74, 3221 (1995)] are not tractable for the large number of structures and pressures necessary for this study.
15. O. Gunnarsson, E. Koch, R. M. Martin, *Phys. Rev. B* **54**, R11026 (1996).
16. Similarly, conventional LDA calculations accurately predict properties of the perovskites, La(Mn,Fe,Co,Ni)O_3 [D. D. Sarma *et al.*, *Phys. Rev. Lett.* **75**, 1126, (1995)], because the transition metal hopping integrals are ~40% larger than in binary oxides and the screened U is ~40% smaller.
17. P. Dufek, P. Blaha, K. Schwarz, *Phys. Rev. B* **51**, 4122 (1995).
18. M. P. Pasternak *et al.*, *Phys. Rev. Lett.* **65**, 790 (1990).
19. W. E. Jackson, E. Knittle, G. E. Brown Jr., R. Jeanloz, *Geophys. Res. Lett.* **14**, 224 (1987).
20. We found the normal perovskite with Fe in the A site to be 4.6 eV (440 kJ/mol) lower in energy than that with Fe in the B site (LAPW, LDA). We also found wüstite plus stishovite to be lower in energy than $FeSiO_3$ at all pressures, which is consistent with the small solubility of Fe in silicate perovskite. Orthorhombic $MgSiO_3$ was lower in energy than stishovite plus periclase at all pressures, which is consistent with the lack of observed breakdown of $MgSiO_3$ at low temperatures. However, cubic $MgSiO_3$ was marginally unstable with respect to periclase plus stishovite.
21. C. B. Bargeron, M. Avinor, H. G. Drickamer, *Inorg. Chem.* **10**, 1338 (1971); H. G. Drickamer and C. W. Frank, *Electronic Transitions and the High Pressure Chemistry and Physics of Solids* (Halsted, New York, 1973), pp. 126–129.
22. Temperatures are about 2000 K in the lower mantle and higher in the core. We can estimate the effects of temperature from the Clapeyron equation $dP/dT = \Delta S/\Delta V$, where P is pressure, T is temperature, S is entropy, and V is volume. We use the upper bound for the magnetic entropy $S_{mag} = R \ln(M + 1)$, where R is the gas constant. For FeO, we estimate a bound of $dP/dT < 0.015$ GPa/K, which is close to the estimate of 0.01 GPa/K of Ohnishi (*7*). Thus, we expect an increase of about 29 GPa in transition pressure at 2000 K.
23. T. Lay and C. J. Young, *Geophys. Res. Lett.* **17**, 2001 (1990).
24. E. Knittle and R. Jeanloz, *Science* **251**, 1438 (1991).
25. D. E. Loper and T. Lay, *J. Geophys. Res.* **100 B4**, 6397 (1995).
26. H. E. Newsom and K. W. W. Sims, *Science* **252**, 926 (1991).
27. A. Holzheid and H. Palme, *Geochim. Cosmochim. Acta* **60**, 1181 (1996).
28. R. Jeanloz and A. Rudy, *J. Geophys. Res.* **92**, 11433 (1987).
29. N. Uchida and S. Saito, *J. Acoust. Soc. Am.* **51**, 1602 (1972).
30. E. Huang, K. Jy, S.-C. Yu, *J. Geophys. Soc. China* **37**, 7 (1994).
31. V. I. Anisimov, J. Zaanen, O. K. Andersen, *Phys. Rev. B* **44**, 943 (1991).
32. We estimated the effective W by taking the total d band width and subtracting Δ, the splitting between e_g and t_{2g}, from our computations.
33. Supported by NSF grant EAR-9418934. Computations were performed on the Cray J90 at the Geophysical Laboratory. D.G.I. was supported in part by the U.S. Office of Naval Research and NSF grant EAR-9405965. We thank O. K. Andersen, Y. Fei, R. Hazen, R. Hemley, H. Krakauer, C. T. Prewitt, D. J. Singh, and L. Stixrude for helpful discussions.

5 July 1996; accepted 26 November 1996

A General Strategy for Selecting High-Affinity Zinc Finger Proteins for Diverse DNA Target Sites

Harvey A. Greisman and Carl O. Pabo*

A method is described for selecting DNA-binding proteins that recognize desired sequences. The protocol involves gradually extending a new zinc finger protein across the desired 9– or 10–base pair target site, adding and optimizing one finger at a time. This procedure was tested with a TATA box, a p53 binding site, and a nuclear receptor element, and proteins were obtained that bind with nanomolar dissociation constants and discriminate effectively (greater than 20,000-fold) against nonspecific DNA. This strategy may provide important information about protein-DNA recognition as well as powerful tools for biomedical research.

Design of DNA-binding proteins that will recognize desired sites on double-stranded DNA has been a challenging problem. Although a number of DNA-binding motifs have yielded variants with altered specificities, zinc finger proteins related to TFIIIA (*1*) and Zif268 (*2*) appear to provide the most versatile framework for design. Modeling, sequence comparisons, and phage display have been used to alter the specificity of an individual zinc finger within a multifinger protein (*3*–*7*), and fingers also have been "mixed and matched" to construct new DNA-binding proteins (*8, 9*). These design and selection studies have assumed that each finger [with its corresponding 3–base pair (bp) subsite] can be treated as an independent unit (Fig. 1B). This assumption has provided a useful starting point for design studies, but crystallographic studies of zinc finger–DNA complexes (*10*–*13*) reveal many examples of contacts that couple neighboring fingers and subsites, and it is evident that context-dependent interactions are important for zinc finger–DNA recognition (*3, 7, 8*). Existing strategies have not taken these interactions into account in the design of multifinger proteins, and this may explain why there has been no effective, general method for designing high-affinity proteins for desired target sites (*14*).

