Crystal Structure of Human Pirin

AN IRON-BINDING NUCLEAR PROTEIN AND TRANSCRIPTION COFACTOR*

Received for publication, September 9, 2003, and in revised form, October 13, 2003
Published, JBC Papers in Press, October 22, 2003, DOI 10.1074/jbc.M310022200

Hai Pang‡§, Mark Bartlam‡§, Qinghong Zeng‡§, Hideyuki Miyatake‡, Tamao Hisano§, Kunio Miki‡, Luet-Lok Wong‡, George F. Gao‡**, and Zihe Rao‡ ‡‡

From the ‡Laboratory of Structural Biology, Tsinghua University and National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Science, School of Life Sciences and Engineering, Beijing, Beijing 100084, China, the *Theoretical Structural Biology Laboratory, RIKEN Harima Institute, SPiring-8, Byogo 679-5148, Japan, and the **Nuffield Department of Clinical Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DU, United Kingdom

Pirin is a newly identified nuclear protein that interacts with the oncoprotein B-cell lymphoma 3-encoded (Bcl-3) and nuclear factor I (NFI). The crystal structure of human Pirin at 2.1-Å resolution shows it to be a member of the functionally diverse cupin superfamily. The structure comprises two β-barrel domains, with an Fe(II) cofactor bound within the cavity of the N-terminal domain. These findings suggest an enzymatic role for Pirin, most likely in biological redox reactions involving oxygen, and provide compelling evidence that Pirin requires the participation of the metal ion for its interaction with Bcl-3 to co-regulate the NF-κB transcription pathway and the interaction with NFI in DNA replication. Substitution of iron by heavy metals thus provides a novel pathway for these metals to directly influence gene transcription. The structure suggests an interesting new role of iron in biology and that Pirin may be involved in novel mechanisms of gene regulation.

Pirin is a newly identified nuclear protein that is widely expressed in dot-like subnuclear structures in human tissues, in particular liver and heart (1). Pirins are highly conserved among mammals, plants, fungi, and even prokaryotic organisms and have been assigned as a sub-family of the cupin superfamily based on both structure and sequence homology (1, 2). The cupin superfamily is among the most functionally diverse of any protein family described to date, with both enzymatic and non-enzymatic members included (2). This cupin superfamily is grouped according to a conserved β-barrel fold and two characteristic sequence motifs. Study of Pirin reveals two cupin domains from its primary sequence that are consistent with other members of the superfamily.

The exact functions of Pirins are not yet known. No enzymatic activity has been described, but human Pirin has been found to bind to the nuclear factor I/CCAAT box transcription factor (NFI)† and to the oncoprotein B cell lymphoma 3-encoded (Bcl-3) in vitro (1, 3), suggesting that it is a transcription cofactor. NFI is known to stimulate DNA replication and RNA polymerase II-driven transcription (4). Bcl-3 is a distinctive member of the IxB family, which inhibits the transcription factor NF-κB by preventing NF-κB nuclear translocation and DNA binding. However, there is also evidence that Bcl-3 preferentially binds to NF-κB p50 or p52 dimers to stimulate transcription (5). The functional nature of this difference between IxB (inhibiting) and Bcl-3 (enhancing) is not known, but it is clear that they bind to different protein partners. Pirin is one of four binding partners of Bcl-3, together with Bard1, Tip60, and Jab1, and can be sequestered into quaternary complexes with Bcl-3, p50, and DNA (3). These four Bcl-3-interacting protein partners, which do not share any sequence homology, might play some crucial role in regulating the function of Bcl-3 and IxB. Indeed, both Pirin and Bcl-3 are localized in the nucleus, and the potential roles of Pirin in NF-κB-dependent transcriptional regulation are implicated in a number of experiments (6–10).

Here we report the crystal structure of human Pirin to 2.1-Å resolution. Understanding of the Pirin structure is of critical importance in elucidating and understanding its function, in particular its interaction with the IxB family of proteins in its role as a transcription cofactor.

