Solution structure of the ubiquitin-like domain of human DC-UbP from dendritic cells

YONG-GUANG GAO,1,3,5 AI-XIN SONG,1,5 YAN-HONG SHI,2,3 YONG-GANG CHANG,1,3 SHU-XUN LIU,4 YI-ZI YU,4 XUE-TAO CAO,4 DONG-HAI LIN,2 AND HONG-YU HU1

1Key Laboratory of Proteomics, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Shanghai 200031, People’s Republic of China
2Institute of Materia Medica, Shanghai Institutes for Biological Sciences, Shanghai 201203, People’s Republic of China
3Graduate School of the Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 201203, People’s Republic of China
4Institute of Immunology, Secondary Military Medical University, Shanghai 200433, People’s Republic of China

(RECEIVED March 11, 2005; FINAL REVISION April 26, 2005; ACCEPTED April 28, 2005)

Abstract

The previously identified dendritic cell-derived ubiquitin-like protein (DC-UbP) was implicated in cellular differentiation and apoptosis. Sequence alignment suggested that it contains a ubiquitin-like (UbL) domain in the C terminus. Here, we present the solution NMR structure and backbone dynamics of the UbL domain of DC-UbP. The overall structure of the domain is very similar to that of Ub despite low similarity (<30%) in amino-acid sequence. One distinct feature of the domain structure is its highly positively charged surface that is different from the corresponding surfaces of the well-known UbL modifiers, Ub, NEDD8, and SUMO-1. The key amino-acid residues responsible for guiding polyubiquitinated proteins to proteasome degradation in Ub are not conserved in the UbL domain. This implies that the UbL domain of DC-UbP may have its own specific interaction partners with other yet unknown cellular functions related to the Ub pathway.

Keywords: solution structure; dynamics; ubiquitin-like domain; DC-UbP; NMR

Supplemental material: see www.proteinscience.org

Ubiquitin (Ub) is a small polypeptide (76 residues) that covalently modifies other cellular proteins and targets them to the proteasome for degradation (Hershko and Ciechanover 1998; Finley 2001; Pickart 2004). Ubiquitination is also a versatile regulatory signal controlling protein activity, function, stability, and intracellular localization (Finley et al. 2004). The ubiquitination process involves Ub-activating enzyme (E1), Ub-conjugating enzyme (E2), Ub-ligase (E3), and the Ub-tagged substrates that are recognized and degraded by the 26S proteasome (Hershko and Ciechanover 1998; Pickart 2001). It has now emerged that UbL proteins are involved in a variety of fundamental cellular processes, such as protein sorting, protein degradation, DNA repair, cell division, autophagy, and apoptosis (Jentsch and Pyrowolakis 2000). UbL proteins can be divided into two separate classes (Jentsch and Pyrowolakis 2000). The class I UbLs function as modifiers in a manner analogous to...
that of Ub (Schwartz and Hoschstrasser 2003). UBLs exist either in a free form or attached covalently to other proteins by their C termini. The conjugation pathways for UBLs closely resemble that for Ub. The class II UDPs contains a domain homologous to Ub as part of an often large, multidomain overall structure. In contrast to UBLs, UDPs are not conjugated to other proteins but implicated in interaction with other regulatory proteins (Jentsch and Pyrowolakis 2000).

Recently, studies about UBLs mainly focus on SUMO, NEDD8, Apg12, and UCRP/ISG15. Post-translational modification with the UbL protein SUMO is involved in transcriptional regulation and genome surveillance (Muller et al. 2004). NEDD8 modification is involved in cell cycle control and in embryogenesis by up-regulating the activities of cullin-based E3 ligases (Pan et al. 2004). Covalent attachment of Apg12 to Apg5 is essential for autophagy (Mizushima et al. 1998, 2003). ISG15 could be conjugated to a variety of proteins when cells are treated with type I interferon or lipopolysaccharide and may be associated with specialized functions in the innate immune system (Kim and Zhang 2003).

UDPs include a structurally and functionally heterogeneous group of proteins (Jentsch and Pyrowolakis 2000). Rad23 is involved in DNA repair and protein degradation. The UbL domain of Rad23 possibly mediates the binding of Rad23 to proteasomes (Schauber et al. 1998). Dsk2, another UDP, competes with Rad23 for proteasome binding through its UbL domain (Elsasser et al. 2002). Ubiquilin may be involved in endoplasmic reticulum stress-associated protein degradation through association of its UbL domain with the 19S proteasome subunits (Ko et al. 2004). UDPs are also involved in the Ub system. For example, UbP6 is a deubiquitinating enzyme (Wyndham et al. 1999), and Elongin B is a subunit of the CBC Ub ligase (Kamura et al. 2002). In addition to its UbL domain, parkin possesses a RBR domain composed of two RING fingers plus an IBR/DRIL domain, which may be part of cullin-containing Ub ligase complexes (Marin et al. 2004).

