Crystal structure of human angiogenin with an engineered loop exhibits conformational flexibility at the functional regions of the molecule

Nethaji Thiyagarajan, K. Ravi Acharya

Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK

**ABSTRACT**

Human angiogenin (ANG) is an angiogenic molecule and a ribonucleolytic enzyme with significant amino acid sequence identity to pancreatic RNase A, plays a critical role in the establishment and growth of tumours. An association between ANG and cancer has been observed in more than 25 clinical studies to date. In addition, ANG has now been shown to be implicated in Amyotrophic Lateral Sclerosis (ALS) and Parkinson’s Disease (PD). Structural and biochemical studies so far have showed several distinguishing features of ANG molecule compared to RNase A and provided details of the putative cell binding site, active site, nuclear translocation sequence and the roles of residues in binding and cleaving RNA. A key finding elucidated from the structural study on ANG is the presence of a ‘blocked’ C-terminus (part of the active site apparatus) compared with RNase A. Here we report the crystal structure of ANG with an ‘engineered-loop’ from eosinophil derived neurotoxin (a homologue of ANG) which has resulted with local perturbations (conformational flexibility) at the cell binding site and at the C-terminus of the molecule. This experimental observation will now provide a new avenue to design compounds (potent inhibitors) through a structure guided drug design route.

**1. Introduction**

Angiogenin (ANG), also known as ribonuclease 5 (RNase 5), is an angiogenic molecule that induces neovascularization [1]. It is a monomeric protein of 14.1 kDa that belongs to the pancreatic ribonuclease A (RNase A; EC 3.1.27.5) superfamily [2]. ANG has similar catalytic properties to RNase A – it cleaves preferentially on the 3’ side of pyrimidine and follows a transphosphorylation/hydrolysis mechanism [3,4]. The enzymatic activity of ANG, *albeit* several orders of magnitude lower than that of RNase A towards conventional RNase substrates [4], is essential to the angiogenic process [3,5].

The known biological properties of ANG encompass induction of angiogenesis to stimulate cell proliferation and promoting cell survival. Under growth conditions, ANG is translocated to nucleus (where it accumulates in nucleolus) and stimulates ribosomal RNA (rRNA) transcription, thus facilitating cell growth and proliferation [6]. Recently it has been established that under stress conditions, ANG accumulates in cytoplasmic compartments and modulates the production of tRNA, a novel class of small RNA that is derived from rRNA and is induced by stress. ANG-mediated tRNA reprogram protein translation, save anabolic energy, and promote cell survival [7–10].

Based on biochemical and structural data it has been established that ANG has three distinct functional sites – (a) a putative binding site for an unknown receptor, (b) a nuclear localization sequence (NLS), and (c) a catalytic site. The crystal structure of ANG has revealed the RNase A fold and the conservation of the catalytic triad – His13, Lys40 and His114 [11]. The loop region from Lys60 to Asn68 forms the receptor-binding site that interacts with a cell surface receptor [12]. Upon binding to the cell surface receptor, ANG is internalized and translocated to the nucleus [13]. The nuclear translocation process is mediated by an NLS located between amino acid residues Met30 and Gly34 [14]. The ribonucleolytic activity of ANG [5] is believed to function in stimulating RNA transcription after ANG is localized in the nucleus. It is now firmly established that all three sites are essential for ANG to have angiogenic and growth stimulating activities.

Apart from the its vascular function, ANG has also been firmly associated with several neurodegenerative conditions. Studies by Greenway et al. [15] on a large cohort of Amyotrophic Lateral Sclerosis (ALS) patients (both familial and sporadic) showed for the first time that ANG is implicated in ALS. Since then several other studies worldwide have also identified ANG mutations in ALS patients. More recently
large multi-site study reported the association of mutations in ANG with ALS as well as Parkinson’s disease (PD) [16,17]. Since the early reports on ANG association with ALS, it has been established that ANG also possesses neurotrophic and neuroprotective functions [18–20]. A recent detailed study has provided for the first time the structural and molecular insights into the mechanism of action of human ANG-ALS variants in neurons [21].

