Molecular basis for modulation of biological function by alternate splicing of the Wilms' tumor suppressor protein

John H. Laity*, H. Jane Dyson*, and Peter E. Wright*†‡

*Department of Molecular Biology and †Skaggs Institute for Chemical Biology, Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037

Edited by Peter H. von Hippel, University of Oregon, Eugene, OR, and approved August 28, 2000 (received for review June 29, 2000)

Alternate splicing, leading to the insertion of the tripeptide KTS in the linker between the third and fourth C_2H_2 zinc fingers, changes both the DNA-binding function and the subnuclear localization of the Wilms' tumor suppressor protein (WT1). We have used NMR relaxation experiments to determine the molecular basis for the differing DNA recognition properties of the WT1–KTS and WT1+KTS isoforms. Our results show that the KTS insertion increases the flexibility of the linker between fingers 3 and 4 and abrogates binding of the fourth zinc finger to its cognate site in the DNA major groove. This represents a mechanism whereby a single zinc-finger gene can be used, through alternate splicing, to fulfill different functions in the cell.

The Wilms' tumor suppressor gene encodes a DNA-binding protein containing four Cys_2His_2 zinc fingers (1, 2) and is essential for normal mammalian urogenital development (3). Mutations or deletions of this gene are associated with Wilms' tumor, one of the most common pediatric solid tumors (4, 5), and frequently are associated with congenital defects of the genitourinary tract such as Denys-Drash syndrome. Alternate splicing, leading to the insertion of the tripeptide KTS in the linker between the third and fourth C_2H_2 zinc fingers (6), changes both the DNA-binding function and the subnuclear localization of the Wilms' tumor suppressor protein (WT1) (7–10).

Four isoforms of WT1 are formed by alternate RNA splicing (6). In one splice variant, an additional 17 aa are inserted N-terminal to the first zinc finger but these have no known effect on protein function. The other alternate splicing event results in the insertion (+KTS isoform) or omission (−KTS isoform) of the KTS tripeptide in the canonical TGEKP linker sequence between zinc fingers 3 and 4 (Fig. 1). All four isoforms are present in cells expressing WT1 and their relative abundance is constant; indeed, developmental abnormalities have been observed in patients with altered ratios of +KTS and −KTS isoforms (4, 11). Insertion of the KTS tripeptide has a profound effect on both the DNA-binding affinity and the specificity of WT1. The WT1–KTS isoform binds to a 9-bp early growth response protein (EGR-1) consensus site with high affinity whereas the +KTS splice variant binds to the same site 10- to 20-fold more weakly (7, 8). The different WT1 isoforms localize to distinct compartments in the nucleus; WT1–KTS colocalizes with other transcription factors, whereas the more abundant +KTS isoform associates with components of the pre-mRNA-splicing machinery (small nuclear ribonucleoproteins) where it potentially functions through RNA binding (9). Thus, the presence or absence of the KTS insert in the third linker modulates both the DNA-binding affinity and the functional distribution of WT1 within the nucleus. Indeed, it has been suggested that differences in DNA-binding affinity of the +KTS and −KTS splice variants might dictate the pattern of nuclear localization, with the tighter-binding −KTS isoform being preferentially compartmentalized with the DNA (10).

In this report, we examine the structural basis for the difference in DNA recognition and binding properties that dictate the respective *in vivo* regulatory roles of the two KTS isoforms of WT1. Understanding of the molecular mechanism by which alternate splicing can modulate the biological function of a zinc-finger transcription factor is clearly of major importance. To this end, we have used NMR relaxation measurements to examine the interactions of both the +KTS and −KTS isoforms of WT1 with its cognate DNA. Of the four WT1 zinc fingers, fingers 2–4 are closely similar in sequence to the three zinc fingers encoded by the early growth response gene, *EGR-1* and bind to the EGR-1 recognition element (GCG GGG GCG). Heteronuclear NMR spectra were recorded for WT1+KTS and WT1–KTS, both free and bound to a 14-bp oligonucleotide duplex containing an EGR-1-binding site. The DNA sequence (Fig. 1) was selected to give high binding affinity for the +KTS isoform, based on the observed sequence preferences for interaction with finger 1 (12, 13).

