UAB
Universitat Autònoma de Barcelona

STRUCTURAL ANALYSIS
OF NUCLEOTIDE BINDING SITES
OF ANTIMICROBIAL RIBONUCLEASES

JOSÉ ANTONIO BLANCO BARRERA

Barcelona, 2015

CHAPTER 4
CONCLUSIONS

1. <u>Architecture of nucleotide binding</u>

a) A statistical study of all the nucleotide binding proteins at the Protein Data Bank has enabled the identification of the interaction patterns for base, ribose and phosphate building blocks.

b) Mono- and dinucleotides ligands that mimic the RNA substrate structure were a good model for the study of subsite interactions and provided a reference for RNase A family members with no reported crystal complexes.

c) Structural analyses in the RNase A superfamily revealed a conservation of the main phosphate and base subsites (p_1, B_1) and a series of non conserved substitutions at additional secondary binding sites, in agreement with previously reported kinetic data.

d) The identified common conserved patterns and distinct peculiarities can assist in the design of selective RNA binders.

2. <u>Structural study of RNase A superfamily members by X-ray crystallography</u>

2.1. <u>Structural study of RNase A</u>

    2.1.1. <u>Crystal structure of the RNase A/H7H10 double mutant in complex with 3'-CMP</u>

e) The insertion of two histidines at RNase A p_2 subsite mimics the enzyme active site, conforming an alternative catalytic centre.

f) The creation of a secondary catalytic site at p_2 induced structural changes at the active site which explain the reduction of the enzyme catalytic efficiency as reported by previous kinetic reports.

g) Arg10→His substitution modifies the interactions at the N-terminus, increasing the region mobility. These changes may explain the different observed unit cell packing in comparison to the wild-type protein.

    2.1.2. <u>Atomic resolution crystal structure of RNase A – 3'-CMP</u>

h) The atomic resolution structure of RNase A – 3'CMP (1.16 Å), in comparison to previously reported data at medium resolution (2.20 Å), provided a better visualisation of the enzyme active site conformation when bound to the catalytic reaction product as well as the protonation state of the catalytic triad.

i) The comparison of RNase A – 3'-CMP with the double mutant RNase A/H7H10 complex confirms the p_2 site influence on the residues of subsite p_1.

2.2. Structural study of RNase 3/ECP

2.2.1. Study of active site in mutant crystal structures

j) RNase 3/ECP active site mutations do not alter the overall 3D structure of the protein. Hence, ECP/H15A and ECP/H128N mutants can be used for comparative functional studies.

k) Structural analyses explained the obtained catalytic results, the total enzymatic activity abolishment upon His15→Ala substitution and the residual activity of the His128→Asn mutant.

l) The comparison with the RNase A mutant counterparts provided a plausible explanation for the potential contribution of Lys38 in catalysis in the ECP/H128N mutant.

2.2.2. High-resolution native RNase 3/ECP complex structures

m) The comparison of the two new high resolution structures of sulphate and citrate RNase 3 complexes highlighted the crystal packing differences and identified the main protein anion binding sites.

2.3. Structural study of RNase 6

n) The first crystal structure of human RNase 6 is reported in this work, setting the basis for further structural and functional studies.

o) Sulphate binding sites were identified, providing a reference for the protein nucleotide phosphate binding sites.

p) A new identified region, unique to RNase 6, corresponds to a previously predicted site for heparin by docking experiments.

q) The comparison to the protein binding mode of other characterised RNase A superfamily members will help to analyse the enzyme catalytic efficiency towards nucleotide substrates.

2.4. Overall comparison of RNases' sulphate interaction sites

r) New interaction subsites have been observed for RNase 3/ECP high resolution crystals.

s) A much higher number of sulphate recognition subsites is observed in RNase 3, corroborating its high affinity for negatively charged polymeric molecules, potentially contributing to its higher cytotoxicity in comparison to RNase A or RNase 6.

t) The new identified binding sites may assist the design of nucleotide or sugar derivative analogues for applied therapeutical drug development.

BIBLIOGRAPHY

1. Raines, R. T., Ribonuclease A. *Chem Rev* **1998,** 98, (3), 1045-1066.
2. Loverix, S.; Steyaert, J., Deciphering the mechanism of RNase T1. *Methods Enzymol* **2001,** 341, 305-23.
3. Ball, T. K.; Wasmuth, C. R.; Braunagel, S. C.; Benedik, M. J., Expression of Serratia marcescens extracellular proteins requires recA. *J Bacteriol* **1990,** 172, (1), 342-9.
4. Kiedrowski, M. R.; Kavanaugh, J. S.; Malone, C. L.; Mootz, J. M.; Voyich, J. M.; Smeltzer, M. S.; Bayles, K. W.; Horswill, A. R., Nuclease modulates biofilm formation in community-associated methicillin-resistant Staphylococcus aureus. *PLoS One* **2011,** 6, (11), e26714.
5. Maekawa, K.; Tsunasawa, S.; Dibo, G.; Sakiyama, F., Primary structure of nuclease P1 from Penicillium citrinum. *Eur J Biochem* **1991,** 200, (3), 651-61.
6. Lawal, A.; Jejelowo, O.; Chopra, A. K.; Rosenzweig, J. A., Ribonucleases and bacterial virulence. *Microb Biotechnol* **2011,** 4, (5), 558-71.
7. Saito, Y.; Takeda, J.; Adachi, K.; Nobe, Y.; Kobayashi, J.; Hirota, K.; Oliveira, D. V.; Taoka, M.; Isobe, T., RNase MRP cleaves pre-tRNASer-Met in the tRNA maturation pathway. *PLoS One* **2014,** 9, (11), e112488.
8. Tadokoro, T.; Kanaya, S., Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes. *Febs J* **2009,** 276, (6), 1482-93.
9. Yoda, M.; Cifuentes, D.; Izumi, N.; Sakaguchi, Y.; Suzuki, T.; Giraldez, A. J.; Tomari, Y., Poly(A)-specific ribonuclease mediates 3'-end trimming of Argonaute2-cleaved precursor microRNAs. *Cell Rep* **2013,** 5, (3), 715-26.
10. Zhabokritsky, A.; Kutky, M.; Burns, L. A.; Karran, R. A.; Hudak, K. A., RNA toxins: mediators of stress adaptation and pathogen defense. *Wiley Interdiscip Rev RNA* **2011,** 2, (6), 890-903.
11. Cooper, D. A.; Jha, B. K.; Silverman, R. H.; Hesselberth, J. R.; Barton, D. J., Ribonuclease L and metal-ion-independent endoribonuclease cleavage sites in host and viral RNAs. *Nucleic Acids Res* **2014**.
12. Olmo, N.; Turnay, J.; Gonzalez de Buitrago, G.; Lopez de Silanes, I.; Gavilanes, J. G.; Lizarbe, M. A., Cytotoxic mechanism of the ribotoxin alpha-sarcin. Induction of cell death via apoptosis. *Eur J Biochem* **2001,** 268, (7), 2113-23.
13. Holloway, D. E.; Singh, U. P.; Shogen, K.; Acharya, K. R., Crystal structure of Onconase at 1.1 A resolution--insights into substrate binding and collective motion. *Febs J* **2011,** 278, (21), 4136-49.
14. Rosenberg, H. F., The eosinophil ribonucleases. *Cell Mol Life Sci* **1998,** 54, (8), 795-803.
15. Sorrentino, S., The eight human "canonical" ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. *FEBS Lett* **2010,** 584, (11), 2194-200.
16. Rosenberg, H. F., RNase A ribonucleases and host defense: an evolving story. *J Leukoc Biol* **2008,** 83, (5), 1079-87.
17. Boix, E.; Nogues, M. V., Mammalian antimicrobial proteins and peptides: overview on the RNase A superfamily members involved in innate host defence. *Mol Biosyst* **2007,** 3, (5), 317-35.
18. Kim, J. S.; Soucek, J.; Matousek, J.; Raines, R. T., Catalytic activity of bovine seminal ribonuclease is essential for its immunosuppressive and other biological activities. *Biochem J* **1995,** 308 (Pt 2), 547-50.
19. Zhou, H. M.; Strydom, D. J., The amino acid sequence of human ribonuclease 4, a highly conserved ribonuclease that cleaves specifically on the 3' side of uridine. *Eur J Biochem* **1993,** 217, (1), 401-10.
20. Cruz-Garcia, F.; Hancock, C. N.; McClure, B., S-RNase complexes and pollen rejection. *J Exp Bot* **2003,** 54, (380), 123-30.
21. Deutscher, M. P., Ribonuclease multiplicity, diversity, and complexity. *J Biol Chem* **1993,** 268, (18), 13011-4.
22. Loverix, S.; Steyaert, J., Ribonucleases: from prototypes to therapeutic targets? *Curr Med Chem* **2003,** 10, (9), 779-85.
23. Eaves, G. N.; Jeffries, C. D., Isolation and Properties of an Exocellular Nuclease of Serratia Marcescens. *J Bacteriol* **1963,** 85, (2), 273-8.
24. Aravind, L.; Koonin, E. V., A natural classification of ribonucleases. *Methods Enzymol* **2001,** 341, 3-28.
25. Boix, E.; Leonidas, D. D.; Nikolovski, Z.; Nogues, M. V.; Cuchillo, C. M.; Acharya, K. R., Crystal structure of eosinophil cationic protein at 2.4 A resolution. *Biochemistry* **1999,** 38, (51), 16794-801.
26. Sierakowska, H.; Shugar, D., Mammalian nucleolytic enzymes. *Prog Nucleic Acid Res Mol Biol* **1977,** 20, 59-130.
27. Sorrentino, S.; Libonati, M., Human pancreatic-type and nonpancreatic-type ribonucleases: a direct side-by-side comparison of their catalytic properties. *Arch Biochem Biophys* **1994,** 312, (2), 340-8.
28. Beintema, J. J.; Schuller, C.; Irie, M.; Carsana, A., Molecular evolution of the ribonuclease superfamily. *Prog Biophys Mol Biol* **1988,** 51, (3), 165-92.
29. Ledoux, L., Action of ribonuclease on certain ascites tumours. *Nature* **1955,** 175, (4449), 258-9.
30. Fitch, W. M.; Beintema, J. J., Correcting parsimonious trees for unseen nucleotide substitutions: the effect of dense branching as exemplified by ribonuclease. *Mol Biol Evol* **1990,** 7, (5), 438-43.
31. Cuchillo, C. M.; Nogues, M. V.; Raines, R. T., Bovine pancreatic ribonuclease: fifty years of the first enzymatic reaction mechanism. *Biochemistry* **2011,** 50, (37), 7835-41.
32. Smyth, D. G.; Stein, W. H.; Moore, S., The sequence of amino acid residues in bovine pancreatic ribonuclease: revisions and confirmations. *J Biol Chem* **1963,** 238, 227-34.
33. D'Alessio, G., New and cryptic biological messages from RNases. *Trends Cell Biol* **1993,** 3, (4), 106-9.
34. Beintema, J. J.; Fitch, W. M.; Carsana, A., Molecular evolution of pancreatic-type ribonucleases. *Mol Biol Evol* **1986,** 3, (3), 262-75.

35. Beintema, J. J.; Gaastra, W.; Lenstra, J. A.; Welling, G. W.; Fitch, W. M., The molecular evolution of pancreatic ribonuclease. *J Mol Evol* **1977,** 10, (1), 49-71.

36. Liu, J.; Li, J.; Wang, H.; Zhang, C.; Li, N.; Lin, Y.; Liu, J.; Wang, W., Cloning, expression and location of RNase9 in human epididymis. *BMC Res Notes* **2008,** 1, 111.

37. Dyer, K. D.; Rosenberg, H. F., The RNase a superfamily: generation of diversity and innate host defense. *Mol Divers* **2006,** 10, (4), 585-97.

38. Cho, S.; Beintema, J. J.; Zhang, J., The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. *Genomics* **2005,** 85, (2), 208-20.

39. Gouet, P.; Courcelle, E.; Stuart, D. I.; Metoz, F., ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **1999,** 15, (4), 305-8.

40. Pous, J.; Canals, A.; Terzyan, S. S.; Guasch, A.; Benito, A.; Ribo, M.; Vilanova, M.; Coll, M., Three-dimensional structure of a human pancreatic ribonuclease variant, a step forward in the design of cytotoxic ribonucleases. *J Mol Biol* **2000,** 303, (1), 49-60.

41. Swaminathan, G. J.; Holloway, D. E.; Veluraja, K.; Acharya, K. R., Atomic resolution (0.98 A) structure of eosinophil-derived neurotoxin. *Biochemistry* **2002,** 41, (10), 3341-52.

42. Terzyan, S. S.; Peracaula, R.; de Llorens, R.; Tsushima, Y.; Yamada, H.; Seno, M.; Gomis-Ruth, F. X.; Coll, M., The three-dimensional structure of human RNase 4, unliganded and complexed with d(Up), reveals the basis for its uridine selectivity. *J Mol Biol* **1999,** 285, (1), 205-14.

43. Leonidas, D. D.; Chavali, G. B.; Jardine, A. M.; Li, S.; Shapiro, R.; Acharya, K. R., Binding of phosphate and pyrophosphate ions at the active site of human angiogenin as revealed by X-ray crystallography. *Protein Sci* **2001,** 10, (8), 1669-76.

44. Huang, Y. C.; Lin, Y. M.; Chang, T. W.; Wu, S. J.; Lee, Y. S.; Chang, M. D.; Chen, C.; Wu, S. H.; Liao, Y. D., The flexible and clustered lysine residues of human ribonuclease 7 are critical for membrane permeability and antimicrobial activity. *J Biol Chem* **2007,** 282, (7), 4626-33.

45. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res* **2000,** 28, (1), 235-42.

46. Beintema, J. J.; Wietzes, P.; Weickmann, J. L.; Glitz, D. G., The amino acid sequence of human pancreatic ribonuclease. *Anal Biochem* **1984,** 136, (1), 48-64.

47. Weickmann, J. L.; Elson, M.; Glitz, D. G., Purification and characterization of human pancreatic ribonuclease. *Biochemistry* **1981,** 20, (5), 1272-8.

48. Landre, J. B.; Hewett, P. W.; Olivot, J. M.; Friedl, P.; Ko, Y.; Sachinidis, A.; Moenner, M., Human endothelial cells selectively express large amounts of pancreatic-type ribonuclease (RNase 1). *J Cell Biochem* **2002,** 86, (3), 540-52.

49. Sorrentino, S.; Libonati, M., Structure-function relationships in human ribonucleases: main distinctive features of the major RNase types. *FEBS Lett* **1997,** 404, (1), 1-5.

50. Gaur, D.; Swaminathan, S.; Batra, J. K., Interaction of human pancreatic ribonuclease with human ribonuclease inhibitor. Generation of inhibitor-resistant cytotoxic variants. *J Biol Chem* **2001,** 276, (27), 24978-84.

51. Zewe, M.; Rybak, S. M.; Dubel, S.; Coy, J. F.; Welschof, M.; Newton, D. L.; Little, M., Cloning and cytotoxicity of a human pancreatic RNase immunofusion. *Immunotechnology* **1997,** 3, (2), 127-36.

52. Suzuki, M.; Saxena, S. K.; Boix, E.; Prill, R. J.; Vasandani, V. M.; Ladner, J. E.; Sung, C.; Youle, R. J., Engineering receptor-mediated cytotoxicity into human ribonucleases by steric blockade of inhibitor interaction. *Nat Biotechnol* **1999,** 17, (3), 265-70.

53. Piccoli, R.; Di Gaetano, S.; De Lorenzo, C.; Grauso, M.; Monaco, C.; Spalletti-Cernia, D.; Laccetti, P.; Cinatl, J.; Matousek, J.; D'Alessio, G., A dimeric mutant of human pancreatic ribonuclease with selective cytotoxicity toward malignant cells. *Proc Natl Acad Sci U S A* **1999,** 96, (14), 7768-73.

54. Gleich, G. J.; Loegering, D. A.; Bell, M. P.; Checkel, J. L.; Ackerman, S. J.; McKean, D. J., Biochemical and functional similarities between human eosinophil-derived neurotoxin and eosinophil cationic protein: homology with ribonuclease. *Proc Natl Acad Sci U S A* **1986,** 83, (10), 3146-50.

55. Boix, E.; Torrent, M.; Sanchez, D.; Nogues, M. V., The antipathogen activities of eosinophil cationic protein. *Curr Pharm Biotechnol* **2008,** 9, (3), 141-52.

56. Venge, P.; Bystrom, J.; Carlson, M.; Hakansson, L.; Karawacjzyk, M.; Peterson, C.; Seveus, L.; Trulson, A., Eosinophil cationic protein (ECP): molecular and biological properties and the use of ECP as a marker of eosinophil activation in disease. *Clin Exp Allergy* **1999,** 29, (9), 1172-86.

57. Bystrom, J.; Amin, K.; Bishop-Bailey, D., Analysing the eosinophil cationic protein--a clue to the function of the eosinophil granulocyte. *Respir Res* **2011,** 12, 10.

58. Rosenberg, H. F.; Dyer, K. D., Eosinophil cationic protein and eosinophil-derived neurotoxin. Evolution of novel function in a primate ribonuclease gene family. *J Biol Chem* **1995,** 270, (50), 30234.

59. Boix, E.; Nikolovski, Z.; Moiseyev, G. P.; Rosenberg, H. F.; Cuchillo, C. M.; Nogues, M. V., Kinetic and product distribution analysis of human eosinophil cationic protein indicates a subsite arrangement that favors exonuclease-type activity. *J Biol Chem* **1999,** 274, (22), 15605-14.

60. Zhang, J.; Rosenberg, H. F.; Nei, M., Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proc Natl Acad Sci U S A* **1998,** 95, (7), 3708-13.

61. Mallorqui-Fernandez, G.; Pous, J.; Peracaula, R.; Aymami, J.; Maeda, T.; Tada, H.; Yamada, H.; Seno, M.; de Llorens, R.; Gomis-Ruth, F. X.; Coll, M., Three-dimensional crystal structure of human eosinophil cationic protein (RNase 3) at 1.75 A resolution. *J Mol Biol* **2000,** 300, (5), 1297-307.

62. Torrent, M.; Cuyas, E.; Carreras, E.; Navarro, S.; Lopez, O.; de la Maza, A.; Nogues, M. V.; Reshetnyak, Y. K.; Boix, E., Topography studies on the membrane interaction mechanism of the eosinophil cationic protein. *Biochemistry* **2007,** 46, (3), 720-33.

63. Boix, E.; Salazar, V. A.; Torrent, M.; Pulido, D.; Nogues, M. V.; Moussaoui, M., Structural determinants of the eosinophil cationic protein antimicrobial activity. *Biol Chem* **2012,** 393, (8), 801-15.

64. Trulson, A.; Bystrom, J.; Engstrom, A.; Larsson, R.; Venge, P., The functional heterogeneity of eosinophil cationic protein is determined by a gene polymorphism and post-translational modifications. *Clin Exp Allergy* **2007,** 37, (2), 208-18.

65. Boix, E.; Pulido, D.; Moussaoui, M.; Nogues, M. V.; Russi, S., The sulfate-binding site structure of the human eosinophil cationic protein as revealed by a new crystal form. *J Struct Biol* **2012,** 179, (1), 1-9.

66. Maeda, T.; Mahara, K.; Kitazoe, M.; Futami, J.; Takidani, A.; Kosaka, M.; Tada, H.; Seno, M.; Yamada, H., RNase 3 (ECP) is an extraordinarily stable protein among human pancreatic-type RNases. *J Biochem* **2002,** 132, (5), 737-42.

67. Nikolovski, Z.; Buzon, V.; Ribo, M.; Moussaoui, M.; Vilanova, M.; Cuchillo, C. M.; Cladera, J.; Nogues, M. V., Thermal unfolding of eosinophil cationic protein/ribonuclease 3: a nonreversible process. *Protein Sci* **2006,** 15, (12), 2816-27.

68. Mohan, C. G.; Boix, E.; Evans, H. R.; Nikolovski, Z.; Nogues, M. V.; Cuchillo, C. M.; Acharya, K. R., The crystal structure of eosinophil cationic protein in complex with 2',5'-ADP at 2.0 A resolution reveals the details of the ribonucleolytic active site. *Biochemistry* **2002,** 41, (40), 12100-6.

69. Fontecilla-Camps, J. C.; de Llorens, R.; le Du, M. H.; Cuchillo, C. M., Crystal structure of ribonuclease A.d(ApTpApApG) complex. Direct evidence for extended substrate recognition. *J Biol Chem* **1994,** 269, (34), 21526-31.

70. Mosimann, S. C.; Newton, D. L.; Youle, R. J.; James, M. N., X-ray crystallographic structure of recombinant eosinophil-derived neurotoxin at 1.83 A resolution. *J Mol Biol* **1996,** 260, (4), 540-52.

71. Leonidas, D. D.; Boix, E.; Prill, R.; Suzuki, M.; Turton, R.; Minson, K.; Swaminathan, G. J.; Youle, R. J.; Acharya, K. R., Mapping the ribonucleolytic active site of eosinophil-derived neurotoxin (EDN). High resolution crystal structures of EDN complexes with adenylic nucleotide inhibitors. *J Biol Chem* **2001,** 276, (18), 15009-17.

72. Baker, M. D.; Holloway, D. E.; Swaminathan, G. J.; Acharya, K. R., Crystal structures of eosinophil-derived neurotoxin (EDN) in complex with the inhibitors 5'-ATP, Ap3A, Ap4A, and Ap5A. *Biochemistry* **2006,** 45, (2), 416-26.

73. Shapiro, R.; Fett, J. W.; Strydom, D. J.; Vallee, B. L., Isolation and characterization of a human colon carcinoma-secreted enzyme with pancreatic ribonuclease-like activity. *Biochemistry* **1986,** 25, (23), 7255-64.

74. Shapiro, R.; Strydom, D. J.; Olson, K. A.; Vallee, B. L., Isolation of angiogenin from normal human plasma. *Biochemistry* **1987,** 26, (16), 5141-6.

75. Moenner, M.; Gusse, M.; Hatzi, E.; Badet, J., The widespread expression of angiogenin in different human cells suggests a biological function not only related to angiogenesis. *Eur J Biochem* **1994,** 226, (2), 483-90.

76. Bond, M. D.; Strydom, D. J.; Vallee, B. L., Characterization and sequencing of rabbit, pig and mouse angiogenins: discernment of functionally important residues and regions. *Biochim Biophys Acta* **1993,** 1162, (1-2), 177-86.

77. Gupta, S. K.; Haigh, B. J.; Griffin, F. J.; Wheeler, T. T., The mammalian secreted RNases: mechanisms of action in host defence. *Innate Immun* **2013,** 19, (1), 86-97.

78. Strydom, D. J.; Fett, J. W.; Lobb, R. R.; Alderman, E. M.; Bethune, J. L.; Riordan, J. F.; Vallee, B. L., Amino acid sequence of human tumor derived angiogenin. *Biochemistry* **1985,** 24, (20), 5486-94.

79. Harper, J. W.; Vallee, B. L., A covalent angiogenin/ribonuclease hybrid with a fourth disulfide bond generated by regional mutagenesis. *Biochemistry* **1989,** 28, (4), 1875-84.

80. Russo, N.; Shapiro, R.; Acharya, K. R.; Riordan, J. F.; Vallee, B. L., Role of glutamine-117 in the ribonucleolytic activity of human angiogenin. *Proc Natl Acad Sci U S A* **1994,** 91, (8), 2920-4.

81. Riordan, J. F., Angiogenin. *Methods Enzymol* **2001,** 341, 263-73.

82. Shapiro, R.; Vallee, B. L., Site-directed mutagenesis of histidine-13 and histidine-114 of human angiogenin. Alanine derivatives inhibit angiogenin-induced angiogenesis. *Biochemistry* **1989,** 28, (18), 7401-8.

83. Hu, G. F.; Riordan, J. F.; Vallee, B. L., A putative angiogenin receptor in angiogenin-responsive human endothelial cells. *Proc Natl Acad Sci U S A* **1997,** 94, (6), 2204-9.

84. Gao, X.; Xu, Z., Mechanisms of action of angiogenin. *Acta Biochim Biophys Sin (Shanghai)* **2008,** 40, (7), 619-24.

85. Rosenberg, H. F.; Dyer, K. D., Molecular cloning and characterization of a novel human ribonuclease (RNase k6): increasing diversity in the enlarging ribonuclease gene family. *Nucleic Acids Res* **1996,** 24, (18), 3507-13.

86. Deming, M. S.; Dyer, K. D.; Bankier, A. T.; Piper, M. B.; Dear, P. H.; Rosenberg, H. F., Ribonuclease k6: chromosomal mapping and divergent rates of evolution within the RNase A gene superfamily. *Genome Res* **1998,** 8, (6), 599-607.

87. Becknell, B.; Eichler, T. E.; Beceiro, S.; Li, B.; Easterling, R. S.; Carpenter, A. R.; James, C. L.; McHugh, K. M.; Hains, D. S.; Partida-Sanchez, S.; Spencer, J. D., Ribonucleases 6 and 7 have antimicrobial function in the human and murine urinary tract. *Kidney Int* **2014**.

88. Harder, J.; Schroder, J. M., RNase 7, a novel innate immune defense antimicrobial protein of healthy human skin. *J Biol Chem* **2002,** 277, (48), 46779-84.

89. Zhang, J.; Dyer, K. D.; Rosenberg, H. F., Human RNase 7: a new cationic ribonuclease of the RNase A superfamily. *Nucleic Acids Res* **2003,** 31, (2), 602-7.

90. Harder, J.; Glaser, R.; Schroder, J. M., Human antimicrobial proteins effectors of innate immunity. *J Endotoxin Res* **2007,** 13, (6), 317-38.

91. Zasloff, M., Antimicrobial RNases of human skin. *J Invest Dermatol* **2009,** 129, (9), 2091-3.

92. Torrent, M.; Sanchez, D.; Buzon, V.; Nogues, M. V.; Cladera, J.; Boix, E., Comparison of the membrane interaction mechanism of two antimicrobial RNases: RNase 3/ECP and RNase 7. *Biochim Biophys Acta* **2009,** 1788, (5), 1116-25.

93. Zhang, J.; Dyer, K. D.; Rosenberg, H. F., RNase 8, a novel RNase A superfamily ribonuclease expressed uniquely in placenta. *Nucleic Acids Res* **2002,** 30, (5), 1169-75.

94. Chan, C. C.; Moser, J. M.; Dyer, K. D.; Percopo, C. M.; Rosenberg, H. F., Genetic diversity of human RNase 8. *BMC Genomics* **2012,** 13, 40.

95. Rudolph, B.; Podschun, R.; Sahly, H.; Schubert, S.; Schroder, J. M.; Harder, J., Identification of RNase 8 as a novel human antimicrobial protein. *Antimicrob Agents Chemother* **2006,** 50, (9), 3194-6.

96. Findlay, D.; Herries, D. G.; Mathias, A. P.; Rabin, B. R.; Ross, C. A., The active site and mechanism of action of bovine pancreatic ribonuclease. 7. The catalytic mechanism. *Biochem J* **1962,** 85, (1), 152-3.

97. Crestfield, A. M.; Stein, W. H.; Moore, S., Alkylation and identification of the histidine residues at the active site of ribonuclease. *J Biol Chem* **1963,** 238, 2413-9.

98. Chatani, E.; Hayashi, R., Functional and structural roles of constituent amino acid residues of bovine pancreatic ribonuclease A. *J Biosci Bioeng* **2001,** 92, (2), 98-107.

99. Kartha, G., Tertiary structure of ribonuclease. *Nature* **1967,** 214, (5085), 234 passim.

100. Avey, H. P.; Boles, M. O.; Carlisle, C. H.; Evans, S. A.; Morris, S. J.; Palmer, R. A.; Woolhouse, B. A.; Shall, S., Structure of ribonuclease. *Nature* **1967,** 213, (5076), 557-62.

101. Soucek, J.; Raines, R. T.; Haugg, M.; Raillard-Yoon, S. A.; Benner, S. A., Structural changes to ribonuclease A and their effects on biological activity. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* **1999,** 123, (2), 103-11.

102. Sica, F.; Pica, A.; Merlino, A.; Russo Krauss, I.; Ercole, C.; Picone, D., The multiple forms of bovine seminal ribonuclease: structure and stability of a C-terminal swapped dimer. *FEBS Lett* **2013,** 587, (23), 3755-62.

103. Sikriwal, D.; Seth, D.; Batra, J. K., Role of catalytic and non-catalytic subsite residues in ribonuclease activity of human eosinophil-derived neurotoxin. *Biol Chem* **2009,** 390, (3), 225-34.

104. Panov, K. I.; Kolbanovskaya, E. Y.; Okorokov, A. L.; Panova, T. B.; Terwisscha van Scheltinga, A. C.; Karpeisky, M.; Beintema, J. J., Ribonuclease A mutant His119 Asn: the role of histidine in catalysis. *FEBS Lett* **1996,** 398, (1), 57-60.

105. Schultz, L. W.; Quirk, D. J.; Raines, R. T., His...Asp catalytic dyad of ribonuclease A: structure and function of the wild-type, D121N, and D121A enzymes. *Biochemistry* **1998,** 37, (25), 8886-98.

106. Cuchillo, C. M.; Pares, X.; Guasch, A.; Barman, T.; Travers, F.; Nogues, M. V., The role of 2',3'-cyclic phosphodiesters in the bovine pancreatic ribonuclease A catalysed cleavage of RNA: intermediates or products? *FEBS Lett* **1993,** 333, (3), 207-10.

107. Nogues, M. V.; Moussaoui, M.; Boix, E.; Vilanova, M.; Ribo, M.; Cuchillo, C. M., The contribution of noncatalytic phosphate-binding subsites to the mechanism of bovine pancreatic ribonuclease A. *Cell Mol Life Sci* **1998,** 54, (8), 766-74.

108. Birdsall, D. L.; McPherson, A., Crystal structure disposition of thymidylic acid tetramer in complex with ribonuclease A. *J Biol Chem* **1992,** 267, (31), 22230-6.

109. McPherson, A.; Brayer, G. D.; Morrison, R. D., Crystal structure of RNase A complexed with d(pA)4. *J Mol Biol* **1986,** 189, (2), 305-27.

110. Boix, E.; Nogues, M. V.; Schein, C. H.; Benner, S. A.; Cuchillo, C. M., Reverse transphosphorylation by ribonuclease A needs an intact p2-binding site. Point mutations at Lys-7 and Arg-10 alter the catalytic properties of the enzyme. *J Biol Chem* **1994,** 269, (4), 2529-34.

111. Moussaoui, M.; Cuchillo, C. M.; Nogues, M. V., A phosphate-binding subsite in bovine pancreatic ribonuclease A can be converted into a very efficient catalytic site. *Protein Sci* **2007,** 16, (1), 99-109.

112. Fisher, B. M.; Grilley, J. E.; Raines, R. T., A new remote subsite in ribonuclease A. *J Biol Chem* **1998,** 273, (51), 34134-8.

113. Rodriguez, M.; Moussaoui, M.; Benito, A.; Cuchillo, C. M.; Nogues, M. V.; Vilanova, M., Human pancreatic ribonuclease presents higher endonucleolytic activity than ribonuclease A. *Arch Biochem Biophys* **2008,** 471, (2), 191-7.

114. Gagne, D.; Charest, L. A.; Morin, S.; Kovrigin, E. L.; Doucet, N., Conservation of flexible residue clusters among structural and functional enzyme homologues. *J Biol Chem* **2012,** 287, (53), 44289-300.

115. Gherardini, P. F.; Ausiello, G.; Russell, R. B.; Helmer-Citterich, M., Modular architecture of nucleotide-binding pockets. *Nucleic Acids Res* **2010,** 38, (11), 3809-16.

116. Brakoulias, A.; Jackson, R. M., Towards a structural classification of phosphate binding sites in protein-nucleotide complexes: an automated all-against-all structural comparison using geometric matching. *Proteins* **2004,** 56, (2), 250-60.

117. Mandel-Gutfreund, Y.; Schueler, O.; Margalit, H., Comprehensive analysis of hydrogen bonds in regulatory protein DNA-complexes: in search of common principles. *J Mol Biol* **1995,** 253, (2), 370-82.

118. Coulocheri, S. A.; Pigis, D. G.; Papavassiliou, K. A.; Papavassiliou, A. G., Hydrogen bonds in protein-DNA complexes: where geometry meets plasticity. *Biochimie* **2007,** 89, (11), 1291-303.

119. Morozova, N.; Allers, J.; Myers, J.; Shamoo, Y., Protein-RNA interactions: exploring binding patterns with a three-dimensional superposition analysis of high resolution structures. *Bioinformatics* **2006,** 22, (22), 2746-52.

120. Kondo, J.; Westhof, E., Classification of pseudo pairs between nucleotide bases and amino acids by analysis of nucleotide-protein complexes. *Nucleic Acids Res* **2011,** 39, (19), 8628-37.

121. Denessiouk, K. A.; Johnson, M. S., "Acceptor-donor-acceptor" motifs recognize the Watson-Crick, Hoogsteen and Sugar "donor-acceptor-donor" edges of adenine and adenosine-containing ligands. *J Mol Biol* **2003,** 333, (5), 1025-43.

122. Allers, J.; Shamoo, Y., Structure-based analysis of protein-RNA interactions using the program ENTANGLE. *J Mol Biol* **2001,** 311, (1), 75-86.

123. Terribilini, M.; Sander, J. D.; Lee, J. H.; Zaback, P.; Jernigan, R. L.; Honavar, V.; Dobbs, D., RNABindR: a server for analyzing and predicting RNA-binding sites in proteins. *Nucleic Acids Res* **2007,** 35, (Web Server issue), W578-84.

124. Draper, D. E., Themes in RNA-protein recognition. *J Mol Biol* **1999,** 293, (2), 255-70.

125. Steyaert, J., A decade of protein engineering on ribonuclease T1--atomic dissection of the enzyme-substrate interactions. *Eur J Biochem* **1997,** 247, (1), 1-11.

126. Zegers, I.; Loris, R.; Dehollander, G.; Fattah Haikal, A.; Poortmans, F.; Steyaert, J.; Wyns, L., Hydrolysis of a slow cyclic thiophosphate substrate of RNase T1 analyzed by time-resolved crystallography. *Nat Struct Biol* **1998,** 5, (4), 280-3.

127. Yakovlev, G. I.; Moiseyev, G. P.; Bezborodova, S. I.; Both, V.; Sevcik, J., A comparative study on the catalytic properties of guanyl-specific ribonucleases. *Eur J Biochem* **1992,** 204, (1), 187-90.

128. Torrent, M.; Badia, M.; Moussaoui, M.; Sanchez, D.; Nogues, M. V.; Boix, E., Comparison of human RNase 3 and RNase 7 bactericidal action at the Gram-negative and Gram-positive bacterial cell wall. *Febs J* **2010,** 277, (7), 1713-25.

129. Torrent, M.; Pulido, D.; Nogues, M. V.; Boix, E., Exploring new biological functions of amyloids: bacteria cell agglutination mediated by host protein aggregation. *PLoS Pathog* **2012,** 8, (11), e1003005.

130. Torrent, M.; Pulido, D.; Valle, J.; Nogues, M. V.; Andreu, D.; Boix, E., Ribonucleases as a host-defence family: evidence of evolutionarily conserved antimicrobial activity at the N-terminus. *Biochem J* **2013,** 456, (1), 99-108.

131. Torrent, M.; Odorizzi, F.; Nogues, M. V.; Boix, E., Eosinophil cationic protein aggregation: identification of an N-terminus amyloid prone region. *Biomacromolecules* **2010,** 11, (8), 1983-90.

132. Torrent, M.; Pulido, D.; de la Torre, B. G.; Garcia-Mayoral, M. F.; Nogues, M. V.; Bruix, M.; Andreu, D.; Boix, E., Refining the eosinophil cationic protein antibacterial pharmacophore by rational structure minimization. *J Med Chem* **2011,** 54, (14), 5237-44.

133. Boix, E.; Wu, Y.; Vasandani, V. M.; Saxena, S. K.; Ardelt, W.; Ladner, J.; Youle, R. J., Role of the N terminus in RNase A homologues: differences in catalytic activity, ribonuclease inhibitor interaction and cytotoxicity. *J Mol Biol* **1996,** 257, (5), 992-1007.

134. Pal-Bhowmick, I.; Pandey, R. P.; Jarori, G. K.; Kar, S.; Sahal, D., Structural and functional studies on Ribonuclease S, retro S and retro-inverso S peptides. *Biochem Biophys Res Commun* **2007,** 364, (3), 608-13.

135. Andreu, D.; Rivas, L., Animal antimicrobial peptides: an overview. *Biopolymers* **1998,** 47, (6), 415-33.

136. Torrent, M.; de la Torre, B. G.; Nogues, V. M.; Andreu, D.; Boix, E., Bactericidal and membrane disruption activities of the eosinophil cationic protein are largely retained in an N-terminal fragment. *Biochem J* **2009,** 421, (3), 425-34.

137. Malik, A.; Batra, J. K., Antimicrobial activity of human eosinophil granule proteins: involvement in host defence against pathogens. *Crit Rev Microbiol* **2012,** 38, (2), 168-81.

138. Gleich, G. J., Mechanisms of eosinophil-associated inflammation. *J Allergy Clin Immunol* **2000,** 105, (4), 651-63.

139. Giembycz, M. A.; Lindsay, M. A., Pharmacology of the eosinophil. *Pharmacol Rev* **1999,** 51, (2), 213-340.

140. Sher, A.; Smithers, S. R.; MacKenzie, P.; Broomfield, K., Schistosoma mansoni: immunoglobulins involved in passive immunization of laboratory mice. *Exp Parasitol* **1977,** 41, (1), 160-6.

