A Novel Anti-tumor Cytokine Contains an RNA Binding Motif Present in Aminoacyl-tRNA Synthetases*

Received for publication, March 28, 2000, and in revised form, June 2, 2000
Published, JBC Papers in Press, June 13, 2000, DOI 10.1074/jbc.C000216200

Youngsoo Kim‡‡, Joongchul Shin‡, Rongbao Li†, Chaejoon Cheong, Kyounghhee Kim‡‡, and Sunghoon Kim**

From the ‡School of Chemical Engineering, Yeungnam University, 214-1 Dae-Dong, Kyungsan Kyungbuk 712-749, South Korea, the †Department of Biological Structure, University of Washington, Seattle, Washington 98195, the §Magnetic Resonance Team, Korea Basic Science Institute, Eon-Dong Daejon 305-333, South Korea, and the **National Creative Research Initiatives Center for ARS Network, Sung Kyun Kwan University, Chunchun-Dong, Suwon Kyunggido 440-746, South Korea

Endothelial monocyte-activating polypeptide II (EMAP II) is a novel pro-apoptotic cytokine that shares sequence homology with the C-terminal regions of several tRNA synthetases. Pro-EMAP II, the precursor of EMAP II, is associated with the multi-tRNA synthetase complex and facilitates aminoacylation activity. The structure of human EMAP II, solved at 1.8 Å resolution, revealed the oligomer-binding fold for binding different tRNAs and a domain that is structurally homologous to other chemo-kinases. The similar structures to the RNA binding motif of EMAP II was previously observed in the anti-codon binding domain of yeast Asp-tRNA synthetase (AspRSSC) and the B2 domain of Thermus thermophilus Phe-tRNA synthetase. The RNA binding pattern of EMAP II is likely to be nonspecific, in contrast to the AspRSSC. The peptide sequence that is responsible for cytokine activity is located, for the most part, in the β1 strand. It is divided into two regions by a neighboring loop.

Endothelial monocyte-activating polypeptide II (EMAP II)† is a novel cytokine that was first isolated from a methylchol-anthrene A-transformed fibrosarcoma. However, it is expressed in normal cells and is involved in apoptosis during development (1). The functional cytokine domain (EMAP II) is located in the C-terminal 166 amino acids of its precursor (pro-EMAP II) and is released by proteolysis (2). Recently, caspase-7 was identified as responsible for the proteolytic mat-

urization of pro-EMAP II, thereby linking apoptosis to inflammation (3). Pro-EMAP II, previously known also as p43, is associated with the multi-tRNA synthetase complex in higher eukaryotes (4).

The N-terminal domain of pro-EMAP II containing 146 amino acids does not show homology to any known proteins and interacts with the N-terminal extension of human arginyl-tRNA synthetase. In contrast, the C-terminal domain of pro-EMAP II can interact with different tRNAs. The interaction with pro-EMAP II stimulated aminoacylation activity of the bound arginyl-tRNA synthetase, but its separated N- or C-terminal domain alone did not (5). This result suggested that pro-EMAP II delivers the tRNA, which is bound to the C-terminal domain, to the active site of the associated tRNA synthetase. Interestingly, despite the cytokine activity of the EMAP II, it is not homologous to any known cytokine in the amino acid sequence. It is homologous, however, to the C-terminal regions of methionyl-tRNA synthetases of prokaryotes, archaeas, and nematodes (4), human tyrosyl-tRNA synthetase (6), and yeast Arc1p (7), which is associated with other tRNA synthetases. Although its homology to the domains of various tRNA synthetases appears to be responsible for its ability to bind tRNA, the property of tRNA binding is not structurally explained.

Although pro-EMAP II facilitates protein synthesis by its stimulatory effect on the catalytic activity of the associated tRNA synthetases, its proteolytic product, EMAP II, accelerates apoptosis, suggesting its dual function in cell viability and death. A similar phenomenon was also reported in the case of human tyrosyl-tRNA synthetase, which is secreted from serum-starved cells and cleaved into two distinct cytokines (8). The N- and C-terminal domains of this enzyme showed interleukin-8- and EMAP II-like functions, respectively.

Although the EMAP II-like domain is widely distributed among different tRNA synthetases and their association factors (such as p43, Arc1p, and Trbp111 (9) isolated from various species), the structural information of this domain is limited. In particular, the lack of structural information on EMAP II makes it difficult to understand its diverse biological activities and maturation process. In this report, we discuss the x-ray structure of human EMAP II solved at 1.8 Å resolution as well as the structural features of the oligomer-binding fold (OB fold) (10) for binding different tRNAs and peptide sequences for cytokine activities.

