Similarity between the picornavirus VP3 capsid polypeptide and the *Saccharomyces cerevisiae* virus capsid polypeptide

J.A.Bruenn, M.E.Diamond and J.J.Dowhanick

Department of Biological Sciences, State University of New York at Buffalo, Buffalo, NY 14260, USA

Received June 13, 1989; Revised and Accepted July 8, 1989

ABSTRACT

We have compared the sequence of the capsid polypeptide of the *Saccharomyces cerevisiae* double-stranded RNA virus, ScV, with those of the picornaviruses. A central region of 245 amino acids in the ScV capsid polypeptide of 680 amino acids has significant similarity to the picornavirus VP3. This similarity is more extensive than that already noted for the alphavirus capsid polypeptide and the picornavirus VP3 (Fuller, S.D. and Argos, P, EMBO J. 6, 1099, 1987). Together with the similarity between the ScV RNA polymerase and the picornavirus RNA polymerases, this result implies an evolutionary relationship between a simple double-stranded RNA virus of fungi and the small plus strand RNA animal viruses.

INTRODUCTION

The *Saccharomyces cerevisiae* viruses (ScV) are non-infectious, double stranded RNA (dsRNA) viruses with segmented genomes. A large viral dsRNA (L) encodes the viral capsid polypeptide (P1) and an RNA-dependent RNA polymerase, or RDRP (1,2). The ScV replicative cycle is similar to that of reovirus: the viral polymerase conservatively synthesizes the viral plus strand which then serves as a template for minus strand synthesis within new viral particles (3,4,5). The ScV RDRP has extensive primary sequence similarity with the RDRPs of the small, plus strand RNA viruses (1,2) which include both plant and animal viruses (6).

In the subset of these viruses with icosahedral symmetry, there is extensive secondary and tertiary structure homology among the viral capsid polypeptides. All have one or more capsid polypeptides with an eight stranded anti-parallel beta-barrel structure (7). In most of the plant viruses, such as tomato bushy stunt virus (TBSV), Southern bean mosaic virus (SBMV), turnip yellow mosaic virus (TYMV), and brome mosaic virus (BMV), the capsid is composed of 180 copies of one such polypeptide of about 30 kilodaltons, in which a central, S domain, of about 190 amino acids adopts the beta-barrel structure. These viruses have a T=3 symmetry. The small plus strand animal viruses, the picornaviruses, such as foot and mouth disease virus (FMDV), poliovirus (polio), human rhinovirus (HRV), encephalomyocarditis virus (EMC), and Mengovirus (Mengo), have a pseudo T=3 structure. Their icosahedral faces have three different capsid polypeptides, each of about 250 amino acids: VP1, VP2, and VP3.

Figure 1. Dotplot of the comparison of ScV P1 with HRV14 VP3. The GCG program COMPARE was used with a window of 30 and a stringency of 14. Arrows indicate the diagonal of similarity detected.

These are of different primary sequence but essentially identical protein fold and replace the three polypeptides of identical sequence present on each face of the T=3 plant virus capsid. There is no detectable primary sequence similarity between the T=3 plant virus capsid polypeptides and the picornavirus capsid polypeptides, but the structural homology between them and the S domain is striking (7,8). The capsid polypeptides of the plant and animal RNA viruses for which X-ray crystal structures exist have been compared by a process of three-dimensional mapping. This analysis demonstrates that the proportion of structurally equivalent residues is highest for VP3 versus the plant virus capsid polypeptides (7). Among the picornaviruses, VP3 also shows the greatest conservation. For instance, the VP3 of Mengovirus is more than 90% structurally identical to the VP3 of HRV14 (7), while they share only about 48% amino acid sequence similarity (for similarity classes see below). Although there is no X-ray structure for the alphavirus Sindbis (SNBV), high resolution electron microscopy has demonstrated a T=3 symmetry for the nucleocapsid. The single viral nucleocapsid polypeptide, of 264 amino acids, has detectable sequence similarity to the picornavirus VP3 (9).

