The Folding Nucleus of the Insulin Superfamily

A FLEXIBLE PEPTIDE MODEL FORESHADOWS THE NATIVE STATE

Qing-xin Hua, John P. Mayer, Wenhua Jia, Jingwen Zhang, and Michael A. Weiss

Oxidative folding of insulin-like growth factor I (IGF-I) and single-chain insulin analogs proceeds via one- and two-disulfide intermediates. A predominant one-disulfide intermediate in each case contains the canonical A20–B19 disulfide bridge (cystines 18–61 in IGF-I and 19–85 in human proinsulin). Here, we describe a disulfide-linked peptide model of this on-pathway intermediate. One peptide fragment (19 amino acids) spans IGF-I residues 7–25 (canonical positions B8–B26 in the insulin superfamily); the other (18 amino acids) spans IGF-I residues 53–70 (positions A12–A21 and D1–D8). Containing only half of the IGF-I sequence, the disulfide-linked polypeptide (designated IGF-p) is not well ordered. Nascent helical elements corresponding to native α-helices are nonetheless observed at 4 °C. Furthermore, 13C-edited nuclear Overhauser effects establish transient formation of a native-like partial core; no non-native nuclear Overhauser effects are observed. Together, these observations suggest that early events in the folding of insulin-related polypeptides are nucleated by a native-like molten subdomain containing CysA20 and CysB19. We propose that nascent interactions within this subdomain orient the A20 and B19 thiolates for disulfide bond formation and stabilize the one-disulfide intermediate once formed. Substitutions in the corresponding region of insulin are associated with inefficient chain combination and impaired biosynthetic expression. The intrinsic conformational propensities of a flexible disulfide-linked peptide thus define a folding nucleus, foreshadowing the structure of the native state.

The vertebrate insulin-related superfamily consists of insulin, insulin-related growth factors (IGF-I and IGF-II), relaxin (3–5), and relaxin-related factors (6–9). Insulin, insulin-related growth factors (IGF-I and IGF-II), relaxin (3–5), and relaxin-related factors (6–9). Insulin and IGF-I's function as ligands for receptor tyrosine kinases (1, 2), relaxin (3–5), and relaxin-related factors (6–9). Insulin and IGF-I are associated with inefficient chain combination and impaired biosynthetic expression. The intrinsic conformational propensities of a flexible disulfide-linked peptide thus define a folding nucleus, foreshadowing the structure of the native state.
is played in each case by formation of the canonical A20–B19 disulfide bridge (cystine 18–61 in IGF-I). In the native states of insulin and IGF-I (Fig. 1) this bridge connects the C-terminal α-helix of the A-domain to the central α-helix of the B-domain, packing within a cluster of conserved aliphatic and aromatic side chains in the hydrophobic core (Fig. 1; Refs. 31–33). In the IGF-I pathway near neutral pH this is the only one-disulfide species to accumulate (34).

Cystine A20–B19 in the insulin-related superfamily is thus proposed to stabilize a specific and polarized folding nucleus (specific folding nucleus (35, 36)).

The structural role of the A20–B19 disulfide bridge in IGF-I intermediates has been investigated through construction of IGF-I analogs containing pairwise substitution of the other cysteines by Ala or Ser (15). Such analogs exhibit partial folds with attenuated but non-negligible α-helix content. As successive disulfide bridges are introduced in these equilibrium models, 1H NMR spectra exhibit a progressive increase in chemical-shift dispersion, suggesting stepwise stabilization of structure (15). Although resonance assignments were not obtained in these studies, comparison to the assigned spectrum of native IGF-I (37, 38) suggests formation of a native-like subdomain near cystine A20–B19 (IGF-I residues 18–61; Ref. 15). Additional evidence for the critical role of this cystine has been obtained through structural analysis of insulin analogs lacking either cystine A6–A11 or A7–B7 (36, 39, 40). Because information required for the folding of proinsulin is contained within the A and B chains (41, 42), these analogs may be regarded as peptide models of corresponding two-disulfide proinsulin intermediates. Although of low thermodynamic stability and partially unfolded, their predominant conformations contain a native-like cluster of non-polar side chains surrounding the A20–B19 disulfide bridge. Structures of analogous one-disulfide insulin analogs have not been described.

In this article we describe the synthesis and characterization of a one-disulfide-linked peptide model of the populated 18–61, IGF-I intermediate (14, 29). The pattern of polar and non-polar residues in the polypeptide (IGF-p; Table 1) is conserved among insulin, IGF-II, relaxin, and relaxin-related proteins (cysteines are boxed, and the critical residues Cys18 and Cys61 are marked by asterisks; top line of Fig. 2, A and B). An analogous one-disulfide PIP analog containing cystine A20–B19 has been described by Feng and co-workers (43). Although the reduced polypeptide could efficiently form the A20–B19 disulfide bridge in vitro, its folding and secretion in Saccharomyces cerevisiae were severely impaired. The PIP analog was found to exhibit marked attenuation of CD-defined helix con-
formation of subsequent disulfide bridges. Implications for the foldability of insulin and related proteins are discussed.

MATERIALS AND METHODS

Peptide Synthesis—The protocol for solid-phase synthesis is as described (44). The unlabeled linear sequences (peptides P8 and P8D; Fig. 2, A and B) were assembled by Fmoc/t-buty1 based methodology (45) using an ABI-431A instrument (Applied Biosystems, Foster City, CA). The manufacturer’s single coupling protocol utilizing preformed N-hydroxybenzotriazole esters was utilized with no modification. Starting with commercially available Wang resins (46), each synthesis was carried out on 0.25-mmol scale using Fmoc-protected amino acids with standard side chain protecting groups. After completion of syntheses, resins were subjected to acidolytic deprotection and cleavage using a mixture consisting of trifluoroacetic acid, thiocresol, anisole, and dicyclohexylcarbodiimide/4-(2-hydroxybenzotriazole esters, resins were subjected to acidolytic deprotection and cleavage using a mixture consisting of trifluoroacetic acid, thiocresol, anisole, and dicyclohexylcarbodiimide/4-(2-hydroxybenzotriazole esters, the sequences were assembled by a manual dicyclohexylcarbodiimide/benzyl methodology to enable real-time monitoring of coupling efficiencies by the ninhydrin reaction (47). Starting with either 4-4-methylbenzylhydrylamine resin (48), the sequences were assembled by a manual dicyclohexylcarbodiimide/N-hydroxybenzotriazole-mediated first coupling followed by a routine second coupling. At labeled residues the first coupling involved activation of the labeled material with dicyclohexylcarbodiimide/N-hydroxybenzotriazole followed by an extended coupling period of several hours. To conserve the labeled material, an additional coupling at that position was carried out using the unlabeled form of the same amino acid. Following removal of the N-terminal t-Boc group, peptide resins were cleaved under high HF conditions for 60 min at 0 °C using 5% p-thiocresol and 5% m-cresol as scavengers. Sites of specific isotopic labeling are summarized in Table 1.

