A Cys_3His Zinc-Binding Domain from Nup475/Tristetraprolin: A Novel Fold with a Disklike Structure†

Barbara T. Amann,‡ Mark T. Worthington,§ and Jeremy M. Berg*,‡

Department of Biophysics and Biophysical Chemistry, 713 WBSB, Johns Hopkins School of Medicine, Baltimore, Maryland 21205, and Digestive Health Center of Excellence, University of Virginia Health Sciences Center, Charlottesville, Virginia 22908-0708

Received October 11, 2002; Revised Manuscript Received November 4, 2002

ABSTRACT: Nup475 (also known as tristetraprolin and TIS11) includes two zinc-binding domains of the form Cys-X_8-Cys-X_5-Cys-X_3-His. These domains are required for rapid degradation of tumor necrosis factor (TNF) and other mRNAs through the interaction with AU-rich elements in their 3′-untranslated regions. The three-dimensional solution structure of the first domain was determined by multidimensional nuclear magnetic resonance spectroscopy, revealing a novel fold around a central zinc ion. The core structure is disklike with a diameter of ∼25 Å and a width of ∼12 Å. This structure provides a basis for evaluating the role of individual residues for structural stability and for nucleic acid binding.

A novel pattern of three Cys residues and one His residue was first identified in the murine protein Nup475/TTP/TIS11 (*1*−*3*). The product of an immediate early response gene induced by a variety of mitogenic growth factors and cytokines, Nup475 contains two adjacent sequences of the form $CysX_8CysX_5CysX_3His$. These Cys_3His domains coordinate zinc through the conserved Cys and His residues (*4*). The sequences in Nup475 are prototypical of over 120 homologous sequences in more than 75 proteins from a variety of eukaryotes, most of which contain two motifs; proteins containing one to five motifs have been identified.

Biological studies have revealed that Nup475 plays an important role in RNA processing. Mice homozygous for a targeted disruption of the Nup475 gene suffer from cachexia, arthritis, conjunctivitis, dermatitis, and myeloid hyperplasia correlated with overexpression of tumor necrosis factor (TNF)[1] (*5*, *6*). This phenotype is prevented by parenteral anti-TNF monoclonal antibodies (*5*). The absence of Nup475 leads to production of excess TNF since Nup475 normally contributes to the cleavage of TNF mRNA, initiated by the binding of Nup475 through its Cys_3His domains to AU-rich elements (AREs) in the 3′-untranslated region of the TNF mRNA (*6*−*10*). Nup475 has also been shown to destabilize granulocyte-macrophage colony stimulating factor (GM-CSF) mRNA through interactions with its AREs (*11*).

The Cys_3His domains of Nup475 bind one zinc per domain with dissociation constants of $<10^{-11}$ M, a value comparable to other well-authenticated zinc-binding domains (*4*). The absorption spectra of cobalt(II)-substituted domains confirm coordination by three Cys and one His (*4*). In initial studies of the Cys_3His domains of Nup475, we characterized a protein fragment that was more soluble than the wild-type two-domain peptide, Nup475(Y143K) (*4*). The mutated fragment has a nonconserved Tyr residue in the second Cys_3-His repeat replaced with Lys. This fragment was selected for study on the basis of its increased solubility and improved one-dimensional NMR spectral properties relative to a wild-type fragment of the same length. Although this peptide displayed an ordered structure for the first domain, it had lost its ability to bind zinc in the second domain (*4*). Thus, we were able to determine the structure for the first metal-binding domain only. We present the structure of this domain from the Nup475(Y143K) bound to zinc. Resonances from residues 5−42 were assigned. The structure of this region reveals a novel fold organized around three cores: the zinc coordination unit and two hydrophobic cores. This structure reveals the basis for several unusually shifted proton resonances that are characteristic of this class of domain.

MATERIALS AND METHODS

NMR Sample Preparation. Unlabeled, uniformly ^{15}N-labeled, and $^{13}C/^{15}N$-labeled protein samples were prepared for the Nup475(Y143K) peptide with residues 91−165 of Nup475 containing a Y143K mutation and a four amino acid thrombin cleavage tag as previously described (*4*). All samples contained 2.2 equiv of zinc at pH 5.8 in deuterated Tris buffer.