We have developed a selection strategy that can accommodate many of these context-dependent interactions between neighboring fingers and subsites. Our strategy involves gradual assembly of a new zinc finger protein at the desired binding site—adding and optimizing one finger at a time as we proceed across the target site. We use the Zif268 structure (*10, 13*) as our framework and randomize six potential base-contacting positions in each finger (Fig. 1, A and D) (*15*). Our protocol includes three selection steps (Fig. 2), one for each finger of the new protein: (i) A finger that recognizes the 3′ end of the target site is selected by phage display (Fig. 2A). At this stage, two wild-type Zif fingers are used as temporary anchors to position the library of randomized fingers over the target site, and we use a hybrid DNA site that has Zif subsites fused to the target site. (ii) The selected finger is retained as part of a "growing" protein and, after the distal Zif finger is discarded, phage display is used to select a new finger that recognizes the central region of the target site (Fig. 2B). (iii) Finally, the remaining Zif finger is discarded, and phage display is used to select a third finger that recognizes the 5′ region of the target site (Fig. 2C). Optimization of this finger yields the new zinc finger protein.

Our strategy ensures that the new fingers are always selected in a relevant structural context. Because an intact binding site is present at every stage, and because our selections are performed in the context of a growing protein-DNA complex, our method readily optimizes context-dependent interactions between neighboring fingers and subsites and naturally selects for fingers that will function well together (*16*). To ensure that the selected proteins will bind tightly and specifically to the desired target sites, we performed all selections in the presence of calf thymus competitor DNA (3 mg/ml) (*17*). This serves to counterselect against any proteins that bind promiscuously or prefer alternative sites, and our protocol thus directly selects for affinity as well as specificity of binding (*18*).

We tested our protocol by performing selections with a TATA box, a p53 binding site, and a nuclear receptor element (NRE) (Fig. 1C). These important regulatory sites were chosen because they normally are recognized by other families of DNA-binding proteins and because these sites are quite

Howard Hughes Medical Institute and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed.

