and produces a giant precursor polypeptide (MW 250,000), representing most of the theoretical coding capacity of the genome. The polyprotein is processed in a remarkable series of proteolytic cleavage steps to yield mature virion capsid proteins, as well as other viral proteins of a noncapsid nature. The map in Figure 1 illustrates the characteristic processing relationships [10]. Among the viruses, the order and rate of the processing cascade is very similar, and it has been proposed that sequential proteolysis may help regulate all phases of the picornaviral life cycle, from genome replication to new particle morphogenesis [11,12]. At least three viral-encoded proteolytic activities are involved in these cleavages. While many details remain to be determined, it is now possible to assign most cleavage responsibilities to specific viral peptides.

The very first cleavage within a polyprotein takes place while the peptide is still nascent on a ribosome and before the sequences encoding the P3 region of the genome have been translated. Originally it was proposed that an unidentified "host"-derived enzyme was responsible for this event [13], but recent work with poliovirus clones has shown that viral peptide 2A (MW 15,000) has proteolytic activity and is probably the agent that catalyzes the first cleavage within this virus (at the P1/2A site). Analogy with other protease amino acid sequences suggest that poliovirus (coxsackie and rhinovirus?) 2A protease may belong to the cysteine-class of enzymes [14]. The demonstrable substrate specificity of 2A seems limited to its own polyprotein, though cross-reactivity among different viruses has not yet been rigorously examined.

Two lines of evidence indicate that cardioviral and aphthoviral 2A peptides may not be functionally equivalent to that of poliovirus, and that analogous nascent cleavage activity for these viruses may be located elsewhere within their genomes. First, the 2AB region of aphthoviruses is very much shorter than in entero- or rhinoviruses, and sequence homologies strongly suggest that the missing or deleted segment(s) correspond to peptide 2A [4,8–10]. Second, the primary (nascent) cleavage within cardioviral (and probably aphthoviral) polyproteins occurs at a different site than in poliovirus (2A/2B, rather than P1/2A) [13,15]. We suspect that the leader

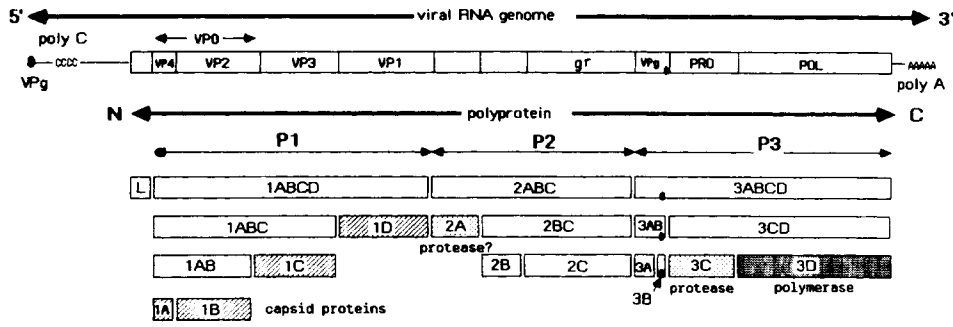

Fig. 1. Structure of the picornaviral genome. Viral RNA, encoded peptides, and their protein precursors are illustrated schematically. Peptide nomenclature is according to standard convention [10].

peptide (L), and/or peptide 2B are alternative candidates for the 2A-equivalent function in these viruses, though this hypothesis remains to be tested.

Most of the remaining cleavages within a picornaviral polyprotein are carried out by viral peptide 3C, an extraordinary cysteine-type protease (MW 20,000), capable of multiple self-cleaving and autocatalytic reactions within the viral precursors. Protease 3C catalyzes cleavages within the P1, P2, and P3 peptides to produce proteins L, 1AB, 1C, 1D, 2A, 2B, 2C, 3A, 3B, and 3D, as well as 3C itself [12,16–18]. It is not clear how the 3C enzyme recognizes or selects the appropriate cleavage sites within its polyprotein to produce the cascade. Aside from the processing sites within the viral polyprotein itself, the enzyme has no other known natural substrates. Our preliminary mixing experiments with polio and EMC, or FMDV and EMC have failed to detect cross-processing, using peptides synthesized in cell-free extracts or expressed from recombinant cDNA clones (ACP, unpublished observation). The specificity is somewhat surprising in view of the similarity among the cleavage sequences (Fig. 2). Most 3C-directed cleavages occur between glutamine (or gluta-