Materials and Methods

Preparation of Protein and DNA. Two uniformly ^{15}N-labeled protein isoforms that encompass the zinc-finger domain (residues 318–438, plus an N-terminal Ala) of WT1, which included (+KTS) or omitted (−KTS) residues 408–410, were expressed and purified as described (14). Complexes were formed with a 14-bp DNA duplex (coding strand sequence 5′-CGCGGGGGCGTCTG-3′) prepared from complementary synthetic oligonucleotides (Operon Technologies, Alameda, CA; Keystone Biosource, Camarillo, CA). A detailed protocol for this procedure has been described (14). All NMR samples were prepared in argon-saturated d_{11}-Tris buffer (10 mM, pH 6.7) in a mixture of 95% H_2O/5% ^2H_2O, containing 20 mM KCl, 50 μM $ZnSO_4$, 2 mM NaN_3, 0.2 mM DSS. The sample concentrations for the free proteins were 0.8 mM, and the WT1–DNA complexes were at equimolar concentrations of 0.3 mM.

NMR Measurements. The [^{15}N] longitudinal relaxation time (T_1), ^{15}N transverse relaxation time (T_2), and [^1H]-[^{15}N] nuclear Overhauser effect (NOE) were measured by using standard inversion-recovery, Carr-Purcell-Meiboom-Gill, and steady-state NOE experiments (15) with water flipback (16). Spectra were recorded at 310 K on a Bruker DRX spectrometer operating at a ^1H frequency of 600.13 MHz and equipped with an 8-mm probe (Nalorec Cryogenics, Martinez, CA). For T_1, spectra were acquired at 13 relaxation delay times for the DNA complexes (30 ms to 3 s), and 15 delay times for the free proteins

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: WT1, Wilms' tumor suppressor protein; EGR-1, early growth response protein; NOE, nuclear Overhauser effect; T_1, longitudinal relaxation time; T_2, transverse relaxation time; τ_m, effective rotational correlation time; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate, sodium salt.

‡To whom reprint requests should be addressed. E-mail: wright@scripps.edu.

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.

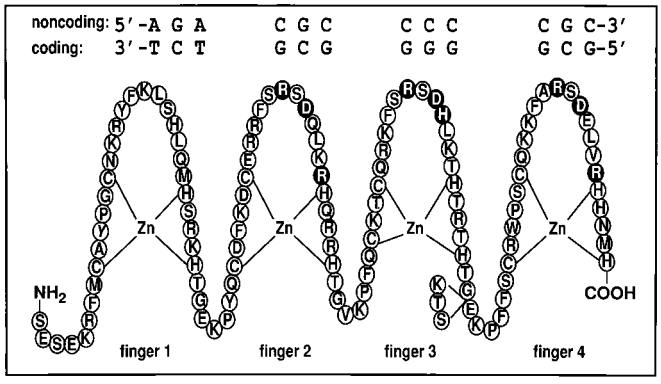

Fig. 1. Schematic representation of the WT1 zinc-finger domain (fingers 1–4; residues 318–438) with the core DNA recognition element showing the ''antiparallel'' orientation in which the complex is formed (the 5′ end of the coding strand of the DNA is aligned with the C terminus of the WT1 protein). The circles representing the putative DNA base-contacting residues (14, 23, 35) are shaded black. Base-contacting residues in finger 1 are unknown, and those in finger 4 occur only in the −KTS isoform (see text). The KTS insertion site between fingers 3 and 4 is indicated.

