FOR THE RECORD

The interaction of eIF4E with 4E-BP1 is an induced fit to a completely disordered protein

C. MARK FLETCHER AND GERHARD WAGNER

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School,
Boston, Massachusetts 02115

(RECEIVED February 10, 1998; ACCEPTED March 18, 1998)

Abstract: 4E binding protein 1 (4E-BP1) inhibits translation by binding to the initiation factor eIF4E and is mostly or completely unstructured in both free and bound states. We wished to determine whether the free protein has local structure that could be involved in eIF4E binding. Assignments were obtained using double and triple resonance NMR methods. Residues 4–10, 43–46, and 56–65 could not be assigned, primarily because of a high degree of ^1H and ^{15}N chemical shift overlap. Steady-state $\{^1\text{H}\}$ - ^{15}N NOEs were measured for 45 residues in the assigned regions. Except for the two C-terminal residues, the NOEs were between -0.77 and -1.14 , indicating a high level of flexibility. Furthermore, the $\{^1\text{H}\}$ - ^{15}N NOE spectrum recorded with presaturation contained no strong positive signals, making it likely that no other residues have positive or smaller negative NOEs. This implies that 4E-BP1 has no regions of local order in the absence of eIF4E. The interaction therefore appears to be an induced fit to a completely disordered protein molecule.

Keywords: induced fit; NMR assignments; relaxation measurements; translational control; unstructured protein

4E binding protein 1 (4E-BP1) is a human phosphoprotein of 118 residues with an important role in the regulation of translation. In the less phosphorylated state, it binds to the translation initiation factor eIF4E, preventing the binding of eIF4G that is required for translation to occur (Sonenberg, 1996). When the cell is stimulated by growth factors, the mTOR signaling pathway mediates hyperphosphorylation of 4E-BP1 (Thomas & Hall, 1997). This causes dissociation from eIF4E and thereby activates translation (Pause et al., 1994). eIF4E is oncogenic and may be involved in transformation by other oncogenes (Sonenberg, 1996). 4E-BP1 is therefore a tumor suppressor, and its interaction with eIF4E is a possible model for tumor suppressing drugs (Rousseau et al., 1996).

NMR studies of 4E-BP1 and its homologue 4E-BP2 have shown that these proteins have little or no folded structure under physi-

ological solution conditions, yet are able to inhibit translation in reticulocyte lysates and to bind purified eIF4E (Fletcher et al., 1998). eIF4E binding and translation inhibition appears to be mediated by a short central region of the 4E-BPs (within residues 49–68 of 4E-BP1) with the rest of the protein remaining unfolded in the bound state (Mader et al., 1995; Fletcher et al., 1998).

The unusual structural nature of these proteins prompted the following alternative hypotheses: does free 4E-BP1 contain regions of preformed structure that are responsible for eIF4E binding, or is the interaction an induced fit to a completely flexible protein chain? Here we present resonance assignments and relaxation data that support the second hypothesis.

4E-BP1 was expressed in *Escherichia coli* and purified as previously described. Our protein construct differs slightly from the native sequence (Fig. 1) but is able to inhibit translation in reticulocyte lysates (Fletcher et al., 1998). Note that this protein is not phosphorylated, which could cause differences in structure and dynamics compared to 4E-BP1 in vivo.

The 4E-BP1 samples degraded significantly during the time required to record 3D NMR spectra, probably because the un-

Fig. 1. Sequence of 4E-BP1 and extent of the backbone assignment. Our protein construct is missing the N-terminal three residues of the native sequence (shown as dashes) and has proline instead of serine at position 6. Asterisks indicate residues for which backbone assignments were obtained. The underlined region contains the binding site defined by deletion, mutagenesis, and peptide experiments (Mader et al., 1995; Fletcher et al., 1998).

Reprint requests to: Mark Fletcher, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115; e-mail: cmf@hugin.med.harvard.edu.

folded protein is susceptible to proteolysis. A number of signals were attributed to degradation products because their intensity increased with time and/or because they were C-terminal or showed sequential connectivities to C-terminal residues (identified by unusually high ^{15}N and ^{13}CO and low NH backbone chemical shifts and an absence of further sequential connectivities). SDS-PAGE was used to check that the major protein species remained intact after each NMR experiment.