141. Ramalho-Pinto, F. J.; McLaren, D. J.; Smithers, S. R., Complement-mediated killing of schistosomula of Schistosoma mansoni by rat eosinophils in vitro. *J Exp Med* **1978,** 147, (1), 147-56.

142. McLaren, D. J.; McKean, J. R.; Olsson, I.; Venges, P.; Kay, A. B., Morphological studies on the killing of schistosomula of Schistosoma mansoni by human eosinophil and neutrophil cationic proteins in vitro. *Parasite Immunol* **1981,** 3, (4), 359-73.

143. Ackerman, S. J.; Gleich, G. J.; Loegering, D. A.; Richardson, B. A.; Butterworth, A. E., Comparative toxicity of purified human eosinophil granule cationic proteins for schistosomula of Schistosoma mansoni. *Am J Trop Med Hyg* **1985,** 34, (4), 735-45.

144. Molina, H. A.; Kierszenbaum, F.; Hamann, K. J.; Gleich, G. J., Toxic effects produced or mediated by human eosinophil granule components on Trypanosoma cruzi. *Am J Trop Med Hyg* **1988,** 38, (2), 327-34.

145. Hamann, K. J.; Barker, R. L.; Loegering, D. A.; Gleich, G. J., Comparative toxicity of purified human eosinophil granule proteins for newborn larvae of Trichinella spiralis. *J Parasitol* **1987,** 73, (3), 523-9.

146. Hamann, K. J.; Gleich, G. J.; Checkel, J. L.; Loegering, D. A.; McCall, J. W.; Barker, R. L., In vitro killing of microfilariae of Brugia pahangi and Brugia malayi by eosinophil granule proteins. *J Immunol* **1990,** 144, (8), 3166-73.

147. Carreras, E.; Boix, E.; Rosenberg, H. F.; Cuchillo, C. M.; Nogues, M. V., Both aromatic and cationic residues contribute to the membrane-lytic and bactericidal activity of eosinophil cationic protein. *Biochemistry* **2003,** 42, (22), 6636-44.

148. Koh, G. C.; Shek, L. P.; Goh, D. Y.; Van Bever, H.; Koh, D. S., Eosinophil cationic protein: is it useful in asthma? A systematic review. *Respir Med* **2007,** 101, (4), 696-705.

149. Hallgren, R.; Terent, A.; Venge, P., Eosinophil cationic protein (ECP) in the cerebrospinal fluid. *J Neurol Sci* **1983,** 58, (1), 57-71.

150. Chusid, M. J.; Dale, D. C.; West, B. C.; Wolff, S. M., The hypereosinophilic syndrome: analysis of fourteen cases with review of the literature. *Medicine (Baltimore)* **1975,** 54, (1), 1-27.

151. Sorrentino, S.; Glitz, D. G.; Hamann, K. J.; Loegering, D. A.; Checkel, J. L.; Gleich, G. J., Eosinophil-derived neurotoxin and human liver ribonuclease. Identity of structure and linkage of neurotoxicity to nuclease activity. *J Biol Chem* **1992,** 267, (21), 14859-65.

152. Weidenmaier, C.; Kristian, S. A.; Peschel, A., Bacterial resistance to antimicrobial host defenses--an emerging target for novel antiinfective strategies? *Curr Drug Targets* **2003,** 4, (8), 643-9.

153. Wiesner, J.; Vilcinskas, A., Antimicrobial peptides: the ancient arm of the human immune system. *Virulence* **2010,** 1, (5), 440-64.

154. Ran, S.; Downes, A.; Thorpe, P. E., Increased exposure of anionic phospholipids on the surface of tumor blood vessels. *Cancer Res* **2002,** 62, (21), 6132-40.

155. Gasset, M.; Martinez del Pozo, A.; Onaderra, M.; Gavilanes, J. G., Study of the interaction between the antitumour protein alpha-sarcin and phospholipid vesicles. *Biochem J* **1989,** 258, (2), 569-75.

156. Nitta, K.; Ozaki, K.; Ishikawa, M.; Furusawa, S.; Hosono, M.; Kawauchi, H.; Sasaki, K.; Takayanagi, Y.; Tsuiki, S.; Hakomori, S., Inhibition of cell proliferation by Rana catesbeiana and Rana japonica lectins belonging to the ribonuclease superfamily. *Cancer Res* **1994,** 54, (4), 920-7.

157. Leland, P. A.; Staniszewski, K. E.; Kim, B. M.; Raines, R. T., Endowing human pancreatic ribonuclease with toxicity for cancer cells. *J Biol Chem* **2001,** 276, (46), 43095-102.

158. Mancheno, J. M.; Gasset, M.; Onaderra, M.; Gavilanes, J. G.; D'Alessio, G., Bovine seminal ribonuclease destabilizes negatively charged membranes. *Biochem Biophys Res Commun* **1994,** 199, (1), 119-24.

159. Dickson, K. A.; Haigis, M. C.; Raines, R. T., Ribonuclease inhibitor: structure and function. *Prog Nucleic Acid Res Mol Biol* **2005,** 80, 349-74.

160. Sandvig, K.; van Deurs, B., Transport of protein toxins into cells: pathways used by ricin, cholera toxin and Shiga toxin. *FEBS Lett* **2002,** 529, (1), 49-53.

161. Bracale, A.; Spalletti-Cernia, D.; Mastronicola, M.; Castaldi, F.; Mannucci, R.; Nitsch, L.; D'Alessio, G., Essential stations in the intracellular pathway of cytotoxic bovine seminal ribonuclease. *Biochem J* **2002,** 362, (Pt 3), 553-60.

162. Dickson, K. A.; Kang, D. K.; Kwon, Y. S.; Kim, J. C.; Leland, P. A.; Kim, B. M.; Chang, S. I.; Raines, R. T., Ribonuclease inhibitor regulates neovascularization by human angiogenin. *Biochemistry* **2009,** 48, (18), 3804-6.

163. Mollenhauer, H. H.; Morre, D. J.; Rowe, L. D., Alteration of intracellular traffic by monensin; mechanism, specificity and relationship to toxicity. *Biochim Biophys Acta* **1990,** 1031, (2), 225-46.

164. Vicentini, A. M.; Kieffer, B.; Matthies, R.; Meyhack, B.; Hemmings, B. A.; Stone, S. R.; Hofsteenge, J., Protein chemical and kinetic characterization of recombinant porcine ribonuclease inhibitor expressed in Saccharomyces cerevisiae. *Biochemistry* **1990,** 29, (37), 8827-34.

165. Iyer, S.; Holloway, D. E.; Kumar, K.; Shapiro, R.; Acharya, K. R., Molecular recognition of human eosinophil-derived neurotoxin (RNase 2) by placental ribonuclease inhibitor. *J Mol Biol* **2005,** 347, (3), 637-55.

166. Blackburn, P.; Wilson, G.; Moore, S., Ribonuclease inhibitor from human placenta. Purification and properties. *J Biol Chem* **1977,** 252, (16), 5904-10.

167. Murthy, B. S.; De Lorenzo, C.; Piccoli, R.; D'Alessio, G.; Sirdeshmukh, R., Effects of protein RNase inhibitor and substrate on the quaternary structures of bovine seminal RNase. *Biochemistry* **1996,** 35, (13), 3880-5.

168. Rutkoski, T. J.; Raines, R. T., Evasion of ribonuclease inhibitor as a determinant of ribonuclease cytotoxicity. *Curr Pharm Biotechnol* **2008,** 9, (3), 185-9.

169. Rosenberg, H. F.; Dyer, K. D., Eosinophil cationic protein and eosinophil-derived neurotoxin. Evolution of novel function in a primate ribonuclease gene family. *J Biol Chem* **1995,** 270, (37), 21539-44.

170. Futami, J.; Maeda, T.; Kitazoe, M.; Nukui, E.; Tada, H.; Seno, M.; Kosaka, M.; Yamada, H., Preparation of potent cytotoxic ribonucleases by cationization: enhanced cellular uptake and decreased interaction with ribonuclease inhibitor by chemical modification of carboxyl groups. *Biochemistry* **2001**, 40, (25), 7518-24.

171. Lee, J. E.; Raines, R. T., Cytotoxicity of bovine seminal ribonuclease: monomer versus dimer. *Biochemistry* **2005,** 44, (48), 15760-7.

172. Lacadena, J.; Mancheno, J. M.; Martinez-Ruiz, A.; Martinez del Pozo, A.; Gasset, M.; Onaderra, M.; Gavilanes, J. G., Substitution of histidine-137 by glutamine abolishes the catalytic activity of the ribosome-inactivating protein alpha-sarcin. *Biochem J* **1995**, 309 (Pt 2), 581-6.

173. Notomista, E.; Catanzano, F.; Graziano, G.; Di Gaetano, S.; Barone, G.; Di Donato, A., Contribution of chain termini to the conformational stability and biological activity of onconase. *Biochemistry* **2001,** 40, (31), 9097-103.

174. Kim, B. M.; Kim, H.; Raines, R. T.; Lee, Y., Glycosylation of onconase increases its conformational stability and toxicity for cancer cells. *Biochem Biophys Res Commun* **2004**, 315, (4), 976-83.

175. Klink, T. A.; Woycechowsky, K. J.; Taylor, K. M.; Raines, R. T., Contribution of disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. *Eur J Biochem* **2000,** 267, (2), 566-72.

176. Ilinskaya, O. N.; Dreyer, F.; Mitkevich, V. A.; Shaw, K. L.; Pace, C. N.; Makarov, A. A., Changing the net charge from negative to positive makes ribonuclease Sa cytotoxic. *Protein Sci* **2002,** 11, (10), 2522-5.

177. Modler, A. J.; Gast, K.; Lutsch, G.; Damaschun, G., Assembly of amyloid protofibrils via critical oligomers--a novel pathway of amyloid formation. *J Mol Biol* **2003,** 325, (1), 135-48.

178. Gotte, G.; Mahmoud Helmy, A.; Ercole, C.; Spadaccini, R.; Laurents, D. V.; Donadelli, M.; Picone, D., Double domain swapping in bovine seminal RNase: formation of distinct N- and C-swapped tetramers and multimers with increasing biological activities. *PLoS One* **2012,** 7, (10), e46804.

179. Kita, Y.; Arakawa, T., Salts and glycine increase reversibility and decrease aggregation during thermal unfolding of ribonuclease-A. *Biosci Biotechnol Biochem* **2002,** 66, (4), 880-2.

180. Sanchez de Groot, N.; Pallares, I.; Aviles, F. X.; Vendrell, J.; Ventura, S., Prediction of "hot spots" of aggregation in disease-linked polypeptides. *BMC Struct Biol* **2005,** 5, 18.

181. Torrent, M.; Navarro, S.; Moussaoui, M.; Nogues, M. V.; Boix, E., Eosinophil cationic protein high-affinity binding to bacteria-wall lipopolysaccharides and peptidoglycans. *Biochemistry* **2008,** 47, (11), 3544-55.

182. Pulido, D.; Moussaoui, M.; Andreu, D.; Nogues, M. V.; Torrent, M.; Boix, E., Antimicrobial action and cell agglutination by the eosinophil cationic protein are modulated by the cell wall lipopolysaccharide structure. *Antimicrob Agents Chemother* **2012,** 56, (5), 2378-85.

183. Fan, T. C.; Fang, S. L.; Hwang, C. S.; Hsu, C. Y.; Lu, X. A.; Hung, S. C.; Lin, S. C.; Chang, M. D., Characterization of molecular interactions between eosinophil cationic protein and heparin. *J Biol Chem* **2008,** 283, (37), 25468-74.

184. Carreras, E.; Boix, E.; Navarro, S.; Rosenberg, H. F.; Cuchillo, C. M.; Nogues, M. V., Surface-exposed amino acids of eosinophil cationic protein play a critical role in the inhibition of mammalian cell proliferation. *Mol Cell Biochem* **2005,** 272, (1-2), 1-7.

185. Suzuki, H.; Parente, A.; Farina, B.; Greco, L.; La Montagna, R.; Leone, E., Complete amino-acid sequence of bovine seminal ribonuclease, a dimeric protein from seminal plasma. *Biol Chem Hoppe Seyler* **1987,** 368, (10), 1305-12.

186. Rybak, S. M.; Newton, D. L., Natural and engineered cytotoxic ribonucleases: therapeutic potential. *Exp Cell Res* **1999,** 253, (2), 325-35.

187. Soucek, J.; Marinov, I.; Benes, J.; Hilgert, I.; Matousek, J.; Raines, R. T., Immunosuppressive activity of bovine seminal ribonuclease and its mode of action. *Immunobiology* **1996,** 195, (3), 271-85.

188. Ardelt, W.; Mikulski, S. M.; Shogen, K., Amino acid sequence of an anti-tumor protein from Rana pipiens oocytes and early embryos. Homology to pancreatic ribonucleases. *J Biol Chem* **1991,** 266, (1), 245-51.

189. Leland, P. A.; Schultz, L. W.; Kim, B. M.; Raines, R. T., Ribonuclease A variants with potent cytotoxic activity. *Proc Natl Acad Sci U S A* **1998,** 95, (18), 10407-12.

190. Newton, D. L.; Walbridge, S.; Mikulski, S. M.; Ardelt, W.; Shogen, K.; Ackerman, S. J.; Rybak, S. M.; Youle, R. J., Toxicity of an antitumor ribonuclease to Purkinje neurons. *J Neurosci* **1994,** 14, (2), 538-44.

191. Saxena, S. K.; Gravell, M.; Wu, Y. N.; Mikulski, S. M.; Shogen, K.; Ardelt, W.; Youle, R. J., Inhibition of HIV-1 production and selective degradation of viral RNA by an amphibian ribonuclease. *J Biol Chem* **1996,** 271, (34), 20783-8.

192. Sakakibara, F.; Takayanagi, G.; Kawauchi, H.; Watanabe, K.; Hakomori, S., An anti-A-like lectin of Rana catesbiana eggs showing unusual reactivity. *Biochim Biophys Acta* **1976,** 444, (2), 386-95.

193. Titani, K.; Takio, K.; Kuwada, M.; Nitta, K.; Sakakibara, F.; Kawauchi, H.; Takayanagi, G.; Hakomori, S., Amino acid sequence of sialic acid binding lectin from frog (Rana catesbeiana) eggs. *Biochemistry* **1987,** 26, (8), 2189-94.

194. Nitta, K.; Ozaki, K.; Tsukamoto, Y.; Furusawa, S.; Ohkubo, Y.; Takimoto, H.; Murata, R.; Hosono, M.; Hikichi, N.; Sasaki, K.; et al., Characterization of a Rana catesbeiana lectin-resistant mutant of leukemia P388 cells. *Cancer Res* **1994,** 54, (4), 928-34.

195. Nitta, R.; Katayama, N.; Okabe, Y.; Iwama, M.; Watanabe, H.; Abe, Y.; Okazaki, T.; Ohgi, K.; Irie, M., Primary structure of a ribonuclease from bullfrog (Rana catesbeiana) liver. *J Biochem* **1989,** 106, (5), 729-35.

196. Pizzo, E.; Varcamonti, M.; Di Maro, A.; Zanfardino, A.; Giancola, C.; D'Alessio, G., Ribonucleases with angiogenic and bactericidal activities from the Atlantic salmon. *Febs J* **2008,** 275, (6), 1283-95.

197. Pizzo, E.; Merlino, A.; Turano, M.; Russo Krauss, I.; Coscia, F.; Zanfardino, A.; Varcamonti, M.; Furia, A.; Giancola, C.; Mazzarella, L.; Sica, F.; D'Alessio, G., A new RNase sheds light on the RNase/angiogenin subfamily from zebrafish. *Biochem J* **2011,** 433, (2), 345-55.

198. Mazzarella, L.; Capasso, S.; Demasi, D.; Di Lorenzo, G.; Mattia, C. A.; Zagari, A., Bovine seminal ribonuclease: structure at 1.9 A resolution. *Acta Crystallogr D Biol Crystallogr* **1993,** 49, (Pt 4), 389-402.

199. Mosimann, S. C.; Ardelt, W.; James, M. N., Refined 1.7 A X-ray crystallographic structure of P-30 protein, an amphibian ribonuclease with anti-tumor activity. *J Mol Biol* **1994,** 236, (4), 1141-53.

200. Perez-Canadillas, J. M.; Santoro, J.; Campos-Olivas, R.; Lacadena, J.; Martinez del Pozo, A.; Gavilanes, J. G.; Rico, M.; Bruix, M., The highly refined solution structure of the cytotoxic ribonuclease alpha-sarcin reveals the structural requirements for substrate recognition and ribonucleolytic activity. *J Mol Biol* **2000,** 299, (4), 1061-73.

201. Leu, Y. J.; Chern, S. S.; Wang, S. C.; Hsiao, Y. Y.; Amiraslanov, I.; Liaw, Y. C.; Liao, Y. D., Residues involved in the catalysis, base specificity, and cytotoxicity of ribonuclease from Rana catesbeiana based upon mutagenesis and X-ray crystallography. *J Biol Chem* **2003,** 278, (9), 7300-9.

202. Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M., Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res* **2001,** 29, (13), 2860-74.

203. Jones, S.; Daley, D. T.; Luscombe, N. M.; Berman, H. M.; Thornton, J. M., Protein-RNA interactions: a structural analysis. *Nucleic Acids Res* **2001,** 29, (4), 943-54.

204. Golovin, A.; Dimitropoulos, D.; Oldfield, T.; Rachedi, A.; Henrick, K., MSDsite: a database search and retrieval system for the analysis and viewing of bound ligands and active sites. *Proteins* **2005,** 58, (1), 190-9.

205. Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N., ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* **2010,** 38, (Web Server issue), W529-33.

206. Studier, F. W.; Moffatt, B. A., Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **1986,** 189, (1), 113-30.

207. Boix, E., Eosinophil cationic protein. *Methods Enzymol* **2001,** 341, 287-305.

208. Burgess, R. R., Refolding solubilized inclusion body proteins. *Methods Enzymol* **2009,** 463, 259-82.

209. Gill, S. C.; von Hippel, P. H., Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* **1989,** 182, (2), 319-26.

210. McPherson, A., Current approaches to macromolecular crystallization. *Eur J Biochem* **1990,** 189, (1), 1-23.

211. Pflugrath, J. W., Macromolecular cryocrystallography--methods for cooling and mounting protein crystals at cryogenic temperatures. *Methods* **2004,** 34, (3), 415-23.

212. Incardona, M. F.; Bourenkov, G. P.; Levik, K.; Pieritz, R. A.; Popov, A. N.; Svensson, O., EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *J Synchrotron Radiat* **2009,** 16, (Pt 6), 872-9.

213. Kabsch, W., Xds. *Acta Crystallogr D Biol Crystallogr* **2010,** 66, (Pt 2), 125-32.

214. Battye, T. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G., iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **2011,** 67, (Pt 4), 271-81.

215. Evans, P. R., An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* **2010,** 67, (Pt 4), 282-92.

216. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2010,** 67, (Pt 4), 235-42.

217. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *J Appl Crystallogr* **2007,** 40, (Pt 4), 658-674.

218. Taylor, G., The phase problem. *Acta Crystallogr D Biol Crystallogr* **2003,** 59, (Pt 11), 1881-90.

219. Evans, P.; McCoy, A., An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr* **2008,** 64, (Pt 1), 1-10.

220. Matthews, B. W., Solvent content of protein crystals. *J Mol Biol* **1968,** 33, (2), 491-7.

221. Dodson, E., The befores and afters of molecular replacement. *Acta Crystallogr D Biol Crystallogr* **2008,** 64, (Pt 1), 17-24.

222. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, (Pt 12 Pt 1), 2126-32.

223. Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A., REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **2011,** 67, (Pt 4), 355-67.

224. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010,** 66, (Pt 2), 213-21.

225. Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010,** 66, (Pt 1), 12-21.

226. Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D., Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **2012,** 68, (Pt 4), 352-67.

227. Howlin, B.; Moss, D. S.; Harris, G. W., Segmented anisotropic refinement of bovine ribonuclease A by the application of the rigid-body TLS model. *Acta Crystallogr A* **1989,** 45 (Pt 12), 851-61.

228. Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M., Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *Febs J* **2008,** 275, (1), 1-21.

229. Berisio, R.; Lamzin, V. S.; Sica, F.; Wilson, K. S.; Zagari, A.; Mazzarella, L., Protein titration in the crystal state. *J Mol Biol* **1999,** 292, (4), 845-54.

230. Afonine, P. V.; Mustyakimov, M.; Grosse-Kunstleve, R. W.; Moriarty, N. W.; Langan, P.; Adams, P. D., Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **2010,** 66, (Pt 11), 1153-63.

231. Riffel, N.; Harlos, K.; Iourin, O.; Rao, Z.; Kingsman, A.; Stuart, D.; Fry, E., Atomic resolution structure of Moloney murine leukemia virus matrix protein and its relationship to other retroviral matrix proteins. *Structure* **2002,** 10, (12), 1627-36.

232. Brunger, A. T., Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **1992,** 355, (6359), 472-5.

233. Merritt, E. A., Expanding the model: anisotropic displacement parameters in protein structure refinement. *Acta Crystallogr D Biol Crystallogr* **1999,** 55, (Pt 6), 1109-17.

234. Vaguine, A. A.; Richelle, J.; Wodak, S. J., SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr* **1999,** 55, (Pt 1), 191-205.

235. Urzhumtseva, L.; Afonine, P. V.; Adams, P. D.; Urzhumtsev, A., Crystallographic model quality at a glance. *Acta Crystallogr D Biol Crystallogr* **2009,** 65, (Pt 3), 297-300.

236. Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M., Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **1991,** 47 (Pt 2), 110-9.

237. Parca, L.; Mangone, I.; Gherardini, P. F.; Ausiello, G.; Helmer-Citterich, M., Phosfinder: a web server for the identification of phosphate-binding sites on protein structures. *Nucleic Acids Res* **2011,** 39, (Web Server issue), W278-82.

238. Lacadena, J.; Alvarez-Garcia, E.; Carreras-Sangra, N.; Herrero-Galan, E.; Alegre-Cebollada, J.; Garcia-Ortega, L.; Onaderra, M.; Gavilanes, J. G.; Martinez del Pozo, A., Fungal ribotoxins: molecular dissection of a family of natural killers. *FEMS Microbiol Rev* **2007,** 31, (2), 212-37.

239. Boix, E.; Blanco, J. A.; Nogues, M. V.; Moussaoui, M., Nucleotide binding architecture for secreted cytotoxic endoribonucleases. *Biochimie* **2013,** 95, (6), 1087-97.

240. Parca, L.; Ferre, F.; Ausiello, G.; Helmer-Citterich, M., Nucleos: a web server for the identification of nucleotide-binding sites in protein structures. *Nucleic Acids Res* **2013,** 41, (Web Server issue), W281-5.

241. Sikriwal, D.; Seth, D.; Dey, P.; Batra, J. K., Human eosinophil-derived neurotoxin: involvement of a putative non-catalytic phosphate-binding subsite in its catalysis. *Mol Cell Biochem* **2007,** 303, (1-2), 175-81.

242. Sorrentino, S., Human extracellular ribonucleases: multiplicity, molecular diversity and catalytic properties of the major RNase types. *Cell Mol Life Sci* **1998,** 54, (8), 785-94.

243. Hofsteenge, J.; Vicentini, A.; Zelenko, O., Ribonuclease 4, an evolutionarily highly conserved member of the superfamily. *Cell Mol Life Sci* **1998,** 54, (8), 804-10.

244. Zhao, W.; Kote-Jarai, Z.; van Santen, Y.; Hofsteenge, J.; Beintema, J. J., Ribonucleases from rat and bovine liver: purification, specificity and structural characterization. *Biochim Biophys Acta* **1998,** 1384, (1), 55-65.

245. Swaminathan, G. J.; Myszka, D. G.; Katsamba, P. S.; Ohnuki, L. E.; Gleich, G. J.; Acharya, K. R., Eosinophil-granule major basic protein, a C-type lectin, binds heparin. *Biochemistry* **2005,** 44, (43), 14152-8.

246. Vicentini, A. M.; Kote-Jarai, Z.; Hofsteenge, J., Structural determinants of the uridine-preferring specificity of RNase PL3. *Biochemistry* **1996,** 35, (28), 9128-32.

247. Zegers, I.; Maes, D.; Dao-Thi, M. H.; Poortmans, F.; Palmer, R.; Wyns, L., The structures of RNase A complexed with 3'-CMP and d(CpA): active site conformation and conserved water molecules. *Protein Sci* **1994,** 3, (12), 2322-39.

248. Merlino, A.; Vitagliano, L.; Sica, F.; Zagari, A.; Mazzarella, L., Population shift vs induced fit: the case of bovine seminal ribonuclease swapping dimer. *Biopolymers* **2004,** 73, (6), 689-95.

249. Lee, J. E.; Bae, E.; Bingman, C. A.; Phillips, G. N., Jr.; Raines, R. T., Structural basis for catalysis by onconase. *J Mol Biol* **2008,** 375, (1), 165-77.

250. Koepke, J.; Maslowska, M.; Heinemann, U.; Saenger, W., Three-dimensional structure of ribonuclease T1 complexed with guanylyl-2',5'-guanosine at 1.8 A resolution. *J Mol Biol* **1989,** 206, (3), 475-88.

251. Vassylyev, D. G.; Katayanagi, K.; Ishikawa, K.; Tsujimoto-Hirano, M.; Danno, M.; Pahler, A.; Matsumoto, O.; Matsushima, M.; Yoshida, H.; Morikawa, K., Crystal structures of ribonuclease F1 of Fusarium moniliforme in its free form and in complex with 2'GMP. *J Mol Biol* **1993,** 230, (3), 979-96.

252. Noguchi, S., Isomerization mechanism of aspartate to isoaspartate implied by structures of Ustilago sphaerogena ribonuclease U2 complexed with adenosine 3'-monophosphate. *Acta Crystallogr D Biol Crystallogr* **2010,** 66, (Pt 7), 843-9.

253. Sevcik, J.; Dodson, E. J.; Dodson, G. G., Determination and restrained least-squares refinement of the structures of ribonuclease Sa and its complex with 3'-guanylic acid at 1.8 A resolution. *Acta Crystallogr B* **1991,** 47 (Pt 2), 240-53.

254. Polyakov, K. M.; Lebedev, A. A.; Okorokov, A. L.; Panov, K. I.; Schulga, A. A.; Pavlovsky, A. G.; Karpeisky, M. Y.; Dodson, G. G., The structure of substrate-free microbial ribonuclease binase and of its complexes with 3'GMP and sulfate ions. *Acta Crystallogr D Biol Crystallogr* **2002,** 58, (Pt 5), 744-50.

255. Wodak, S. Y., The structure of cytidilyl(2',5')adenosine when bound to pancreatic ribonuclease S. *J Mol Biol* **1977,** 116, (4), 855-75.

256. Toiron, C.; Gonzalez, C.; Bruix, M.; Rico, M., Three-dimensional structure of the complexes of ribonuclease A with 2',5'-CpA and 3',5'-d(CpA) in aqueous solution, as obtained by NMR and restrained molecular dynamics. *Protein Sci* **1996,** 5, (8), 1633-47.

257. Tarragona-Fiol, A.; Eggelte, H. J.; Harbron, S.; Sanchez, E.; Taylorson, C. J.; Ward, J. M.; Rabin, B. R., Identification by site-directed mutagenesis of amino acids in the B2 subsite of bovine pancreatic ribonuclease A. *Protein Eng* **1993,** 6, (8), 901-6.

258. Vitagliano, L.; Merlino, A.; Zagari, A.; Mazzarella, L., Productive and nonproductive binding to ribonuclease A: X-ray structure of two complexes with uridylyl(2',5')guanosine. *Protein Sci* **2000,** 9, (6), 1217-25.

259. Listgarten, J. N.; Maes, D.; Wyns, L.; Aguilar, C. F.; Palmer, R. A., Structure of the crystalline complex of deoxycytidylyl-3',5'-guanosine (3',5'-dCpdG) cocrystallized with ribonuclease at 1.9 A resolution. *Acta Crystallogr D Biol Crystallogr* **1995,** 51, (Pt 5), 767-71.

260. Ercole, C.; Colamarino, R. A.; Pizzo, E.; Fogolari, F.; Spadaccini, R.; Picone, D., Comparison of the structural and functional properties of RNase A and BS-RNase: a stepwise mutagenesis approach. *Biopolymers* **2009,** 91, (12), 1009-17.

261. Irie, M.; Nitta, K.; Nonaka, T., Biochemistry of frog ribonucleases. *Cell Mol Life Sci* **1998,** 54, (8), 775-84.

262. Singh, U. P.; Ardelt, W.; Saxena, S. K.; Holloway, D. E.; Vidunas, E.; Lee, H. S.; Saxena, A.; Shogen, K.; Acharya, K. R., Enzymatic and structural characterisation of amphinase, a novel cytotoxic ribonuclease from Rana pipiens oocytes. *J Mol Biol* **2007,** 371, (1), 93-111.

263. Chang, C. F.; Chen, C.; Chen, Y. C.; Hom, K.; Huang, R. F.; Huang, T. H., The solution structure of a cytotoxic ribonuclease from the oocytes of Rana catesbeiana (bullfrog). *J Mol Biol* **1998,** 283, (1), 231-44.

264. Acharya, K. R.; Shapiro, R.; Allen, S. C.; Riordan, J. F.; Vallee, B. L., Crystal structure of human angiogenin reveals the structural basis for its functional divergence from ribonuclease. *Proc Natl Acad Sci U S A* **1994,** 91, (8), 2915-9.

265. Makarov, A. A.; Ilinskaya, O. N., Cytotoxic ribonucleases: molecular weapons and their targets. *FEBS Lett* **2003,** 540, (1-3), 15-20.

266. Leonidas, D. D.; Shapiro, R.; Irons, L. I.; Russo, N.; Acharya, K. R., Crystal structures of ribonuclease A complexes with 5'-diphosphoadenosine 3'-phosphate and 5'-diphosphoadenosine 2'-phosphate at 1.7 A resolution. *Biochemistry* **1997,** 36, (18), 5578-88.

267. Papageorgiou, A. C.; Shapiro, R.; Acharya, K. R., Molecular recognition of human angiogenin by placental ribonuclease inhibitor--an X-ray crystallographic study at 2.0 A resolution. *Embo J* **1997,** 16, (17), 5162-77.

268. Moussaoui, M.; Guasch, A.; Boix, E.; Cuchillo, C.; Nogues, M., The role of non-catalytic binding subsites in the endonuclease activity of bovine pancreatic ribonuclease A. *J Biol Chem* **1996,** 271, (9), 4687-92.

269. Cuchillo, C. M.; Moussaoui, M.; Barman, T.; Travers, F.; Nogues, M. V., The exo- or endonucleolytic preference of bovine pancreatic ribonuclease A depends on its subsites structure and on the substrate size. *Protein Sci* **2002,** 11, (1), 117-28.

270. Berisio, R.; Sica, F.; Lamzin, V. S.; Wilson, K. S.; Zagari, A.; Mazzarella, L., Atomic resolution structures of ribonuclease A at six pH values. *Acta Crystallogr D Biol Crystallogr* **2002,** 58, (Pt 3), 441-50.

271. Park, C.; Schultz, L. W.; Raines, R. T., Contribution of the active site histidine residues of ribonuclease A to nucleic acid binding. *Biochemistry* **2001,** 40, (16), 4949-56.

272. Fedorov, A. A.; Joseph-McCarthy, D.; Fedorov, E.; Sirakova, D.; Graf, I.; Almo, S. C., Ionic interactions in crystalline bovine pancreatic ribonuclease A. *Biochemistry* **1996,** 35, (50), 15962-79.

273. Mueller-Dieckmann, C.; Panjikar, S.; Schmidt, A.; Mueller, S.; Kuper, J.; Geerlof, A.; Wilmanns, M.; Singh, R. K.; Tucker, P. A.; Weiss, M. S., On the routine use of soft X-rays in macromolecular crystallography. Part IV. Efficient determination of anomalous substructures in biomacromolecules using longer X-ray wavelengths. *Acta Crystallogr D Biol Crystallogr* **2007,** 63, (Pt 3), 366-80.

274. Kurpiewska, K.; Font, J.; Ribo, M.; Vilanova, M.; Lewinski, K., X-ray crystallographic studies of RNase A variants engineered at the most destabilizing positions of the main hydrophobic core: further insight into protein stability. *Proteins* **2009,** 77, (3), 658-69.

275. Boerema, D. J.; Tereshko, V. A.; Kent, S. B., Total synthesis by modern chemical ligation methods and high resolution (1.1 A) X-ray structure of ribonuclease A. *Biopolymers* **2008,** 90, (3), 278-86.

276. Hooft, R. W.; Vriend, G.; Sander, C.; Abola, E. E., Errors in protein structures. *Nature* **1996,** 381, (6580), 272.

277. Dimitropoulos, D.; Ionides, J.; Henrick, K., Using MSDchem to search the PDB ligand dictionary. *Curr Protoc Bioinformatics* **2006,** Chapter 14, Unit14 3.

278. Leonidas, D. D.; Shapiro, R.; Allen, S. C.; Subbarao, G. V.; Veluraja, K.; Acharya, K. R., Refined crystal structures of native human angiogenin and two active site variants: implications for the unique functional properties of an enzyme involved in neovascularisation during tumour growth. *J Mol Biol* **1999,** 285, (3), 1209-33.

279. Fisher, B. M.; Schultz, L. W.; Raines, R. T., Coulombic effects of remote subsites on the active site of ribonuclease A. *Biochemistry* **1998,** 37, (50), 17386-401.

280. Boque, L.; Gracia Coll, M.; Vilanova, M.; Cuchillo, C. M.; Fita, I., Structure of ribonuclease A derivative II at 2.1-A resolution. *J Biol Chem* **1994,** 269, (31), 19707-12.

281. Irie, M.; Ohgi, K.; Yoshinaga, M.; Yanagida, T.; Okada, Y.; Teno, N., Roles of lysine1 and lysine7 residues of bovine pancreatic ribonuclease in the enzymatic activity. *J Biochem* **1986,** 100, (4), 1057-63.

282. Doncheva, N. T.; Klein, K.; Domingues, F. S.; Albrecht, M., Analyzing and visualizing residue networks of protein structures. *Trends Biochem Sci* **2011,** 36, (4), 179-82.

283. Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A., PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **2004,** 32, (Web Server issue), W665-7.

284. Li, H.; Robertson, A. D.; Jensen, J. H., Very fast empirical prediction and rationalization of protein pKa values. *Proteins* **2005,** 61, (4), 704-21.

285. Markley, J. L., Correlation proton magnetic resonance studies at 250 MHz of bovine pancreatic ribonuclease. I. Reinvestigation of the histidine peak assignments. *Biochemistry* **1975,** 14, (16), 3546-54.

286. Fairman, R.; Shoemaker, K. R.; York, E. J.; Stewart, J. M.; Baldwin, R. L., The Glu 2- ... Arg 10+ side-chain interaction in the C-peptide helix of ribonuclease A. *Biophys Chem* **1990,** 37, (1-3), 107-19.

287. Borkakoti, N., The active site of ribonuclease A from the crystallographic studies of ribonuclease-A-inhibitor complexes. *Eur J Biochem* **1983,** 132, (1), 89-94.

288. deMel, V. S.; Martin, P. D.; Doscher, M. S.; Edwards, B. F., Structural changes that accompany the reduced catalytic efficiency of two semisynthetic ribonuclease analogs. *J Biol Chem* **1992,** 267, (1), 247-56.

289. de Llorens, R.; Arus, C.; Pares, X.; Cuchillo, C. M., Chemical and computer graphics studies on the topography of the ribonuclease A active site cleft. A model of the enzyme-pentanucleotide substrate complex. *Protein Eng* **1989,** 2, (6), 417-29.

290. Wlodawer, A.; Svensson, L. A.; Sjolin, L.; Gilliland, G. L., Structure of phosphate-free ribonuclease A refined at 1.26 A. *Biochemistry* **1988,** 27, (8), 2705-17.

291. delCardayre, S. B.; Raines, R. T., A residue to residue hydrogen bond mediates the nucleotide specificity of ribonuclease A. *J Mol Biol* **1995,** 252, (3), 328-36.

292. Ladner, J. E.; Wladkowski, B. D.; Svensson, L. A.; Sjolin, L.; Gilliland, G. L., X-ray structure of a ribonuclease A-uridine vanadate complex at 1.3 A resolution. *Acta Crystallogr D Biol Crystallogr* **1997,** 53, (Pt 3), 290-301.

293. Thompson, J. E.; Kutateladze, T. G.; Schuster, M. C.; Venegas, F. D.; Messmore, J. M.; Raines, R. T., Limits to Catalysis by Ribonuclease A. *Bioorg Chem* **1995,** 23, (4), 471-481.

294. Veenstra, T. D.; Lee, L., NMR study of the positions of His-12 and His-119 in the ribonuclease A-uridine vanadate complex. *Biophys J* **1994,** 67, (1), 331-5.

295. Laurents, D. V.; Bruix, M.; Jimenez, M. A.; Santoro, J.; Boix, E.; Moussaoui, M.; Nogues, M. V.; Rico, M., The (1)H, (13)C, (15)N resonance assignment, solution structure, and residue level stability of eosinophil cationic protein/RNase 3 determined by NMR spectroscopy. *Biopolymers* **2009,** 91, (12), 1018-28.

296. Singh, A.; Batra, J. K., Role of unique basic residues in cytotoxic, antibacterial and antiparasitic activities of human eosinophil cationic protein. *Biol Chem* **2011,** 392, (4), 337-46.