(30 ms to 2.5 s); for T_2, 17 delay times were acquired for both complexes (6–482 ms) and the free proteins (6–642 ms). Duplicate data sets were recorded at three time points for both T_1 and T_2 to allow estimation of errors. The refocusing delay in the Carr-Purcell-Meiboom-Gill experiment was 1 ms. The [^{1}H]-[^{15}N] NOE experiment was repeated five times for each complex and three times for the free proteins, and data sets with and without ^{1}H saturation were interleaved. Recovery delays of 3.5, 3.5, and 2.5 s were used for the T_1, T_2, and NOE experiments, respectively. All spectra were processed by using FELIX (Molecular Simulations, San Diego, CA), and all T_1 and T_2 time constants were calculated from the decay curves by using the program CURVEFIT (A.G. Palmer, III, Columbia University, New York). Published backbone resonance assignments were used (14), although those of the free proteins were extrapolated from 293 to 310 K by using a series of [^{15}N] heteronuclear single quantum coherence spectra recorded at intervening temperatures.

Analysis of Relaxation Results. Reduced spectral densities were calculated by using published methods, assuming a $1/\omega^2$ dependence at high frequencies (17). An effective rotational correlation time, τ_m, for each residue was calculated from the spectral densities (18). A mean correlation time for each zinc finger was calculated by averaging individual τ_m values for residues in elements of regular secondary structure (β-hairpin, α-helix).

Results and Discussion

The dynamics of the WT1 polypeptide chain were investigated by measurement of the [^{15}N] T_1, [^{15}N] T_2, and the [^{1}H]-[^{15}N] heteronuclear NOE by using two-dimensional inverse-detected NMR experiments. For each [^{15}N]–[^{1}H] pair, the data provide information on the fast internal dynamics as well as the overall tumbling of the macromolecule. The heteronuclear [^{1}H]-[^{15}N] NOE of the backbone amides, shown as a function of residue number in Fig. 2, is highly sensitive to motions of the polypeptide backbone on a ps to ns time scale. NOE values smaller than ≈0.6 indicate large amplitude backbone fluctuations on a ps to ns time scale. For the free protein, pronounced minima are observed in the heteronuclear NOE values at each linker (Fig. 2), showing that the linkers are all highly flexible. Insertion of the KTS tripeptide in the linker between fingers 3 and 4 significantly increases the backbone flexibility.