**EXPERIMENTAL PROCEDURES**

Cloning, Expression, Purification, and Crystallization—The complete gene fragment encoding human Pirin protein was subcloned into PET-28a expression vector from a human liver cDNA library, and the human Pirin was highly expressed as a soluble protein in Escherichia coli strain BL21(DE3) with a 6-residue His tag attached to its N terminus. Purification of the Pirin protein was carried out through an affinity chromatography Co-NTA His Bind column (Qiagen) followed by size-exclusion chromatography on a Superdex 75 column (Amersham Biosciences). Crystals of Pirin were grown using the hanging drop vapor diffusion method from a solution containing 25 mg/ml Pirin in 14% polyethylene glycol 20000, 0.1 M MES, pH 6.5, at 16 °C. For phase determination, a selenomethionyl derivative was produced and crystallized under similar conditions. The Pirin crystals belong to the space group P2₁2₁2₁, with unit cell parameters a = 42.28, b = 67.12, c = 106.39 Å, α = β = γ = 90°.

Structure Determination—Data sets were collected at three wavelengths from a single selenomethionyl derivative crystal at 100 K on
The numbers in parentheses correspond to the highest resolution shell.

<table>
<thead>
<tr>
<th></th>
<th>MAD peak</th>
<th>MAD edge</th>
<th>MAD remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.9795</td>
<td>0.9799</td>
<td>0.9817</td>
</tr>
<tr>
<td>Resolution limit (Å)</td>
<td>50.0–2.1 (2.18–2.10)</td>
<td>50.0–2.1 (2.18–2.10)</td>
<td>50.0–2.1 (2.18–2.10)</td>
</tr>
<tr>
<td>Total reflections</td>
<td>95,485</td>
<td>95,271</td>
<td>94,836</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>30,975</td>
<td>30,927</td>
<td>30,894</td>
</tr>
<tr>
<td>Completeness</td>
<td>90.5 (81.5)</td>
<td>90.4 (81.4)</td>
<td>89.9 (79.8)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>5.0 (27.5)</td>
<td>5.0 (28.8)</td>
<td>5.0 (29.3)</td>
</tr>
<tr>
<td>l(1)</td>
<td>7.3 (1.9)</td>
<td>7.0 (1.8)</td>
<td>6.8 (1.8)</td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range</td>
<td>50.0–2.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rwork</td>
<td>20.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rfree</td>
<td>24.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of protein atoms</td>
<td>2,244</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of water molecules</td>
<td>187</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r.m.s. deviation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angles (°)</td>
<td>1.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average B factor (Å²)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>21.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>29.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favored (%)</td>
<td>88.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>11.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Rmerge = ΣΣ |I(b) - I(a)| / ΣΣ |I(a)|, where I(a) is the mean of the observations I(a) of reflection b.

**Structural Overview**—Pirin is composed of two structurally similar domains arranged face to face (Fig. 1, b and c). The core of each domain comprises two antiparallel β-sheets, with eight strands forming a β-sandwich. The fold of each domain is very similar, and the two domains can be superimposed with an r.m.s. difference of 1.3 Å for 64 equivalent residues. The N-terminal domain (residues 3–134) additionally features four ß-strands, and the C-terminal domain (residues 143–290) also includes four additional ß-strands and a short ß-helix. The two domains are cross-linked, with ß1 forming part of one sheet of the C-terminal domain, and strands ß25 and ß26 forming an extension of one sheet of the N-terminal domain. The C-terminal ß-helix packs against the outer surface of the N-terminal domain ß-barrel. The two domains are joined by a short linker of 10 amino acids (residues 134–143) that contains a single turn of a 3_10 helix. Four additional 3_10 helices are located in the structure.

Similar cavities are found in each domain. In the N-terminal domain, the cavity contains a metal binding site with a single Fe²⁺ ion located at about 6 Å from the protein surface (Fig. 1d). The C-terminal domain cavity is more compact than the N-terminal domain and is closed by the ß-strand formed by residues 6–12. The C-terminal domain does not contain any metal binding site.