In the previous study, a cDNA encoding a UbL protein was identified from dendritic cells (namely, DC-Ubp) (Liu et al. 2003). The UbL domain of DC-Ubp shares 28.6% identity and 55% similarity to Ub in the amino-acid sequence. In the present paper, we report the three-dimensional solution structure and backbone dynamics of the UbL domain of DC-Ubp solved by NMR spectroscopy. Although the UbL domain and Ub share low sequence similarity, we find that the two proteins take a similar fold. The structure may provide clues for identification of its interaction partners and further exploration of the biological functions of DC-Ubp protein.

Results

Structure determination

DC-Ubp protein contains a UbL domain (residues 85–190, numbered 1–106) in the C-terminal region. The sequence alignment of the UbL domain of DC-Ubp with Ub, SUMO-1, and NEDD8 is shown in Figure 1. Circular dichroism (CD) study revealed that the UbL domain is compactly folded with possibly α/β structures (data not shown). We then used heteronuclear multidimensional NMR techniques to elucidate the three-dimensional structure of the UbL domain in solution.

A summary of the NMR data used for structure calculation is presented in Table 1. All 15 final structures converged with no NOE or dihedral angle violation > 0.3 Å or 5°, respectively. The average RMSD for the 15 structures within the secondary structure elements of residues Cys24–Val94 (including residues Cys24–Leu30, Lys34–Val40, Val46–Ala56, Arg65–Phe68, Arg71–Pro72, Lys78–Glu81, Val89–Ile93) for

![Figure 1. Sequence alignment of the UbL domain of DC-Ubp, Ub, NEDD8, and SUMO-1. All the sequences are originated from Homo sapiens. The numerals are labeled according to the sequence of the UbL domain and the secondary structures are displayed based on Ub. T and η denote hydrogen-bonded turn and 3_10 helix, respectively. The figures were prepared using the program ESPript (Gouet et al. 1999).](image-url)
the backbone atoms and for all heavy atoms are 0.56 and 1.62 Å, respectively. Figure 2 depicts a superimposition of the 15 lowest energy structures and a ribbon representation of one of the 15 NMR structures. Ramachandran plot analysis (Table 1) of the domain within the secondary structural elements indicates 82.7% residues in most favored regions and no residues in disallowed regions. Moreover, both termini (residues Met1–Ser20 and Pro101–Asn106) are not included in the structural calculation process for their conformational flexibility. These flexible regions were identified based on diverse structural data: backbone $^1$H-$^{15}$N NOE (Jin et al. 2001; Huang et al. 2003), CSI (Wishart et al. 1992), NOE patterns (Cavanagh et al. 1995), and the Clean SEA-HSQC experiments (Lin et al. 2002). The peak intensity ratios in the Clean SEA-HSQC experiments indicate the amide solvent-exposed degrees of the UbL domain in DC-UbP (Fig. S1), suggesting that, besides N- and C-terminal flexible residues, the residues within the UbL domain such as Thr32, Thr45, Glu60, Ser69–Gly70, Leu73–Asp75, and Glu81 might be also exposed to solvent. These data are in good agreement with the structure determined. The overall fold of the UbL domain in DC-UbP is very similar to that of Ub (Cornilescu et al. 1998), NEDD8 (Whitby et al. 1998), and SUMO-1 (Bayer et al. 1998; Jin et al. 2001) despite low similarity in amino-acid sequences. Like Ub, the UbL domain contains a five-strand β-sheet and two helices. Four strands (β1, β2, β3, and β5) form a main plane, but the β4 strand is twisted against the main plane. Strands β1 and β2 as well as β3 and β5 are anti-parallel aligned, while strands β1 and β5 are parallel. The first α-helix is long (11 residues) and rotates $\sim$45° relative to the β-sheet formed by β1 and β2. The second helix is short with only four residues. This arrangement also appears in SUMO-1 and NEDD8. In the UbL domain of DC-UbP, residues Thr32–Gly33 and Glu60–Gly62 are located in loops, which are flexible, as few long- or medium-range NOEs were obtained. Residues Pro85–Asp87 are also flexible, probably due to lack of resonance assignments for residues Pro85 and Lys86.