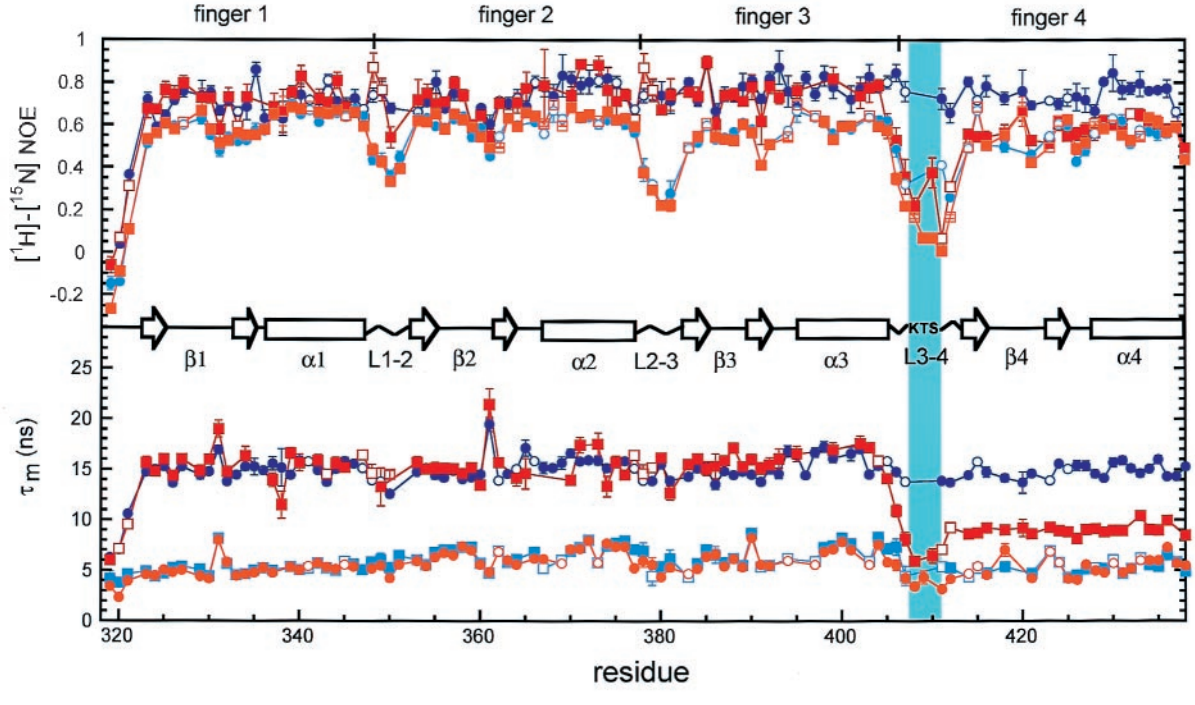

Fig. 2. Backbone [^{1}H]-[^{15}N] NOE and effective τ_m values for free (light colors) and DNA-bound (dark colors) +KTS (squares) and −KTS (circles) as a function of residue number. The β-sheet (open arrows), α-helix (open rectangles), linker regions (zigzag lines), and alternative KTS splice site (vertical cyan bar) in the zinc-finger domain of WT1 are indicated schematically. Error bars on the [^{1}H]-[^{15}N] NOE data points represent the SEM in the respective measurement, whereas the error bars on the τ_m plot account for uncertainties in both the measured data points and the respective exponential fits used to obtain T_1 and T_2 for each residue (67% confidence). Open symbols indicate resonances that are overlapped and for which there are consequently large uncertainties in the relaxation parameters.

On binding to DNA, the average NOE increases for all four zinc fingers of WT1–KTS because of loss of the domain flexibility that is a characteristic property of free zinc-finger proteins (19, 20). Furthermore, the short time scale motions in all three linkers are damped and the heteronuclear NOEs in these regions become comparable in magnitude to those in the finger domains (Fig. 2). Damping of linker motions on DNA binding also has been observed for other zinc-finger proteins and is caused by DNA-induced helix capping interactions and formation of a highly conserved structure by the linker residues (20, 21). This DNA-induced folding of the linker has been likened to a "snap lock" that, once activated, helps to restrict the motions between adjacent fingers and position them for optimal binding interactions in the major groove (21).

Very different behavior is observed for the WT1+KTS isoform. On DNA binding, increases are observed in the heteronuclear NOE for residues in fingers 1–3 and in the linkers between fingers 1 and 2 and between fingers 2 and 3; these changes are very similar to those seen on forming the WT1–KTS complex with DNA and indicate similar restriction of backbone motions on DNA binding. However, the linker containing the KTS insertion, between fingers 3 and 4, retains the high degree of backbone flexibility seen in the free WT1+KTS isoform and the NOEs in finger 4 closely parallel those of the free protein.

The most revealing insights into the effect of alternate splicing on the DNA-binding properties come from a consideration of the τ_m calculated from the relaxation data. These data are shown for both isoforms in Fig. 2. Not surprisingly, the τ_m is considerably smaller in both WT1 isoforms free in solution (6.0 ± 1.1 ns and 5.8 ± 1.0 ns for the WT1–KTS and WT1+KTS isoforms, respectively) compared with the same proteins bound to the DNA recognition element. The average τ_m for zinc fingers 1–4 in the DNA complex of WT1–KTS is 15 ± 1.0 ns, reflecting the higher molecular mass (≈25 kDa, vs. ≈15 kDa for the free protein) and slower molecular tumbling of the complex. For the WT1+KTS complex, the average τ_m for zinc fingers 1–3 (15 ± 1.4 ns) is the same as for the −KTS isoform, confirming that these three fingers are tightly bound to the DNA. Remarkably, finger 4 tumbles in solution with an effective τ_m that is very much shorter (τ_m = 9.2 ± 0.6 ns) than the τ_m for fingers 1–3. Indeed, this τ_m value is much closer to that of the free protein than that of the protein–DNA complex, providing unequivocal evidence that binding of WT1+KTS to DNA does not significantly restrict the molecular tumbling of finger 4. This can only occur if finger 4 makes very weak interactions or no interactions with the DNA. Thus, the [^{15}N] relaxation data show clearly that the effect of the KTS insert in the linker between fingers 3 and 4 is to prevent finger 4 from making optimal interactions with its cognate binding site in the DNA major groove. The short τ_m value associated with finger 4 indicates that, at best, it makes only transient interactions with DNA and tumbles largely independently of the finger 1–3–DNA complex, as shown schematically in Fig. 3.