**Pirin Is a Member of the Cupin Superfamily**—It has been predicted that Pirin belongs to the cupin superfamily on the basis of primary sequence (2). From a Dali (17) search for structural similarity, the Pirin structure was confirmed to closely resemble members of the cupin superfamily (Fig. 2, a–f), particularly the structures of quercetin 2,3-dioxygenase (18), glycinin g1 (19), and phosphomannose isomerase (20). As with Pirin, these three proteins are bicupins with two germin-like ß-barrel domains. Pirin also contains the two characteristic sequences of the cupin superfamily, namely PG-DX_y3-HXXH-(X)y3-E-DX_y3-G and G-(X)y3-PXG-(X)y3-H-(X)y3-N separated by a variable stretch of 15–50 amino acids (Fig. 2g). These motifs are best conserved in the N-terminal domain (residues 49–71 and residues 88–105) where the conserved histidine and glutamic acid residues correspond to the metal-coordinating residues. The C-terminal domain motifs (residues 178–199 and residues 216–230) lack the metal binding residues normally associated with the cupin fold.
The cupin superfamily has possibly the widest range of biochemical functions of any superfamily identified to date (2). It is comprised of both enzymatic and non-enzymatic members, which have either one or two cupin domains. The cupin fold comprises a motif of six to eight antiparallel β-strands located within a conserved β-barrel structure (2). The variety of biochemical functions is reflected by the low sequence homology shared among the cupin superfamily members (Fig. 2g). A BLAST search for proteins homologous to human Pirin revealed significant conservation between mammals, plants, and prokaryotes, particularly within the N-terminal domain (1). Interestingly, sequence alignment of human Pirin and related proteins reveals two clusters that are highly conserved throughout all aligned sequences. These sequence clusters, corresponding to residues 52–70 (cluster 1) and 88–106 (cluster 2), include the four metal coordinating residues of human Pirin, which are strictly conserved among all aligned sequences.

Pirin (Fig. 2a) and the bicupin metalloenzymes quercetin 2,3-dioxygenase (Fig. 2b) and phosphomannose isomerase (Fig. 2d) show similarities in the overall fold. Quercetin 2,3-dioxygenase is a copper-containing enzyme, and its N-terminal domain can be superimposed onto the N-terminal domain of Pirin with an r.m.s. difference of 1.5 Å for 84 equivalent residues. The Cu-binding site of quercetin 2,3-dioxygenase, formed by 3 histidines, a glutamic acid residue and a single water molecule, matches the metal site of Pirin. The three histidines of Pirin (His56, His58, and His101) are structurally equivalent to those of quercetin 2,3-dioxygenase (His66, His68, and His112). Similarly, the phosphomannose isomerase structure can be superimposed onto that of Pirin with an r.m.s. difference of 1.3 Å for 71 equivalent residues. Phosphomannose isomerase is a zinc-containing enzyme, and its metal binding site, comprising two histidines and two glutamic acid residues, also matches the metal site of Pirin.

Regardless of the overall structural similarity between its two domains, the C-terminal domain of Pirin shows interesting differences from the N-terminal domain. Notably, the C-terminal domain of Pirin does not contain a metal binding site and its sequence does not contain the conserved metal-coordinating residues. Several members of the cupin superfamily do not contain the metal-coordinating residues, including the transcription factor araC from Escherichia coli (21). araC is a single
**FIG. 2.** A comparison between Pirin and similar structures. **a**, Pirin structure; **b**, quercetin 2,3-dioxygenase structure (PDB code: 1JUH); **c**, glycinin g1 structure (PDB code: 1FXZ); **d**, phosphomannose isomerase structure (PDB code: 1PMI); **e**, oxalate oxidase (germin) structure (PDB code: 1FI2); **f**, araC dimer structure (PDB code: 2ARC). **g**, structure-based sequence alignment of Pirin with related structures. **PIRIN-N**, N-terminal domain of Pirin; **PIRIN-C**, C-terminal domain of Pirin; **2,3QD-N**, N-terminal domain of quercetin 2,3-dioxygenase; **2,3QD-C**, C-terminal...
domain member of the cupin family and binds arabinose for activation of transcription. There are notable similarities between the C-terminal domain of Pirin and araC. In araC, the arabinose binding site is located within the β-barrel and is closed by an N-terminal domain arm. The β-barrel of the C-terminal domain of Pirin is also closed by the N-terminal strand. However, in the absence of arabinose, the N-terminal arm of araC becomes disordered, suggesting that it is important for ligand stability. A detailed analysis of the arabinose binding residues in the araC structure shows that the C-terminal domain of Pirin is unlikely to bind arabinose.