The role of ANG as a tumour angiogenic factor has been well documented. ANG has been shown to be upregulated in a variety of human cancers [22]. Elevated serum level and/or enhanced tissue expression of ANG have been noticed in all types of solid and blood cancers so far examined. A major mechanism by which ANG induces angiogenesis is related to its activity in stimulating ribosomal RNA (rRNA) transcription [22], a common downstream event of tumour angiogenesis [23]. In addition, ANG inhibitors have been shown to inhibit not only ANG-induced angiogenesis but also those induced by other angiogenic factors such as VEGF, FGF, and EGF [24]. ANG inhibitors would therefore have a profound effect in inhibiting tumour angiogenesis. Indeed, ANG inhibitors such as its siRNA [25], antisense [26], monoclonal antibodies [27], soluble binding proteins [28], enzymatic inhibitors [29], and nuclear translocation blockers [24,30–32] have all been shown to inhibit tumour angiogenesis and cancer progression in various animal models.

We have identified different types of inhibitors of ANG based on our structural, mutagenesis and biophysical studies. These will be used to probe the mechanistic details of ANG. In particular, our aim is to study the nucleotide inhibitors that target the enzymatic active site of ANG and also monoclonal antibodies against ANG. The angiogenic activity of ANG is critically dependent on its enzymatic activity and inhibitors that bind to its enzymatic active site also inhibit its angiogenic activity. Our results from ANG inhibitors have demonstrated proof of principle, but design of structure-based small molecule nucleotide inhibitors so far proved to be elusive due to the blockage of the active site predominantly by Gin I17 and the remaining C-terminal segment of the ANG molecule [11,33–37]. Mutation of Gin I17 to Gly/Ala has shown to increase the ribonucleolytic activity of ANG [38]. However, the structures of ANG in complex with – (a) a leucine rich, natural protein inhibitor (50 kDa) of ANG (the ribonuclease inhibitor – RI; [39]; PDB ID: 1A4Y) and (b) an antibody fragment 26-2F (shown to inhibit ANG) ([40]; PDB id: 1H0D) have shown significant conformational changes at the cell-binding region of ANG. This region is close to the active site of ANG and thus it was predicted that induction of similar conformational change would allow the C-terminal β-strand (Fig. 1A) to undergo small movement to accommodate the substrate/inhibitor without significant alteration in the catalytic property of the molecule.

In this study we have substituted HGGSPWPPG1 (a loop that interacts with RI) of ANG [39] with 86TTPSQNISN86 of eosinophil-derived neurotoxin (EDN, also shown to interact with RI) [41] and constructed an ANG–EDN hybrid (AEH) protein molecule of molecular weight ~14.5 kDa. EDN is also a member of RNase A superfamily (known as RNase 2), possesses weak ribonucleolytic activity, antiviral property [42–45] and neurotoxicity [46,47]. Based on our previous crystal structure analysis of ANG–RI [39] and EDN–RI [41] complexes it was realistic to assume that AEH molecule would also retain RI binding property. The experimental data presented in this report shows that the design of AEH molecule proved to be useful to study the relationship between the RI binding loop and the structural changes in the cell-binding region. In addition, the high resolution crystal structure of the AEH molecule reported here shows that the ‘engineered loop’ causes conformational changes that would perturb the network of interactions from the nuclear localization sequence to cell binding region extending to the active site (through inclusion of the C-terminal segment of the molecule) without altering the enzymatic activity. This information may prove useful for the design of ‘ANG–nucleotide inhibitor complexes’. Similar studies have been reported previously amongst RNase A family members and also other proteins [48–52].

2. Materials and methods

2.1. Cloning, expression and purification of AEH

The ANG–EDN hybrid protein (AEH) was prepared using recombinant DNA technology where the loop region (HGGSPWPPG1) of ANG was substituted with the corresponding segment (loop region) (TTPSQNISN86) of EDN. The pET22b(+)/Escherichia coli expression plasmid was provided by Dr. Robert Shapiro, Harvard Medical School (USA). The protein expression trials were carried out in BL21-DE3 Codon plus (RIPL) cells in the culture medium terrific broth. The required AEH protein was purified from inclusion bodies using the protocol described by Holloway et al. [53]. Purity of the protein was analysed on an SDS–PAGE.
2.2. Mass spectrometry

