Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 6511–6514, July 1994
Cell Biology

Loss of a histidine residue at the active site of S-locus ribonuclease is associated with self-compatibility in *Lycopersicon peruvianum*

JOAQUIN ROYO*, CAROLINE KUNZ†, YASUO KOWYAMA‡, MARILYN ANDERSON*, ADRIENNE E. CLARKE*§, AND ED NEWBIGIN*

*Plant Cell Biology Research Centre, School of Botany, University of Melbourne, Parkville, Victoria 3052, Australia; †Laboratoire du Métabolisme, Physiologie et Biochimie Végétales, Institut National de la Recherche Agronomique, 78026 Versailles, France; and ‡Faculty of Bioresources, Mie University, Tsu 514, Japan

Contributed by Adrienne E. Clarke, March 25, 1994

ABSTRACT **Gametophytic self-incompatibility in the Solanaceae is controlled by a single, multiallelic locus, the *S* locus. We have recently described an allele of the *S* locus of *Lycopersicon peruvianum* that caused this normally self-incompatible plant to become self-compatible. We have now characterized two glycoproteins present in the styles of self-compatible and self-incompatible accessions of *L. peruvianum*: one is a ribonuclease that cosegregates with a functional self-incompatibility allele (S_6 allele); the other cosegregates with the self-compatible allele (S_c allele) but has no ribonuclease activity. The derived amino acid sequences of the cDNAs encoding the S_6 and S_c glycoproteins resemble sequences of other ribonucleases encoded by the *S* locus. The derived sequence for the S_c glycoprotein differs from the others by lacking one of the histidine residues found in all other *S*-locus ribonucleases. These findings demonstrate the essential role of ribonuclease activity in self-incompatibility and lend further weight to evidence that this histidine residue is involved in the catalytic site of the enzyme.**

Self-incompatibility is a major factor affecting mating systems in flowering plants (1, 2). In plants with gametophytic self-incompatibility such as members of the Solanaceae, rejection or acceptance of pollen tubes by the style is controlled by a single, multiallelic locus, the *S* locus. Pollen expresses its haploid *S* genotype, and matings are incompatible if the *S* allele of the pollen is matched by one of the two alleles expressed in the pistil. Thus, self-incompatibility is an example of recognition between plant cells; the underlying mechanism may be similar to other recognition systems in plants such as those involved in host-pathogen interactions (3, 4). The products of the *S* locus are a class of extracellular glycoproteins with RNase activity called S-RNases (5, 6). The genes that encode these proteins cosegregate with alleles of the *S* locus (7, 8). S-RNases are abundant proteins found in high concentrations in the transmitting tract of the style, the site at which inhibition of pollen tubes occurs during incompatible matings (9). Sequences of S-RNase alleles from different solanaceous species share a characteristic structure that includes five short stretches of highly conserved sequence (10). Two of these conserved regions correspond to the sequences surrounding the catalytic domains of fungal RNases and include both of the histidine residues essential for catalytic activity (11).

Recently, Lee *et al.* (12) and Murfett *et al.* (13) have shown that manipulating the expression of S-RNases in *Petunia inflata* or a hybrid *Nicotiana* affected the self-incompatibility phenotype of the style. This finding provided direct evidence for the role of the stylar S-RNase protein in self-incompatibility. We have worked with *Lycopersicon peruvianum*, a species that is almost entirely self-incompatible except for a single self-compatible accession (LA2157) found in Peru growing near a population of self-incompatible plants (accession LA2163; refs. 14 and 15). Previously, we demonstrated that self-compatibility in LA2157 is controlled by a single gene that behaves as an allele of the *S* locus in a series of crosses with LA2163 (16), a result also reported by Bernatzky and Miller (17). We called this nonfunctional allele, S_c. Here, we report the purification of the protein encoded by the S_c allele, the cloning and sequencing of a cDNA corresponding to this protein, and the comparison of this sequence to that derived from the S_6 allele isolated similarly from the LA2163 accession. We demonstrate that the S_c allele encodes a protein that lacks RNase activity presumably due to the lack of a histidine that is thought to be at the active site of the enzyme; these findings reinforce other evidence that the RNase activity of stylar S glycoproteins is essential for the expression of self-incompatibility.¶

MATERIALS AND METHODS

Plant Materials. Two accessions of *L. peruvianum* (LA2157 and LA2163) were collected in the province of Cajamarca, northern Peru, by Charles Rick (Tomato Genetics Resource Center, University of California, Davis). LA2157 is self-compatible and has *S* genotype S_cS_c; LA2163 is self-incompatible and has the *S* genotype S_6S_7 (16). *L. peruvianum* plants homozygous for the S_6 allele were produced by self-pollinating heterozygous individuals at the green bud stage as described (16).

