Supplementary Information

Dual binding site of angiogenin and its inhibition mechanism: the crystal structure of the rat angiogenin-heparin complex

Kwon Joo Yeo, Eunha Hwang, Kyong-Mi Min, Chung-Kyung Lee, Kwang Yeon Hwang, Young Ho Jeon, Soo-Ik Chang, Hae-Kap Cheong**

Expression, purification, crystallization, data collection and structure solving

The catalytic domain of rAng (Q1-L121, numbered from the catalytic domain in this report) was expressed and purified as reported previously.[1] A monomeric form of rAng (13.5 kDa) was obtained by the refolding of inclusion bodies from *Escherichia coli*.[1] Crystals of rAng were obtained using the sitting-drop vapor-diffusion method at 293 K. The final protein concentration was 16 mg/ml in 50 mM MES pH 6.0, 100 mM NaCl. For the rAng-heparin complex, the mixture (protein: heparin (8 mer) = 1 : 2, molar ratio) was used in the same buffer. The crystals of apo rAng were grown in 24% PEG 8,000, 5 mM zinc acetate, 100 mM sodium cacodylate (pH 6.5). The crystals of the complex were obtained from a solution containing 9% PEG 8,000, 5 mM zinc acetate, 100 mM sodium cacodylate (pH 6.5). The apo rAng and complex data sets were collected using a synchrotron radiation source at beam-line 4A in the Pohang Accelerator Laboratory (Pohang, Korea) and beam-line 17A at the Photon Factory (Tsukuba, Japan). The crystals belonged to space group $p2_12_12_1$ and contained 2 molecules in the asymmetric unit. The wavelength of the X-ray beam was 1.00 Å, and the crystal-to-detector distance was 259 mm. The diffraction images were integrated and scaled using HKL-2000.[2] Crystals of apo rAng and the rAng complex diffracted to 1.8 Å and 2.2 Å resolution, respectively. The structures were solved using molecular replacement; the search

model used for all data sets was 2BWK (PDB code). Refinement of the apo structures was carried out using PHENIX.[3] All of the final models were validated by PROCHECK.[4] The data collection and refinement statistics are summarized in Table S1.

Table S1. Data-collection and structure-refinement statistics

	Apo	**Complex**
Data collection		
Resolution (Å)	50~1.8 (1.83~1.80)	50~2.2 (2.24~2.2)
Space group	$P2_12_12_1$	$P2_12_12_1$
Wavelength	0.97880	1.00000
Unit-cell parameters		
a, b, c (Å)	43.093, 43.674, 154.422	39.415, 42.420, 154.312
α, β, γ (°)	90, 90, 90	90, 90, 90
Molecules per asymmetric unit	2	2
Redundancy	5.8 (4.3)	9.9 (9.7)
R_{merge} (%)	6.0 (31.9)	9.5 (36.7)
Mean $I/\sigma I$	9.7 (2.173)	11.9 (10.0)
Completeness (%)	98.7 (98.5)	99.3 (99.9)
Refinement		
R_{work}/R_{free} [a]	0.2315/0.2534	0.2141/ 0.2655
RMSD bond length (Å)	0.008	0.010
RMSD bond angle (°)	1.058	1.334
Ramachandran plot		
Favored regions (%)	96.5	97.3
Allowed regions (%)	3.5	2.7
Disallowed regions (%)	0	0
Average B factor (Å^2)		
Chain A	31.51	27.62
Chain B	32.06	33.27
ligand		37.21
Water	31.79	32.50
PDB code	4QFI	4QFJ

The crystal structures of apo rAng and 2rAng-heparin complex were solved at resolutions of 1.8 Å and 2.2 Å, respectively (Table S1). Two rAng molecules (apo A and apo B) were present in the asymmetric unit with ionic interactions between the side chain of the E66

residue of apo A and the side chains of R31 and R92 residues of apo B (Fig. S1a). Similar to murine angiogenin (mAng),[5] rAng contains one Zn^{2+} ion, which is coordinated by E41, H82, acetate, and the carbonyl oxygen of the G85 backbone. The acetate molecule likely came from the crystallization solution (Fig. S1b).