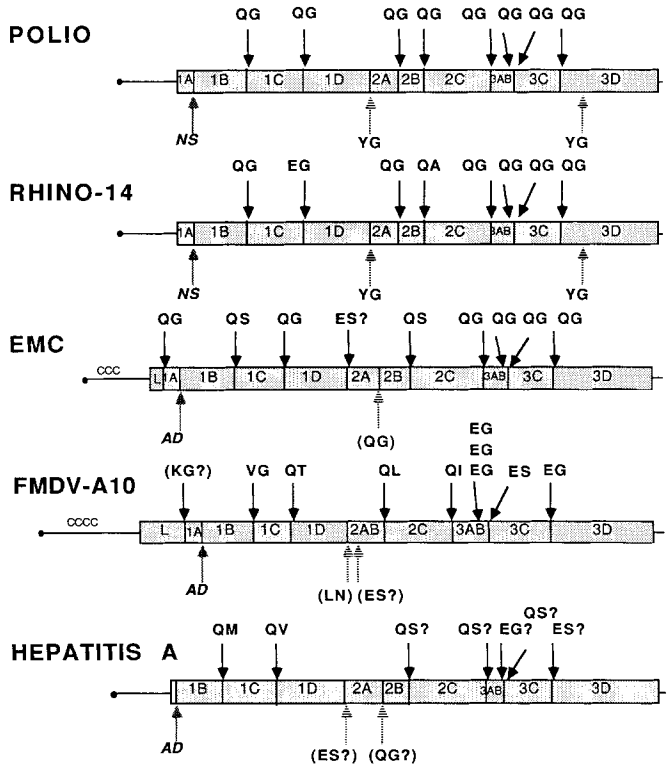

Fig. 2. Cleavage sequences within viral polyproteins. Representative picornaviral genomes (polio, rhino-14, EMC, FMDV-A10, and hepatitis A) are presented schematically. Proteolytic cleavage occurs between the indicated amino acid pairs (4,20,23,24,26,32). FMDV cleavages between tandemly linked 3B proteins occur at identical EG dipeptides (20–22). Question marks denote sites whose sequence is not precisely known, or whose position is suggested by the author on the basis of homologous protein alignments. Protease 3C catalyzed sites are shown above each genome. Protease 2A or VPO (1AB) sites are shown below the genomes. Sites for which the proteolytic agent has not been specifically identified are enclosed in parentheses.

mate) and glycine (or serine) dipeptide pairs. The surrounding sequences (not shown) have a high frequency of helix-breaking residues, such as proline and threonine, which leads to the speculation that protein secondary structure may play a role in cleavage site identification. (Cleavage sequences are reviewed in ref. [4]; see also refs. [19–32].

Because of its obvious importance in the viral processing scheme, the 3C enzyme is the subject of several current scientific investigations. The enzyme encoded within EMC virus has been isolated in an active form from virus-infected HeLa cells. DNA plasmid clones capable of expressing EMC 3C peptide in bacterial or eukaryotic cell-free translation reactions have also been developed [33]. These systems are being used to produce and examine the unique specific cleavage mechanism of the 3C protease.

The final cleavage within picornaviral polyproteins, maturation processing of the 1AB peptide (also called VPO), is not catalyzed by 3C or by 2A. These reactions are observed in vivo only during the final stages of virion morphogenesis and are probably concomitant with RNA association with large capsid assembly structures [34]. Figure 3 diagrams the capsid region processing cascade and the parallel capsid assembly intermediates that are generated at each cleavage step. An agent responsible for VPO proteolysis has never been isolated, but recent crystallographic resolution of rhino and polio capsid protein structures has allowed renewed speculation on the nature of this event [1,2,35,36].

Within each capsid subunit of these viruses, the side chain of a VP2 (1B) serine residue is positioned near, and points directly toward the carboxy-end of its homologous VP4 (1A) peptide chain. In rhinovirus, these amino acids are linked by a hydrogen bond [36]. Aspartate and arginine residues are located near the serine, and an adjacent beta-pleated sheet (beta-A1/beta A2) helps orient and stabilize the entire region. Figure 4 illustrates the relevent portion of the rhinovirus VP4-VP2 structure.