Peptide Purification—Reverse-phase high-performance liquid chromatography (RP-HPLC) was used to purify both Fmoc and t-Boc-derived materials. Crude peptides were applied to a Vydac C-18 column and eluted with a linear acetonitrile gradient in 0.1% trifluoroacetic acid. Fractions exhibiting the best analytical profile were pooled and lyophilized. Yields of purified peptides were in the 10–20% range, based on initial resin substitution levels.

Disulfide Pairing—To obtain IGF-p, chain combination reactions were performed by mixing equimolar amounts of the individual chains at a concentration of 1.0–1.5 mg/ml in distilled water/acetonitrile (2:1) at pH 9.5 for 48–72 h. HPLC analysis of reaction mixtures revealed three peaks (each ~30% of the total), corresponding to the two homodimers as well as the desired heterodimer. To optimize the yield of labeled material following preparative RP-HPLC isolation of the heterodimer, labeled homodimers were reduced in the presence of dithiothreitol, re-purified, and used in another chain combination. Recycling labeled material resulted in an eventual yield of 10–20 mg of each heterodimer.

**Figure 2.** Homologous protein sequences. A and B, sequence of IGF-p in relation to human IGF-I, IGF-II, insulin, relaxin-1, -2, and -3, and insulin-like peptides 3–6. Cysteines are boxed; asterisks indicate Cys18 and Cys61 (canonical positions A20 and B19 in insulin).

---

5 Deconvolution of CD spectra suggest that native PIP exhibits a helix content of 48%, consistent with crystal structures of related analogs (83), the one-disulfide PIP analog exhibits a helix content of about 9% (43), similar to that observed on removal of cystine A7–B7 (84).
An Insulin-related Folding Intermediate

was subjected to amino acid analysis and mass spectroscopy; results coincided closely with expected values. As a control for disulfide-dependent structural features, spectroscopic studies were in part repeated following reduction of the inter-chain disulfide bridge by dithiothreitol; perdeuterated dithiothreitol was purchased from Cambridge Isotopes, Inc. (Woburn, MA).

**Gel Permeation Chromatography**—To assess its oligomeric status, IGF-p was fractionated with Superdex<sup>®</sup> Peptide column HR 10/30 (30 cm × 10 mm; Amersham Biosciences) using a Waters 515 HPLC system. The column was equilibrated in 5% acetic acid (pH 2.5) in 50 mM KCl and run at a flow rate of 0.5 ml/min. IGF-p and control samples (25 μl) were introduced through a Waters 717-plus autoinjector. A range of molecular weight standards was employed to calibrate the column: melilitin (2.9 kDa), a 34-residue fragment of PTH-rp (a monomeric and disordered peptide of similar length to IGF-p), the dimer of HNF-p1 (the dimerization domain of hepatocyte nuclear factor-1α); human proinsulin (86 residues), and cytochrome c (12.5 kDa). IGF-p was loaded at a polypeptide concentration of 200 μM; elution was monitored at 280 nm. Samples and running buffer were cooled to 4 °C prior to loading; at elution the temperature was about 15 °C.

**Circular Dichroism**—Far ultraviolet CD spectra were acquired using an Aviv spectropolarimeter equipped with thermister control. Samples were made 30–50 μM in dilute HCl (pH 2); spectra were acquired at 4 and 25 °C in quartz cuvettes with 1-mm path length. Spectra were also obtained after addition of 20% ethanol and 20% trifluoroethanol (TFE), chosen as helicogenic organic cosolvents to mimic possible effects of 20% deuterioacetic acid in NMR studies. Spectra were also obtained in phosphate-buffered saline (pH 7.4). Helix contents were estimated by Selcon-3 (49).

**NMR Studies of IGF-p**—Spectra were obtained at 600 and 700 MHz in 20% deuterioacetic acid at 4–25 °C. These conditions, previously employed in NMR studies of IGF-I analogs (15) and insulin (33, 35, 50), enhance solubility relative to neutral pH. Resonance assignment was based on homonuclear two-dimensional NOESY (mixing times 80 and 250 ms), total correlation spectroscopy (TOCSY; mixing time 55 ms), and double-quantum filtered correlated spectroscopy spectra. In key cases resonance overlap was resolved based on selective <sup>2</sup>H and <sup>13</sup>C labeling; <sup>1</sup>H–<sup>13</sup>C HSQC and <sup>13</sup>C-edited NOESY spectra of labeled samples (Table 1) were obtained at 700 MHz. Control spectra were obtained in the absence of acetic acid at pH 2 (0.01 N HCl) and on progressive dilution in 20% acetic acid to a polypeptide concentration of 50 μM. Amide proton exchange was monitored in a D<sub>2</sub>O solution containing 20% deuterioacetic acid (51).

**Natural Abundance <sup>13</sup>C NMR Studies**—<sup>1</sup>H–<sup>13</sup>C HSQC spectra of unlabeled IGF-p were obtained at a polypeptide concentration of 1.5 mM at 700 MHz with cryogenic probe. The acquisition time was 18 h per spectrum. Spectra were obtained in (i) 20% deuterioacetic acid at 7, 15, 25, and 35 °C; (ii) 100% D<sub>2</sub>O at pH 2 and 7 °C; and (iii) 20% deuterio-TFE at pH 2 and 7 °C. Due to low digital resolution, line broadening is manifested as attenuated signal intensity resulting from transverse relaxation during the pulse sequence. <sup>1</sup>H–<sup>1</sup>H TOCSY experiments were obtained with 80 and 250 ms mixing times, and steady state data were acquired with mixing times of 55 ms.