Purification and N-Terminal Sequencing of the S_6 and S_c Glycoproteins. Extracts from 50 styles of plants homozygous for the S_6 or S_c alleles were prepared and fractionated by cation-exchange chromatography as described (18). S glycoproteins recovered from the extract after ammonium sulfate fractionation (50–95%) were desalted on a Sephadex G-25 column and applied to a Mono S (HR 5/5) cation-exchange column (Pharmacia). The column was washed with a linear NaCl gradient (0–0.5 M) and the eluant was continuously monitored for absorbance at 280 nm. Each fraction was assayed for RNase activity as described (6).

Following initial enrichment by cation-exchange chromatography, fractions containing S glycoproteins were rechromatographed on a reverse-phase column (C-8, Brownlee Lab) using a gradient of acetonitrile (0–60%) in 0.1% trifluoroacetic acid. The eluant from the column was continuously monitored at 280 nm, the peak containing S glycoproteins was collected, and the N-terminal sequence of the

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.

Abbreviations: *S* locus, self-incompatibility locus; S-RNase, *S*-locus ribonuclease; S_c allele, self-compatible allele.

§To whom reprint requests should be addressed.

¶The sequences reported in this paper have been deposited in the GenBank data base [accession nos. Z26581 (S_c glycoprotein) and Z26583 (S_6 glycoprotein)].

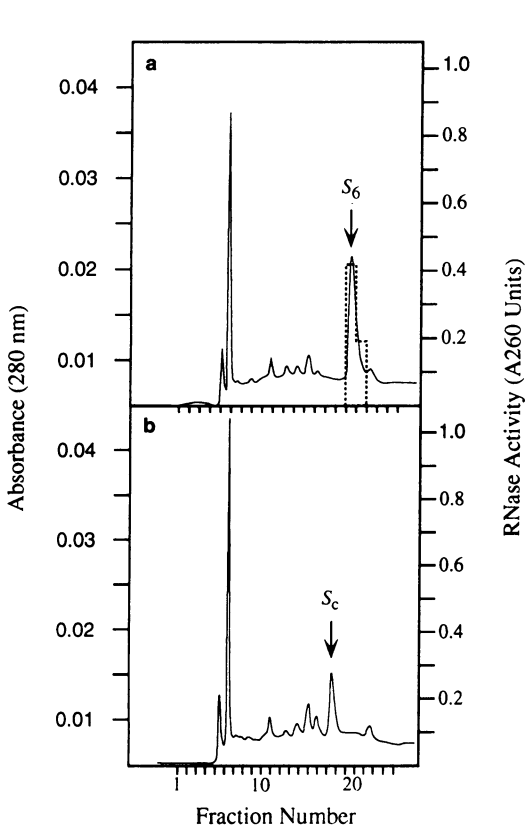

FIG. 1. (*A*) Elution profiles of style extracts from self-compatible and self-incompatible accessions of *L. peruvianum*. (*a*) Profile after cation-exchange chromatography of style extracts obtained from plants homozygous for the S_6 allele. (*b*) Profile after cation-exchange chromatography of style extracts obtained from plants homozygous for the S_c allele. Chromatography was performed on a Mono S cation-exchange column washed with a linear salt gradient (0–0.5 M). The eluant was continuously monitored for absorbance at 280 nm (solid line) and RNase activity of each fraction was measured (dotted line). The S_6 glycoprotein co-eluted with a major peak of RNase activity; no RNase activity was detected in the peak containing the S_c glycoprotein. (*B*) N-terminal amino acid sequence of the S_6 and S_c glycoproteins. Fractions containing S glycoproteins (*A*) were rechromatographed on a reverse-phase column prior to automated sequencing as described by Mau *et al.* (19). The amino acid sequences of both glycoproteins can be aligned with those of other S-RNases (see Fig. 3*B*).

proteins was determined by automated Edman degradation on a gas-phase sequencer as described (19).