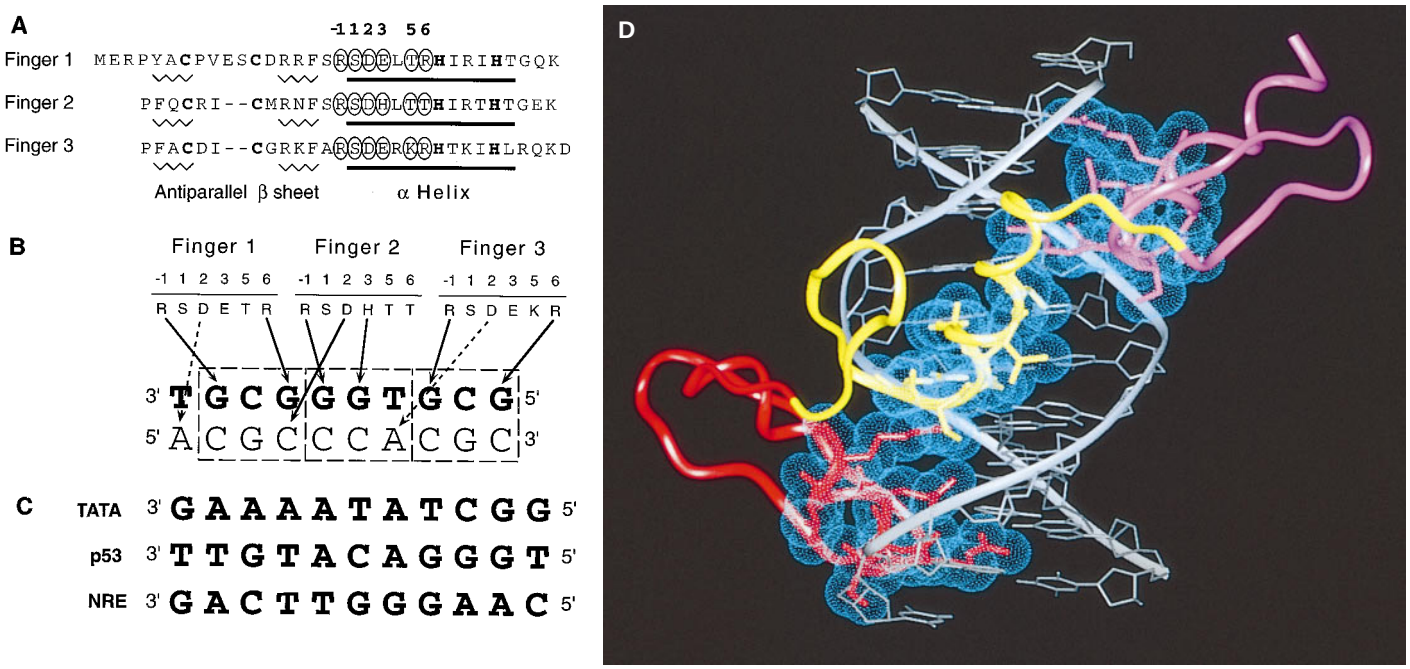

Fig. 1. (**A**) Amino acid sequence and secondary structure of the Zif268 zinc fingers. [Adapted from (*10*)] Randomized positions (circled) correspond to residues −1, 1, 2, 3, 5, and 6 in each of the α helices and include every position that makes a base contact in one of the known zinc finger–DNA complexes (*10–13, 30*). The wild-type Zif268 sequence was retained at all other positions in the new proteins. (**B**) Key base contacts (solid arrows) in the Zif268-DNA complex (*10, 13*). Most of the bases contacted are located on the primary (guanine-rich) strand (boldface). Each finger makes several base contacts with its 3-bp subsite (dashed boxes), but also makes important base and phosphate contacts in flanking subsites. The 1.6 Å structure (*13*) shows that the aspartic acid at position 2 in finger 2 contacts a cytosine that is just outside the canonical 3-bp subsite. Analogous contacts from position 2 in the other fingers (dashed arrows) have less favorable hydrogen-bonding geometry, but binding site selections (*32*) suggest that these contacts may contribute to recognition. Contacts made by Tramtrack (*11*) and GLI (*12*) also include bases and phosphates outside the canonical 3-bp subsites. (**C**) DNA sequences of the sites used in our selections. The TATA box is from the adenovirus major late promoter (*33*), the p53 binding site is from the human $p21^{WAF1/CIP1}$ promoter (*31*), and the NRE is from the human apolipoprotein AI promoter (*34*). One strand of each duplex site is shown. (**D**) Structure of the wild-type Zif268 zinc finger–DNA complex (*10, 13*). The DNA is gray, and a ribbon trace of the three zinc fingers is shown in red (finger 1), yellow (finger 2), and purple (finger 3). The 18 residues that were randomized in this study (van der Waals surfaces shown in blue) occupy the major groove of the DNA and span the entire length of the binding site. [Image created with Insight II (Biosym Technologies, San Diego, California)]

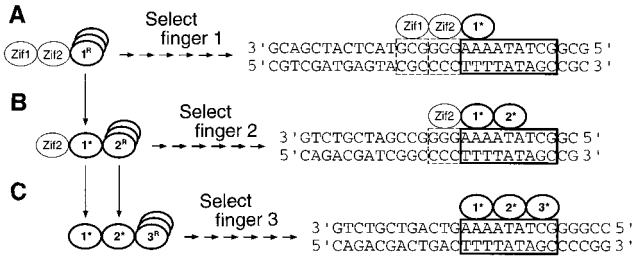

Fig. 2. Overview of protocol that successively selects finger 1, finger 2, and finger 3 to create a new zinc finger protein. Fingers that are present in the phage libraries used in these steps (*15*) are indicated on the left side of each panel. Zif1 and Zif2, wild-type Zif268 fingers; R, a randomized finger library; and asterisk, a selected finger. Small horizontal arrows indicate the multiple cycles of selection and amplification used when selecting each finger by phage display (*35*). The right side of each panel shows the binding sites used in selections with the TATA site and indicates the overall binding mode for the selected fingers [each DNA duplex has biotin (not shown) attached at the 3′ end of the upper strand]. Vertical arrows indicate how fingers selected in earlier steps are incorporated into the phage libraries used in later steps and reselected to optimize affinity and specificity in the new context (*16*). (**A**) A randomized finger 1 library was cloned into the pZif12 phagemid display vector (*36*), and selections with this library were performed in parallel at the TATA, p53, and NRE sites (*17*). (**B**) The wild-type Zif1 finger was removed, and a randomized finger 2 cassette was ligated to the appropriate vector pool and optimized by phage display (*29*). (**C**) The remaining wild-type finger was removed, and a randomized finger 3 cassette was added and optimized by phage display. To construct the sites used in these selections, we fused the target strand with the higher purine content to the guanine-rich strand of the Zif268 site. Because of the overlapping base contacts that can occur at the junction of neighboring subsites (Fig. 1B), the 3′ end of the target site (Fig. 1C) was aligned so that it overlapped with the Zif2 subsite.

different from the guanine-rich Zif268 site and from sites that have been successfully targeted in previous design studies (*14*). After the multiple rounds of selections (Fig. 2) were completed, the final phage pools bound tightly to their respective target sites. DNA sequencing of eight clones from each pool revealed marked patterns of conserved residues (Fig. 3) (*19*), and many of the selected residues (Arg, Asn, Gln, His, and Lys) could readily contribute to base recognition (*20*).

Because of the marked sequence conservation within each of the final phage pools, we used a single clone from each set for further analysis. The corresponding peptides were overexpressed in *Escherichia coli* and purified (*21*). Affinities of the peptides for their respective target sites were determined by electrophoretic mobility shift analysis (*22*), and the measured dissociation constants (K_d's) were 0.12 nM for the TATA box, 0.11 nM for the p53 binding site, and 0.038 nM for the NRE. These new complexes are almost as stable as the wild-type Zif268-DNA complex (K_d of 0.010 nM under these buffer conditions).

Apparent K_d's for nonspecific DNA were estimated by competition experiments with calf thymus DNA (*23*). Ratios of the nonspecific to specific dissociation constants (K_d^{ns}/K_d) indicate that the peptides selected for the TATA box, p53

