COMMUNICATION

Crystal Structure of GCN4-pI QI, a Trimeric Coiled Coil with Buried Polar Residues

Debra M. Eckert, Vladimir N. Malashkevich and Peter S. Kim*

Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research
Department of Biology
Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge MA 02142, USA

Coiled coils consist of two or more $\alpha$-helices wrapped around each other with a superhelical twist. The interfaces between helices of a coiled coil are formed by hydrophobic amino acid residues packed in a "knobs-into-holes" arrangement. Most naturally occurring coiled coils, however, also contain buried polar residues, as do the cores of the majority of naturally occurring globular proteins. Two common buried polar residues in both dimeric and trimeric coiled coils are asparagine and glutamine. In dimeric coiled coils, buried asparagine, but not glutamine, residues have been shown to confer specificity of oligomerization. We have placed a glutamine residue in the otherwise hydrophobic interior of a stable trimeric coiled coil, GCN4-pII, to study the effect of this buried polar residue in a trimeric coiled-coil environment. The resulting peptide, GCN4-pIQI, is a discrete, trimeric coiled coil with a lower stability than GCN4-pII. The crystal structure determined to 1.8 Å shows that GCN4-pIQI is a trimeric coiled coil with a chloride ion coordinated by one buried glutamine residue from each monomer.

#1998 Academic Press

Keywords: coiled coil; GCN4; buried polar residues; crystallography; ion binding

The coiled coil is a protein structural motif consisting of two to five $\alpha$-helices wrapped around each other with a superhelical twist (Cohen & Parry, 1990; Malashkevich et al., 1996). The fold of a coiled coil is determined by a simple pattern of amino acid residues, in which there is a characteristic heptad repeat designated by the letters a through g (McLachlan & Stewart, 1975). The first and fourth positions of the repeat, the a and d positions, respectively, form the interior of the interacting strands of the coiled coil and are generally hydrophobic (Hodges et al., 1972; McLachlan & Stewart, 1975; Parry, 1982; Lupas et al., 1991; Berger et al., 1995). Nevertheless, most coiled coils also contain polar residues at some of the a and d positions (Parry, 1982; Woolfson & Alber, 1995).

Interestingly, buried polar residues are also a common feature of naturally occurring globular proteins (Barlow & Thornton, 1983; Rashin & Honig, 1984), even though it is energetically unfavorable to sequester a polar residue in a hydrophobic protein core (Sauer & Lim, 1992; Harbury et al., 1993; Lumb & Kim, 1995; Hendsch et al., 1996).

Buried polar residues have been shown to contribute to the specificity of the protein fold at the expense of stability in the GCN4 coiled-coil model system (Harbury et al., 1993). GCN4-p1, a dimeric peptide model of the coiled-coil domain of the yeast transcription activator, GCN4 (O’Shea et al., 1989), contains an asparagine residue in an a position that forms hydrogen bonds with the corresponding asparagine residue in the other subunit of the homodimer (O’Shea et al., 1991). When this asparagine residue is replaced with valine, the coiled coil becomes more stable, but oligomerization specificity is lost, as both dimers and trimers are formed (Harbury et al., 1993).

In a designed heterodimeric coiled coil, “Peptide Velcro,” a buried asparagine residue promotes specificity of both helix orientation and oligomerization state. Peptide Velcro has a buried aspara-
when the asparagine residue of the dimeric GCN4-pI is changed to glutamine, oligomeric specificity is lost, and a mixture of dimers and trimers is formed (Gonzalez et al., 1996). Although glutamine, in the GCN4-p1 backbone, does not specify a discrete oligomeric state, there is a strong preference for glutamine residues in coiled-coil trimers; glutamine is found approximately three times more frequently than asparagine in the a position (Woolfson & Alber, 1995; Wolf et al., 1997). Because of its predominance in naturally occurring coiled-coil trimers, we investigated the effect of burying a glutamine residue in the hydrophobic core of a well-characterized, trimeric version of the GCN4 coiled coil.

The oligomeric state of a coiled coil is determined by both the hydrophobic and polar residues in the a and d positions. In GCN4-p1, the oligomeric state can be altered by substituting other hydrophobic residues at the a and d positions (Harbury et al., 1993). For example, when all the a and d positions are changed to isoleucine residues, the peptide folds into an extremely stable, discretely trimeric state (Harbury et al., 1993, 1994). This peptide, GCN4-pII, provides the backbone for the current studies. We inserted a glutamine residue into the interior of GCN4-pII at position 16, an a position. The resulting peptide, GCN4-pIQI, has seven isoleucine residues and one glutamine residue in the core. A helical wheel projection of GCN4-pIQI is shown in Figure 1.

