24. Molina, T. J. *et al. Nature* **357,** 161–164 (1992).
25. Mombaerts, P., Anderson, S. J., Perlmutter, R. M., Mak, T. W. & Tonegawa, S. *Immunity* **1,** 261–267 (1994).
26. Arpaia, E., Shahar, M., Dadi, H., Cohen, A. & Roifman, C. *Cell* **76,** 947–958 (1994).
27. Chan, A. *et al. Science* **264,** 1599–1601 (1994).
28. Elder, M. *et al. Science* **264,** 1596–1599 (1994).
29. Kishihara, K. *et al. Cell* **74,** 143–156 (1993).
30. Danska, J. S. *et al. Science* **266,** 450–455 (1994).

ACKNOWLEDGEMENTS. We thank G. Knowles and Z. Madden for technical support, A. Nagy for advice on ES cell technology, and J. Danska and S. Egan for helpful discussions and critical reading of the manuscript. This work was supported by a fellowship from the Human Frontiers Science Program (K.-D.F.), and grants from the Medical Research Council of Canada and the National Cancer Institute of Canada (A.B. and C.G.), and by Bristol-Myers Squibb (A. B.). A.B. is an International Research Scholar of the Howard Hughes Medical Institute.

Direct demonstration of an intramolecular SH2–phosphotyrosine interaction in the Crk protein

Michael K. Rosen*†, Toshio Yamazaki*, Gerald D. Gish†, Cyril M. Kay‡, Tony Pawson† & Lewis E. Kay*

* Protein Engineering Network Centres of Excellence and Departments of Medical Genetics, Biochemistry and Chemistry, Medical Sciences Building, University of Toronto, Toronto, Ontario M5S 1A8, Canada
† Programme in Molecular Biology and Cancer, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario M5G 1X5, Canada
‡ MRC Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

MANY signal transduction processes are mediated by the binding of Src-homology-2 (SH2) domains to phosphotyrosine (pTyr)-containing proteins[1]. Although most SH2–pTyr interactions occur between two different types of molecules, some appear to involve only a single molecular type. It has been proposed that the enzymatic activity and substrate recognition of the Src-family kinases[2–4], and the protein-binding and transforming activity of Crk-family adaptor proteins[5], are regulated by intramolecular SH2–pTyr interactions. In addition, the DNA-binding activity of Stat transcription factors seems to be regulated by SH2-mediated homodimerization[6]. Here we examine the phosphorylated and non-phosphorylated forms of murine Crk II (p-mCrk and mCrk, respectively)[7–9] using a combination of physical techniques. The Crk protein contains a single SH2 domain and two SH3 domains in the order SH2-SH3-SH3. There is a tyrosine-phosphorylation site between the two SH3 domains at residue 221 which is phosphorylated *in vivo* by the Abl tyrosine kinase[5]. Using NMR spectroscopic analysis, we show here that the SH2 domain of purified p-mCrk is bound to pTyr, and by hydrodynamic measurements that the phosphorylated protein is monomeric. These results provide direct demonstration of an intramolecular SH2–pTyr interaction in a signalling molecule.

To investigate the possible interaction of pTyr221 with the Crk SH2 domain, full-length mCrk, and an amino-terminal fragment truncated immediately after the first SH3 domain (mCrk23, residues 1–197) were expressed in bacteria. Figure 1 illustrates a gradient, enhanced-sensitivity ^{1}H–^{15}N heteronuclear single quantum correlation (HSQC) spectrum[10,11] of recombinant mCrk23. Crosspeaks in this spectrum are virtually an exact subset of those in spectra of the full-length protein (Fig. 2). Titration of phosphopeptide I (Fig. 2*a*), which is derived from the sequence of mCrk surrounding Tyr 221, into an NMR sample of ^{15}N-mCrk23 causes significant (at saturating peptide, >25 Hz change in ^{1}H and/or ^{15}N chemical shifts) and progressive changes in the positions of 39 resolvable $^{1}H/^{15}N$ crosspeaks in HSQC spectra of the protein (Fig. 2). A similar titration of a proteolytically generated SH2-domain fragment (residues 1–128) shows that all 39 of these HSQC peaks represent amides in the SH2 domain (data not shown). Thus, the observed shifts are due to a specific interaction of the phosphopeptide with the SH2 domain. Crosspeaks in spectra of mCrk23 plus saturating amounts of phosphopeptide I are a close subset of those observed in spectra of mCrk that has been phosphorylated stoichiometrically on Tyr 221 by the Abl kinase (p-mCrk). This can be seen in Fig. 2*a*, which shows the results of the mCrk23 titration, and in Fig. 2*b–e*, which shows expanded regions of HSQC spectra of mCrk, p-mCrk, mCrk23 and the mCrk23-phosphopeptide I complex, respectively. The spectra indicate that the changes that occur to the ^{1}H and ^{15}N chemical shifts of mCrk23 on addition of phosphopeptide are nearly identical to those that occur in full-length mCrk on phosphorylation. This suggests that phosphorylation of mCrk on Tyr 221 results in binding of the phosphorylated segment of the molecule to the SH2 domain.