297. Becknell, B.; Eichler, T. E.; Beceiro, S.; Li, B.; Easterling, R. S.; Carpenter, A. R.; James, C. L.; McHugh, K. M.; Hains, D. S.; Partida-Sanchez, S.; Spencer, J. D., Ribonucleases 6 and 7 have antimicrobial function in the human and murine urinary tract. *Kidney Int* **2015,** 87, (1), 151-61.

298. Rico, M.; Gallego, E.; Santoro, J.; Bermejo, F. J.; Nieto, J. L.; Herranz, J., On the fundamental role of the Glu 2- ... Arg 10+ salt bridge in the folding of isolated ribonuclease A S-peptide. *Biochem Biophys Res Commun* **1984,** 123, (2), 757-63.

299. Lovell, S. C.; Davis, I. W.; Arendall, W. B., 3rd; de Bakker, P. I.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C., Structure validation by Cα geometry: φ/ψ and Cβ deviation. *Proteins* **2003,** 50, (3), 437-50.

300. Krissinel, E.; Henrick, K., Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **2007,** 372, (3), 774-97.

301. Bell, J. A., X-ray crystal structures of a severely desiccated protein. *Protein Sci* **1999,** 8, (10), 2033-40.

302. Holloway, D. E.; Chavali, G. B.; Hares, M. C.; Subramanian, V.; Acharya, K. R., Structure of murine angiogenin: features of the substrate- and cell-binding regions and prospects for inhibitor-binding studies. *Acta Crystallogr D Biol Crystallogr* **2005,** 61, (Pt 12), 1568-78.

303. Garcia-Mayoral, M. F.; Canales, A.; Diaz, D.; Lopez-Prados, J.; Moussaoui, M.; de Paz, J. L.; Angulo, J.; Nieto, P. M.; Jimenez-Barbero, J.; Boix, E.; Bruix, M., Insights into the glycosaminoglycan-mediated cytotoxic mechanism of eosinophil cationic protein revealed by NMR. *ACS Chem Biol* **2013,** 8, (1), 144-51.

304. Faham, S.; Hileman, R. E.; Fromm, J. R.; Linhardt, R. J.; Rees, D. C., Heparin structure and interactions with basic fibroblast growth factor. *Science* **1996,** 271, (5252), 1116-20.

ANNEX

1. **Papers related to the PhD thesis**

- Nucleotide binding architecture for secreted cytotoxic endoribonucleases

Boix, E.; Blanco, J. A.; Nogués, M. V.; Moussaoui, M.

(2013) *Biochimie* 96:1087-1097

- New structure of human RNase 6 in complex with sulphate ions

Blanco, J. A.; Arranz, J.; Pulido, D.; Moussaoui, M.; Boix, E.

Data in Brief (to be submitted)

2. **Validation reports for all the crystal structures submitted to the Protein Data Bank**

2.1. Structure of the RNase A double mutant (RNase A/H7H10) in complex with 3'-CMP at 2.10 Å
2.2. Structure of RNase A at high resolution in complex with 3'-CMP at 1.16 Å (PDB ID: 4U7R)
2.3. Structure of ECP with sulphate anions at 1.50 Å (PDB ID: 4OXB)
2.4. Structure of ECP with citrate ions at 1.50 Å (PDB ID: 4OXF)
2.5. Structure of ECP/H15A mutant at 1.47 Å (PDB ID: 4OWZ)
2.6. Structure of ECP/H128N mutant with sulphate anions at 1.34 Å (PDB ID: 4X08)
2.7. Structure of human RNase 6 in complex with sulphate anions at 1.72 Å (PDB ID: 4X09)

Validation reports were elaborated by the Validation Server of the Protein Data Bank (wwwpdb-validation.wwwpbd.org/validservice.

Contents lists available at SciVerse ScienceDirect

Biochimie

Review

Nucleotide binding architecture for secreted cytotoxic endoribonucleases

Ester Boix*, Jose A. Blanco, M. Victòria Nogués, Mohammed Moussaoui

Department of Biochemistry and Molecular Biology, Biosciences Faculty, Universitat Autònoma de Barcelona, E-08193 Cerdanyola del Vallès, Spain

ABSTRACT

Vertebrate secreted RNases are small cationic protein endowed with an endoribonuclease activity that belong to the RNase A superfamily and display diverse cytotoxic activities. In an effort to unravel their mechanism of action, we have analysed their nucleotide binding recognition patterns. General shared features with other nucleotide binding proteins were deduced from overall statistics on the available structure complexes at the Protein Data Bank and compared with the particularities of selected representative endoribonuclease families. Results were compared with other endoribonuclease representative families and with the overall protein–nucleotide interaction features. Preferred amino acids and atom types involved in pair bonding interactions were identified, defining the spatial motives for phosphate, base and ribose building blocks. Together with the conserved catalytic triad at the active site, variability was observed for secondary binding subsites that may contribute to the proper substrate alignment and could explain the distinct substrate preference patterns. Highly conserved binding patterns were identified for the pyrimidine and purine subsites at the main and secondary base subsites. Particular substitution could be ascribed to specific adenine or guanine specificities. Distribution of evolutionary conserved residues were compared to search for the structure determinants that underlie their diverse catalytic efficiency and those that may account for putative physiological substrate targets or other non-catalytic biological activities that contribute to the anti-pathogen role of the RNases involved in the host defence system. A side by side comparison with another endoribonuclease superfamily of secreted cytotoxic proteins, the microbial RNases, was carried on to analyse the common features and peculiarities that rule their substrate recognition. The data provides the structural basis for the development of applied therapies targeting cellular nucleotide polymers.

© 2012 Elsevier Masson SAS. All rights reserved.

1. Introduction

The vertebrate secreted RNase superfamily comprises small secreted proteins showing very diverse catalytic and biological properties [1]. A variety of biological functions have been attributed to some family members, ranging from angiogenesis to host defence [2–6]. Mammalian homologues are grouped in eight lineages which are referred as the canonical RNases [5] (Fig. 1A). The family members were first gathered together as pancreatic type RNases, in honour to the family reference prototype, the bovine pancreatic RNase A, conforming the so-called "RNase A

superfamily". The catalytic mechanism of RNase A was already proposed in the 60s decade prior to the three-dimensional structure knowledge [7]. RNase A, is nowadays one of the best studied enzyme and represents an ideal model to understand the endoribonuclease catalytic mechanism and polymeric substrate binding mode [8–11]. The protein active site architecture reveals several phosphate binding subsites adjacent to the main catalytic site which contributes to align the RNA substrate (Fig. 1B). Together with a primary role in RNA digestion in ruminants for RNase A, a variety of non-catalytic biological properties were described for the other family members [2,12,13]. Other RNases were identified in many organs and tissues, and found not only in mammals, but in reptiles, birds, amphibians and fishes, setting the basis to define a properly vertebrate secreted RNase superfamily [5,14,15]. The antibacterial activity of distant related RNases suggested that the family evolved from an ancestral host–defence function [4]. Another suggestive hypothesis based on the angiogenic properties of low order family members, such as bird and fish RNases, ascribed

---
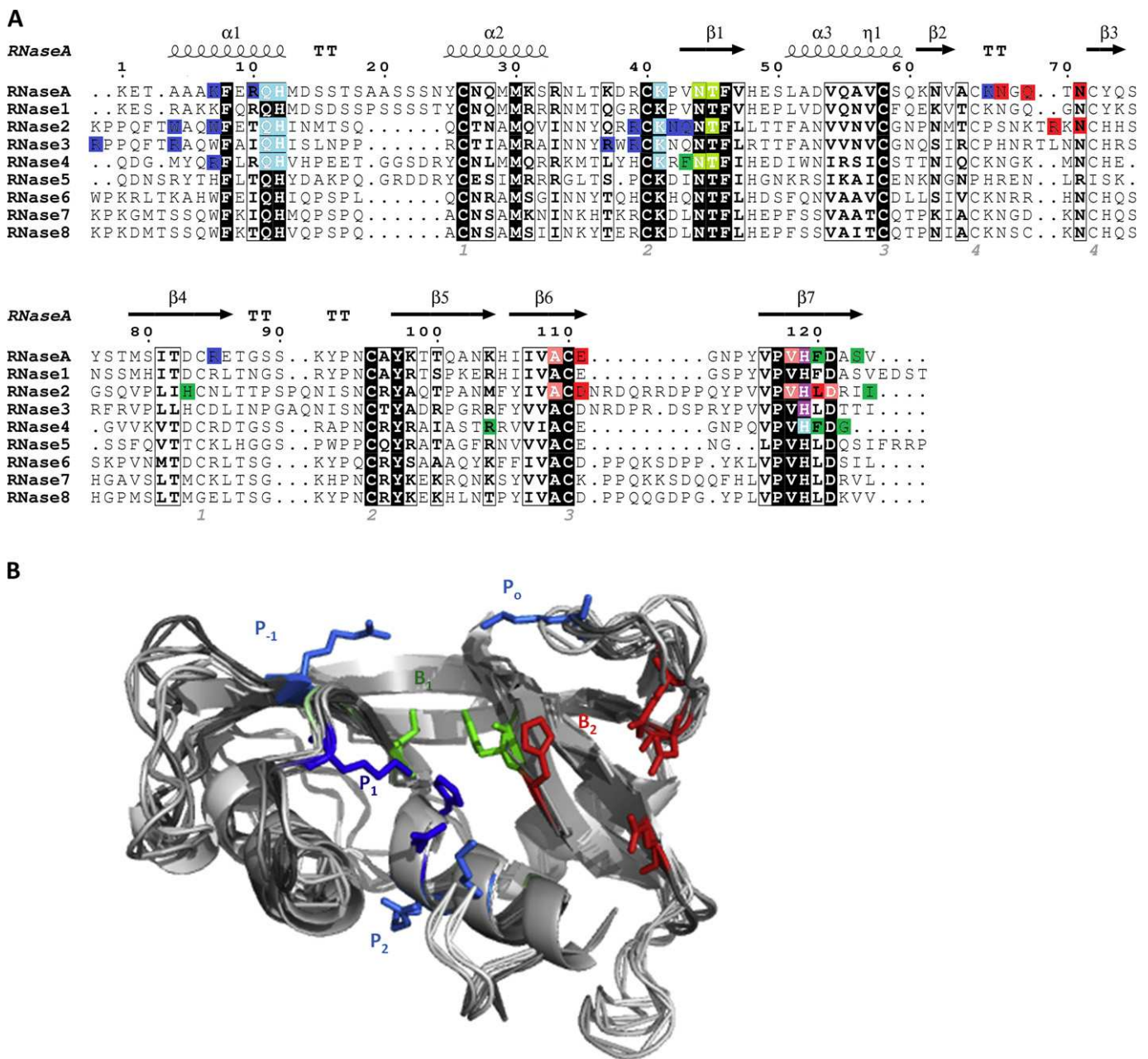
Fig. 1. (A) Sequence alignment of human RNase A superfamily members. Secondary structure elements of RNase A are depicted at the top. Strictly conserved residues are boxed in black and conserved residues, as calculated by a similarity score, are boxed in white. Coloured residues in RNase 1, RNase 2 (EDN), RNase 3 (ECP) and RNase 4 refer to those identified in protein complexes (see Table 2), and ascribed to phosphate/ribose (blue), pyrimidine (green) and purine (red) bases. Cysteine pairings for disulfide bridges are numbered below. The figure was created using the *ESPript* software [100]. (B) Representation of the superimposed three-dimensional structures of the RNases showing the subsites location and corresponding residue side chains for RNase A, coloured according to the same criteria as above.

an ancestral angiogenesis role [15,16]. All members, even with a low primary sequence identity, around 30%, share a common three-dimensional fold. All the so-called canonical RNases conserve the catalytic triad identified in the RNase A reference enzyme by His12, His119 and Lys41, where Lys is located at the family signature CKXXNTF, together with a main base subsite for pyrimidines. On the contrary, variability is found for the residues ascribed to secondary binding sites and a huge amount of work is still pending to accurately understand the enzymatic properties of most of the vertebrate secreted RNases.

The RNase A superfamily is an $\alpha + \beta$ protein type and one of the representative classified RNases superfamilies [17]. According to its

catalytic mechanism (EC 3.1.27.5), the enzyme cleaves ester bonds with an endoribonuclease activity. Endoribonuclease families embrace prokaryote and eukaryote members that can participate in cell RNA processing or are secreted proteins with a cytotoxic defence role. In particular, together with the RNase A superfamily, microbial RNases represent another reference superfamily, which includes bacterial and fungal RNases [18]. Some representative members have also been thoroughly characterized, as RNase T1, binase or RNase Sa, and provide a wealth of information on the nucleotide binding mode patterns [19].

In a wider context, nucleotide binding motives have been analysed by the statistical analysis of databases of nucleic acid–protein

V

complexes [20,21]. Nucleotides are recognized by a great variety of proteins: together with RNases we find proteins involved in cell replication and expression machinery, such as RNA polymerases or transcription factors and enzymes dependant on nucleotide-based cofactors. A modular architecture can be defined for bases, ribose and phosphate elements [22]. Small spatial three-dimensional motives were drawn, outlining pockets for each individual module, which could arise from both divergent and convergent paths [22]. In particular, structural motives for phosphate binding were previously reported [23] and searching tools were developed as the *Phosfinder* server [24]. Likewise, nucleobase recognition patterns have been defined by screening DNA/RNA–protein complexes [25–28] or adenine-based ligand protein complexes [29] and prediction softwares were proposed [30,31]. General rules for nucleic acid recognition have been inferred and RNA and DNA binding proteins are found to apply equivalent strategies, suggested to share a common evolution [32].

The exponential growth of the total number of structure complexes submitted to the protein data bank, together with the availability of statistical and analysis tools to closely inspect the ligand environments, such as the *PDBe motif* provided utilities [33,34], offer now the opportunity to review and rediscover the structural determinants for RNases substrate specificity. Main residues participating in nucleotide pairing and their interaction geometries have been screened and classified according to the overall defined patterns [22,28]. In this context, we have compared the overall patterns for protein–nucleotide binding with the specific requirements of the RNase A superfamily, as an endoribonuclease model for substrate recognition. Many secreted endonucleases have cytotoxic antitumoral properties [35–37]. Mastering the rules that govern the cellular RNA recognition and cleavage could set the basis for the development of targeted therapeutic agents [38–40].

2. Overall recognition patterns for nucleotide binding proteins

General patterns for protein–nucleotide interaction were inferred by applying the analysis tools supplied by the online server "PDBe Motif" at the PDBe Data Base (http://www.ebi.ac.uk/pdbe-site/pdbemotif/) [41]. The server provides statistical analysis of protein ligand interactions [34,42]. Protein complexes with nucleotide-type ligands were selected to define the interacting amino acids and the molecule pair bonding patterns. For each protein–ligand complex the total of hydrogen bonds and van der Waals interactions were counted and the main representative residues were listed. The percentage value of each type of interaction was then taken into account and selection criteria were then applied based on statistical interaction significance. For each residue the significant pair bonding was defined at atomic level.

CONSURF server was applied to identify the evolutionary conserved residues and correlate the data to the key ligand interacting residues [43]. For each RNase A member a multiple sequence alignment was carried on selecting the first 50 sequences from the Swiss Protein database to compute their evolutionary conservation score. Results were always considered with caution bearing in mind the potential output bias derived from the variable availability of sequences at the database for each RNase type. Analysis of protein complexes was further performed manually using *COOT* software [44] to check the relevant interaction distances within the ligand environment. A three-dimensional model was built for RNase A family members when no reference structure was available using the Automated Modelling Server *Swiss Model* [45].

The overall binding preference for nucleotide binding proteins has extensively been analysed in the literature. However, as the number of deposited structures at the protein data bank is exponentially growing, we have recalculated the statistical distribution from a considerably larger database to take it as the best reliable reference for further comparison studies. For our purpose, to obtain an adequate reference to analyse the RNase binding mode, we have selected for a further detailed analysis only the nucleotide-type ligands that could better mimic a potential RNA substrate. Small nucleotide ligands were proven to be a good model to elucidate the key interactions for polymeric substrate, as long and flexible single stranded RNA would be recognized by the protein modular motives [24].

From all the available protein–nucleotide complexes at the PDB, only the nucleotide-type ligands showing statistical significance were selected. When evaluating the residues within hydrogen bond and van der Waals distances, a generous cut-off distance was chosen, to account for the intrinsic marginal error within solved structure complexes, and inherent side chain mobility. We must also bear in mind that many complex structures were solved in non-physiological conditions. Table S1 lists for each studied ligand the most frequent hydrogen bond interacting residues. Atomic pair bonding details are subdivided according to the nucleotide elements to highlight the protein atom pairing predominance for base, ribose and phosphate recognition. To better interpret the statistical significance, the number of studied protein structures and complexes are indicated below each amino acid. For each of the best binding amino acid, the ligand interacting atom was plotted against its frequency (Figure S1).

To analyse the tendency of each amino acid to interact with every part of the nucleotide structures, we have also searched for the significant pair bonding for the respective individual building block (Figure S2). Only the data retrieved from complexes represented in a significant number in the database are listed. We are also aware that overrepresentation of some physiological relevant ligands will also bias the final conclusions, as adenine nucleotide predominance in the database, due to their important role in metabolic paths. Comparatively, fewer complexes for guanine derivates are reported, and scarce data is found for pyrimidine bases, so output results should always be considered with caution. Figure S3 shows the percentage of total, hydrogen bond and van der Waals interactions for the selected most significant residues for bases, ribose and phosphate groups. The results show that Arg, Lys and His are the favoured residues for phosphate recognition. In a second group we find polar residues, as Thr and Ser. Likewise, anionic residues, as Glu and Asp show a preferential interaction with the ribose ring. On its side, adenine recognition is driven by both hydrogen bond and van der Waals forces. The importance of van der Waals contacts was highlighted by Luscombe et al. [20], who reported that these type of interactions constituted two-third of the DNA–protein total binding. Residues Val and Ile are found to conform hydrophobic pockets for adenine and Phe for uracil base. Selected basic and acid residues contributing to specific interactions with potential donors or acceptors base groups at a physiological pH were analysed. Amino acid–base pairing were considered according to the number of bonds, as single, bidentate or multiple interactions as previously classified [20]. Particular attention was paid to discriminating groups between adenine and guanine bases, as N1 or N6/O6 atoms. As an example, Arg is one of the favoured residues for both purine bases, as NH1 and NH2 atoms bind to N6, N7 and N1 of adenine, and can also contribute by a bidentate interaction, to anchor both N7 and O6 guanine atoms [20]. Combination of Thr side chain with main chain interactions is also a characteristic of a planar binding to adenine base, as described by Kondo and Westhof [28]. Contribution of acidic residues is also favoured for base binding, where Asp OD1 or OD2 hydrogen bonding to N6 is the most observed pairing for adenine,

and Asp and Glu acceptor oxygen atoms to guanine N1 and N2 groups. Ser OG group can also interact with the guanine O6 atom. As for uracil binding pattern, N3, O2 and O4 atoms are the potential candidates for hydrogen bond interactions. Following, Asn and Gln as polar uncharged residues can also provide hydrogen bonding to uracil atoms. Indeed, Kondo and Westhof classification of nucleotide base recognition mode [28] finds Asn and Gln as the main residues for specific uracil binding by simultaneous anchoring of N3 and either O2 or O4 atoms. Similar binding mode is reported for thymine. Frequently, the peptide bond CO and NH groups are also contributing in a similar fashion, being even one of the most frequent base pairing for adenine binding proteins [28]. As a general rule, all the residues that can provide simultaneously a pair of hydrogen bonds are favoured for base recognition: Asn, Gln, Asp, Glu and Arg, fixing the base in a coplanar orientation and providing more specificity to the site pocket.

On its turn, phosphate groups are mainly recognized by spatial structural motives composed by 3 residues [24]. Statistical analysis for phosphate groups reveals a clear predominance of Arg binding, followed by Lys, fixing the phosphate groups by electrostatic interactions (Figure S2). When analysing the phosphate binding mode we would also need to consider the protein role. Location of a phosphate at the enzyme active site, where the phosphodiester bond is cleaved would be ruled by the catalytic requirements. That would justify the overrepresentation of residues as His, Asp and Glu, that can act as base catalysts.

As for ribose, the binding statistics is both biased by steric hindrance and by its proximity to the phosphate group. Among all ribose-containing ligand structures studied, the O2′ and O3′ atoms are the ones interacting the most and Asp and Glu are often observed at hydrogen bond distance. Notwithstanding, difference when comparing isolated elements will always be dependent on the particular location of each group inside the ligand that will determine its solvent exposure.

3. Binding patterns for the RNase a superfamily

The statistical analysis of nucleotide binding protein complexes was then applied to characterize the RNase A superfamily, as a representative endoribonuclease family group. The results were analysed based on the characterized catalytic mechanism of action of the family prototype RNase A. Substrate subsites for nitrogenous bases, ribose and phosphate groups were defined as Bn, Rn and Pn respectively, p1 corresponding to the main phosphate binding site where the 3′-5′ phosphodiester bond is cleaved [8—10].

A list of the RNase complexes submitted at the PDB has been taken to define the binding patterns for each of the nucleotide building blocks. Table 1 lists all the residues interacting for all nucleotide-type ligand complexes for canonical RNases and Table S2 shows the pair bonding at atomic level for the most representative complexes with substrate analogues. Additional information on other family related members is annexed in Table S3. Main interacting residues for base, ribose and phosphate are summarized in Table 2. The subsite binding architecture is well defined for RNase A (Figs. 2 and 3), where the main base site (B1) is specific for pyrimidines and the secondary site (B2) favours purine binding [8]. Together with the main phosphate binding site (p1), other secondary subsites facilitate the RNA binding and would contribute to the enzyme endonuclease cleavage pattern [10,46,47]. Residues ascribed to each defined site are labelled in the sequence alignment, where conservation of the main binding sites is visualized (Fig. 1). Variability at the secondary binding sites could explain the RNase homologues distinct catalytic efficiencies and substrate specificities [5,12,48]. Table 3 summarizes the available experimental data to

Table 1
List of protein complexes with nucleotide-type ligands from the RNase A superfamily. Residues involved in potential hydrogen bonds and van der Waals interactions are included, as calculated by the *PDBe motif* server. Peptide backbone interactions are indicated in brackets as PB. Atom detail involved in molecule pair bonding from selected nucleotide substrate analogues is included in Table S2.

Protein name	Ligand name	Ligand code	PDB code	Interacting residue
RNase A	d(CpA)	CPA	1rpg	Q11, H12, K41, V43, N44, T45, C65, Q69, N71, A109, V118, H119, F120
	5′ADP	ADP	1o0h	K7, Q11, H12, K41, C65, N67, Q69, N71, A109, E111, V118, H119, F120
	3′UMP	U3P	1o0n	Q11, H12, K41, V43 (PB), N44 (PB), T45 (PB), H119, F120 (PB)
	3′,5′ADP	A3P	1o0f	K7, Q11, H12, K41, C65, N67, Q69, N71. A109, E111, V118 (PB), H119, F120 (PB)
	2′,5′ADP	A2P	1o0o	Q11, H12, K41, C65, N67, Q69, N71, A109, E111, V118, H119, F120 (PB)
	2′UMP	U2P	1o0m	Q11, H12, K41, V43, N44 (PB), T45 (PB), H119, F120 (PB), D121 (PB), A122
	dUp	UM3	1w4p	Q11, H12, K41, V43 (PB), N44, T45 (PB), H119, F120 (PB)
	5′IMP	IMP	1z6d	A4, K7, Q11, H12, K41, V43, N44 (PB), T45, K66, R85, H119, F120 (PB), A122, S123, V124
	5′CMP	C5P	1rnn	H12, K41, V43 (PB), N44 (PB), T45, H119, F120 (PB)
	2′-5′UpG	U2G	1eos	Q11, H12, K41, N44, T45 (PB), H119, F120 (PB)
	ppAp	PAP	1afk	A4, K7, Q11, H12, N67, Q69, N71, A109, V118 (PB), H119, F120 (PB)
	2′CMP	C2P	1jvu	Q11, H12, K41, V43 (PB), N44 (PB), T45, H119, F120 (PB)
	dUpAp	PUA	1qhc	A4, K7, Q11, H12, N44 (PB), T45, N67, Q69, N71, A109, V118, H119, D121
	Purine-5′P	P5P	1rbn	K1, E2 (PB), T3, A4
	3′,5′d(CpG)	CGP	1rca	H12, K41, V43, N44 (PB), T45, R85, F120, D121 (PB), A122 (PB)
	3′CMP	C3P	1rpf	Q11, H12, K41, V43 (PB), N44 (PB), T45, H119, F120
	2′,3′UV	UVC	1ruv	Q11, H12, K41, V43, N44 (PB), T45, H119, F120 (PB)
	Uarabinose 3′P	UA3	1w4o	Q11, H12, K41, V43, N44 (PB), T45, H119, F120
	5′AMP	AMP	1z6s	Q11, H12, K41, C65, N67, Q69, N71, A109, E111, H119, F120 (PB)
	5′UMP	U5P	3jw1	H12, K41, V43 (PB), N44, T45, K66
	3′TMP	T3P	3lxo	Q11, H12, N44 (PB), T45, K66, H119
	5′ATP	ATP	2w5g	Q11, H12, K41, C65, N67, Q69, N71, A109, E111, V118, H119, F120
	5′UDP	UDP	3dxh	E2, R10, H12, N34, K41, V43 (PB), N44 (PB), T45, N67, Q69, A109, E111, V118, H119, F120
	dGp	DGP	2qca	K7, Q11, H12, K41, V43, N44 (PB), T45, K66, Q69, R85, E111, V118, H119, F120 (PB)
RNase 2 (EDN)	3′,5′ADP	A3P	1hi4	W7, Q14, H15, C62, R68, N70, A110, H129, L130
	ApppppA	AP5	2bzz	W7, Q14, H15, R36, K38, N39, Q40, C62, N70, H82, A110, D112, V128, H129, L130 (PB)
	ApppA	B4P	2c05	W7, Q14, H15, K38, R68, N70, A110, D112, V128, H129, L130
	2′,5′ADP	A2P	1hi3	Q14, H15, H129, L130 (PB), D131
	5′ADP	ADP	1hi5	Q14, H15, K38, Q40, H82, Q40, H82, H129, L130 (PB)
RNase 3 (ECP)	2′,5′ADP	A2P	1h1h	Q14, H15, R34, K38, H128
RNase 4	dUp	UM3	2rnf	R7, Q11, H12, F42, N43 (PB), T44, R101, H116, F117, G119

Table 2
Summary of the identified interacting residues in structure complexes from RNase A superfamily members with nucleotide-type ligands, as detailed in Tables 1 and S2. Residues were classified according to the interacting nucleotide bases, ribose and phosphate groups.

Protein	Interaction block				
	Base		Ribose	Phosphate	
	Pyr	Pur		Main	Others
RNase A	N44	N67	H12	Q11	K7
	T45	Q69	H119	H12	R10
	F120	N71	K41	K41	K66
	S123	A109		H119	R85
		E111			
		V118			
		H119			
RNase 2 (EDN)		R68	W7	Q14	W7
		N70	H129	H15	W10
		A110		K38	R36
		D112		H129	N39
		V128			Q40
		H129			
		L130			
		D131			

assess the base preferences for the RNase homologues. Non-conserved substitutions at some secondary sites might explain the lower catalytic activity of some family members, as RNase 3 or RNase 5, which have acquired other non-catalytic biological properties. Previous docking analysis using di- and tetranucleotides [12] confirmed that main B1 and p1 sites are conserved in all members. Some variability at B2 and P2 sites was already observed by
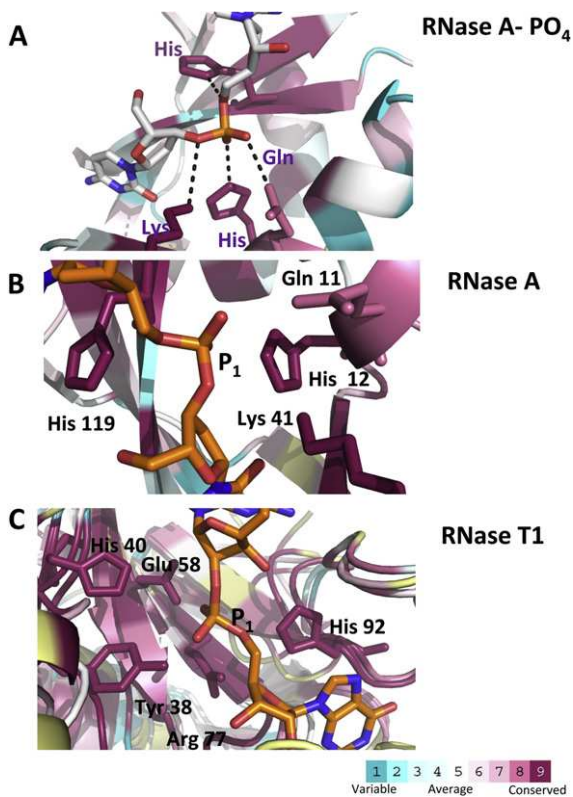
Fig. 2. Cartoon representation showing residues involved in the catalytic mechanism and the main phosphate recognition at P1 site for RNase A (A and B) and RNase T1 (C). Conserved residues are labelled. Pictures were drawn with *PYMOL* (DeLano scientific) and coloured using the *CONSURF* web server (http://consurf.tau.ac.il/) according to their evolutionary conservation score using the colour-coding bar at the bottom image. When not enough information was available, residues were coloured in yellow.
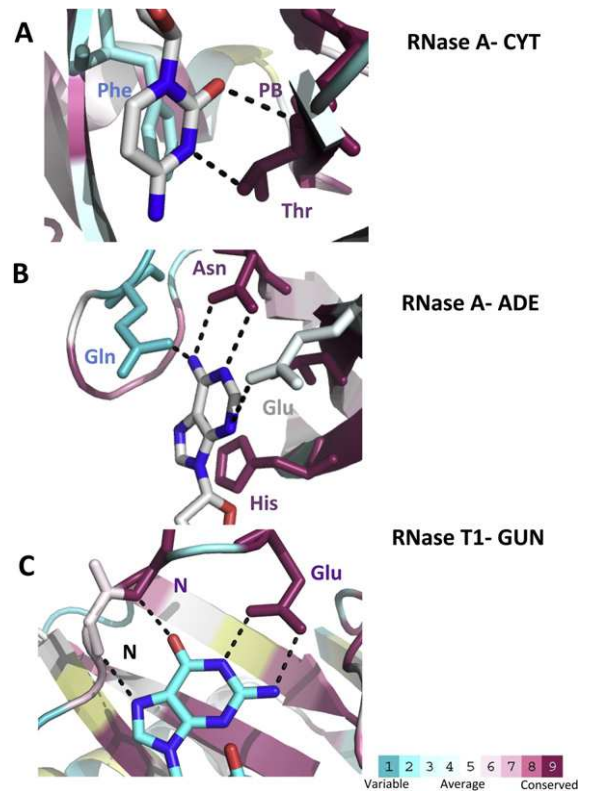
Fig. 3. Representation model for pair bonding interactions for (A) RNase A–Cytosine, (B) RNase A–Adenine and (C) RNase T1–Guanine. Structural data was taken from the RNase A–d(CpA) and the RNase T1–3GP complexes. Common hydrogen bonds shared with the other superfamily members are indicated with dashed lines. Pictures were drawn with *PYMOL* (DeLano scientific) and coloured using the *CONSURF* web server (http://consurf.tau.ac.il/) according to their evolutionary conservation score using the colour-coding bar at the bottom image. When not enough information was available, residues were coloured in yellow.

inspection of RNase 2 and RNase 3 nucleotide complexes [49–51]. Reduced interaction at the phosphate secondary site p2 in the eosinophil RNase 3 was attributed to a particular substrate preference and cleavage pattern [52]. Mutagenesis studies on the other eosinophil RNase homologue, RNase 2 [53,54], identified the contribution of secondary phosphate sites on the cleavage of long polymeric substrates. In particular both eosinophil RNases expose at the 5′ side a cationic site, named p1, which has first been identified by protein–sulphate crystal complexes [55–57]. Residues Arg36, Asn39 and Gln40 in RNase 2 were ascribed to p1 site by kinetic

Table 3
Base preference for studied RNase A superfamily members at B1 and B2 subsites, as deduced from reported kinetic results. Estimated ratio for base preference is indicated in parenthesis.

Protein	Substrate	B1 site	B2 site	Reference
RNase A	CpA, UpA, CpG, UpG Poly(C), Poly(U)	C > U (2:1)	A >>> G (10:1)	[71,101–104]
RNase 1	Poly(C), Poly(U)	C > U (2:1)	A	[104]
RNase 2 (EDN)	UpA, UpG Poly(C), Poly(U)	U > C (2:1)	A > G (4:1)	[53,104]
RNase 3 (ECP)	CpA, UpA, UpG Poly(C), Poly(U)	U > C (2:1)	A	[52,104]
RNase 4	UpA, CpA Poly(C), Poly(U)	U >>> C (10:1)	A	[62,104–106]
Bullfrog RNase	UpX, CpX CpG, UpG	U > C (2:1)	G	[81,106]
Onconase	UpG, CpG	U > C	G	[107]

comparison of single mutants [53]. A later study identified the residues contributing to B1 (Thr42, Leu130 and Ile133) and B2 (Asn70 and Asp112) [54]. Interestingly, the authors observed that removal of p0 and p2 putative sites was not only altering the enzyme substrate affinity but was also reducing the catalytic constant parameter, implying that these secondary sites are also contributing to define the main enzyme catalytic centre. This scenario was previously described for RNase A, where removal of the positive charges at p2 residues (Lys7 and Arg10) was reducing the enzyme catalytic constant even for dinucleotides [58]. The results were interpreted as a long range influence through electrostatic interactions on the pKa values of the enzyme main p1 active site residues [59].

Differences at the distribution of cationic clusters on eosinophil RNases (RNases 2 and 3) may not only reflect distinct physiological nucleic acid substrates but also affinity to other ligand partners that may modulate their biological properties. The protein exposed cationic clusters may correspond to sulphate rather than phosphate binding sites, and may facilitate the recognition of other anionic polymers as heterosaccharides [12,57]. Indeed, sulphate anions are bound at phosphate binding sites in several RNase crystal complexes [56,57] and overlap with the corresponding anion positions of either saccharide or nucleotide probes [12,60,61].

Following, we have compared the conserved interacting residues found in RNase A members complexes with the most common nucleotide binding patterns defined for all the database proteins. The available protein complexes with representative substrate analogues were overlapped and manually inspected and residues involved in each defined site were coloured according to their evolutionary conserved score, as implemented in the *CONSURF* server. Shared common traits were then depicted for recognition of each building block taking RNase A as a reference (Figs. 2 and 3). We observe that the RNase A family applies one of the most frequent strategy for base binding [28]. Fig. 3 illustrates the Thr main chain peptide backbone pairing at the pyrimidine specific B1 site. The Thr residue (Thr45) is conserved among the protein homologues (Fig. 4) and Thr45 pair bonding is observed in both cytosine and uracil nucleotide complexes (Table S2). On the other hand, the Phe contribution through stacking interactions to the pyrimidine ring is not conserved in many family homologues (Fig. 3) and may account for the higher catalytic efficiency of RNase A and RNase 4. In any case, a conservative substitution Phe to Leu for the other members can also contribute to the B1 hydrophobic pocket.

We have also considered the structural determinants for uracil/cytosine preference. Although the B1 pocket can accommodate both cytosine and uracil bases, significant differences are found
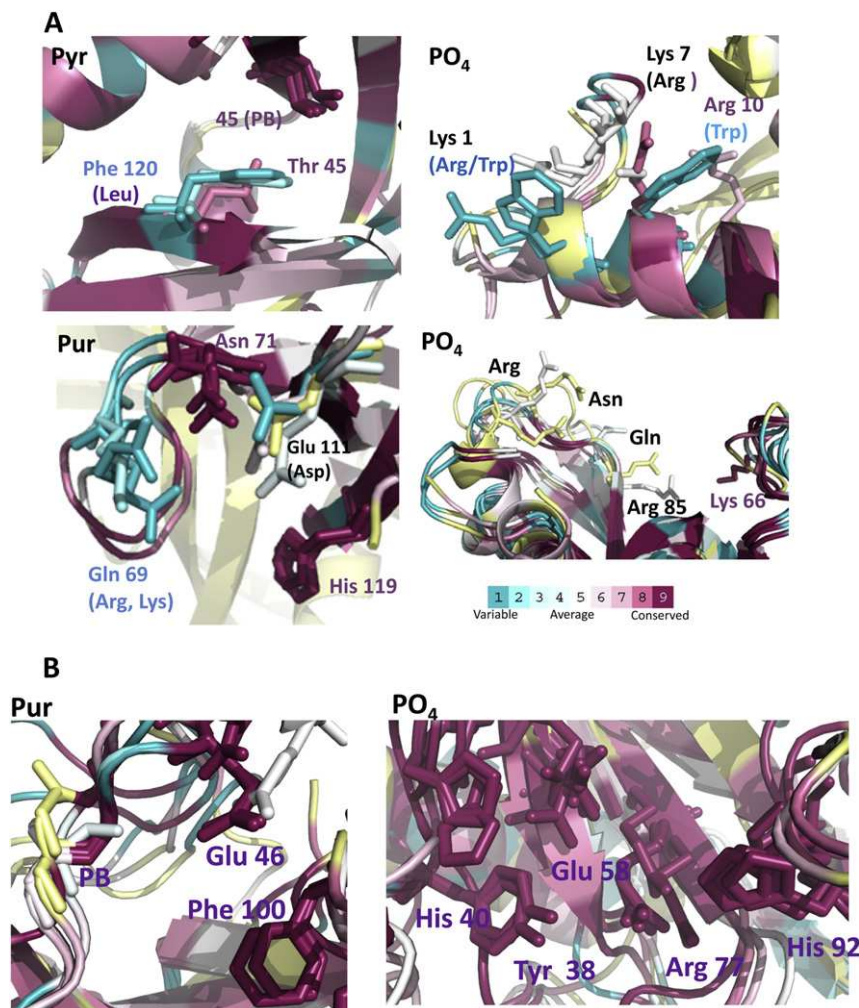
Fig. 4. Cartoon representation showing residues involved in nucleotide recognition in the RNase A superfamily (A) and the microbial RNase superfamily (B) for pyrimidine (Pyr), purine (Pur), and phosphate (PO₄) elements. Residue numbers belonging to RNase A and RNase T1 are labelled, and equivalent counterparts in other family members are indicated. Tables S2–S4 are taken as a reference. Pictures were drawn with *PYMOL* (DeLano scientific) and coloured according to their evolutionary conservation score using the *CONSURF* web server (http://consurf.tau.ac.il/). When not enough information was available, residues were coloured in yellow.