The Metal Binding Site—Surprisingly, a single Fe²⁺ ion is located in the Pirin N-terminal domain where it is coordinated by 3 histidine residues (His⁵⁶, His⁵₈, and His¹⁰₁) through their Ne₂ atoms, and one glutamic acid (Glu¹⁰₃) through the Oe₂ atom (Fig. 3a). The metal binding site is exposed to the solvent, and the octahedral coordination environment is completed by two water molecules, with metal-ligand distances of 2.24 and 2.15 Å (Table II). A structure-based sequence alignment shows that the metal binding motif is highly conserved within a number of other cupin superfamily members (Fig. 2g). The metal binding domain of Pirin is most structurally similar to that in germin, a Mn²⁺-containing oxalate oxidase in which the metal is also octahedrally coordinated by three histidines, a glutamic acid and two water molecules (22). As with quercetin 2,3-dioxygenase, the three histidines in oxalate oxidase (His⁸₈, His⁹₀, and His¹₃₇) are structurally equivalent to those in Pirin (His⁵₆, His⁵₈, and His¹⁰₁). One notable difference between Pirin and related metalloproteins of the cupin superfamily is the location of the coordinating Glu residue. Quercetin 2,3-dioxygenase (Fig. 3b), phosphomannose isomerase (Fig. 3c), and oxalate oxidase (Fig. 3d) all contain a conserved Glu residue in the equivalent structural position to residue 63 of Pirin. However, this residue is not conserved in Pirin and the coordinating Glu residue is instead located in position 103 on β-strand 10.

The structure of a nickel-binding dioxygenase, acireductone

domain of quercetin 2,3-dioxygenase; PROGLYCININ, glycinin g1; PMI-N, N-terminal domain of phosphomannose isomerase; PMI-C, C-terminal domain of phosphomannose isomerase; OXALATE, oxalate oxidase (germin); ARAC, araC. Only the regions adjacent to the two characteristic conserved cupin sequences are shown. The two conserved cupin sequence motifs are shaded in gray, and metal coordinating residues are shown in green boxes. The secondary structure elements relate to the Pirin N-terminal domain structure.
dioxxygenase (Ni-ARD) from *Klebsiella pneumoniae*, was recently determined by NMR and its active site modeled by comparative homology modeling (23). It was suggested that ARD might also be a member of the cupin superfamily. The Ni$^{2+}$ ion in ARD was predicted to be coordinated by six ligands, namely three histidine residues, a glutamic acid, and two water molecules. The spatial arrangement of ligand groups in the Ni-ARD model is similar to that of Pirin and germin, with an average distance of 2.1 Å between the metal and donor atoms.

**Potential Functions of Pirin**—The bound metal ion may play an important role in Pirin function by stabilizing the N-terminal cupin domain structure and/or by imparting enzymatic activity to the protein. The iron found in the structure is the metal cofactor in numerous enzymes spanning a wide spectrum of activities.

**Metal-dependent Transcriptional Regulation**—Pirins are putative transcription cofactors. Heavy metals are known to play an important role in the transcriptional regulation of eukaryotic and prokaryotic genes (24–29). Pirin is newly identified in this study to be a metal-binding protein, and, interestingly, the metal-binding residues of Pirins are highly conserved across mammals, plants, fungi, and prokaryotic organisms. Pirin acts as a cofactor for the transcription factor NFI, the regulatory protein in NF-κB, that binds to the transcription factor NFI, the regulatory protein in NF-κB, that binds to the transcription factor NF-κB DNA complex. Metals have been implicated directly or indirectly in the NF-κB/p50 homodimer is shown as a ribbon model colored orange and the other colored yellow.

**TABLE II**

<table>
<thead>
<tr>
<th>Metal-donor atom distances and temperature factors</th>
<th>Distance to metal ion</th>
<th>Temperature factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe$^{2+}$</td>
<td>2.07</td>
<td>17.5</td>
</tr>
<tr>
<td>His$^{2+}$/Ne2</td>
<td>2.12</td>
<td>10.8</td>
</tr>
<tr>
<td>His$^{2+}$/Ne2</td>
<td>2.32</td>
<td>18.1</td>
</tr>
<tr>
<td>Glu$^{2+}$/Oe2</td>
<td>2.28</td>
<td>12.9</td>
</tr>
<tr>
<td>Water 1</td>
<td>2.24</td>
<td>13.8</td>
</tr>
<tr>
<td>Water 2</td>
<td>2.15</td>
<td>19.1</td>
</tr>
</tbody>
</table>

**FIG. 4. A model of the Pirin-Bcl-3(p50)$_2$ complex (A and B).** The ankyrin repeat domain of Bcl-3 is shown as a ribbon model colored red. The Pirin structure is shown as a ribbon model with the N-terminal domain colored blue and the C-terminal domain colored green. The (p50)$_2$ homodimer is shown as a ribbon model with one p50 molecule colored orange and the other colored yellow.
quired for interaction with Bcl-3, and the only common properties between them is that they are nuclear proteins and associate with gene regulators.

