Conformational analysis of the interdomain linker of the central homology region of chloroplast initiation factor IF3 supports a structural model of two compact domains connected by a flexible tether

Yuxin Hua[a], Daniel P. Raleigh[a,b,*]

[a] *Department of Chemistry, State University of New York at Stony Brook, Stony Brook, NY 11794-3400, USA*
[b] *Graduate Program in Biophysics and Graduate Program in Molecular and Cellular Biology, State University of New York at Stony Brook, Stony Brook, NY 11794-3400, USA*

Received 29 May 1998; revised version received 6 July 1998

Abstract A peptide corresponding to the interdomain linker of chloroplast IF3 has been synthesized and its structure studied by NMR and CD as a function of temperature and pH. At low temperature and neutral pH the apparent helical content is 25%. pH and ionic strength dependent CD studies demonstrate that sidechain-sidechain interactions stabilize the structure observed at low temperature. The helicity decreases with temperature and above 25°C the peptide is less than 15% helical. These results indicate that the peptide has little intrinsic tendency to form helical structure at physiologically relevant temperatures and strongly suggests that the linker region is flexible in intact chloroplast IF3.

Key words: Initiation factor 3; Alpha-helix; Helix-coil transition; Protein structure; *Euglena gracilis*

1. Introduction

Initiation factor 3 (IF3) plays a role in the initiation stage of protein synthesis by binding to the 30S subunit of the ribosome and preventing its association with the 50S subunit [1]. IF3 also enhances the binding of IF1 and IF2 to the 30S subunit. In addition, IF3 helps to ensure the accuracy of the initiation of translation by promoting the dissociation of initiation complexes that involve a non-initiator tRNA or a non-initiation codon [2,3]. Prokaryotic IF3 is a two domain protein comprised of compact N- and C-terminal domains connected by an interdomain linker [4–6]. Chloroplast IF3 isolated from *Euglena gracilis* is the only organellar IF3 which has been purified and characterized [7,8] and there is currently no structure of chloroplast IF3. Chloroplast IF3 is approximately twice as large as prokaryotic IF3s, however the central region of this molecule shows some homology to prokaryotic IF3. A model of part of the central region of chloroplast IF3 has been constructed using the crystal structures of the N- and C-terminal domains of *Bacillus stearothermophilus* as a starting point [8]. The model suggests that the C-terminal domain of the central region of chloroplast IF3 is similar to the C-terminal domain of the *B. stearothermophilus* protein. The model also predicts that the N-terminal domain of the central region of the chloroplast protein resembles the N-terminal domain of *B. stearothermophilus* IF3 but contains an additional β-strand. The relative orientation of these two domains, which is likely to be of functional importance, will be determined by the conformation of the interdomain linker region. Studies of chloroplast IF3 have shown that mutation of several conserved lysine residues in the linker decreases the activity of the protein, suggesting that the linker region is not just a passive connecting element [8]. The orientation of these residues will depend upon the structure of the linker. The relative spatial organization of the domains is also likely to be important for the interactions of IF3 with the ribosome. A rigid linker would ensure that the relative orientation and distance between the two domains remained constant, while a disordered linker would allow more flexibility. Unfortunately, the conformation of the chloroplast IF3 linker region is difficult to model since its primary sequence is significantly different from that of both *Escherichia coli* and *B. stearothermophilus* IF3, the two IF3s for which structural information is available. This is a particularly important point since there are conflicting structural models for these two proteins which differ in the conformation of the linker region. The crystal structure of the N-terminal domain of the *B. stearothermophilus* protein shows that the protein forms a compact domain with the C-terminal helix projecting away from the body of the protein leading to the suggestion that the two domains are connected by a structured helix [9]. Solution phase small angle neutron scattering experiments are consistent with this model [10]. In contrast, NMR studies of intact *E. coli* protein show that the linker is highly flexible in solution [11]. The NMR studies of the *E. coli* protein support a model of two compact domains linked by a flexible tether while the crystallographic and solution scattering studies of the *B. stearothermophilus* protein support a structure in which the linker region forms a rigid helical rod. The linker regions of prokaryotic IF3s have minimal interactions with the remainder of the protein indicating that the conformation is dominated by local interactions. We have previously demonstrated that the structure of the isolated linker domains from *E. coli* and *B. stearothermophilus* IF3 are very different and that these differences correlate extremely well with the structure of the intact proteins [12]. Here we report the results of a detailed CD and NMR study of the conformation of the isolated linker region of *E. gracilis* chloroplast IF3. We also compare our results to the