The sequence of rAng is 74% identical to hAng (Fig. S2a), and as shown in Fig. S2b, the overall structure of the apo rAng (apo B) superimposes well onto hAng (PDB code: 1A4Y). The primary residues of rAng responsible for the ribonucleolytic activity (H13, K40, and H113) are conserved and are located in the same positions as in hAng. Another acetate molecule is observed near the side chain of the H113 residue (Fig. S1b). The ^{31}RRRGL^{35} motif in rAng is conserved in hAng (Fig. S2a). However, the ^{60}KNGNPHREN^{68} motif, which is part of the cell-binding site, contains differences between the two proteins in 4 out of the 9 residues. The differences in this region may confer differing binding properties between the two proteins.

Ribonucleolytic activity of the native rat angiogenin versus the double mutants R31A/R32S and K50L/K54Q using Zymogram electrophoresis.

The double mutants of R31A/R32S and K50L/K54Q of rat angiogenin were obtained from Mr. Murali Krishna and Mr. Sundar Hengoju (Department of Bichemistry, Chungbuk National University, Cheongju, Republic of Korea). SDS-PAGE was performed using a running gel that was co-polymerized with poly(C) (0.6 mg/ml).[6] After electrophoresis, SDS was removed from the gel by washing with 10 mM Tris-HCl buffer (pH 7.5) containing 2-propanol (20% v/v). Native rat angiogenin or the double mutants R31A/R32S and K50L/K54Q were renatured by first washing the gel for 10 min with 10 mM Tris-HCl buffer (pH 7.5), then for 10 min with 0.10 M Tris-HCl buffer (pH 7.5), and finally for 10 min with 10 mM Tris-HCl buffer (pH 7.5). The gel was stained with 10 mM Tris-HCl buffer (pH 7.5)

containing toluidine blue (0.2% w/v) and destained with distilled and deionized H_2O.

Nuclear translocation of the native rat angiogenin versus the double mutants R31A/R32S and K50L/K54Q in HeLa cells

HeLa cells were maintained in MEM/EBSS (Thermo Scientific, USA) with 10% FBS (Thermo Scientific, USA) in the presence of penicillin (100 units/ml) and streptomycin (100 mg/ml). HeLa cells were incubated with Cy3-labeled native rat angiogenin (Cy3-rAng) or Cy3-labeled double mutants R31A/R32S and K50L/K54Q (Cy3-R31A/R32S and Cy3-K50L/K54Q) of rat angiogenin (5 μg/ml) for 30 min at 310 K. The cells were washed with PBS and fixed with paraformaldehyde for 15 min at 298 K. Then, the cells were washed with cold PBS and TritonX-100 for 10 min at 298 K and stained with DAPI 9(4',6-Diamidino-2-Phenylindole, Dilactate) for 1 hr at 298 K. The cells were imaged at original 60X magnification using a fluorescence microscope (Olympus, USA). The images were analyzed with Image J (Microsoft, USA).

Molecular dynamics simulations for the rAng-heparin complex

In order to know a more preferential binding site of heparin toward rAng, we performed MM-PBSA analyses based on 50 ns trajectories of molecular dynamics simulations with AMBER package.[7] The end-state binding free energies in complex A and B were -50.1 ± 7.6 and -27.1 ± 13.1 (kcal/mol), respectively, indicating that the interaction of rAng and heparin in complex A is stronger. The result is reasonably consistent with the observation that the number of intermolecular contacts in complex A is more than that in complex B.