**Backbone dynamics**

For the UbL domain of DC-UBP, the relaxation rates for 70 residues (excluding prolines and the resonances that were severely overlapped) out of the 80 residues were obtained (Fig. 3). The backbone $^{15}$N resonances of flexible residues Met1–Ser20 and Pro101–His114 (including histidine tag) show negative NOE values.

---

**Table 1. Structural restraints used and structural statistics**

<table>
<thead>
<tr>
<th>No. of experimental restraints</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total unambiguous distance restraints</td>
<td>851</td>
</tr>
<tr>
<td>Intraresidual</td>
<td>301</td>
</tr>
<tr>
<td>Sequential (i,j-1)</td>
<td>247</td>
</tr>
<tr>
<td>Medium (2 ≤ j-i ≤ 4)</td>
<td>149</td>
</tr>
<tr>
<td>Long range (j-i ≥ 5)</td>
<td>148</td>
</tr>
<tr>
<td>Hydrogen bond restraints</td>
<td>60</td>
</tr>
<tr>
<td>Dihedral angle restraints</td>
<td>83</td>
</tr>
<tr>
<td>φ:</td>
<td>44</td>
</tr>
<tr>
<td>Ψ:</td>
<td>39</td>
</tr>
</tbody>
</table>

Structural statistics

| Coordinate precision (Å), secondary structure |  |
| Backbone (N, Cα, C) | 0.56 ± 0.11 |
| Heavy atoms | 1.62 ± 0.17 |
| Coordinate precision (Å), residues 24–94 |  |
| Backbone (N, Cα, C) | 0.87 ± 0.20 |
| Heavy atoms | 1.82 ± 0.26 |
| RMSD from experimental restraints |  |
| NOE distances (Å) | 0.0250 ± 0.0012 |
| Dihedral angles (deg.) | 0.3786 ± 0.0290 |
| RMSD from idealized geometry |  |
| Bonds (Å) | 0.0030 ± 0.000014 |
| Angles (deg.) | 0.4802 ± 0.0118 |
| Impropers (deg.) | 0.4309 ± 0.0201 |

Ramachandran analysis

<table>
<thead>
<tr>
<th>Residues in most favored regions</th>
<th>24–94</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residues in additionally allowed regions</td>
<td>17.3% 28.2%</td>
</tr>
<tr>
<td>Residues in generously allowed regions</td>
<td>0.0% 2.9%</td>
</tr>
<tr>
<td>Residues in disallowed regions</td>
<td>0.0% 2.8%</td>
</tr>
</tbody>
</table>

*The coordinate precision is defined as the average RMS deviation among the 15 final structures.

*Exception of Pro and Gly residues.
which are not shown in the figure. The average heteronuclear NOE value for residues Gly21–Asn100 is 0.82 (±0.07), and the average $R_1$ and $R_2$ values are 1.55 (±0.05) sec$^{-1}$ and 14.40 (±1.46) sec$^{-1}$, respectively. To better illustrate the backbone dynamics, the plot of $R_2/R_1$ ratio as a function of residue number is also shown in Figure 3. By combining the relaxation data and the structural coordinates (Tjandra et al. 1995), the ratio of two principal axes ($D_1/D_2$) of the diffusion tensor for the UbL domain was estimated to be 1.32, suggesting that the domain has the attribute of anisotropic tumbling.

The backbone $^{15}$N relaxation data provide useful information about the internal conformational dynamics of a protein (Bruschweiler 2003). Residues Gly21, Tyr22, and Glu96 show smaller NOE values than the average. In addition, these residues also exhibit lower $R_1$ values and higher $R_2$ values than the respective average. These data indicate that both termini of the UbL domain are conformationally flexible. Within the core structure of the UbL domain, some residues like Cys24, Ser42, Arg51, Gly58, Glu81, and Leu82 exhibit higher $R_2$ values (also high $R_2/R_1$) than the average, which are often associated with conformational exchange broadening due to intermediate timescale dynamics. Structurally, residues Cys24, Ser42, Glu81, and Leu82 form a cluster on the surface of the structure of the UbL domain. These cluster residues reside mainly in the N-terminal α2-helix (residues Glu81 and Leu82), the loop linking β2-strand and α1-helix (Ser42), and the N-terminal β1-strand (Cys24). This kind of dynamically specific cluster was also revealed in a cold-shock adaptation protein, RbfA (Huang et al. 2003).