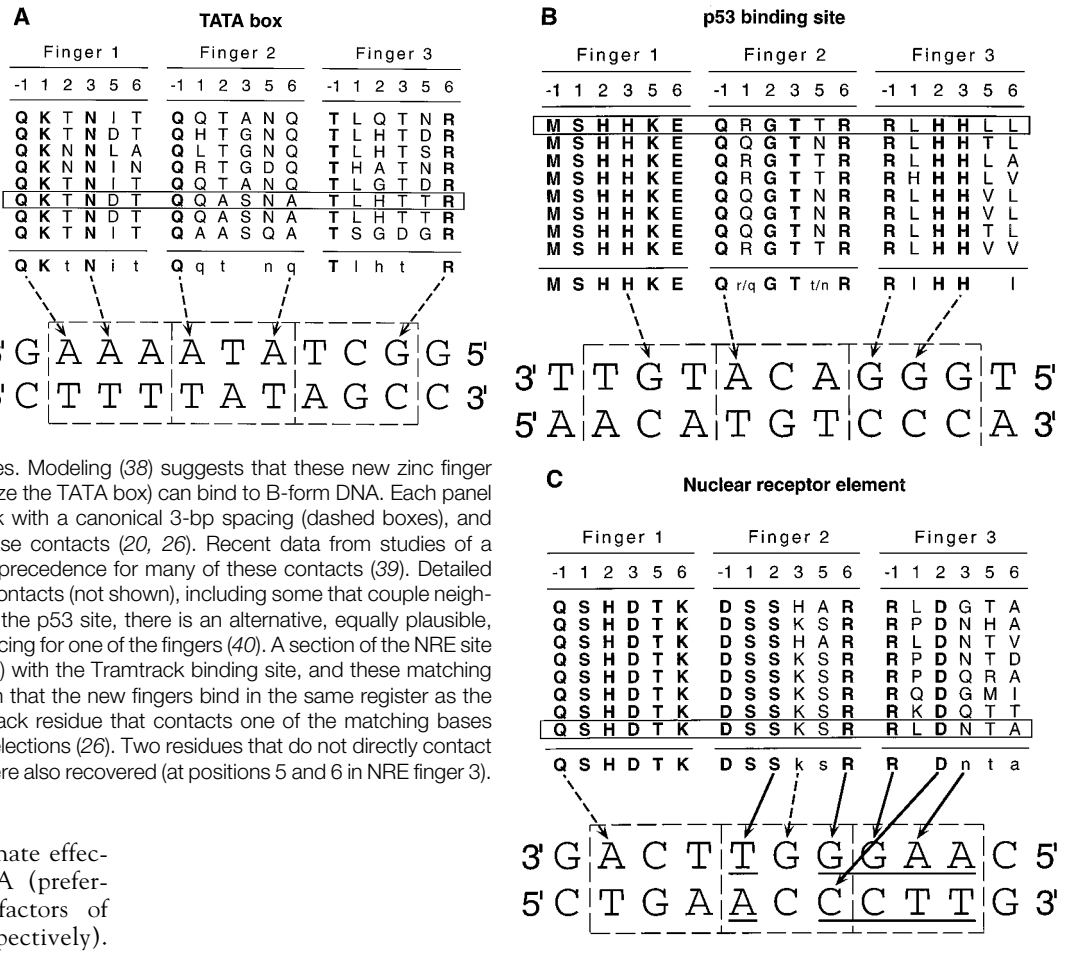

Fig. 3. Amino acid sequences of new zinc finger proteins that recognize (**A**) the TATA box, (**B**) the p53 binding site, and (**C**) the NRE. Residues selected at each of the six randomized positions are shown (*37*). Six or more of the eight clones in each phage pool encode unique zinc finger proteins (*16, 19*). A box indicates the clone that was overexpressed and used for binding studies. Residues that are fully conserved (eight of eight clones) are shown in boldface; residues that are partially conserved (four or more of eight) are denoted by lowercase letters in the consensus sequence below the set of clones. Modeling (*38*) suggests that these new zinc finger proteins (including those that recognize the TATA box) can bind to B-form DNA. Each panel indicates how the fingers could dock with a canonical 3-bp spacing (dashed boxes), and dashed arrows indicate plausible base contacts (*20, 26*). Recent data from studies of a designed zinc finger protein provide precedence for many of these contacts (*39*). Detailed modeling suggests many additional contacts (not shown), including some that couple neighboring fingers and subsites (*38*). For the p53 site, there is an alternative, equally plausible, docking arrangement with a 4-bp spacing for one of the fingers (*40*). A section of the NRE site shows a 5 of 6 bp match (underlined) with the Tramtrack binding site, and these matching segments happen to be aligned such that the new fingers bind in the same register as the Tramtrack fingers (*11*). Every Tramtrack residue that contacts one of the matching bases (solid arrows) was recovered in our selections (*26*). Two residues that do not directly contact the DNA in the Tramtrack complex were also recovered (at positions 5 and 6 in NRE finger 3).

binding site, and NRE discriminate effectively against nonspecific DNA (preferring their specific sites by factors of 25,000, 54,000, and 36,000, respectively). These ratios are similar to the specificity ratio of 31,000 that we measured for wild-type Zif268. Taken together, the affinities and specificities of the new proteins indicate that they bind as well as many natural DNA-binding proteins.

Many discussions of zinc finger–DNA recognition have considered the idea of a "code" that specifies which positions along the α helix contact the DNA and which side chain–base interactions are most favorable at each position (*5, 24*). There are recurring patterns of contacts in some zinc finger proteins (*10, 11*), and similar patterns are apparent in the proteins we selected (Fig. 3). Thus, when adenine or guanine occurs in the primary strand of one of our binding sites (the strand corresponding to the guanine-rich strand of the Zif268 site), there often is a conserved residue at position −1, 3, or 6 of the α helix that could form hydrogen bonds with this base (*20*). Related patterns have been discussed in previous design and selection studies (*3–6*). There also are strong "homologies" between the zinc fingers we have selected and natural zinc fingers that may recognize the same subsites (Fig. 3) (*25*).