**TABLE 1**

| Sample | P<sub>n</sub> | P<sub>25</sub> | P<sub>53</sub> | P<sub>AD</sub> |
|--------|-------------|-------------|-------------|-------------|
| 1      | GAEL VDALQF VCGD RGFYF DLRRLEMY CAPLKPAKSA | DLRRLEMY CAPLKPAKSA | DLRRLEMY CAPLKPAKSA | DLRRLEMY CAPLKPAKSA |
| 2      | GAEL VDALQF VCGD RGFYF DLRRLEMY CAPLKPAKSA | DLRRLEMY CAPLKPAKSA | DLRRLEMY CAPLKPAKSA | DLRRLEMY CAPLKPAKSA |
| 3      | GAEL VDALQF VCGD RGFYF DLRRLEMY CAPLKPAKSA | DLRRLEMY CAPLKPAKSA | DLRRLEMY CAPLKPAKSA | DLRRLEMY CAPLKPAKSA |

**FIGURE 3.** CD studies of IGF-I and analogs. A and B, far UV CD spectra of IGF-p (solid line), IGF-(18–61) (broken solid line), and IGF-I (dashed line) at pH 2.0 (A) and 7.4 (B). C and D, enhancement of helical features on addition of TFE at pH 2 and 4 °C. C, CD spectra of IGF-p in the absence of TFE (solid line), 20% TFE (long dashed line; center), 40% TFE (dashed line; triangles), 60% TFE (dashed-dotted line; filled circles), and 80% TFE (dotted line; open circles). D, corresponding CD spectra of native IGF-I. Because the apparent helix content of IGF-I in 80% TFE is consistent with crystal structures, its successive increase in helix content on addition of TFE may in large part reflect progressive stabilization of native α-helices in otherwise molten structure at pH 2.
following each HSQC experiments to enable partial resonance assignment.

NMR Studies of IGF-(18–61)—$^1$H NMR spectra of the corresponding one-disulfide analog of IGF-I were obtained in 10 and 20% deuteroacetic acid as described (15). Presumptive selected resonance assignments, proposed by Narhi et al. (15) by analogy to the assigned spectrum of native IGF-I, were verified by explicit sequential assignment. Assignments were obtained in the regions of IGF-(18–61) spanned by IGF-p but incomplete elsewhere due to degeneracy. Spectra were obtained at 600, 700, and 750 MHz, the latter were acquired at Varian Instruments Inc. (Palo Alto, CA).

RESULTS

IGF-p and three labeled analogs were prepared by chemical synthesis (Table 1). Residue numbers refer to the corresponding positions in native IGF-I; disulfide bridges and structural elements are also designated by canonical insulin nomenclature (e.g. the central $\alpha$-helix of the B-domain spans residues 8–18 in IGF-I, corresponding to canonical helix B9–B19 in insulin). Peptide $P_B$ (19 amino acids; IGF-I residues 7–25) thus spans canonical residues B8–B26. In the structure of native IGF-I (37, 38, 52, 53) this region forms a U-shaped helix-turn-strand super-secondary structure. Peptide $P_{AD}$ (18 amino acids) spans IGF-I residues 53–70 (positions A12–A21 and D1–D8). In native IGF-I this region comprises the C-terminal A-chain helix (residues 53–60) and less well organized D-domain extension (52, 53). Gel-permeation chromatography under conditions similar to those employed for NMR studies indicates that IGF-p is monomeric at a polypeptide concentration of 200 $\mu$M. $^1$H NMR spectra are unaffected by polypeptide concentrations in the 50 $\mu$M to 1.5 mM range.

CD Studies Suggest Disorder with Latent Helical Propensity—Far UV CD spectra of IGF-p, obtained at 4 °C at either pH 2 or 7.4 (solid lines in Fig. 3, A and B, respectively), suggest that the polypeptide contains little organized structure at either pH. The estimated mean helix content is at most 8–10% (corresponding to one helical turn). The spectrum of native IGF-I under these conditions (dotted lines in Fig. 3, A and B) by contrast exhibits a distinct helix-related maximum near 195 nm, minimum at 208 nm, and shoulder at 222 nm. These features correspond to an apparent helix content of 16–20%, less than the actual extent of $\alpha$-helix in crystal structures of IGF-I (25 of 70 residues, or 36%; Ref. 52 and 53). This discrepancy, which presumably reflects the flexibility of IGF-I in solution (37, 38), highlights the limitations of CD deconvolution algorithms trained on the basis of well ordered globular proteins (49). CD spectra of the one-disulfide analog IGF-(18–61) has features intermediate between the spectra of IGF-p and native IGF-I (dashed lines in Fig. 3, A and B).

Helical CD features of both IGF-p and IGF-I are markedly enhanced on addition of TFE (Fig. 3, C and D, respectively). In 20% TFE (v/v) the spectrum of IGF-p exhibits a distinct minimum at 208 nm and shoulder at 222 nm (dashed line in Fig. 3C). These features, similar to those of IGF-(16–81) in the absence of TFE (dashed lines in Fig. 3, A and B), correspond to an estimated helix content of about 13%. Helical CD features are further accentuated in 40, 60, and 80% TFE; these spectra, nearly identical (Fig. 3C), correspond to helix contents of 25–
27%. The CD spectrum of IGF-I also exhibits accentuated helical features in the presence of TFE (Fig. 3D). Unlike in studies of IGF-p, successive addition of TFE in the 40–80% range is associated with a progressive increase in the magnitude of helix-related features at 195, 208, and 222 nm. In 80% TFE the CD spectrum of IGF-I is consistent with its crystal structure. Comparison of the TFE dependence of the CD spectra of IGF-p and IGF-I suggest that the peptide model may contain a latent helical propensity analogous to that of the native protein.

**NMR Studies Demonstrate Native-like Conformational Propensities**—In light of the low resolution of CD spectroscopy, NMR studies were undertaken to provide a residue-specific probe of possible conformational propensities. One-dimensional $^1$H NMR spectra of IGF-p are shown in Fig. 4 (A and B, 4 and 25 °C, respectively) in relation to spectra of IGF-(18–61) (C) and native IGF-I (D). The one-disulfide polypeptides (panels A–C) exhibit a marked reduction in chemical shift dispersion relative to native IGF-I (panel D), especially evident in the methyl region (0–1 ppm). The majority of $^1$H NMR resonances in IGF-p are near random coil values, consistent with the absence of significant CD-detectable structure in the absence of TFE. As expected, the spectrum of IGF-p (37 residues) contains less than half of the overall spectral amplitude of the corresponding IGF-(18–61) (70 residues). Line widths at 4 (Fig. 4) and 40 °C (not shown) are in each case as expected for a monomeric polypeptide of the expected molecular masses. Resonance assignments of IGF-p and IGF-(18–61) are provided as supplemental materials (Tables S1 and S2, respectively). Resonances assignments of native IGF-I are given in Table S3.