NMR Spectroscopy and Structural Calculations. NMR data were collected on a Varian UNITYplus 500 MHz spectrometer at 20 °C, processed using Felix software (Biosym Technologies), and sorted using in-house software. The connectivity between amino acids 5−42 was established using an HNCACB experiment (*12*, *13*) as well as other standard methods. NOE distance constraints were generated, and ^1H NOESY in D_2O and 10% D_2O/90% H_2O, $^1H-^{15}N$ NOESY-HSQC, and $^1H-^{13}C$ NOESY-HSQC spectra were

† This work was supported by grants from the National Institutes of Health (J.M.B., M.T.W.) and an ADHF Industry Research Scholar Award (M.T.W.).

* Address correspondence to this author: phone, (410) 955-7322; fax, (410) 502-6910; e-mail, jberg@jhmi.edu.

‡ Johns Hopkins School of Medicine.

§ University of Virginia Health Sciences Center.

[1] Abbreviations: TTP, tristetraprolin; TNF, tumor necrosis factor; ARE, AU-rich element.

10.1021/bi026988m CCC: $25.00
Published on Web 12/10/2002

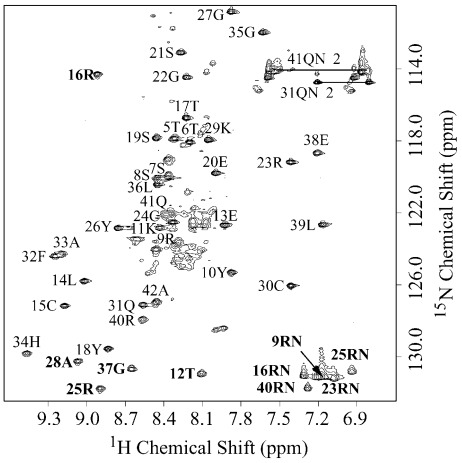

FIGURE 1: $^{1}H-^{15}N$ HSQC spectra of the Nup475(Y143K) peptide bound to zinc with residues 9−42 indicated. Resonances folded into the spectrum from outside the spectral window are shown in bold. The data were collected with a spectral window in the ^{15}N direction of 21 ppm centered around 121 ppm.

optimized with a 200 ms mixing time (*14*, *15*). Interproton distance restraints were calculated from peak volumes scaled using appropriate reference distances (AMX $\beta-\beta$, Gly $\alpha-\alpha$, ring protons) and classified into three distance levels (<2.7, 2.7−3.3, and 3.3−4.3 Å). Peptide ϕ torsional angle restraints were extracted from a $^{1}H-J$ HSQC spectrum (*16*). CNS was used for structure calculations starting with a peptide in an extended conformation. Proton and ^{13}C chemical shifts were used in the final structure calculations.

RESULTS

Structure Determination. The solution structure of Nup475-(Y143K) was solved with use of a combination of 2D and 3D experiments on nonlabeled, ^{15}N-labeled, and $^{15}N/^{13}C$-labeled protein samples. Although the fragment was 77 amino acids in length, intraresidue NOEs were observed only for residues 9−42, and resonance assignments are reported for residues 5−42. This region has the amino acid sequence -T_5TSSRYKTEL**C$_{15}$**RTYSESGR**C$_{24}$**RYGAK**C$_{30}$**QFA**H$_{34}$**-GLGELRQA$_{42}$- and corresponds to residues 91−128 in the murine protein. The assigned ^{15}N HSQC of the peptide demonstrates that nearly all of the shifted resonances correspond to the well-structured domain, while the unassigned resonances are sharp, poorly dispersed, and occur in the region of the spectrum normally associated with unstructured elements (Figure 1).

Initial structures were generated from unambiguously assigned NOE and torsional angle constraints. Further NOEs were assigned as the structures were refined, with 442 NOESY peaks unambiguously assigned. Initially, only the lengths of the three Cys S−Zn bonds were constrained. After several iterations, the Nϵ of His 34 was established as a zinc ligand, and the Nϵ−Zn bond length was constrained as well. No angle restraints for the zinc center were used until all assigned NOESY peaks were included in the structure calculations. Eleven angle constraints were then added to ensure an approximately tetrahedral geometry around the metal center. Such a tetrahedral geometry is indicated by the absorption spectrum of the cobalt(II) complex, which reveals extinction coefficients greater than 800 M^{-1} cm^{-1} for the d−d transitions centered around 680 nm (*4*). The