Such simple patterns are not seen at other positions in our selected proteins. Thus, we found no simple patterns of residues at positions 1, 2, and 5 of the α helix, and when thymine or cytosine occurs on the primary strand (Fig. 3), we found no simple pattern of potential contacts from residues at positions −1, 3, and 6. However, there still are numerous instances in which residues at these positions are highly conserved within a particular set of proteins (Fig. 3), and we infer that many of these conserved residues make energetically significant contributions to folding or binding (*26*). Because no readily predicted pattern of coded contacts is apparent, we surmise that residues at these positions may be involved in more subtle, context-dependent interactions. In short, there still is no general code that can be used to design optimal zinc finger proteins for any desired target sequence or that can predict the preferred binding site of every zinc finger protein (*27*). Nonetheless, our sequential selection strategy should provide valuable information about potential patterns in zinc finger–DNA recognition, because it (i) makes few assumptions about the preferred spacing, docking, or contacts of the individual fingers; (ii) yields proteins with essentially wild-type affinities and specificities; (iii) yields sequences that match very well with those of natural zinc finger proteins that recognize similar subsites (*25*); and (iv) can readily be adapted to pursue analogous studies with other TFIIIA-like zinc finger proteins.

The sequential selection strategy provides a general and effective method for design of new zinc finger proteins, and our success with a diverse set of target sites suggests that it should be possible to select zinc finger proteins for many important regulatory sequences. These proteins could then be fused with appropriate regulatory or effector domains for a variety of applications. The protocol also could be adapted to allow selection of proteins with four, five, or six fingers or to allow optimization of zinc fingers fused to other DNA-binding domains (*28*). Related selection methods might be developed for other families of multidomain proteins, including other DNA- and RNA-binding proteins, and possibly even modular domains involved in protein-protein recognition. The sequential selection strategy should open the field to a host of applications and studies, including tests to see how designer zinc finger proteins can be used in gene therapy.

REFERENCES AND NOTES

1. J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **4**, 1609 (1985).
2. B. A. Christy, L. F. Lau, D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7857 (1988).
3. J. Nardelli, T. Gibson, P. Charnay, *Nucleic Acids Res.* **20**, 4137 (1992); S. K. Thukral, M. L. Morrison, E. T. Young, *Mol. Cell. Biol.* **12**, 2784 (1992); J. R. Desjarlais and J. M. Berg, *Proteins* **12**, 101 (1992); *ibid.* **13**, 272 (1992); *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7345 (1992); *ibid.* **91**, 11099 (1994).
4. E. J. Rebar and C. O. Pabo, *Science* **263**, 671 (1994).
5. Y. Choo and A. Klug, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11163 (1994); *ibid.*, p. 11168.
6. A. C. Jamieson, S. H. Kim, J. A. Wells, *Biochemistry* **33**, 5689 (1994); H. Wu, W. P. Yang, C. F. Barbas, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 344 (1995).
7. W. E. Taylor *et al.*, *Biochemistry* **34**, 3222 (1995); C. Cheng and E. T. Young, *J. Mol. Biol.* **251**, 1 (1995).
8. J. R. Desjarlais and J. M. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2256 (1993).
9. Y. Choo, I. Sánchez-García, A. Klug, *Nature* **372**, 642 (1994).
10. N. P. Pavletich and C. O. Pabo, *Science* **252**, 809 (1991).
11. L. Fairall, J. W. R. Schwabe, L. Chapman, J. T. Finch, D. Rhodes, *Nature* **366**, 483 (1993).
12. N. P. Pavletich and C. O. Pabo, *Science* **261**, 1701 (1993).
13. M. Elrod-Erickson, M. A. Rould, L. Nekludova, C. O. Pabo, *Structure* **4**, 1171 (1996).
14. "Mix and match" design strategies have, so far, been limited to binding sites in which the primary strand (Fig. 1B) contains at least one guanine within each 3-bp subsite (*3, 8, 9*). The affinities of designed zinc finger proteins also have varied widely, and some K_d's have been in the micromolar range (*8, 9*). Subtle, context-dependent interactions (which provided the motivation for our protocol) may have a critical, cumulative effect when optimizing multifinger proteins: A modest (10-fold) increase in affinity for each finger may yield a substantial (1000-fold) increase in affinity for a three-finger protein.