Despite the overall paucity of chemical shift dispersion, aromatic chemical shifts in IGF-p and IGF-(18–61) are distinct from random coil values. Dispersion is particularly marked between the respective ortho aromatic resonances of Tyr$^{60}$ and Tyr$^{24}$ and likewise between their respective meta resonances. The chemical shifts of Tyr$^{60}$ (7.16 and 6.78 ppm in IGF-p at 4 °C and similarly 7.17 and 6.74 ppm in IGF-(18–61) at 40 °C) resemble those of native IGF-I (7.31 and 6.75), presumably due in each case to core packing (see below). Of particular interest are the non-negligible secondary shifts of corresponding Tyr$^{60}$ ortho resonances: in IGF-p this shift ($\Delta$ $\delta$) is 0.17 ppm relative to free tyrosine in 20% acetic acid. Corresponding values of $\Delta$ $\delta$ are 0.21 ppm in IGF-(18–61) and 0.20 ppm in IGF-I. By contrast, the meta and ortho chemical shifts of Tyr$^{24}$, which flexibly projects into solvent, are near random coil values: 7.02 and 6.78 ppm in IGF-p, 6.94 and 6.73 ppm in IGF-(18–61), and 7.04 and 6.77 in native IGF-I.

Upfield methyl chemical shifts, a distinctive NMR signature of the native state, are attenuated in the one-disulfide species, presumably reflecting the absence of fixed ring-current shifts in molten or disordered conformations. Significant differences are observed, for example, between corresponding Leu$^{14}$ methyl resonances, which are sensitive to the ring current of Phe$^{23}$. In native IGF-I this interaction leads to a significant upfield secondary shift (chemical shifts 0.70 and 0.48 ppm in IGF-I at 40 °C relative to reference position 0.98 ppm of free leucine in 20% acetic acid). By contrast, in IGF-p the methyl resonances of Leu$^{14}$ exhibit chemical shifts of 0.91 and 0.84 ppm at 4 °C; corresponding chemical shifts in IGF-(18–61) are 0.73 and 0.63 ppm at 40 °C. Progressive attenuation of Leu$^{14}$ secondary shifts may reflect trends in the extent of long-range order (see “Discussion”).

**FIGURE 5. Two-dimensional NMR analysis of IGF-p.** A, amide region, and B, fingerprint region of NOESY spectrum at 4 °C. Sequential assignment of residues 8–13 is illustrated. The sequential $d_{i,i+3}$ connectivities indicate nascent helix. In B, the following $d_{i,i+3}$ cross-peaks are labeled: (a) residues 8–11, (b) 9–12, and (c) 10–13. TOCSY (C) and NOESY (D) spectra of IGF-p: long-range NOEs are observed between aromatic and methyl resonances. Of particular interest are native-like B contacts involving Tyr$^{60}$, Phe$^{23}$, and Leu$^{14}$. An Insulin-related Folding Intermediate
The NOESY spectrum of IGF-p at 4 °C (Fig. 5) contains an unexpected richness of inter-residue cross-peaks, suggesting transient formation of non-random local and non-local contacts in the ensemble. Key NOEs are observed with similar intensities (relative to diagonal resonances) at polypeptide concentrations of 1 mM and 50 μM (supplemental Materials Fig. S1), suggesting that these contacts are intrinsic to the ensemble of monomeric conformations and not a consequence of weak dimerization. To explore such conformational preferences, a mixing time (250 ms) was chosen to allow spin diffusion at low temperature (such mixing time is suitable for a 37-residue polypeptide). The results are in striking contrast to the low CD-detectable helical structure of IGF-I and insulin, some (β-turns) exhibit intensities similar to those at 35 °C as expected of a random coil. These anomalous features provide evidence for conformational exchange within a disordered ensemble. Addition of 20% deutero-TFE leads to further broadening of these conformationally sensitive resonances and a shift in the H-α-Cα cross-peak of Val11 toward a helical fingerprint position (supplemental materials Fig. S2, D). Traces through representative cross-peaks in the natural abundance 1H-13C HSQC spectra, assigned to regions of disorder or nascent structure, are provided in the supplemental materials (Fig. S2, E and F, respectively). The greater sensitivity of 1H-13C HSQC spectroscopy (relative to homonuclear TOCSY and NOESY spectroscopy) to conformational broadening is due to a combination of the breadth of the 13C chemical shift scale (Δω between substates) and the marked dependence of HSQC cross-peak amplitude on 13C-α T2 relaxation times.

To investigate the dependence of the nascent structure on maintenance of the 18–61 disulfide bridge, IGF-p was reduced with 25 mM deuterated dithiothreitol in 25 mM ammonium bicarbonate (pH 8). Following lyophilization, the reduced sample was reconstituted in 20% deuteroacetic acid. Its natural abundance 1H-13C HSQC spectrum at 7 °C contained random coil PAD resonances without evidence of conformational broadening; the NOESY spectrum likewise contained few non-local contacts. P_B cross-peaks were not observed in either spectrum, presumably due to aggregation or fibrillation of the reduced peptide.

13C-Edited NMR Studies Demonstrate Nascent Tertiary Structure—Comparison of the NOESY spectrum of IGF-p with spectra of native IGF-I (37, 38) and native insulin (50) indicates that P_B contains a nascent native-like supersecondary structure: NOEs following the helix (Arg21 H_N/Gly22 H_N, Gly22 H_N/Asp20 H_N, H_N/Asp20 H_δ, and Phe23 H_δ/Leu14 δ₁₂-CH₃) provide evidence of a β-turn. Although the nascent helix in P_AD also corresponds to an α-helix in native IGF-I and insulin, some (i, i+3) NOEs in P_AD are anomalous (supplemental materials Table S5). Such contacts, not characteristic of canonical α-helices, presumably reflect the flexibility of the nascent helix. The disulfide bridge in IGF-p gives rise to characteristic inter-peptide NOEs between Cys18 and Cys61 (supplemental materials Table S5). The features are independent of the polypeptide concentration.
in the range 50 μM to 1 mM at 4–7 °C but are not observed at 25 °C.

