1.88 Å crystal structure of the C domain of hCyP33: A novel domain of peptidyl-prolyl cis–trans isomerase

Tao Wang, Cai-Hong Yun, Shen-Yan Gu, Wen-Rui Chang, Dong-Cai Liang

A National Key Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang District, Beijing 100101, PR China
b Graduate School of the Chinese Academy of Sciences, PR China

Received 25 May 2005
Available online 13 June 2005

Abstract

Cyclophilins (CyPs) are a widespread protein family in living organisms and possess the activity of peptidyl-prolyl cis–trans isomerase (PPIase), which is inhibited by cyclosporin A (CsA). The human nuclear cyclophilin (hCyP33) is the first protein which was found to contain two RNA binding domains at the amino-terminus and a PPIase domain at the carboxyl-terminus. We isolated the hCyP33 gene from the human hematopoietic stem/progenitor cells and expressed it in Escherichia coli, and determined the crystal structure of the C domain of hCyP33 at 1.88 Å resolution. The core structure is a β-barrel covered by two α-helices. Superposition of the structure of the C domain of hCyP33 with the structure of CypA suggests that the C domain contains PPIase active site which binds to CsA. Furthermore, C domain seems to be able to bind with the Gag-encoded capsid (CA) of HIV-1 and may affect the viral replication of HIV-1. A key residue of the active site is changed from Ala-103-CypA to Ser-239-hCyP33, which may affect the PPIase domain/substrates interactions.

Keywords: Human nuclear cyclophilin; Cyclophilin A; Peptidyl-prolyl cis–trans isomerase; Cyclosporin A

Among peptide bonds in proteins, rotation of prolyl imide bond is the slowest and is thought to be the rate-limiting step in the folding process of proteins. An enzyme which catalyses this slow rotation is called peptidyl prolyl cis–trans isomerase (PPIase) [1,2]. There are three PPIases families: cyclophilin (CyP), FK506-binding protein (FKBP), and parvulin (Pvn) [1,2]. Cyclophilins (CyPs) bind to the immunosuppressive drug cyclosporine A (CsA) and are widespread in prokaryotes and eukaryotes at different subcellular locations [3,4]. CyPs play important roles in living organisms and possess the PPIase activity, which accelerate protein folding [1,2,5,6], and CyP–CsA complexes act as inhibitors for signal transduction pathways in the immunological reactions [7,8]. CypA is one of the well-studied cyclophilins which possess PPIase activity and is inhibited by CsA [1–4]. CypA performs an essential role in HIV-1 replication and is affected by CsA. The formation of CA–CypA complex can be competitively inhibited by CsA’s binding to the active site of CypA [9–12]. Although some of the functions of CypA have been largely known, the mechanisms of these functions are still elusive.

Human nuclear cyclophilin (hCyP33) (GenBank Accession No.: NM_006112) contains two RNA-binding domains at the amino-terminus and a PPIase domain at the carboxyl-terminus (C domain). The C domain of hCyP33 is homologous to CypA in amino acid sequence [13]. We isolated the hCyP33 gene from the hematopoietic stem/progenitor cells and expressed it in Escherichia coli. We also purified and crystallized hCyP33 protein. Molecular weight of the protein in
Materials and methods

Cloning, expression, and purification. Cloning of the gene encoding hCyP33 was carried out by the Shanghai Institute of Hematology. The hCyP33 gene was cloned into pET22b (+) (Novagen) between the restriction sites NdeI and XhoI, but it was expressed as an insoluble protein. We, then, cloned this gene into PQE30 (Qiagen) between the restriction sites BamHI and HindIII, and got soluble expression of it. To the protein expressed by the vector PQE30-hCyP33, added a His6-tag to the N-terminus. This construction was confirmed by DNA sequencing. hCyP33 was expressed in E. coli strain M15 (Qiagen) after induction with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 6 h at 30 °C. The cells were harvested by centrifugation at 5000 rpm for 10 min and were lysed by ultrasonic cell disrupter. The soluble fraction was loaded on a 5 ml Chelating Sepharose Fast Flow resin column preloaded with NiSO4 and pre-equilibrated in 50 mM Tris–HCl (pH 8.0), 500 mM NaCl, 20 mM imidazole, and 2 mM β-mercaptoethanol. After three washes with increasing concentrations of imidazole the protein was eluted with 150 mM imidazole and pre-equilibrated in 50 mM Tris–HCl (pH 8.0), 500 mM NaCl, 20 mM imidazole, and 2 mM β-mercaptoethanol. After three washes with increasing concentrations of imidazole the protein was eluted with 150 mM imidazole. Proteins eluted from the resin were further purified by ion-exchange chromatography on Resource Q, hydrophobic interaction chromatography on Resource ISO, and size exclusion chromatography on Superdex 200. The molecular weight of the purified hCyP33 protein is about 34 kDa measured by SDS–PAGE.