We employed AMBER package (ver. 12) for atomistic molecular dynamics (MD) simulation. Force fields for a sulfated heparin, IdoA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS(6S), were generated by LEaP program with the help of GLYCAM06.[8]

Separation of a protein unit from the complex produced two 1:1 complexes of rAng and heparin. Each system was solvated with TIP3P water molecules in a periodic truncated box such that its walls were at least 10 Å away from the solute. Sodium and chloride ions were added to constitute 150 mM ionic strength and neutralized total charge. All the simulations were performed with PMEMD module using ff99SB-ILDN and GLYCAM_06j-1 force fields. [9] SHAKE constrained all bonds involving hydrogen atoms, permitting a time step of 2 fs. The particle mesh Ewald method was used to calculate long-range electrostatic interactions under periodic boundary conditions with non-bond interactions truncated at 10 Å. The simulations consisted of five stages: 1,000 steps minimization; 50 ps run for heating 0.1 to 310 K; 50 ps run under constant pressure of 1 atm and temperature of 310 K; 500 ps run for equilibration; 55 ns run for production. Positional restraints were applied for the first 3 stages. Unrestrained equilibration and production runs were carried out at constant temperature (310 K) using a Langevin thermostat with a collision frequency of 2 ps^{-1} and constant pressure (1 atm) using a Berendsen barostat with a pressure relaxation time of 2 ps. Trajectory was saved at every 10 ps for the later 50 ns of production run, leading to 5,000 in total. MMPBSA.py analyzed the trajectories to calculate the end-state free energy. [10] MD simulations were repeated three times in a complex with different random seeds for statistical analyses.