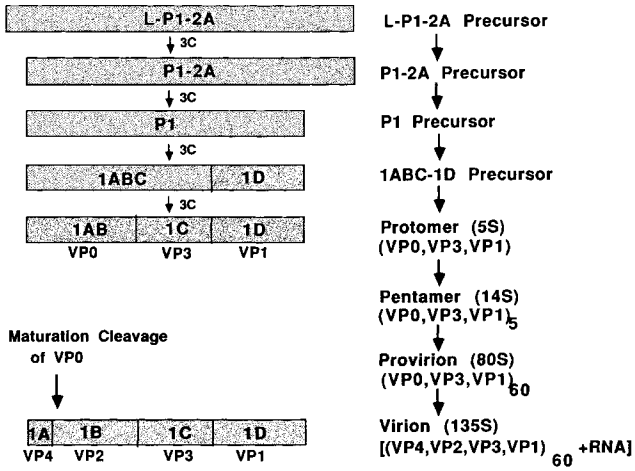

Fig. 3. Capsid region processing and assembly. The capsid region processing cascade and accompanying morphogenic assembly steps are illustrated schematically. The "S" values (5, 14, 80, and 135) refer to approximate sedimentation coefficients for particular structures [8,9]. The L-P1-2A → P1-2A step occurs only in cardio- and aphthoviruses. Cleavage of P1 from the polyprotein is catalyzed by 2A in entero- and rhinoviruses.

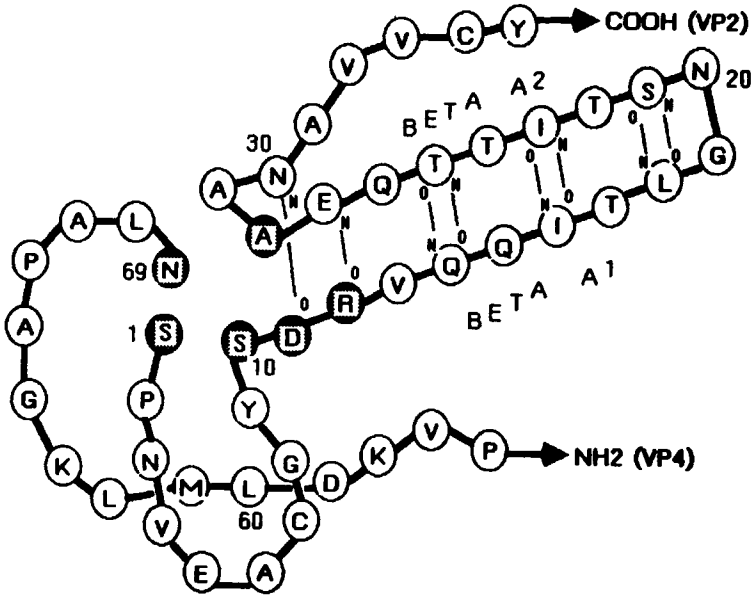

Fig. 4. Structure of the VP4/VP2 cleavage site. The amino acid sequence and probable conformation near the cleavage site between VP4 and VP2 of rhino-14 are depicted schematically. The complete structure is described in ref. [2], the rhino-14 sequence is from ref. [32].

The spacial arrangement evokes analogy to the active site conformations found in normal serine proteases such as trypsin and chymotrypsin. (For reviews of serine protease structures, see refs [37–39].) However, while these enzymes usually also have a proton-accepting histidine residue located near the active serine, this feature is completely missing from the viral structures. The histidine-equivalent position in the serine protease-like active site is instead occupied by viral RNA, packaged inside the capsid (that area behind the Fig. 4 structure). We postulate that during virion assembly, the RNA may act in place of histidine, accepting protons and thereby helping to catalyze VPO cleavages [2,35,36].