There is as yet no data on the arrangement of P1 subunits in the ScV virion, but the virion has "spherical" symmetry and is slightly larger than the picornavirus virion: 35-40 nm (10) as opposed to 30 nm (11). The number of monomers of P1 (680 amino acids) present in virions has been variously estimated at 60 (12) or 120 (10, 13). Given that the ScV RDRP has more extensive sequence similarity to the picornavirus and alphavirus RDRPs than to the plant virus RDRPs, we sought primary sequence similarity between P1 and the picornavirus and alphavirus capsid polypeptides. We find significant sequence similarity between the central 245 amino acids of ScV P1 and VP3 of the picornaviruses. This primary sequence similarity is considerably more extensive than that reported for SNBV.

Table 1. Similarity between picornavirus VP3, ScV P1, and SNBV C

	% similarity to				
	HRV14 VP3	EMC VP3	FMDV VP3	ScV P1(203-448)	SNBV C
Polio 1 VP3	67.4	50.9	41.8	37.9	28.9
HRV14 VP3		48.0	46.6	36.2	26.2
EMC VP3			54.8	29.7	28.6
FMDV VP3				33.3	29.5
ScV P1 (203-448)					29.0

Sequences are from the following references: polio 1 (18), HRV14 (19), EMC (20), FMDV (21), SNBV (17), and ScV (2). Similarity is essentially a measurement of evolutionary distance (see Materials and Methods).

MATERIALS AND METHODS

Computer analysis used the programs provided by the University of Wisconsin Genetics Computer Group (14). The program GAP, which we have used to quantify similarity, makes an alignment to maximize "quality," where the quality of an alignment is the sum of the values of the matches less the "gap weight" times the number of internal gaps less the "gap length weight" times the total length of the internal gaps (15). The symbol comparison table from which the value of a match (%similarity) is calculated is from Gribskov (16), in which non-identical residues are compared by a table of mutational differences.

RESULTS

Using the University of Wisconsin GCG program COMPARE, we were able to find significant similarity between VP3 of poliovirus (types 1, 2, and 3) or HRV (types 2, 14 and 89) and the ScV capsid polypeptide (P1). This comparison of HRV14 VP3 and ScV P1, is shown as a DOTPLOT in Fig. 1. The similarity extends approximately between residues 200 and 450 of P1 and throughout VP3. The same alignment is detectable with the picornavirus consensus VP3 sequence (see below), although it is not as dramatic, since ScV P1 is more closely related to polio and Hrv 14 than to EMC or FMDV. Neither VP1 nor VP2 of any of the picornaviruses showed this much similarity to P1. Similar comparisons at the same stringency of the SNBV capsid polypeptide C (17) with all the picornavirus capsid polypeptides failed to demonstrate any detectable similarity. We also failed to detect any similarity between P1 and any of the plant virus capsid polypeptides with this method. The degree of similarity among the picornavirus VP3 polypeptides and their similarity to P1 (residues 203-448) and to the SNBV capsid polypeptide was quantified using the programs GAP and BESTFIT. The results are shown in Table 1.

Only one example of each family of sequences known is shown. Within the picornaviruses, the pair of HRV and polio and the pair of EMC and FMDV are most closely related. Clearly, the region of ScV P1 identified by similarity to the polio or HRV VP3 is more