EXPERIMENTAL PROCEDURES

Purification of EMAP II—The human EMAP II gene was cloned into a pET28a vector as described previously (5), except that two restriction
Crystal Structure of EMAP II

**TABLE I**

Summary of crystallographic data

<table>
<thead>
<tr>
<th></th>
<th>Native</th>
<th>CH$_3$HgCl</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diffraction data statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>$C_{2}$</td>
<td>$C_{2}$</td>
</tr>
<tr>
<td>Resolution</td>
<td>20 – 1.80 Å (1.89–1.80 Å)</td>
<td>20 – 2.1 (2.21–2.10 Å)</td>
</tr>
<tr>
<td>Overall observations</td>
<td>96,335</td>
<td>48,451</td>
</tr>
<tr>
<td>Unique observations</td>
<td>31,502</td>
<td>20,817</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>88.2 (37.9)</td>
<td>91.9 (85.2)</td>
</tr>
<tr>
<td>$R_{merge}$ (%)</td>
<td>6.8 (28.2)</td>
<td>6.2 (28.0)</td>
</tr>
<tr>
<td>($I/2I$)</td>
<td>31.3 (1.8)</td>
<td>20.6 (2.3)</td>
</tr>
<tr>
<td><strong>Lattice constants:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a$ = 134.01 Å, $b$ = 38.34 Å, $c$ = 80.99 Å, $\alpha$ = $\gamma$ = 90°, $\beta$ = 112.90°</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phasing statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heavy atom sites</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Phasing power ($a$)</td>
<td>2.13/1.79</td>
<td></td>
</tr>
<tr>
<td>$R_{cal}$(a) (acentrics/centrics)</td>
<td>0.64/0.66</td>
<td></td>
</tr>
<tr>
<td><strong>Refinement statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>20 – 1.80 Å</td>
<td></td>
</tr>
<tr>
<td>No. of reflections</td>
<td>30,839</td>
<td></td>
</tr>
<tr>
<td>No. of reflections ($R_{free}$ calculation)</td>
<td>2,329</td>
<td></td>
</tr>
<tr>
<td>$R_{free}$</td>
<td>20.8</td>
<td></td>
</tr>
<tr>
<td>$R_{free}$</td>
<td>23.2</td>
<td></td>
</tr>
<tr>
<td>No. of non-hydrogen atoms</td>
<td>2,510</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>195</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>r.m.s.d. /$\sigma$</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>Average B-factors</td>
<td>1.58</td>
<td></td>
</tr>
<tr>
<td>Ramachandran analysis</td>
<td>31.7</td>
<td></td>
</tr>
<tr>
<td>Most favored</td>
<td>88.4%</td>
<td></td>
</tr>
<tr>
<td>Additionally favored</td>
<td>11.6%</td>
<td></td>
</tr>
</tbody>
</table>

$\alpha$ $R_{merge}$ = $\sum$ |i| - $\langle$F\rangle|/$\sum$ |i| $\langle$F\rangle.
$\beta$ Phasing power = (r.m.s. $F_{h}$)/(r.m.s. $E$), where $F_{h}$ is the heavy atom structure amplitude and $E$ is the residual lack of closure error.
$\gamma$ $R_{cal}$ = $\sum$$F_{h}$$\cdot$$|F_{o}|$/$\sum$$|F_{o}|$, where $F_{h}$ are test set amplitudes (2,329 reflections) not used in refinement.
$r_{rms}$/$\sigma$ = root-mean-square deviations.