Discussion

Although having an overall fold similar to that of Ub, the UbL domain of DC-UbP appears different from Ub in respect to biological function. Being a class II UbL protein, DC-UbP does not possess two conserved Gly residues in its C-terminal UbL domain as that in Ub. In addition, the domain lacks the conserved Lys residues corresponding to Lys29 and Lys48 in Ub, which are responsible for the formation of Ub polymers. The Lys86 residue of the UbL domain is conserved, corresponding to Lys63 in Ub (Fig. 1), which has recently been found to be involved in activation of NF-κB transcription factor (Chen and Fuchs 2004). The three residues (Leu8, Ile44, and Val70) of Ub that have been demonstrated to be important for binding to the proteasome are not conserved in the UbL domain of DC-UbP. Instead, these residues are replaced by Ser31, Phe67, and Ile93 in the domain. The difference in the key residues between the UbL domain and Ub implies that DC-UbP protein may own specific interaction partners for its function. Our in vitro pull-down experiments showed that the UbL domain of DC-UbP does not interact with some Ub binding domains, such as UBA (Ub-associated domain) and UIM (Ub interacting motif) (data not shown). The NMR structure of the UbL domain of DC-UbP may provide structural basis for functional annotation of the DC-UbP protein.

Many UbLs and UDPs possess a UbL fold (McNally et al. 2003; Mueller and Feigon 2003; Ramelot et al. 2003; Sakata et al. 2003). However, they may have different binding partners and diverse biological function. One distinct feature of the UbL domain of DC-UbP is the positively charged surface that is different from Ub, NEDD8, and SUMO-1. As shown in Figure 4, the first surface view of the UbL domain (A) displays a large hydrophobic patch surrounded by positively charged epitopes (contributed by Arg27, Arg29, Lys37, Arg65, Arg71, and Lys86), which is similar to Ub (B) and
NEDD8 (C), but dissimilar to SUMO-1 (D). The region is also the binding interface of Ub and NEDD8 with their specific enzymes or interaction factors (Goettsch and Bayer 2002), suggesting that this region is also a binding interface of the UbL domain in DC-UbP with specific interaction factors. Opposite the positively charged region, on the other side of the β-sheet of the UbL domain (E), there is a large positively charged surface formed by Arg41, His48, Lys50, Arg51, Arg52, His54, and Lys76 in the domain, which is absent on the corresponding surfaces of Ub (F), NEDD8 (G), and SUMO-1 (H). This charge topology difference might implicate an alternative interaction pattern or biological function for the UbL domain of DC-UbP, which remains to be discovered.

In summary, the solution structure of the UbL domain of DC-UbP shares a common topology of UbL fold. The distinct charged surfaces and internal dynamics suggest its specific binding partners with biological significance.

Materials and methods

Protein expression and purification

The DNA fragment encoding the UbL domain of DC-UbP (residues 85–190, numbered 1–106) from human dendritic cell was cloned into the pET22b (+) vector. The plasmid was transformed into *Escherichia coli* BL21(DE3) cells. The fragment protein was separated by affinity chromatography using Ni-NTA agarose resin (QIAGEN) and further purified by FPLC gel-filtration on a Superose 12 column (Amersham Biosciences). Uniformly labeled proteins with 15N and/or 13C were prepared in M9 medium with 15NH4Cl (1 g/L, Cambridge Isotope Laboratories, Inc.) and/or [U-13C]-glucose (2 g/L, Cambridge Isotope Laboratories, Inc.) as sole nitrogen and carbon sources. The purified protein was concentrated to 1.2 mM in a sodium acetate buffer (10 mM NaAc, 100 mM NaCl [pH 5.6]) (92% H2O/8% D2O or 100% D2O) for NMR experiments.

Circular dichroism measurements

The far-UV CD experiments were performed on a Jasco-J715 spectropolarimeter at room temperature. The spectra of 190- to 250-nm wavelength range were recorded with a bandwidth of 1 nm, a speed of 20 nm/min, and a response time of 0.25 sec. Each spectrum was the average of three consecutive scans, followed by subtracting the baseline for buffer solutions (20 mM phosphate, 100 mM NaCl [pH 7.0]). The protein concentration for far-UV CD measurements was about 0.2 mg/mL.