	0		49
Con	***.....*.....*.....*.....*.....*		***.*..*
Fmdv	GIFPVACADG YGGLVTTDPK TADPVYGVY NPPKNTNYPGR FTNLLDVAEA		
Emc	SPIPVITIREH AGTWYSTLPD STVPIYGKTP VAPSNYMVGE YKDFLEIAQI		
Hrv14	GLPTTTTLPG SQGFLTDDR QSPSALPNYE PTPRIHIPGK VHNLLLEIQV		
Polio1	GLPVMNTPG SNQYLADNF QSPCALPEFD VTPPIDIPGE VKNMMEIAEI		
Scv	VRPLTVTEV NFVLMMSKW HRRTNLDAIDY EAPQLADKFA YRAQLTVQDA		
Vp3	.glPvt..pg sqgwltdpk qspgalgkfe vtPridipGe fknlleiaei		
Beta	AAAAAAA.....		
	50		99
Con	*..*..*..*..*.....*.....*.....*.....*.....*		*..*..*..*..*
Fmdv	CPTFLRF.D. D...G..KPY VVTRA.D.D. TRLLAKFDVS LAAKHMSNTY		
Emc	PTFIGNK.IP N...A..VPY IEASN.TAVK TQPLATYQVT LSCSCLANTF		
Hrv14	DTLIPMNNT. HTKDEVNSYL IPLNANR.Q. NEQVFGTNNL IGDGVFKTTL		
Polio1	DTMIFDLSA TKKNTMEMYR VRLSDKPHTD DPILCLSLSP ASDPRLSHTM		
Scv	DEWIEGDRT. DDQFRPPSSK VMLSALRKIV NRRLNYNQFY YRAQLTVQIM		
Vp3	dtlipmd.s. ..kn.mn.yy v.lsa..... deila..dvs lsc.clsnTl		
Beta	..BBB.....		BBBBB.....CCCC.....
	100		149
Con	*..*..*..*..*.....*.....*.....*.....*.....*		*..*..*..*..*
Fmdv	LSGIAQYYTQ YSGTINLHEM FT...GSTDS KARYMVA..Y IPPGVETPPD		
Emc	LAALSRNFAQ YRGSIVYTFV FT...GTAMM KGKFLIA..Y TPGAG.KPT		
Hrv14	LGEIVQYYTH WSGSLRFSIM YT...GPALS SAKLILA..Y TPGAR.GPQ		
Polio1	LGEILNYYTH WAGSLKFTFL FC...GFMMMA TGKLLVS..Y APPGAD.PPK		
Scv	MKPVPNCAEG YAWLMHDALV NIPKFGSIRG RYPFLLSGDA ALIQAT.ALE		
Vp3	Lgeivqyyth wsGslrftfm ft...Gsams kgkflva..Y iPPGa..pP		
Beta	DD DDDDDDDDD DDDDDDD...EEEEEEEEEE E.....		
	150		199
Con	*.....*.....*.....*.....*.....*		*.....*
Fmdv	TPEEAAHCII AEWDTLGNSK ...FTFSIP YVSAADYAYT ASDTAETTNV		
Emc	SRDQAMQATY AIWDLGLNSS ...YSFTVP FISPTHFRMV GTDQVNIINA		
Hrv14	DRREAMLGTH VVWDIGLQST ...IVMTIP WTSQVQFRT DPDTYT...S		
Polio1	KRKEAMLGTH VIWDIGLQSS ...CTMVVP WISNTTYRQT IDDSFT...E		
Scv	DWSAIMAKEPE LVFTYAMQVS VALNTGLYLR RVKKTGFETT IDDSYE....		
Vp3	tr.eAmlgth viWDiGLnSs ...itftvp wiSqt.yryt isDsyn....		
Beta	FF FFFFF...GG GGGGGGGGGG.GGG.....		
	200		249
Con	*..*..*..*..*.....*.....*.....*.....*.....*		*..*..*..*..*
Fmdv	QGWVCVYQIT .HGK.....AENDT ..LLVSASAG KDFELRLPID		
Emc	DGWTVWQLT PLTYPPGCG.TSAK ..ILTMVSAG KDFSLRMPIS		
Hrv14	AGFLSCWYQT .SLI.....L PPETTQ.QVY ..LLSFISAC PDKFLRLMD		
Polio1	GGYISVFYQT .RIV.....V PLSTPR.EMD ..ILGFVSAC NDFSVRLLRD		
Scv	DGA.FLQPET FVQAALACCT QDAPL.NGM SDVYVTPDL LEFDAVTQVP		
Vp3	dGwsvwyqT ..li.....l p..t...e... ..lSfvsSAC kDfslrlpid		
Beta	.HHHHHHHHH HHHH.....		II IIIIIIIIII IIIIIII
	250	262	
Con	...*		
Fmdv	PRIQ		
Emc	PAPWSQ		
Hrv14	TQTISQTVAl TEG		
Polio1	THIEQKALA Q		
Scv	ITVIE		
Vp3	t.tisq.aaa		

Figure 2. Similarities between the picornavirus VP3 and ScV P1 (203-448). The picornavirus VP3 sequences were aligned by structural and primary sequence similarity as described in the text and a consensus sequence generated. These were aligned to ScV P1 as described in the text. The top line (con) indicates by asterisks similarity between ScV P1 and the picornavirus consensus sequence. The bottom line (beta) indicates the location of beta strands of HRV14, FMDV, and EMC (8, 11). Sequences are from the same references as in Table 1. Similarity classes of amino acids were (P,G), (S,T), (E,D,Q,N), (L,V,A,I,M,C), (F,W,Y,H), (K,R); these are different than those used by the COMPARE and GAP programs.