As determined by circular dichroism, GCN4-pIQI has greater than 90% helix content in PBS (pH 7.0), 4°C. At a concentration of 10 μM, GCN4-pIQI undergoes a reversible, cooperative thermal unfolding transition with a midpoint of 68°C, as compared to a midpoint exceeding 100°C for GCN4-pII at the same concentration (Figure 2). Sedimentation equilibrium experiments indicate that GCN4-pIQI exists in a discretely trimeric state across a concentration range from 10 to 100 μM (Figure 3). Therefore, the introduction of buried glutamine residues into the core of the trimeric coiled coil does not alter the oligomeric state of the coiled coil but does cause a decrease in stability. In order to determine if there were any structural perturbations not detectable by the above experiments, GCN4-pIQI was crystallized, and its structure was determined to 1.8 Å resolution.

The crystals of GCN4-pIQI belong to the space group P321, and contain a monomer in the asymmetric unit, with the trimer formed around the crystallographic 3-fold axis. An initial model was determined by molecular replacement methods (Navaza, 1994) using GCN4-pII (Ile16Ala) as the search model. The identical solution was revealed when a poly-alanine version of GCN4-pII was used as the search model. The glutamine residue side-chain was added after the first round of refinement, as its electron density was apparent in a 2F_o-F_c map. The structure was refined to a crystallographic R-factor of 21.1% with an R_free of 23.6% (Bruenger, 1996) over a resolution range of...
The final model consists of 31 amino acid residues and incorporates 47 water molecules and a chloride ion, which is located on the crystallographic 3-fold axis. The model exhibits excellent geometry as determined by PROCHECK (Laskowski et al., 1993). Since molecular replacement methods can cause the final model to be biased by the search model, the final model was verified by simulated-annealing omit maps. The details of crystal growth, data collection, molecular replacement, and refinement are in the legend to Table 1.

GCN4-pIQI, like GCN4-pII, forms a trimeric, parallel coiled coil with a left-handed superhelical twist (Figure 4). The width of the trimer is 24 Å and the length is 47 Å. The isoleucine residues at the core positions show the characteristic “knobs-into-holes” packing (Crick, 1953), in the “acute” mode (Harbury et al., 1993), as seen in the crystal structure of GCN4-pII (Harbury et al., 1994). Interhelical electrostatic interactions are formed between the residues in the g position of one heptad (gₙ) and the residues in the e position of the following heptad in the neighboring strand of the coiled coil (eₙ₊₁), a trait common to coiled-coil structures. All three of the possible interhelical ionic interactions per monomer (between residues Arg1 and Glu6’, Lys15 and Glu20’, and Glu22 and Lys27’) are visible in the final model. The high B-factors, however, of the Lys15 N’ atom and Glu20 O’ atoms (~37 Å² as compared to an average B-factor for the model of 29 Å²) imply that these residues are flexible. Structural superposition of the final model of GCN4-pIQI and the crystal structure of GCN4-pII results in a root-mean-square deviation (r.m.s.d.) of 0.82 Å for Cα atoms, with the largest deviations occurring at the ends of the helices. The r.m.s.d. using only the Cα atoms of the two models without three residues on each end of each helix is 0.35 Å, indicating that the I16Q mutation has a very minor effect on the overall structure of the trimeric coiled coil.

Figure 2. Thermal denaturation experiments of GCN4-pIQI at 10 μM, monitored by circular dichroism, show that it undergoes a reversible cooperative unfolding transition upon heating, with a midpoint of 68°C. The midpoint for GCN4-pII at the same concentration exceeds 100°C. Experiments were performed on an Aviv 62A DS circular dichroism spectrometer. Measurements at 222 nm were performed on 10 μM solutions of peptides in PBS (pH 7.0) in a 10 mm path-length cuvette. Samples were heated from 4°C to 84°C, with equilibration times of 1.5 minutes and averaging times of 60 seconds. To obtain a 10 μM sample of the peptide, lyophilized peptide was resuspended first to approximately 500 μM in water. The concentration of the peptide stock was determined by tyrosine absorbance at 280 nm in 6 M GuHCl (Edelhoch, 1967), and the stock was diluted to 10 μM in PBS.