IX

between some studied members (Figs. 4 and 5A). While the pancreatic RNase A has no specificity, RNase 4 strongly prefers uracil [62] (Table 3). The RNase A versatility is attributed to Thr45, which can act as a donor or acceptor for base bonding [8,63,64]. The close Asp83 side chain can adopt two rotamer conformations and thereby switch the role of the Thr hydroxyl group. In contrast, in RNase 4, as visualized by NMR complexes [65], the Asp residue counterpart (Asp80) cannot shift freely due to close contacts with Phe42, and the Thr is then fixed, working as an acceptor group interacting with the uracil N3H group. Mutagenesis studies on RNase 4 [66,67] also highlighted the key residues for uracil specificity. Structural analysis also indicates that Arg101 in RNase 4 (substituting Lys104 in RNase A) would also modify the B1 preference, interfering with the cytosine base. Lys to Arg substitution at that position is also observed for RNase 3, while RNases 6 and 7 retain the Lys residue. Unfortunately no detailed kinetic data is available for some of these RNases.

A lower conservative pattern inside the family is observed at the B2 site (Fig. 4), although the main purine recognition trait is shared by all members. Early structural studies by X-ray crystallography reported the contribution of residues Gln69, Asn71 and Glu111 at adenine binding [68]. However, later data by X-ray and NMR only observed the contribution of Asn71 [69,70]. Mutagenesis studies

confirmed that only substitutions at Asn71 reduced the catalytic rates [71]. A double pairing in a coplanar orientation is provided by Asn71 (Fig. 3) binding to the N1 and N6 adenine atoms (Table S2) that can contribute as acceptor and donor groups. The strategy matches again the overall preferred binding mode for adenine [28]. Asn71 is conserved by all human RNase homologues except for RNase 5 (Fig. 1). On the other hand, we observe more variability at the other two purine binding residues (Gln69 and Glu111) (Figs. 1 and 4). Gln69 is only found in RNase A and human RNase 1 (Fig. 5A). Cationic residues at equivalent positions in other homologues may offer a distinct paring. While Gln OE1 atom can fix the N6 donor group in adenine base, a cationic Lys or Arg side chain may have a preference to anchor guanine bases at B2 site. This might be the case of the studied fish family members, as zebrafish RNase, where a Lys residue is oriented towards the B2 pocket. A more conserved substitution is found at Glu111 counterparts (Fig. 4), although Asp substitution was attributed to a lower binding efficiency at B2 for eosinophil RNases 2 and 3. The only exception at position 111 for human family members is observed for RNase 7 (Fig. 1), where a Lys residue close to the N3 acceptor atom would facilitate the purine binding. Unfortunately, no kinetic studies are available to check the substrate base specificity for some of the recent new family members.
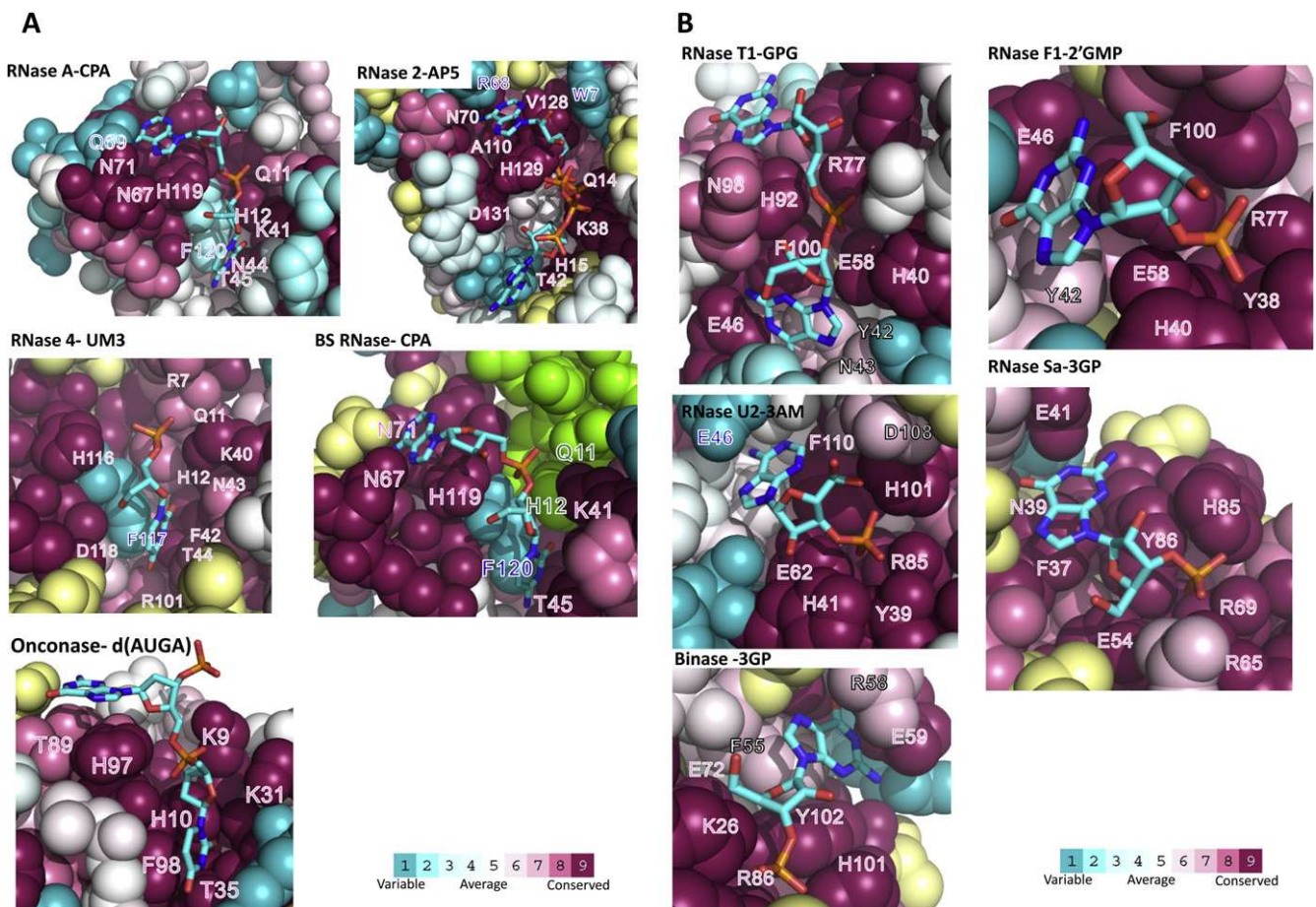
Fig. 5. Three-dimensional structure surface representation of RNase representative members of (A) RNase A superfamily and (B) microbial RNases. Pictures were drawn with *PYMOL* (DeLano scientific) and coloured using the *CONSURF* web server (http://consurf.tau.ac.il/) featuring the relationships among the evolutionary conservation of amino acid positions. Residues were coloured by their conservation score using the colour-coding bar at the bottom image. When not enough information was available, residues were coloured in yellow. Residues for bovine seminal RNase belonging to the swapped molecule are coloured in lemon green. Nucleotide ligands are drawn in sticks and coloured according to the atom type. Residues interacting with nucleotide ligands are labelled. Selected PDB complexes are: RNase A–CPA (1RPG); RNase 2–AP5 (2BZZ); RNase 4–UM3 (2RNF); BS–RNase (1R5C); Onconase–d(AUGA) (2I5S); RNase T1–GPG (2RNT); RNase F1–2′GMP (1FUT); RNase U2–3AM (3AGN); RNase Sa–3GP (2SAR); Binase–3GP (1GOY). RMS deviation for all overlapped structures is below 3 Å.

X

Kinetic data highlight that canonical RNases have a clear preference for adenine for the secondary base site (Table 3). Indeed, structural data is also supporting this substrate preference, as all nucleotide complexes locate an adenine at B2 site (Tables S2 and S3). Moreover, guanine-containing dinucleotides bind in a non-productive mode or very weakly to RNase A [72]. As an example, the RNase A–d(CpG) complex [73] shows an anomalous nucleotide orientation, where the guanine is positioned at B1. We can predict that Asn71 would interfere with the guanine O6 group. Glu111 on its side, while favouring adenine binding, may also be able to accommodate guanine. Indeed Glu and Asn interaction to N6 are one of the selected for adenine binding (Figure S2).

Finally, to analyse the particularities from the RNase members where no structural information is currently available we have compared modelled complexes using a di- and tetranucleotide docked ligand [12]. Overlapped structures onto the RNase A–d(CpA) [69] and RNase A–d(ATAA)p [74] highlight the non-conserved substitutions. Equivalent interactions were predicted at p1 and B1 sites for the 8 human RNase members. Differences are mostly observed at B2 site, where the introduction of a negative charge at the 64–71 loop in RNase 7 may explain the absence of local contacts at the purine base for the docked ligand [12]. On the contrary an increase in the number of contacts is patent for the ligand docked to the RNase 6 homologue, which can be attributed to the insertion of an additional Arg at the corresponding loop. Indeed, the counterpart Arg68 in RNase 2 is contributing to the base binding in the crystal complex (Table S2). A drastic structural modification is expected for RNase 8, an RNase uniquely expressed in placenta, where a new Cys determines an alternate disulfide bonding at this loop [75]. Preliminary kinetic assays using tRNA indicated a very reduced catalytic activity in comparison to most of the other members [75]. Recent discovery of RNase 8 polymorphisms even suggest that the RNase activity is not involved in the real protein physiological role [76]. However, lack of a systematic comparative kinetic data precludes a thorough analysis of these RNase A homologues.

Together with the 8 human canonical RNases we have analysed other representative family members (Table S3). The bovine seminal RNase, is a peculiar member with cytotoxic and antitumoral properties [77] that adopts a natural dimeric structure by domain swapping [78], where each monomer provides one of the catalytic residue to a hybrid active site and the nucleotide gets fixed by the two swapped domains [79] (Fig. 5A). The quaternary structure retains the catalytic properties of the parental RNase A [80] and all the critical nucleotide binding residues are conserved, with the only exception of the anionic residue at 111 (Table S3).

A different scenario is presented by frog RNases, which have been extensively studied because of their antitumoral properties [36,81]. In particular RNase purified from *Rana pipiens* (onconase) is currently applied in anticancer therapy. Structural studies are still ongoing to understand its poor kinetic efficiency and cytotoxicity mechanism [82]. Characterized frog RNases, as onconase or bullfrog RNase 4, share non-conserved substitutions and a deletion at the B2 loop (residues 64–71). The absence of the Asn71 loop favours the guanine binding, as confirmed by kinetic and structural studies [83]. Indeed, the onconase–tetranucleotide complex confirms the contribution of a Glu residue (corresponding to RNase A Glu111) to fix the guanine at N1 and N2 (Fig. 5A; Table S3). B1 on its side is conserved with equivalent interactions at main chain and Thr residue. Kinetic results using dinucleotides for bullfrog RNase 4 confirm the guanine preference at B2 [84] (Table 3). Onconase and bullfrog RNase also share the peculiarity of having a Pyr1 residue at the N-terminus which is oriented towards the active site and is essential for the enzyme activity [83,85,86].

The later discovery of fish homologous RNases, isolated from the zebrafish (*Danio rerio*) and the Atlantic salmon RNase [87] provided new insight into the family origin. Inspection of the structural features of zebrafish RNase 1 also suggest a guanine preference at B2, although no nucleotide complex or kinetic results can corroborate it [15]. Phylogenetic analysis associate fish RNases with RNase 5 (angiogenin) [14,87], considered the ancestral canonical member. Their structural similarity is also supported by related biological properties, as antimicrobial and angiogenic activities [15]. Interestingly RNase 5 (angiogenin) has a low catalytic activity and presents a deletion at the B2 loop. Moreover, insertion of the corresponding RNase A loop increased by about 500× fold the protein catalytic activity, while abolishing the corresponding angiogenic activity [88]. Another striking similarity between zebrafish RNase and angiogenin is the blocking of the B1 base by the protein C-terminal. In particular, Gln117 at RNase 5 C-terminus is oriented towards the active site in the solved crystal structure and a conformational rearrangement would be required to provide and alternate positioning to allow the main base binding [89].

We can conclude that all the canonical RNase A family members provide a common double bonding base pairing for both main and secondary bases which is classified as one of the main protein nucleotide recognition patterns [20,28]. While residues contributing to B1 and p1 are mostly conserved, non-conservative substitutions are found for B2 and secondary phosphate binding sites. In particular, low order RNases, as frog and fish RNases, present significant differences at the secondary base site, shifting the base specificity from adenine to guanine. Unfortunately, only scarce kinetic data is available to better interpret the structural particularities of each RNase. In any case, the physiological role of their substrate specificity remains unknown.

Notwithstanding, recent finding also suggest that surface cationic residues may also be involved in the protein binding to other polyanionic biomolecules, explaining other non-catalytic properties. As an example, cationic clusters in RNase 3 were identified to contribute to heparin binding [90–93] and to lipopolysaccharides at the Gram-negative outer membrane [94]. The protein high affinity to the exposed heterosaccharides would mediate its further toxic effect on either bacterial or eukaryote cell [12,94,95].

4. Binding patterns for microbial RNases

Microbial RNases is another well characterized superfamily of endoribonucleases that share an α + β fold and includes small cytotoxic RNases secreted by a diversity of prokaryote and eukaryote microorganisms, as bacteria and fungi [96]. Here we have gathered all the protein nucleotide complexes corresponding to bacterial RNases (Barnase, RNase Sa, RNase Sa2) and fungal RNases (RNase F1, RNase T1, RNase U2, RNase MS; Restrictocin) (see Table S4).

All the structural analysis has been referred to RNase T1 (EC 3.1.27.3), the most studied superfamily member [97]. RNase T1 is a secreted RNase from the slime mould, *Aspergillus oryzae*, and was first purified from an enzyme extract obtained from the sake brewing process. RNase T1 also shows a general acid base catalytic mechanism of action [98], where the cleavage of the phosphodiester bond is catalysed by an His and a Glu residue (Glu58 and His92 acting as a base and acid groups respectively). Additional residues are contributing to anchor the substrate at p1, as Tyr38, His40, Arg77 (Fig. 2C), which are suggested to contribute to stabilize the transition state and facilitate the transphosphorylation mechanism [97].

We have studied the substrate binding site architecture for microbial RNases by analysing the available nucleotide complexes structures (Table S4). As most protein complexes submitted to the protein database are from mononucleotides, we could only analyse the conserved binding patterns for the main base and phosphate

XI

Table 4
Summary of the common identified active site residues in structure complexes from microbial RNases with nucleotide-type ligands, as detailed in Table S4. Residues were classified according to the interacting main nucleotide base and phosphate groups. Residues not present in all structures are indicated in brackets. PDB codes are indicated together with the residual mean standard deviation values in relation to RNase T1–GPG complex. PB refers to peptide bond.

Protein (PDB code)	Interaction block	
	Base	Phosphate
RNase T1–GPG (2rnt)	E46	Y38
	N43	H40
	43 (PB)	E58
	44 (PB)	H92
	F100	R77
		[N36]
RNase F1–2GP (1fut/0.88 Å)	E46	Y38
	N43	H40
	43 (PB)	E58
	44 (PB)	H92
		R77
		[N36]
RNase U2–3AM (3agn/1.43 Å)	E49	Y39
	44 (PB)	H41
	F110	E62
		H101
		R85
RNase Sa–3GP (2sar/2.79 Å)	E41	E54
	38 (PB)	H85
	39 (PB)	R69
	F37	[Y86]
		[Q32]
		R65
Binase–3GP (1goy/3.5 Å)	E59	[K26]
	56 (PB)	E72
	57 (PB)	[R82]
	F55	R86
		H101
		[Y102]

sites (B1 and p1). The conserved patterns for all analysed microbial RNases were considered inside each representative family (Fig. 5B). Microbial RNases show a base specificity for guanine at the main base site [99]. The common recognition patterns, taking RNase T1 as a reference, is illustrated in Fig. 3C. The observed peptide bond pair binding to guanine corresponds to one of the representative classified patterns [28]. Also, an anionic residue proving double binding to guanine N1 and N2 donor groups is observed, matching one of the preferred pairing. Another conserved residue at the base pocket is Phe. As for the main phosphate binding site, His, Glu and Tyr counterparts are conserved in all members. Together with the reference residues, the figure illustrates the contribution of complementary residues, as Arg77 and His40 counterparts. Table 4 illustrates the particularities that contribute to recognition of each building block. Manual inspection of overlapped RNase-3GP complexes confirms the matching of the key conserved residues within each representative family (Figs. 4B and 5B).

5. Conclusions

Overall screening of nucleotide complexes at the protein database outlines the main recognition binding strategies. A further systematic analysis of the available nucleotide complex structures from two representative secreted endoribonuclease superfamilies revealed their main common binding recognition features. Although the available protein–nucleotide complexes present an overrepresentation of mono- and dinucleotides ligand types, their individual binding mode can mimic the substrate flexibility and provide therefore a good model to study the particular interactions of each nucleotide building block at the substrate

binding sites. A thorough analysis of the structure complexes of RNase A superfamily members and the identification of evolutionary conserved residues revealed the main shared patterns and the divergence at secondary interacting regions. While the residues contributing to the principal phosphate and base binding sites are conserved, more freedom is offered by the secondary sites. Main recognition patterns for base selectivity were outlined and related to the reported kinetic data. The common conserved patterns were applied as analytical tools to interpret the peculiarities of homologous family members where no crystal complex is available, contributing to identify their putative physiological substrates that may account for their cytotoxic activities. The data provides a reference starting point for applied bioengineering towards the design of selective RNA binders.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.biochi.2012.12.015.

References

[1] J.J. Beintema, R.G. Kleineidam, The ribonuclease A superfamily: general discussion, Cellular and Molecular Life Sciences 54 (1998) 825–832.
[2] E. Boix, M.V. Nogues, Mammalian antimicrobial proteins and peptides: overview on the RNase A superfamily members involved in innate host defence, Molecular BioSystems 3 (2007) 317–335.
[3] A. Malik, J.K. Batra, Antimicrobial activity of human eosinophil granule proteins: involvement in host defence against pathogens, Critical Reviews in Microbiology (2012).
[4] H.F. Rosenberg, RNase A ribonucleases and host defense: an evolving story, Journal of Leukocyte Biology 83 (2008) 1079–1087.
[5] S. Sorrentino, The eight human "canonical" ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins, FEBS Letters 584 (2010) 2194–2200.
[6] M. Simanski, B. Köten, J.M. Schröder, R. Gläser, J. Harder, Antimicrobial RNases in cutaneous defense, Journal of Innate Immunity 4 (2012) 241–247.
[7] D. Findlay, D.G. Herries, C.A. Ross, B.R. Rabin, A.P. Mathias, Active site and mechanism of action of bovine pancreatic ribonuclease, Nature 190 (1961) 781.
[8] R.T. Raines, Ribonuclease A, Chemical Reviews 98 (1998) 1045–1066.
[9] M.V. Nogues, M. Moussaoui, E. Boix, M. Vilanova, M. Ribo, C.M. Cuchillo, The contribution of noncatalytic phosphate-binding subsites to the mechanism of bovine pancreatic ribonuclease A, Cellular and Molecular Life Sciences 54 (1998) 766–774.
[10] C.M. Cuchillo, M.V. Nogues, R.T. Raines, Bovine pancreatic ribonuclease: fifty years of the first enzymatic reaction mechanism, Biochemistry 50 (2011) 7835–7841.
[11] G.R. Marshall, J.A. Feng, D.J. Kuster, Back to the future: ribonuclease A, Biopolymers 90 (2008) 259–277.
[12] E. Boix, V.A. Salazar, M. Torrent, D. Pulido, M.V. Nogués, M. Moussaoui, Structural determinants of the eosinophil cationic protein antimicrobial activity, Biological Chemistry 393 (2012) 801–815.
[13] S.K. Gupta, B.J. Haigh, F.J. Griffin, T.T. Wheeler, The mammalian secreted RNases: mechanisms of action in host defence, Innate Immunity (2012).
[14] S. Cho, J. Zhang, Zebrafish ribonucleases are bactericidal: implications for the origin of the vertebrate RNase A superfamily, Molecular Biology E 24 (2007) 1259–1268.
[15] E. Pizzo, A. Merlino, M. Turano, I.R. Krauss, F. Coscia, A. Zanfardino, M. Varcamonti, A. Furia, C. Giancola, L. Mallarella, F. Sica, G. D'Alessio, A new RNase sheds light on the RNase/angiogenin subfamily form zebrafish, Biochemical Journal 433 (2011) 345–355.
[16] E. Pizzo, G. D'Alessio, The success of the RNase scaffold in the advance of biosciences and in evolution, Gene 406 (2007) 8–12.
[17] L. Aravind, E.V. Koonin, A natural classification of ribonucleases, Ribonucleases, Part A 341 (2001) 3–28.
[18] S. Loverix, J. Steyaert, Ribonucleases: from prototypes to therapeutic targets? Current Medicinal Chemistry 10 (2003) 779–785.

[19] J. Steyaert, A decade of protein engineering on ribonuclease T-1-atomic dissection of the enzyme–substrate interactions, European Journal of Biochemistry 247 (1997) 1–11.

[20] N.M. Luscombe, R.A. Laskowski, J.M. Thornton, Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level, Nucleic Acids Research 29 (2001) 2860–2874.

[21] S. Jones, D.T.A. Daley, N.M. Luscombe, H.M. Berman, J.M. Thornton, Protein–RNA interactions: a structural analysis, Nucleic Acids Research 29 (2001) 943–954.

[22] P.F. Gherardini, G. Ausiello, R.B. Russell, M. Helmer-Citterich, Modular architecture of nucleotide-binding pockets, Nucleic Acids Research 38 (2010) 3809–3816.

[23] A. Brakoulias, R.M. Jackson, Towards a structural classification of phosphate binding sites in protein–nucleotide complexes: an automated all-against-all structural comparison using geometric matching, Proteins-Structure Function and Bioinformatics 56 (2004) 250–260.

[24] L. Parca, I. Mangone, P.F. Gherardini, G. Ausiello, M. Helmer-Citterich, Phosfinder: a web server for the identification of phosphate-binding sites on protein structures, Nucleic Acids Research 39 (2011) W278–W282.

[25] Y. Mandelgutfreund, O. Schueler, H. Margalit, Comprehensive analysis of hydrogen-bonds in regulatory protein DNA-complexes – in search of common principles, Journal of Molecular Biology 253 (1995) 370–382.

[26] S.A. Coulocheri, D.G. Pigis, K.A. Papavassiliou, A.G. Papavassiliou, Hydrogen bonds in protein–DNA complexes: where geometry meets plasticity, Biochimie 89 (2007) 1291–1303.

[27] N. Morozova, J. Allers, J. Myers, Y. Shamoo, Protein–RNA interactions: exploring binding patterns with a three-dimensional superposition analysis of high resolution structures, Bioinformatics 22 (2006) 2746–2752.

[28] J. Kondo, E. Westhof, Classification of pseudo pairs between nucleotide bases and amino acids by analysis of nucleotide–protein complexes, Nucleic Acids Research 39 (2011) 8628.

[29] K.A. Denessiouk, M.S. Johnson, "Acceptor-donor-acceptor" motifs recognize the Watson-Crick, Hoogsteen and sugar "donor-acceptor-donor" edges of adenine and adenosine-containing ligands, Journal of Molecular Biology 333 (2003) 1025–1043.

[30] J. Allers, Y. Shamoo, Structure-based analysis of protein–RNA interactions using the program ENTANGLE, Journal of Molecular Biology 311 (2001) 75–86.

[31] M. Terribilini, J.D. Sander, J.H. Lee, P. Zaback, R.L. Jernigan, V. Honavar, D. Dobbs, RNABindR: a server for analyzing and predicting RNA-binding sites in proteins, Nucleic Acids Research 35 (2007) W578–W584.

[32] D.E. Draper, Themes in RNA–protein recognition, Journal of Molecular Biology 293 (1999) 255–270.

[33] S. Velankar, G.J. Kleywegt, The Protein Data Bank in Europe (PDBe): bringing structure to biology, Acta Crystallographica Section D-Biological Crystallography 67 (2011) 324–330.

[34] S. Velankar, Y. Alhroub, C. Best, S. Caboche, M.J. Conroy, J.M. Dana, M.A. Fernandez Montecelo, G. van Ginkel, A. Golovin, S.P. Gore, A. Gutmanas, P. Haslam, P.M.S. Hendrickx, E. Heuson, M. Hirshberg, M. John, I. Lagerstedt, S. Mir, L.E. Newman, T.J. Oldfield, A. Patwardhan, L. Rinaldi, G. Sahni, E. Sanz-Garcia, S. Sen, R. Slowley, A. Suarez-Uruena, G.J. Swaminathan, M.F. Symmons, W.F. Vranken, M. Wainwright, G.J. Kleywegt, PDBe: Protein Data Bank in Europe, Nucleic Acids Research 40 (2012) D445–D452.

[35] S.M. Rybak, M.A.E. Arndt, T. Schirrmann, S. Dubel, J. Krauss, Ribonucleases and ImmunoRNases as anticancer drugs, Current Pharmaceutical Design 15 (2009) 2665–2675.

[36] W. Ardelt, B. Ardelt, Z. Darzynkiewicz, Ribonucleases as potential modalities in anticancer therapy, European Journal of Pharmacology 625 (2009) 181–189.

[37] A.A. Makarov, A. Kolchinsky, O.N. Ilinskaya, Binase and other microbial RNases as potential anticancer agents, Bioessays 30 (2008) 781–790.

[38] C.H. Schein, From housekeeper to microsurgeon: the diagnostic and therapeutic potential of ribonucleases, Nature Biotechnology 15 (1997) 529–536.

[39] T. Schirrmann, J. Krauss, M.A.E. Arndt, S.M. Rybak, S. Dubel, Targeted therapeutic RNases (ImmunoRNases), Expert Opinion on Biological Therapy 9 (2009) 79–95.

[40] E.F. Fang, T.B. Ng, Ribonucleases of different origins with a wide spectrum of medicinal applications, Biochimica et Biophysica Acta-Reviews on Cancer 1815 (2011) 65–74.

[41] A. Golovin, D. Dimitropoulos, T. Oldfield, A. Rachedi, K. Henrick, MSDsite: a database search and retrieval system for the analysis and viewing of bound ligands and active sites, Proteins-Structure Function and Bioinformatics 58 (2005) 190–199.

[42] A. Golovin, K. Henrick, MSDmotif: exploring protein sites and motifs, BMC Bioinformatics 9 (2008) 312.

[43] H. Ashkenazy, E. Erez, E. Martz, T. Pupko, N. Ben-Tal, ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids, Nucleic Acids Research 38 (2010) W529–W533.

[44] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallographica Section D-Biological Crystallography 60 (2004) 2126–2132.

[45] K. Arnold, L. Bordoli, J. Kopp, T. Schwede, The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, Bioinformatics 22 (2006) 195–201.

[46] M. Irie, F. Mikami, K. Monma, K. Ohgi, H. Watanabe, R. Yamaguchi, H. Nagase, Kinetic-studies on the cleavage of oligouridylic acids and poly-U by bovine pancreatic ribonuclease-A, Journal of Biochemistry 96 (1984) 89–96.

[47] M. Moussaoui, M.V. Nogues, A. Guasch, T. Barman, F. Travers, C.M. Cuchillo, The subsites structure of bovine pancreatic ribonuclease A accounts for the abnormal kinetic behavior with cytidine 2′,3′-cyclic phosphate, Journal of Biological Chemistry 273 (1998) 25565–25572.

[48] S. Sorrentino, M. Libonati, Human pancreatic-type and nonpancreatic-type ribonucleases: a direct side-by-side comparison of their catalytic properties, Archives of Biochemistry and Biophysics 312 (1994) 340–348.

[49] C.G. Mohan, E. Boix, H.R. Evans, Z. Nikolovski, M.V. Nogues, C.M. Cuchillo, K.R. Acharya, The crystal structure of eosinophil cationic protein in complex with 2′,5′-ADP at 2.0 A resolution reveals the details of the ribonucleolytic active site, Biochemistry 41 (2002) 12100–12106.

[50] D.D. Leonidas, E. Boix, R. Prill, M. Suzuki, R. Turton, K. Minson, G.J. Swaminathan, R.J. Youle, K.R. Acharya, Mapping the ribonucleolytic active site of eosinophil-derived neurotoxin (EDN). High resolution crystal structures of EDN complexes with adenylic nucleotide inhibitors, Journal of Biological Chemistry 276 (2001) 15009–15017.

[51] M.D. Baker, D.E. Holloway, G.J. Swaminathan, K.R. Acharya, Crystal structures of eosinophil-derived neurotoxin (EDN) in complex with the inhibitors 5′-ATP, Ap3A, Ap4A, and Ap5A, Biochemistry 45 (2006) 416–426.

[52] E. Boix, Z. Nikolovski, G.P. Moiseyev, H.F. Rosenberg, C.M. Cuchillo, M.V. Nogues, Kinetic and product distribution analysis of human eosinophil cationic protein indicates a subsite arrangement that favors exonuclease-type activity, Journal of Biological Chemistry 274 (1999) 15605–15614.

[53] D. Sikriwal, D. Seth, P. Dey, J.K. Batra, Human eosinophil-derived neurotoxin: involvement of a putative non-catalytic phosphate-binding subsite in its catalysis, Molecular and Cellular Biochemistry 303 (2007) 175–181.

[54] D. Sikriwal, D. Seth, J.K. Batra, Role of catalytic and non-catalytic subsite residues in ribonuclease activity of human eosinophil-derived neurotoxin, Biological Chemistry 390 (2009) 225–234.

[55] S.C. Mosimann, D.L. Newton, R.J. Youle, M.N. James, X-ray crystallographic structure of recombinant eosinophil-derived neurotoxin at 1.83 A resolution, Journal of Molecular Biology 260 (1996) 540–552.

[56] G.J. Swaminathan, D.E. Holloway, K. Veluraja, K.R. Acharya, Atomic resolution (0.98 A) structure of eosinophil-derived neurotoxin, Biochemistry 41 (2002) 3341–3352.

[57] E. Boix, D. Pulido, M. Moussaoui, M.V. Nogues, S. Russi, The sulfate-binding site structure of the human eosinophil cationic protein as revealed by a new crystal form, Journal of Structural Biology 179 (2012) 1–9.

[58] E. Boix, M.V. Nogues, C.H. Schein, S.A. Benner, C.M. Cuchillo, Reverse transphosphorylation by ribonuclease A needs an intact p2-binding site. Point mutations at Lys-7 and Arg-10 alter the catalytic properties of the enzyme, Journal of Biological Chemistry 269 (1994) 2529–2534.

[59] B.M. Fisher, J.E. Grilley, R.T. Raines, A new remote subsite in ribonuclease A, Journal of Biological Chemistry 273 (1998) 34134–34138.

[60] G.J. Swaminathan, D.G. Myszka, P.S. Katsamba, L.E. Ohnuki, G.J. Gleich, K.R. Acharya, Eosinophil-granule major basic protein, a C-type lectin, binds heparin, Biochemistry 44 (2005) 14152–14158.

[61] E. Boix, M. Torrent, M. Nogués, V. Salazar, Searching for heparin binding patterns, in: D. Piyathilake, R. Liang (Eds.), Heparin: Properties, Uses and Side Effects, Nova Sciences Publishers, Inc., 2011.

[62] R. Shapiro, J.W. Fett, D.J. Strydom, B.L. Vallee, Isolation and characterization of a human-colon carcinoma-secreted enzyme with pancreatic ribonuclease-like activity, Biochemistry 25 (1986) 7255–7264.

[63] M.E. Eftink, R.L. Biltonen, Pancreatic ribonuclease A: the most studied endoribonuclease, in: A. Neuberger, K. Brocklehurst (Eds.), Hydrolytic Enzymes, Elsevier Science Publishers B.V., Amsterdam, 1987, pp. 333–376.

[64] A. Fersht, Structures and Mechanism of Selected Enzymes, , In: Ribonuclease. Enzyme Structure and Mechanism, W. H. Freeman and Co., New York, 1985, pp. 426–431.

[65] S.S. Terzyan, R. Peracaula, R. de Llorens, Y. Tsushima, H. Yamada, M. Seno, F.X. Gomis-Ruth, M. Coll, The three-dimensional structure of human RNase 4, unliganded and complexed with d(Up), reveals the basis for its uridine selectivity, Journal of Molecular Biology 285 (1999) 205–214.

[66] H.M. Zhou, D.J. Strydom, The amino-acid-sequence of human ribonuclease 4, a highly conserved ribonuclease that cleaves specifically on the 3′ side of uridine, European Journal of Biochemistry 217 (1993) 401–410.

[67] A.M. Vicentini, Z. KoteJarai, J. Hofsteenge, Structural determinants of the uridine-preferring specificity of RNase PL3, Biochemistry 35 (1996) 9128–9132.

[68] S.Y. Wodak, M.Y. Liu, H.W. Wyckoff, Structure of cytidilyl(2′,5′)adenosine when bound to pancreatic ribonuclease-S, Journal of Molecular Biology 116 (1977) 855–875.

[69] I. Zegers, D. Maes, M.H. Dao-Thi, F. Poortmans, R. Palmer, L. Wyns, The structures of RNase A complexed with 3′-CMP and d(CpA): active site conformation and conserved water molecules, Protein Science 3 (1994) 2322–2339.

[70] C. Toiron, C. Gonzalez, M. Bruix, M. Rico, Three-dimensional structure of the complexes of ribonuclease A with 2′,5′-CpA and 3′,5′-d(CpA) in aqueous solution, as obtained by NMR and restrained molecular dynamics, Protein Science 5 (1996) 1633–1647.

[71] A. Tarragonafiol, H.J. Eggelte, S. Harbron, E. Sanchez, C.J. Taylorson, J.M. Ward, B.R. Rabin, Identification by site-directed mutagenesis of amino-acids in the

B2 subsite of bovine pancreatic ribonuclease-A, Protein Engineering 6 (1993) 901–906.

[72] L. Vitagliano, A. Merlino, A. Zagari, L. Mazzarella, Productive and nonproductive binding to ribonuclease A: X-ray structure of two complexes with uridylyl(2′,5′)guanosine, Protein Science 9 (2000) 1217–1225.

[73] J.N. Lisgarten, D. Maes, L. Wyns, C.F. Aguilar, R.A. Palmer, Structure of the crystalline complex of deoxycytidylyl-3′,5′-guanosine (3′,5′-Dcpdg) cocrystallized with ribonuclease at 1.9-Angstrom resolution, Acta Crystallographica Section D-Biological Crystallography 51 (1995) 767–771.

[74] J.C. Fontecilla-Camps, R. de Llorens, M.H. le Du, C.M. Cuchillo, Crystal structure of ribonuclease A.d(ApTpApApG) complex. Direct evidence for extended substrate recognition, Journal of Biological Chemistry 269 (1994) 21526–21531.

[75] J. Zhang, K.D. Dyer, H.F. Rosenberg, RNase 8, a novel RNase A superfamily ribonuclease expressed uniquely in placenta, Nucleic Acids Research 30 (2002) 1169–1175.

[76] C.C. Chan, J.M. Moser, K.D. Dyer, C.M. Percopo, H.F. Rosenberg, Genetic diversity of human RNase 8, BMC Genomics 13 (2012).

[77] G. D'Alessio, A. Di Donato, R. Piccoli, N. Russo, Seminal ribonuclease: preparation of natural and recombinant enzyme, quaternary isoforms, isoenzymes, monomeric forms; assay for selective cytotoxicity of the enzyme, Ribonucleases, Part A 341 (2001) 248–263.

[78] R. Spadaccini, C. Ercole, M.A. Gentile, D. Sanfelice, R. Boelens, R. Wechselberger, G. Batta, A. Bernini, N. Niccolai, D. Picone, NMR studies on structure and dynamics of the monomeric derivative of BS-RNase: new insights for 3D domain swapping, PloS One 7 (2012).

[79] K. Dossi, V.G. Tsirkone, J.M. Hayes, J. Matousek, P. Pouckova, J. Soucek, M. Zadinova, S.E. Zographos, D.D. Leonidas, Mapping the ribonucleolytic active site of bovine seminal ribonuclease. The binding of pyrimidinyl phosphonucleotide inhibitors, European Journal of Medicinal Chemistry 44 (2009) 4496–4508.

[80] C. Ercole, R.A. Colamarino, E. Pizzo, F. Fogolari, R. Spadaccini, D. Picone, Comparison of the structural and functional properties of RNase a and BS-RNase: a stepwise mutagenesis approach, Biopolymers 91 (2009) 1009–1017.

[81] C.H. Hsu, Y.D. Liao, Y.R. Pan, L.W. Chen, S.H. Wu, Y.J. Leu, C.P. Chen, Solution structure of the cytotoxic RNase 4 from oocytes of bullfrog Rana catesbeiana, Journal of Molecular Biology 326 (2003) 1189–1201.