RNA Extraction and PCR. RNA was isolated from styles of S_6S_6 or S_cS_c plants using the hot-phenol method (20) and first-strand cDNA was synthesized from 3 μg of total RNA by avian myeloblastosis virus reverse transcriptase (BRL) using oligo(dT) as a primer. PCR was performed on the cDNA template using *Taq* DNA polymerase (Perkin-Elmer) according to the manufacturer's recommended protocol and primers pC2f [5′-AA(T/C)TT(T/C)AC(A/C/G/T)(A/G)T(A/C/G/T)CA(T/C)GG(A/C/G/T)(C/T)T(A/C/G/T)TGGC] and pC5 [5′-GTCGAAACATATACCTATCTCC] that correspond in sequence to conserved regions C2 and C5 of S-RNases (10, 21). The amplified products were ligated into plasmid pCRII (Invitrogen) and sequenced to confirm similarity to S-RNase sequences. Fragment 2B20 was derived by the above protocol using RNA from S_6S_6 styles as starting material and fragment FB33 was obtained using RNA from S_cS_c styles.

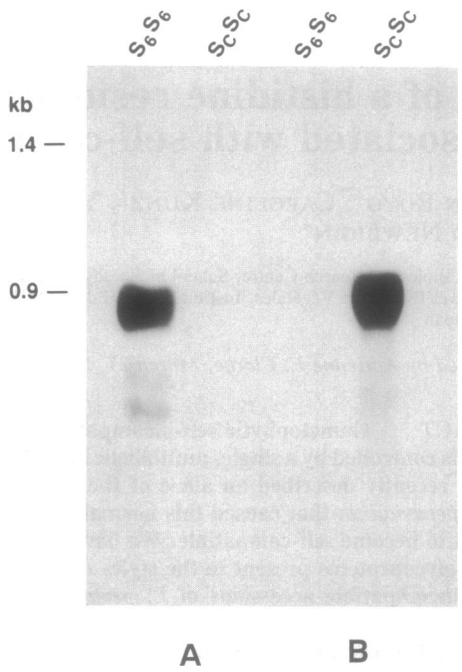

FIG. 2. Northern analysis of RNA from styles of S_6 or S_c homozygous plants probed with DNA fragments specific to each genotype. The DNA fragments were obtained by reverse transcription of style RNA, followed by PCR using as primers, oligonucleotides based on conserved domains found in S-RNases (10). The probe 2B20 was obtained by this procedure using RNA from styles of S_6S_6 plants as a template; 2B20 detects a message of ≈1 kb present in the S_6S_6 genotype but not the S_cS_c genotype (*A*). The probe FB33 was obtained by this procedure using RNA from styles of S_cS_c plants as a template; FB33 detects a message of ≈1 kb present in the S_cS_c genotype but not the S_6S_6 genotype (*B*). The sizes of the molecular mass markers (BRL) are in kb.

Northern Blot Analyses. Samples containing 5 μg of RNA were electrophoresed on formaldehyde/1.5% (wt/vol) agarose gels and transferred to Hybond-N filters (Amersham). The PCR-derived DNA fragments, 2B20 and FB33, were labeled with [^{32}P]dCTP and hybridized to the filters, which were then washed in 0.1× SSC/0.1% SDS at 50°C for 20 min before being exposed to film.

Isolation of cDNA Clones. Double-stranded cDNA, synthesized from 2.5 μg of poly(A)$^+$ RNA isolated from styles of S_cS_c or S_6S_6 individuals, was used in the preparation of cDNA libraries using a commercial kit (Stratagene). Libraries were screened with ^{32}P-labeled 2B20 or FB33, plaques that hybridized to these probes were isolated, and plasmid DNA was purified by standard procedures. The sequence of the cDNA inserts was determined using a *Taq* DyeDeoxy sequencing kit (Applied Biosystems) and analyzed on an Applied Biosystems 373A DNA sequencer.

RESULTS AND DISCUSSION

Although *L. peruvianum* is a highly polymorphic species, the accessions LA2163 and LA2157 are morphologically identical, indicating that they are closely related (16). The two accessions differ in that LA2163 is self-incompatible, as are all other available accessions of *L. peruvianum*, whereas LA2157 is self-fertile. Styles of S_c homozygous plants lack RNase activity and the S_c allele cosegregates with a protein that resembles other S-RNases in its biochemical properties (16).

