Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 4471–4473, April 1996
Cell Biology

On signal sequence polymorphisms and diseases of distribution

JONATHAN S. ROSENBLUM, NORTON B. GILULA, AND RICHARD A. LERNER

Departments of Chemistry, Molecular Biology and Cell Biology, and The Skaggs Institute of Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037

Contributed by Richard A. Lerner, January 2, 1996

ABSTRACT **We report a previously unappreciated property of the signals that target organelle-specific proteins to their subcellular sites of action. Such targeting sequences are shown to be polymorphic. We discovered this polymorphism when we cloned the mitochondrial manganese-containing superoxide dismutase from cell lines of normal individuals and patients with genetic diseases of premature aging and compared their sequences to each other and to those previously reported. The polymorphism consists of a single nucleotide change in the region of the DNA that encodes the signal sequence such that either an alanine or valine is present. Subsequently, eight cell lines were analyzed and all three possible combinations of the two signal sequences were observed. Such signal sequence polymorphisms could result in diseases of distribution, where essential proteins are not properly targeted, thereby leading to absolute or relative deficiencies of critical enzymes within specific cellular compartments. Progeria and related syndromes may be diseases of distribution.**

Among the most unusual diseases are those whose hallmarks include clinical manifestations of normal aging, but in youngsters. Progerioid diseases include the syndromes of Werner, Cockayne, Hutchinson–Gilford, and others. Symptoms vary within and between diseases, but affected individuals have much in common with each other and with individuals decades older. Cataracts, heart disease, and atherosclerosis are common, and mortality often strikes within 20 years of birth.

The molecular mechanisms of precocious aging diseases may illuminate aspects of normal aging. It was suggested that Werner syndrome was the result of mutations in DNA polymerase β, but this has been questioned (1). Recently, two Cockayne syndrome-associated genes have been identified (2). Progerioid diseases, like normal aging, may be the culmination of the interaction of several factors. As part of an effort to delineate molecular bases for such diseases, we have cloned and sequenced genes for proteins thought to be important in the aging process. In the course of this analysis, we have identified a polymorphism in the signal sequence of mitochondrial manganese-containing superoxide dismutase (MnSOD; EC 1.15.1.1). MnSOD is involved in controlling dioxygen toxicity in the mitochondria, an organelle of extreme oxidative load (3, 4). The presence of more than one signal sequence for this vital enzyme suggests a combinatorial mechanism determining rates of targeting, membrane translocation, and/or signal sequence cleavage with concomitant folding of MnSOD and perhaps other organelle-specific proteins. Also, the presence of polymorphisms in signal sequences raises the possibility that there exist diseases of distribution, where allocation, not activity, of essential proteins is faulty.

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.

MATERIALS AND METHODS

Reagents. TRIzol was obtained from Molecular Research Center (Cincinnati). The first-strand cDNA synthesis kit was from Pharmacia. Primers were prepared on a Beckman 1000M DNA synthesizer. pBluescript was from Stratagene.

Cell Lines and Culture. Fibroblast cell lines listed in Table 1 were obtained from the National Institute of Aging Cell Culture Repository at the Coriell Institute for Medical Research (Camden, NJ) and cultured according to their protocols.

Cloning of MnSOD. Total RNA was isolated (5) and cDNA was prepared with a primer that anneals to the polyadenylylated region of actively transcribed genes. Primers that flank the translated region (sense, CATCAGCGGTctagaAGCACTAGCAGCATG; and antisense, GGCCTCACTctcgagCGATCGTGGTTTA; restriction sites are in lowercase letters) were used to amplify the MnSOD gene, which was subsequently cloned into pBluescript KS(−). Four clones from three cell lines (AG00780G, AGO6269, and AG07721C) were sequenced in both directions. The overall rate of PCR mutation was <0.02%.