Thus, the molecular basis for the decreased DNA-binding affinity of the WT1+KTS isoform lies in the increased length and flexibility of the linker between fingers 3 and 4, which prevents finger 4 from binding to its cognate site on the DNA, even though its structure and DNA contact residues are unchanged by the splicing event (14). With the exception of potential phosphate backbone contacts involving the Lys side chain (22), the residues in the highly conserved TGEKP linker are distant from the DNA in the available structures of zinc-finger protein complexes (22–26). Loss of finger 4 binding is unlikely to arise solely from disruption of the Lys-phosphate backbone contacts, because mutagenesis studies of transcription factor IIIA suggest that these contribute relatively little to the overall DNA-binding affinity (27). A more likely explanation is that insertion of the KTS tripeptide prevents the linker from adopting the unique and highly conserved structure observed for all TGEKP linkers in the available x-ray and NMR structures of zinc finger–DNA complexes (21). In addition, the increased entropic cost of ordering the longer linker would militate against localization and binding of finger 4 in the DNA major groove.

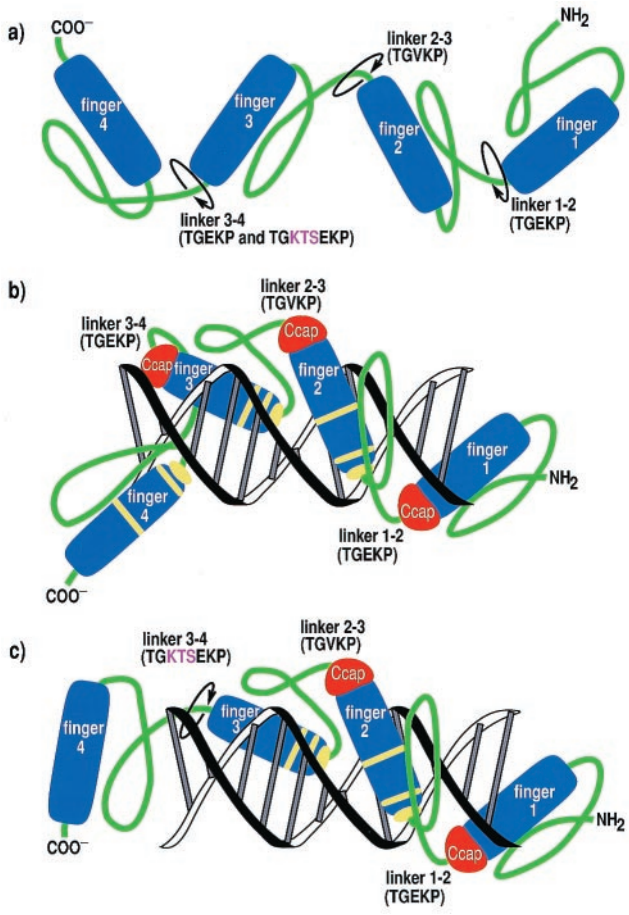

Fig. 3. Cartoon drawing illustrating the structural and dynamic consequences of the KTS insertion on DNA binding by WT1. (*a*) Free WT1 isoforms and DNA-bound (*b*) −KTS and (*c*) +KTS proteins have different degrees of interdomain tumbling motions. Circular arrows in *a* and *c* illustrate the sites of linker flexibility between all four fingers in the free proteins (*a*) and in linker 3–4 of the +KTS complex with DNA (*c*). Yellow bands in *b* and *c* indicate putative base-contacting residues. Note that finger 4 does not contact the DNA in the +KTS complex (*c*).