Protein extracts from styles of plants of the LA2157 accession (genotype S_cS_c) and plants derived from the LA2163 accession (genotype S_6S_6) were fractionated by cation-exchange chromatography, and the individual fractions were assayed for RNase activity (Fig. 1*A*). In both extracts, a major protein peak that eluted from the column with ≈0.4 M NaCl was detected. This peak coeluted with a single peak of RNase activity in the extracts from S_6S_6 styles (Fig. 1*A, a*), while no RNase activity was detected in the protein fractions from S_cS_c styles. The proteins in these major peaks from S_6S_6 or S_cS_c stylar extracts were examined by SDS gel electrophoresis and contained a single major polypeptide of 30.4 and 29.7 kDa, respectively (data not shown). These polypeptides comigrated with the stylar proteins that were previously shown to be associated with the S_6 or S_c alleles (16). The N-terminal sequences of these proteins were determined (Fig. 1*B*) and could be aligned with the sequence of conserved amino acids found at the N terminus of S-RNases (see Fig. 3*B*).

Two probes for genes that encode S glycoproteins were obtained by PCR using cDNA derived from stylar RNA as template and oligonucleotides that matched regions of conserved sequence in S-RNases as primers. The region of DNA amplified by these primers includes the "hypervariable" domain of S-RNases, a region of sequence rich in variation between different alleles (2, 10). One of the probes (FB33) was produced by using RNA from S_cS_c styles and the other (2B20) by using RNA from S_6S_6 styles. The probes do not hybridize with each other (data not shown) and only hybridized to stylar RNA of the genotype from which they were derived (Fig. 2). Both probes detected a mRNA of ≈1 kb. These probes were used to screen cDNA libraries prepared from styles of either the S_cS_c or the S_6S_6 genotype, and the sequence of representative clones from each library is shown in Fig. 3. The derived amino acid sequences at the N termini of these clones correspond exactly with those obtained from the purified S_6 and S_c glycoproteins, and overall, the two sequences share 78% positional identity. Significantly, the S_c allele has an asparagine residue at position 33, a position occupied by a histidine residue in all other S-RNases sequenced to date (Fig. 3*B*). This substitution in the S_c allele results from a single nucleotide alteration (C → A) at the first base of the codon. The replacement of this histidine is particularly interesting because the sequence motifs surrounding the histidine residues at positions 33 and 96 (C2 and C3 in Fig. 3*B*) are common to all members of the RNase T_2 family (2, 6, 23) and are essential for enzymatic activity of S-RNases (Y. H. Liu and A.E.C., unpublished results) and fungal RNases (11).

Because of the predominance of self-incompatibility in *L. peruvianum*, Rick (15) considered that self-compatibility in LA2157 must be derived from a self-incompatible progenitor. We have observed that self-compatibility in this accession results from a nonfunctional allele of the *S* locus and, furthermore, this allele encodes a protein that lacks RNase activity (16). Neither the self-incompatible progenitor of LA2157 nor the functional allele from which the S_c allele was derived is available. Thus, we cannot exclude the possibility that changes other than the histidine-to-asparagine change at position 33 were involved in the conversion of the self-incompatible ancestral line into the modern self-compatible accession. Further testing is necessary to determine whether this single amino acid change is sufficient to change the self-incompatibility phenotype to a self-compatible phenotype by expressing a modified version of the S_c allele in transgenic *L. peruvianum*. Notwithstanding this limitation, we have shown that the RNase activity of these style glycoproteins is an essential part of self-incompatibility.

Secreted RNases are distributed widely in eukaryotic organisms (24, 25). Some RNases such as angiogenin, a potent inducer of blood vessel growth (26), have hormonal functions with the RNase activity playing some role in the biological action (27, 28). In our study of self-incompatibility in plants, the availability of a natural mutant that had lost both stylar RNase activity and self-incompatibility provided a system to demonstrate the direct involvement of this enzymic