Figure 3. Sedimentation equilibrium studies of GCN4-pIQI indicate that it is trimeric across a concentration range from 10 to 100 μM. Representative data for 10 μM GCN4-pIQI are plotted as ln( absorbance ) at 231 nm against the square of the radius from the axis of rotation divided by 2. The slope is proportional to molecular mass. Continuous lines indicate the calculated lines for the indicated oligomeric states. The deviation in the data from the line calculated for the trimeric species is plotted in the upper panel. Sedimentation equilibrium studies were performed at 4°C on a Beckman XL-A analytical ultracentrifuge using an An-60 Ti rotor at speeds of 28,000 and 32,000 rpm and protein concentrations of 10, 20, 50, and 100 μM. The protein solutions were dialyzed overnight against PBS (pH 7.0) before spinning in Beckman 6-sector cells.
Table 1. Data processing and refinement statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P321</td>
</tr>
<tr>
<td>Cell dimensions (Å)</td>
<td>a = b = 59.66, c = 47.47</td>
</tr>
<tr>
<td>Cell angles (%)</td>
<td>α = β = 90, γ = 120</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>3.6</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>6.7</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>94.9</td>
</tr>
<tr>
<td>No. of reflections (working/free)</td>
<td>3652/376</td>
</tr>
<tr>
<td>Rcryst (20.0-1.8 Å) (%)</td>
<td>21.1</td>
</tr>
<tr>
<td>Rfree (20.0-1.8 Å) (%)</td>
<td>23.6</td>
</tr>
<tr>
<td>Average B-factor (Å)</td>
<td>29.0</td>
</tr>
<tr>
<td>r.m.s. standard deviations from target values</td>
<td>Bond length (Å) 0.011</td>
</tr>
<tr>
<td></td>
<td>Bond angles (%) 1.2</td>
</tr>
<tr>
<td></td>
<td>Dihedrals (%) 14.4</td>
</tr>
</tbody>
</table>

* Rmerge = ΣΣ |Ij| - |ß|Ij| / ΣΣIj|, where |Ij| is the intensity measurement for reflection j and |ß| is the mean intensity for multiply recorded reflections.
* Rcryst, Rfree = Σ|Iobs| - |Fcalc| / Σ|Iobs|, where the crystallographic and free B-factors are calculated using the working and test reflection sets, respectively. The test reflections included 10% of the total reflections which were chosen before refinement of the initial model and were not used during refinement.

GCN4-pII-J was crystallized at 20°C using vapor diffusion. A 10 mg/ml protein solution was mixed 1:1 with a reservoir solution containing 75 mM sodium acetate (pH 4.5), 20 mM MgCl₂, and 23% (v/v) methyl pentanediol (MPD) and equilibrated against the reservoir solution for several days. Most of the crystals appeared as long, thin needles and diffracted to only 3.0 Å. However, crystals grown using a macroseeded protocol (Thaller et al., 1985) diffracted to 1.8 Å. Crystals belong to the space group P321 with one monomer per asymmetric unit and a solvent content of 54%. A crystal with dimensions 0.3 mm x 0.2 mm x 0.2 mm was flash-frozen using an X-stream cryogenic crystal cooler ( Molecular Structure Corporation, MSC, The Woodlands, Texas). Diffraction data were collected using a Raxis IV detector mounted on a Rigaku RU-300 rotating anode generator (MSC). The diffraction data were indexed using DENZO and scaled using SCALEPACK (Otwinowski, 1993). Initial phases were determined by molecular replacement using GCN4-pII (Ile6Ala) as the search model in the program AMoRe (Navaza, 1994). Interestingly, using a trimer as a search model produced an unambiguous rotation/translation solution, whereas a search with a monomer did not. This observation is probably due to the non-globular shape of the molecule and the interacting atom of Gln 16 (Oe1) of the glutamine residue side-chain was then built into the electron density using O (Jones et al., 1991). Additional refinement steps included rounds of simulated annealing, positional refinement, and individual B-factor refinement, as well as model rebuilding using O. When the Rmerge reached approximately 35%, water molecules were added manually. Bulk solvent and anisotropic B-factor corrections were applied, followed by additional positional and individual B-factor refinement. The final structure was verified using simulated annealing omit maps, omitting one residue at a time and heating the structure to 3000 K before cooling. The B-factor for the final model is 21.1% with an Rmerge of 23.6%. According to PROCHECK (Laskowski et al., 1993), all bond lengths and angles are good, and all residues lie within the most preferred region of the Ramachandran plot for α-helices.

During refinement, a strong spherical peak of electron density appeared on the coiled-coil tri-mer axis near Gln16 (Figure 4b). This peak is modeled as a chloride ion, present in the crystallization buffer, for two reasons. First, modeling this peak as a water molecule resulted in an anomalously low individual B-factor value of 8 Å² as compared with a value of 31 Å² for the interacting Oe1 atom of Gln 16, when the Oe1 atom of the glutamine residue was facing into the center of the coiled coil, and an individual B-factor value of 2 Å² as compared to 19 Å² for the Nε2 atom, when the Nε2 atom is facing in. In contrast, the B-factor of the modeled chloride ion, 24.4 Å², is quite close to the 24.9 Å² B-factor of the interacting Nε2 atom of Gln16. Second, the distance between the initially modeled water molecule and the interacting atom of Gln 16 (Oe1 or Nε2) was too long (>3.2 Å) for optimal hydrogen bonding. The refined distance between the chloride ion and the Nε2 atom of the glutamine residue is 3.35 Å, compared to the sum of the van der Waals radii for NH₃⁺ and Cl⁻ of 3.24 Å (Weast, 1980). This distance is also consistent with NH₃⁺ to Cl⁻ distances previously observed crystallographically in coiled coils (Fass et al., 1996; Malashkevich et al., 1996). In the trimeric coiled coil from the ectodomain of Moloney murine leukemia virus, a chloride ion is coordinated by three asparagine residues (Fass et al., 1996).