We have also studied SH2-dependent changes in the phosphotyrosine ligand through ^{31}P NMR spectroscopy. At pH 6.8, free phosphopeptide I has a ^{31}P chemical shift of −0.05 p.p.m. (Fig. 3*a*), a value similar to that of pTyr at this pH (data not shown). In Fig. 3*b*, addition of excess mCrk23 has caused a large downfield shift in the phosphotyrosine ^{31}P resonance to 0.82 p.p.m. This change is probably due to deprotonation of the phosphate group[12] by Arg 38, which should lie at the bottom of the SH2 binding pocket[13–18]. In support of this explanation, the largest ^{1}H chemical shift change (0.40 p.p.m.) on titration of mCrk23 with phosphopeptide I occurs to the HN^{ε} of an arginine side chain (data not shown). The ^{31}P spectrum of p-mCrk is shown in Fig. 3*c*. Clearly, the ^{31}P chemical shift at 0.77 p.p.m. represents

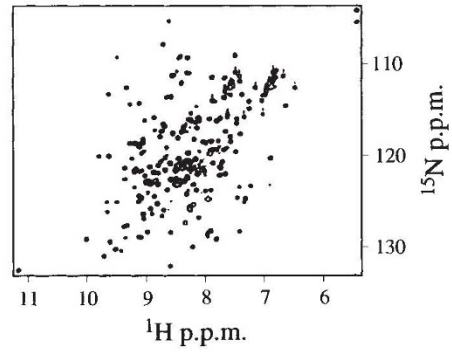


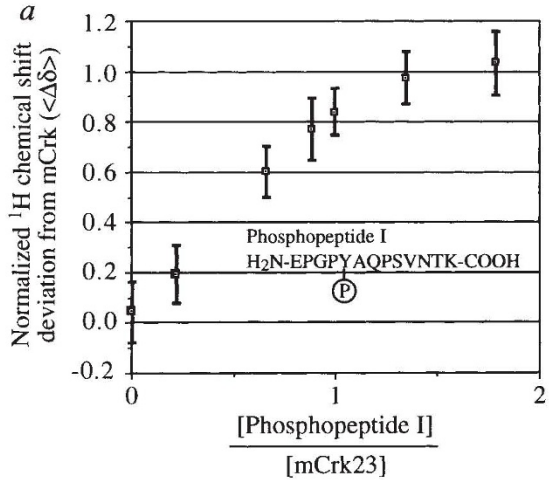
FIG. 1 $^{1}H/^{15}N$ HSQC spectrum of mCrk23. The protein was 0.8 mM in 50 mM MOPS buffer, pH 6.8, 1 mM sodium phosphate, 200 mM NaCl, 2 mM EDTA, 1 mM DTT, 2 mM benzamidine, 0.02% NaN_3, 35 °C. The spectrum was recorded on a Varian UNITY-500 MHz spectrometer equipped with a pulsed-field gradient unit, using reported enhanced sensitivity, pulsed-field gradient pulse sequences[10,11]. ^{1}H and ^{15}N chemical shifts are referenced to external TSP and liquid ammonia, respectively at 0 p.p.m.
METHODS. mCrk23, consisting of residues 1–197 of mCrk, was subcloned into plasmid pET11d (Novagen), and expressed in *E. coli* strain BL21 (DE3). Cells were grown in ^{15}N-labelled M9 minimal medium ($^{15}NH_4Cl$ from Cambridge Isotope Laboratories) to an absorbance at 600 nm of 0.6, and protein synthesis was induced with 0.5 mM IPTG. After 3 h, cells were collected and lysed by sonication. mCrk23 was purified from the cleared lysate by successive anion exchange (Pharmacia DEAE-CL6B resin), hydrophobic exchange (Pharmacia phenyl–Sepharose CL-4B resin) and gel filtration (Pharmacia Sephacryl S-100 resin) chromatography.

an SH2-bound species. This analysis of the phosphate group and the SH2 domain of p-mCrk indicates that there is an interaction between the SH2 domain and pTyr221 in this molecule.