Crystalization and data collection. The crystals were obtained by the hanging-drop, vapor diffusion method. Screening experiments were performed with commercial Screen kits from Hampton Research, microcrystals were obtained in 0.1 M Hepes (pH 7.5) and 25% w/v PEG 3350. The sample of hCyP33 was presented at a concentration of about 15 mg/ml in a storage buffer consisting of 10 mM Tris–HCl (pH 8.0), 10 mM NaCl, and 2 mM β-mercaptoethanol. The well solution contains 0.1 M Hepes (pH 7.5) and 20% (w/v) PEG 3350. We got crystals by mixing protein solution and well solution in 1:1 ratio at 20 °C within 2 weeks. Molecular weight of the protein in the crystal is 18.2 kDa measured by MALDI-TOF-MS.

Diffraction data were collected at Mar345 with seal-tube X-ray source in the National Key Laboratory of Biomacromolecules, Institute of Biophysics, CAS. Diffraction data were scaled and integrated with DENZO and SCALEPACK [14]. The results showed that the crystal belongs to space group P42121, with one molecule in the crystallographic asymmetric unit. The diffraction data statistics are summarized in Table 1.

Structure determination and refinement. The structure was solved by molecular replacement with MolRep [15] in the CCP4 suite [16]. The search model was the structure of CypA (PDB entry 1AWV), which displays 67% sequence similarity to the C domain of hCyP33. The initial model was manually rebuilt with the program O [17] and refined by AR/wARP [18] and refmac5 [19]. After several cycles of refinement, final model was obtained. A summary of the refinement statistics is given in Table 1.

Results and discussion

Overall structure of the C domain of hCyP33

The PPIase domain of hCyP33 has 165 residues. The model displays the structure of hCyP33 from Asn-139 to Tyr-300 and is shown in Fig. 1. In this model, two amino-acid residues at the N-terminus and one at the C-terminus are not included for their poor electron density. The core structure is a β-barrel similar to those of CyPs. The eight β-strands are arranged in an anti-parallel fashion around the barrel surface and two α-helices sit on the top and the bottom of the barrel, respectively. The β-turns and loops in the strands are probably responsible for the flexibility of the barrel. This β-barrel motif

Table 1

<table>
<thead>
<tr>
<th>Crystal data, data collection, and refinement statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Unit cell parameters (Å)</td>
</tr>
<tr>
<td>a</td>
</tr>
<tr>
<td>b</td>
</tr>
<tr>
<td>c</td>
</tr>
<tr>
<td>Molecules (a.u.)</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
</tr>
<tr>
<td>Unique reflections</td>
</tr>
<tr>
<td>Average redundancy</td>
</tr>
<tr>
<td>Completeness</td>
</tr>
<tr>
<td>Rmerge (%)</td>
</tr>
<tr>
<td>I/σ(I)</td>
</tr>
<tr>
<td>Non-hydrogen atoms</td>
</tr>
<tr>
<td>Amino-acid residues</td>
</tr>
<tr>
<td>Water molecules</td>
</tr>
<tr>
<td>Rwork (%)</td>
</tr>
<tr>
<td>Rfree (%)</td>
</tr>
<tr>
<td>Average B-values (Å²) for main chain/side chain/water</td>
</tr>
</tbody>
</table>

