Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 8596–8601, July 1998
Biophysics

Identification of viral mutants by mass spectrometry

J. KATHLEEN LEWIS*†, MOHAMMED BENDAHMANE†‡, THOMAS J. SMITH§, ROGER N. BEACHY‡¶, AND GARY SIUZDAK*¶

*Departments of Molecular Biology and Chemistry and ‡Department of Cell Biology, The Scripps Research Institute, La Jolla, CA 92037; and §Department of Biological Sciences, Purdue University, West Lafayette, IN 49707

Contributed by Roger N. Beachy, May 22, 1998

ABSTRACT **A method to identify mutations of virus proteins by using protein mass mapping is described. Comparative mass mapping was applied to a structural protein of the human rhinovirus Cys1199 → Tyr mutant and to genetically engineered mutants of tobacco mosaic virus. The information generated from this approach can rapidly identify the peptide or protein containing the mutation and, in cases when nucleic acid sequencing is required, significantly narrows the region of the genome that must be sequenced. High-resolution matrix-assisted laser desorption/ionization (MALDI) mass spectrometry and tandem mass spectrometry were used to identify amino acid substitutions. This method provides valuable information for those analyzing viral variants and, in some cases, offers a rapid and accurate alternative to nucleotide sequencing.**

Identification of virus mutants currently requires sequencing part or all of the genome to determine the nature of the mutation. However, localizing the mutation prior to sequencing may be possible if the virus is well characterized biologically as well as structurally. For example, if a new disease phenotype is characterized by altered interactions of the virus with cell receptors, it is likely that a structural protein has been mutated. Thus, if the function of the viral proteins is known, a mutation can be narrowed to a particular protein or proteins.

Automated DNA sequencing is a well established method for identifying mutant proteins and is used to pinpoint specific regions that undergo mutation. In fact, high-throughput DNA sequencing often allows for expeditious identification of viral mutants and offers complete and unambiguous sequence information. However, some disadvantages of this approach include technical limitations in sequencing viral RNAs and the inability to map post-translational modifications. A complementary approach to nucleic acid sequencing of viral mutants is protein mass mapping using matrix-assisted laser desorption/ionization (MALDI) and/or electrospray ionization (ESI) mass spectrometry (1). The development of these two techniques has revolutionized the mass analysis of small and large, thermally labile biomolecules. Indeed, mass spectrometry has recently been used to characterize capsid proteins and post-translational modifications (2) and, in combination with protein digestions, has been used to investigate capsid mobility (3, 4).

Our approach to identify mutant proteins employs protein mass mapping. Protein mass mapping consists of enzymatic digestion of a protein(s) followed by mass analysis of the resulting peptide mixture. By comparing differences in the mass of peptides that are released by such treatment, one is able to identify peptides in which amino acid differences occur. This information defines the region containing a mutation and, in cases when nucleotide sequencing is required, significantly narrows the region of the genome that must be sequenced. Accurate mass measurements and tandem mass spectrometry can then be used to definitively identify the amino acid substitution; a schematic of this approach is shown in Fig. 1.

MATERIALS AND METHODS

Tobacco Mosaic Virus (TMV) Constructs. The known atomic structure (5) of TMV was used to select two amino acids in the coat protein (CP) for mutation. Glu-50 and Asp-77 are located on the right slewed (RS) and right radial (RR) α-helices of the CP, respectively. The CP mutants Glu-50 → Gln, Glu-50 → Met, Asp-77 → Asn, and Asp-77 → Arg (M.B. and R.N.B., unpublished work) were generated by PCR-based site-directed mutagenesis using the plasmid pKN2 (6) containing the full-length TMV CP. The mutant CP genes were then used to replace their homologue in the TMV cDNA clone U3/12-4 (7) to generate infectious full-length cDNA clones: pTMV-Glu-50Gln, pTMV-Glu50Met, pTMV-Asp77Asn, and pTMV-Asp77Arg.

Purification of Virus Particles from Infected Plants. Full-length transcripts were produced with T7 RNA polymerase from the wild-type pTMV (U3/12-4) and mutant pTMV cDNA clones as described previously (7) and used to inoculate 4-week-old *Nicotiana tabacum* Linnaeus cv. Xanthi nn plants. Ten to 15 days after inoculation, virus particles were purified from systemically infected leaves as described previously (8). Infected leaf material was ground in liquid N_2 and homogenized in the extraction buffer (0.5 M Na_2HPO_4/0.5% sodium ascorbate). Cellular debris was removed by centrifugation and chlorophyll was removed by extraction with diatomaceous earth (grade III, Sigma). Virus particles were precipitated twice in the presence of 3% PEG_{8000} and 1% NaCl, washed with 5% Triton X-100, and collected by centrifugation at 90,000 × *g*. Virus particles were further washed (two times) with 200 mM sodium phosphate, pH 7, for 4 hr at 37°C and collected by centrifugation at 90,000 × *g*. The pure virus particles were then resuspended in water or in 10 mM Tris·HCl, pH 7.2/1 mM EDTA to a final concentration of 1 mg/ml and stored at 4°C. Mutant virus particles were indistinguishable from wild-type virus particles by electron microscopy (data not shown).

Production of Mutant Human Rhinovirus 14 (HRV14). The HRV14 mutant (Cys-199 → Tyr in structural protein 1), a naturally occurring spontaneous mutant of HRV14, was selected for by isolating plaques that developed when wild-type

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/958596-6$2.00/0
PNAS is available online at http://www.pnas.org.

Abbreviations: MALDI, matrix-assisted laser desorption/ionization; TMV, tobacco mosaic virus; HRV14, human rhinovirus 14; FTMS, Fourier transform mass spectrometry; TOF, time-of-flight; MS^2, tandem mass spectrometry.

†J.K.L. and M.B. contributed equally to this work.
¶To whom reprint requests and requests for materials should be addressed at: for R.N.B., 10550 North Torrey Pines Road, BCC206, La Jolla, CA 92037 (e-mail: beachy@scripps.edu); and for G.S., 10550 North Torrey Pines Road, BCC157, La Jolla, CA 92037 (e-mail: siuzdak@scripps.edu).