Figure 3. Hydrophobicity of VP3 polypeptides. The Kyte-Doolittle "hydropathy" (24) with a window of 9 amino acids is plotted as a function of residue for the HRV14 and EMC VP3 and for the ScV P1 (203-448). The top graph is the hydrophobicity of HRV14 and the lower graphs superimpose this curve on that for EMC or ScV P1. Maximum value on the ordinate is 3 and minimum value -3.

closely related to all the picornavirus VP3 sequences than is the SNBV C polypeptide. In fact, this 245 amino acid long region in P1 is almost as closely related to the polio VP3 as is the FMDV VP3. As reported (9), SNBV C is most closely related to FMDV VP3. Even this best match of SNBV C with a picornavirus VP3 (with FMDV) is not as good as the worst match of ScV P1 (with EMC).

We performed an alignment of the picornavirus VP3 protein sequences and ScV P1 using the LINEUP and BESTFIT programs in the following way. First, EMC, FMDV, and HRV14 were aligned as previously determined by structural identities (8, 11). For convenience, the EMC VP3 was used in place of the Mengo VP3; these are more than 95% identical (8). Next, polio 1 and HRV14 (67.4% similar) were aligned by BESTFIT. The result was aligned to P1 using BESTFIT on polio VP3 and ScV P1. The consensus VP3 sequence, calculated only from the picornavirus VP3 sequences and shown on the next to last line, is now 67.9% similar to HRV14, 67.1% similar to polio 1, 58.6% similar to FMDV, 54.5% similar to EMC, 33.7% similar to ScV, and 30.4% similar to SNBV as calculated by GAP. The residues in ScV P1 similar to those in the picornavirus VP3 consensus sequence are shown as asterisks on the first line of the figure. Similarity groups of amino acids are taken as (F, W, Y, H), (L, I, V, A, C, M), (S, T), (E, D, Q, N), (K, R), and (P, G). A search of the NBRF data bank with the FASTP program for sequences similar to the consensus VP3 sequence found only picornavirus VP3 sequences among the 18 most similar sequences, indicating that this is an accurate consensus.

The "beta" line of the figure shows the locations of the beta strands of HRV14, FMDV, and EMC (8, 11). The eight-stranded beta-barrel is composed of strands B through I. If the ScV P1 has a similar secondary and tertiary structure in this region, the location of its beta strands should be predicted by this alignment. We used the program PEPTIDESTRUCTURE to predict regions of beta strand in ScV P1 and HRV14. Both the Chou-Fasman (22) and the Garnier et al (23) algorithms found 8 of the 9 beta strands in HRV14; regions G2 and H were

found only by the Chou-Fasman algorithm. In some cases, the regions were not very accurately located by these predictions, especially for regions D and I, which were minimized by the computer algorithms, and B and C, which were exaggerated by the computer algorithms. The Chou-Fasman algorithm predicted two additional non-existent beta strands and the Garnier et al algorithm three. The locations of the beta strands in ScV P1 predicted by the alignment of Fig. 2 corresponded to Chou-Fasman and/or Garnier et al predictions for all but region B1. As with HRV14, putative regions D and I were again minimized and B and C exaggerated by both algorithms.

The alignment of ScV P1 with VP3 is also supported by examination of the hydrophobicity along the polypeptide. Although the HRV14 and EMC VP3 polypeptides are only 48% similar in primary sequence (Table 1), their hydrophobicity plots are remarkably similar (Fig. 3). ScV P1 preserves much of this similarity, as might be expected from the secondary structure predictions.