[82] D.E. Holloway, U.P. Singh, K. Shogen, K.R. Acharya, Crystal structure of onconase at 1.1 angstrom resolution – insights into substrate binding and collective motion, FEBS Journal 278 (2011) 4136–4149.

[83] M. Irie, K. Nitta, T. Nonaka, Biochemistry of frog ribonucleases, Cellular and Molecular Life Sciences 54 (1998) 775–784.

[84] U.P. Singh, W. Ardelt, S.K. Saxena, D.E. Holloway, E. Vidunas, H.S. Lee, A. Saxena, K. Shogen, K.R. Acharya, Enzymatic and structural characterisation of amphinase, a novel cytotoxic ribonuclease from Rana pipiens oocytes, Journal of Molecular Biology 371 (2007) 93–111.

[85] E. Boix, Y. Wu, V.M. Vasandani, S.K. Saxena, W. Ardelt, J. Ladner, R.J. Youle, Role of the N terminus in RNase A homologues: differences in catalytic activity, ribonuclease inhibitor interaction and cytotoxicity, Journal of Molecular Biology 257 (1996) 992–1007.

[86] C.F. Chang, C.P. Chen, Y.C. Chen, K. Hom, R.F. Huang, T.H. Huang, The solution structure of a cytotoxic ribonuclease from the oocytes of Rana catesbeiana (bullfrog), Journal of Molecular Biology 283 (1998) 231–244.

[87] E. Pizzo, M. Varcamonti, A. Di Maro, A. Zanfardino, C. Giancola, G. D'Alessio, Ribonucleases with angiogenic and bactericidal activities from the Atlantic salmon, FEBS Journal 275 (2008) 1283–1295.

[88] J.F. Riordan, Angiogenin, Ribonucleases, Part A 341 (2001) 263–273.

[89] K.R. Acharya, R. Shapiro, S.C. Allen, J.F. Riordan, B.L. Vallee, Crystal-structure of human angiogenin reveals the structural basis for its functional divergence from ribonuclease, Proceedings of the National Academy of Sciences of the United States of America 91 (1994) 2915–2919.

[90] T.C. Fan, S.L. Fang, C.S. Hwang, C.Y. Hsu, X.A. Lu, S.C. Hung, S.C. Lin, D.T. Chang, Characterization of molecular interactions between eosinophil cationic protein and heparin, Journal of Biological Chemistry 283 (2008) 25468–25474.

[91] M. Torrent, M.V. Nogues, E. Boix, Eosinophil cationic protein (ECP) can bind heparin and other glycosaminoglycans through its RNase active site, Journal of Molecular Recognition 24 (2011) 90–100.

[92] M.F. García-Mayoral, M. Moussaoui, B.G. de la Torre, D. Andreu, E. Boix, M.V. Nogués, M. Rico, D.V. Laurents, M. Bruix, NMR structural determinants of eosinophil cationic protein binding to membrane and heparin mimetics, Biophysical Journal 98 (2010) 2702–2711.

[93] M.F. García-Mayoral, A. Canales, D. Díaz, J. López-Prados, M. Moussaoui, J.L. de Paz, J. Angulo, P.M. Nieto, J. Jiménez-Barbero, E. Boix, M. Bruix, Insights into the glycosaminoglycan-mediated cytotoxic mechanism of eosinophil cationic protein revealed by NMR, ACS Chemical Biology (in press).

[94] D. Pulido, M. Moussaoui, D. Andreu, M.V. Nogues, M. Torrent, E. Boix, Antimicrobial action and cell agglutination by the eosinophil cationic protein are modulated by the cell wall lipopolysaccharide structure, Antimicrobial Agents and Chemotherapy 56 (2012) 2378–2385.

[95] S. Navarro, J. Aleu, M. Jimenez, E. Boix, C.M. Cuchillo, M.V. Nogues, The cytotoxicity of eosinophil cationic protein/ribonuclease 3 on eukaryotic cell lines takes place through its aggregation on the cell membrane, Cellular and Molecular Life Sciences 65 (2008) 324–337.

[96] A.A. Makarov, O.N. Ilinskaya, Cytotoxic ribonucleases: molecular weapons and their targets, FEBS Letters 540 (2003) 15–20.

[97] S. Loverix, J. Steyaert, Deciphering the mechanism of RNase T1, Ribonucleases, Part A 341 (2001) 305–323.

[98] I. Zegers, R. Loris, G. Dehollander, A.F. Haikal, F. Poortmans, J. Steyaert, L. Wyns, Hydrolysis of a slow cyclic thiophosphate substrate of RNase T1 analyzed by time-resolved crystallography, Nature Structural Biology 5 (1998) 280–283.

[99] G.I. Yakovlev, G.P. Moiseyev, S.I. Bezborodova, V. Both, J. Sevcik, A comparative-study on the catalytic properties of guanyl-specific ribonucleases, European Journal of Biochemistry 204 (1992) 187–190.

[100] P. Gouet, E. Courcelle, D.I. Stuart, F. Metoz, ESPript: analysis of multiple sequence alignments in PostScript, Bioinformatics 15 (1999) 305–308.

[101] F.M. Richards, H.W. Wyckoff, Bovine pancreatic ribonuclease, in: P.B. Boyer (Ed.), Enzymes, Academic Press, New York, 1971, pp. 647–806.

[102] H. Witzel, E.A. Barnard, Mechanism and binding sites in ribonuclease reaction. 1. Kinetic studies on second step of reaction, Biochemical and Biophysical Research Communications 7 (1962) 289.

[103] P.L. Ipata, R.A. Feliciol, A spectrophotometric assay for ribonuclease activity using cytidylyl-(3′,5′)-adenosine and uridylyl-(3′,5′)-adenosine as substrates, FEBS Letters 1 (1968) 29–31.

[104] S. Sorrentino, Human extracellular ribonucleases: multiplicity, molecular diversity and catalytic properties of the major RNase types, Cellular and Molecular Life Sciences 54 (1998) 785–794.

[105] J. Hofsteenge, A. Vicentini, O. Zelenko, Ribonuclease 4, an evolutionarily highly conserved member of the superfamily, Cellular and Molecular Life Sciences 54 (1998) 804–810.

[106] W. Zhao, Z. Kote-Jarai, Y. van Santen, J. Hofsteenge, J.J. Beintema, Ribonucleases from rat and bovine liver: purification, specificity and structural characterization, Biochimica et Biophysica Acta-Protein Structure and Molecular Enzymology 1384 (1998) 55–65.

[107] W. Ardelt, H.S. Lee, G. Randolph, A. Viera, S.M. Mikulski, K. Shogen, Enzymatic characterization of onconase, a novel ribonuclease with anti-tumor activity, Protein Science (1994) 137–147.
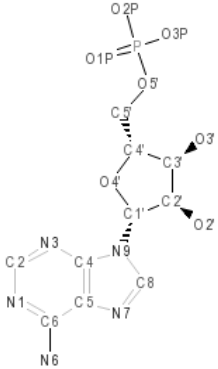
XIV

Nucleotide binding architecture for secreted cytotoxic endoribonucleases

Boix, E.; Blanco, J. A.; Nogués, M. V.; Moussaoui, M.

Department of Biochemistry and Molecular Biology, Biosciences Faculty, Universitat Autònoma de Barcelona, E-08193 Cerdanyola del Vallès, Spain

Supplemental material

Table S1: Pair binding statistics for protein interaction with nucleotide-type ligands. The data was retrieved using the pair binding statistics search tool implemented in the *PDBeMotif* web server. Below each interacting residue the total number of protein structure complexes is indicated (number of corresponding UniProt entries/ total number of analysed structures from the *Protein Data Bank*. Only ligands with significant amount of statistical data available have been selected. For each ligand, the first 5 residues with the highest number of interactions were listed, and for each amino acid, only the atomic interactions with significance values over 80% have been included.

Code	Ligand Name and Image	Hydrogen bond interaction information			
		Interacting residue	*Ligand interacting atom*	*Interacting atom in residue side chain*	*Significance*
AMP	Adenosine monophosphate	**R** 1386/22325	N1	NH1	**97**
				NH2	**90**
			N3	NH1	**97**
				NH2	**93**
			N7	NH1	**97**
				NH2	**80**
			O2'	NE	**93**
				NH1	**99**
				NH2	**98**
			O3'	NH1	**99**
				NH2	**98**
			O4'	NH1	**80**
				NH2	**80**
			O5'	NH1	**97**
				NH2	**98**
			O1P	NE	**98**
				NH1	**99**
			O2P	NH2	**99**
				NE	**98**
				NH1	**99**
			O3P	NH2	**99**
				NE	**99**
				NH1	**99**
		K 105/682	N1	NZ	**90**
			N7	NZ	**95**
			O3'	NZ	**90**
			O4'	NZ	**96**
			O5'	NZ	**98**
			O1P	NZ	**99**
			O2P	NZ	**99**
			O3P	NZ	**99**
		H 85/932	N7	ND1	**91**
			O4'	ND1	**91**
			O5'	ND1	**94**
			O1P	ND1	**98**
				NE2	**99**
			O2P	ND1	**94**
				NE2	**99**
			O3P	ND1	**94**
				NE2	**99**
		E 69/111	N6	OE1	**97**
				OE2	**96**
			O2'	OE1	**96**
				OE2	**96**
			O3'	OE1	**97**
				OE2	**96**
			O2P	OE2	**86**
		S 151/1385	O3'	OG	**92**
			O5'	OG	**95**
			O1P	OG	**98**
			O2P	OG	**99**
			O3P	OG	**98**

A2P	Adenosine-2'-5'-diphosphate	Q 6/9	O2P	NE2	92
		H 4/116	O1P	ND1	81
			O3P	NE2	81
			O5P	NE2	94
		K 5/47	O2P	NZ	94
			O3P	NZ	82
			O5P	NZ	82
			O6P	NZ	82
		T 5/47	O3'	OG1	87
			O2P	OG1	87
			O3P	OG1	87
		S 5/18	O3'	OG	82
			O2P	OG	82
			O3P	OG	94

A3P	Adenosine-3'-5'-diphosphate	K 204/2707	O4P	NZ	98
			O5P	NZ	97
			O6P	NZ	99
		S 181/1962	O1P	OG	99
			O2P	OG	98
			O3P	OG	96
		T 106/904	O4P	OG1	99
			O5P	OG1	99
			O6P	OG1	99
		Y 18/19	N3	OH	92
		H 68/852	O3'	NE2	85
			O4'	NE2	85
			O5'	ND1	95
			O1P	NE2	97
			O2P	NE2	97
			O5P	ND1	95
			O6P	ND1	85

ADP		N1	NZ	**91**
		N3	NZ	**86**
		N7	NZ	**99**
		O2′	NZ	**100**
		O3′	NZ	**99**
	K	O4′	NZ	**95**
	1194/14064	O5	NZ	**91**
		O1A	NZ	**100**
		O2A	NZ	**100**
		O3A	NZ	**100**
		O1B	NZ	**100**
		O2B	NZ	**100**
		O3B	NZ	**100**
		O2′	OG1	**92**
		O5′	OG1	**91**
	T	O2A	OG1	**99**
	378/3856	O3A	OG1	**93**
		O1B	OG1	**99**
		O2B	OG1	**99**
		O3B	OG1	**98**
		N1	OG	**99**
		N6	OG	**99**
		N7	OG	**96**
		O2′	OG	**98**
		O3′	OG	**98**
		O4′	OG	**92**
	S	O5′	OG	**96**
	1668/23265	O1A	OG	**99**
		O2A	OG	**99**
		O3A	OG	**98**
		O1B	OG	**100**
		O2B	OG	**100**
		O3B	OG	**100**
		N1	ND2	**80**
		N6	OD1	**99**
		N7	ND2	**99**
		O2′	ND2	**97**
			OD1	**92**
		O3′	ND2	**98**
	N	O4′	ND2	**92**
	155/302	O1A	ND2	**95**
		O2A	ND2	**99**
		O2B	ND2	**96**
		O3B	ND2	**98**
		N6	OD1	**97**
			OD2	**98**
		O2′	OD1	**99**
			OD2	**99**
		O3′	OD1	**99**
	D		OD2	**99**
	170/310	O1A	OD2	**89**
		O1B	OD2	**83**
		O2B	OD2	**92**
		O3B	OD1	**83**
			OD2	**92**

Adenosine-5'-diphosphate

ATP	Residue	ATP atom	Residue atom	%
ATP *Adenosine-5'-triphosphate*	**K** 264/3194	N7	NZ	98
		O2'	NZ	98
		O3'	NZ	98
		O1A	NZ	99
		O2A	NZ	99
		O3A	NZ	98
		O1B	NZ	99
		O2B	NZ	99
		O3B	NZ	98
		O1G	NZ	99
		O2G	NZ	99
		O3G	NZ	99
	S 255/3588	N6	OG	91
		O2'	OG	95
		O3'	OG	83
		O1A	OG	97
		O2A	OG	97
		O3A	OG	87
		O1B	OG	99
		O2B	OG	98
		O3B	OG	99
		O1G	OG	99
		O2G	OG	99
		O3G	OG	99
	R 232/340	N7	NH1	95
		O2'	NH1	96
			NH2	90
		O5'	NH1	84
			NH2	95
		O1A	NH1	98
			NH2	99
		O2A	NH2	99
		O3A	NH2	96
		O1B	NE	91
			NH1	97
			NH2	98
		O3B	NH1	96
			NH2	98
		O1G	NE	87
			NH1	99
			NH2	99
		O2G	NE	96
			NH1	99
			NH2	99
		O3G	NE	97
			NH1	99
			NH2	99
	T 197/1707	N3	OG1	86
		N6	OG1	89
		O2'	OG1	97
		O3'	OG1	93
		O4'	OG1	86
		O1A	OG1	98
		O2A	OG1	99
		O1B	OG1	98
		O2B	OG1	99
		O3B	OG1	86
		O1G	OG1	98
		O2G	OG1	98
		O3G	OG1	97
	D 143/304	N6	OD2	93
		O2'	OD1	98
			OD2	98
		O3'	OD1	99
			OD2	98
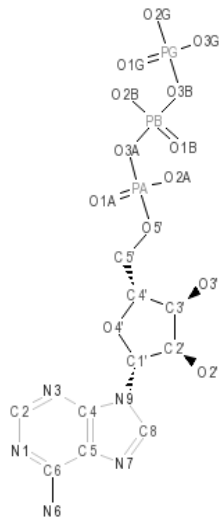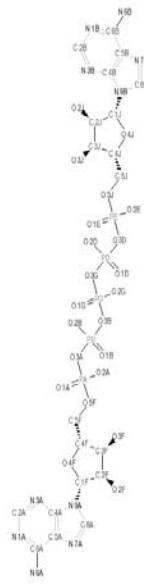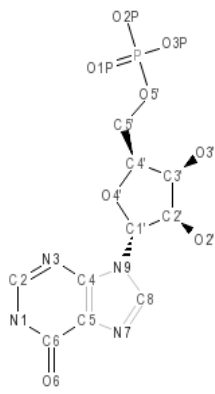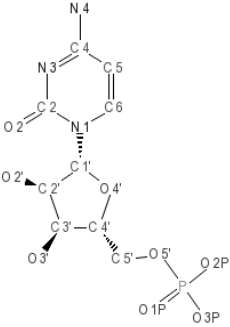
		O2A	OD2	**95**
		O1B	OD2	**95**
		O2B	OD2	**93**
		O2G	OD1	**91**
			OD2	**98**
		O3G	OD1	**86**
			OD2	**98**
	E	N6	OE1	**96**
			OE2	**98**
	127/246	O2'	OE1	**98**
			OE2	**99**
		O3'	OE2	**99**
		O1B	OE2	**88**
		O2B	OE2	**92**
		O1G	OE2	**88**
		O3G	OE1	**88**
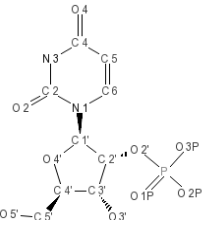
PAP	*3'-phosphate-adenosine-5'-diphosphate*	**H** **39/78**	O1A	NE2	**97**
		K **682/1359**	O1A	NZ	**97**
			O2A	NZ	**99**
			O2B	NZ	**99**
			O3B	NZ	**93**
		R **2/10**	O2'	NH2	**94**
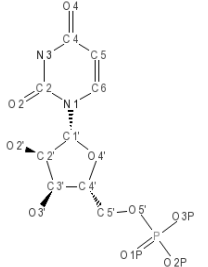
B4P	*Bis(adenosine)-5'-tetraphosphate*	**R** **232/3401**	O1A	NH1	**86**
				NH2	**97**
			O1B	NH1	**97**
				NH2	**99**
			O1D	NH2	**97**
			O1G	NH1	**97**
				NH2	**98**
			O2A	NH1	**95**
				NH2	**86**
			O3A	NH2	**95**
		K **6/55**	O2A	NZ	**91**
			O2G	NZ	**96**
		S **6/72**	O2E	OG	**94**
			O2F	OG	**94**

AP5 — Bis(adenosine)-5'-pentaphosphate

Residue	Ligand atom	Protein atom	%
R 14/623	O1D	NH1	97
		NH2	98
	O1E	NH1	98
		NH2	98
	O1G	NH1	99
		NH2	99
	O2A	NH1	98
		NH2	97
	O2D	NH1	94
		NH2	96
	O2E	NH1	91
		NH2	97
	O2G	NH1	98
		NH2	97
	O3D	NH1	96
		NH2	94
	O3G	NH1	80
T 12/174	N7A	OG1	98
	N7B	OG1	98
	O1A	OG1	93
	O2A	OG1	87
	O2E	OG1	87
K 13/165	O1B	NZ	98
	O1D	NZ	99
	O2D	NZ	85
	O3G	NZ	85
H 13/92	N6A	OG1	83
	N7A	OG1	96
	N7B	OG1	98
	O1A	OG1	99
	O2A	OG1	93
		OG	83
	O2E	OG1	93

IMP — Iosinic acid

Residue	Ligand atom	Protein atom	%
D 18/42	N1	OD1	94
	O2'	OD2	98
	O3'	OD1	98
K 11/19	O6	NZ	97
	O3'	NZ	84
	O2P	NZ	95
R 11/27	N7	NH2	95
	O6	NH1	87
		NH2	87
	O2'	NH2	87
	O1P	NH2	87
	O2P	NH1	87
		NH2	87
	O3P	NH2	87
S 16/59	O1P	OG1	96
	O2P	OG1	94
	O3P	OG	94
		OG1	94
T 10/31	O1P	OG1	95
	O2P	OG1	92
	O3P	OG1	92
Y 8/15	O1P	OH	97
	O3P	OH	97

C5P

Cytidine 5'-monophosphate

Residue	Atom	Partner	%
K **41/74**	O2'	NZ	93
	O2P	NZ	93
	O3P	NZ	98
R **35/146**	N3	NH2	97
	O2	NH1	95
	O2'	NH1	87
		NH2	87
	O3'	NH1	87
	O1P	NH1	87
		NH2	87
	O2P	NH1	87
		NH2	97
	O3P	NE	87
		NH1	97
		NH2	98
D **33/69**	O2'	OD1	95
		OD2	95
	O3'	OD1	98
		OD2	95
	O2P	OD2	87
T **37/99**	N3	OG1	95
	O2	OG1	88
	O2'	OG1	88
	O1P	OG1	95
	O2P	OG1	88
	O3P	OG1	97
S **54/66**	N3	OG	97
		OG1	96
	O2	OG1	89
	O2'	OG1	96
	O1P	OG1	89
	O2P	OG	89
		OG1	89
	O3P	OG	97
		OG1	97
H **19/39**	O1P	NE2	93
	O2P	NE2	93

U2P

Uridine-2'-phosphate

Residue	Atom	Partner	%
H **1/4**	O3P	NE2	93
K **2/2**	O2'	NZ	96
	O2P	NZ	96
T **2/5**	N3	OG1	97

U5P

Uridine-5'-phosphate

Residue	Atom	Partner	%
R **24/65**	O5'	NH2	83
	O1P	NE	93
		NH1	96
		NH2	97
	O2P	NE	83
		NH1	96
		NH2	96
	O3P	NE	83
		NH1	83
		NH2	83

Uridine-5'-dinophosphate

UDP

Residue	Atom 1	Atom 2	%
K **39/360**	O2'	NZ	97
	O3'	NZ	90
	O1A	NZ	97
	O2A	NZ	98
	O3A	NZ	95
	O1B	NZ	99
	O2B	NZ	99
	O3B	NZ	99
D **31/59**	N3	OD1	92
		OD2	97
	O2'	OD1	97
		OD2	96
	O3'	OD2	92
E **31/59**	O2'	OE1	99
		OE2	91
	O3'	OE1	91
		OE2	98
S **48/113**	N3	OG1	92
	O4	OG	92
		OG1	80
	O2'	OG	95
		OG1	80
	O3'	OG	80
	O5'	OG1	80
	O2A	OG	98
	O1B	OG	98
		OG1	92
	O2B	OG	80
		OG1	92
	O3B	OG	80
T **40/294**	N3	OG1	95
	O2	OG1	93
	O4	OG1	93
	O2'	OG1	87
	O5'	OG1	96
	O2A	OG1	96
	O1B	OG1	95
	O2B	OG1	97
	O3B	OG1	93

XXIII
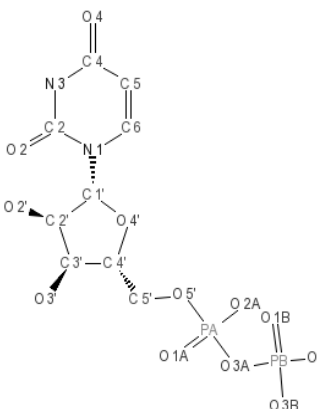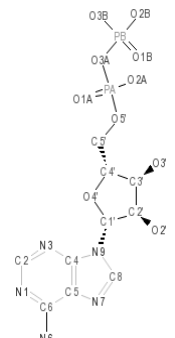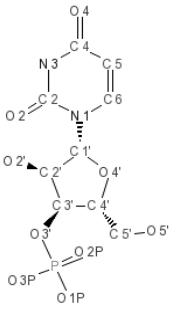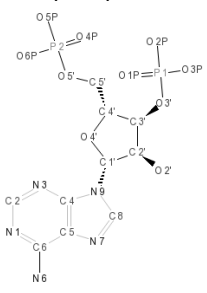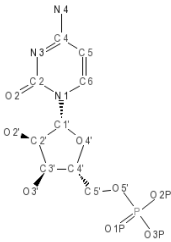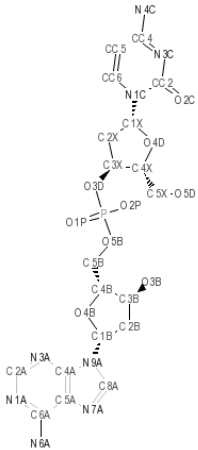
Table S2: Pair binding statistics for RNase A family canonical members with selected nucleotide-type ligands. Atomic interactions retrieved using the *PDBeMotif* web server are listed for both *Van der Waals* and hydrogen bonds. Protein atoms are labeled according to each molecule at the asymmetric unit.

Protein Name	Ligand Information	PDB Code	Pair bonding interacting atoms			
RNase A	**ADP** *Adenosine-5'-diphosphate*	1o0h	**PB**	NE2.HIS 12A	**PB**	NE2.GLN 11B
				ND1.HIS 119A		ND1.HIS 11B
			O1B	CD2.HIS 12A	**O1B**	CD2.HIS 12B
				CE1.HIS 12A		NE2.HIS 12B
				NE2.HIS 12A		CA.HIS 119B
				CA.HIS 119A	**O2B**	ND1.HIS 11B
			O2B	CE.LYS 41A		CE1.HIS 119B
				NZ.LYS 41A		CG.HIS 119B
			O3B	ND1.HIS 119A	**O3B**	NE2.GLN 11B
				CE1.HIS 119A	**PA**	NZ.LYS 41B
				CG.HIS 119A	**O1A**	CE1.HIS 12B
			PA	NZ.LYS 7A		NE2.HIS 12B
				NE2.GLN 11A		CE.LYS 41B
			O1A	NE2.GLN 11A		NZ.LYS 41B
				NZ.LYS 7A	**C8**	ND1.HIS 11B
			O3A	NE2.GLN 11A	**N7**	CB.HIS 119B
			O5'	NZ.LYS 7A		CG.HIS 119B
			C5'	CB.HIS 119A		ND1.HIS 11B
			C8	CD.GLU 111A	**C5**	CB.HIS 119B
				OE2.GLU 111A		CE1.HIS 119B
				CG2.VAL 118A		CG.HIS 119B
			N7	ND2.ASN 71A		ND1.HIS 119B
				CB.ALA 109A	**C6**	ND2.ASN 67B
			C5'	OE1.GLN 69A		CD2.HIS 119B
				CB.ALA 109A		CG.HIS 119B
			C6	ND2.ASN 67A	**N6**	CD2.HIS 119B
				OE1.GLN 69A		CG.HIS 119B
				CB.ALA 109A	**N1**	CG.ASN 67B
			N6	OD1.ASN 71A		ND2.ASN 67B
				CB.CYS 65A	**C2**	ND2.ASN 67B
				SG.CYS 65A	**N3**	CE1.HIS 119B
				OD1.ASN 67A	**C4**	CE1.HIS 119B
				OE1.GLN 69A		ND1.HIS 11B
				CB.ALA 109A		
			N1	CG.ASN 67A		
				ND2.ASN 67A		
				CD2.HIS 119A		
			C2	CD2.HIS 119A		
				CE1.HIS 119A		
				CG.HIS 119A		
				ND1.HIS 119A		
				NE2.HIS 119A		
			N3	CG.HIS 119A		
			C4	CB.HIS 119A		
				CG.HIS 119A		
	U3P *Uridine-3'-phosphate*	1o0n	**C2**	O1.THR 45A	**C2**	CD1.PHE 120B
				CD2.PHE 120A	**N3**	CB.THR 45B
				CB.THR 45A		OG1.THR 45B
			N3	OG1.THR 45A		CD1.PHE 120B
				CD2.PHE 120A		CE1.PHE 120B
				CE2.PHE 120A	**C4**	CD1.PHE 120B
			C4	OG1.THR 45A		CE1.PHE 120B
				CE2.PHE 120A	**O2**	CA.ASN 44B
				CE1.HIS 12A	**O4**	CE1.PHE 120B
			O2	CA.ASN 44A	**P**	CE1.HIS 119B

XXIV

Top entry (continued)

Atom	Contact (A)	Atom	Contact (B)
	OG1.THR 45A		CG.HIS 119B
O4	CE1.HIS 119A		ND1.HIS 119B
	ND1.HIS 119A	O1P	CA.HIS 119B
	NE2.GLN 11A		CD2.HIS 119B
O1P	NZ.LYS 41A		CE1.HIS 119B
	CE.LYS 41A		CG.HIS 119B
	NZ.LYS 41A		ND1.HIS 119B
C2'	CE1.HIS 119A		NE2.HIS 119B
	CE1.HIS 12A		CA.PHE 120B
	CE1.HIS 119A	O2P	NE2.GLN 11B
O2'	CB.PHE 120A		CE1.HIS 12B
			NE2.HIS 12B
		O3P	ND1.HIS 119B
			CE1.HIS 119B
			CG.HIS 119B
		O3'	CE1.HIS 119B
		C2'	CE1.HIS 12B
			NZ.LYS 41B
		O2'	NE2.HIS 12B
			CE.LYS 41B
			NZ.LYS 41B

A3P — *Adenosine-3'-5'-diphosphate* — 1o0f

Atom	Contact (A)	Atom	Contact (B)
P2	NE2.HIS 12A	P1	NZ.LYS 7B
	ND1.HIS 119A	O1P	NZ.LYS 7B
O4P	CD2.HIS 12A		CE.LYS 7B
	CE1.HIS 12A	P2	NE2.HIS 12B
	NE2.HIS 12A		CD2.HIS 119B
	CA.HIS 119A	O4P	CD2.HIS 12B
O5P	NE2.GLN 11A		NE2.HIS 12B
	NZ.LYS 41A		CA.HIS 119B
	CD.GLN 11A	O5P	CD2.HIS 119B
O6P	CE1.HIS 119A	O6P	NE2.GLN 11B
	ND1.HIS 119A	C5'	CA.HIS 119B
O5'	ND1.HIS 119A		CB.HIS 119B
	CE1.HIS 119A		CD2.HIS 119B
C5'	CB.HIS 119A		CG.HIS 119B
	ND1.HIS 119A	C4'	CD2.HIS 119B
O4'	CB.HIS 119A	O4'	CD2.HIS 119B
N9	CG.HIS 119A		CG.HIS 119B
C8	CD2.HIS 119A	C2'	CD2.HIS 119B
	CE1.HIS 119A	C1'	CD2.HIS 119B
	CG.HIS 119A		CG.HIS 119B
	ND1.HIS 119A	N9	CB.HIS 119B
	NE2.HIS 119A		CD2.HIS 119B
N7	ND2.ASN 67A		CG.HIS 119B
	CD2.HIS 119A	C8	CD2.HIS 119B
	NE2.HIS 119A		CE1.HIS 119B
C5	CD2.HIS 119A		CG.HIS 119B
C6	OE1.GLN 69A		ND1.HIS 119B
	CB.ALA 109A		NE2.HIS 119B
N6	OD1.ASN 71A	N7	ND2.ASN 67B
	SG.CYS 65A		CE1.HIS 119B
	OE1.GLN 69A		ND1.HIS 119B
	CB.ALA 109A	C5'	ND2.ASN 67B
N1	ND2.ASN 71A	C6	ND2.ASN 67B
	ND2.ASN 71A		OE1.GLN 69B
C2	CB.ALA 109A		CB.ALA 109B
C4	OE1.GLU 111A	N6	SG.CYS 65B
	CB.HIS 119A		ND2.ASN 67B
	CG.HIS 119A		OE1.GLN 69B
			OD1.ASN 71B
			CB.ALA 109B
		N1	ND2.ASN 71B
			OE1.GLN 69B
			CB.ALA 109B
		C2	CB.ALA 109B
			CG2.VAL 118B
		C4	CB.HIS 119B

					CG.HIS 119B
	C5P *Cytidine-5'-monophosphate*	1rnn	C3'	CE1.HIS 119E	
			O3'	CE1.HIS 119E	
				ND1.HIS 119E	
			C2'	CE1.HIS 12E	
			O2'	CE1.HIS 12E	
				NE2.HIS 12E	
				NZ.LYS 41E	
			C2	OG1.THR 45E	
				CD1.PHE 120E	
			N3	OG1.THR 45E	
				CD1.PHE 120E	
			C4	OG1.THR 45E	
			O2	CA.ASN 44E	
			N4	OG1.THR 45E	
	CPA *2'-deoxycytidine-2'-deoxyadenosine-3',5'-monophosphate*	1rpg	P	NE2.HIS 12A	
			O1P	NE2.HIS 12A	
				CD2.HIS 12A	
			O2P	NE2.GLN 11A	
			O5D	CE1.HIS 119A	
			C1X	NZ B.LYS 41A	
			CC2	OG1.THR 45A	
				CD1.PHE 120A	
			N3C	OG1.THR 45A	
				CB.THR 45A	
				CD1.PHE 120A	
				CE1.PHE 120A	
			CC4	CG1.VAL 43A	
				OG1.THR 45A	
			CC5	CG1.VAL 43A	
			O2C	CA.ASN 44A	
			O5B	ND1.HIS 119A	
				CE1.HIS 119A	
				CG.HIS 119A	
			C5B	ND1.HIS 119A	
			O4B	CB.HIS 119A	
				CG.HIS 119A	
				ND1.HIS 119A	
			C8A	CE1.HIS 119A	
				CG.HIS 119A	
				ND1.HIS 119A	
			N7A	CD2.HIS 119A	
			C5A	CD B.GLN 69A	
				NE2 B.GLN 69A	
				OE1 B.GLN 69A	
			C6A	CD B.GLN 69A	
				NE2 B.GLN 69A	
				OE1 B.GLN 69A	
				OD1.ASN 71A	
				CB.ALA 109A	
			N6A	SG.CYS 65A	
				CB.GLN 69A	
				CD B.GLN 69A	
				CG B.GLN 69A	
				CG A.GLN 69A	
				NE2 B.GLN 69A	
				CG.ASN 71A	
				OD1.ASN 71A	
				CB.ALA 109A	
			N1A	ND2.ASN 71A	

Compound	Code	Atom	Residue	Atom	Residue
			OE1 B.GLN 69A		
			CB.ALA 109A		
			OE2.GLU 111A		
		C2A	CD.GLU 111A		
			CG2.VAL 118A		
		N3A	OE2.GLU 111A		
		C4A	CB.HIS 119A		
CGP *2'-deoxycytidine-2'-deoxyguanosine-3',5'-monophosphate*	1rca	O5D	NH1.ARG 85A		
		C5D	NH1.ARG 85A		
		C2D	CA.ALA 122A		
		C8G	CG1.VAL 43A		
			OG1.THR 45A		
		N7G	OG1.THR 45A		
		C5G	CD1.PHE 120A		
		C6G	CE1.HIS 12A		
			CD1.PHE 120A		
		O6G	ND1.HIS 12A		
			CA.ASN 44A		
		P	OE1.GLN 11A		
			NE2.HIS 12A		
			NZ.LYS 41A		
		O1P	OE1.GLN 11A		
			CE1.HIS 12A		
		O3B	NE2.HIS 12A		
			NZ.LYS 41A		
		O5D	ND1.HIS 119A		
		C5D	CE1.HIS 119A		
			CG.HIS 119A		
			CE1.HIS 12A		
			CA.HIS 119A		
			CB.HIS 119A		
			CG.HIS 119A		
			ND1.HIS 119A		
			CA.PHE 120A		
			CB.PHE 120A		
C3P *Cytidine-3'-monophosphate*	1rpf	P	ND1.HIS 119A		
		O1P	CA.HIS 119A		
			CB.HIS 119A		
			CG.HIS 119A		
			ND1.HIS 119A		
		O2P	NE2.GLN 11A		
			CD2.HIS 12A		
			CE1.HIS 12A		
			NE2.HIS 12A		
		O3P	ND1.HIS 119A		
		C2'	CE1.HIS 12A		
			NZ.LYS 41A		
		O2'	NZ.LYS 41A		
			NE2.HIS 12A		
		C2	CA.ASN 44A		
			CD1.PHE 120A		
		N3	OG1.THR 45A		
			CB.THR 45A		
			CD1.PHE 120A		
		C4	CD1.PHE 120A		
		O2	CA.ASN 44A		
		P A	NE2.GLN 11A	P	NE2.HIS 12B
			NE2.HIS 12A		CD2.HIS 119B
			ND1.HIS 119A	O1P	CD2.HIS 12B

XXVII

Ligand	PDB	Atom	Contacts	Atom	Contacts
AMP *Adenosine monophosphate*	1z6s	**O1P A**	CD2.HIS 12A		NE2.HIS 12B
			NE2.HIS 12A		CA.HIS 119B
			CA.HIS 119A		CD2.HIS 119B
		O2P A	NE2.GLN 11A	**O2P**	NE2.GLN 11B
			NZ.LYS 41A		NE2.HIS 12B
		O3P A	CE1.HIS 119A	**O3P**	CD2.HIS 119B
			ND1.HIS 119A		NE2.HIS 119B
		O5' A	ND1.HIS 119A	**C5'**	CD2.HIS 119B
			CE1.HIS 119A	**C4'**	CD2.HIS 119B
			CG.HIS 119A	**O4'**	CB.HIS 119B
		O5' B	CE1.HIS 119A		CD2.HIS 119B
			ND1.HIS 119A		CG.HIS 119B
		C5' A	CB.HIS 119A	**C3'**	CD2.HIS 119B
			ND1.HIS 119A	**C2'**	CD2.HIS 119B
		C5' B	ND1.HIS 119A	**C1'**	CD2.HIS 119B
		O4' A	CB.HIS 119A		CG.HIS 119B
			CG.HIS 119A	**N9**	CD2.HIS 119B
			ND1.HIS 119A		CG.HIS 119B
		O4' B	CB.HIS 119A	**C8**	CD2.HIS 119B
			CE1.HIS 119A		CE1.HIS 119B
			CG.HIS 119A		CG.HIS 119B
		C2' A	CE1.HIS 119A		ND1.HIS 119B
			ND1.HIS 119A		NE2.HIS 119B
		C1' A	ND1.HIS 119A	**N7**	CG.ASN 67B
		C1' B	ND1.HIS 119A		ND2.ASN 67B
		N9	CG.HIS 119A		CE1.HIS 119B
		C8	CD2.HIS 119A		ND1.HIS 119B
			CE1.HIS 119A	**C5**	ND2.ASN 67B
			CG.HIS 119A	**C6**	SG.CYS 65B
			ND1.HIS 119A		CD.GLN 69B
			NE2.HIS 119A		OE1.GLN 69B
		N7	CG.ASN 67A		CB.ALA 109B
			ND2.ASN 67A	**N6**	OD1.ASN 71B
			CD2.HIS 119A		CB.CYS 65B
		C5	CD2.HIS 119A		SG.CYS 65B
		C6	OE1.GLN 69A		CB.GLN 69B
			OD1.ASN 71A		CD.GLN 69B
			CB.ALA 109A		OE1.GLN 69B
		N6	OD1.ASN 71A		CB.ALA 109B
			SG.CYS 65A	**N1**	ND2.ASN 71B
			CB.GLN 69A		OE1.GLN 69B
			CD.GLN 69A		CB.ALA 109B
			OE1.GLN 69A	**C2**	OE1.GLU 111B
			CG.ASN 71A		CG2.VAL 118B
			CB.ALA 109A	**C4**	CB.HIS 119B
		N1	ND2.ASN 71A		CG.HIS 119B
			CB.ALA 109A		
		C2	OE1.GLU 111A		
			CG2.VAL 118A		
		C4	CB.HIS 119A		
U5P *Uridine-5'-phosphate*	3jw1	**C2**	OG1.THR 45A	**C2**	OG1.THR 45B
			CD1.PHE 120A		CD1.PHE 120B
		N3	OG1.THR 45A	**N3**	OG1.THR 45B
			CB.THR 45A		CB.THR 45B
			CD1.PHE 120A		CD1.PHE 120B
			CE1.PHE 120A		CE1.PHE 120B
		C4	OG1.THR 45A	**C4**	OG1.THR 45B
		C6	CG1.VAL 43A	**O2**	CE1.HIS 12B
		O2	CE1.HIS 12A		CA.ASN 44B
			CA.ASN 44A		OG1.THR 45B
			CB.THR 45A	**O4**	OG1.THR 45B
			OG1.THR 45A	**C1'**	CE.LYS 41B
		C1'	CE.LYS 41A		NZ.LYS 41B
		O2'	NZ.LYS 41A	**O2'**	NZ.LYS 41B
			CE1.HIS 12A		CE1.HIS 12B
			CE.LYS 41A	**O3'**	CE1.HIS 119B
		O3'	ND1B.HIS 119A		ND1.HIS 119B
				O4'	CE.LYS 41B

						O1P	NZ.LYS 66B
						O1P	NZ.LYS 66B
							CD.LYS 66B
							CE.LYS 66B
						O2P	CE1.HIS 119B
EDN	A3P *Adenosine-3'-5'-diphosphate*	1hi4	P1	CH2.TRP 7A
				CZ3.TRP 7A
			O1P	CZ3.TRP 7A
			P2	NE2.GLN 14A
				ND1.HIS 129A
			O4P	CD2.HIS 15A
				NE2.HIS 15A
				CA.HIS 129A
			O5P	NE2.GLN 14A
			O6P	ND1.HIS 129A
				CE1.HIS 129A
				CG.HIS 129A
			C5'	CA.HIS 129A
				CB.HIS 129A
				CG.HIS 129A
				ND1.HIS 129A
			C4'	CZ3.TRP 7A
			O4'	CB.HIS 129A
			C2'	CE1.HIS 129A
				ND1.HIS 129A
			N9	CG.HIS 129A
				ND1.HIS 129A
			C8	CD2.HIS 129A
				CE1.HIS 129A
				CG.HIS 129A
				ND1.HIS 129A
				NE2.HIS 129A
			N7	CD2.HIS 129A
				NE2.HIS 129A
			C5	CZ.ARG 68A
				NH2.ARG 68A
				CD2.HIS 129A
			C6	CZ.ARG 68A
				NE.ARG 68A
				NH2.ARG 68A
				ND2.ASN 70A
				OD1.ASN 70A
				CB.ALA 110A
			N6	OD1.ASN 70A
				SG.CYS 62A
				CB.ARG 68A
				CG.ARG 68A
				CZ.ARG 68A
				NE.ARG 68A
				CG.ASN 70A
			N1	ND2.ASN 70A
				NH2.ARG 68A
				CG.ASN 70A
				OD1.ASN 70A
				CB.ALA 110A
			C2	NH2.ARG 68A
				ND2.ASN 70A
				CG2.VAL 128A
			C4	NH2.ARG 68A

Table S3: List of selected structure complexes of nucleotide-type ligands with representative RNase A superfamily members. Residues involved in potential hydrogen –bonds and van der Waals interactions for each nucleotide building block are included, as calculated using the *PDBe motif* server. PB indicates peptide backbone interactions and PCA refers to pyroglutamic acid.