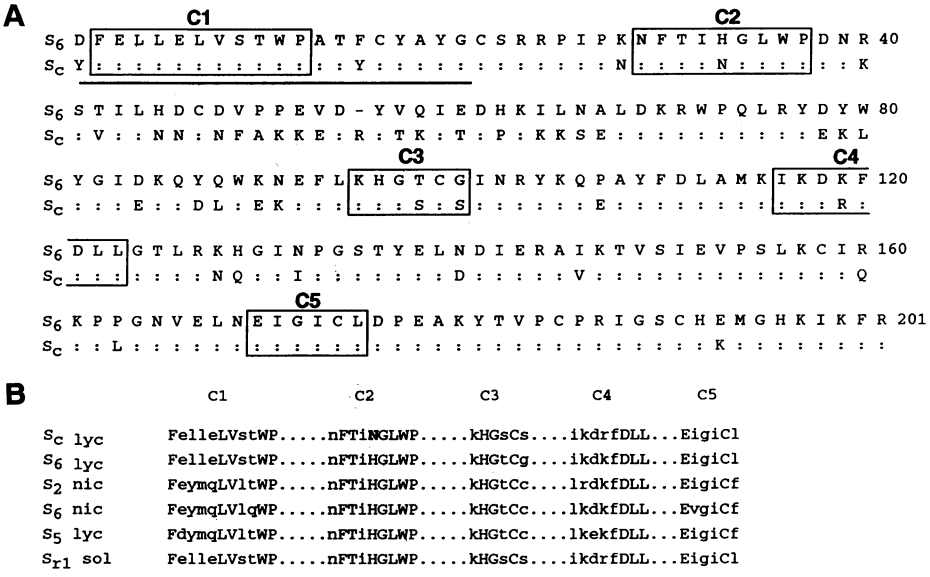

FIG. 3. (*A*) Derived amino acid sequence of the S_6 allele from *L. peruvianum* (shown in the one-letter code) and comparison of this to the S_c allele. Identical amino acids in the S_c allele are indicated as: and gaps as –. Only the amino acids present in the mature S_6 and S_c glycoproteins are considered in this comparison, and the N terminus determined from the purified proteins (Fig. 1*B*) is underlined. Boxed regions show the conserved domains, C1–C5 (10). All residues normally conserved in S-RNases (10) are present in the S_c glycoprotein except the histidine residue at position 33, which is replaced by an asparagine. This amino acid change results from a single change in the codon (CAC → AAC). (*B*) Alignment of the conserved domains (C1–C5) from five representative S-RNases and the S_c glycoprotein. Uppercase letters indicate amino acids present in all sequenced S-RNases (10). The S_c lyc and S_6 lyc sequences are *L. peruvianum* sequences taken from this paper; S_2 nic and S_6 nic are sequences of two alleles from *Nicotiana alata* (7), S_5 lyc is another allele from *L. peruvianum* (10), and S_{r1} sol is from *Solanum tuberosum* (22). The asparagine residue present in the S_c glycoprotein sequence is indicated in bold-face type.

function in the physiological role of S-RNases. However, the enzymic function of the protein is unlikely to be the sole determinant of self-incompatibility; in particular, the hypervariable region of S-RNases is believed to be required for allelic specificity (10). Furthermore, mutational studies indicate that other genes within the *S* locus are essential for self-incompatibility (for example, see ref. 29). This is confirmed by the presence of RNases similar to S-RNases in the pistils of a number of self-compatible solanaceous plants (ref. 30; J. Golz and M.A., unpublished results). Thus, it appears that the RNase activity of S-RNases is necessary but not sufficient to determine self-incompatibility in solanaceous plants.

We are grateful to Bruce McGinness for his assistance and care in the glasshouse and to Dr. Richard Simpson (Ludwig Institute for Cancer Research, Melbourne, Australia) for N-terminal sequencing of the S_c and S_6 glycoproteins. J.R. is supported by a Formación de Personal Investigador postdoctoral fellowship from the Spanish Ministerio de Educación y Ciencia and C.K. is supported by the Swiss Foundation for Scientific Research.