It might be expected that the longer linker in the WT1+KTS isoform could alter the DNA-binding specificity to favor a target with increased spacing between the finger 3- and 4-binding sites. This appears not to be the case, however; screening of complete genomic DNA libraries failed to identify efficient binding targets for WT1+KTS (8). For all known recognition elements, the −KTS isoform binds DNA with higher affinity than the +KTS isoform (7, 8, 12). Thus, abrogation of finger 4–DNA interactions by the KTS insert is likely to be universal to all DNA sequences that WT1 binds. We note that construction of a molecular model suggests that the inserted linker residues could easily be accommodated structurally without displacement of either of the neighboring fingers from their DNA-binding sites. This provides added support for our contention that changes in linker conformation and increased linker flexibility play a major role in mediating the DNA-binding properties of WT1+KTS.

Our data thus provide insights into the molecular mechanism by which the DNA-binding properties, subnuclear localization, and function of the Wilms' tumor gene product are mediated by alternate splicing. Whereas WT1–KTS binds EGR-like recognition elements with high affinity and is clearly involved in transcriptional regulation, the +KTS isoform binds DNA weakly and is preferentially associated with the splicing machinery where, it is postulated, it may interact with RNA (9). Although RNA sequences have been identified that bind with high affinity to WT1+KTS (28), a physiologically relevant target sequence has yet to be identified. It is also possible that WT1+KTS may function by binding protein components of the splicing machinery. Indeed, it has recently been shown that WT1 interacts with the splicing factor U2AF65 *in vitro*, and that the +KTS isoform binds more strongly than WT1–KTS (29). Further characterization of the detailed molecular function and interactions of WT1+KTS is of particular interest because variations in the ratio of ± KTS splice isoforms are associated with human disease. Frasier Syndrome is caused by a mutation in an intronic region of the WT1 gene that prevents production of the +KTS isoform (30, 31). Although many questions still remain about the physiological interactions mediated by the different isoforms, the studies presented here reveal the fundamental molecular mechanism by which alternate splicing modulates DNA binding and hence determines the function and cellular localization of WT1.

Finally, we note that multiplicity of function is not without precedent in zinc-finger proteins. Transcription factor IIIA, with nine zinc fingers, binds both the internal control region of the 5S RNA gene and the 5S RNA transcript itself (32, 33). In this case, different sets of zinc fingers (fingers 1–3 and fingers 4–7, respectively) mediate the high-affinity interactions with DNA and RNA (34). Alternate splicing to alter the length of a linker represents a further simple but very effective mechanism by which the DNA-binding properties and hence the biological function of a zinc-finger protein can be modulated.

We thank J. Chung for assistance with NMR measurements and B. Duggan for computer programs used in the reduced spectral mapping and many helpful discussions. This work was supported by Grant GM36643 from the National Institutes of Health. J.H.L. is supported by Fellowship CA79192 from the National Institutes of Health.