The phenomenon of VPO cleavage is believed to be universal among picornaviruses. Even hepatitis A, in which the putative VP4 peptide is only 23 amino acids long, probably still undergoes maturation processing [19,40]. A proposed mechanism for VPO processing must therefore be compatible with and accountable for any differences (sequence or structural) among the viruses. The protein sequences of many picornaviruses are now available [19–32]. The strains share a striking degree of sequence and structural similarity throughout the entire capsid region, but the homology is particularly strong near the amino-end of VP2. Figure 5 shows a small segment of a computer-generated alignment, representing the VP4/VP2 junction from 17 viral strains. According to all structural considerations, the aligned residues should hydrogen-bond in orientations exactly analogous to those of rhino and polio, forming congruent beta-A1/beta-A2 structures [1,2]. Thus, the overall conformational arrangement near the VP4/VP2 junction is probably common and characteristic of all picornaviruses, an important requirement for the serine protease-like model of VPO cleavage.

	Alignment of VP4 Sequences (COOH end only)	Alignment of VP2 Sequences (NH2 end only)
	...(COOH)	(NH2)...
FMDA10	TtnTQnndwfSklassafTglfGAlLa	dkKtEettlleDRIlttrnGHTTsTTQsSvgvtyGYS
FMDA12	TtnTQnndwfSklassafTglfGAlLa	dkKtEettlleDRIlttrnGHTTsTTQsSvgvtyGYS
FMD01K	TtnTQnndwfSklassafSglfGAlLa	dkKtEettlleDRIlttrnGHTTsTTQsSvgvtyGYA
EMC	Ttv.QfsnlfSg.aVnafSnml.PlLa	dqNtEemenlSDRVsqdTaGNTvtnTQsTvgrLVGYG
Mengo	T.YgQfsnllSg.aVnafSnml.PlLa	dqNtEemenlSDRVsqdTaGNTvtnTQsTvgrLVGYG
HAVChi	GiF.Q..tvgSglD...h..ilS..La	diEeEqmiqsvDRtavtgaSyfTsvdQsSvHTaev.G
HAVHm175	GiF.Q..tvgSglD...h..ilS..La	diEeEqmiqsvDRtavtgaSyfTsvdQsSvHTaev.G
HAVCr326	GiF.Q..tvgSglD...h..ilS..La	diEeEqmiqsvvRtavtgaSyfTsvdQsSvHTaev.G
Polio1M	SkFTE.....PikDVlikT...APmLn	spNiEa.cgySDRVlqlTlGNSTiTTQeAaNSVVAYG
Polio1S	SkFTE.....PikDVlikT...SPmLn	spNiEa.cgySDRVlqlTlGNSTiTTQeAaNSVVAYG
Polio2La	SkFTE.....PikDVlikT...APtLn	spNiEa.cgySvRVmqlTlGNSTiTTQeAaNSVVAYG
Polio2S	SkFTE.....PikDVlikT...APmLn	spNiEa.cgySDRVmqlTlGNSTiTTQeAaNSVVAYG
Polio3Le	SkFTE.....PlkDVlikT...APaLn	spNvEa.cgySDRVlqlTlGNSTiTTQeAaNSVVAYG
Polio3S	SkFTE.....PlkDVlikT...APaLn	spNvEa.cgySDRVlqlTlGNSTiTTQeAaNSVVAYG
CoxB3	GkFTE.....PvkDImikS.l..PaLn	sptvEe.cgySDRVrsiTlGNSTiTTQecaNvVVGYG
Rhino2	SkFTD.....PvkDVlek.gi..PtLq	sptvEa.cgySDRIiqiTrGDSTiTSQdvaNAIVAYG
Rhino14	SkFTE.....PvkDLmlk.g..APaLn	spNvEa.cgySDRVqqiTlGNSTiTTQeAaNAVVcYA
		bbbA1bbbbbbA2bbb

Fig. 5. Sequences near VP4/VP2 cleavage sites. Amino acid sequences near the VP4/VP2 cleavage site for 17 strains of picornaviruses have been aligned by computer analysis [19–32]. Gaps (...) have been inserted into the sequences to maximize homology among aligned residues. Upper case letters denote positions where a majority of the sequences share homology. The bottom line (bbb) shows beta-A1/beta-A2 amino acids of rhino-14 [35]. Serines (S) in boldface are putative catalytic residues.