*Corresponding author. Fax: (1) (516) 632-7960.
E-mail: draleigh@ccmail.sunysb.edu

Abbreviations: CD, circular dichroism; FMOC, 9-fluorenylmethyloxycarbonyl; HPLC, high pressure liquid chromatography; IF3, initiation factor 3; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; PAL, 5-(4′-Fmoc-aminomethyl-3′,5′-dimethoxyphenoxy) valeric acid; TBTU, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate; TFA, trifluoroacetic acid; TOCSY, total correlated spectroscopy; TSP, 3-(trimethylsilyl) propionate

0014-5793/98/$19.00
PII: S0014-5793(98)00901-6

results of a previous study of the linker region of *E. coli* and *B. stearothermophilus* IF3 and discuss the implications for the domain organization of chloroplast IF3.

2. Materials and methods

2.1. Peptide synthesis and purification

The chloroplast linker peptide was synthesized using a Perseptive BioSystem 9050 Plus Peptide Synthesizer. Fluorenylmethyloxycarbonyl (Fmoc)-protected amino acids were purchased from Advanced ChemTech and PerSeptive BioSystems. The activating reagent 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (TBTU) was purchased from PerSeptive BioSystem. All solvents used for peptide synthesis and purification were from Fisher Scientific. Fmoc-protected 5-(4′-Fmoc-aminomethyl-3′,5′-dimethoxyphenoxy) valeric acid-PEG resin was used as the support. Peptides were acetylated at the N-terminus and amidated at the C-terminus. The peptides were cleaved from the resin using a solution of 3% anisole, 3% thioanisole and 3% ethanedithiol in trifluoroacetic acid solution for 2 h. The cleaved crude peptide was purified by reverse phase HPLC on a C-18 column using an A-B gradient. Buffer A consisted of 0.1% trifluoroacetic acid in water and buffer B was 0.1% trifluoroacetic acid and 10% water in acetonitrile. The peptide was $>95\%$ pure as judged by analytical HPLC. The composition of the pure peptide was verified by MALDI-TOF mass spectroscopy (calculated 2411.8 Da, observed 2409.7 Da).

2.2. CD spectroscopy

All CD studies were performed on an Aviv model 62A DS CD spectrometer. A buffer containing 2 mM sodium phosphate, 2 mM sodium citrate, 2 mM sodium borate and 10 mM NaCl was used for the pH dependence studies. The initial solution was prepared at low pH and titrated to higher pH with small amounts of concentrated NaOH. A total of 50 μl NaOH was added to 3.0 ml of sample. The concentration and temperature dependent experiments were performed using a 30 mM phosphate buffer. Temperature dependent experiments were performed using a 1 mm cuvette over the range 1°C to 94°C. The temperature was increased in steps of 3°C. The sample was allowed to equilibrate for 2 min at each new temperature before recording the signal at 222 nm. A 10 mM MOPS buffer containing varying amounts of NaCl was used for the ionic strength studies. Peptide concentration was determined by UV absorbance in the presence of 6M guanidine hydrochloride using an extinction coefficient of 1450 M^{-1} cm^{-1} for tyrosine at 275 nm [13]. In order to calculate the fraction helicity it is necessary to have a reliable model for the temperature dependence of the ellipticity of the random coil state. The temperature dependence of the mean residue ellipticity of the peptide was measured in 6 M and in 3 M guanidinium hydrochloride. The temperature dependent curves are identical, strongly suggesting that these measurements provide a good estimate of the value of θ_{222} for the unfolded state.

2.3. NMR studies

NMR spectra were acquired on a Varian INOVA 600 MHz spectrometer. The experiments were carried out in 90% H_2O and 10% D_2O at pH 5.0. Presaturation was used to suppress the water signal in all pulse sequences. Chemical shifts are referenced to 3-(trimethylsilyl)propionic 2,2,3,3-d_4 acid sodium salt (TSP). A TOCSY and a NOESY spectrum were used to make assignments. A mixing time of 80 ms was used for the TOCSY experiment and 400 ms for the NOESY experiment. Data sets were acquired with 2048 points in t2 for the TOCSY and NOESY experiments. All 2D spectra were acquired using TPPI [14] with 512 increments and zero filled to 2048 points. A 90° phase shifted sine-square window function was applied in both dimensions for the TOCSY and NOESY experiments. The deviation between the measured CαH chemical shifts and random coil values was calculated using the experimental values determined at 3°C and the random coil values derived from Merutka et al. which were measured at 4°C and pH 5.0 [15].