Protein Name	Ligand Information	PDB Code	Interaction block			
			Pyr base	Pur base	Ribose	Phosphate
Onconase (*Rana pipiens*)	**DU DG DA** d(AUGA)	2i5s	T35 F98	PCA1 (PB) E91 T89 H97		K9 H10 F98 (PB) K31
RC- RNase 6 (*Rana catesbeiana*, Bullfrog)	**CG2** *Cytidil-2',5'-phosphoguanosine*	1oj1		H10 N34 (PB) T35 F98	N69 A99 (PB) G100 (PB) V101 (PB)	Q67
Bovine Seminal RNase (*Bos Taurus*)	**CPA** *2'-deoxycytidine-2'-deoxyadenosine-3',5'-monophos-phate*	1r5c	T45 V43 N44 (PB) F120	C65 N67 N71 A109 V118 H119	H12 K41	Q11 H12 H119 F120 (PB)
	U3P *Uridine-3'-monophosphate*	1n3z	N44 (PB) T45 F120 D121 (PB)		H12 K41	Q11 K41 H119 F120 (PB)
	UPA *Uridylyl-2',5'-phosphoadenosine*	11ba	H12 V43 (PB) N44 (PB) T45 F120	C65 N67 Q69 N71 A109 V118 H119	H119	H12 Q11 K41 H119 F120 (PB)
	CPA *2'-deoxycytidine-2'-deoxyadenosine-3',5'-monophos-phate*	1tq9	H12 N44 (PB) T45 F120	C65 Q69 N71 A109 V118 H119	H12 K41 V43 (PB) F120 (PB)	K7 Q11 H12 K41 F120 (PB)
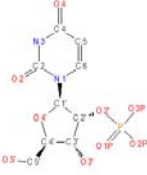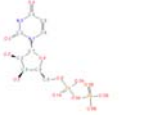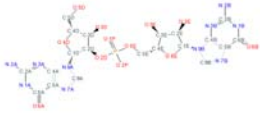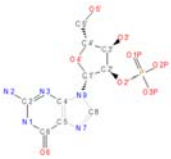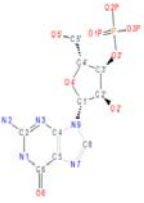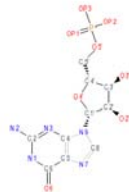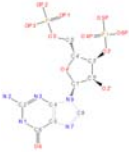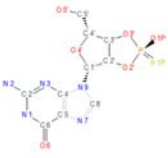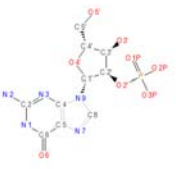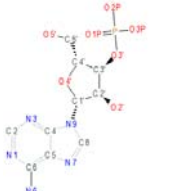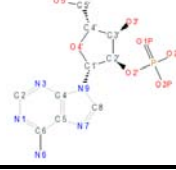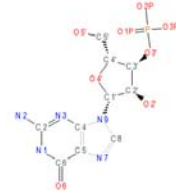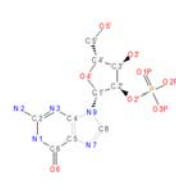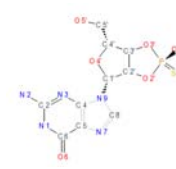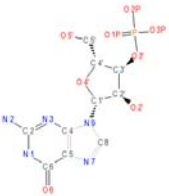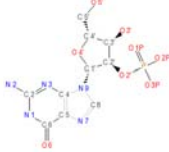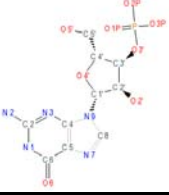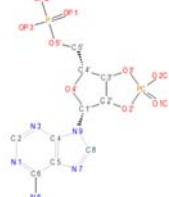
U2P *Uridine 2'-phosphate*		3djo	T45 N44 (PB) F120		K41 V43 (PB) K66 H119 F120 (PB) D121 (PB)	Q11 H12 K41 K66 R85 H119
UDP *Uridine 5'-diphosphate*		3djq	V43 N44 (PB) T45 F120		H12 K41 V43 (PB) H119	

Table S4. List of selected structure complexes of nucleotide-type ligands with representative members of the microbial RNases superfamily. Residues involved in potential hydrogen –bonds and van der Waals interactions for each nucleotide building block are included, as calculated using the *PDBe motif* server. PB indicates peptide backbone interactions.

Protein Name	Ligand Information	PDB Code	Interaction block		
			Pur base	*Ribose*	*Phosphate*
RNase T1 *(Aspergillus oryzae)*	**GPG** *Guanylyl-2,5'phosphoguanosine*	2rnt	K41 (PB) Y42 N43 Y45 E46 H92 N98 (PB) F100	H40 E58 R77 H92 N98	N36 Y38 E58 R77 H92 F100
	2GP *Guanosine 2'-monophosphate*	1i0v	H40 K41 (PB) Y42 N43 N44 (PB) Y45 E46 E58 F100	H40 N98	Y38 H40 E58 R77 H92 F100
		1i0x	K41 (PB) Y42 N43 N44 (PB) Y45 E46 P73 G74 R77 N98 (PB) F100	H40 N98	N36 Y38 H40 E58 R77 H92 N98 F100
		1bvi	K41 (PB) Y42 N43 (PB) N44 (PB) Y45 E46 S72 P73 N98 N99 F100	N36 (PB) H40 N98	Y38 H40 Y45 E58 R77 H92 N98 F100
		2bu4	K41 (PB) Y42 N43 N44 (PB) Y45 E46 E58 N98 (PB) N99 F100	K40 E58	Y38 K40 E58 R77 H92 F100

XXXII

Compound	Structure	PDB			
		1rnt	H40 K41 (PB) Y42 N43 N44 (PB) Y45 E46 N98 (PB) F100	N98	Y38 H40 E58 H92 F100
3GP *Guanosine 3'-monophosphate*		1rgc	K41 (PB) Y42 N43 (PB) N44 Y45 E46 N98 (PB) N99 (PB) F100	H40 E58 N98	R38 E58 R77 H92 F100
		1rls	K41 (PB) Y42 N43 N44 Y45 E46 N98 (PB) F100	H40 E58 N98 (PB)	Y38 E58 R77 H92 F100
		1rga	K41 (PB) Y42 N43 N44 Y45 E46 N98 (PB) F100	N36 H40 E58 N98	N36 Y38 R77 H92
5GP *Guanosine 5'-monophos-phate*		1rga	A75 H92 N98	S72 P73 R77 H92	
PGP *Guanosine-3',5'-diphosphate*		5rnt	K41 (PB) Y42 N43 N44 (PB) Y45 N98 (PB) F100	H40 E58	T45
SGP *Guanosine-2',3'-cyclophospho-rothioate*		1gsp	K41 (PB) Y42 N43 N44 (PB) Y45 E46 N98 (PB) F100	H40 E58 N98	Y38 E58 R77 H92 F100

Enzyme	Ligand	PDB			
	2AM *2'-Adenosine monophosphate*	6rnt	A75 R77 H92	N36 H92	Y38 H40 E58 R77 H92 F100
RNase F1 *(Fusarium moniliforme)*	**2GP** *Guanosine 2'-monophosphate*	1fut	T41 (PB) Y42 N43 (PB) N44 (PB) Y45 E46 N98 (PB) F100	H40 E58 N98 (PB)	Y38 H40 E58
RNase U2 *(Ustilago sphaerogena)*	**3AM** *3'-Adenosine monophosphate*	3agn	Y43 Y44 E46 E49 F110	H41 E62 D108	Y39 H41 E62 R85 H101 F110
		3ago	E46 E49 F110	H41 E62 D108	Y39 H41 E62 R85 H101
	2AM *2'-adenosine monophosphate*	3ahw	P81 P83 (PB) H101	H101	Y39 H41 E62 R85 H101 F110
RNase Sa *(Streptomyces aureofaciens)*	**3GP** *Guanosine 3'-monophosphate*	2sar	F37 Q38 N39 (PB) R40 E41 Y86	N32 E54 Y86	N32 R65 R69 H85 Y86
	2GP *Guanosine 2'-monophosphate*	1rge	F37 Q38 R40 (PB) E41 Y86	H85	E54 R65 R69 Y86
		1gmr	F37 Q38 N39 (PB) R40 E41 E54 Y86	R40 R85 Y86	E54 R65 R69 H85 Y86
	SGP *Guanosine 2',3'-cyclophospho-rotioate*	1rsn	F37 Q38 N39 (PB) R40 E41 Y86	V35 E54 R65 Y86	R65 R69 H85 Y86

Binase (Bacillus intermedius)	**3GP** Guanosine 3'-monophos-phate		1goy	F55 S56 N57 (PB) R58 E59	K26 V54 (PB) E72 Y102	K26 R82 Y86 H101 Y102
RNase Sa2 (Streptomyces aureofaciens)	**3GP** Guanosine 3'-monophos-phate		3d4a	V38 (PB) F39 E40 N41 (PB) R42 E43 Y87	V37 E56 Y87	R67 R71 H86
			3dh2	F39 E40 N41 (PB) R42 E43 Y87	V37 E56 Y87	R67 R71 H86 Y87
	2GP Guanosine 2'-monophos-phate		3dgy	V38 (PB) F39 E40 N41 (PB) R42 E43 Y87	R34 R42 E56 H86	R34 E56 R67 R71 H86
RNase Ms (Aspergillus saitoi)	**3GP** Guanosine 3'-monophos-phate		1rms	Y41 H42 N43 (PB) Y44 E45 D97 (PB) F99	H39 D97 F99	Y37 E57 R76 H91
Restrictocin (Aspergillus restrictus)	**A23** Adenosine -5'-phosphate-2',3'-cyclic phosphate		1jbr	W50 (PB) F51 (PB) T52 N53 (PB) R65 L144	E95	Y47 H49
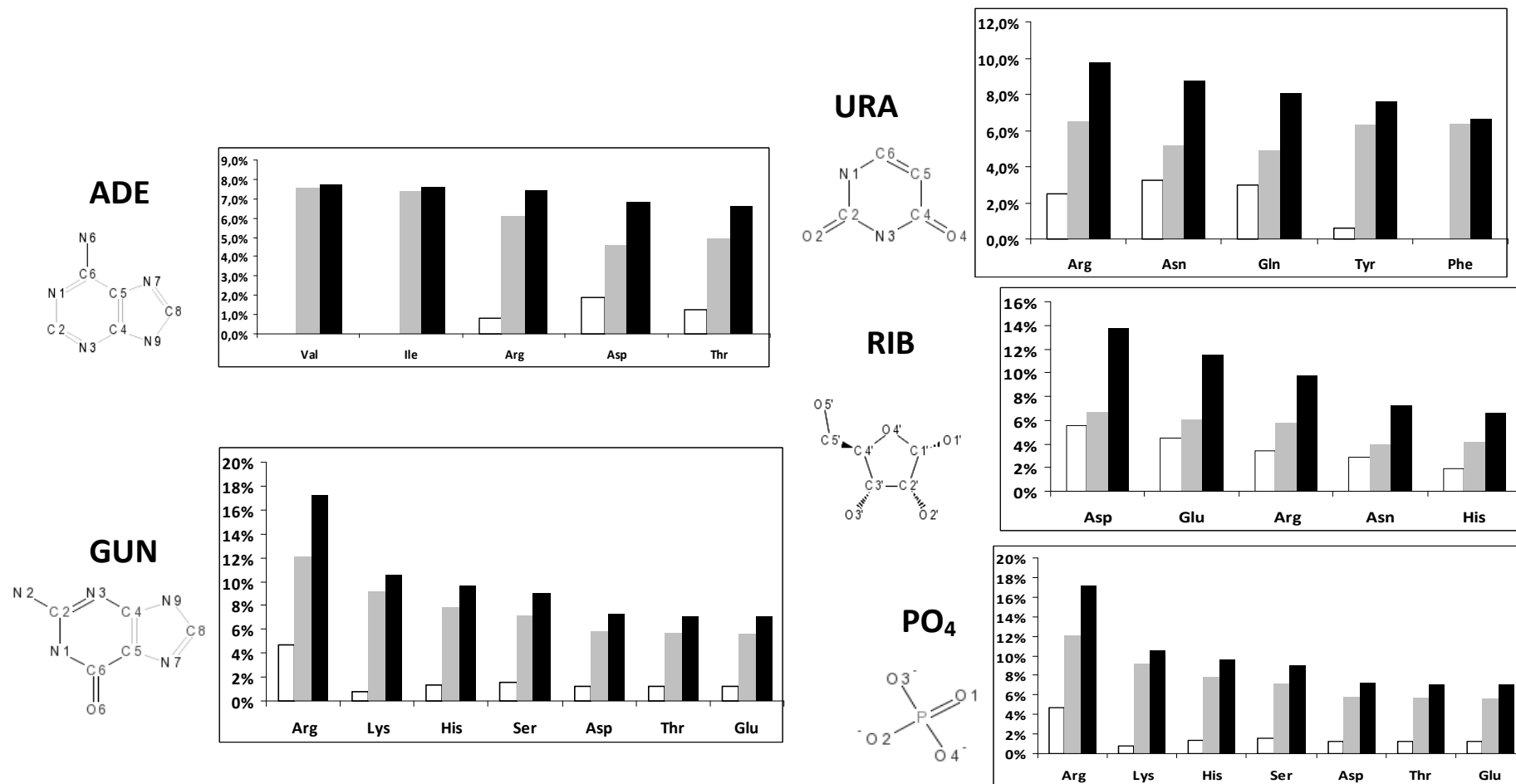
Figure S1: Graphical representation of interactions between aminoacids and nitrogenous bases, ribose and phosphate group. Bars represent hydrogen bond (white) interactions, van der Waals forces (grey) and the sum of both (black). Only aminoacids with significance over 80% and have been considered. The amount of protein structures and complexes where these interactions have been found are, respectively, 2735 and 14483 for Adenine (ADE), 18 and 161 for Guanine (GUN), 450 and 2526 for Uracil (URA), 254 and 1748 for Ribose (RIB) and 1393 and 7285 for Phosphate (PO4). Due to the poor amount of data for Cytosine (CYT), no further studies have been considered for this base.
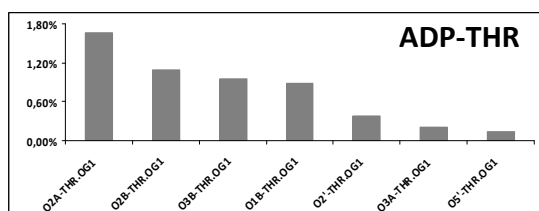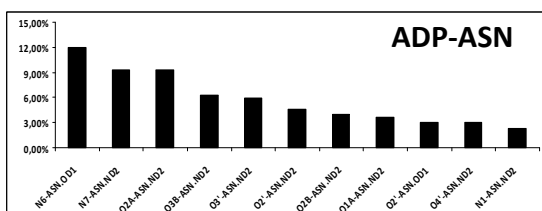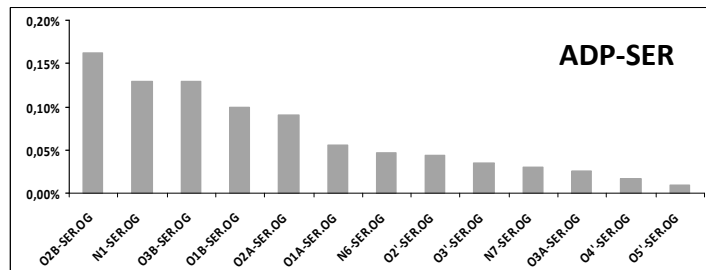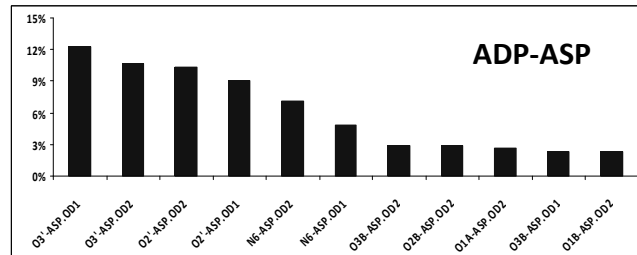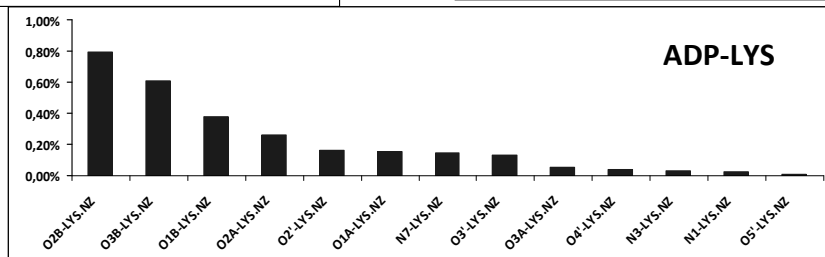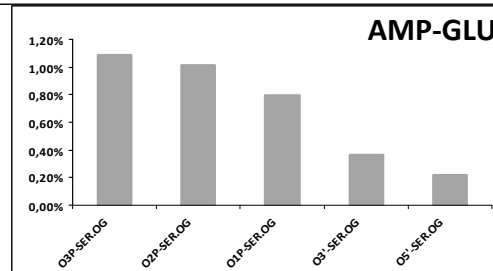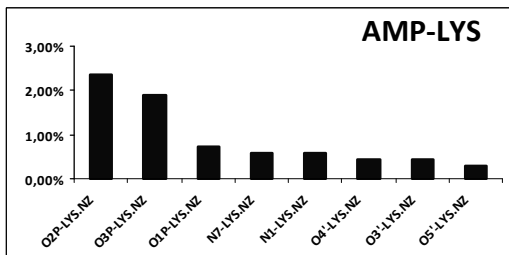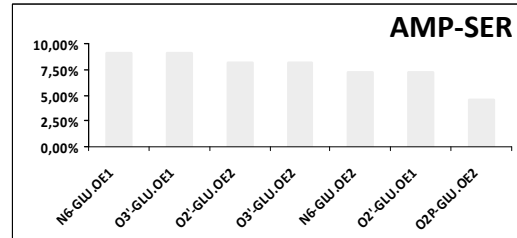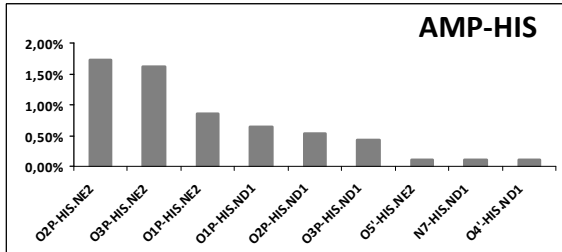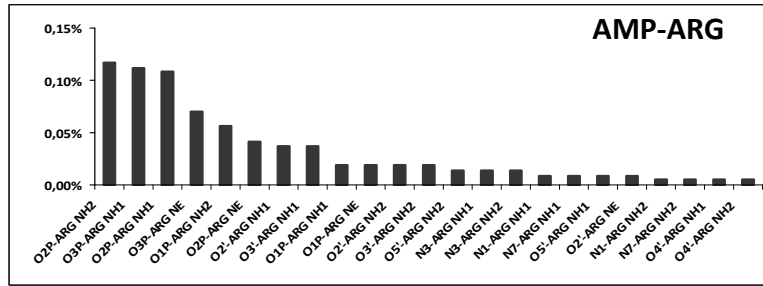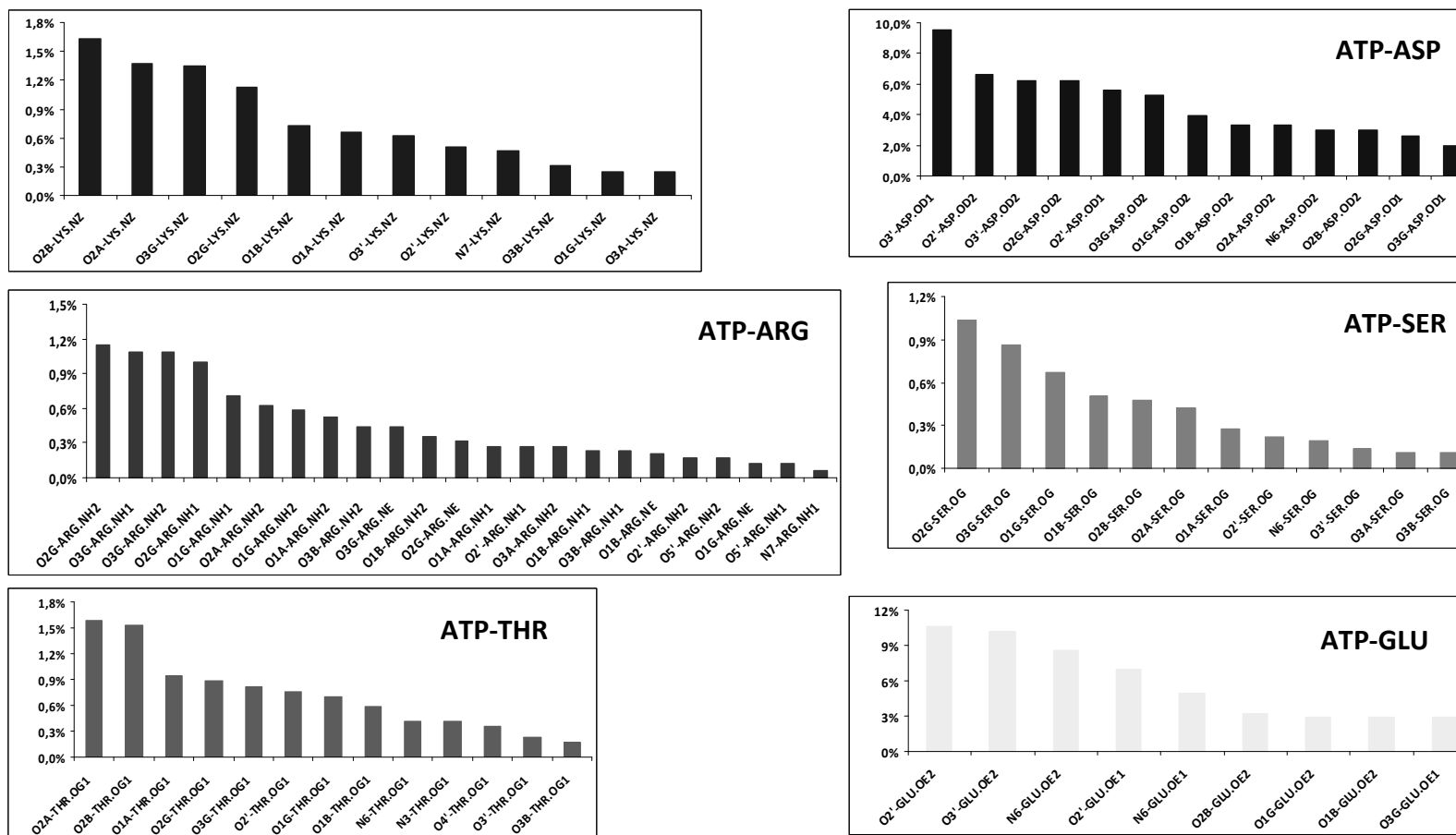
Figure S2: Graphic representation of hydrogen bond interactions for selected nucleotides with the best amino acids binders. Nucleotide-type ligands have been selected according to statistical significance criteria. Structures and complexes where these interactions were found are: 1386/22325 for AMP-Arg; 105/682 for AMP-Lys; 85/932 for AMP-His; 69/111 for AMP-Glu; 151/1385 for AMP-Ser; 1194/14064 for ADP-Lys; 378/3856 for ADP-Thr; 1668/23265 for ADP-Ser; 155/302 for ADP-Asn; 170/310 for ADP-Asp; 264/3194 for ATP-Lys; 255/3588 for ATP-Ser; 232/3401 for ATP-Arg; 197/1707 for ATP-Thr; 143/304 for ATP-Asp and 127/246 for ATP-Glu.

XL

XLI
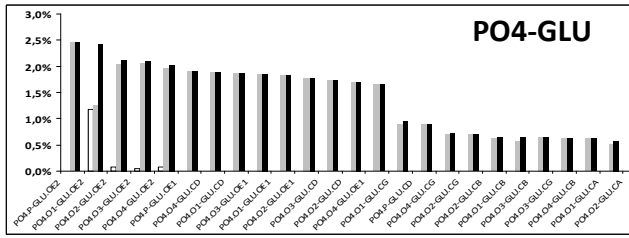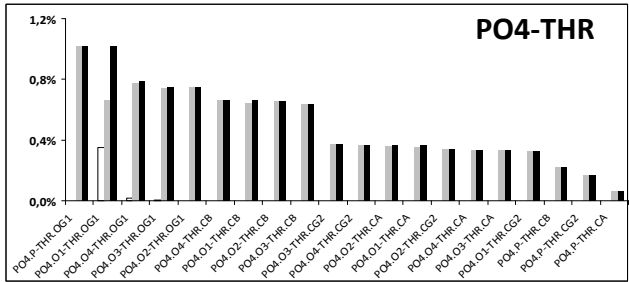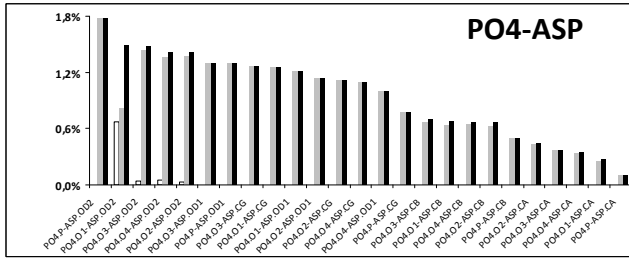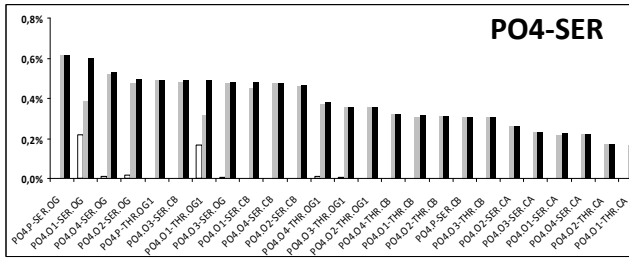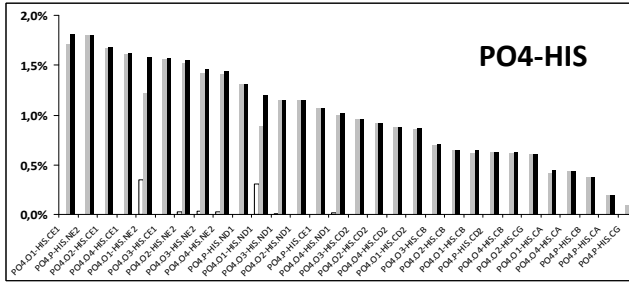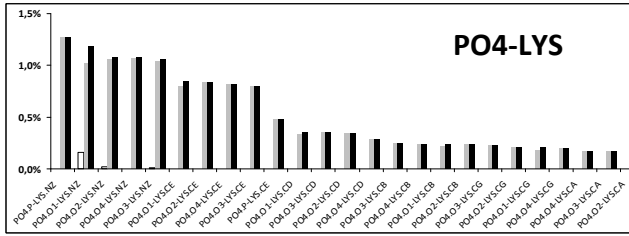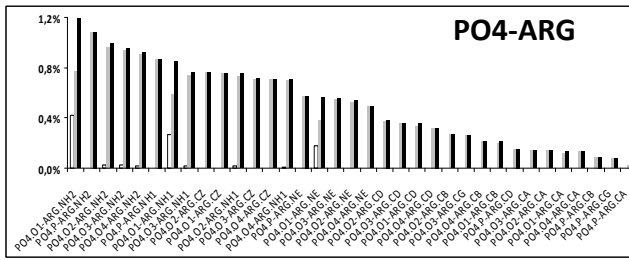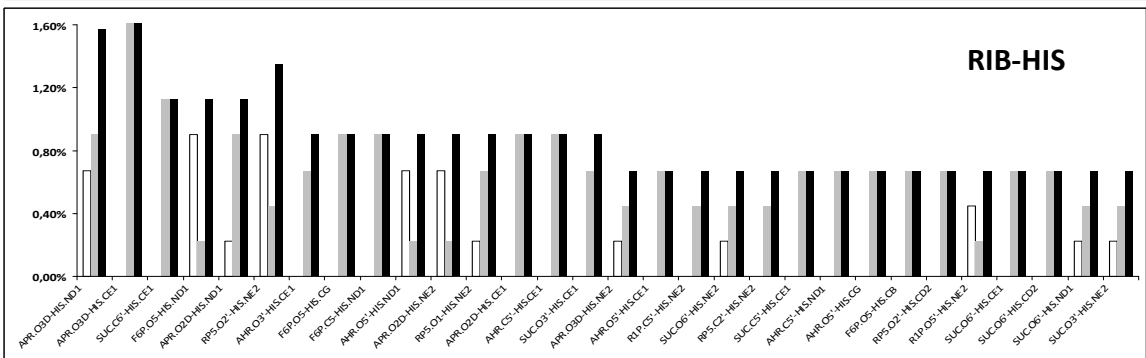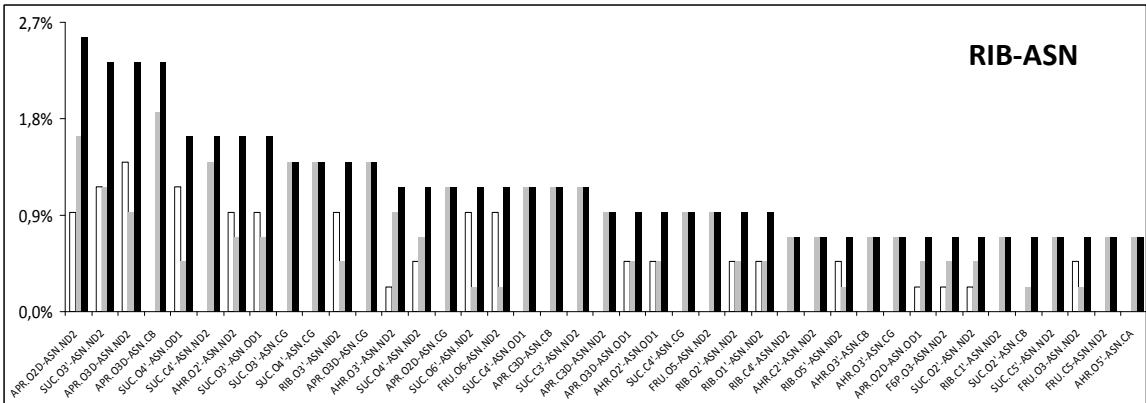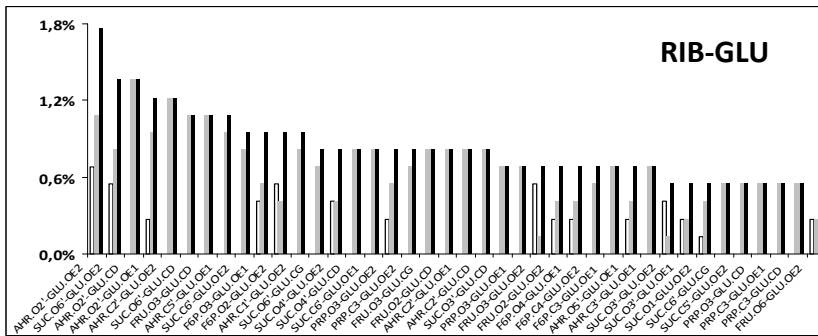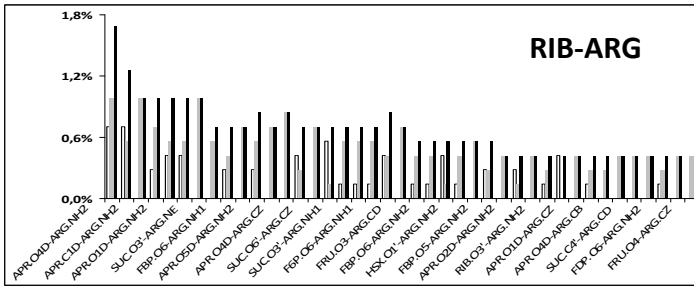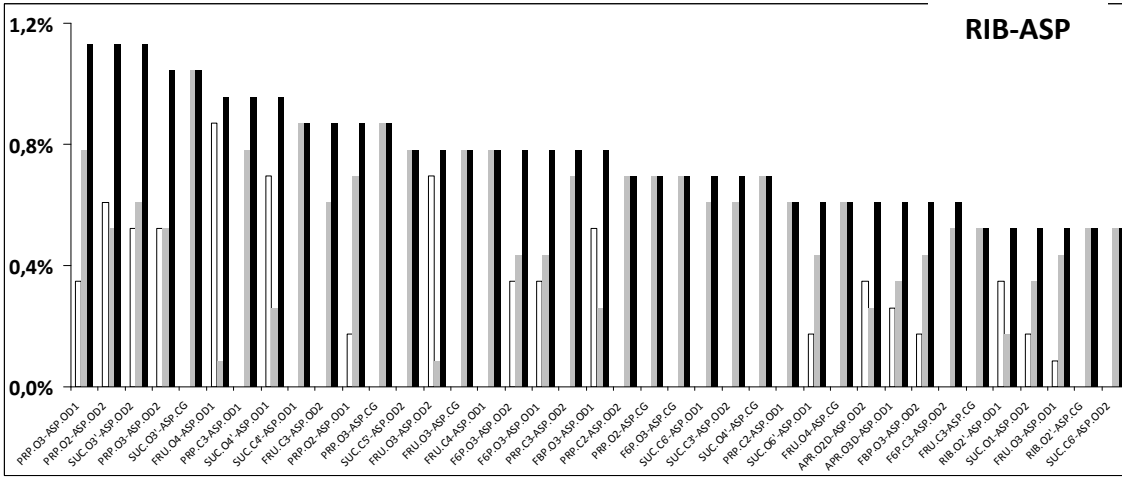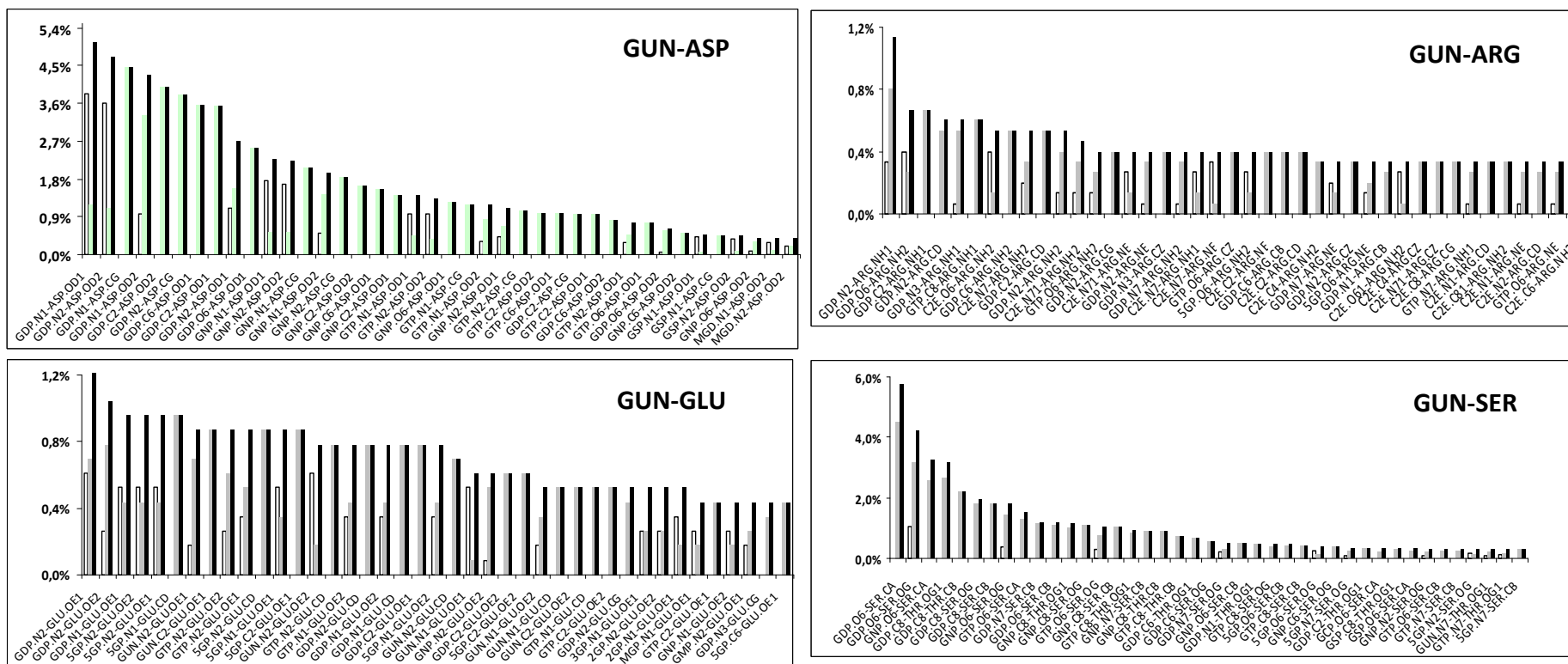
Figure S3: Graphical representation of interactions between amino acids and phosphate (PO4), ribose (RIB) and nitrogenous bases (ADE: adenine, GUN: guanine, URA: uracil). Bars represent hydrogen bond (white) interactions, *van der Waals* forces (grey) and the sum of both (black). Only the most significant amino acids (values over 80%), with highest total interaction percentage and interacting to all structures have been considered. For statistical data comprising many interaction values, a cut-off limit of only the 40 highest total significance value has been taken. Structures and complexes for which these interactions (PO4, RIB, ADE, GUN and URA) can be found are, respectively, 2942/45850 for PO4-ARG, 2598/29728 for PO4-LYS, 1497/14185 for PO4-HIS, 3339/51107 for PO4-SER, 1370/11676 for PO4-ASP, 2184/24546 for PO4-THR, 984/7141 for PO4-GLU; 129/1151 for RIB-ASP, 115/736 for RIB-GLU, 115/734 for RIB-ARG, 77/430 for RIB-ASN, 78/445 for RIB-HIS; 981/11194 for ADE-ARG, 923/5135 for ADE-ASP, 848/5090 for ADE-THR; 112/1149 for GUN-GLU, 346/3989 for GUN-ASP, 137/1495 for GUN-ARG, 359/2393 for GUN-SER; 174/1677 for URA-ARG, 157/1416 for URA-ASN and 139/1321 for URA-GLN.