The hypothetically active serine residue at the amino terminus of the hairpin is found only in coxsackie B3, EMC, Mengo, the polio strains, and rhinoviruses. Hepatitis A (HAV) and foot-and-mouth disease (FMD) do not have a serine at this alignment position. Instead, the first VP2 serine residue in FMD is part of a Gln-Ser-Ser-Val sequence, which is common also to the hepatitis A strains. The equivalent rhino-14 sequence, Gln-Glu-Ala-Ala, is spacially located at the carboxy-end of the beta-A1/beta-A2 structure. In other words, each of the 17 sequenced picornaviruses displays a serine residue at either the extreme carboxy- or extreme amino-end of the VP2 beta-hairpin. We believe these amino acid positions (boldfaced in Fig. 5) would be functionally equivalent in a proteolytic mechanism. A serine in either position could potentially catalyze VPO cleavage, since the processing site between VP4 and VP2 would always be in direct proximity of the side chain.

The adjacent aspartate residue (amino acid 11 in rhino-14), which is characteristic of serine-protease reactive sites [37–39], is highly conserved in its position and orientation among the viruses. Several other nearby amino acids are also common to all sequences (VP4: Leu-67, VP2: Glu-5 Arg-12, Gln-26), suggesting that these too may play a role in the catalytic mechanism or formation of the reactive environment.

However, proton-accepting histidine residues are not conspicuously located near the potential cleavage site in any picornavirus. VPO processing in vivo takes place deep within a maturing particle, inaccessible to external agents and contiguous only with the packaged RNA [1,2]. This has led to the speculation that RNA, or ions associated with the RNA, may supply this required function [2,35,36].

Study of the maturation cleavage has, until now, been hindered by lack of an appropriate in vitro system capable of studying the activity. We have recently discovered that addition of certain small diamino compounds to cell-free reactions pro-

grammed with EMC RNA results in the apparent cleavage of VPO into VP4+VP2 [36]. Hydrazine and ethylenediamine were the only tested compounds to induce this in vitro effect. Monoamine reagents such as methylamine, ethylamine, propylamine, hydroxylamine, and semicarbazide were inactive in our experiments. Presumably, hydrazine and ethylenediamine have "catalytic" activity in vivo because they can function as a viral RNA would in vivo, serving as strong proton-acceptors during VPO catalysis. The mechanism by which diamines, but not monoamines, can replace RNA in these reactions is completely unknown. We hope in the future to develop the in vitro system into a sensitive assay for genetic manipulation experiments, allowing closer study of the required elements for VPO processing and the serine protease-like hypothesis.

The overall picornaviral proteolytic cascade is a distinguishing feature of these viruses. The work cited above has helped to define the scheme and to clarify some mechanistic elements by which proteolysis may contribute to regulation of the viral life cycle. It is interesting to note that all identifiable cleavages within a polyprotein are catalyzed by viral-encoded sequences, 2A (L or 2B?), 3C, and VPO. No required cellular components seem to be necessary for this process. At least, none have yet been isolated. Enzymatic self-sufficiency is undoubtedly an advantage to a pathogen and may explain, in part, why the picornaviruses can efficiently infect a wide range of hosts with such obvious success.

ACKNOWLEDGMENTS

I thank Michael Rossmann and his colleagues for helpful discussions on the VPO mechanism. ACP is supported by Public Health Service grant AI-17331 from the National Institutes of Health.

REFERENCES

1. Hogle JM, Chow M, Filman DJ: Science 229:1358, 1985.
2. Rossmann MJ, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht H-J, Johnson JE, Kamer G, Luo M, Mosser AG, Rueckert RR, Sherry B, Vriend G: Nature 317:145, 1985.
3. Ahlquist P, Kaesberg P: Nucleic Acids Res 7:1195, 1979.
4. Palmenberg AC: In Rowlands DJ, Mahy BWJ, Mayo M (eds): "The Molecular Biology of Positive Strand RNA Viruses." London: Academic Press, 1986.
5. Black DN, Stephenson P, Rowlands DJ, Brown F: Nuc Acid Res 6:2381, 1979.
6. Costa Giomi MP, Bergmann IE, Scodeller EA, Auge de Mello P, Gomez I, La Torre JL: J Virol 51:799, 1984.
7. Harris TJR, Brown F: J Gen Virol 33:493, 1976.
8. Rueckert RR: In Fraenkel-Conrat H, Wagner RR (eds): "Comprehensive Virology," Vol 6. New York: Plenum, 1976, pp 131–212.
9. Rueckert RR: In Fields B, Knipe DM, Chanock RM, Melnick JL, Roizman B, Shope RE (eds): "Virology." New York: Raven Press, 1985, pp 705–738.
10. Rueckert RR, Wimmer E: J Virol 50:957, 1984.
11. Palmenberg AC: J Virol 44:900, 1982.
12. Palmenberg AC, Pallansch MA, Rueckert RR: J Virol 32:770, 1979.
13. Shih DS, Shih CT, Zimmern D, Rueckert RR, Kaesberg P: J Virol 30:472, 1979.
14. Toyoda H, Nicklin MJH, Murray MG, Anderson CW, Dunn JJ, Studier FW, Wimmer E: Cell 1986, in press.
15. Grubman MJ, Baxt B: Virology 116:19, 1982.
16. Gorbalenya AE, Svitkin YV, Kazachkov YA, Agol VI: FEBS Lett 108:1, 1979.