NMR spectra of 4E-BP1 displayed very high sensitivity due to long relaxation times, typical of unfolded proteins. A ^{15}N TOCSY-HSQC spectrum with 100 ms mixing time provided many intra-residue assignments. Amino acids were easily identified because of the large number of crosspeaks observed for each spin system and because the chemical shifts were very close to random coil values (Wishart et al., 1995). ^{15}N NOESY- and ROESY-HSQC spectra were simple to analyze because crosspeaks were only observed for intraresidue and sequential connectivities. However, it was difficult to obtain sequence-specific assignments from these data because ^1H chemical shifts for residues of the same amino acid type were very similar. Many additional sequential assignments were obtained from HNCO and HN(CA)CO spectra.

Backbone ^1H , ^{15}N , and ^{13}CO assignments were obtained for 77 out of 101 nonproline, nonterminal residues and ^{13}CO assignments for 12 out of 16 prolines (Figs. 1, 2). Except for the three proline-proline dipeptides, gaps in the backbone assignment are found only at residues 4–10, 43–46, and 56–65. For the assigned nonproline residues, approximately 90% of the side-chain ^1H assignments were obtained.

In an unfolded protein, regions with the same amino acid sequence are likely to have identical chemical shifts. This seems to occur for the repeated sequence STTPGGT at residues 35–50 in 4E-BP1. Thus, the chemical shifts of P38-G39-G40-T41 are almost certainly identical to those of P47-G48-G49-T50, supported by sequential assignments for both L42 and R51. The resonances of S44-T45-T46 may be overlapped with those of S35-T36-T37, although there is no evidence for this assignment. The same may occur for R63-N64-S65 (overlapping the signals assigned to R99-N100-S101).

Even if the last point is true, a large section of the sequence from 56–65 remains unassigned. This region is of particular interest because it is involved in eIF4E binding. The primary reason for the lack of assignments is the high degree of ^1H and ^{15}N chemical shift overlap, particularly because this section of the sequence contains many residue types that are expected to lie in the most crowded region of the spectrum (arginine, lysine, phenylalanine, and methionine). An additional problem could be signal broadening due to solvent exchange catalyzed by the basic side chains of arginine and lysine or due to conformational exchange in partly structured regions. ^{15}N HSQC spectra recorded at 17 °C and 7 °C showed no additional dispersed signals, so if broadening is occurring the signals are also likely to be overlapped. Completion of the assignments will probably require protein samples labeled at specific amino acid types.

Steady-state $\{^1\text{H}\}$ - ^{15}N NOEs were measured for 45 of the assigned residues (Fig. 3). Their values lie between -0.77 and -1.14 , except for the two C-terminal residues, which are more negative. These values indicate a high level of flexibility with no transient or local order. Other measures of flexibility are in agreement. First, we observed no homonuclear NOE or ROE crosspeaks that could not be explained by intraresidue or sequential interactions. Second, the backbone chemical shifts are very close to random coil values

Fig. 2. ^1H - ^{15}N correlation spectrum of 4E-BP1 with backbone assignments. This is a projection of the HNCO spectrum along the ^{13}C dimension. NMR samples of ^{15}N - and $(^{13}\text{C}, ^{15}\text{N})$ -labeled 4E-BP1 were prepared as previously described (Fletcher et al., 1998) except that the buffer was (50 mM $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ pH 6.5, 100 mM NaCl, 10 mM DTT, 1 mM EDTA, 0.1 mM NaN_3 , 0.1 mM PMSF, 1 Boehringer Mannheim protease inhibitor tablet per liter). Assignments were obtained from 500 MHz HNCO and HN(CA)CO spectra (Kay et al., 1994; Matsuo et al., 1996), 750 MHz ^{15}N TOCSY- and NOESY-HSQC spectra with mixing times of 100 and 250 ms, respectively (Zhang et al., 1994; Talluri & Wagner, 1996) and a 400 MHz ^{15}N ROESY-HSQC spectrum with 100 ms mixing time (C.M. Fletcher & G. Wagner, unpubl. data). Spectra were recorded at 27 °C unless otherwise stated.