Purified protein mass of AEH was experimentally confirmed (in addition to amino acid sequence verification) using electrospray ionization mass spectrometry (ESMS) service at University of Bath. Protein solutions were prepared for mass spectrometry by dilution of a desalted standard stock solution of protein in water (1 mg/ml, 70 μM) in a 50:50 ratio with HPLC-grade acetonitrile containing 0.1% formic acid. ESMS was performed in positive ion mode on a QStar XL System (Applied Biosystems), fitted with a NanoMate (Advion Biosciences) automated nanospray source. A delivery pressure of 0.3 psi of nitrogen gas and a spray voltage of 1.6 kV were applied to the sample to generate the nanospray plume. Transformed mass spectra were calculated by use of the Bayesian protein reconstruct algorithm provided as part of the MDS Sciex Analyst 1.1 software package.

2.3. RNase activity assay

The activity of AEH protein toward yeast tRNA was determined by measuring the formation of acid-soluble fragments as described by Shapiro et al. [4]. Assay mixtures contained 2 mg/ml yeast tRNA (Sigma), 0.1 mg/ml bovine serum albumin (BSA), and 0.05, 0.1, 0.2, 0.3, 0.4, or 0.5 μM test protein in 33 mM Na-Hepes and 33 mM NaCl, pH 7.0. After 2 h of incubation at 37 °C, reactions were terminated by the addition of 2.3 volumes of ice-cold 3.4% perchloric acid. The mixture was then centrifuged at 13,000 rpm for 10 min at 4 °C. The absorbance of the supernatant was measured at 260 nm.

2.4. X-ray crystallography

X-ray diffraction data were recorded (at 100 K) from a single crystal of AEH protein grown in 1.4–2.5 M NaCl at 16 °C on a Mar CCD 165 at station PX 10.1, Synchrotron Radiation Source, Daresbury, UK. No cryoprotectant was used. The protein was crystallized using hanging-drop vapour diffusion method at 1:1 drop ratio of protein (10 mg/ml in water) to mother liquor. The data were processed and scaled in primitive orthorhombic P2₁2₁2₁ space group (2 molecules/asymmetric unit) using HKL2000 software suite [54] to a 1.76 Å resolution. The data were complete to 91.9% overall with a merging R-value of 0.082 (Table 1). Initial phases for AEH molecule were obtained by molecular replacement method using the human ANG structure (PDB ID: 1ANG; [11]) as the starting model. The software used for molecular replacement trials were MOLREP [55] and Phaser [56]. The resultant model of AEH protein was taken through a rigid body refinement (after deletion of the substituted loop region [HGGSPWPPP]) in the Ramachandran plot statistics using MolProbity validation tool [59,60]. There were no residues in the disallowed region of the Ramachandran plot. Crystallographic data statistics are summarized in Table 1. All figures were drawn with PyMOL (DeLano Scientific LLC, San Carlos, CA, USA) and rendered with POV-ray.

Table 1

<table>
<thead>
<tr>
<th>Space group</th>
<th>Orthorhombic – P2₁2₁2₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell dimensions</td>
<td>a = 38.4 Å; b = 54.7 Å; c = 94.7 Å</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>50–1.76</td>
</tr>
<tr>
<td>R cryst (%)</td>
<td>0.082 (0.299)</td>
</tr>
<tr>
<td>wR (outer shell)</td>
<td>14.5 (2.9)</td>
</tr>
<tr>
<td>Completeness (outer shell) (%)</td>
<td>91.9 (65.8)</td>
</tr>
<tr>
<td>Total number of reflections</td>
<td>230,027</td>
</tr>
<tr>
<td>Number of unique reflections</td>
<td>20,481</td>
</tr>
<tr>
<td>Redundancy (outer shell)</td>
<td>4.8 (2.5)</td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>29.0</td>
</tr>
<tr>
<td>R cryst (%) / R free (%)</td>
<td>0.239/0.285</td>
</tr>
<tr>
<td>Average B-factor (Å²)</td>
<td>34.7</td>
</tr>
<tr>
<td>Overall</td>
<td>33.6, 35.5</td>
</tr>
<tr>
<td>Protein (chain A, B)</td>
<td>36.9</td>
</tr>
<tr>
<td>Chloride ion</td>
<td>37.1</td>
</tr>
<tr>
<td>Solvent</td>
<td>37.1</td>
</tr>
<tr>
<td>RMS deviation</td>
<td>0.007</td>
</tr>
<tr>
<td>Bond angle (°)</td>
<td>1.012</td>
</tr>
<tr>
<td>Bond angle (%)</td>
<td>97.71</td>
</tr>
<tr>
<td>Additional favoured (%)</td>
<td>2.29</td>
</tr>
<tr>
<td>PDB ID</td>
<td>4B36</td>
</tr>
</tbody>
</table>