3. Results and discussion

A peptide corresponding to the nineteen residue linker which connects the N- and C-terminal domains of the central homology region of chloroplast IF3 was chemically synthesized. The peptide was prepared with an acetylated N-terminus and an amidated C-terminus to avoid any complications due to charge helix-dipole interactions. The sequence of the peptide is listed in Table 1 which also includes, for comparison, the sequences of the *E. coli* and *B. stearothermophilus* linker regions [7,16]. Concentration-dependent CD studies were performed to test for association and aggregation. The mean residue ellipticity at 222 nm is independent of concentration over the range of 2 μM to 1 mM. The linewidths and chemical shifts of the proton NMR spectra are also independent of concentration over the range of 0.5–3 mM. The CD spectrum of the chloroplast peptide recorded at pH 6.5, 1°C is shown in Fig. 1 and clearly does not correspond to that expected for a random coil conformation. There is a pronounced shoulder centered near 222 nm and the low wavelength minimum is observed at 201 nm which is higher than expected for a random coil peptide. The mean residue ellipticity measured at 222 nm is −8400. The structure is abolished in 3 M guanidine hydrochloride, and the mean residue ellipticity at 222 nm is +2000. If this value of θ_{222} is taken as that expected for a random coil peptide, then the mean residue ellipticity recorded in the absence of denaturant corresponds to 25% helicity. This value is considerably less than that previously observed for a peptide corresponding to the linker region of *B. stearothermophilus* at low temperature 53%, but is noticeably higher than that measured for a peptide corresponding to the linker region of *E. coli* IF3, 8% [12]. The CD spectrum of the *B. stearothermophilus* peptide is included in Fig. 1 for comparison.

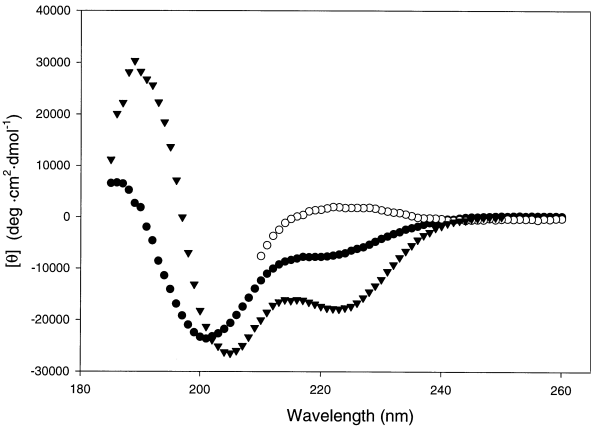

Fig. 1. The far UV CD spectrum of the choloroplast peptide recorded at 1°C and pH 6.5 in H_2O (filled circles) and in 3 M guanidine hydrochloride (open circles). The CD spectrum of the *B. stearothermophilus* peptide (filled triangles) is included for comparison. Samples were prepared in a buffer containing 2 mM sodium phosphate, 2 mM sodium citrate, 2 mM sodium borate and 10 mM NaCl.

Table 1
The primary sequence of the linker peptides from *E. gracilis* chloroplast IF3, *B. stearothermophilus* IF3, and *E. coli* IF3

Organism	Sequence
E. gracilis	NYSKLKYESEKKKKDSHKK
B. stearothermophilus	DYGKFRFEQQKKEKEARKK
E. coli	DYGKFLYEKSKSSKEQKKK

Two dimensional ^{1}H NMR experiments provide additional evidence that the chloroplast linker peptide is less structured than the *B. stearothermophilus* linker. There is a great deal of overlap in the fingerprint region of the spectra making sequential assignments difficult. In particular, the large number of lysine residues makes it impossible to unambiguously assign those residues in a sequence specific manner. Nevertheless, the NMR data are entirely consistent with the CD studies. The CαH chemical shifts of all of the residues are close to random coil values. The largest upfield shift is only 0.14 ppm from random coil values. In contrast, the CαH chemical shifts of all but two of the residues of the *B. stearothermophilus* peptide have been shown to be shifted upfield of the random coil values by greater than 0.15 ppm and several are shifted by more than 0.3 ppm.