Table 1: Structural Statistics for Nup475(Y143K) Restraints

NOE-derived restraints (total)	442
intraresidue	251
sequential ($i-j=1$)	111
medium ($1<i-j\leq 4$)	25
long ($i-j>4$)	55
torsional angle restraints	15
metal restraints (total)	15
S−Zn, N−Zn bond length[a]	4
ligand−Zn angles[b]	5
tetrahedral distance restraints[c]	6
carbon chemical shift restraints	64
hydrogen chemical shift restraints	174
structural statistics (23 low-energy structures out of 30 generated, residues 10−40)	
RMSD of backbone atoms (Å)	0.47
RMSD of all heavy atoms (Å)	1.37
RMSD of distance restraints (Å)	0.043
RMSD from idealized geometry	
bonds (Å)	0.006
angles (deg)	0.67
impropers (deg)	0.54
CNS energies (kcal mol^{-1})	
total	532
bond	19
angle	77
improper	15
VDW	71
NOE	64
^{13}C shifts	27
^{1}H shifts	256
Ramachandran analysis (%)[d]	
favored regions	29
additionally allowed regions	55
generously allowed region	14
disallowed regions	2

[a] The S−Zn and N−Zn bond lengths were set to 2.30 and 2.00 Å, respectively, with force constants of 500. [b] Ligand−Zn angles His(Cδ2−Nϵ2)−Zn, His(Cδ2−Nϵ2)−Zn, and Cys(Cβ−S)−Zn were set to 126°, 126°, and 110°, respectively, with force constants of 500. [c] Distance restraints between the nitrogen and sulfurs of the metal-binding ligands were used at values of S−S 3.76 ± 0.2 Å and S−N 3.52 ± 0.2 Å to maintain the tetrahedral Zn site. [d] Analyzed by PROCHECK.

imposition of these constraints did not induce any additional NOE violations. Finally, chemical shifts were included as restraints, and stereospecific assignments were made, including pseudo-stereospecific assignments for the Tyr and Phe ring protons. The final parameters are summarized in Table 1. The structure determination is heavily reliant on constraints derived from the NH protons. The overlap introduced by the peaks due to the unstructured domain made the $^{13}C/^{15}N$ HSQC-NOESY and TOCSY spectra difficult to assign unambiguously. These limitations as well as the lack of regular secondary structure may account for the relatively large number of residues that do not lie in the fully favored regions of the Ramachandran diagram (Table 1). The coordinates for *m*-Nup475(Y143K) have been deposited to the Protein Data Bank (accession code 1M90) and to the Biomagnetic Resonance Bank (entry BMRB 5525).

Solution Structure. Thirty structures including residues 9−42 were initially generated. Of these, 23 structures were selected for further refinement, leading to a set of structures with a pairwise RMSD of 0.47 Å for all backbone atoms and 1.37 Å for all non-H atoms (Figure 2 and Table 1). An overlay shows good agreement between the structures, especially in the core of the molecule. The overall structure is disklike with a diameter of ~25 Å and a thickness of ~12 Å. The fold of this domain is distinct from that observed for