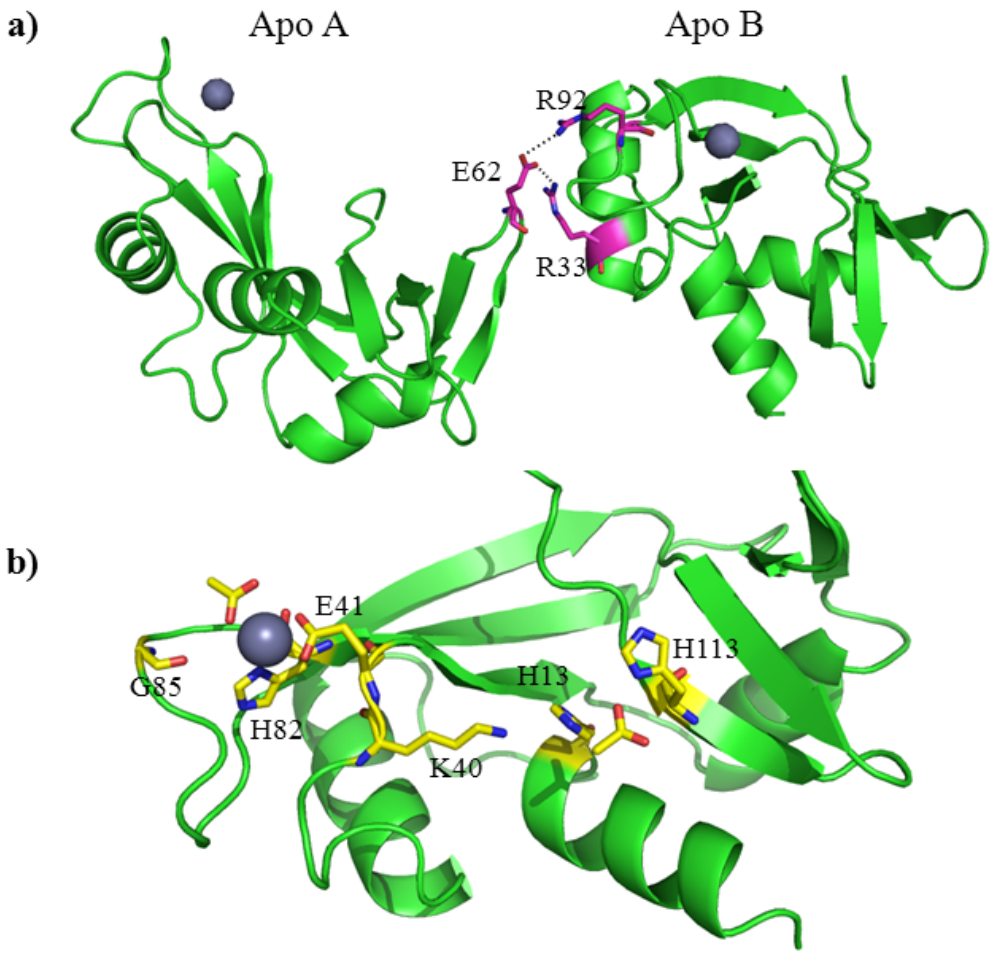

Figure S1. Two apo rAng molecules in the asymmetric unit (a). Each apo rAng molecule, denoted as apo A and apo B, interacts via three residues (magenta sticks). H13, K40, and H113 represent the important residues for ribonucleolytic activity and E41, H82, G85, and acetate are coordinated to the zinc ion (b). One acetate molecule interacts with the imidazole ring of the H113 residue.