measured in short peptides (Wishart et al., 1995). For ^{13}CO the mean absolute deviation from random coil values is 0.41 (the standard deviation σ is 0.30 and N is 88) and the maximum absolute deviation is 1.52. For $^1\text{H}\alpha$ the mean is 0.046 (σ is 0.04 and N is 77) and the maximum is 0.20. In comparison, the folded protein eIF1 gives the following values: for ^{13}CO a mean of 1.15 (σ is 0.92 and N is 111) and a maximum of 4.08; for $^1\text{H}\alpha$ a mean of 0.273 (σ is 0.285 and N is 90) and a maximum of 1.20 (C.M. Fletcher & G. Wagner, unpubl. data). Furthermore, the distribution of backbone chemical shift deviations along the sequence of 4E-BP1 shows no region of consistently larger deviations that might indicate the presence of structure (data not shown).

Since the binding site is likely to be ordered in the complex, we might expect this region to have some preformed structure in the free protein. The only heteronuclear NOE data available in this region are for isoleucine residues 52 and 53. The values are -1.05 and -1.13 , respectively, indicating that these residues are actually

Fig. 3. Steady state $\{^1\text{H}\}$ - ^{15}N NOEs for backbone amide groups of 4E-BP1. Spectra were recorded at 750 MHz and NOEs calculated as previously described (Farrow et al., 1994) except that peak intensities were measured by volume integration using the program XEASY (Bartels et al., 1995). NOEs were determined only from reasonably separated crosspeaks for which multiple assignments were not suspected.

among the most flexible. Chemical shift overlap prevents further assignment and quantitative NOE analysis at this stage. However, inspection of the NMR data allows us to draw further conclusions. The $\{^1\text{H}\}$ - ^{15}N NOE is given by $(I - I_0)/I_0$, where I is the intensity of the crosspeak in a spectrum recorded with presaturation of the ^1H frequency, and I_0 is the intensity in a reference spectrum recorded without presaturation. Any regions of structure in the protein are likely to have positive or smaller negative NOEs. It follows that crosspeaks for these residues in the spectrum with presaturation must be positive and have an intensity comparable to the crosspeaks in the reference spectrum.

Inspection of the 4E-BP1 presaturation spectrum shows no large positive signals, even among crosspeaks that have not been assigned. The maximum peak heights present are 47,420 for a crosspeak known to contain multiple signals and 28,790 for a dispersed crosspeak (this is assigned to E103 and gives rise to an NOE of -0.78). In contrast, the heights of dispersed peaks in the reference spectrum range from 29,440 to 201,400 with a mean of 97,723. It is possible that strong positive signals in the presaturation spectrum are reduced by signal broadening or by overlap with strong negative signals. Nevertheless, it seems likely that no positive or smaller negative NOEs are to be found in free 4E-BP1. This implies that the entire binding site, like residues 52 and 53, is flexible and has no preformed structure in the absence of eIF4E. We can therefore conclude that the interaction with eIF4E is an induced fit to a completely disordered protein chain.

There are now many examples of proteins that are both biologically active and unstructured under physiological solution conditions, several of which are known to adopt varying amounts of folded structure upon binding to their ligands (Penkett et al., 1997; Plaxco & Groß, 1997; Uesugi et al., 1997 and references therein). In two cases, the unfolded state has been analyzed using $\{^1\text{H}\}$ - ^{15}N NOE data that can be compared to the results presented here. The

activation domains of yeast heat shock transcription factors have uniformly negative NOEs of -0.3 or less, and are therefore similar to 4E-BP1 in being completely disordered in the absence of ligand (Cho et al., 1996). In contrast, the C-terminal domain of the bacterial protein FlgM has NOEs between 0.5 and -0.5 , indicating a significant amount of transient or local structure (Daughdrill et al., 1997). This may reflect more extensive folding upon binding to its ligand.