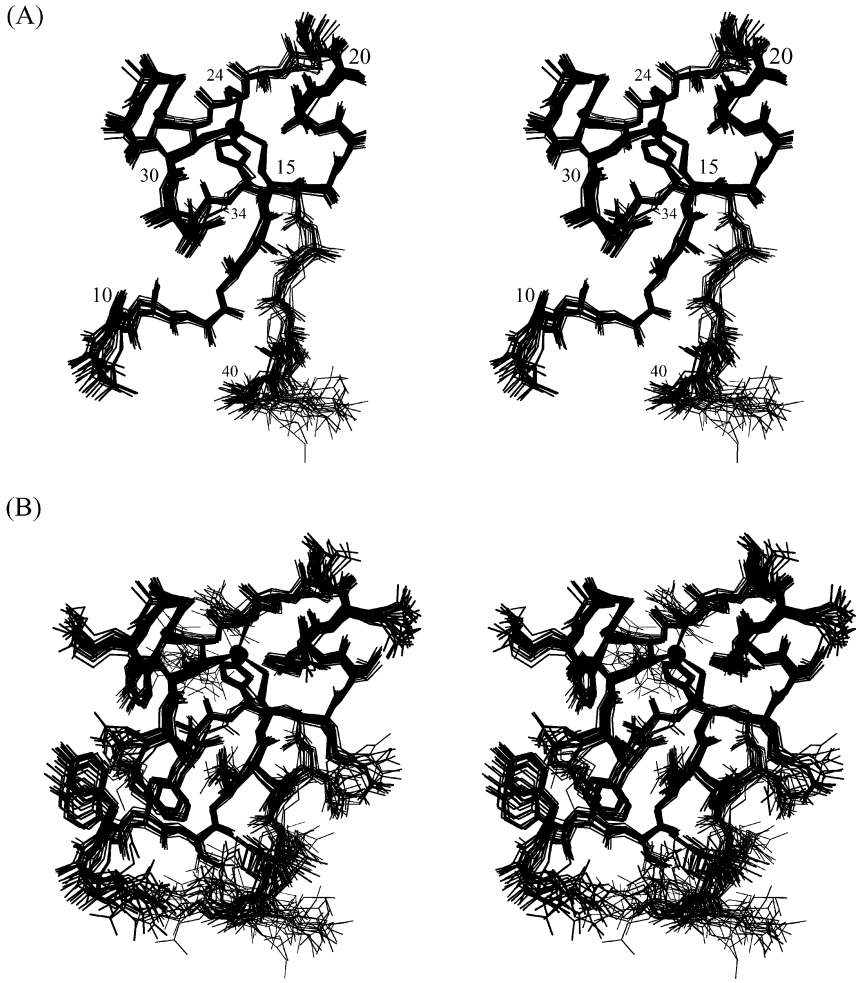

FIGURE 2: The first zinc-binding domain of Nup475. (A) Stereoview of a superposition of the backbone atoms of 23 refined structures with the metal ligands and residue numbers shown. (B) Stereoview with all non-hydrogen atoms shown.

all previously characterized metal-binding domains (*17*, *18*). A DALI search revealed that this fold has not been reported for any other proteins (*19*). The only regular secondary structure is a small segment of helix that includes residues 17−20. The zinc lies near one end of the structure and is somewhat solvent-exposed from one face but inaccessible from the opposite face. Three Cys residues are coordinated to the zinc. As is seen in many metal−cysteinate complexes, the cysteinate sulfur atoms accept hydrogen bonds from peptide NH groups (Cys 15, NH-Ser 19; Cys 24, NH-Tyr 26, Gly 27; Cys 30, NH-Ala 28). His 34 is coordinated to the zinc through its ϵ nitrogen. The δ NH group of His 34 appears to be hydrogen bonded to the carbonyl group of Ala 33. These primary and secondary interactions around the metal center clearly play a major role in organizing the structure.

The structure is also stabilized by two hydrophobic cores (Figure 3). The first consists primarily of Leu 14, Leu 36, and Leu 39. These residues, particularly Leu 14 and Leu 39, are conserved in most members of this family. The role of the C-terminal extension including Leu 39 in stabilizing the structure may explain the low solubility and structural stability of single- and double-domain fragments of Nup475 that were truncated prior to this position (*4*).

The second hydrophobic core is nearly adjacent to the first and includes three aromatic residues, Tyr 10, Tyr 26, and Phe 32. These residues surround the tetramethylene chain

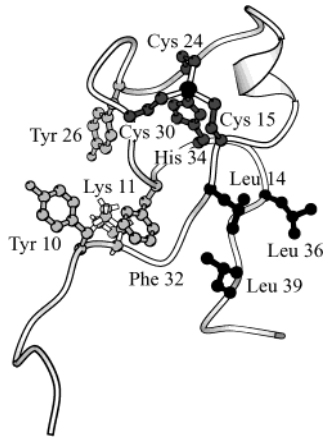

FIGURE 3: Ribbon diagram showing the metal-binding site (dark gray), the leucine-rich core (black), and the aromatic core including Lys 11 (light gray).

of Lys 11. The relative positions of the aromatic residues with respect to this side chain are responsible for the unusual chemical shifts found for the protons of this lysine residue. Both the γ (0.34, 0.03 ppm) and δ (1.17, 0.98 ppm) protons resonate substantially upfield of their random coil positions due to the combined ring current effects of the surrounding aromatic residues. Similar chemical shifts were observed for a conserved lysine in the first zinc-binding domain of the *Caenorhabditis elegans* protein PIE-1 (*20*) (unpublished