The oligomerization state of p-mCrk was examined by sedimentation equilibrium centrifugation (Fig. 4). The molecular mass calculated using data in the concentration range 1.54–4.86 mg ml^{-1} is 34.0K ± 0.5K (33.4 ± 0.4K for mCrk), in agreement with the actual monomeric molecular mass of 34.181K (34.101K for mCrk). The linear relationship between r^2 and the natural logarithm of optical density indicates that only a single species is present in solution. Dynamic light scattering also indicated that p-mCrk was monomeric at 2.7 mg ml^{-1}, with a weight-average molecular mass of 35K ± 2K (36K ± 2K for mCrk). ^{1}H–^{15}N HSQC spectra of p-mCrk recorded at 2.7 mg ml^{-1} were identical to those shown in Fig. 2 (at 20.5 mg ml^{-1}), indicating the absence of concentration effects on the SH2–pTyr interaction. Further, the agreement of the measured molecular masses of p-mCrk and mCrk shows that phosphorylation does not affect the aggregation state of the protein. The combination of these results with the NMR data demonstrates the existence of an intramolecular SH2–pTyr interaction in p-mCrk.

The NMR experiments described here could be applied to molecules with molecular masses of up to 50–60K. Many signalling proteins in this size range may be regulated by intramolecular interactions like those described here. Our approach, in which this is investigated using intermolecular association of selectively labelled protein fragments, may therefore be generally applicable.

The SH2 domains of proteins in the Crk family bind a number of tyrosine-phosphorylated proteins, including p130^{cas} and paxillin[19–21]. Phosphorylation of Crk on Tyr 221 could regulate these interactions through intramolecular competition for the SH2 binding site. The N-terminal SH3 domains of Crk family members can also bind a variety of signalling molecules, including c-Abl, Sos and C3G (refs 5, 22–24). The full-length proteins,