The NOESY spectrum of IGF-p contains two key long-range NOEs, one between side chains Phe23–Leu14 (Fig. 5D and supplemental materials Table S5), and the other between side chains Tyr60–Leu14 (and/or Tyr60–Leu10). The first NOE provides evidence for a turn in P8 leading to a native-like secondary structure (see above); the second provides evidence for a native-like inter-chain helix-helix interaction. Because these and other diagnostic resonances are only partially resolved, site-specific 13C and 2H labels were introduced (Table 1). 1H-13C HSQC spectra verify sites of 13C labeling (supplemental materials Fig. S4). 13C-Edited NOESY spectra (Fig. 8) confirm assignments of cross-peaks that would otherwise be ambiguous in the spectrum of the unlabeled polypeptide (Fig. 5). An example is provided by the overlapping methyl resonance of Leu69 and Leu10. In one of the labeled samples the three leucine residues in P15 are perdeuterated, making unambiguous the assignment of the remaining cross-peak between Tyr60 and Leu14 (Fig. 8A). This contact corresponds to canonical packing of TyrA19 against LeuB15 in the core of insulin (13). A long-range NOE between Phe23 and Leu14 was likewise resolved by isotope editing (Fig. 8B), ambiguous in the unlabeled spectrum (Fig. 5) due to overlap between the aromatic resonances of Phe23 and Tyr24/Phe25. A potential non-native NOE between Tyr24 and Leu14 was excluded by its absence in an appropriately labeled sample (Fig. 8C). The latter spectrum also identified a medium-range Tyr24/Arg21 contact (Fig. 8C).

Other regions of this spectrum (supplemental materials Fig. S5) contain native-like helix-related (i, i+3) NOEs in PAD and a chemical shifts of the two aromatic rings in the nascent microdomain (Phe23 and Tyr60; canonical positions B24 and A19, respectively) are essentially identical in the two samples (Fig. 5D and supplemental materials Table S2).

DISCUSSION

How the native state of a globular protein is encoded by its sequence defines a major unsolved problem (54). Although folding may be assisted by chaperones and enzymes such as protein-disulfide isomerase (55), structures ordinarily represent the ground state of a multidimensional energy landscape (56). Globular proteins form cooperatively folded structures, stabilized primarily by the hydrophobic effect and multiple weak interactions simultaneously permitted in the native state (entropic cooperativity; Ref. 20). The conformational search leading to the ground state is of broad biomedical importance as folding defects underlie diverse human diseases (57, 58). This search is envisaged to occur in a funnel-like landscape (59) wherein the native state is reached from an initially large distribution of conformers (60). The efficiency of this search may be enhanced by preferred trajectories, one limit of which is the population of transient intermediates. Off-pathway intermediates may also accumulate, leading to kinetic barriers (61).

These concepts have been extensively investigated in studies of oxidative protein folding. Of interest in relation to disulfide pairing in the endoplasmic reticulum (62), oxidative folding is central to the biosynthesis of transmembrane and secretory proteins (63). In vitro the relative reactivities of thiol groups provide site-specific kinetic probes, enabling

FIGURE 7. Natural abundance 1H-13C HSQC spectrum of IGF-p. Chemical shifts are near random coil values. Whereas N- and C-terminal residues in each chain exhibit motional narrowing, evidence of a conformational equilibrium is provided by selective broadening of cross-peaks in regions of nascent structure. Assignment of selected Cα-Hα cross-peaks are indicated (Gly9, Val11, Ala13, Val15, Gly19, and Ser69); Cα-Hα cross-peaks of Ser69 and two prolines (Pro53 and Pro66) are also indicated. The spectrum was observed in 20% deuteroacetic acid at 7 °C; additional natural abundance spectra are provided as supplemental materials (Fig. S2).

Relationship of IGF-p to Corresponding One-disulfide Intermediate—IGF-(18–61) at 40 °C yields a pattern of inter-residue NOEs similar to that of IGF-p at 4 °C (supplemental materials Fig. S3 and Table S5). Such similarity suggests that cooling IGF-p enhances intrinsic conformational preferences that in IGF-(18–61) are stabilized at higher temperature by flanking regions of the 70-residue polypeptide. Sequential assignment of the A and B-domains of IGF-(18–61) demonstrates corresponding helical segments and long-range NOEs (supplemental materials Fig. S3 and Table S3), including Phe23–Leu14 and Tyr60–Leu14. Significantly, the methyl chemical shifts of Leu14 are upfield of those in IGF-p (and downfield of those in native IGF-I), suggesting that the nascent core of the longer polypeptide is more ordered. Of particular interest, the non-random long-range NOE (Ala62–Val72), whose proximity is constrained by the adjoining 18–61 disulfide bridge.

An Insulin-related Folding Intermediate
reaction intermediates to be trapped (64). The time course of formation and disappearance of free cysteines and specific pairing arrangements thus provides a chemical map of the oxidative folding pathway (65, 66). A paradigm has been provided by bovine pancreatic trypsin inhibitor (23, 24, 67). Like IGF-I and insulin, bovine pancreatic trypsin inhibitor is a small globular protein containing three disulfide bridges, and so its folding pathway provides a context for comparison. The subject of vigorous debate (36), refolding is characterized by a preferred sequence of one- and two-disulfide species. Both native and non-native pairings may be observed whose relative ratios are sensitive to experimental conditions (68). Unlike IGF-I intermediates (15, 35), bovine pancreatic trypsin inhibitor intermediates that are well populated at neutral pH exhibit native-like structures (23, 24, 67). Large kinetic barriers among such folded intermediates lead to a preferred final step: formation of an external disulfide bridge between solvent-exposed loops. Kinetic barriers are thus central to the logic of the overall disulfide pathway (27, 69). Given that a populated one-disulfide species exhibits a native-like fold (67), the kinetics of disulfide formation and rearrangement are unrelated to the logic of the initial conformational search from the unfolded state ensemble.

Refolding of IGF-I—The insulin-related superfamily provides a complementary system for studies of oxidative folding. Profound differences from bovine pancreatic trypsin inhibitor are observed. IGF-I and related proteins (including proinsulin and insulin) are significantly less stable (35, 68); kinetic barriers among populated two-disulfide species are also lower (35, 39). Furthermore, and most significant, the conformational search is fundamentally intertwined with formation of disulfide bridges: analogs of IGF-I and insulin exhibit stepwise stabilization of native structural elements with successive disulfide pairing (29, 32, 36). Oxidative folding of IGF-I and related proteins may thus be viewed as occurring on a sequence of successive energy landscapes whose topography is constrained by disulfide pairing (35).