The helical content is, as expected strongly temperature dependent. A plot of the measured fraction helix versus temperature is displayed in Fig. 2. At 25°C the apparent helical content decreases to 15% and it is even smaller at higher temperatures. This value is comparable to that measured for the *B. stearothermophilus* and *E. coli* linker peptides at their respective temperatures of optimum growth. This result strongly suggests that the linker region of chloroplast IF3 from *E. gracilis* is flexible in the intact protein and offers excellent evidence in support of the model of Spremulli and coworkers [8].

The chloroplast linker peptide is largely unstructured at and above room temperature but the helical content measured at low temperature is higher than is typically observed for isolated peptides and it is interesting to examine the factors which contribute to this residual structure. pH dependence studies provide direct experimental evidence that interactions involving charged sidechains play a role in stabilizing the helical structure observed at low temperature. Fig. 3 shows the helix content of the peptide as a function of pH from 2.2 to 11.2. The maximum helical content of 40% is found near pH 10.4. The helical content of the peptide decreases dramatically below pH 4.5, due to protonation of the Asp and Glu residues. A noticeable loss of helical character is also observed above pH 10.4, presumably because at these extremes of pH the helix stabilizing ion pairs are not formed. The peptide still retains significant helicity at the highest pH measured (25% at pH 11.2). This is likely due to favorable sidechain to sidechain hydrogen bonding interactions which can still form at high pH [17,18]. If electrostatic interactions play an important role in stabilizing the structure of the peptide observed at low temperature, the helical content should depend upon ionic strength. This is precisely what is observed. At high salt, 1.5 M and above, the helicity is reduced by roughly 50%. The ionic strength and pH dependent studies clearly indicate that the helix is stabilized by favorable electrostatic interactions.

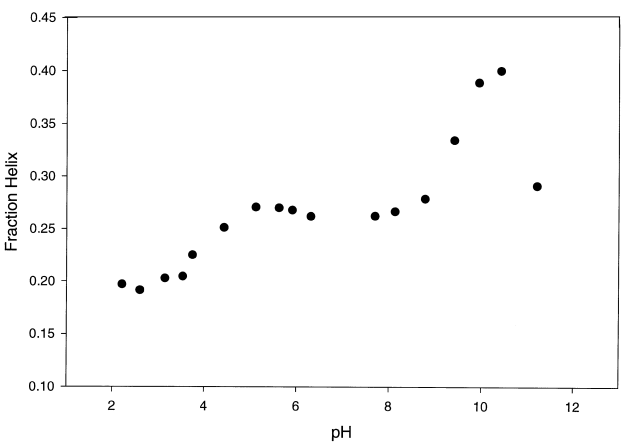

Fig. 2. The pH dependence of the apparent helical content of the chloroplast peptide measured at 1°C in a buffer containing 2 mM sodium phosphate, 2 mM sodium citrate, 2 mM sodium borate and 10 mM NaCl.

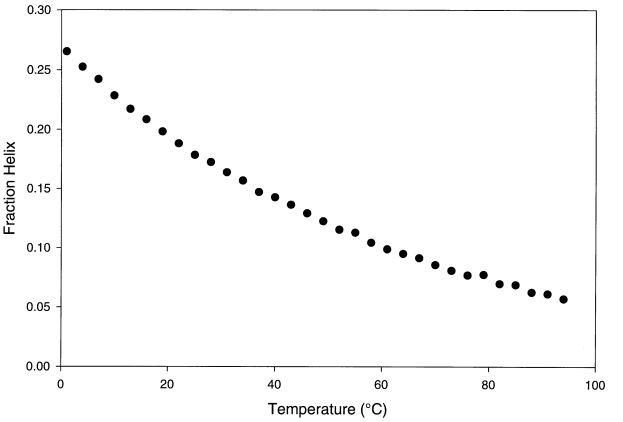

Fig. 3. The temperature dependence of the apparent helical content of the chloroplast linker peptide as a function of temperature. The data were collected at pH 6.5 in 30 mM phosphate.