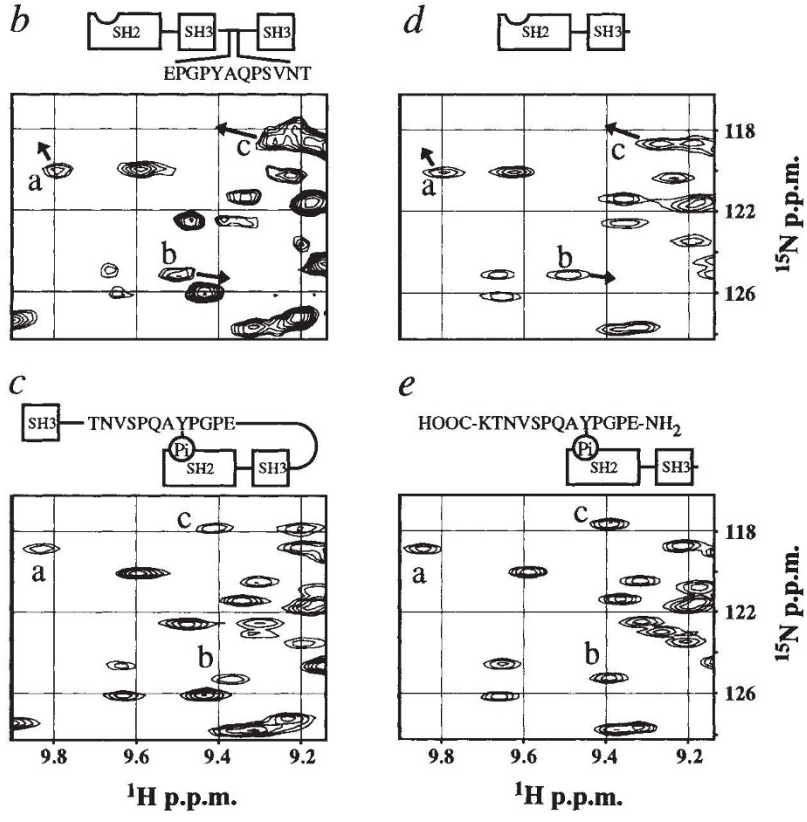


FIG. 2 Chemical shift changes in the Crk SH2 domain induced by phosphorylation or binding of phosphopeptide I. *a*, Plot of $\langle\Delta\delta\rangle$ versus phosphopeptide I concentration, where $\Delta\delta = (\delta - \delta_{mCrk})/(\delta_{p\text{-}mCrk} - \delta_{mCrk})$, δ is the ^{1}H chemical shift of a crosspeak in the mCrk23 sample during the course of the titration (fast exchange on the NMR timescale), and δ_{mCrk} and $\delta_{p\text{-}mCrk}$ are the ^{1}H shifts of the corresponding cross peaks in the spectra of mCrk and p-mCrk, respectively. Data are plotted as the average, $\langle\Delta\delta\rangle$, and standard deviation (represented by error bars) of $\Delta\delta$ for the 17 amide ^{1}H/^{15}N peaks that (1) could be clearly followed throughout the titration; (2) changed more than 25 Hz in the ^{1}H dimension as a result of peptide addition; and (3) were in non-overlapping regions in spectra of both mCrk and p-mCrk. *b*–*e*, Expanded regions of ^{1}H–^{15}N HSQC spectra of mCrk (*b*), p-mCrk (*c*), mCrk23 (*d*) and mCrk23 plus a 40% excess of peptide I (*e*) recorded and referenced as described in Fig. 1 legend. The letters *a*–*c* highlight peaks that shift on phosphorylation or phosphopeptide I binding. Arrows in *b* and *d* show the direction of shift.

METHODS. Phosphopeptide I was synthesized using standard solid-phase methods and purified by reverse-phase HPLC. mCrk was expressed and purified as described for mCrk23 in Fig. 1 legend. Phosphorylation of ~15 mg mCrk was done over 5 days at room temperature in 3.5 ml 120 mM Tris, 30 mM HEPES, pH 7.5, 30 mM NaCl, 4.8 mM DTT, 60 mM $MgCl_2$, 50 mM ATP, using a total of 33 units of p43$^{v\text{-}Abl}$ (Oncogene Science) added in two portions (28 units initially, and 5 additional units at 75 h). Native polyacrylamide gel electrophoresis indicated > 95% conversion to a faster-migrating, phosphotyrosine-containing species. p-mCrk was purified by anion-exchange chromatography. Mass spectroscopic analysis of the ^{15}N-labelled NMR samples of mCrk and p-mCrk showed molecular masses of 34,101 ± 6 and 34,181 ± 3, respectively, indicating only a single phosphorylation site in p-mCrk. Mass spectra of p-mCrk showed no evidence for non-phosphorylated protein. Tryptic digest of the p-mCrk NMR sample followed by liquid chromatography–mass spectroscopic (LC-MS) analysis showed only a single phosphorylated peptide, corresponding to residues Tyr 189 to Arg 240. Isolation of the phosphorylated 189–240 peptide, and partial digestion with thermolysin produced only two phosphorylated species corresponding to residues Leu 213 to pTyr 221 and Leu 213 to Ser 225. As each of these peptides contains only a single tyrosine residue, Tyr 221, we conclude that p-mCrk is phosphorylated uniquely at this residue. In support of this conclusion, mutation of Tyr 221 to Phe resulted in a dramatic decrease in ^{32}P incorporation in p43$^{v\text{-}Abl}$ kinase reactions using [^{32}P]ATP (G.D.G., unpublished results). Concentrations of mCrk, p-mCrk and mCrk23 NMR samples were 0.8, 0.6 and 0.8 mM protein, respectively, in the buffer described in Fig. 1 legend.