sites, NcoI and SalI, were used to remove the His tag.$^2$ EMAP II was expressed in *Escherichia coli* BL21(DE3) harboring pET28a-EMAP II. The cells were grown at 37 °C in an LB-kanamycin (30 $\mu$g/ml) broth with 1 mM isopropyl-1-thio-$\beta$-galactopyranoside to the cell density corresponding to 0.5 of $A_{600}$. Cells were harvested by centrifugation (6000 $\times$ g for 5 min), resuspended in an extraction buffer (20 mM MES, pH 5.5, 1 mM EDTA, 1 mM $\beta$-mercaptoethanol, 50 mM KCl, 0.5% Tween 20, 0.5% Nonidet P-40, and 50 mM glucose), and lysed by sonication. After centrifugation at 10000 $\times$ g for 15 min, the lysate was loaded on a CM-Sepharose ion exchange column, which was equilibrated with CM buffer (20 mM MES, pH 5.5, 1 mM EDTA, 1 mM $\beta$-mercaptoethanol). The proteins bound to the column were eluted by a five-column-volume linear gradient of NaCl from 0 to 1000 mM in CM buffer. Fractions containing EMAP II were pooled, concentrated to 5 ml, and loaded on a Sephadex G-75 gel filtration column, which was equilibrated with a G-75 buffer (20 mM Tris-HCl, pH 8.0, 1 mM $\beta$-mercaptoethanol, 1 mM EDTA, and 100 mM KCl). EMAP II, eluted from a gel filtration column, was equilibrated with 1 M EDTA, and 100 mM KCl. The crystals of EMAP II were grown over a period of 3 days at 21 °C in hanging drops formed by mixing 1 ml of the protein solution and 3 ml of a reservoir solution, which consisted of 100 mM sodium acetate, pH 4.6, 20% (v/v) polyethylene glycol 4000, and 15 mM MgCl$_2$. Cell dimensions are: $a$ = 134.01 Å, $b$ = 38.34 Å, $c$ = 80.99 Å, $\alpha$ = $\gamma$ = 90°, and $\beta$ = 112.90°. Crystal contains 2 molecules/asymmetric unit and 64% solvent. These crystals have a non-crystallographic 2-fold symmetry at $\omega$ = 36.7°, $\phi$ = 286.4° and $\kappa$ = 180°, described in POLRFN (11). Native and heavy atom derivative data were collected at room temperature on a RAXIS IV detector with a Rigaku RU200 rotating anode (Cu-K$_{\alpha}$) x-ray generator operated at 50 kV and 100 mA. All data were processed with DENZO and SCALEPACK (12).

**RESULTS AND DISCUSSION**

**Overall Structure**—The crystal structure of EMAP II was solved using the single isomorphous replacement method (Table I). The final model contains two molecules per asymmetric unit (residues 3–166, respectively) and 193 water molecules refined at 1.8 Å resolution (Fig. 1, A–C). The EMAP II structure consists of 11 $\beta$-strands forming a structural core and three flanking $\alpha$-helices (Fig. 1C): the strands $\beta 1$ (residues 10–21), $\beta 2$ (residues 28–34), $\beta 3$ (residues 40–46), $\beta 4$ (residues 59–66), $\beta 5i$ (residues 70–72), $\beta 6i$ (residues 75–77), $\beta 7$ (residues 79–85), $\beta 8$ (residues 90–92), $\beta 9$ (residues 103–106), $\beta 10$ (residues 132–134), $\beta 11$ (residues 140–142), and the $\alpha$-helices $\alpha 1$ (residues 53–56), $\alpha 2$ (residues 119–123), $\alpha 3$ (residues 124–130).

---

**Structure Determination**—For heavy atom derivative analysis, the initial 6 mercury sites of CH$_3$HgCl were determined using SOLVE (13). Heavy atom parameters were refined using SHARP (14). A high quality electron density map was obtained using single isomorphous replacement phases. The map was further improved by density modification with SOLONOM (15). The phases were calculated to 2.1 Å using SHARP (14) and then extended to 1.8 Å using solvent flipping (15). Using O (16), the electron density map was readily traceable through the whole molecule, even to side chains without any ambiguous regions (Fig. 1, A and B). Thus, 2-fold non-crystallographic symmetry averaging was not needed to improve the electron density map. The EMAP II model includes residues 3–166 for each molecule in the asymmetric unit. The 2 residues in the N terminus were disordered, and the third residue is registered as Ala.

All crystallographic refinements were carried out using CNS (17) with maximum likelihood refinement, bulk solvent correction, and corrections for anisotropic diffraction. Ramachandran plot analysis with PROCHECK (18) showed that all residues in the current model are in the most favorable and additionally allowed regions. Data phasing and refinement statistics are shown in Table I. Figures were generated with the programs Molscript (19), Bobscript (20), and Raster3D (21).

---

$^2$ S. Kim, unpublished data.
**Fig. 1.** Overall structure and experimental electron density map of EMAP II. 

A, stereo view of the EMAP II structure in a Ca trace. The N-terminal region shows the OB fold in black and the loop-rich C-terminal region in gray. N represents the location of Pro-3 and C indicates the C terminus. Every 20th Ca is labeled with a dark circle. 

B, stereo view of the experimental electron density map. The map calculated after solvent flipping is contoured at 1.2σ. The electron density map is superimposed upon refined coordinates of residues 131–136. Three residues (Asp-108, His-133, and Asp-136 of one molecule from the non-crystallographic 2-fold symmetry in the asymmetric unit) are directly hydrogen-bonded to...
The structure can be divided into the N- and C-terminal regions. The N-terminal region, consisting of the strands $\beta_{1-7}$ and the helix $\alpha_1$, forms a distinct structural motif called the OB fold (10), which is known to bind oligonucleotide and oligosaccharide. The OB fold of EMAP II has a five-stranded Greek key $\beta$-barrel (strands $\beta_{1-3}$, $\beta_4$, and $\beta_7$) that is capped by the short helix $\alpha_1$, located between strands $\beta_3$ and $\beta_4$. The C-terminal region contains strands $\beta_{8-11}$, helices $\alpha_{2-3}$, and several long loops. This C-terminal region contains longer loops than the N-terminal region. This region does not share homology with any known structure.