Allele-Specific PCR. We used an allele-specific PCR assay (ASP) (6) to distinguish thymidine at nucleotide (nt) 47 from cytidine at nt 47. First, the entire MnSOD gene was amplified with the sense and antisense primers given above (94°C for 30 sec, 50°C for 1 min, 72°C for 1 min for 20 cycles) providing a 715-bp product. This product was then used as a template for a second round of PCR that used a sense primer specific for either cytidine at nt 47 (ASP C, GCAGGCAGCTGGCTCCGAC) or thymidine at nt 47 (ASP T, GCAGGCAGCTGGCTCCGAT) and an antisense primer (ASP back, GTTCTCCACCACCGTTAGGG) (94°C for 30 sec, 60°C for 1 min, 72°C for 1 min for 21 cycles), providing a 311-bp product.

RESULTS

Initially, we cloned and sequenced MnSOD from AG00780G (Werner syndrome), AG06269 (Cockayne syndrome), and AG07721C (clinically normal) lines. Four clones were sequenced from each cell line. Comparison of the sequences to one another and to published sequences (7) revealed a polymorphism at nt 47 (counting from the adenosine of the initial methionine codon) (Fig. 1). At this position, thymidine was found in one of the AG07721C clones and all four of the AG06269 clones. All remaining clones had cytidine at this position. Thus, a GCT codon and a GTT codon exist at this location, corresponding to alanine and valine, respectively. The alanine/valine polymorphism occurs at amino acid 16, which is toward the carboxyl end of the 24-residue mitochondrial signal sequence.

We then used an ASP (6) to determine if AG06269 and AG00780G were homozygous, and to screen other cell lines (Fig. 2, Table 1). Two of the 8 cell lines examined were from clinically normal individuals, and both were heterozygous at

Abbreviations: MnSOD, manganese-containing superoxide dismutase; ASP, allele-specific PCR assay; mtDNA, mitochondrial DNA.

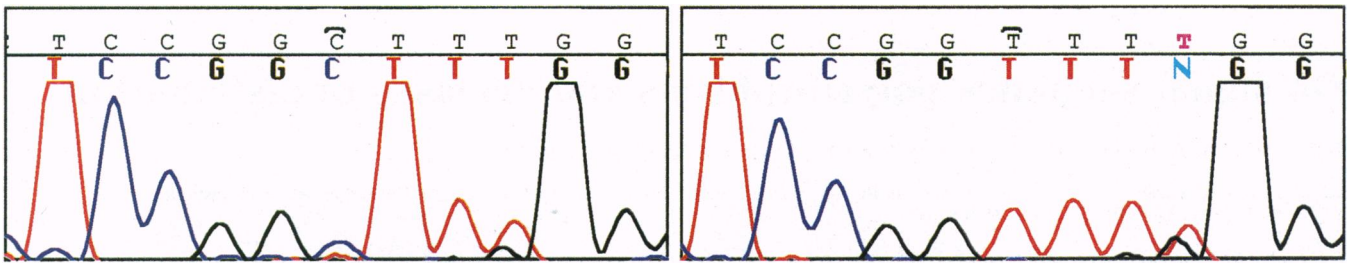

FIG. 1. There is a cytidine/thymidine polymorphism at nt 47 of mitochondrial MnSOD. The top row of letters is the manual base call, the bottom row is the computerized base call. The chromatogram was generated by cycle sequencing plasmid templates with dye-labeled dideoxynucleotides.

this position. One progeria-derived line was homozygotic for cytidine at nt 47, and two others were heterozygotic. Two cell lines, one Cockayne syndrome-derived line and one progeria-derived line, were homozygotic for thymidine at nt 47.

A data base search (8) for the polymorphism revealed two entries that contained thymidine at nt 47, as well as other differences from the "canonical" sequence (9, 10). However, neither report commented on the differences.

DISCUSSION

According to the signal hypothesis of Blobel and colleagues (11–13), signal sequences target nuclear-encoded proteins to their site of action in the cell. Many different sequences are used to target proteins to a particular organelle. However, the concept that one protein can use more than one signal sequence (even in the same cell in the heterozygotes) has not been explored. Our results suggest that this polymorphism may be fairly common, and several lines of evidence support our conclusion that it could be important. Further, we can extrapolate from our finding to make some predictions concerning the importance of intracellular targeting in normal aging.