15. Each cassette encodes one of the Zif268 fingers (Fig. 1A), and randomized codons have A/C/G at the first position, A/C/G/T at the second position, and C/G at the third position. These randomized codons allow 16 side chains at each position (all residues except Cys, Phe, Tyr, and Trp) and they do not give any termination codons. Each cassette encodes a maximum of 16^6 ($\approx 1.7 \times 10^7$) different zinc finger sequences, represented by 24^6 ($\approx 1.9 \times 10^8$) different DNA sequences. All phage display libraries contained between 5.6×10^8 and 1.9×10^9 clones. After the finger 1 selections (Fig. 2A), double-stranded DNA was purified from $\geq 10^5$ optimized phagemids, and the first wild-type Zif finger was removed; transformed colonies ($\geq 10^7$) were pooled, and purified DNA from this pool was used to construct the finger 2 library. After the second selection step (Fig. 2B), a similar protocol was used to remove the remaining wild-type finger from the selected pool and to construct the finger 3 library. To accommodate the restriction sites used in these cloning steps (*29*), we changed residues in the COOH-terminal linker of each randomized finger to TGESR (*30*) for one round of selections; wild-type residues were restored when the next cassette was added.
16. Our protocol actually was designed so that a sublibrary of successful zinc finger sequences could be carried over from one selection step (Fig. 2, A or B) to the next (*15*). Preliminary sequencing data to analyze the "evolutionary history" of our selections (*29*) indicates that a set of finger 1 sequences was carried over into the step in Fig. 2B and that this step then selects for combinations of fingers that work well together (*19*).
17. E. J. Rebar, H. A. Greisman, C. O. Pabo, *Methods Enzymol.* **267**, 129 (1996).
18. Assuming that the calf thymus DNA has one potential binding site per base (that is, binding could conceivably occur in any register on either strand), a 3 mg/ml solution of DNA corresponds to a 0.01 M solution of potential binding sites. (Our specific site is present at 40 nM). If the DNA sequence of this competitor were random, each of the 4^9 (= 262,144) possible 9-bp sites would be present, with an average concentration of about 40 nM.
19. Each set of proteins exhibits a clear gradient of sequence diversity across the three fingers (Fig. 3), but the finger 1 and finger 2 sequences were more diverse at intermediate stages of the optimization protocol (*16, 29*). For example, after the first step (Fig. 2A), many of the TATA clones had Asn residues at position −1 or position 6 or in both locations. After the selections indicated in Fig. 2B, most clones had Gln at position −1 and Thr at position 6 of finger 1, and these residues also are present in a homologous natural finger that recognizes the same subsite (*25*).
20. Based on the Zif268 (Fig. 1B) and Tramtrack (*11*) structures, our alignments assume that residues at position −1 can contact the 3′ base on the primary strand of the subsite, residues at position 3 can contact the central base, and residues at position 6 can contact the 5′ base. Guanine bases in our sites appear to prefer Arg at positions −1 and 6, but His or Lys at position 3. Adenine bases appear to prefer Asn at position 3, but prefer Gln at position −1 and, to some extent, at position 6.
21. Zinc finger regions were subcloned into pET21d (Novagen), and the corresponding peptides (with end points as in Fig. 1A) were expressed in *E. coli* BL21 (DE3) and purified as described (*4*).
22. Dissociation constants were determined essentially as described (*4*). However: (i) each K_d was determined in the absence of competitor DNA; (ii) binding buffer contained 15 mM Hepes-NaOH (pH 7.9), 50 mM KCl, 50 mM potassium glutamate, 50 mM potassium acetate, 5 mM $MgCl_2$, 20 μM $ZnSO_4$, acetylated bovine serum albumin (100 μg/ml), 5% (v/v) glycerol, and 0.1% (w/v) NP-40; (iii) binding reactions contained 2 or 4 pM of the labeled site and were equilibrated for 1 hour; (iv) K_d values were calculated from the slopes of Scatchard plots and represent the average of three independent experiments (SD values were all <60%); and (v) mobility shift assays were performed with double-stranded oligonucleotides containing TTT overhangs at the 5′ end of each strand. The sequences of the primary strands within the duplex regions were 5′-AGGGGGGCTATAAAAGGGGGT-3′ (TATA box), 5′-GCTGTTGGGACATGTTCGTGA-3′ (p53 site), 5′-GCCGTCAAGGGTTCAGTGGGG-3′ (NRE site), and 5′-CCAGTAGCGGGGGCGTCCTCG-3′ (Zif268 site).
23. For competition experiments, 8 pM of labeled specific oligonucleotide (*22*) was mixed with binding buffer containing successive twofold dilutions of calf thymus competitor DNA. An equal volume of binding buffer that contained a fixed amount of protein (sufficient for a 50 to 80% mobility shift in the absence of competitor DNA) was added, after which the reaction mixtures were incubated for ≥1 hour and then subjected to gel electrophoresis (*4*). K_d^{ns} (in μg/ml) was calculated from the slope of a $C_t\theta$ versus plot, using the equation