3. Results and discussion

3.1. General features

All ANG structures determined so far [11,33–37] have a highly conserved RNase A fold – a characteristic feature of proteins belonging to this family (Fig. 1A). Briefly, the RNase A fold (a kidney-shaped molecule) consists of α and β secondary structure elements forming the core of the molecule connected by loop structures. In ANG molecule there are three α-helices, seven β-strands and a 310 helix (at the C-terminus) (Fig. 1A). The core of the molecule is stabilized by three disulphide bridges. The catalytic site [formed by residues His13 (α1), Lys40 (loop connecting α2 and β1) and His114 (β7)] is located on the concave structure of the molecule. The NLS residues lie on helix α2, RI binding residues in the loop connecting β4 and β5, and receptor binding site includes the loop connecting helix α3 and strand β2 and β3. The AEH structure retains these features in both molecules of the asymmetric unit but displays significant conformational flexibility in the loop regions of the molecule (Fig. 1B–D, see below).

Comparison of native ANG structure [11] with the AEH structure (present study) result with a root mean square deviation of 1.34 and 1.71 Å (Cα atoms) for molecule A and B (for the two copies in the hybrid molecule in the asymmetric unit) respectively. The active site of AEH molecule comprises the catalytic triad residues His13, Lys40 and His116 (His13, Lys40 and His114 in ANG) which form links with residues Glu12, Asn43, 117Leu-Asp-Gln119 (115Leu-Asp-Gln117 in ANG) either through direct hydrogen bonding or through a water-mediated network. A detailed analyses of the catalytic site of AEH against a few members of the RNase A superfamily determined at atomic resolution – the prototypic molecule RNase A (PDB: 2E3W, 1.05 Å resolution [61]) showed the presence of a highly conserved solvent network involving four water molecules (labelled as 1–4 in Fig. 2A–E). It is likely that these solvent molecules might play an important role in substrate recognition and/or binding. For example, in an ANG-ALS variant Glu12Leu structure [33], the enzyme lost both its ribonucleolytic and angiogenic properties (required for
Fig. 2. Stereo representation of conserved solvent network in (A) AEH – molecule A; (B) AEH – molecule B; (C) ANG; (D) EDN (solvent molecule equivalent to 4 or chloride ions are coloured in magenta) and (E) RNase A. Conserved water molecules are labelled as 1–4. Chloride ions are coloured in orange. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Conformational flexibility

3.2.1 RI-binding loop

A large movement in the engineered loop region (86TTPSPQNISN95) was observed where residue Pro88 undergoes a major shift of $\sim 90^\circ$ in comparison to its position in the native ANG structure (Figs. 1B, 3). In the ANG structure this region (also called as the RI-binding loop) is involved in binding to RI as observed in the crystal structure of ANG–RI complex [39] through several main- and side-chain interactions. In both the molecules of AEH, the RI-binding loop adopts similar conformation, but significantly different from that observed in the ANG–RI complex or in the native ANG structure (Figs. 1B, 3). Hence we predict that this change would perturb RI binding even though AEH molecule seems to possess similar ribonuclease activity as native ANG (Fig. 1E). The $K_i$ value of RI for ANG and EDN are $7.1 \times 10^{-16}$ M [63,64] and $6.7 \times 10^{-15}$ M [41] respectively. The structural superposition of native EDN (PDB: 1GQV [62]), EDN-RI (PDB: 2BEX [41]), native ANG (PDB: 4AOH [21]), ANG–RI (PDB: 1A4Y [39]) against AEH molecules show that the conformational changes observed are unique to AEH molecules. However, the nuclear translocation sequence31RRRGL35 in the AEH molecule does not undergo any conformational change and retains similar conformation as in native ANG.