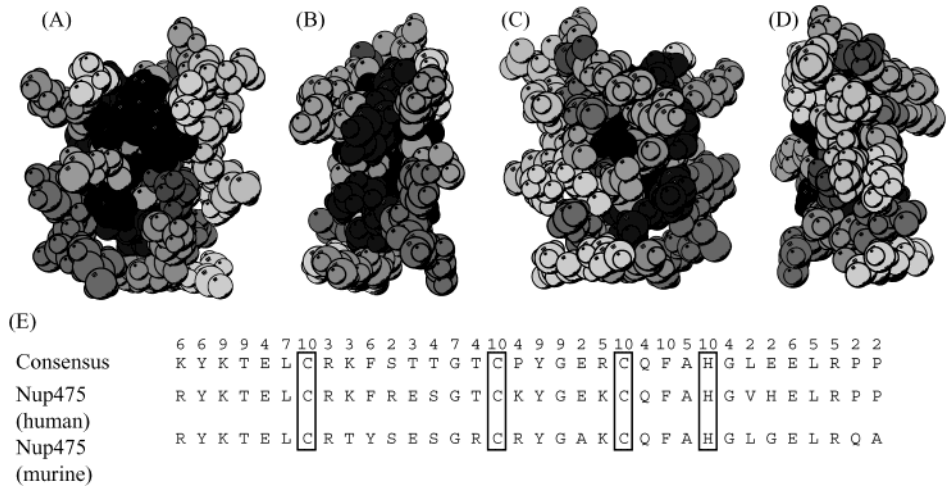

FIGURE 4: Space-filling models showing the level of sequence conservation within the domain family mapped onto the structure. (A−D) Four views related by 90° rotations around the vertical axis. More conserved residues are shown in darker gray. The Cys and His residues and the zinc are shown in black and lie toward the top of the structure. The view in (A) corresponds to the view in Figures 2 and 3. (E) Consensus sequence for this domain family with the level of sequence conservation (1, least conserved, through 10, most conserved) shown above the sequence. The consensus sequence was derived from 72 domains from proteins containing exactly two domains. When highly similar sequences were available from multiple species, the human sequence was used for the consensus. The human and murine sequences for the same region are shown for comparison.

data), providing direct evidence that the observed structure is conserved in at least some other members of this domain family.

DISCUSSION

The structure that we have determined serves as a framework for examining sequence conservation within this family of domains. Examination of the level of conservation among the Nup475/TTP/TIS11 domain family reveals that the regions including the structural cores are more conserved than are the solvent-exposed residues (Figure 4). This observation is consistent with the hypothesis that the structures of these domains are conserved among different family members and that surface residues involved in interactions with other molecules are more variable. Recently, studies of a set of site-directed mutants of human Nup475 have been reported (*9*). The effects of mutation of a number of residues on nucleic acid binding and the ability to promote TNF mRNA degradation were examined. Mutations examined within the first domain (for which we have determined the three-dimensional structure) include (using our numbering system) R16E, R16L, **(F,Y)18Q**, G22E, R23E, R23L, Y26F, **G27T**, G27R, K29D, **C30H**, **C30R**, F32Y, **F32N**, and **H34K**. Some of these mutations (shown in bold) have dramatic effects on RNA binding and TNF mRNA degradation. These mutations involve either changing of metal-binding residues (C30H, C30R, H34K), changing conserved aromatic residues to nonaromatic residues [(F,Y)18Q, F32N], or changing a conserved glycine to a β-branched residue (G27T). The effects of these mutations on nucleic acid binding may well be due to disruption of the overall structure of the domains rather than modification of a specific nucleic acid-binding surface. A similar set of mutations was made in the second metal-binding domain (for which we have no direct structural information but which is more than 50% identical in sequence to the first domain). Again, many of these mutations involved nonconservative substitutions of conserved residues. The sole exception appears to be mutation of an Arg residue that occurs immediately prior to the third Cys residue of the domain to Leu. The residue in this position in the first domain is Lys, and this residue does not appear to play an obvious role required for the stability of the domain. Thus, the large decrease in RNA-binding activity and in the stimulation of TNF mRNA degradation for this substitution may be due to a specific modification of the RNA-binding surface. More extensive mutagenesis in concert with the structure and direct stability studies should be useful in mapping the nucleic acid-binding surface in more detail. Much of the surface of the domain includes basic residues that could interact electrostatically with the nucleic acid backbone. Examination of the structure reveals one potential binding pocket defined by residues Tyr 10, Lys 11, Tyr 26, Lys 29, and Phe 32. The elucidation of the role of this pocket must await further studies.