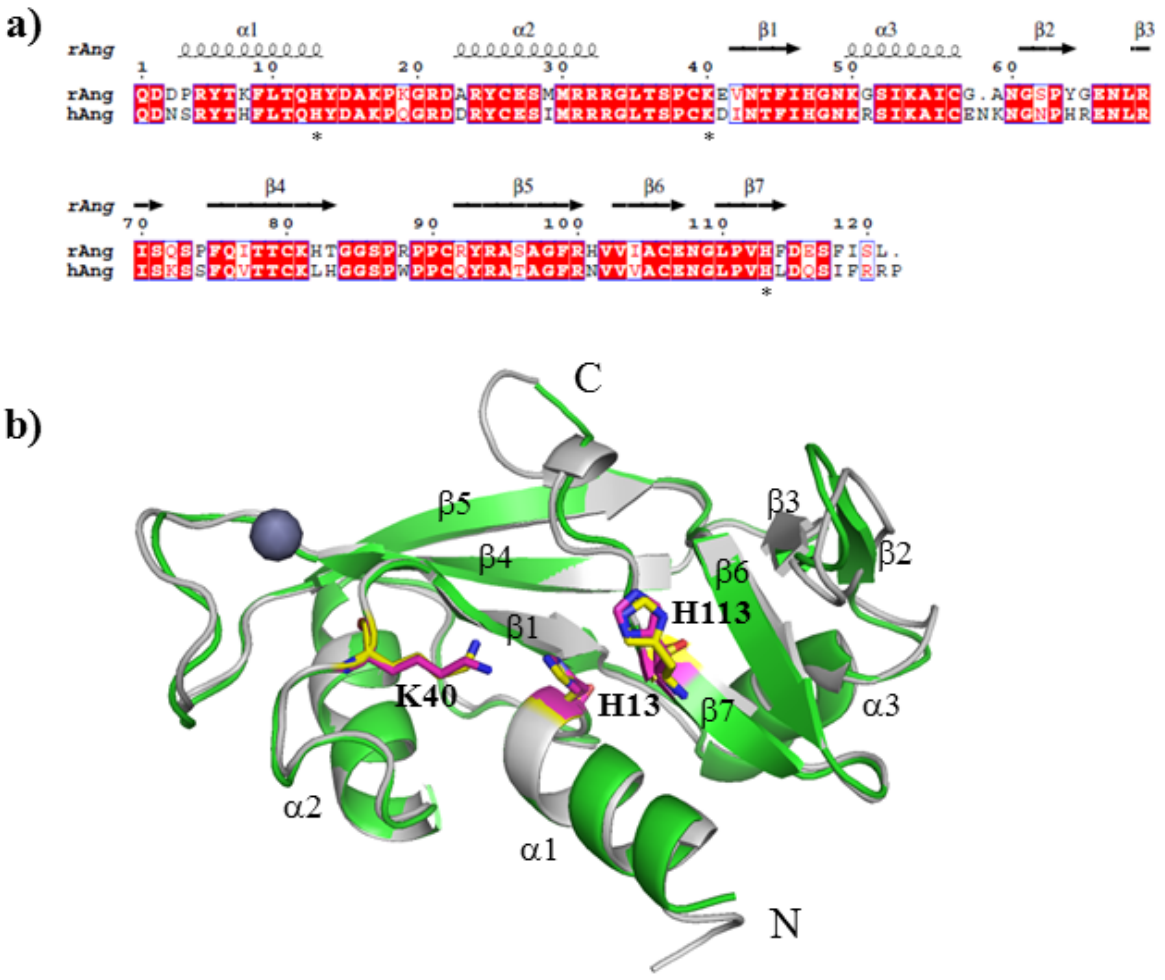

Figure S2. Sequence alignment between rAng and hAng (a) and the overall structure of apo B rAng (green) superposed on hAng (gray) (b). The red blocks denote regions of sequence identity across the two proteins, and the blue boxes indicate partially conserved residues (a). The important residues for ribonucleolytic activity are highlighted with star below the sequence alignment (a) and are represented as sticks (magenta, rAng; yellow, hAng) (b). The gray sphere indicates a zinc ion.

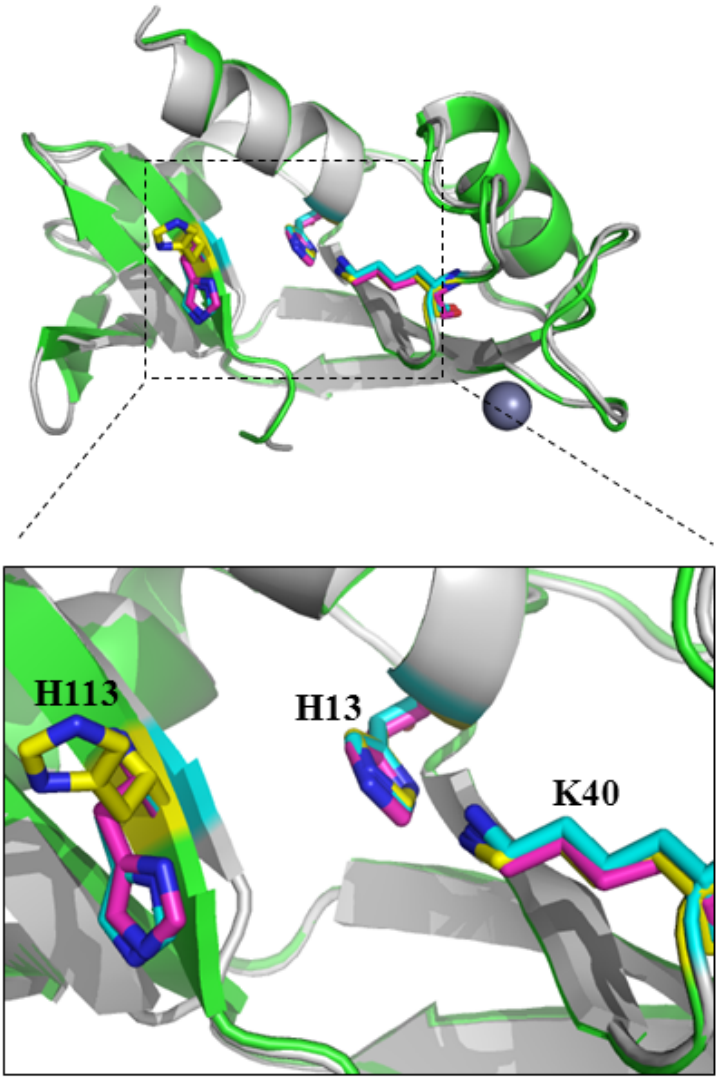

Figure S3. Comparison of the structures of apo rAng (apo A, gray) and the rAng-heparin complex (complex A, green). The enlarged box indicates the active site of each protein. The important residues for ribonucleolytic activity are represented by sticks (yellow for complex A, magenta for complex B, and cyan for apo A).