1. Gessler, M., Poustka, A., Cavenee, W., Neve, R. L., Orkin, S. H. & Bruns, G. A. P. (1990) *Nature (London)* **343,** 774–778.
2. Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., *et al.* (1990) *Cell* **60,** 509–520.
3. Kreidberg, J. A., Sariola, H., Loring, J. M., Maeda, M., Pelletier, J., Housman, D. & Jaenisch, R. (1993) *Cell* **74,** 679–691.
4. Coppes, M. J., Campbell, C. E. & Williams, B. R. G. (1993) *FASEB J.* **7,** 886–894.
5. Hastie, N. D. (1994) *Annu. Rev. Genet.* **28,** 523–558.
6. Haber, D. A., Sohn, R. L., Buckler, A. J., Pelletier, J., Call, K. M. & Housman, D. E. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 9618–9622.
7. Rauscher, F. J., Morris, J. F., Tournay, O. E., Cook, D. M. & Curran, T. (1990) *Science* **250,** 1259–1262.
8. Bickmore, W. A., Oghene, K., Little, M. H., Seawright, A., van Heyningen, V. & Hastie, N. D. (1992) *Science* **257,** 235–237.
9. Larsson, S. H., Charlieu, J.-P., Miyagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., van Heyningen, V. & Hastie, N. D. (1995) *Cell* **81,** 391–401.
10. Charlieu, J.-P., Larsson, S., Miyagawa, K., van Heyningen, V. & Hastie, N. D. (1995) *J. Cell Sci.* **19,** 95–99.
11. Reddy, J. C. & Licht, J. D. (1996) *Biochim. Biophys. Acta* **1287,** 1–28.
12. Drummond, I. A., Rupprecht, H. D., Rohwer-Nutter, P., Lopez-Guisa, J. M., Madden, S. L., Rauscher, F. J., III & Sukhatme, V. P. (1994) *Mol. Cell Biol.* **14,** 3800–3809.
13. Hamilton, T. B., Borel, F. & Romaniuk, P. J. (1998) *Biochemistry* **37,** 2051–2058.
14. Laity, J. H., Dyson, H. J. & Wright, P. E. (2000) *Biochemistry* **39,** 5341–5348.
15. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. (1995) *Biochemistry* **34,** 868–878.
16. Grzesiek, S. & Bax, A. (1993) *J. Am. Chem. Soc.* **115,** 12593–12594.
17. Farrow, N. A., Zhang, O., Szabo, A., Torchia, D. A. & Kay, L. E. (1995) *J. Biomol. NMR* **6,** 153–162.
18. Bracken, C., Carr, P. A., Cavanagh, J. & Palmer, A. G., III (1999) *J. Mol. Biol.* **285,** 2133–2146.
19. Brüschweiler, R., Liao, X. & Wright, P. E. (1995) *Science* **268,** 886–889.
20. Foster, M. P., Wuttke, D. S., Radhakrishnan, I., Case, D. A., Gottesfeld, J. M. & Wright, P. E. (1997) *Nat. Struct. Biol.* **4,** 605–608.
21. Laity, J. H., Dyson, H. J. & Wright, P. E. (2000) *J. Mol. Biol.* **295,** 719–727.
22. Wuttke, D. S., Foster, M. P., Case, D. A., Gottesfeld, J. M. & Wright, P. E. (1997) *J. Mol. Biol.* **273,** 183–206.
23. Pavletich, N. P. & Pabo, C. O. (1991) *Science* **252,** 809–817.
24. Pavletich, N. P. & Pabo, C. O. (1993) *Science* **261,** 1701–1707.
25. Houbaviy, H. B., Usheva, A., Shenk, T. & Burley, S. K. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 13577–13582.
26. Kim, C. A. & Berg, J. M. (1996) *Nat. Struct. Biol.* **3,** 940–945.
27. Choo, Y. & Klug, A. (1993) *Nucleic Acids Res.* **21,** 3341–3346.
28. Caricasole, A., Duarte, A., Larsson, S. H., Hastie, N. D., Little, M., Holmes, G., Todorov, I. & Ward, A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 7562–7566.
29. Davies, R. C., Calvio, C., Bratt, E., Larsson, S. H., Lamond, A. I. & Hastie, N. D. (1998) *Genes Dev.* **12,** 3217–3225.
30. Barbaux, S., Niaudet, P., Gubler, M. C., Grunfeld, J. P., Jaubert, F., Kuttenn, F., Fekete, C. N., Souleyreau-Therville, N., Thibaud, E., Fellous, M., *et al.* (1997) *Nat. Genet.* **17,** 467–470.
31. Klamt, B., Koziell, A., Poulat, F., Wieacker, P., Scambler, P., Berta, P. & Gessler, M. (1998) *Hum. Mol. Genet.* **7,** 709–714.
32. Honda, B. M. & Roeder, R. G. (1980) *Cell* **22,** 119–126.
33. Pelham, H. R. & Brown, D. D. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 4170–4174.
34. Clemens, K. R., Wolf, V., McBryant, S. J., Zhang, P., Liao, X., Wright, P. E. & Gottesfeld, J. M. (1993) *Science* **260,** 530–533.
35. Elrod-Erickson, M., Rould, M. A., Nekludova, L. & Pabo, C. O. (1996) *Structure (London)* **4,** 1171–1180.