NMR spectroscopy

All NMR measurements were carried out at 25°C on a Varian INOVA 600 spectrometer equipped with three RF channels and...
a triple-resonance pulsed-field gradient probe. The sequential backbone assignment of the UbL domain of DC-UbP was obtained from standard heteronuclear-correlation experiments including 3D HNCA, HN(CO)CA, HNCO, HN(CA)CO, HN(CA)CB, and CBCA(CO)NH. The side-chain assignments were obtained from a 3D HCCH-TOSCY experiment. Distance information between protons was derived from 3D 13\textsuperscript{N}-edited and 13\textsuperscript{C}-edited NOESY recorded on the 13\textsuperscript{N}- or 13\textsuperscript{C}+-labeled sample. Torsion angles were derived from 3J (H\textsubscript{o}, H\textsubscript{\textalpha}) coupling constants measured by HNHA experiment. Solvent-exposed amides were mapped using Clean SEA-HSQC experiments (Lin et al. 2002).

**Structure calculation**

The NMRPipe software suite (Delaglio et al. 1995) was applied to process the NMR data, and NMRView (Johnson and Blevins 1994) software packages were used for picking resonance peaks and for data analysis. The CNS program (Brünger et al. 1998) with the ARIA module (Nile et al. 1997) was adopted to assign NOE peaks and to calculate structures. The protein structures were displayed by using the MOLMOL program (Koradi et al. 1996). Hydrogen bond restraints (two per hydrogen bond) were generated by a combination of H/D exchange data, medium-range NOEs, and CSI. Backbone dihedral angle restraints (φ and Ψ) were derived from the 3J(H\textsuperscript{o}, H\textsubscript{\textalpha}) coupling constants measured with 3D HNHA experiment (Vuister et al. 1993) and the chemical shifts using by TALOS program (Cornilescu et al. 1999). The restraint range was extended to ±50° for the residues located in β-strands and to ±30° for the residues located in other regions. The restraints used for structural calculation are summarized in Table 1.

The structural calculation in combination with iterative NOE peak assignments were performed for nine cycles and a total of 150 structures were finally obtained. The solution NMR structure generated from the calculation is represented by the 15 lowest energy structures out of the 150 structures. The 1H, 13\textsuperscript{N}, and 13\textsuperscript{C} chemical shifts have been deposited in the NMR2Bank (accession no. 6609). The coordinates of the UbL domain of DC-UbP (residues 21–100) and the supporting data have been deposited in the Protein Data Bank (accession code 1TTN).

**Determination of backbone relaxation parameters**

All 13\textsuperscript{N} relaxation data were acquired at 25°C using two-dimensional proton-detected heteronuclear NMR experiments implementing the standard pulse sequences (Farrow et al. 1994). Spin-lattice relaxation rates (R\textsubscript{1}) were measured in a series of spectra with relaxation delays of 10, 20, 400 (×2), 600, 800, 1000, 1200, 1400, 1600 (×2), and 1800 msec. Spin–spin relaxation rates (R\textsubscript{2}) were determined with relaxation delays of 10, 30 (×2), 50, 70, 90, 110 (×2), 130, 150, 170, and 190 msec. A recycle delay of 1.5 sec was used for measuring R\textsubscript{1} and R\textsubscript{2} relaxation rates. The spectra for 1H, 13\textsuperscript{N} heteronuclear NOE measurements were recorded with a 2-sec relaxation delay followed by a 3-sec period of proton saturation. The spectra recorded in the absence of proton saturation employed a relaxation delay of 5 sec. All spectra were processed using NMRPipe and NMRView software. The relaxation rates, R\textsubscript{1} and R\textsubscript{2}, were obtained by fitting measured peak intensity to a two-parameter function using a nonlinear least-square fitting algorithm (Press et al. 1989). Uncertainties in peak intensity were determined from duplicated spectra. The steady-state 1H–13\textsuperscript{N} NOE enhancements were calculated as the ratio of peak intensity in spectra recorded with or without proton saturation. The standard errors were determined from two data sets.

**Acknowledgments**

We thank Dr. Wang W.N. and Miss Lin X.J. for technical help. This work was supported by grants from the 863 Hi-Tech program (2002BA711A13), the Chinese Academy of Sciences (KSCX1-SW-17), Shanghai Commission of Science and Technology (03JC14081), and One Hundred Talents Program of the Chinese Academy of Sciences.

**References**


NMR structure of a ubiquitin-like domain