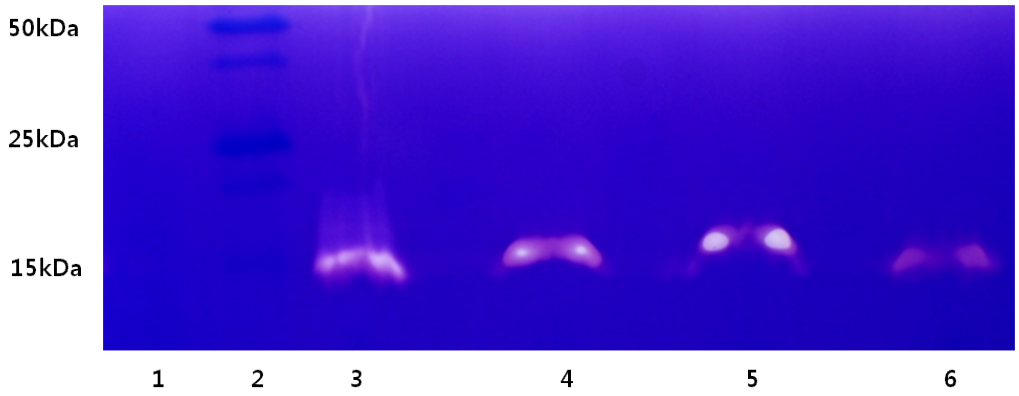

Figure S4. Ribonucleolytic activity of the native rat angiogenin (rAng) versus the double mutants R31A/R32S and K50 L/K54Q of rAng using Zymogram electrophoresis. Lane 1, lysozyme (2 μg); lane 2, molecular weight marker; lane 3, RNaseA (1 ng); lane 4, native rAng (1 μg); lane 5, double mutant R31A and R32S of rAng (1 μg); lane 6, double mutant K50L and K54Q of rAng (1 μg).

Zymogram electrophoresis is a sensitive method for assaying the ribonucleolytic activity of rAng. Because the running gel was co-polymerized with a substrate for rAng, poly(C), regions in the gel that contain ribonucleolytic activity will appear as a white band in a dark purple background. As shown in Figure S4, the double mutant R31A/R32S of rAng (lane 5) has the same enzymatic activity as the native rAng (lane 4) but the double mutant K50L/K54Q (lane 6) has a slightly lower activity than that of the native rAng. In addition, the ribonucleolytic activity of rAng is far lower than that of RNase A (a positive control, lane 3).