1. Nettancourt, D. de (1977) *Incompatibility in Angiosperms* (Springer, New York).
2. Newbigin, E. J., Anderson, M. A. & Clarke, A. E. (1993) *Plant Cell* **5**, 1315–1324.
3. Hogkin, T., Lyon, G. D. & Dickinson, H. G. (1988) *New Phytol.* **110**, 557–569.
4. Matton, D. P., Nass, N., Clarke, A. E. & Newbigin, E. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 1992–1997.
5. Anderson, M. A., Cornish, E. C., Mau, S.-L., Williams, E. G., Hoggart, R., Atkinson, A., Bönig, I., Grego, B., Simpson, R., Roche, P. J., Haley, J. D., Penschow, J. D., Niall, H. D., Tregear, G. W., Coghlan, J. P., Crawford, R. J. & Clarke, A. E. (1986) *Nature (London)* **321**, 38–44.
6. McClure, B. A., Haring, V., Ebert, P. R., Anderson, M. A., Simpson, R. J., Sakiyama, F. & Clarke, A. E. (1989) *Nature (London)* **342**, 955–957.
7. Anderson, M. A., McFadden, G. I., Bernatzky, R., Atkinson, A., Orpin, T., Dedman, H., Tregear, G., Fernley, R. & Clarke, A. E. (1989) *Plant Cell* **1**, 483–491.
8. Rivers, B. A., Bernatzky, R., Robinson, S. J. & Jahnen-Dechent, W. (1993) *Mol. Gen. Genet.* **238**, 419–427.
9. Cornish, E. C., Pettitt, J. M., Bönig, I. & Clarke, A. E. (1987) *Nature (London)* **326**, 99–102.
10. Tsai, D.-S., Lee, H. S., Post, L. C., Kreiling, K. M. & Kao, T.-h. (1992) *Sex. Plant Reprod.* **5**, 256–263.
11. Kawata, Y., Sakiyama, F., Hayashi, F. & Kyogoku, Y. (1990) *Eur. J. Biochem.* **187**, 255–262.
12. Lee, H.-S., Huang, S. & Kao, T.-h. (1994) *Nature (London)* **367**, 560–563.
13. Murfett, J., Atherton, T. L., Mou, B., Gasser, C. S. & McClure, B. A. (1994) *Nature (London)* **367**, 563–566.
14. Rick, C. M. (1982) *TGC Rep.* **32**, 43–44.
15. Rick, C. M. (1986) in *Solanaceae: Biology and Systematics*, ed. D'Arcy, W. G. (Columbia Univ. Press, New York), pp. 477–495.
16. Kowyama, Y., Kunz, C., Lewis, I., Newbigin, E., Clarke, A. E. & Anderson, M. A. (1994) *Theor. Appl. Genet.*, in press.
17. Bernatzky, R. & Miller, D. D. (1994) *Sex. Plant Reprod.*, in press.
18. Jahnen, W., Batterham, M. P., Clarke, A. E., Moritz, R. L. & Simpson, R. J. (1989) *Plant Cell* **1**, 493–499.
19. Mau, S.-L., Williams, E. G., Atkinson, A., Anderson, M. A., Cornish, E. C., Simpson, R. J., Kheyr-Pour, A. & Clarke, A. E. (1986) *Planta* **169**, 184–191.
20. De Vries, S., Hoge, H. & Bisseling, T. (1988) in *Plant Molecular Biology Manual*, eds. Gelvin, S. B., Schilperoot, R. A. & Verma, D. P. S. (Kluwer, Dordrecht, The Netherlands), pp. 1–13.
21. Dodds, P. N., Bönig, I., Du, H., Rödin, J., Anderson, M. A., Newbigin, E. & Clarke, A. E. (1993) *Plant Cell* **5**, 1771–1782.
22. Kaufmann, H., Salamini, F. & Thompson, R. D. (1991) *Mol. Gen. Genet.* **226**, 457–466.
23. Schneider, R., Unger, G., Stark R., Schneider-Scherzer, E. & Theil, H.-J. (1993) *Science* **261**, 1169–1171.
24. D'Alessio, G., Di Donato, A., Parente, A. & Piccoli, R. (1991) *Trends Biochem. Sci.* **16**, 104–106.
25. Youle, R. J., Newton, D., Wu, Y.-N., Gadina, M. & Rybak, S. M. (1993) *Crit. Rev. Ther. Drug Carrier Syst.* **10**, 1–28.
26. Fett, J. W., Strydom, D. L., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F. & Vallee, B. L. (1985) *Biochemistry* **24**, 5480–5486.
27. Shapiro, R. & Vallee, B. L. (1989) *Biochemistry* **28**, 7401–7408.
28. Shapiro, R., Fox, E. A. & Riordan, J. F. (1989) *Biochemistry* **28**, 1726–1732.
29. Lewis, D. (1949) *Heredity* **3**, 339–355.
30. Ai, Y., Kron, E. & Kao, T.-h. (1992) *Mol. Gen. Genet.* **230**, 353–358.