The structure of the designed ABC heterotrimer (Nautiyal et al., 1995), which has a core similar to that in GCN4-pII-J, has recently been solved to 1.8 Å and also shows a chloride ion coordinated by the buried glutamine residues (S. Nautiyal & T. Alber, in press). Based on these results, it seems likely that anion binding is a common feature of trimeric coiled coils with buried polar residues. Nonetheless, several other crystal structures of trimeric coiled coil domains containing buried glutamine residues, including the HIV gp41 core (Chan et al., 1997; Weissenhorn et al., 1997b; Tan et al., 1997), the SIV gp41 core (Malashkevich et al., 1998), and a GCN4 variant (Gonzalez et al., 1996), do not indicate chloride binding.

GCN4-pII-J should serve as a valuable addition to the repertoire of coiled coils being used to assemble soluble domains of membrane proteins for structural and functional studies. There are many examples where dimeric coiled coils have been used for this purpose (Chang et al., 1994; Muir et al., 1994; Cochran & Kim, 1996; Kalandadze et al., 1996). For example, Peptide Velcro was used to replace the transmembrane domain of the T-cell receptor (TCR) in order to express soluble zβ₁ TCR extracellular domains (Chang et al., 1994). Because many of the membrane-fusion proteins of viruses are trimeric (Wilson et al., 1981; Allison et al., 1995; Fass & Kim, 1995; Lu et al., 1995; Blacklow et al., 1995; Weissenhorn et al., 1998), GCN4-pII-J will be a useful tool for constructing soluble domains aimed at understanding viral membrane fusion. A potential concern in fusing stable coiled coils to other pro-
tein domains is that the native structure of the soluble domain could be distorted by the attached coiled coil. Indeed, the conformation of the cytoplasmic signaling domain of the *Escherichia coli* aspartate receptor, and therefore the extent of activation of the soluble model of the receptor, can be altered by the length of the linker between that domain and the attached dimeric coiled coil (Cochran & Kim, 1996). GCN4-pII has already been utilized for constructing peptide models of trimeric membrane proteins (Weissenhorn et al., 1997a, 1998), but the lower stability of GCN4-pIQL makes it less likely to impose structure on the attached domain.

**Figure 4.** Crystal structure of GCN4-pIQL. a, A portion of the 20.0 to 1.8 Å $2F_o - F_c$ electron density map contoured at 1.5σ is superposed on the final model. The side view of the helix covers residues 16 to 23. b, Another region of the same map, superposed on the final model, showing the cross section of the trimer at residue 16. A chloride ion (green sphere) is located in the center of the trimer, on the crystallographic 3-fold axis, 3.35 Å from the Nε2 atom of Gln16. The Oδ1 atom of Gln16 forms a hydrogen bond with a water molecule (red sphere). c, A view of the GCN4-pIQL structure, looking down the helices from the N to the C terminus. All side-chains are displayed and the chloride ion is depicted as a yellow ball in the center of the trimer. GCN4-pIQL is a trimeric coiled coil with three right-handed α-helices wrapped around each other in a left-handed superhelical twist. d, A side, stereo view of GCN4-pIQL. Side-chains of the residues in the α and δ positions are displayed. The chloride ion is represented as a yellow ball. Figures a and b were generated with O (Jones et al., 1991), and Figures c and d were generated with InsightII (Biosym).
Protein Data Bank Accession Number

The coordinates for the GCN4-pI\textsubscript{QI} structure have been deposited in the Brookhaven Protein Data Bank with accession number 1piq, and are available immediately at the website http://www.wi.mit.edu/kim/home.html.

Acknowledgments

We thank Mike Burgess and James Pang for peptide synthesis and James Berger and Lawren Wu for technical advice. We also thank Deborah Fass for generosity of advice. We also thank Deborah Fass for generosity of time and encouragement. Thanks to members of the Kim laboratory for critically reviewing the manuscript. D.M.E. is a National Science Foundation Predoctoral Fellow. This research was supported by the National Institutes of Health (GM44162).

References


neoprotein structural model for the cytoplasmic domain of a multisubunit cell-surface receptor: integrin $\alpha_{IIb}\beta_3$ (platelet GPIIb-IIIa). Biochemistry, 33, 7701–7708.


Edited by P. E. Wright

(Received 17 June 1998; accepted 11 September 1998)