$$C_t\theta = \left[\frac{-K_d^{ns}}{1-\theta_o}\right]\theta + \left[\frac{K_d^{ns}}{(1-\theta_o)/\theta_o}\right]$$

where θ is the fraction of specific site bound by protein in the presence of competitor DNA (at concentration C_t), and θ_o is the fraction bound in the absence of competitor. This equation was derived from equation 3 of S. Y. Lin and A. D. Riggs [*J. Mol. Biol.* **72**, 671 (1972)]. Each K_d^{ns} value represents the average of six plots (three plots in two independent experiments). All SD values were <25%. When calculating K_d^{ns}/K_d, we assumed that each base in the calf thymus DNA represents the beginning of a potential binding site. A simple estimate for the specificity of these new zinc finger proteins can be made by taking various powers of 4^n and comparing these numbers with the measured specificity ratios. All of our new proteins have specificity ratios that lie between 4^7 (= 16,384) and 4^8 (= 65,536). This indicates that our proteins—like Zif268 itself—can effectively specify 7 to 8 bp in the target DNA sites.
24. J. M. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11109 (1992); R. E. Klevit, *Science* **253**, 1367 (1991); M. Suzuki, M. Gerstein, N. Yagi, *Nucleic Acids Res.* **22**, 3397 (1994).
25. Several of the subsites recognized by our optimized fingers (Fig. 3) also happen to appear in binding sites for the Tramtrack (*11*) and Gfi-1 zinc finger proteins [P. A. Zweidler-McKay, H. L. Grimes, M. M. Flubacher, P. N. Tsichlis, *Mol. Cell. Biol.* **16**, 4024 (1996)], and we find remarkable similarities in the amino acid sequences of the corresponding recognition helices. These homologies include, but are not limited to, the canonical base-contacting residues (*20*) at positions −1, 3, and 6. For example, finger 4 of the Gfi-1 protein and finger 1 of our NRE proteins appear to recognize the subsite 3′-ACT-5′, and the Gfi-1 residues at positions −1, 1, 2, 3, 5, and 6 are QKSDKK (underlined residues match the consensus in the selected fingers) (*30*). Finger 5 of Gfi-1 and finger 1 of the TATA proteins appear to recognize the subsite 3′-AAA-5′, and the corresponding Gfi-1 residues are QSSNIT (*30*).
26. Given the remarkable homology with Tramtrack (Fig. 3), it seems plausible that the Ser and Asp residues at position 2 in NRE fingers 2 and 3 may make the same contacts that corresponding residues make in Tramtrack fingers 1 and 2 (*11*). We also anticipate that the Lys at position 1 in finger 1 of the TATA box proteins may make a phosphate contact analogous to the contact made by Tramtrack finger 2.
27. There are several examples of zinc fingers that have appropriate residues (Arg, His, Asn, or Gln) at positions −1, 3, and 6, but do not make the expected coded contacts with their 3-bp subsites. Examples include some natural fingers, such as finger 3 of GLI (*12*) and finger 2 of ADR1 (*7*), as well as synthetic fingers designed to recognize particular subsites (*3*). As noted by others (*3, 7*), context-dependent interactions may explain these effects.
28. J. L. Pomerantz, P. A. Sharp, C. O. Pabo, *Science* **267**, 93 (1995).
29. H. A. Greisman, thesis, Massachusetts Institute of Technology, Cambridge, MA (1997).
30. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
31. W. S. El-Deiry *et al.*, *Cell* **75**, 817 (1993); W. S. El-Deiry *et al.*, *Cancer Res.* **55**, 2910 (1995).
32. A. H. Swirnoff and J. Milbrandt, *Mol. Cell. Biol.* **15**, 2275 (1995).
33. E. B. Ziff and R. M. Evans, *Cell* **15**, 1463 (1978).
34. J. A. A. Ladias and S. K. Karathanasis, *Science* **251**, 561 (1991).
35. Phage display was performed in an anaerobic chamber to ensure proper folding of the zinc fingers (*4, 17*). Five to eight cycles of selection and amplification were performed for each finger, and retention efficiencies plateaued at values ranging from ~0.2 to 3% of input phage (*17, 29*). Binding reactions for the p53 finger 3 selections contained the nonbiotinylated duplex competitor 5′-CCCTTGGAACATGTTCCTGATCGCGG-3′ (*29*). [The p53 target site is pseudosymmetric (Fig. 1C) (*31*), and we wanted to avoid inadvertently selecting a zinc finger protein that would bind to the opposite strand] The biotinylated sites used in the TATA box selections are shown in Fig. 2, and the sites used for the other selections (*29*) were designed in a similar manner; we altered the Zif268 subsites when they were no longer needed (Fig. 2, B and C) and removed any cryptic binding sites that resembled the binding site of interest.
36. The pZif12 phagemid display vector (*17*) encodes a fusion protein that contains (i) Zif268 fingers 1 and 2 [residues 327 to 391 of the intact protein (*2*)]; (ii) a linker that introduces an amber codon; and (iii) residues 23 to 424 of the M13 gene III protein. The zinc finger region contains a set of restriction sites that were designed to facilitate the multiple cloning steps in our protocol (*29*).
37. Four of the eight p53 clones had a conservative Ser → Thr mutation at position −2 in finger 2; in all other clones, residues outside the randomized regions were identical to those in wild-type Zif268.
38. L. Nekludova, personal communication.
39. C. A. Kim and J. M. Berg, *Nature Struct. Biol.* **3**, 940 (1996).

40. In the alternative arrangement, p53 finger 2 spans a 4-bp subsite (3′-ACAG-5′) and finger 3 recognizes the adjacent 3′-GGT-5′ subsite. A similar spacing occurs at one point in the GLI-DNA complex (*12*).
41. We thank E. Rebar for support, encouragement, reagents, and advice that made this project possible; M. Elrod-Erickson for sharing refined coordinates of the Zif268-DNA complex and for advice on purification of zinc finger peptides; L. Nekludova for extensive discussions about modeling studies with these new zinc finger proteins and for help with computer graphics; W. El-Deiry for providing sequences before publication; and J. Pomerantz, S. Wolfe, E. Fraenkel, M. Elrod-Erickson, and E. Rebar for insightful comments on the manuscript. H.A.G. was supported by a predoctoral fellowship from the Howard Hughes Medical Institute and by a Massachusetts Institute of Technology Department of Biology fellowship from the Centocor Corporation; C.O.P. was supported by the Howard Hughes Medical Institute.

10 July 1996; accepted 20 November 1996

Regulation of Neuronal Survival by the Serine-Threonine Protein Kinase Akt

Henryk Dudek, Sandeep Robert Datta,* Thomas F. Franke,* Morris J. Birnbaum, Ryoji Yao, Geoffrey M. Cooper, Rosalind A. Segal, David R. Kaplan, Michael E. Greenberg†

A signaling pathway was delineated by which insulin-like growth factor 1 (IGF-1) promotes the survival of cerebellar neurons. IGF-1 activation of phosphoinositide 3-kinase (PI3-K) triggered the activation of two protein kinases, the serine-threonine kinase Akt and the p70 ribosomal protein S6 kinase ($p70^{S6K}$). Experiments with pharmacological inhibitors, as well as expression of wild-type and dominant-inhibitory forms of Akt, demonstrated that Akt but not $p70^{S6K}$ mediates PI3-K–dependent survival. These findings suggest that in the developing nervous system, Akt is a critical mediator of growth factor–induced neuronal survival.