Title: Crystal structure of human RNase 6 in complex with sulphate anions

Authors: Blanco, José Antonio[1]; Arranz, Javier[1]; Pulido, David[1,2]; Moussaoui, Mohammed[1]; Boix, Ester[1, *]

[1] Department of Biochemistry and Molecular Biology, Universitat Autònoma de Barcelona, E-08193 Cerdanyola del Vallès, Spain

[2] Present address: Imperial College London, South Kensington Campus London, London SW7 2AZ, United Kingdom

* Author to whom correspondence should be addressed; E-Mail ester.boix@uab.cat; Tel.: +34-93-581-4147; Fax: +34-93-581-1264

Abstract

Human ribonuclease 6 (RNase k6) is a basic secretory protein that belongs to the ribonuclease A (EC 3.1.27.5) superfamily. Its expression and induction in neutrophils and monocytes has suggested a role in host defence. A crystal structure of this protein has been obtained at a resolution of 1.72 Å, being the first report for the protein 3D structure and thereby setting the basis for structural and functional studies. The structure was solved by molecular replacement using the NMR structure of human RNase 7, its closest homologue, sharing a 55% of sequence identity. Comparative analyses with nucleotide analogue complexes of other members of the RNase A superfamily have been applied to understand the RNase 6 substrate subsite base selectivity. The presence of sulphate anions enables additional understanding of the protein ligand recognition sites and binding interaction mode, with potential applications in the design of nucleotide-type inhibitors or other types of drugs.
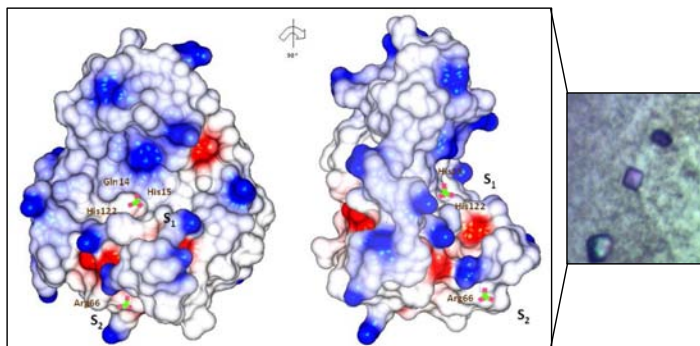
Keywords: Crystal structure, human RNase 6, RNase k6, RNase A superfamily, sulphate anion

Chemical compounds studied in this article: Sulphate (PubChem CID: 1117)

Abbreviations: XRD: X-ray diffraction

Specifications Table

Subject area	*Biochemistry*
More specific subject area	*Protein Crystallography*
Type of data	*X-ray diffraction data, coordinates file*
How data was acquired	*X-ray diffraction. Protein crystal was diffracted at 100K using a λ_{XRD} = 0.9795 Å and a Pilatus 6M detector (Dectris®, Switzerland). 800 images were taken at t_{exp} = 0.2 s, $\Delta\varphi$ = 0.2°.*
Data format	*Structure factor amplitudes file (MTZ format) and atomic coordinate file (PDB format)*
Experimental factors	*Recombinant human RNase 6 was resuspended at 10 mg/mL in 20 mM sodium cacodylate pH 5.0. Crystals were grown by vapour diffusion using a crystallization buffer of 2.0 M $(NH_4)_2SO_4$, 0.1 M sodium cacodylate pH 6.5 and 50 mM NaCl. One microlitre of the sample was mixed with an equal volume of the reservoir solution. Cubic shaped crystals appeared after 10 days of incubation at 20 °C and were soaked in the cryoprotectant solution prior to X-ray exposure.*
Experimental features	*The crystal diffracted at 1.72 Å. The RNase 6 3D structure was solved using RNase 7 as a model.*
Data source location	*Beamline BL13 (XALOC), ALBA Synchrotron Light Facility, Cerdanyola del Vallès, Spain*
Data accessibility	*Structure coordinate file has been deposited at the Protein Data Bank (PDB code 4X09).*

Value of the data

- First reported 3D structure of human RNase 6.
- The polypeptide chain interacts with two sulphate anions, which facilitates further structural studies of the protein substrate binding mode or other potential ligand recognition sites.
- Analysis of RNase 6 3D structure would assist the understanding of the protein antimicrobial mechanism of action.

Experimental Design, Materials and Methods

A plasmid containing the gene of recombinant human RNase 6 was transformed in a prokaryote expression system. The cDNA encoding RNase6 sequence was a kind gift provided by Dr. Helene Rosenberg (National Institute of Health, Bethesda, MD, USA). The gene was subcloned in plasmid pET11c for prokaryote high yield expression. *E. coli* cells –BL21(DE3) strain– were transformed with the pET11c/RNase 6 plasmis. The expressed protein was purified by a modification of the described expression procedure (Boix et al., 1999), using 100mM TRIS-HCl buffer, pH 7.0-7.5, for refolding and cationic exchange purification. An additional reverse phase chromatography was applied and the purified

protein was checked by MALDI-TOF spectrometry. (Boix et al., 2012a). A sample was lyophilised and resuspended at 10 mg/mL in 20 mM sodium cacodylate pH 5.0 and equilibrated against 2.0 M $(NH_4)_2SO_4$, 0.1 M sodium cacodylate pH 6.5 and 50 mM NaCl. One microlitre of the sample was mixed with an equal volume of the reservoir solution and set to incubation at 20 °C. After 5 to 10 days, cubic shaped crystals appeared and were soaked using 15% glycerol as cryofreezing agent. Data were captured at 100K using a λ_{XRD} = 0.9795 Å at the BL13(XALOC) beamline of the ALBA Synchrotron Light Facility (Cerdanyola del Vallès, Spain). For data processing, the EDNA platform (Incardona et al., 2009) was used to predict crystal symmetry and unit cell parameters and XDS (Kabsch, 2010) was used for data indexing and scaling. The PHENIX software (Adams et al., 2010) was further utilised for model phasing by means of molecular replacement (McCoy et al., 2007) using the RNase 7 NMR structure (PDB coordinate file 2HKY (Huang et al., 2007)) as a model. Iterative cycles of refinement and manual structure fitting were performed with PHENIX (Adams et al., 2010) and COOT (Emsley and Cowtan, 2004) until R_{free} could not be further improved (Brunger, 1992). Finally, the stereochemistry of the structure was validated with SFCHECK (Vaguine et al., 1999) and WHATCHECK (Hooft et al., 1996). Table 1 shows all the data collection and structure refinement statistics.

Table 1: *Structure refinement parameters of the RNase 6 crystal.*

Data collection		Refinement	
Space group	P2₁2₁2₁	Resolution range (Å)	48.98 – 1.72
Unit cell		R_{cryst}[c] / R_{free}[d] (%)	20.85 / 25.54
a, b, c (Å)	27.73 38.86 97.97	Number of protein atoms	1049
α, β, γ (°)	90.0 90.0 90.0	Number of water molecules	164
No. of molecules in the asymmetric unit	1	Number of anions (sulphate)	2
Resolution (Å)	1.72	Rms deviation from ideal geometry	
No. of total reflections	22981	Bond lengths (Å)	0.009
No. of unique reflections	11717	Bond angles (deg)	1.345
R_{merge} [a, b] (%)	2.8 (23.4)	B factors of protein atoms (Å²)	
I/σ_I [b]	13.0 (2.4)	All	28.60
Completess for range [b] (%)	99.2 (99.0)	Main chain	25.90
Wilson B factor (Å²)	21.98	Side chain	31.13
Matthews coefficient (Å³/Da)	1.80	B factors of sulphate anions (Å²)	40.95
Solvent content (%)	31.71	B factors of water molecules (Å²)	40.43

[a] $R_{merge} = \Sigma_{hkl} \Sigma_{j-1}$ to $N|I_{hkl}-I_{hkl}(j)|/ \Sigma_{hkl} \Sigma_{j-1}$ to $N\ I_{hkl}(j)$, where N is the redundancy of the data.
[b] Outermost shell is 1.78-1.72 Å.
[c] $R_{crystal} = \Sigma_h|F_o-F_c|/\Sigma_hF_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h, respectively.
[d] R_{free} is equal to R_{cryst} for a randomly selected 5% subset of reflections not used in the refinement.

Analysis of the RNase 6 three-dimensional structure

Human RNase 6 is a small cationic protein that belongs to the pancreatic ribonuclease (RNase A) superfamily. It is expressed in neutrophils and monocytes (Rosenberg and Dyer, 1996). Recent findings describe suggest that RNase 6 antimicrobial activity protects the urinary tract from infection (Becknell et al., 2015). Here we present its first 3D structure, which may contribute to understand the biological properties of the protein

Overall features: The unit cell belongs to the space group P2₁2₁2₁, with one protein molecule in the asymmetric unit. The RNase 6 three-dimensional structure complies the RNase A superfamily overall conformation, with a kidney shaped structure formed by 7 β-strands and 3 α-helices as listed in Figure 1 and Table 2.
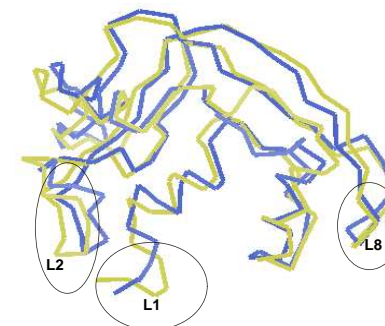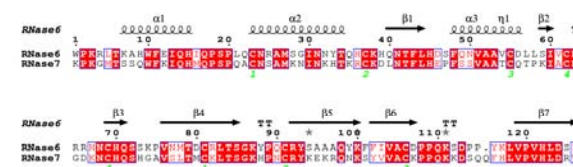
Figure 1: *Top, comparison of the blast alignment of RNase 6 and RNase 7 primary sequences. Secondary structure of RNase 6 is depicted. Strictly conserved residues are boxed in red and conserved residues, as calculated by a similarity score, are boxed in white. Cysteine pairings for disulphide bridges are numbered below. The figure was created using the ESPript software (Gouet et al., 1999). Bottom, overlay of the backbone structures of RNase 6 (yellow) and RNase 7 (blue, PDB ID: 2HKY (Huang et al., 2007)). Loops L1, L2 and L8 are indicated, too, showing the higher divergence between the two structures.*

Loop residues Trp1-Lys3, Gln17-Leu21, Lys63-Arg66, Gly86-Gln90 and Pro108-Ser112 were partially disordered. In particular, practically no electron density was visualised for residues Pro2, Lys3, Gln17, Leu21, Gly86 and Lys87, which could not be not properly modelled. Alternate side chain conformations were modelled for Arg4, Ser59, Arg92 and Lys111. Arg4, the RNase A counterpart of Glu2 in RNase A, cannot provide the corresponding Glu⋯Arg salt bridge observed in RNase A, which stabilises

the α1 helix (Chatani and Hayashi, 2001; Rico et al., 1984). The high motion values of some RNase 7 loop residues are also observed in the RNase 6 structure, particularly in loops L1, L2 and L8 (see Figure 1), where 4% of the residues are disordered and could not be modelled properly.

Table 2: *RNase 6 secondary structure elements.*

loop residues		helix residues		strand residues	
L1	W1 – T6	α1	K7 – H15	β1	Q40 – L44
L2	I16 – Q22	α2	C23 – T34	β2	S59 – I60
L3	Q35 – H39	α3	F48 – D56	β3	C69 – Q71
L4	H45 – S47			β4	V76 – S85
L5	L57 – L58			β5	R92 – K100
L6	V61 – N68			β6	F102 – D107
L7	S72 – P75			β7	V119 – I126
L8	G86 – C91				
L9	P108– L118				

Crystal packing: Residues involved in crystal packing were analysed by the PISA web server (Krissinel and Henrick, 2007). The intermolecular contacts are illustrated in Figure 2 and listed in Table 3. Interactions are found mostly between β3, β4, β5 and β7 strand residues (Gln71, Arg82, Ala97-Tyr99, Ser125, Ile126) and loop residues. No packing contacts are seen in the environment of the active site, therefore enabling further substrate analogue studies.

Active site: Active site architecture is conserved with respect to RNase A. Residues His15, Lys38 and His122 (His12, Lys41 and His119 RNase A counterparts) build the active site groove, with His122 adopting the so-called *inactive* orientation (Borkakoti, 1983), a conformation reported to be favoured in aqueous ionic salt solutions (Berisio et al., 1999). The imidazole ring of RNase 6 His122 shows a related rotation of $\chi_1 \sim 137°$ and $\chi_2 \sim 127°$. The other ring orientation (*active*) is found predominant in other ribonuclease crystal structures 1RPG (Zegers et al., 1994) (deMel et al., 1992). The potential *active* orientation of this histidine ring would be also influenced by the interaction with the vicinal Asp124 residue, whose interaction would account for the correct His122 tautomer in catalysis (Schultz et al., 1998). Interestingly, the *inactive* histidine orientation would be hindered upon presence of a purine base at the B_2 subsite (Zegers et al., 1994).
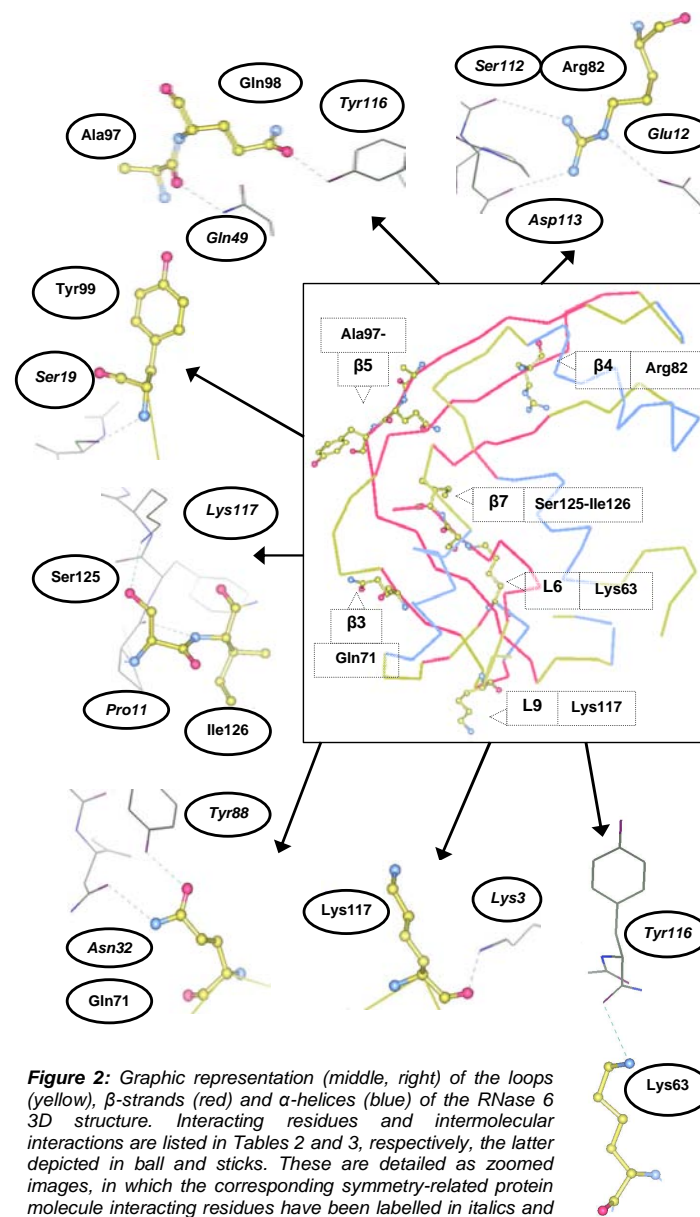
Figure 2: *Graphic representation (middle, right) of the loops (yellow), β-strands (red) and α-helices (blue) of the RNase 6 3D structure. Interacting residues and intermolecular interactions are listed in Tables 2 and 3, respectively, the latter depicted in ball and sticks. These are detailed as zoomed images, in which the corresponding symmetry-related protein molecule interacting residues have been labelled in italics and represented in dark. Interactions have been drawn in green.*

XLV

Table 3: *Intermolecular packing interactions between symmetry related molecules in the RNase 6 crystal. Only hydrogen bond interactions are included, taking a cut-off of 3.4 Å as a reference, as calculated by the PDBePISA server (Krissinel and Henrick, 2007).*

symmetry operation	crystal molecule interacting atom	symmetry related molecule interacting atom	distance (Å)
x-1, y, z	Arg82 Nε	Glu12 Oε1	3.32
	Arg82 Nη1	Ser112 O	3.30
	Arg82 Nη2	Asp113 Oδ1	3.12
	Ile126 N	Pro115 O	3.10
	Lys63 Nζ	Tyr116 O	3.27
	Ala97 O	Gln49 Nε2	2.90
	Gln98 Oε1	Tyr116 Oη	2.47
	Ser125 Oγ	Lys117 Nζ	3.03
x-½, -y-½, -z	Tyr99 N	Ser19 Oγ	2.73
x, y-1, z	Gln71 Nε2	Asn32 Oδ1	3.11
	Gln71 Oε1	Tyr88 Oη	2.86
-x+1, y-½, -z-½	Lys117 O	Lys3 Nζ	2.81

Table 4: *Atomic interactions between sulphate anions and RNase 6 residues. Potential hydrogen bond distances have been considered using a cut-off distance of 3.4 Å.*

sulphate interaction site and atom		interacting protein atom	distance (Å)
S1	O1	His122 Nδ1	2.73
	O3	His15 Nε2	2.98
		Leu123 N	2.73
	O4	Gln14 Oε1	3.30
		His15 Nε2	3.26
S2	O3	Arg66 Nε	2.89
		His67 N	3.16

Sulphate binding sites: Two sulphate ions were located in the solved protein structure (see Figure 3) Interactions with nearby residues are listed in Table 4 and illustrated in Figure 3. S1 corresponds to the active site of the enzyme and has also been identified in substrate analogue complexes of RNase A (Bell, 1999; Berisio et al., 2002; Fedorov et al., 1996; Mueller-Dieckmann et al., 2007; Zegers et al., 1994), EDN/RNase 2 (Leonidas et al., 2001; Mosimann et al., 1996), ECP/RNase 3 (Boix et al., 2012a) and angiogenin/RNase 5 (Holloway et al., 2005). On the other hand, the second sulphate binding site (S2) has not been reported in any other superfamily member so far. Interestingly, docking studies have predicted potential interactions of RNase 6 residue Arg66 with heparin sulphate and nucleotide phosphate moieties of ligand analogues (Boix et al., 2012b). Comparative studies with other family member complexes with substrate analogue or heparin derivatives (Fontecilla-Camps et al., 1994) (Garcia-Mayoral et al., 2013) suggest that this site may represent a distinct anion interaction subsite. Further work is envisaged to identify putative related RNase 6 biological properties.

Figure 3: *Representation of sulphate anions S1 and S2. Each anion has been depicted together with its own |2F$_o$-F$_c$| (blue) and |F$_o$-F$_c$| (magenta) electron density maps at 1.5σ and 5σ, respectively. Interactions with nearby residues (see Table 4) and water molecules (red crosses) have been represented in purple lines.*

Database linking

Crystal structure atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB), see www.pdb.org (PDB ID: 4X09).

Conflicts of interest

The authors declare no conflict of interest.

References

Adams, P.D., 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, 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66: 213-21.

Becknell, B., T.E. Eichler, S. Beceiro, B. Li, R.S. Easterling, A.R. Carpenter, C.L. James, K.M. McHugh, D.S. Hains, S. Partida-Sanchez, and J.D. Spencer, 2015. Ribonucleases 6 and 7 have antimicrobial function in the human and murine urinary tract. Kidney Int 87: 151-61.

Bell, J.A., 1999. X-ray crystal structures of a severely desiccated protein. Protein Sci 8: 2033-40.

Berisio, R., V.S. Lamzin, F. Sica, K.S. Wilson, A. Zagari, and L. Mazzarella, 1999. Protein titration in the crystal state. J Mol Biol 292: 845-54.

Berisio, R., F. Sica, V.S. Lamzin, K.S. Wilson, A. Zagari, and L. Mazzarella, 2002. Atomic resolution structures of ribonuclease A at six pH values. Acta Crystallogr D Biol Crystallogr 58: 441-50.

Boix, E., D. Pulido, M. Moussaoui, M.V. Nogues, and S. Russi, 2012a. The sulfate-binding site structure of the human eosinophil cationic protein as revealed by a new crystal form. J Struct Biol 179: 1-9.

Boix, E., Z. Nikolovski, G.P. Moiseyev, H.F. Rosenberg, C.M. Cuchillo, and M.V. Nogues, 1999. Kinetic and product distribution analysis of human eosinophil cationic protein indicates a subsite arrangement that favors exonuclease-type activity. J Biol Chem 274: 15605-14.

Boix, E., V.A. Salazar, M. Torrent, D. Pulido, M.V. Nogues, and M. Moussaoui, 2012b. Structural determinants of the eosinophil cationic protein antimicrobial activity. Biol Chem 393: 801-15.

Borkakoti, N., 1983. The active site of ribonuclease A from the crystallographic studies of ribonuclease-A-inhibitor complexes. Eur J Biochem 132: 89-94.

Brunger, A.T., 1992. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355: 472-5.

Chatani, E., and R. Hayashi, 2001. Functional and structural roles of constituent amino acid residues of bovine pancreatic ribonuclease A. J Biosci Bioeng 92: 98-107.

deMel, V.S., P.D. Martin, M.S. Doscher, and B.F. Edwards, 1992. Structural changes that accompany the reduced catalytic efficiency of two semisynthetic ribonuclease analogs. J Biol Chem 267: 247-56.

Emsley, P., and K. Cowtan, 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126-32.

Fedorov, A.A., D. Joseph-McCarthy, E. Fedorov, D. Sirakova, I. Graf, and S.C. Almo, 1996. Ionic interactions in crystalline bovine pancreatic ribonuclease A. Biochemistry 35: 15962-79.

Fontecilla-Camps, J.C., R. de Llorens, M.H. le Du, and C.M. Cuchillo, 1994. Crystal structure of ribonuclease A.d(ApTpApApG) complex. Direct evidence for extended substrate recognition. J Biol Chem 269: 21526-31.

Garcia-Mayoral, M.F., A. Canales, D. Diaz, J. Lopez-Prados, M. Moussaoui, J.L. de Paz, J. Angulo, P.M. Nieto, J. Jimenez-Barbero, E. Boix, and M. Bruix, 2013. Insights into the glycosaminoglycan-mediated cytotoxic mechanism of eosinophil cationic protein revealed by NMR. ACS Chem Biol 8: 144-51.

Gouet, P., E. Courcelle, D.I. Stuart, and F. Metoz, 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15: 305-8.

Holloway, D.E., G.B. Chavali, M.C. Hares, V. Subramanian, and K.R. Acharya, 2005. Structure of murine angiogenin: features of the substrate- and cell-binding regions and prospects for inhibitor-binding studies. Acta Crystallogr D Biol Crystallogr 61: 1568-78.

Hooft, R.W., G. Vriend, C. Sander, and E.E. Abola, 1996. Errors in protein structures. Nature 381: 272.

Huang, Y.C., Y.M. Lin, T.W. Chang, S.J. Wu, Y.S. Lee, M.D. Chang, C. Chen, S.H. Wu, and Y.D. Liao, 2007. The flexible and clustered lysine residues of human ribonuclease 7 are critical for membrane permeability and antimicrobial activity. J Biol Chem 282: 4626-33.

Incardona, M.F., G.P. Bourenkov, K. Levik, R.A. Pieritz, A.N. Popov, and O. Svensson, 2009. EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. J Synchrotron Radiat 16: 872-9.

Kabsch, W., 2010. Xds. Acta Crystallogr D Biol Crystallogr 66: 125-32.

Krissinel, E., and K. Henrick, 2007. Inference of macromolecular assemblies from crystalline state. J Mol Biol 372: 774-97.

Leonidas, D.D., E. Boix, R. Prill, M. Suzuki, R. Turton, K. Minson, G.J. Swaminathan, R.J. Youle, and K.R. Acharya, 2001. Mapping the ribonucleolytic active site of eosinophil-derived neurotoxin (EDN). High resolution crystal structures of EDN complexes with adenylic nucleotide inhibitors. J Biol Chem 276: 15009-17.

McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read, 2007. Phaser crystallographic software. J Appl Crystallogr 40: 658-674.

Mosimann, S.C., D.L. Newton, R.J. Youle, and M.N. James, 1996. X-ray crystallographic structure of recombinant eosinophil-derived neurotoxin at 1.83 A resolution. J Mol Biol 260: 540-52.

Mueller-Dieckmann, C., S. Panjikar, A. Schmidt, S. Mueller, J. Kuper, A. Geerlof, M. Wilmanns, R.K. Singh, P.A. Tucker, and M.S. Weiss, 2007. On the routine use of soft X-rays in macromolecular crystallography. Part IV. Efficient determination of anomalous substructures in biomacromolecules using longer X-ray wavelengths. Acta Crystallogr D Biol Crystallogr 63: 366-80.

Rico, M., E. Gallego, J. Santoro, F.J. Bermejo, J.L. Nieto, and J. Herranz, 1984. On the fundamental role of the Glu 2- ... Arg 10+ salt bridge in the folding of isolated ribonuclease A S-peptide. Biochem Biophys Res Commun 123: 757-63.

Rosenberg, H.F., and K.D. Dyer, 1996. Molecular cloning and characterization of a novel human ribonuclease (RNase k6): increasing diversity in the enlarging ribonuclease gene family. Nucleic Acids Res 24: 3507-13.

Schultz, L.W., D.J. Quirk, and R.T. Raines, 1998. His...Asp catalytic dyad of ribonuclease A: structure and function of the wild-type, D121N, and D121A enzymes. Biochemistry 37: 8886-98.

Vaguine, A.A., J. Richelle, and S.J. Wodak, 1999. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. Acta Crystallogr D Biol Crystallogr 55: 191-205.

Zegers, I., D. Maes, M.H. Dao-Thi, F. Poortmans, R. Palmer, and L. Wyns, 1994. The structures of RNase A complexed with 3'-CMP and d(CpA): active site conformation and conserved water molecules. Protein Sci 3: 2322-39.

2.1. Structure of the RNase A double mutant (RNase A/H7H10) in complex with 3'-CMP at 2.10 Å

Preliminary Full wwPDB X-ray Structure Validation Report (i)

Oct 15, 2014 – 11:39 AM EDT

The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable23828
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable23828

1   Overall quality at a glance (i)

The reported resolution of this entry is 2.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
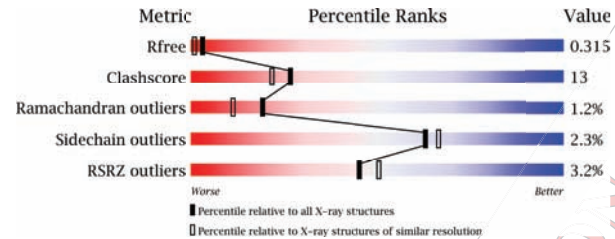
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	3012 (2.10-2.10)
Clashscore	79885	3649 (2.10-2.10)
Ramachandran outliers	78287	3610 (2.10-2.10)
Sidechain outliers	78261	3611 (2.10-2.10)
RSRZ outliers	66119	3013 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	125	
1	B	125	
1	C	125	
2	D	125	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4338 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	125	Total	C	N	O	P	S	0	0	0
			972	584	174	201	1	12			
1	B	125	Total	C	N	O	P	S	0	0	0
			972	584	174	201	1	12			
1	C	125	Total	C	N	O	P	S	0	0	0
			972	584	174	201	1	12			

- Molecule 2 is a protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	D	125	Total	C	N	O	P	S	0	0	0
			971	584	174	200	1	12			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	F	451	Total	O	0	0
			451	451		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1:

Chain A:

- Molecule 1:

Chain B:

- Molecule 1:

Chain C:

- Molecule 2:

Chain D:

4   Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants	160.37Å 32.28Å 106.96Å	Depositor
a, b, c, α, β, γ	90.00° 125.71° 90.00°	
Resolution (Å)	29.12 – 2.10	Depositor
	29.12 – 2.10	EDS
% Data completeness	99.1 (29.12-2.10)	Depositor
(in resolution range)	99.1 (29.12-2.10)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.224 , 0.310	Depositor
	0.227 , 0.315	DCC
R_{free} test set	1333 reflections (5.31%)	DCC
Wilson B-factor (Å^2)	(Not available)	Xtriage
Anisotropy	(Not available)	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 39.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$</L/> =$ (Not available), $< L^2 > =$ (Not available)	Xtriage
Outliers	(Not available)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4338	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *(Not available)*

5   Model quality (i)

5.1   Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: C3P

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#|Z| > 5$	RMSZ	$\#|Z| > 5$
1	A	0.44	0/969	0.60	0/1309
1	B	0.45	0/969	0.59	0/1309
1	C	0.40	0/969	0.58	0/1309
2	D	0.42	0/968	0.56	0/1308
All	All	0.43	0/3875	0.58	0/5235

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	4	0
1	B	3	0
1	C	4	0
2	D	4	0
All	All	15	0

There are no bond length outliers.

There are no bond angle outliers.

All (15) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	500	C3P	C2',C4',C1',C3'
1	B	500	C3P	C2',C4',C3'
1	C	500	C3P	C2',C4',C1',C3'
2	D	500	C3P	C2',C4',C1',C3'

There are no planarity outliers.

LI

5.2   Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	972	0	901	33	0
1	B	972	0	901	15	1
1	C	972	0	901	44	1
2	D	971	0	901	9	0
3	F	451	0	0	44	0
All	All	4338	0	3604	99	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 13.