Recent models of mitochondrial import include several discrete stages of interaction between signal sequences and subunits of the import machinery (14). Proper recognition of the signal sequence by mitochondrial import stimulation factor, the Mas37p–Mas70p receptor, and the Mas20p–Mas22p receptor may all be necessary for proper import through just the outer membrane of mitochondria. After passing through the outer membrane, the signal sequence must be recognized by similar receptors in the inner membrane, and then cleaved from the mature enzyme by matrix-processing peptidase (14). This chain of protein transport is only as strong as its weakest link. Poor recognition of a signal at any stage of its transport could result in mistargeting. In addition, inefficient cleavage of a particular signal may interfere with proper folding of accurately targeted protein (15). It is quite possible that the signal sequence mutation has no effect on the enzymatic activity of MnSOD. Rather, the cellular allocation of the enzyme is likely altered. Perhaps the diversity of signal sequences permits different rates of organelle import.

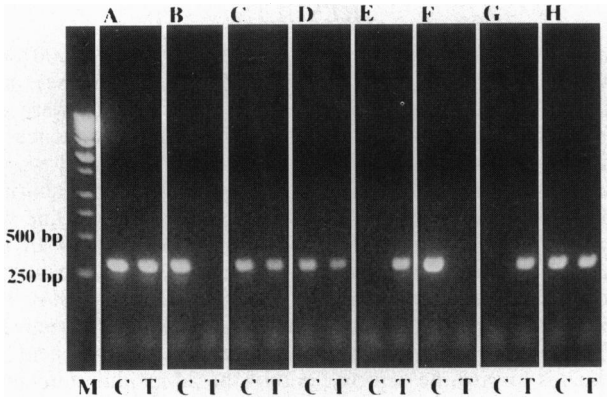

FIG. 2. Eight cell lines contain all combinations of the polymorphism. An ASP (see *Materials and Methods*) was performed. The 311-bp product indicates the presence of either a cytidine or a thymidine at nt 47. M indicates DNA size markers, with relevant bands indicated. C lanes were primed with nt 47 C-specific primer; T lanes were primed with nt 47 T-specific primer. Templates used were AG00780G (*A*), AG06269 (*B*), AG07721C (*C*), GM02037B (*D*), AG03513D (*E*), AG06297B (*F*), AG06917A (*G*), and AG010750 (*H*).

In the mitochondria, MnSOD is involved in the metabolism of superoxide radicals arising from several mechanisms, including normal aerobic respiration. Its mitochondrial localization has recently been shown to be essential for protection from ionizing radiation (16). MnSOD without its signal sequence could not protect cells from radiation, whereas redirection of the normally cytosolic Cu/ZnSOD to the mitochondria with a signal sequence does protect (16). Even with optimal functioning of enzymes responsible for oxygen metabolism, oxidative load in the mitochondria is great. Mitochondrial DNA (mtDNA) from rat liver has been shown to contain 10 times the oxidative damage of nuclear DNA from the same tissue (17). Even though mitochondria are constantly turned over, mtDNA oxidation appears to increase with age (18). This increase in oxidation may be the cause of the increase in mtDNA mutations that occur with age (19, 20). Inefficient targeting of MnSOD could leave mitochondria without their full defense against superoxide radicals. This could lead to protein oxidation, as well as mtDNA mutations. Several diseases that result from mtDNA mutations are already known (21) and may include common aging related diseases such as Alzheimer and Parkinson disease (21). In this manner, a subset of progerioid diseases may mimic normal aging in providing an alternative mechanism for the accumulation of oxidative damage in the mitochondria.