Evidence for non-random initial folding trajectories is provided by the trapping of a unique one-disulfide intermediate containing the A20–B19 disulfide bridge (cystine (18–61) in IGF-I; Refs. 15 and 43). Although several subsequent two-disulfide species have been characterized, this species appears to be an obligatory on-pathway intermediate. The initial conformational search leading to formation of this disulfide bridge is thus of central importance. Here, we have constructed and characterized a peptide model of this key intermediate. Although highly flexible, analysis of inter-molecular NOEs at low temperature indicates transient formation of a native-like microdomain surrounding the A20–B19 disulfide bridge. We thus view IGF-I as adopting an ensemble of disordered structures in equilibrium with native-like conformational substates (Fig. 9). We imagine that transient formation of a native-like subdomain provides an “internal template” constraining subsequent folding trajectories and possible disulfide pairings.

The richness of the NOESY spectrum of IGF-I is remarkable in light of its uninformative CD spectrum. This apparent contradiction presumably reflects the differing physical mechanisms of these probes: whereas CD detects the mean handedness of the ensemble of main chain conformations, NOEs can arise through the transient proximity of protons in a flexible polypeptide. Because the CD spectrum excludes formation of well ordered α-helices, we interpret non-random NOE patterns as evidence of nascent conformational preferences rather than fixed structure. Despite such NOEs, chemical shift dispersion is limited. We presume that conformational fluctuations lead to averaging of ring-current shifts and other sources of chemical

---

**FIGURE 8.** 13C-Edited NOESY spectra of IGF-p. Sites of labeling are given in Table 1: sample 2 (spectrum A), sample 3 (B), and sample 4 (C). In each case NOEs between aromatic and aliphatic resonance are shown in the upper box, and intra-residue aromatic NOEs in the lower box.
shift anisotropy that could otherwise lead to non-random secondary shifts (70). The low apparent helix content implied by the CD spectrum of IGF-p is also in accord with the absence of protected amide resonances in D$_2$O within NOE-defined nascent helical segments. Evidence for conformational fluctuations in a related analog IGF-(18–61) has previously been obtained from tyrosine fluorescence studies (15).

Although the NOESY spectrum of IGF-p contains far fewer cross-peaks than the spectrum of native IGF-I, no non-native NOEs are observed, i.e. contacts not in accord with the crystal structure of IGF-I. Comparison of IGF-p and IGF-(18–61) (15) suggest that structural relationships in the peptide are retained and extended in the physiological intermediate, presumably due to interactions involving segments of IGF-I not represented in the peptide model. By analogy to two-disulfide insulin analogs (35, 36), we speculate that the non-polar side chain of Ile$^{13}$ contributes to desolvation of a hydrophobic mini-core even in the absence of A-domain helix 42–49. The conformational preferences of IGF-p nevertheless demonstrate that essential folding information is intrinsic to a subset of A- and B-domain residues. These results significantly extend previous NMR studies of IGF-I disulfide analogs (15, 35).

An unusual feature of IGF-I refolding is its bifurcation to yield two products (14, 29). Designated native IGF-I and IGF-swap (14), the two products share the 18–61 disulfide bridge but differ by interchange of Cys$^{10}$ and Cys$^{7}$; respective pairing schemes are (18–61, 6–48, and 47–52) and (18–61, 6–47 and 48–52). These isomers exhibit near equal thermodynamic stabilities. Proinsulin and single chain insulin analogs by contrast refold to form a unique native state; alternative pairing schemes are metastable, and if present, rearrange to the native state. That the number of products is precisely two demonstrates that a non-random folding pathway is encoded but must be “ambiguous” following formation of cystine-(18–61). Studies of chimeric PIP analogs indicate that the respective B-domains of proinsulin and IGF-I are responsible for determining the relative stability of the swapped isomer (72). Specification of native IGF-I disulfide pairing in vivo is attributed to specific IGF-I-binding proteins, present in equimolar proportions. Such partner proteins bind native IGF-I but not IGF-swap, thus favoring formation of the native isomer (73).

Application to Insulin and Proinsulin—Insulin chain combination provides a peptide model for the folding of proinsulin (31). Although the mechanism of disulfide pairing is not well characterized, evidence for the kinetic importance of an homologous A20–B19-related microdomain is provided by the contrasting effects of N- and C-terminal A chain substitutions. Whereas diverse substitutions in the N-terminal segment of the A chain are well tolerated (39), substitutions in the C-terminal α-helix can profoundly impair efficiency of chain combination (74). Because the A20–B19-related microdomain is proposed to function as a specific folding nucleus, destabilization of this substructure would also destabilize the transition state. An example of blocked chain combination is provided by mutations of Leu$^{A16}$ (Leu$^{57}$ in IGF-I). This side chain packs against the A20–B19 disulfide bridge and is proposed to assist in orienting Cys$^{A20}$ and Cys$^{B19}$ for proper disulfide pairing. Off-nucleus contacts involving the N-terminal segment of the A chain are optional in the proposed polarized transition state and so may be bypassed. Unstable N-terminal analogs are thus readily prepared. Substitutions in B-domain residues involved in the A20–B19-related microdomain are also associated with impaired chain combination (75) and decreased efficiency of PIP expression in yeast (43).

The robustness of insulin chain combination to drastic changes in the sequence and structure of the N-terminal A chain segment demonstrates asymmetric encoding of folding information. A polarized kinetic model rationalizes why it is possible to achieve efficient biosynthetic expression and folding of PIP variants containing pairwise substitution of either “off nucleus” cystine (A6–A11 or A7–B7) but not of cystine A20–B19 (76). Likewise, syntheses of two-disulfide insulin analogs lacking either cystine A6–A11 or A7–B7 were accomplished with high yield (36, 39). Efficient pairing of such analogs seems remarkable in light of their marked instabilities. In transfected mammalian cells a variant proinsulin analog lacking cysteines A6 and A11 exhibits an efficiency of folding and secretion similar to that of wild-type proinsulin (39). Expression and secretion are impaired on pairwise substitution of either A7–B7 or A20–B19, presumably due to misfolding and/or degradation. Populated one- and two-disulfide intermediates in the folding of proinsulin have not been characterized. This is due in part to a technical barrier. Whereas the refolding of reduced IGF-I can be investigated near neutral pH (14, 29), studies of proinsulin are restricted to basic conditions (pH 9.5–10.5) due to pH-dependent aggregation of the unfolded polypeptide. Under these conditions the protein rapidly oxidizes to form an assortment of three-disulfide isomers, which then rearrange to the native state (31). The relevance of this phenomenon to in vitro folding at neutral pH or physiological folding in the ER (pH 6) is unclear. Because aggregation of reduced proinsulin at neutral pH can be circumvented by the addition of protein-disulfide isomerase (and in particular of its chaperonin domain; Ref. 77), it would be of future interest to investigate the kinetic