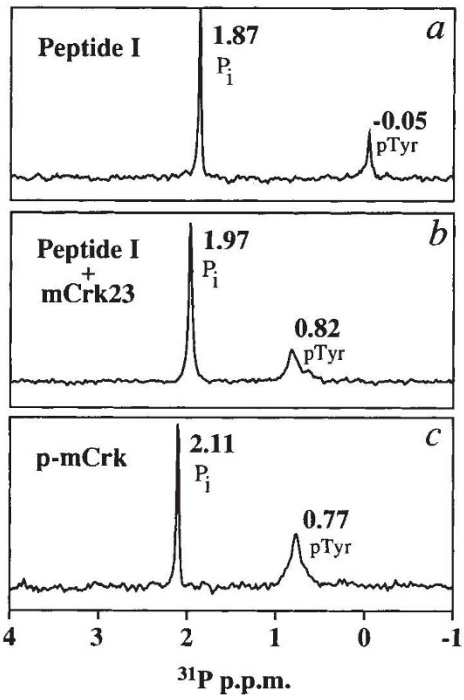


FIG. 3 One-dimensional ^{31}P spectra of peptide I (*a*), peptide I plus a 50% excess of mCrk23 (*b*), and p-mCrk (*c*) recorded in 50 mM MOPS buffer, pH 6.8, 1 mM sodium phosphate (as an internal pH reference), 200 mM NaCl, 2 mM EDTA, 1 mM DTT, 2 mM benzamidine, 0.02% NaN_3. Spectra were recorded at 35 °C at a spectrometer frequency of 202 MHz. Chemical shifts are referenced to external 85% phosphoric acid at 0 p.p.m. The change in chemical shift of the phosphate peak (near 2 p.p.m.) across the three spectra represents a difference of ~0.2 pH units[12].

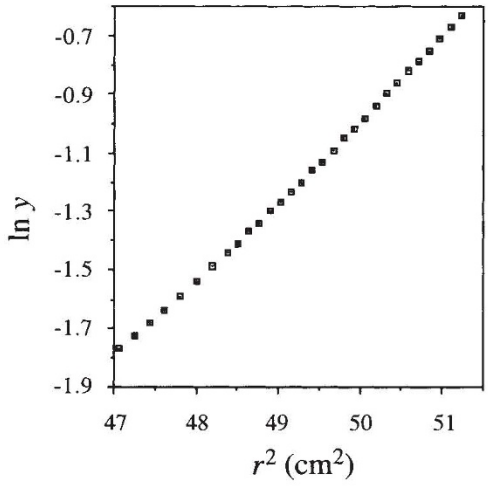


FIG. 4 Phosphorylated Crk is monomeric. A plot of the natural logarithm of the optical density (*y*) as a function of the square of the distance moved in the centrifuge cell (r^2). Data were recorded at 37 °C. Similar results were obtained at 20 °C.

METHODS. All hydrodynamic measurements were made in the buffer used in the NMR experiments (Fig. 1 legend). Analytical ultracentrifugation and data analysis are described in ref. 26. Initial sample concentration was determined from a fringe count assuming an average refractive increment of 4.1 fringes mg^{-1} ml^{-1} according to ref. 27. Determination of molecular mass by light-scattering at 20 °C was essentially as described[26]. A series of protein solutions at 1.06–2.73 mg ml^{-1} were injected directly into a Dawn F multiangle laser light-scattering photometer (Wyatt Technology) using a manual injector. The common intercept on a Zimm plot of the extrapolations to zero angle and zero concentration yielded the reciprocal molecular mass; a value of 0.185 was assumed for d*n*/d*c* for the sample.

however, appear to be able to bind these species only in the non-phosphorylated state[5,25]. Thus, the intramolecular SH2–pTyr interaction may also inhibit intermolecular interactions involving the intervening N-terminal SH3 domain. Phosphorylation then provides a mechanism for regulating the SH2/SH3 adaptor function of Crk. The direct identification of an intramolecular SH2–pTyr interaction in p-mCrk also supports the idea that this is a general mechanism by which the activities of SH2-containing proteins, including Src-family kinases, may be regulated. □

Received 23 November 1994; accepted 6 February 1995.