The intracellular signaling pathways by which growth factors promote survival—in particular, survival of neurons of the central nervous system—are not well characterized. The survival of certain subsets of neurons of the peripheral nervous system can be promoted by the activation of a pathway that includes the guanosine triphosphate (GTP)–binding protein Ras and a series of protein kinases leading to mitogen-activated protein kinase (MAPK) (*1, 2*). In addition, a pathway that includes the lipid kinase PI3-K is important for the survival of several cell lines (*3, 4*), although the mechanisms by which PI3-K promotes survival are unclear. We investigated the contribution of two targets of PI3-K, the serine-threonine kinase Akt (*5–7*) and $p70^{S6K}$ (*8*), to the IGF-1–mediated survival of cerebellar neurons. We found that Akt has a critical role in promoting IGF-1–dependent survival.

For these studies, we used a well-characterized culture system of cerebellar neurons (*9, 10*). Large numbers of neurons of relatively homogeneous composition (consisting primarily of granule neurons) can be obtained, thus allowing biochemical analyses (*9*). Withdrawal of survival factors leads to the rapid and synchronous apoptosis of cerebellar neurons (*9*). About 50% of the cells were apoptotic within 1 day (Fig. 1), and almost all of the cells died within 3 to 4 days. The dying neurons showed characteristic features of apoptosis, including nuclear condensation and cleavage of chromatin into oligonucleosomal fragments (Fig. 1E) (*9, 11*). The apoptosis could be inhibited by defined trophic factors (*9*), including IGF-1 (Fig. 1C); insulin at superphysiological concentrations (Fig. 1D), which is believed to act through the IGF-1 receptor (*10*); and membrane-depolarizing concentrations of KCl, which lead to increased concentrations of intracellular calcium and may therefore simulate neuronal activity (*9*). The effects of IGF-1 and high concentrations of insulin on cerebellar neuron survival may reflect an in vivo function of IGF-1, because both IGF-1 and its receptor are expressed in the cerebellum, and transgenic mice overexpressing IGF-1 show increases in cell number in the brain (*12*).

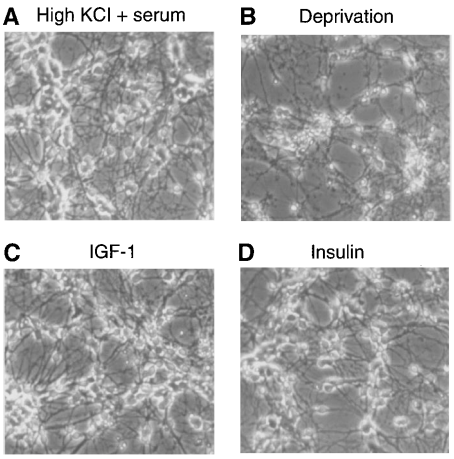

Fig. 1. Induction of apoptosis or survival of cerebellar neurons (*10*). (**A**) Phase-contrast photomicrograph of cerebellar neurons in complete medium [containing 25 mM KCl (high KCl) and 10% calf serum]. Cells were healthy, as indicated by large size and phase-brightness. (**B**) After 22 hours of deprivation (5 mM KCl, no serum), many neurons died. Death of cells was also confirmed by staining with trypan blue (*11*). (**C** and **D**) Death was inhibited by IGF-1 (25 ng/ml) (C) or insulin (10 μg/ml) (D). (**E**) Deprivation of cerebellar neurons induces chromatin cleavage. Starting at 4 to 8 hours after deprivation of trophic factors, extensive DNA laddering was present (*2*). U, untreated; positions of molecular size markers (in kilobases) are indicated on the left.

We first identified the signal transduction pathways that are activated in cerebellar neurons by IGF-1 or insulin. We examined activation of the Ras-MAPK pathway, because this pathway has been implicated in the control of nerve growth factor (NGF)–induced survival (*1, 2*). Insulin and IGF-1 failed to activate MAPK, although brain-derived neurotrophic factor (BDNF) efficiently activated MAPK (Fig. 2, A and B) (*13*). These results suggest that activation of the Ras-MAPK pathway is not critical for IGF-1–promoted or insulin-promoted survival of cerebellar neurons, and they also corroborate reports that insulin fails to activate MAPK in certain cell types, including chick forebrain neurons (*14*).

Activation of PI3-K is required for growth factor–induced survival of the PC12 neuronal cell line (*3*). By several criteria, we established that IGF-1 and insulin efficiently activate PI3-K in cerebellar neurons. First, IGF-1 and insulin induced binding of PI3-K to the receptor-associated protein IRS-1 (insulin receptor substrate 1) (Fig. 2C) (*13*), an interaction that mediates activation of PI3-K (*15*). Second, IGF-1 and insulin caused increased lipid

H. Dudek, S. R. Datta, M. E. Greenberg, Division of Neuroscience, Department of Neurology, Children's Hospital, and Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA.
T. F. Franke, Montreal Neurological Institute, McGill University, PQ H3A 2B4, Canada; and Division of Signal Transduction, Beth Israel Hospital, and Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA.
M. J. Birnbaum, Howard Hughes Medical Institute, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA.
R. Yao and G. M. Cooper, Division of Molecular Genetics, Dana-Farber Cancer Institute, and Department of Pathology, Harvard Medical School, Boston, MA 02115, USA.
R. A. Segal, Department of Neurology, Beth Israel Hospital, and Harvard Medical School, Boston, MA 02115, USA.
D. R. Kaplan, Montreal Neurological Institute, McGill University, PQ H3A 2B4, Canada.
*These authors made comparable contributions to this report.
†To whom correspondence should be addressed.