The long terminal sections of 4E-BP1 that flank the binding region (residues 1–48 and 69–118) have no proven function, although it has been suggested that PEST motifs in these regions regulate the turnover of 4E-BP1 (Fletcher et al., 1998). For this purpose a high degree of flexibility may be advantageous, since it will make the protein susceptible to rapid proteolytic degradation when this is required.

Acknowledgments: We thank Greg Heffron and Hiroshi Matsuo for help with NMR spectroscopy. This work was supported by NIH grant CA68262 to G.W.

References

- Bartels C, Xia T, Billeter M, Güntert P, Wüthrich K. 1995. The program XEASY for computer-supported NMR spectral analysis of biological molecules. *J Biomol NMR* 6:1–10.
- Cho HS, Liu CW, Damberger FF, Pelton JG, Nelson HC, Wemmer DE. 1996. Yeast heat shock transcription factor N-terminal activation domains are unstructured as probed by heteronuclear NMR spectroscopy. *Protein Sci* 5:262–269.
- Daughdrill GW, Chadsey MS, Karlinsky JE, Hughes KT, Dahlquist FW. 1997. The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, σ^{28} . *Nature Struct Biol* 4:285–291.
- Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. 1994. Backbone dynamics of a free and a phosphopeptide complexed src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* 33:5984–6003.

- Fletcher CM, McGuire AM, Li H, Gingras A-C, Matsuo H, Sonenberg N, Wagner G. 1998. 4E binding proteins inhibit the translation factor eIF4E without folded structure. *Biochemistry* 37:9–15.
- Kay LE, Xu GY, Yamazaki T. 1994. Enhanced sensitivity triple resonance spectroscopy with minimal H₂O saturation. *J Magn Reson A* 109:129–133.
- Mader S, Lee H, Pause A, Sonenberg N. 1995. The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 γ and the translational repressors 4E binding proteins. *Mol Cell Biol* 15:4990–4997.
- Matsuo H, Li H, Wagner G. 1996. A sensitive HN(CA)CO experiment for deuterated proteins. *J Magn Reson B* 110:112–115.
- Pause A, Belsham GJ, Gingras A-C, Donzé O, Lin T-A, Lawrence JC Jr, Sonenberg N. 1994. Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature* 371:762–767.
- Penkett CJ, Redfield C, Dodd I, Hubbard J, McBay DL, Mossakowska DE, Smith RAG, Dobson CM, Smith LJ. 1997. NMR analysis of main-chain conformational preferences in an unfolded fibronectin binding protein. *J Mol Biol* 274:152–159.
- Plaxco KW, Groß M. 1997. The importance of being unfolded. *Nature* 386:657–659.
- Rousseau D, Gingras A-C, Pause A, Sonenberg N. 1996. The eIF4E-binding proteins 1 and 2 are negative regulators of cell growth. *Oncogene* 13:2415–2420.
- Sonenberg N. 1996. mRNA 5' cap-binding protein eIF4E and control of cell growth. In: Hershey JWB, Mathews MB, Sonenberg N, eds. *Translational control*. Plainview, New York: Cold Spring Harbor Laboratory Press. pp 245–269.
- Talluri S, Wagner G. 1996. An optimized 3D NOESY-HSQC. *J Magn Reson A* 112:200–205.
- Thomas G, Hall MN. 1997. TOR signalling and control of cell growth. *Curr Opin Cell Biol* 9:782–787.
- Uesugi M, Nyanguile O, Lu H, Levine AJ, Verdine GL. 1997. Induced α helix in the VP16 activation domain upon binding to a human TAF. *Science* 277:1310–1313.
- Wishart DS, Bigam C, Holm A, Hodges RS, Sykes BD. 1995. ¹H, ¹³C, and ¹⁵N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest neighbour effects. *J Biomol NMR* 5:67–81.
- Zhang O, Kay LE, Olivier JP, Forman-Kay JD. 1994. Backbone ¹H and ¹⁵N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced sensitivity pulsed field gradient NMR techniques. *J Biomol NMR* 4:845–858.