---

7 The methyl resonances of Leu$^{14}$, expected to be shifted upfield by the ring current of Phe$^{23}$, exhibit near random coil chemical shifts of 0.91 and 0.84 ppm in IGF-p, in contrast to upfield positions 0.70 and 0.48 ppm in native IGF-I.
An Insulin-related Folding Intermediate

Acknowledgments—We are grateful to P. G. Katsoyannis for advice and encouragement throughout this study. We thank G. Gray (Varian Associates, Inc.) for use of a 750 MHz NMR spectrometer, M. DeFellippas and B. H. Frank (Eli Lilly & Co.) for human insulin and proinsulin, T. Gardella and H. Kronenberg (Massachusetts General Hospital) for PTH-rp, N. F. B. Phillips (Case Western Reserve University) for assistance with chromatography and discussion, reviewers for insightful suggestions, and S. Price for assistance with the manuscript.

REFERENCES

1. Rotwein, P., Pollock, K. M., Didier, D. K., and Krivi, G. G. (1986) J. Biol. Chem. 261, 4828–4832
2. Dull, T. J., Gray, A., Hayflick, J. S., and Ullrich, A. (1984) Nature 310, 777–781
3. Hudson, P., Haley, J., John, M., Cronk, M., Crawford, R., Haralambidis, I., Tregear, G., Shine, J., and Niall, H. (1983) Nature 301, 628–631
4. Hudson, P., John, M., Crawford, R., Haralambidis, J., Scanlon, D., Gorman, J., Tregear, G., Shine, J., and Niall, H. (1984) EMBO J. 3, 2333–2339
5. Bathgate, R. A., Samuel, C. S., Burzyn, T. C., Layfield, S., Claass, A. A., Raytoms, I. G., Dawson, N. F., Zhao, C., Bond, C., Summers, R. I., Parry, L. J., Wade, J. D., and Tregear, G. W. (2002) J. Biol. Chem. 277, 1148–1157
6. Burkhardt, E., Adham, I. M., Hobohm, U., Murphy, D., Sander, C., and Engel, W. (1994) Hum. Genet. 94, 91–94
7. Chassin, D., Laurent, A., Janneau, J.-L., Berger, R., and Bellet, D. (1995) Genomics 29, 465–470
8. Conklin, D., Lofton-Day, C. E., Haldeman, B. A., Ching, A., Whitmore, T. E., Lok, S., and Jaspers, S. (1999) Genomics 60, 50–56
9. Hsu, S. Y. (1999) Mol. Endocrinol. 13, 2163–2174
10. Adams, T. E., Epa, V. C., Garrett, T. P., and Ward, C. W. (2000) Cell Mol. Life Sci. 57, 1050–1093
11. De Meyts, P., and Whittaker, J. (2002) Nat. Rev. Drug Discov. 1, 769–783
12. Katugampola, S., and Davenport, A. (2003) Trends Pharmacol. Sci. 24, 30–35
13. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., and Reynolds, C. D. (1988) Philos. Trans. R. Soc. Lond. Ser. B 319, 369–456
14. Miller, J. A., Narhi, L. O., Hua, Q. X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestrelski, S., Lauren, S., Stoney, K. S., Tsai, L., and Weiss, M. A. (1993) Biochemistry 32, 5203–5213
15. Narhi, L. O., Hua, Q. X., Arakawa, T., Fox, G. M., Tsai, L., Rosenfeld, R., Holst, P., Miller, J. A., and Weiss, M. A. (1993) Biochemistry 32, 5214–5221
16. Matagne, A., Radford, S. E., and Dobson, C. M. (1997) J. Mol. Biol. 267, 1068–1074
17. Yokota, A., Izutani, K., Takai, M., Kubo, Y., Noda, Y., Koumoto, Y., Tachibana, H., and Segawa, S. (2000) J. Mol. Biol. 295, 1275–1288
18. Pitsyn, O. B. (1991) FEBS Lett. 285, 176–181
19. Dill, K. A., and Chan, H. S. (1997) Nat. Struct. Biol. 4, 10–19
20. Lazaridis, T., and Karplus, M. (1997) Science 278, 1928–1931
21. Raleigh, D. P., and Plaxo, K. W. (2005) Protein Pept. Lett. 12, 117–122
22. Krishna, M. M., Hoang, L., Lin, Y., and Englander, S. W. (2004) Methods Mol. Biol. 34, 51–64
23. Creighton, T. E. (1992) Bioessays 14, 195–199
24. Creighton, T. E., Darby, N. J., and Kemmink, J. (1996) FEBS Lett. 110, 110–118
25. Oas, T. G., and Kim, P. S. (1988) Nature 336, 42–48
26. Dyson, H. J., and Wright, P. E. (1996) Annu. Rev. Phys. Chem. 47, 369–395
27. Baldwin, T. O., Ziegler, M. M., Chaffotte, A. F., and Goldberg, M. E. (1993) J. Biol. Chem. 268, 10766–10772
28. Zhang, Y., Hu, H., Cai, R., Feng, Y., Zhu, S., He, Q., Tang, Y., Xu, M., Xu, Y., Zhang, X., Liv, B., and Liang, Z. (1996) Sci. China C Life Sci. 39, 225–233
29. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. (1992) Biochemistry 31, 1749–1756
30. Qiao, Z. S., Guo, Z. Y., and Feng, Y. M. (2001) Biochemistry 40, 2662–2668
An Insulin-related Folding Intermediate