17. Palmenberg AC, Rueckert RR: J Virol 41:244, 1982.
18. Svitkin YV, Gorbalenya AE, Kazachkov YA, Agol VI: FEBS Lett 108:6, 1979.
19. Baroudy BM, Ticehurst JR, Miele TA, Maizel JV Jr, Purcell RH, Feinstone SM: Proc Natl Acad Sci USA 82:2143, 1985.
20. Boothroyd JC, Highfield PE, Cross GAM, Rowlands DJ, Lowe PA, Brown F, Harris TJR: Nature 290:800, 1981.
21. Robertson BH, Grubman MJ, Wenddell GN, Moore DM, Welsh JD, Fischer T, Dowbenko DJ, Yanssura DG, Small B, Kleid DG: J Virol 54:651, 1985.
22. Forss S, Strebel K, Beck E, Schaller H: Nucleic Acids Res 12:6587, 1984.
23. Palmenberg AC, Kirby EM, Janda MJ, Drake NL, Duke GM, Potratz KF, Collett MS: Nucleic Acids Res 12:2969, 1984.
24. Najarian R, Caput D, Gee W, Potter S, Renard A, Merryweather J, Van Nest G, Dina D: Proc Natl Acad Sci USA 82:2627, 1985.
25. Linemeyer DL, Menke JG, Martin-Gallardo A, Hughes JV, Young A, Mitra SW: J Virol 54:247, 1985.
26. Kitimura N, Semler B, Rothberg PG, Larsen GR, Adler CJ, Dorner AJ, Emini EA, Hanecak R, Lee JJ, van derf Werf S, Anderson CW, Wimmer E: Nature 291:547, 1981.
27. Toyoda H, Kohara M, Kataoka Y, Suganuma T, Omata T, Imura N, Nomoto A: J Mol Biol 174:561, 1984.
28. La Monica N, Meriam C, Racaniello VR: J Virol 57:515, 1986.
29. Stanway G, Hughes PJ, Mountford RC, Reeve P, Schield GC, Almond JW: Proc Natl Acad Sci USA 81:1539, 1984.
30. Tracy S, Liu H-L, Chapman NM: Virus Res 3:263, 1985.
31. Skern T, Sommergruber W, Blass D, Gruendler P, Fraundorfer F, Pieler C, Fogy I, Kuechler E: Nucleic Acids Res 13:2111, 1985.
32. Stanway G, Hughes P, Mountford R, Minor P, Almond J: Nucleic Acids Res 12:7859, 1984.
33. Parks GD, Duke GM, Palmenberg AC: J Virol 60:376, 1986.
34. Jacobson MF, Baltimore D: J Mol Biol 33:369, 1968.
35. Rossmann MG, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht H-J, Jonson JE, Kamer G, Luo M, Mosser AG, Palmenberg AC, Rueckert RR, Sherry B, Vriend G: Proc Swed Acad Sci 1986 in press.
36. Arnold A, Luo M, Vriend G, Rossmann MG, Palmenberg AC, Parks GD, Nicklin M, Wimmer E: Proc Natl Acad Sci USA 83:in press, 1986.
37. Robertus JD, Kraut J, Alden RA, Birktoft JJ: Biochemistry 11:4293, 1972.
38. Stroud RM: Sci Am 231:74, 1974.
39. Kraut J: Annu Rev Biochem 46:331, 1977.
40. Ticehurst J: Semin Liver Dis 6:46, 1986.