Fig. 1. The stereo view of the structure of the C domain of hCyP33. Red: N-terminus of the PPIase domain of hCyP33; blue: C-terminus of the PPIase domain of hCyP33. The core structure is displayed as a β-barrel similar to those of CyPs. The eight β-strands are arranged in an anti-parallel fashion around the barrel surface and two α-helices sit on the top and the bottom of the barrel, respectively. The β-turns and loops in the strands are probably responsible for the flexibility of the barrel. This β-barrel motif
is highly conserved in CyPs and has a conserved hydrophobic core [20]. The conformation of the PPIase domain of hCyP33 is very similar to CypA (the r.m.s.d. of whole structure is 0.58 Å). The C domain of hCyP33 is highly conserved in key residues among the CyPs family. In the structure of the C domain of hCyP33, the loop of Gly-149-Asn-150-Lys-151-Pro-152 (compared with CypA, the r.m.s.d. is 1.96 Å) and the loop of Tyr-215-Gly-216-Lys-217-Lys-218 (compared with CypA, the r.m.s.d. is 1.5 Å) have remarkable difference with CypA (Fig. 2).

The PPIase active sites of the C domain of hCyP33

The PPIase domain of hCyP33 possesses the activity of peptidyl-prolyl cis–trans isomerase (PPIase) [13]. The PPIase active sites of the C domain of hCyP33 can be predicted by comparing with CypA in amino acid sequence and structure. Residues Arg-191, Phe-196, Met-197, Gln-199, Ala-237, Asp-238, Phe-249, Leu-258, and His-262 of the C domain are involved in hydrogen bonds formation or van der Wall interactions with its substrates. The side chains of Phe-196, Met-197, Phe-249, Leu-258, and His-262 in the C domain make up the hydrophobic pocket (Fig. 3), to which the proline residue of the substrates bind [21]. The residues Arg191 and Asn238 form hydrogen bonds with proline residue in the substrates. The underlying mechanism is unclear, and the structure of the C domain–substrate complex will give some insights into it.

The PPIase activity of the C domain of hCyP33 can be inhibited by CsA [13], and the CypA-CsA complex acts as inhibitor in the immunological reactions by binding to protein phosphatase calcineulin [7,8]. The superposition of the C domain and CypA-CsA complex suggests that the CsA-binding sites of the C domain are in fact its PPIase activity site. Those residues contributing to the CsA binding are similar to CypA [22–24], including Arg-191, Phe-196, Met-197, Gly-199, Gly-208, Ala-237, Asn-238, Ser-239, Gln-247, Phe-249, Trp-257, Leu-258, and His-262. The mutation of Trp-121 of CypA leads to dramatically decreased sensitivity to CsA [25]. So we predict that the Trp-257 of the C domain of hCyP33 plays a similar role in recognizing CsA.

Analysis of the key residues and the binding site

Arg-55, Asn-102, and Ala-103 of CypA are the key residues for the interactions with the substrates and these interactions are affected by the neighboring residues [26]. Alignment of the amino acid sequence of the C domain of hCyP33 with CypA indicates that two of the three key residues (Arg and Asn) are the same, but Ala-103 in CypA is changed to Ser-239 in hCyP33 (Fig. 4). Comparing with the structure of CypA, Arg-191, and Asn-238 of the C domain sit in the same sites and adopt similar conformations. We predict that these three residues of the C domain play an important role as in CypA in the interactions with the substrates, and these interactions may also be affected by the neighboring residues. The movements of the neighboring

---

T. Wang et al. / Biochemical and Biophysical Research Communications 333 (2005) 845–849

847
backbone will affect the enzymatic reaction. So the displacement of Ala with Ser may alter the PPIase/substrates interactions and influence the nature of the key residue Asn-238. The conformational fluctuation of this region where the three key residues lie may change the C domain/substrates interactions and therefore affect the placement of Ala with Ser may alter the PPIase/substrate interactions and influence the nature of the key residues for the hydrogen bond formation with the substrates. The Ala-103 in CypA is changed to Ser-239 in hCyP33. Asn-238 and Ser-239 are located in a loop region. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) Nos. 30130080 and 30200046; 973 Project Nos. 2004CB520801 and 2002CB713801; 863 Program Nos. 2002BA711A13, KSCX1-SW-17, and KJCX2-SW-N06 from the Ministry of Science and Technology of China (MOST) as well as grants from CAS.

References