31. Hua, Q.-X., Chu, Y.-C., Jia, W., Phillips, N. F. B., Wang, R.-Y., Katsoyannis, P. G., and Weiss, M. A. (2002) J. Biol. Chem. 277, 43443–43453
32. Hua, Q. X., Jia, W., Frank, B. H., Phillips, N. F., and Weiss, M. A. (2002) Biochemistry 41, 14700–14715
33. Hua, Q. X., Gozani, S. N., Chance, R. E., Hoffmann, J. A., Frank, B. H., and Weiss, M. A. (1991) Nat. Struct. Biol. 2, 129–138
34. Milner, S. J., Carver, J. A., Ballard, F. J., and Francis, G. L. (1999) Biotechnol. Bioeng. 62, 693–703
35. Hua, Q. X., Narhi, L., Jia, W., Arakawa, T., Rosenfeld, R., Hawkins, N., Miller, J. A., and Weiss, M. A. (1996) J. Mol. Biol. 259, 297–313
36. Hua, Q.-X., Nakagawa, S. H., Jia, W., Hu, S. Q., Chu, Y.-C., Katsoyannis, P. G., and Weiss, M. A. (2001) Biochemistry 40, 12299–12311
37. Cooke, R. M., Harvey, T. S., and Campbell, I. D. (1991) Biochemistry 30, 5484–5491
38. Sato, A., Nishimura, S., Ohkubo, T., Kyogoku, Y., Koyama, S., Kobayashi, M., Yasuda, T., and Kobayashi, Y. (1993) Int. J. Pept. Protein Res. 41, 433–440
39. Hua, Q. X., Hu, S. Q., Frank, B. H., Jia, W., Chu, Y.-C., Wang, S. H., Burke, G. T., Katsoyannis, P. G., and Weiss, M. A. (1996) J. Mol. Biol. 264, 390–403
40. Weiss, M. A., Hua, Q.-X., Jia, W., Chu, Y.-C., Wang, R.-Y., and Katsoyannis, P. G. (2000) Biochemistry 39, 15429–15440
41. Tang, J. G., and Tsou, C. L. (1990) Biochem. J. 268, 429–435
42. Wang, C. C., and Tsou, C. L. (1991) Trends Biochem. Sci. 16, 279–281
43. Yan, H., Guo, Z. Y., Gong, X. W., Xi, D., and Feng, Y. M. (2003) Protein Sci. 12, 768–775
44. Merrifield, R. B., Vizioli, L. D., and Boman, H. G. (1982) Biochemistry 21, 5020–5031
45. Atherton, E., and Shepard, R. C. (1989) Solid Phase Peptide Synthesis: A Practical Approach, IRL Press, Oxford
46. Wang, S. S. (1973) J. Am. Chem. Soc. 93, 287–394
47. Vajdos, F. F., Ultsch, M., Schaffer, M. L., Deshayes, K. D., Liu, J., Skelton, N. J., and de Vos, A. M. (2001) Biochemistry 40, 11022–11029
48. Brzozowski, A. M., Dodson, E. J., Dodson, G. G., Murshudov, G. N., Verma, C., Turkenburg, J. P., de Bree, F. M., and Dauter, Z. (2002) Biochemistry 41, 9389–9397
49. Anfinsen, C. B. (1973) Science 181, 223–230
50. Creighton, T. E. (1997) Biocatal. Chem. 378, 731–744
51. Sali, A., Shaknhovich, E., and Karplus, M. (1994) Nature 369, 248–251
52. Bucciante, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., and Stefani, M. (2002) Nature 416, 507–511
53. Dobson, C. M. (1999) Trends Biochem. Sci. 24, 329–332
54. Onuchic, J. N., Luthey-Schulten, Z., and Wolynes, P. G. (1997) Annu. Rev. Phys. Chem. 48, 545–600
55. Baldwin, R. L. (1995) J. Biol. Chem. 270, 103–109
56. Rumbley, J., Hoang, L., Mayne, L., and Englander, S. W. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 105–112
57. Harding, H. P., and Ron, D. (2002) Diabetes 51, 5455–5461
58. Zinszner, H., Kuroda, M., Wang, X., Batchvarova, N., Lightfoot, R. T., Remotti, H., Stevens, J. L., and Ron, D. (1998) Genes Dev. 12, 982–995
59. Weissman, J. S., and Kim, P. S. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 9900–9904
60. Creighton, T. E. (1990) Biochem. J. 270, 1–16
61. Staley, J. P., and Kim, P. S. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 1519–1523
62. Jacoby, E., Hua, Q. X., Stern, A. S., Frank, B. H., and Weiss, M. A. (1996) J. Mol. Biol. 258, 136–157
63. Wuthrich, K. (1986) NMR of Proteins and Nucleic Acids, John Wiley & Sons, New York
64. Chen, Y., You, Y., Jin, R., Guo, Z. Y., and Feng, Y. M. (2004) Biochemistry 43, 9225–9233
65. Jansson, M., Andersson, G., Uhlen, M., Nilsson, B., and Kordel, J. (1998) J. Biol. Chem. 273, 24701–24707
66. Weiss, M. A., Nakagawa, S. H., Jia, W., Xu, B., Hua, Q. X., Chu, Y.-C., Wang, R. Y., and Katsoyannis, P. G. (2002) Biochemistry 41, 809–819
67. Huang, K., Xu, B., Hu, S. Q., Chu, Y. C., Hua, Q. X., Qu, Y., Li, B., Wang, S., Wang, R. Y., Nakagawa, S. H., Theede, A. M., Whittaker, J., De Meyts, P., Katsoyannis, P. G., and Weiss, M. A. (2004) J. Mol. Biol. 341, 529–550
68. Guo, Z.-Y., and Feng, Y.-M. (2001) Biochem. 382, 443–448
69. Hua, Q. X. (2004) Prog. Biochem. Biophys. 31, 1–16
70. Liu, M., Li, Y., Cavender, D., and Arvan, P. (2005) J. Biol. Chem. 280, 13209–13212
71. Markussen, J. (1985) Int. J. Pept. Protein Res. 25, 431–434
72. Ohnishi, S., Lee, A. L., Edgell, M. H., and Shortle, D. (2004) Biochemistry 43, 4064–4070
73. Jha, A. K., Colubri, A., Freed, K. F., and Sosnick, T. R. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 13099–13104
74. Bernado, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R. W., and Blackledge, M. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 17002–17007
75. Bertolotti, A., Zhang, Y., Hendershot, L. M., Harding, H. P., and Ron, D. (2001) Nat. Cell. Biol. 2, 326–332
76. Harding, H. P., Zeng, H., Zhang, Y., Jungries, R., Chung, P., Plesken, H., Sabatini, D. D., and Ron, D. (2001) Mol. Cell 7, 1153–1163