Note that this structure is distinct from those of all other zinc-binding domains characterized to date. The nearly complete lack of regular secondary structural elements stands in contrast to the classical or TFIIIA-type zinc-binding domains whose structures comprise a β-hairpin followed by a helix (*21*−*23*). These domains also differ from the Nup475 domain with regard to the zinc coordination sphere utilizing two cysteine residues and two histidine residues. The structure is also distinct from that observed for the CCHC-box domains present in retroviral nucleocapsid proteins (*24*). The CCHC-box domains also show three cysteine−one histidine metal coordination, but the order of the histidine within the sequence is different and the loops are all smaller than those in the Nup475 domain. The CCHC-box domains also bind single-stranded nucleic acids. Once a structure of one or more Nup475 domains bound to a nucleic acid target has been determined, it will be possible to compare the modes of nucleic acid binding between these two classes of zinc-binding domain proteins to reveal any similarities apparent from the structures of the free domains.

REFERENCES

1. DuBois, R. N., McLane, M. W., Ryder, K., Lau, L. F., and Nathans, D. (1990) *J. Biol. Chem. 265*, 19185−19191.

2. Lai, W. S., Stumpo, D. J., and Blackshear, P. J. (1990) *J. Biol. Chem. 265*, 16556−16563.
3. Varnum, B. C., Ma, Q. F., Chi, T. H., Fletcher, B., and Herschman, H. R. (1991) *Mol. Cell. Biol. 11*, 1754−1758.
4. Worthington, M. T., Amann, B. T., Nathans, D., and Berg, J. M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 13754−13759.
5. Taylor, G. A., Carballo, E., Lee, D. M., Lai, W. S., Thompson, M. J., Patel, D. D., Schenkman, D. I., Gilkeson, G. S., Broxmeyer, H. E., Haynes, B. F., and Blackshear, P. J. (1996) *Immunity 4*, 445−454.
6. Carballo, E., Lai, W. S., and Blackshear, P. J. (1998) *Science 281*, 1001−1005.
7. Lai, W. S., Carballo, E., Thorn, J. M., Kennington, E. A., and Blackshear, P. J. (2000) *J. Biol. Chem. 275*, 17827−17837.
8. Lai, W. S., and Blackshear, P. J. (2001) *J. Biol. Chem. 276*, 23144−23154.
9. Lai, W. S., Kennington, E. A., and Blackshear, P. J. (2002) *J. Biol. Chem. 277*, 9606−9613.
10. Worthington, M. T., Pelo, J. W., Sachedina, M. A., Applegate, J. L., Arseneau, K. O., and Pizarro, T. T. (2002) *J. Biol. Chem. 277*, 48558−48564.
11. Carballo, E., Lai, W. S., and Blackshear, P. J. (2000) *Blood 95*, 1891−1899.
12. Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson. B101*, 201−205.
13. Mahandirum, D. R., and Kay, L. E. (1994) *J. Magn. Reson. B103*, 203−216.
14. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) *Biochemistry 28*, 6150−6164.
15. Mori, S., Abeygunawardana, C., van Zijl, P. C., and Berg, J. M. (1996) *J. Magn. Reson. B110*, 96−101.
16. Kay, L. E., and Bax, A. (1990) *J. Magn. Reson. 86*, 110−126.
17. Klug, A., and Schwabe, J. W. (1995) *FASEB J. 9*, 597−604.
18. Berg, J. M., and Shi, Y. (1996) *Science 271*, 1081−1085.
19. Holm, L., and Sander, C. (1997) *Nucleic Acids Res. 25*, 231−234.
20. Mello, C. C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J. R. (1996) *Nature 382*, 710−712.
21. Berg, J. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 99−102.
22. Párraga, G., Horvath, S. J., Eisen, A., Taylor, W. E., Hood, L., Young, E. T., and Klevit, R. E. (1988) *Science 241*, 1489−1492.
23. Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., and Wright, P. E. (1989) *Science 245*, 635−637.
24. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) *Science 279*, 384−388.

BI026988M