1. Pawson, A. J. *Nature* **373,** 573–580 (1995).
2. Matsuda, M., Mayer, B. J., Fukui, Y. & Hanafusa, H. *Science* **248,** 1537–1539 (1990).
3. Roussel, R. R., Brodeur, S. R., Shalloway, D. & Laudano, A. P. *Proc. natn. Acad. Sci. U.S.A.* **88,** 10696–10700 (1991).
4. Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S. A. & Draetta, G. *EMBO J.* **12,** 2625–2634 (1993).
5. Feller, S. M., Knudsen, B. & Hanafusa, H. *EMBO J.* **13,** 2341–2351 (1994).
6. Shuai, K. *et al. Cell* **76,** 821–828 (1994).
7. Matsuda, M. *et al. Molec. cell. Biol.* **12,** 3482–3489 (1992).
8. Reichman, C. T., Mayer, B. J., Keshav, S. & Hanafusa, H. *Cell Growth Different.* **3,** 451–460 (1992).
9. Ogawa, S. *et al. Oncogene* **9,** 1669–1678 (1994).
10. Kay, L. E., Keifer, P. & Saarinen, T. *J. Am. chem. Soc.* **114,** 10663–10664 (1992).
11. Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. *J. Biomol. NMR* **4,** 845–858 (1994).
12. Vogel, H. J. *Meth. Enz.* **177,** 263–282 (1989).
13. Kuryian, J. & Cowburn, D. *Curr. Opin. struct. Biol.* **3,** 828–837 (1993).
14. Yu., H. & Schreiber, S. L. *Nature struct. Biol.* **1,** 417–420 (1994).
15. Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D. & Kuryian, J. *Cell* **72,** 779–790 (1993).
16. Eck, M. J., Shoelson, S. E. & Harrison, S. C. *Nature* **362,** 87–91 (1993).
17. Pascal, S. M. *et al. Cell* **77,** 461–472 (1994).
18. Koch, C. A., Anderson, D., Moran, M. F., Ellis, C. & Pawson, T. *Science* **252,** 668–674 (1991).
19. Matsuda, M., Mayer, B. J. & Hanafusa, H. *Molec. cell. Biol.* **11,** 1607–1613 (1991).
20. Birge, R. B. *et al. Molec. cell. Biol.* **13,** 4648–4656 (1993).
21. Sakai, R. *et al. EMBO J.* **13,** 3748–3756 (1994).
22. Ren, R., Ye, Z.-S. & Baltimore, D. *Genes Dev.* **8,** 783–795 (1994).
23. Tanaka, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **91,** 3443–3447 (1994).
24. Matsuda, M. *et al. Molec. cell. Biol.* **14,** 5495–5500 (1994).
25. Oda, T. *et al. J. biol. Chem.* **269,** 22925–22928 (1994).
26. Farrow, N. A. *et al. Biochemistry* **33,** 5984–6003 (1994).
27. Babul, J. & Stellwagon, E. *Analyt. Biochem.* **28,** 216–221 (1969).

ACKNOWLEDGEMENTS. We thank H. Pang for mass spectroscopic analysis of p-mCrk proteolytic fragments; L. Hicks for assistance with sedimentation equilibrium centrifugation and light-scattering measurements; J. McGlade for help in cloning mCrk II; Y. M. Chook for critically reading the manuscript; and Y.M.C. and M. Anafi for discussion. Supported by grants from the National Cancer Institute of Canada (T.P. and L.E.K.) and the Medical Research Council of Canada (T.P., C.M.K. and L.E.K.). T.P. is a Terry Fox Cancer Research Scientist of the NCIC. M.K.R. is supported by the Cancer Research Fund of the Damon Runyon–Walter Winchell Foundation Fellowship. T.Y. is a recipient of a Human Frontiers Science Program Fellowship.

CORRECTION

Positional cloning of the mouse *obese* gene and its human homologue

Yiying Zhang, Ricardo Proenca, Margherita Maffei, Marisa Barone, Lori Leopold & Jeffrey M. Friedman

Nature **372,** 425–432 (1994)

DURING the course of experiments to define the structure of the mouse *obese* gene, it has come to our attention that the first 58 base pairs of the published 5′ untranslated sequence are not found in the genomic DNA corresponding to the first exon of the gene. It is assumed that the presence of this sequence in one of the complementary DNA clones was the result of a cloning artefact. These nucleotides are in part identical to a sequence in the 3′ untranslated sequence. The ambiguities in the restriction map that resulted from an apparent direct repeat in the 5′ and 3′ untranslated region led us to conclude that the polymorphic *Bgl*II site in SM/Ckc+^{Dac}*ob*^{2J} mice was upsteam of the RNA start site. Current data indicate that this polymorphism is the result of genomic alteration in an intron of the *ob* gene (manuscript in preparation). These new data do not affect the coding sequence of the mouse *ob* gene or any of the conclusions in the paper. □