In summary, our results suggest that individuals differ in the sequences used to target certain proteins to the organelles in which they function. As such, diseases of distribution must be considered viable alternatives to the well understood diseases that result from a decrease in protein stability or enzymatic activity. Indeed, it was recently shown that the BRCA1 gene product is aberrantly localized in breast and ovarian cancers (22), raising the possibility that its mistargeting is pathogenic, although in this case it is apparently not the result of a mutation in a signal sequence. Diseases of distribution are analogous to

Table 1. Cell lines, clinical diagnosis, and genotype

Panel in Fig. 2	Cell line	Clinical diagnosis	Genotype
A	GM02037B	Normal	Heterozygote
B	AG03513D	Progeria	Homozygote C
C	AG06297B	Progeria	Heterozygote
D	AG06917A	Progeria	Heterozygote
E	AG010750	Progeria	Homozygote T
F	AG00780G	Werner syndrome	Homozygote C
G	AG06269	Cockayne syndrome	Homozygote T
H	AG07721C	Normal	Heterozygote

disorders of carbohydrate or lipid metabolism where ultimately the absolute amount of a metabolite exerts its effect over a long period of time. Although different signal sequences can be expected to be somewhat degenerate, there is probably an optimal sequence for a given protein. Any variation from the best sequence can be expected to have subtle influences on protein distribution. Resulting inefficiencies in processes such as oxygen radical metabolism may exert profound cumulative effects. Whether progerioid diseases, or the normal differences in aging itself, result from variations in protein targeting remains to be proven. Nevertheless, we speculate that signal sequence polymorphisms in these and other proteins have some effect over an entire lifetime.

Note Added in Proof. In a subsequent analysis of genomic DNA from 20 clinically normal individuals, a nearly random distribution of the two signal sequences was observed.

We thank Janet Dickerson for superb technical assistance. This work was supported by a grant from the Lucille P. Markey Charitable Trust (N.B.G.) and the National Science Foundation (predoctoral fellowship to J.S.R.).

1. Chang, M., Burmer, G. C., Sweasy, J., Loeb, L. A., Edelhoff, S., Disteche, C. M., Yu, C.-E., Anderson, L., Oshima, J., Nakura, J., Miki, T., Kamino, K., Ogihara, T., Schellenberg, G. D. & Martin, G. M. (1994) *Hum. Genet.* **93,** 507–512.
2. Henning, K. A., Li, L., Iyer, N., McDaniel, L. D., Reagan, M. S., Legerski, R., Schultz, R. A., Stefanini, M., Lehmann, A. R., Mayne, L. V. & Friedberg, E. C. (1995) *Cell* **82,** 555–564.
3. Fridovich, I. (1995) *Annu. Rev. Biochem.* **64,** 97–112.
4. Richter, C., Gogvadze, V., Laffranchi, R., Schlepbach, R., Scweizer, M., Suter, M., Walter, P. & Yaffee, M. (1995) *Biochim. Biophys. Acta* **1271,** 67–74.
5. Chomczynski, P. (1993) *BioTechniques* **15,** 163–164.
6. Cha, R. S., Zarbl, H., Keohavong, P. & Thilly, W. G. (1992) *PCR Met. Appl.* **2,** 14–20.
7. Church, S. L. (1990) *Biochim. Biophys. Acta* **1087,** 250–252.
8. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
9. Heckl, K. (1988) *Nucleic Acids Res.* **16,** 6224.
10. Church, S. L., Grant, J. W., Meese, E. U. & Trent, J. M. (1992) *Genomics* **14,** 823–825.
11. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol.* **67,** 835–851.
12. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol.* **67,** 852–862.
13. Blobel, G. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 1496–1500.
14. Ryan, K. R. & Jensen, R. E. (1995) *Cell* **83,** 517–519.
15. Eilers, M. & Schatz, G. (1986) *Nature (London)* **322,** 228–232.
16. Wong, G. H. W. (1995) *Biochim. Biophys. Acta* **1271,** 205–209.
17. Richter, C., Park, J.-W. & Ames, B. N. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6465–6467.
18. Ames, B. N., Shigenaga, M. K. & Hagen, T. M. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7915–7922.
19. Cortopassi, G. A., Shibata, D., Soong, N.-W. & Arnheim, N. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7370–7374.
20. Wallace, D. C. (1992) *Science* **256,** 628–632.
21. Wallace, D. C. (1992) *Annu. Rev. Biochem.* **61,** 1175–1212.
22. Chen, Y., Chen, C.-F., Riley, D. J., Allred, D. C., Chen, P.-L., Von Hoff, D., Osborne, C. K. & Lee, W.-H. (1995) *Science* **270,** 789–791.