The structure of EMAP II is described in detail through a structural comparison of OB folds in AspRSSC (23), EMAP II, and PheRSTT (28). A superposition of the OB folds is shown, with ball-and-stick symbols for specific interactions. The residue Glu-188 is not labeled for a clear view. EMAP II is shown in purple and PheRSTT in green. The letters $N$ represent the locations of the secondary structure elements, indicating the position of the OB fold. The nomenclature of the secondary structure elements is based on the same convention with minor variations as described for the OB fold (10). The locations of the secondary structure elements are indicated above and below the aligned sequences. The yellow boxes represent $\beta$-strands, and the green box represents $\alpha$-helix. If there are no corresponding residues in the sequences, they are marked with a "-". The secondary structure elements of AspRSSC are marked on the top, and those of EMAP II and PheRSTT follow the notation at the bottom. EMAP II and PheRSTT have some variations in the assignments of secondary structure elements compared with AspRSSC. For instance, PheRSTT lacks the helix $\alpha_1$, and both EMAP II and PheRSTT have two inserted short strands ($\beta_{5i}$ and $\beta_{6i}$). Therefore, the notations were changed accordingly. The residue numbers for each sequence are shown at the beginning of the sequence. The uppercase letters represent secondary structure elements, and the lowercase letters represent loop regions between the secondary structure elements. The arrows refer to $\beta$-strands and the cylinder to $\alpha$-helix.
The OB fold was first recognized as a common structural motif among four proteins: staphylococcal nuclease (22), aspartyl-tRNA synthetase (AspRS) (23), and the B-subunits of cholera-like (24) and shiga-like cytotoxins (25) from *E. coli*. Since then, 14 proteins have been identified as having the OB fold. Among the aminoacyl-tRNA synthetases, aspartyl-(23), lysyl-(26), asparaginyl-(27), and phenylalanyl-tRNA (28) synthetases (PheRS) contain the OB fold. In the first three tRNA synthetases, the OB fold is located in the N-terminal domain of about 90 residues and has been shown to be responsible for specific recognition of the anticodon loop of tRNA. In the PheRS, the OB fold is present in the B2 domain, whose function is not known (28). The OB folds from the yeast AspRS (AspRSSC), EMAP II, and *Thermus thermophilus* PheRS (PheRSTT) structures are aligned in the Ca backbone (Fig. 2A), and their corresponding residues in the structural alignment can be identified (Fig. 2B). The OB fold for the AspRSSC structure is superimposed onto that of the EMAP II structure at 1.69 Å for 41 Ca atoms, whereas the PheRSTT is at 1.18 Å for 66 atoms. The structural alignments show that the β-barrel of AspRSSC is closed between strands β3 and β5, whereas the corresponding β-barrels of EMAP II and PheRSTT (β3 and β5) are partially open (Fig. 2A and B). The alignments also show that the helix α1 and loop L5 regions of AspRSSC are the least conserved among the three structures. EMAP II has a shorter α1 helix consisting of 4 residues, whereas AspRSSC has 9 residues in this helix. In contrast, the PheRSTT structure does not carry the α-helix. In addition, EMAP II and PheRSTT have two inserted short β-strands (β5i and β6i) in the region corresponding to loop L5 of AspRSSC.