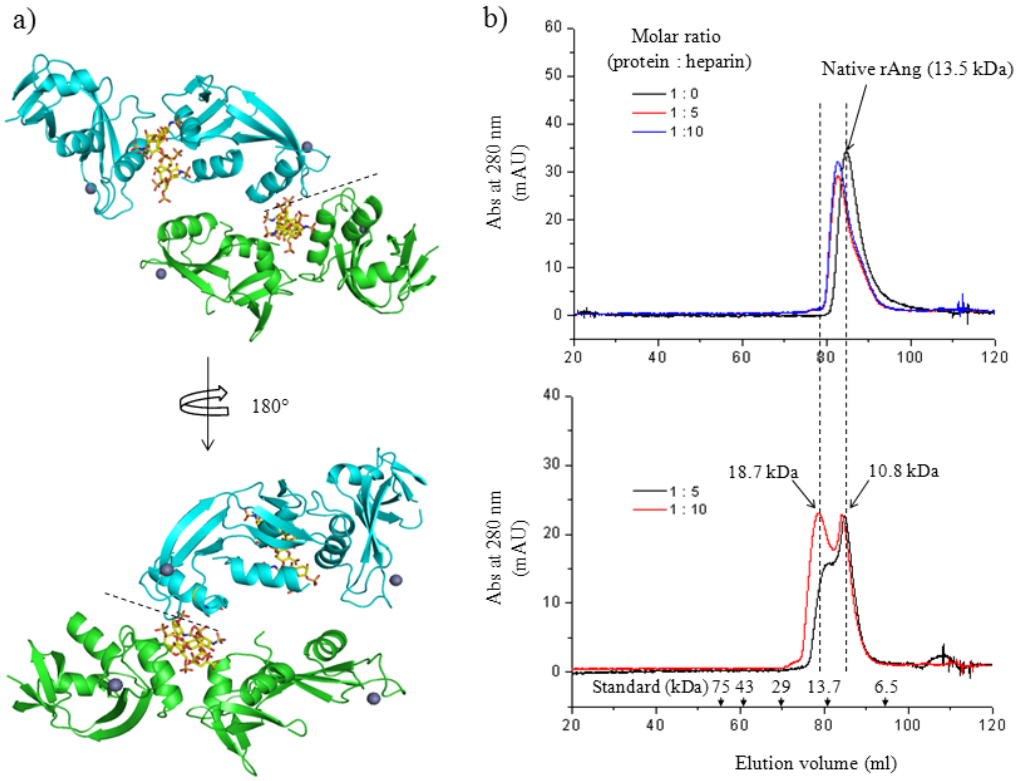

Figure S5. Crystal contacts (dashed line) for the complex (a) and size-exclusion chromatography (SEC) for rAng with varying concentrations of heparin (8 mer) in 20 mM HEPES pH 7.0, 200 mM NaCl (b, upper) and 100 mM NaCl (b, lower) using a prep-grade HiLoad 16/60 Superdex 75 column (GE Healthcare). The molecular weights described in 100 mM NaCl are calculated from standard proteins (b, lower).

Another interaction between rAng and the heparin molecule is observed between asymmetric units; however, it appears that the interaction is the result of crystal packing because tetrameric or oligomeric forms are not detected in solution. The maximum peak is slightly shifted a higher molecular weight in 200 mM NaCl (b, upper) and shows an intermediate size between a monomeric and dimeric form, despite the excess of heparin molecule. The SEC profile for apo rAng (1:0) in 100 mM NaCl (b, lower) could not be obtained because the

protein is adsorbed on the column in the absence of heparin, likely because of nonspecific interactions with the column. Nevertheless, in the presence of heparin, both monomer and dimer are observed in 100 mM NaCl and the dimeric form increases with increasing heparin concentrations (b, lower), indicating that heparin interacts with both the monomer and the dimer and interferes with interactions between rAng and the column. All of these results suggest that rAng exists in a monomer-dimer equilibrium in solution and that the binding affinity of the complex can be affected by salt concentration because the complex forms primarily by ionic interactions.

References

1 C. K. Lee, K. J. Yeo, E. Hwang and H. K. Cheong, *Biomo. NMR Aassign* 2013, **7**, 89-92.

2 Z. Otwinowski and W. Minor, *Methods Enzymol,* 1997, **276**, 307-326.

3 P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta crystallographica Section D Biological crystallography,* 2010, **66**, 213-221.

4 R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thronton, *J Appl Crystallogr,* 1993, **26**, 283-291.

5 S. Iyer, D. E. Holloway and K. R. Acharya, *FEBS J*, 2013, **280**, 302-318.

6 P. A. Leland, K. E. Staniszewski, C. Park, B. R. Kelemen and R. T. Raines, *Biochemistry,* 2002, **41**, 1343-1350.

7 D. A. Case, T. E. Cheatham, 3rd, T. Darden, H. Gohlke, R. Luo, K.M. Merz, Jr, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, *J Comput Chem*, 2005, **26**, 1668-1688.

8 K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. Gonzalez-Outeirino, C.R. Daniels,

B.L. Foley, and R.J. Woods, *J Comput Chem*, 2008, **29**, 622-655.

9 K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, *Proteins*, 2010, **78**, 1950-1958.

10 B.R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, and A. E. Roitberg, *J Chem Theory Comput*, 2012, **8**, 3314-3321.