The structure of the Bcl-3 ankyrin repeat domain (ARD) is elongated and is comprised of seven ankyrin (ANK) repeats (36). Its central ankyrin repeats are similar to those of IκBα, with the key difference that Bcl-3 has a seventh ankyrin repeat at the C terminus in place of the PEST region of IκBα. Mutagenesis studies of Bcl-3 indicated that all seven ankyrin repeats are required for binary interaction with Pirin, whereas interactions with Jαb1 or Bard1 require only five repeats (3). Pirin and Bcl-3 can also be sequestered into quaternary complexes with p50 and DNA, and Pirin is known to increase the DNA binding by Bcl-3/p50. As with Pirin, all seven ankyrin repeats are required for the Bcl-3/p50 interaction (3). The stoichiometry of the complex is not known, but Bcl-3 is able to recognize at least two proteins simultaneously (3).

Because the structures of IκBα/NF-κB (p65/p50 heterodimer) (37, 38) complex and NF-κB p50 homodimer (39) are known, and to understand the interaction between Birin, Bcl-3, and p50, we modeled the complex of these two proteins using protein-protein docking techniques (Fig. 4, A and B). A Bcl-3/p50α2 complex was constructed using the structures of the IκBα-NF-κB complex and the ankyrin repeat domain (ARD) of Bcl-3. The Bcl-3 ARD structure and the (p50)α2 homodimer were superimposed onto the IκBα-NF-κB complex. Pirin was then docked to the Bcl-3/p50α2 complex using the program FTDOCK (40). The surface of Pirin has a large acidic patch on the N-terminal domain surface formed by residues 77–82, 97–103, and 124–128. We suggest that this acidic patch could interact with the large basic patch on ankyrin repeats 6 and 7 of Bcl-3. Additional contacts to ankyrin repeat 5 of Bcl-3 could be formed by residues 160–166 of Pirin. Notably, the previous observation that Pirin binding to Bcl-3 requires all seven repeats of the ARD was based on C-terminal ARD deletion (ΔARD6–7 or ΔARD5–7) mutants, therefore, ARD4–7 of Bcl-3 might be the only domain interacting with Pirin (3), which is consistent with our model. Further analysis involving N-terminal ARD deletion mutants should provide further insight into the Pirin-Bcl-3 recognition interactions.

In our model of the complex, residues 265–284 of the C-terminal arm of Pirin could make contacts with one of the p50 molecules. One possibility is that Bcl-3 binding to the (p50)α2 homodimer could induce conformational changes that lead to Pirin binding to p50, although gel retardation experiments showed that there is no direct binding between Pirin and p50 (3). Although we emphasize that this model is only hypotheti-
cal, the extensive contacts formed by Pirin with Bcl-3 and p50 represent the first crystal structure of the Pirin family and is an important step toward understanding the function of these proteins. Our work shows that human Pirin is a monomer with two cross-linked β-barrel domains. Pirin is confirmed to be a member of the cupin superfamily on the basis of primary sequence and structural similarity. The presence of a metal binding site in the N-terminal β-barrel of Pirin, which is highly conserved among Pirins, may be significant in its role in regulating NF-κB DNA replication and NF-κB transcription factor activity. Pirin was surprisingly found to be bound to Pirin and could impart enzymatic activity to the protein. Substitution of iron by other heavy metals also offers a direct mechanism for heavy metals to influence gene transcription. This novel pathway may be relevant in the toxicity and other effects of these metals. Finally, the structure of Pirin strongly suggests that a new role of iron in biology, namely regulating DNA replication and gene transcription at the level of the DNA complexes, should be added to the list of the many vital functions of this metal in living organisms.

Acknowledgments—We are grateful to Feng Gao, Fei Sun, and Zhiyong Lou from the Rao laboratory for assistance, and to Rong Zhang and Andrzej Joachimiak from the Advanced Photon Source (Argonne National Laboratory) for help with data collection. We thank Zhi Xing of the Tsinghua Analysis Center for help with atomic absorption spectrophotometry. Critical discussions with Neil Isaacs (Glasgow University, UK) and David Stuart (Oxford University, UK) are greatly appreciated.

REFERENCES
29. Laang, R., Igarashi, H., Tsuzuki, T., Nakabeppu, Y., Sekiguchi, M., Kasprzak,
Crystal Structure of Human Pirin