DISCUSSION

The similarity we have detected between ScV P1 and the picornavirus VP3 is consistent with the evolutionary conservation of the eight stranded beta-barrel in the capsid polypeptides of all the small RNA viruses. If the similarity to the picornavirus capsid polypeptides were complete, P1 would have four domains, corresponding to VP4, VP2, VP3, and VP1, in that order from amino to carboxy-terminus. We have not, however, found any domains of P1 similar to any of the picornavirus capsid polypeptides except for the central region similar to VP3. An alternative model is that ScV has adopted the strategy of the small RNA plant viruses, in which a central region of the capsid polypeptide adopts the beta-barrel structure, but in this case only about a third of the polypeptide would be so engaged. In addition to the already noted similarity between the putative ScV RDRP and that of the picornaviruses, the similarity in capsid polypeptides supports the existence of an evolutionary relationship between the picornaviruses and this fungal dsRNA virus.

ACKNOWLEDGMENTS

We thank Michael Holland for the hospitality one of us (J.A.B) enjoyed during some of this work and Alan Pepper and Kate Willett for use of their VAX accounts. This work was supported by grants from the NIH (GM22200 and AI25721), the USDA (87-CRCR-1-2368), and the Center for Applied Molecular Biology and Immunology of SUNY/Buffalo.

REFERENCES

1. Pietras, D.F., Diamond, M.E., and Bruenn, J.A. (1988). Nucl. Acids Res. 16, 6226.
2. Diamond, M.E., Dowhanick, J.J., Nemeroff, M.E., Pietras, D.F., Tu, C.-L. and Bruenn, J.A. (1989). J. Virol. 63, in press.
3. Nemeroff, M.E. and Bruenn, J.A. (1986). J. Virol. 57, 754-758.
4. Fujimura, T., Esteban, R., and Wickner, R.B. (1986). Proc. Natl. Acad. Sci. USA 83, 4433-4437.

5. Williams, T.L. and Leibowitz, M.J. (1987). *Virology* 158, 231-234.
6. Kamer, G. and Argos, P. (1984). *Nucl. Acids Res.* 12, 7269-7282.
7. Rossmann, M. (1987). *Bioessays* 7, 99-103.
8. Luo, M., Vriend, G., Kamer, G., Minor, I., Arnold, E., Rossmann, M.G., Boege, U., . Scraba, D.G., Duke, G.M., and Palmenberg, A.C. (1987). *Science* 235, 182-191.
9. Fuller, S.D. and Argos, P. (1987). *EMBO J.* 6, 1099-1105.
10. Bruenn J.A. (1980). *Ann. Rev. Microbiol.* 34, 49-68.
11. Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D., and Brown, F. (1989). *Nature* 337, 709-716.
12. Reilly, J.D., Bruenn, J., and Held, W. (1984). *Biochem. Biophys. Res. Comm.* 121, 619-25.
13. Fujimura, T., and Wickner, R.B. (1988). *Cell* 55, 663-671.
14. Devereux, J., Haeberli, P., and Smithies, O. (1984). *Nucl. Acids Res.* 12, 387-395.
15. Needleman, S.B. and Wunsch, C.D. (1970). *J. Molec. Biol.* 48, 443-453.
16. Gribskov, M. and Burgess, R.R. (1986). *Nucl. Acids Res.* 14, 6745-6763.
17. Strauss, E.G., Rice, C.M., and Strauss, J.H. (1984). *Virology* 133, 92-110.
18. Racaniello, v.R. and Baltimore, D. (1981). *Proc. Natl. Acad. Sci. USA* 78, 4887-4891.
19. Callahan, P.L., Mizutani, S., and Colonno, R.J. (1985). *Proc. Natl. Acad. Sci. US* 82, 732-736.
20. Palmenberg, A.C., Kirby, E.M., Janda, M.R., Drake, N.L., Duke, G.M., Potratz, K.F., and Collett, M.S. (1984). *Nucl. Acids Res.* 12, 2969-2985.
21. Boothroyd, J.C., Harris, T.J.R., Rowlands, D.J. and Lowe, P.A. (1982). *Gene* 17, 153-161.
22. Chou, P.Y. and Fasman, G.D. (1978). *Adv. in Enzym.* 47, 45-147.
23. Garnier, J., Osguthorpe, D.J., and Robson, B. (1978). *J. Molec. Biol.* 120, 79-120.
24. Kyte, J. and Doolittle, R.F. (1982). *J. Molec. Biol.* 157, 105-132.