Gel mobility shift assays indicate that EMAP II can bind several tRNA species with similar dissociation constants (4). The OB fold of EMAP II is most likely to be responsible for this nonspecific tRNA binding. In the structure of AspRSSC that is complexed with tRNA (23), anticodon bases are bound to a valley formed on one side of the barrel. Residues from strands β1, β2, β3, and β5 form the floor of the valley, whereas residues from loops L1 and L5 line the edges. Specific interactions are formed from the conserved residues Phe-127, which stacks against the anticodon base U35, and both Gln-78 and Gln-138, which form hydrogen bonds to the anticodon bases (Figs. 2A and 3A). If we assume that a tRNA approaches the OB fold from the same direction in both AspRS and EMAP II, then their binding patterns would be somewhat different. In EMAP II, the close proximity of loops L1, L4, β5i, and β6i provides a smooth surface for the interaction with the incoming tRNA instead of a specific binding pocket of AspRSSC (Figs. 2A and 3B). For example, there are no aromatic side chains exposed on the side of the β-barrel, as found in Phe-127 of AspRSSC. Instead, several residues, such as Asp-24, Glu-32, Thr-42, Arg-73, and Gln-78, are exposed to solvent. These 5 residues are highly conserved among the proteins with the EMAP II-like domain (4, 9). Thus, the nonspecificity in the tRNA binding of the EMAP II-like domain is likely to occur through these residues on the surface. The structural similarity between the OB fold of EMAP II and the B2 domain of PheRSTT probably predicts that the B2 domain may also function as a nonspecific binder of region by three hydrogen bonds of residues 108, 133, and 136 (Fig. 1B). These surface contacts occur at the opposite end of EMAP II from the putative tRNA binding and cytokine activity sites and will be discussed later. Interestingly, a structure-specific tRNA-binding protein (Trbp111) from *Aquifex aeolicus* (9) and the C-terminal domain of *E. coli* methionyl-tRNA synthetase (data not shown) (2), which share sequence homologies with the OB fold of EMAP II, form dimers. However, the C-terminal region of EMAP II, which is involved in the dimer contacts in the asymmetric unit of the crystal, is not present in Trbp111 and the C-terminal domain of *E. coli* methionyl-tRNA synthetase (9). In addition, a gel filtration analysis showed that hamster EMAP II exists as a monomer in the solution (4).
EMAP II, strands b of monomers of these chemokines (Fig. 4) consists of three strands (β1–β3) and one short helix (α1). The peptide sequence (residues 12–18) involved in chemotaxis of EMAP II is positioned at the beginning of this homologous domain. Three residues, Cys-15, Ile-17, and Thr-18, which were exposed to solvent, are shown by ball-and-stick symbols close to the end of strand β1. The other functional peptide sequence (residues 6–11) is located at the other side of a loop formed by residues 99–102. The hydrophobic residues Val-6, Leu-8, and Leu-11 form a shallow hydrophobic pocket with Phe-107 and Leu-132. The Cas of residues 12, 99, and 102 are shown as black circles.

FIG. 4. Domain with cytokine activity. Stereo views of EMAP II are shown as Cα traces, truncated at residue 133 for a clear view. The view is in the same orientation as Fig. 3B. The domain, consisting of the residues 13–57, which is homologous to the chemokines, is shown in gray and consists of three strands (β1–β3) and one short helix (α1). The peptide sequence (residues 12–18) involved in chemotaxis of EMAP II is positioned at the beginning of this homologous domain. Three residues, Cys-15, Ile-17, and Thr-18, which were exposed to solvent, are shown by ball-and-stick symbols close to the end of strand β1. The other functional peptide sequence (residues 6–11) is located at the other side of a loop formed by residues 99–102. The hydrophobic residues Val-6, Leu-8, and Leu-11 form a shallow hydrophobic pocket with Phe-107 and Leu-132. N represents the location of Pro-3 and C the location of His-133. The Cas of residues 12, 99, and 102 are shown as black circles.

tRNAs. Presumably, all of the proteins containing the EMAP II-like domain will carry the partially open OB fold that is shown in the EMAP II and PherSTTT structures. This may provide a capability to function as a nonspecific binder of tRNAs.

Implications for Cytokine Function—The structural homology search program DALI (29) identified that the N-terminal 45 residues of EMAP II (residues 13 to 57) have limited structural homologies with monomers of chemokines, such as RANTES (residues 22–61) (30), human monocyte chemooattractant protein (residues 25–65) (31), and neutrophil-activating peptide-2 (residues 38–79) (32). They share a similar structural motif consisting of a three-stranded β-sheet with an α-helix. In EMAP II, strands β1–β3 and helix α1 form a homologous structural domain of monomers of these chemokines (Fig. 4) even though the primary sequence of EMAP II shows no homology. The Cas of those chemokine structural motifs in RANTES, human monocyte chemooattractant protein, and neutrophil-activating peptide-2 could be superimposed onto the corresponding region of EMAP II with root-mean-square deviations of 1.78, 1.75, and 2.5 Å for 30, 23, and 21 atoms, respectively.

Studies on murine and human EMAP II (33, 34) suggest that the residues responsible for cytokine activity in human EMAP II are the residues from 6 to 18 (VSRLDLRIGCIIT). This peptide stimulated chemotaxis and increased cytosolic calcium in the presence of CHO cells, as previously reported (13). The x-ray data were collected at the Yeungnam University Instrumental Analysis Center.

REFERENCES

Crystal Structure of EMAP II

27068