3.2.2 Cell receptor binding loop

The second structural change was observed at the secondary structural element $\beta$2 strand of ANG, which transformed to a loop region (Gly62–His65) in the AEH structure (Fig. 1A, D), and the loop connecting the strand $\beta$2 and $\alpha$3 could not be modelled due to lack of visible electron density. This region is a part of the putative cell binding segment of ANG [12,65,66]. This segment has shown to undergo significant conformational change by adopting an extended conformation based on the structural details observed from the crystal structures of ANG–RI [39] and ANG–Fab antibody [40] complex reported previously from our laboratory.

3.2.3 C-terminal region

An unexpected conformational flexibility was observed at the C-terminal segment of AEH structure starting from Leu117 (Leu115 based on ANG numbering). This structural movement at the active site displaces the Gln119 of AEH (Gln117 of ANG) by $\sim 90^\circ$ inwards (away from the active site) (Figs. 1C, 4). In the most predominantly
observed state (as observed in the native structure of ANG [11]), the side chain atoms OE1 and NE2 of Gln117/Gln119 of ANG/AEH respectively are involved in direct hydrogen bonding interactions with active site residue His13 and Thr44, thus blocking substrate binding and the availability of His13 and Thr44 for catalysis. As observed in one of the molecule of AEH structure, the displacement of Gln119 (Gln117 of ANG) opens the active site and enhances the binding of the substrate. The side chain atoms OE1 and NE2 of Gln119 (Gln117 of ANG) of AEH interact with Arg103, Arg104 and Val106 through hydrogen bonding interactions. A chloride ion (due to high concentration of NaCl used in crystallization medium) was bound in the place of Gln119 following the displacement of Gln119 (Fig. 4). We tested the AEH protein for its catalytic activity in the presence of 0.1–2.5 M NaCl concentration and did not observe any alteration in the RNase activity and hence conclude that the observed chloride ion at this site has no direct role in catalysis. Comparison of electrostatic potential between ANG and the AEH (in both molecules) (Fig. 5) clearly shows the apparent change in potential charge at the active site due to the conformational change at the C-terminal region – resulting in larger basic surface area for the nucleotide substrate to bind. Thus for the first time the AEH structure has provided a glimpse of conformational flexibilities at the functional regions of the parent molecule.

4. Conclusions

We have compared the solution structure of ANG (determined by nuclear magnetic resonance) (PDB: 1AWZ [67]) with the high resolution structure of ANG in the crystalline state (determined by X-ray crystallography, PDB: 4AOH [21]) and found that the loop regions retain similar conformation in both structures. Thus, substitution of ANG loop region (86TTPSPQNN55) in the AEH structure seem to have resulted in conformational changes at the functional regions of the molecule (that are observed for the first time in the ANG structure determined so far). The hybrid molecule AEH retains the full complement of ribonucleolytic activity and do not exhibit any perturbation at the catalytic centre (involving the catalytic triad). However, significant movement of the cell-binding segment of AEH molecule was evident. In addition, the C-terminus of AEH molecule (in one copy of the asymmetric unit) adopts a non-obstructive conformation (hitherto unseen) in the hybrid molecule. These observations will have implications for the future design of ANG nucleotide based inhibitors.

The structural flexibility observed in the AEH molecule gives us some insight about the network of interactions involving the NLS and the cell-binding sites through the active site and C-terminal segments. These changes appear to be a sequential event during catalysis. The perturbation of the C-terminus moves the Gln119 residue away from its obstructive state and Arg121 points outwards – leading to an enhanced surface accessibility and gain in net positive charge. In order to assess the feasibility of substrate binding, we have docked an oligomer of DNA backbone (containing only the ribose and the phosphate moiety) to this cavity (Fig. 5D). The length of the DNA backbone used was a 10-mer. From the modelling studies we predict that Arg5, His8, Gln12, Lys40, Asp41, Lys82, His116 (equivalent to H114 of ANG) and Arg124 make hydrogen bonding interactions with the phosphate and ribose molecules.

Database

The atomic coordinates and structure factors for AEH (with code 4B36) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

Acknowledgements

This work was supported by the Wellcome Trust (UK) through a programme grant (number 083191) and an equipment grant (number 088464) to K.R.A. We thank Dr. Robert Shapiro for providing the plasmid carrying the engineered segment of human angiogenin; Dr. Matthew Baker, Jodie Gibson and Claire Padgett for their contribution to structural and protein biochemistry part of the project.

References