There are several possible explanations for the differences in the helical content of the *E. gracilis* chloroplast and the *B. stearothermophilus* linker peptides. There are differences in the number, type and location of the ionizable residues and this almost certainly contributes to the different propensities to fold. There are additional differences between the peptides which likely contribute to their different conformational propensities. The *B. stearothermophilus* peptide is stabilized by an N-capping interaction which is likely weaker in the *E. gracilis* chloroplast peptide. The capping residue at the start of the *B. stearothermophilus* linker is an Asp residue. This residue is an Asn in the *E. gracilis* chloroplast IF3 linker. Asp is expected to be more favorable at the N-terminus of a helix since the negative charge should be able to interact favorably with the positive end of the helix dipole [19–21]. The chloroplast peptide also contains three serine residues while the *B. stearothermophilus* peptide contains none. Serine is known to have a low intrinsic helix propensity and this may contribute to the differing tendencies to adopt a helical structure.

In summary, the results reported in this paper offer excellent evidence that the linker region *E. gracilis* chloroplast IF3 is disordered at physiologically relevant temperatures. Our results strongly support the model of Yu and Spremulli which postulates that the linker region is highly flexible and largely disordered in *E. gracilis* chloroplast IF3 [8]. It is interesting to compare the conformational tendencies of the chloroplast linker peptide to the *B. stearothermophilus* and *E. coli* linker peptides. There are clear differences in the helical structure of the three peptides at low temperature but the helical content is striking similar if they are all compared at physiologically relevant temperatures, suggesting that the structure of the proteins may, in fact, be similar when compared at their respective physiologically relevant temperatures.

Acknowledgements: This work was supported in part by NSF Grant MCB 9600866 to D.P.R. D.P.R. is a Pew Scholar in the Biomedical Sciences. We thank Ms. Shari Spector for critically reading the manuscript and for helpful comments. We thank Mr. Daniel Moriarty for recording the MALDI-TOF mass spectra.

References

[1] Gualerzi, C.O. and Pon, C.L. (1990) Biochemistry 29, 5881–5889.
[2] Weiel, J. and Hershey, J.W.B. (1981) Biochemistry 20, 5859–5865.
[3] Zucker, F.H. and Hershey, J.W.B. (1986) Biochemistry 25, 3682–3690.
[4] Bellis, D.D., Liveris, D., Goss, D., Ringquist, S. and Schwartz, I. (1992) Biochemistry 31, 11984–11990.
[5] Fortier, P.L., Schmitter, J.M., Garcia, C. and Dardel, E. (1994) Biochimie 76, 376–383.
[6] Garcia, C., Fortier, P., Blanquet, S., Lallemand, J.Y. and Dardel, F. (1995) J. Mol. Biol. 254, 247–259.
[7] Lin, Q., Ma, L., Burkhart, W. and Spremulli, L.L. (1994) J. Biol. Chem. 269, 9436–9444.
[8] Yu, N.J. and Spremulli, L.L. (1997) Biochemistry 36, 14827–14835.
[9] Biou, V., Shu, F. and V, R. (1995) EMBO J. 14, 4056–4064.
[10] Kycia, J., Biou, V., Shou, F., Gerchman, S., Graziano, V. and Ramakrishnan, V. (1995) Biochemistry 34, 6183–6187.
[11] Moreau, M., Coch, E., Fortier, P., Garcia, C., Albaret, C., Blanquet, S., Lallemand, J. and Dardel, F. (1997) J. Mol. Biol. 266, 15–22.
[12] Hua, Y. and Raleigh, D.P. (1998) J. Mol. Biol. 278, 871–878.
[13] Brandts, J.R. and Kaplan, K.J. (1973) Biochemistry 12, 2011–2024.
[14] Marion, D. and Wüthrich, K. (1983) Biochem. Biophys. Res. Commun. 113, 967–974.
[15] Merutka, G., Dyson, H.J. and Wright, P.E. (1995) J. Bio. NMR 5, 14–24.
[16] Lin, Q., Ma, L., Burkhart, W. and Spremulli, L.L. (1994) J. Biol. Chem. 269, 16984.
[17] Scholtz, J.M., Qian, H., Robbins, V.H. and Baldwin, R.L. (1993) Biochemistry 32, 9668–9676.
[18] Huyghues-Despointes, B.M.P. and Baldwin, R.L. (1997) Biochemistry 36, 1965–1970.
[19] Aurora, R. and Rose, G.D. (1998) Protein Sci. 7, 21–38.
[20] Dasgupta, S. and Bell, J.A. (1993) Int. J. Peptide Protein Res. 41, 499–5111.
[21] Richardson, J.S. and Richardson, D.C. (1988) Science 240, 1648–1652.