All (99) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:84:CYS:SG	3:F:420:HOH:O	2.24	0.94
1:C:48:HIS:HE2	1:C:82:THR:HG1	1.16	0.89
2:D:500:C3P:O3P	3:F:9:HOH:O	1.90	0.89
1:C:90:SER:HA	1:C:96:ALA:H	1.41	0.86
1:C:83:ASP:OD2	1:C:85:ARG:NE	2.10	0.85
1:C:26:CYS:SG	3:F:370:HOH:O	2.42	0.77
1:A:1:LYS:HG3	1:A:2:GLU:H	1.55	0.71
1:C:100:THR:OG1	3:F:279:HOH:O	2.06	0.71
1:B:47:VAL:HB	3:F:156:HOH:O	1.89	0.71
1:A:23:SER:O	3:F:253:HOH:O	2.09	0.69
1:A:16:SER:OG	3:F:333:HOH:O	2.14	0.66
1:A:105:HIS:ND1	3:F:155:HOH:O	2.28	0.66
1:C:20:ALA:O	3:F:382:HOH:O	2.13	0.66
1:A:91:LYS:NZ	3:F:335:HOH:O	2.29	0.65
1:A:78:THR:HG22	3:F:155:HOH:O	1.96	0.64
1:A:33:ARG:HB3	3:F:235:HOH:O	1.98	0.64
1:C:91:LYS:NZ	3:F:368:HOH:O	2.26	0.63
1:C:26:CYS:N	3:F:370:HOH:O	2.30	0.63
1:B:9:GLU:OE1	3:F:122:HOH:O	2.16	0.62
1:A:28:GLN:HE22	1:C:24:ASN:HD21	1.48	0.61
1:C:98:LYS:HG3	3:F:279:HOH:O	2.02	0.60
2:D:87:THR:HG22	2:D:89:SER:H	1.66	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:85:ARG:N	3:F:381:HOH:O	2.35	0.60
1:C:90:SER:HA	1:C:96:ALA:N	2.16	0.59
1:C:41:LYS:HD3	1:C:44:ASN:HB2	1.84	0.59
1:C:97:TYR:OH	3:F:371:HOH:O	2.17	0.57
1:C:29:MET:HG2	3:F:435:HOH:O	2.04	0.57
1:A:105:HIS:N	3:F:155:HOH:O	2.36	0.57
1:B:1:LYS:N	1:B:2:GLU:HA	2.19	0.57
1:C:14:ASP:OD2	1:C:25:TYR:OH	2.21	0.57
1:A:97:TYR:HB3	3:F:260:HOH:O	2.04	0.56
1:C:98:LYS:N	3:F:381:HOH:O	2.39	0.55
1:A:1:LYS:HG3	1:A:2:GLU:N	2.21	0.55
1:C:40:CYS:SG	1:C:92:TYR:HB2	2.47	0.55
1:C:38:ASP:HB3	3:F:244:HOH:O	2.08	0.54
1:C:9:GLU:OE1	3:F:321:HOH:O	2.18	0.53
1:C:25:TYR:HD2	3:F:370:HOH:O	1.92	0.53
1:A:20:ALA:O	1:A:21:SER:HB2	2.09	0.53
1:B:12:HIS:CE1	1:B:500:C3P:O2'	2.63	0.52
1:B:66:LYS:HD3	3:F:82:HOH:O	2.09	0.52
1:C:12:HIS:CE1	1:C:500:C3P:O2'	2.63	0.51
1:A:20:ALA:O	3:F:210:HOH:O	2.19	0.51
1:A:85:ARG:O	3:F:13:HOH:O	2.19	0.51
1:A:21:SER:HA	1:A:22:SER:C	2.32	0.51
1:A:88:GLY:HA2	3:F:377:HOH:O	2.12	0.50
2:D:45:THR:HG21	2:D:120:PHE:CZ	2.47	0.50
1:A:114:PRO:HD2	3:F:284:HOH:O	2.10	0.50
1:C:22:SER:OG	3:F:227:HOH:O	2.19	0.50
1:C:89:SER:OG	1:C:90:SER:N	2.44	0.50
1:A:22:SER:C	1:A:24:ASN:H	2.14	0.49
1:C:67:ASN:ND2	1:C:69:GLN:OE1	2.42	0.49
1:B:98:LYS:HE2	3:F:307:HOH:O	2.12	0.48
2:D:87:THR:HG22	2:D:89:SER:N	2.29	0.48
1:A:104:LYS:HD3	3:F:139:HOH:O	2.13	0.48
1:A:78:THR:HA	3:F:155:HOH:O	2.14	0.48
1:C:87:THR:HG21	3:F:362:HOH:O	2.13	0.48
2:D:105:HIS:HB2	2:D:124:VAL:O	2.12	0.48
1:A:43:VAL:O	3:F:125:HOH:O	2.20	0.48
1:A:14:ASP:OD2	1:A:17:THR:HG22	2.14	0.48
1:B:118:VAL:HG23	1:B:119:HIS:CD2	2.49	0.48
1:C:20:ALA:HB2	1:C:82:THR:HG21	1.96	0.48
1:C:2:GLU:HG2	1:C:7:HIS:HB2	1.96	0.47
1:A:12:HIS:CE1	1:A:500:C3P:O2'	2.68	0.47
1:A:21:SER:HA	1:A:22:SER:O	2.16	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:66:LYS:HD2	1:B:66:LYS:HA	1.54	0.46
2:D:83:ASP:HA	3:F:61:HOH:O	2.17	0.45
2:D:86:GLU:HG3	2:D:90:SER:HB3	1.97	0.45
1:B:105:HIS:HB2	1:B:124:VAL:HG13	1.99	0.45
1:B:24:ASN:O	1:B:28:GLN:HG3	2.17	0.45
1:A:29:MET:O	1:A:33:ARG:HB2	2.17	0.45
1:B:81:ILE:HD11	1:B:104:LYS:HD3	1.98	0.45
1:A:25:TYR:OH	1:A:48:HIS:HE1	1.99	0.44
1:A:29:MET:HA	1:A:32:SER:OG	2.17	0.44
1:A:109:ALA:O	1:A:117:PRO:HA	2.17	0.44
1:C:92:TYR:CG	1:C:93:PRO:HA	2.53	0.44
1:C:27:ASN:OD1	1:C:97:TYR:N	2.48	0.44
1:C:38:ASP:N	3:F:169:HOH:O	2.50	0.44
1:C:92:TYR:CD2	1:C:93:PRO:HA	2.52	0.44
2:D:105:HIS:HB2	2:D:124:VAL:HG23	1.99	0.44
1:A:118:VAL:HG23	1:A:119:HIS:CD2	2.53	0.43
1:C:66:LYS:HE3	1:C:122:ALA:HB2	2.00	0.43
1:C:31:LYS:HA	1:C:36:THR:OG1	2.18	0.43
1:A:109:ALA:HB3	1:A:119:HIS:HB2	2.00	0.43
2:D:12:HIS:CE1	2:D:500:C3P:O2'	2.71	0.43
1:A:28:GLN:HE22	1:C:24:ASN:ND2	2.16	0.42
1:A:37:LYS:HG3	1:A:38:ASP:OD1	2.19	0.42
1:B:500:C3P:H5'1	3:F:363:HOH:O	2.18	0.42
1:C:103:ASN:ND2	3:F:59:HOH:O	2.50	0.42
1:B:31:LYS:HG3	3:F:220:HOH:O	2.19	0.42
1:C:85:ARG:HH12	1:C:98:LYS:HG3	1.86	0.41
1:C:25:TYR:CZ	1:C:29:MET:HG3	2.55	0.41
1:C:19:ALA:O	3:F:437:HOH:O	2.22	0.41
1:A:10:HIS:CD2	3:F:235:HOH:O	2.74	0.41
1:C:14:ASP:O	1:C:48:HIS:HA	2.21	0.41
1:B:92:TYR:CD1	1:B:93:PRO:HA	2.56	0.41
1:C:12:HIS:HE1	1:C:500:C3P:O2'	2.04	0.40
1:B:36:THR:HA	1:B:39:ARG:O	2.20	0.40
1:C:17:THR:HB	3:F:301:HOH:O	2.21	0.40
1:C:42:PRO:HA	1:C:86:GLU:HB3	2.04	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:77:SER:OG	1:C:68:GLY:O[4_647]	2.16	0.04

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	122/125 (98%)	112 (92%)	6 (5%)	4 (3%)	6	1
1	B	122/125 (98%)	117 (96%)	5 (4%)	0	100	100
1	C	122/125 (98%)	109 (89%)	11 (9%)	2 (2%)	14	7
2	D	122/125 (98%)	117 (96%)	5 (4%)	0	100	100
All	All	488/500 (98%)	455 (93%)	27 (6%)	6 (1%)	19	11

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	21	SER
1	C	89	SER
1	A	2	GLU
1	A	22	SER
1	A	19	ALA
1	C	91	LYS

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	109/109 (100%)	106 (97%)	3 (3%)	56	59
1	B	109/109 (100%)	107 (98%)	2 (2%)	71	75
1	C	109/109 (100%)	106 (97%)	3 (3%)	56	59
2	D	108/108 (100%)	106 (98%)	2 (2%)	69	73
All	All	435/435 (100%)	425 (98%)	10 (2%)	63	66

LIII

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	LYS
1	A	29	MET
1	A	39	ARG
1	B	2	GLU
1	B	66	LYS
1	C	15	SER
1	C	86	GLU
1	C	89	SER
2	D	87	THR
2	D	124	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	28	GLN
1	B	24	ASN
1	C	12	HIS

5.3.3   RNA ⓘ

There are no RNA chains in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6   Ligand geometry ⓘ

There are no ligands in this entry.

5.7   Other polymers ⓘ

There are no such residues in this entry.

5.8   Polymer linkage issues

There are no chain breaks in this entry.

LIV

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	125/125 (100%)	0.23	5 (4%)	36	41	18, 30, 51, 61	0
1	B	125/125 (100%)	0.01	4 (3%)	45	50	19, 28, 44, 65	0
1	C	125/125 (100%)	0.32	7 (5%)	24	26	15, 31, 56, 70	1 (0%)
2	D	125/125 (100%)	-0.19	0	100	100	17, 28, 38, 50	0
All	All	500/500 (100%)	0.09	16 (3%)	45	50	15, 29, 51, 70	1 (0%)

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2	GLU	3.6
1	C	1	LYS	3.4
1	A	22	SER	3.3
1	C	85	ARG	3.3
1	C	19	ALA	3.1
1	A	37	LYS	2.5
1	B	23	SER	2.5
1	A	34	ASN	2.4
1	C	22	SER	2.3
1	C	90	SER	2.3
1	C	92	TYR	2.3
1	A	21	SER	2.1
1	B	37	LYS	2.1
1	B	1	LYS	2.1
1	C	91	LYS	2.1
1	A	39	ARG	2.1

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.

LV

2.2. Structure of RNase A at high resolution in complex with 3-CMP at 1.16 Å (PDB ID: 4U7R)

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Preliminary Full wwPDB X-ray Structure Validation Report (i)

May 18, 2015 – 01:41 PM EDT

1 Overall quality at a glance (i)

The reported resolution of this entry is 1.16 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
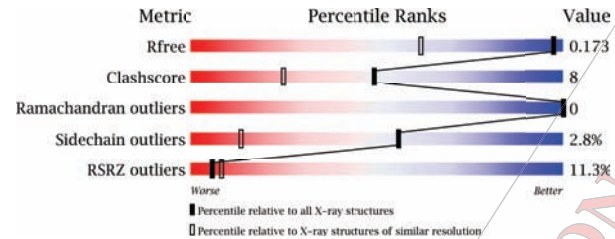
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1042 (1.22-1.10)
Clashscore	79885	1154 (1.22-1.10)
Ramachandran outliers	78287	1097 (1.22-1.10)
Sidechain outliers	78261	1093 (1.22-1.10)
RSRZ outliers	66119	1042 (1.22-1.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	124	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	B	2	-	X
3	CL	C	2	-	X

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2   Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 2302 atoms, of which 1007 are hydrogens and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	124	Total 2040	C 626	H 1007	N 184	O 210	S 13	0	19	0

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: unknown).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: unknown).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	1	Total 1	Cl 1	0	0
3	C	1	Total 1	Cl 1	0	0
3	C	1	Total 1	Cl 1	0	0
3	C	1	Total 1	Cl 1	0	0

- Molecule 4 is CYTIDINE-3'-MONOPHOSPHATE (three-letter code: C3P) (formula: unknown).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	D	1	Total 21	C 9	N 3	O 8	P 1	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: unknown).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	E	1	Total 1	Na 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	F	205	Total 221	O 221	0	16

LVIII

3   Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1:

Chain A:

4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	64.06Å 64.06Å 64.06Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.65 – 1.16 32.03 – 1.16	Depositor EDS
% Data completeness (in resolution range)	100.0 (28.65-1.16) 100.0 (32.03-1.16)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.63 (at 1.16Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.144 , 0.172 0.145 , 0.173	Depositor DCC
R_{free} test set	2695 reflections (5.09%)	DCC
Wilson B-factor (Å²)	14.5	Xtriage
Anisotropy	0.086	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.42 , 74.1	EDS
Estimated twinning fraction	0.034 for -h,-k,l	Xtriage
L-test for twinning	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Outliers	0 of 52958 reflections	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2302	wwPDB-VP
Average B, all atoms (Å²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.04% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NA, CL, C3P, SO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
1	A	0.42	0/1125	0.60	0/1515

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1033	1007	981	14	9
2	B	15	0	0	1	1
3	C	4	0	0	0	1
4	D	21	0	12	1	0
5	E	1	0	0	0	0
6	F	221	0	0	7	2
All	All	1295	1007	993	16	10

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 8.

All (16) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:119[B]:HIS:ND1	4:D:1:C3P:O3P	2.17	0.78
1:A:104:LYS:HE3	1:A:104:LYS:HA	1.64	0.77
1:A:14:ASP:OD1	6:F:273:HOH:O	2.14	0.65
2:B:2:SO4:O1	6:F:296[B]:HOH:O	2.13	0.62
1:A:101[B]:GLN:NE2	6:F:112:HOH:O	2.15	0.62
1:A:103:ASN:O	1:A:104:LYS:HD2	2.01	0.60
1:A:41[B]:LYS:NZ	1:A:44:ASN:OD1	2.40	0.54
1:A:25:TYR:CZ	1:A:29[A]:MET:HG3	2.43	0.53
1:A:103:ASN:C	1:A:104:LYS:HD2	2.30	0.52
1:A:89:SER:O	1:A:89:SER:OG	2.30	0.49
1:A:29[B]:MET:HG3	1:A:46:PHE:CZ	2.52	0.44
1:A:69:GLN:NE2	6:F:92:HOH:O	2.51	0.43
1:A:113[B]:ASN:ND2	6:F:186:HOH:O	2.52	0.43
1:A:1[A]:LYS:NZ	6:F:87:HOH:O	2.48	0.42
1:A:43:VAL:HG22	1:A:85[B]:ARG:HG2	2.02	0.41

All (10) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:2:SO4:O4	6:F:296[B]:HOH:O[4_555]	2.13	0.07
1:A:16[B]:SER:HG	1:A:59[B]:SER:HG[2_654]	1.65	-0.05
1:A:66:LYS:HZ1	1:A:85[B]:ARG:HH12[4_555]	1.79	-0.19
1:A:104:LYS:HZ1	6:F:29[A]:HOH:O[4_555]	1.89	-0.29
1:A:16[B]:SER:OG	1:A:59[B]:SER:HG[2_654]	1.96	-0.36
1:A:16[B]:SER:HG	1:A:59[B]:SER:OG[2_654]	1.97	-0.37
1:A:1[B]:LYS:HB2	1:A:101[B]:GLN:HE21[6_654]	2.02	-0.42
1:A:66:LYS:HZ2	1:A:85[A]:ARG:HH22[4_555]	2.13	-0.53
1:A:16[B]:SER:HG	3:C:3:CL:CL[2_654]	2.17	-0.57
1:A:85[B]:ARG:HH12	1:A:121[B]:ASP:OD2[4_555]	2.17	-0.57

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	141/124 (114%)	138 (98%)	3 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	129/109 (118%)	124 (96%)	5 (4%)	43	7

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1[A]	LYS
1	A	1[B]	LYS
1	A	89	SER
1	A	121[A]	ASP
1	A	121[B]	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 5 are modelled with single atom - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	SO4	B	1	-	4,?,?	0.10	0	6,?,?	0.13	0
2	SO4	B	2	-	4,?,?	0.08	0	6,?,?	0.18	0
2	SO4	B	3	-	4,?,?	0.08	0	6,?,?	0.06	0
4	C3P	D	1	-	22,?,?	2.15	4 (18%)	33,?,?	1.49	8 (24%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SO4	B	1	-	-	0/0/?/?	0/0/?/?
2	SO4	B	2	-	-	0/0/?/?	0/0/?/?
2	SO4	B	3	-	-	0/0/?/?	0/0/?/?
4	C3P	D	1	-	-	0/11/?/?	0/2/?/?

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1	C3P	P-O3'	7.16	1.81	1.60
4	D	1	C3P	O2-C2	-3.82	1.16	1.23
4	D	1	C3P	O3'-C3'	-3.55	1.33	1.44
4	D	1	C3P	C2-N1	2.16	1.44	1.40

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	1	C3P	P-O3'-C3'	-3.11	114.10	121.56
4	D	1	C3P	C5-C4-N3	2.89	126.40	121.34
4	D	1	C3P	O2P-P-O3P	2.61	117.34	107.38

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	1	C3P	C4-N3-C2	-2.61	116.02	120.29
4	D	1	C3P	N4-C4-N3	-2.47	113.76	118.02
4	D	1	C3P	O3'-P-O1P	-2.27	101.45	107.11
4	D	1	C3P	O2P-P-O1P	2.21	117.68	110.58
4	D	1	C3P	O4'-C1'-C2'	2.11	111.49	106.58

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7   Other polymers ⓘ

There are no such residues in this entry.

5.8   Polymer linkage issues

There are no chain breaks in this entry.

6   Fit of model and data ⓘ

6.1   Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	124/124 (100%)	0.52	14 (11%)	6	8	12, 17, 41, 56	12 (9%)

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	92	TYR	7.7
1	A	21	SER	6.2
1	A	38	ASP	5.5
1	A	93	PRO	4.6
1	A	89	SER	3.5
1	A	91	LYS	3.4
1	A	39	ARG	3.4
1	A	1[A]	LYS	3.2
1	A	37	LYS	2.9
1	A	94	ASN	2.8
1	A	113[A]	ASN	2.8
1	A	40	CYS	2.3
1	A	19	ALA	2.0
1	A	68	GLY	2.0

6.2   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3   Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4   Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors($Å^2$)	Q<0.9
2	SO4	B	2	5/?	0.43	11.04	46,46,57,101	5
3	CL	C	2	1/?	0.12	2.11	13,13,13,13	1
4	C3P	D	1	21/?	0.13	1.42	13,23,46,53	9
5	NA	E	1	1/?	0.07	1.23	20,20,20,20	0
3	CL	C	4	1/?	0.07	0.98	20,20,20,20	1
2	SO4	B	1	5/?	0.15	0.53	17,21,42,76	5
3	CL	C	1	1/?	0.07	-0.20	16,16,16,16	1
3	CL	C	3	1/?	0.05	-0.72	20,20,20,20	1
2	SO4	B	3	5/?	0.13	-	23,26,27,34	5

6.5   Other polymers ⓘ

There are no such residues in this entry.

2.3. Structure of ECP with sulphate anions at 1.50 Å (PDB ID: 4OXB)
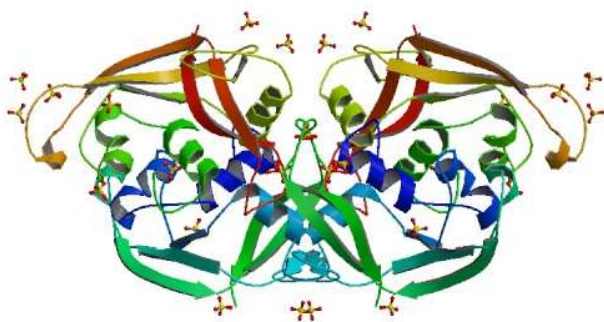
Full wwPDB X-ray Structure Validation Report ⓘ

Feb 26, 2014 – 02:03 PM EST

PDB ID : 4OXB
Title : Structure of ECP with sulphate anions at 1.50 Angstroms
Authors : Blanco, J.A.; Boix, E.; Moussaoui, M.
Deposited on : 2014-02-05
Resolution : 1.50 Å(reported)

The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable22501
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22501

---

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 1.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1513 (1.50-1.50)
Clashscore	79885	1768 (1.50-1.50)
Ramachandran outliers	78287	1720 (1.50-1.50)
Sidechain outliers	78261	1718 (1.50-1.50)
RSRZ outliers	66119	1514 (1.50-1.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	134	
1	B	134	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	A	301	-	X
2	SO4	A	302[A]	-	X
2	SO4	A	302[B]	-	X
2	SO4	A	303	-	X
2	SO4	A	304	-	X
2	SO4	A	305	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	A	306	-	X
2	SO4	B	301	-	X
2	SO4	B	302	-	X
2	SO4	B	303	-	X
2	SO4	B	304	-	X
2	SO4	B	305	-	X
2	SO4	B	307	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2765 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Eosinophil cationic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	133	Total 1165	C 725	N 243	O 188	S 9	0	13	0
1	B	133	Total 1159	C 718	N 243	O 187	S 11	0	13	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P12724
A	97	ARG	THR	variant	UNP P12724
B	0	MET	-	initiating methionine	UNP P12724
B	97	ARG	THR	variant	UNP P12724

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 10	O 8	S 2	0	1
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	203	Total 203	O 203	0	0
3	B	168	Total 168	O 168	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Eosinophil cationic protein

Chain A:

- Molecule 1: Eosinophil cationic protein

Chain B:

LXVII

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	92.73Å 51.29Å 55.62Å 90.00° 111.25° 90.00°	Depositor
Resolution (Å)	44.10 – 1.50 44.10 – 1.50	Depositor EDS
% Data completeness (in resolution range)	98.8 (44.10-1.50) 95.5 (44.10-1.50)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 1.50Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.179 , 0.232 0.179 , 0.231	Depositor DCC
R_{free} test set	1882 reflections (5.04%)	DCC
Wilson B-factor (Å^2)	15.3	Xtriage
Anisotropy	0.122	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 38.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.48, < L^2 > = 0.32$	Xtriage
Outliers	1 of 38673 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2765	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.50% of the height of the origin peak. No significant pseudotranslation is detected.*

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$ >2	RMSZ	#$|Z|$ >2
1	A	0.50	6/880 (0.7%)	0.59	9/995 (0.9%)
1	B	0.42	2/887 (0.2%)	0.55	5/1006 (0.5%)
All	All	0.46	8/1767 (0.5%)	0.57	14/2001 (0.7%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	19	ASN	CG-OD1	4.60	1.34	1.24
1	A	92	ASN	CG-OD1	4.29	1.33	1.24
1	B	92	ASN	CG-OD1	4.23	1.33	1.24
1	A	19	ASN	CG-ND2	-4.16	1.22	1.32
1	A	64	HIS	CG-CD2	3.85	1.41	1.35
1	B	92	ASN	CG-ND2	-3.68	1.23	1.32
1	A	92	ASN	CG-ND2	-3.62	1.23	1.32
1	A	33	TYR	CD1-CE1	-2.07	1.36	1.39

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	92	ASN	CB-CG-OD1	-2.56	116.49	121.60
1	B	92	ASN	CB-CG-OD1	-2.56	116.49	121.60
1	A	15	HIS	N-CA-C	2.52	117.81	111.00

Continued on next page...

LXVIII

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	64	HIS	CA-C-O	2.48	125.32	120.10
1	A	114	ARG	NE-CZ-NH2	-2.40	119.10	120.30
1	A	66	ARG	NE-CZ-NH2	-2.37	119.11	120.30
1	B	34	ARG	NE-CZ-NH2	-2.32	119.14	120.30
1	A	91	GLN	N-CA-CB	2.24	114.63	110.60
1	B	15	HIS	N-CA-C	2.24	117.04	111.00
1	B	104	ARG	N-CA-C	-2.24	104.96	111.00
1	B	97	ARG	NE-CZ-NH2	-2.23	119.18	120.30
1	A	91	GLN	CB-CA-C	2.23	114.86	110.40
1	A	19	ASN	CB-CG-OD1	-2.13	117.33	121.60
1	A	96	CYS	CA-CB-SG	-2.11	110.20	114.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	90	ALA	Peptide

5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1165	0	1149	18	0
1	B	1159	0	1141	11	0
2	A	35	0	0	0	0
2	B	35	0	0	2	0
3	A	203	0	0	5	0
3	B	168	0	0	6	0
All	All	2765	0	2290	27	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

All (27) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:91[A]:GLN:NE2	3:B:401:HOH:O	2.25	0.69
1:B:77[A]:ARG:NH1	2:B:303:SO4:O4	2.27	0.66
1:B:4:GLN:NE2	3:B:402:HOH:O	2.36	0.58
1:A:53[A]:ASN:ND2	3:A:585:HOH:O	2.33	0.57
1:B:7[B]:ARG:NH1	3:B:404:HOH:O	2.38	0.55
1:A:91:GLN:HG3	1:A:95:ASN:HD22	1.73	0.54
1:A:42:THR:OG1	1:A:82[A]:HIS:HD2	1.93	0.51
1:A:122[B]:TYR:OH	1:B:49:ALA:HB2	2.11	0.51
1:A:86[A]:ILE:HG12	1:A:97:ARG:O	2.11	0.50
1:A:91:GLN:CG	1:A:92:ASN:H	2.23	0.50
1:A:18:LEU:HD12	3:B:560:HOH:O	2.12	0.50
1:A:7[B]:ARG:NH2	3:A:580:HOH:O	2.25	0.49
1:A:42:THR:OG1	1:A:82[B]:HIS:HD2	1.97	0.48
1:A:91:GLN:HG2	1:A:92:ASN:H	1.80	0.47
3:A:564:HOH:O	1:B:18:LEU:HD12	2.15	0.46
1:A:1:ARG:NH1	1:A:5:PHE:O	2.41	0.46
1:B:57:ASN:HB3	2:B:307:SO4:O3	2.18	0.44
1:A:114:ARG:HB2	1:A:118:ASP:HB2	1.99	0.44
1:A:122[B]:TYR:HH	1:B:122:TYR:HH	1.53	0.44
1:A:7[A]:ARG:NH1	3:A:558:HOH:O	2.52	0.43
1:A:82[B]:HIS:HB2	1:A:101[B]:ARG:HB3	2.01	0.43
1:B:82[B]:HIS:HD2	3:B:522:HOH:O	2.02	0.42
1:A:18:LEU:HD23	1:A:18:LEU:HA	1.91	0.42
1:A:73:ARG:NH1	3:A:411:HOH:O	2.54	0.41
1:B:37[B]:CYS:SG	3:B:503:HOH:O	2.62	0.41
1:A:42:THR:OG1	1:A:82[B]:HIS:CD2	2.74	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	144/134 (108%)	138 (96%)	4 (3%)	2 (1%)	16	2
1	B	144/134 (108%)	143 (99%)	1 (1%)	0	100	100
All	All	288/268 (108%)	281 (98%)	5 (2%)	2 (1%)	27	7

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	91	GLN
1	A	89	GLY

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	135/123 (110%)	132 (98%)	3 (2%)	64	28
1	B	135/123 (110%)	131 (97%)	4 (3%)	53	16
All	All	270/246 (110%)	263 (97%)	7 (3%)	75	21

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	91	GLN
1	A	122[A]	TYR
1	A	122[B]	TYR
1	B	86[A]	ILE
1	B	86[B]	ILE
1	B	91[A]	GLN
1	B	91[B]	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	19	ASN
1	A	95	ASN

5.3.3 RNA (i)

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

14 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	SO4	A	301	-	4,4,4	0.14	0	6,6,6	0.10	0
2	SO4	A	302[A]	-	4,4,4	0.19	0	6,6,6	0.09	0
2	SO4	A	302[B]	-	4,4,4	0.22	0	6,6,6	0.09	0
2	SO4	A	303	-	4,4,4	0.19	0	6,6,6	0.07	0
2	SO4	A	304	-	4,4,4	0.20	0	6,6,6	0.07	0
2	SO4	A	305	-	4,4,4	0.24	0	6,6,6	0.29	0
2	SO4	A	306	-	4,4,4	0.26	0	6,6,6	0.08	0
2	SO4	B	301	-	4,4,4	0.13	0	6,6,6	0.08	0
2	SO4	B	302	-	4,4,4	0.21	0	6,6,6	0.10	0
2	SO4	B	303	-	4,4,4	0.16	0	6,6,6	0.14	0
2	SO4	B	304	-	4,4,4	0.16	0	6,6,6	0.10	0
2	SO4	B	305	-	4,4,4	0.17	0	6,6,6	0.08	0
2	SO4	B	306	-	4,4,4	0.21	0	6,6,6	0.12	0
2	SO4	B	307	-	4,4,4	0.19	0	6,6,6	0.14	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SO4	A	301	-	-	0/0/0/0	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SO4	A	302[A]	-	-	0/0/0/0	0/0/0/0
2	SO4	A	302[B]	-	-	0/0/0/0	0/0/0/0
2	SO4	A	303	-	-	0/0/0/0	0/0/0/0
2	SO4	A	304	-	-	0/0/0/0	0/0/0/0
2	SO4	A	305	-	-	0/0/0/0	0/0/0/0
2	SO4	A	306	-	-	0/0/0/0	0/0/0/0
2	SO4	B	301	-	-	0/0/0/0	0/0/0/0
2	SO4	B	302	-	-	0/0/0/0	0/0/0/0
2	SO4	B	303	-	-	0/0/0/0	0/0/0/0
2	SO4	B	304	-	-	0/0/0/0	0/0/0/0
2	SO4	B	305	-	-	0/0/0/0	0/0/0/0
2	SO4	B	306	-	-	0/0/0/0	0/0/0/0
2	SO4	B	307	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	\<RSRZ\>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	133/134 (99%)	0.15	4 (3%)	48	53	12, 17, 31, 53	0
1	B	133/134 (99%)	0.06	2 (1%)	70	77	12, 17, 27, 38	0
All	All	266/268 (99%)	0.10	6 (2%)	54	64	12, 17, 30, 53	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	35	TRP	7.2
1	B	35	TRP	3.1
1	A	86[A]	ILE	2.8
1	A	133	ILE	2.8
1	A	90	ALA	2.8
1	B	91[A]	GLN	2.6

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors($Å^2$)	Q<0.9
2	SO4	A	302[A]	5/5	0.35	28.84	32,35,37,39	5
2	SO4	A	302[B]	5/5	0.35	28.78	33,33,38,38	5
2	SO4	B	301	5/5	0.24	9.82	24,45,56,62	0
2	SO4	A	301	5/5	0.19	8.16	24,45,52,68	0
2	SO4	A	305	5/5	0.25	7.50	23,28,41,41	0
2	SO4	B	302	5/5	0.22	4.74	29,31,33,36	5
2	SO4	A	304	5/5	0.15	4.56	23,24,26,27	5
2	SO4	B	304	5/5	0.24	4.55	35,52,55,57	0
2	SO4	B	305	5/5	0.14	3.87	28,31,36,37	0
2	SO4	A	303	5/5	0.23	2.67	46,50,55,57	0
2	SO4	B	307	5/5	0.16	2.29	24,26,29,36	0
2	SO4	A	306	5/5	0.14	2.11	25,30,33,41	0
2	SO4	B	303	5/5	0.30	2.06	34,45,58,69	0
2	SO4	B	306	5/5	0.12	0.06	23,26,30,30	0

6.5 Other polymers (i)

There are no such residues in this entry.

2.4. Structure of ECP with citrate anions at 1.50 Å (PDB ID: 4OXF)

![WORLDWIDE PDB PROTEIN DATA BANK]
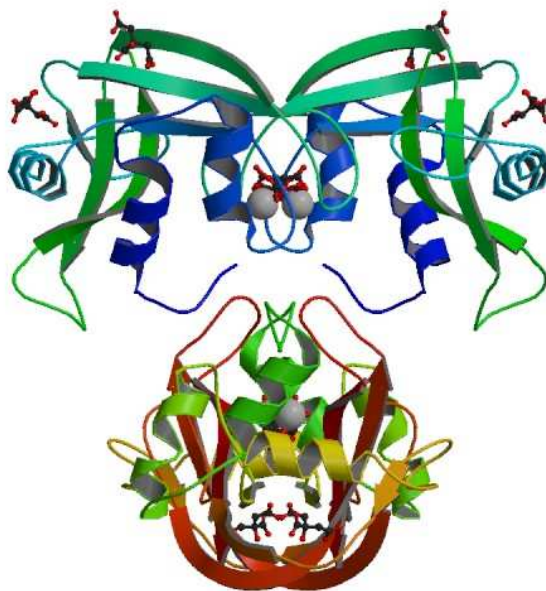
Full wwPDB X-ray Structure Validation Report (i)

Feb 26, 2014 – 01:38 PM EST

PDB ID   :   4OXF
Title   :   Structure of ECP in complex with citrate ions at 1.50 Angstroms
Authors   :   Blanco, J.A.; Boix, E.; Moussaoui, M.; Salazar, V.A.
Deposited on :   2014-02-05
Resolution :   1.50 Å(reported)

The following versions of software and data (see references) were used in the production of this report:

MolProbity   :   4.02b-467
Mogul   :   1.15 2013
Xtriage (Phenix)   :   dev-1439
EDS   :   stable22501
Percentile statistics :   21963
Refmac   :   5.8.0049
CCP4   :   6.3.0 (Settle)
Ideal geometry (proteins)   :   Engh & Huber (2001)
Ideal geometry (DNA, RNA)   :   Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)   :   stable22501

1   Overall quality at a glance (i)

The reported resolution of this entry is 1.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1513 (1.50-1.50)
Clashscore	79885	1768 (1.50-1.50)
Ramachandran outliers	78287	1720 (1.50-1.50)
Sidechain outliers	78261	1718 (1.50-1.50)
RSRZ outliers	66119	1514 (1.50-1.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	134	
1	B	134	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	CIT	A	301	-	X
2	CIT	A	302	-	X
2	CIT	A	303	-	X
2	CIT	B	302	-	X

![WORLDWIDE PDB PROTEIN DATA BANK]

2　Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2798 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Eosinophil cationic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	134	Total	C	N	O	S	0	14	0
			1182	732	249	189	12			
1	B	133	Total	C	N	O	S	0	10	0
			1149	716	238	186	9			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P12724
A	97	ARG	THR	variant	UNP P12724
B	0	MET	-	initiating methionine	UNP P12724
B	97	ARG	THR	variant	UNP P12724

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	6	7		
2	A	1	Total	C	O	0	0
			13	6	7		
2	A	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Fe	0	0
			1	1		
3	A	1	Total	Fe	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	212	Total	O	0	0
			212	212		
4	B	188	Total	O	0	0
			188	188		

3   Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Eosinophil cationic protein

Chain A:

- Molecule 1: Eosinophil cationic protein

Chain B:

4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, α, β, γ	62.53Å 62.53Å 175.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.88 – 1.50 62.53 – 1.50	Depositor EDS
% Data completeness (in resolution range)	98.8 (50.88-1.50) 96.5 (62.53-1.50)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.04 (at 1.50Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.197 , 0.231 0.197 , 0.228	Depositor DCC
R_{free} test set	2762 reflections (5.05%)	DCC
Wilson B-factor (Å^2)	19.6	Xtriage
Anisotropy	0.199	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 41.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< \|L\| > = 0.48, < L^2 > = 0.32$	Xtriage
Outliers	0 of 55985 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2798	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.33% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: FE, CIT

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >2	RMSZ	#|Z| >2
1	A	0.46	2/889 (0.2%)	0.62	11/1007 (1.1%)
1	B	0.42	2/908 (0.2%)	0.62	8/1028 (0.8%)
All	All	0.44	4/1797 (0.2%)	0.62	19/2035 (0.9%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	40	GLN	C-O	-4.30	1.15	1.23
1	A	58	GLN	C-O	-4.28	1.15	1.23
1	B	36	ARG	CZ-NH1	-3.56	1.28	1.33
1	B	36	ARG	C-O	-2.94	1.17	1.23

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	34	ARG	NE-CZ-NH2	-5.97	117.31	120.30
1	A	101	ARG	NE-CZ-NH1	-5.69	117.46	120.30
1	B	36	ARG	NE-CZ-NH2	4.40	122.50	120.30
1	B	36	ARG	NE-CZ-NH1	-4.22	118.19	120.30
1	B	34	ARG	NE-CZ-NH1	3.26	121.93	120.30
1	B	115	ASP	CB-CG-OD2	2.73	120.76	118.30
1	A	107	TYR	CA-CB-CG	2.72	118.57	113.40
1	A	34	ARG	NE-CZ-NH2	-2.65	118.97	120.30
1	A	34	ARG	NE-CZ-NH1	2.55	121.57	120.30
1	A	45	ARG	NE-CZ-NH2	-2.51	119.04	120.30
1	A	101	ARG	NE-CZ-NH2	2.49	121.54	120.30
1	B	107	TYR	CA-CB-CG	2.39	117.94	113.40
1	B	80	LEU	CA-CB-CG	2.24	120.45	115.30
1	B	15	HIS	N-CA-C	2.19	116.92	111.00
1	A	101	ARG	CB-CG-CD	-2.10	106.15	111.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	18	LEU	CB-CG-CD2	-2.06	107.50	111.00
1	A	61	ARG	NE-CZ-NH2	-2.04	119.28	120.30
1	A	125	VAL	CB-CA-C	-2.02	107.57	111.40
1	A	80	LEU	CA-CB-CG	2.01	119.92	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1182	0	1165	12	1
1	B	1149	0	1144	10	0
2	A	39	0	14	1	4
2	B	26	0	8	2	2
3	A	1	0	0	0	3
3	B	1	0	0	0	2
4	A	212	0	0	9	0
4	B	188	0	0	12	0
All	All	2798	0	2331	25	6

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

All (25) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:101:ARG:HG3	4:A:591:HOH:O	1.52	1.09
1:A:101:ARG:CG	4:A:591:HOH:O	1.97	1.08
1:A:101:ARG:HD3	4:A:612:HOH:O	1.52	1.07
2:B:302:CIT:O2	4:B:401:HOH:O	1.92	0.86
1:A:101:ARG:HG2	4:A:591:HOH:O	1.71	0.71
1:B:58[A]:GLN:OE1	4:B:402:HOH:O	2.09	0.70
1:A:22:ARG:HE	1:A:24:THR:HB	1.58	0.69
1:A:95[B]:ASN:ND2	4:A:402:HOH:O	2.27	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:A:302:CIT:O6	4:A:401:HOH:O	2.13	0.66
1:B:101:ARG:CD	4:B:406:HOH:O	2.43	0.65
2:B:302:CIT:O6	4:B:403:HOH:O	2.15	0.64
1:B:101:ARG:NE	4:B:406:HOH:O	2.33	0.61
1:B:36:ARG:NH2	4:B:408:HOH:O	2.39	0.54
1:A:93:ILE:O	4:A:536:HOH:O	2.18	0.53
1:A:22:ARG:HG3	4:A:499:HOH:O	2.08	0.52
1:B:22:ARG:NH2	4:B:477:HOH:O	2.42	0.52
1:B:75[A]:ARG:NH1	4:B:577:HOH:O	2.40	0.51
1:A:25[B]:ILE:HD13	1:A:28:ARG:NH1	2.25	0.51
1:A:22:ARG:NH2	4:A:508:HOH:O	2.45	0.48
1:A:82:HIS:HB3	1:A:101:ARG:HB2	1.94	0.48
1:B:97:ARG:NH1	4:B:407:HOH:O	2.38	0.46
1:B:58[B]:GLN:OE1	4:B:404:HOH:O	2.21	0.45
1:B:1:ARG:N	4:B:410:HOH:O	2.48	0.44
1:A:1[A]:ARG:HB2	1:A:10:TRP:CD1	2.52	0.44
1:B:58[B]:GLN:NE2	4:B:414:HOH:O	2.55	0.40

All (6) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:302:CIT:C6	3:B:303:FE:FE[5_455]	1.57	0.63
1:A:34:ARG:NH2	2:A:302:CIT:O7[5_455]	1.96	0.24
2:A:302:CIT:C4	3:A:304:FE:FE[5_455]	1.96	0.24
2:A:302:CIT:C6	3:A:304:FE:FE[5_455]	1.97	0.23
2:B:302:CIT:C3	3:B:303:FE:FE[5_455]	2.00	0.20
2:A:302:CIT:C3	3:A:304:FE:FE[5_455]	2.17	0.03

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	148/134 (110%)	145 (98%)	3 (2%)	0	100	100
1	B	141/134 (105%)	139 (99%)	2 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	289/268 (108%)	284 (98%)	5 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	139/123 (113%)	136 (98%)	3 (2%)	64	28
1	B	132/123 (107%)	128 (97%)	4 (3%)	53	16
All	All	271/246 (110%)	264 (97%)	7 (3%)	75	21

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1[A]	ARG
1	A	1[B]	ARG
1	A	22	ARG
1	B	73[A]	ARG
1	B	73[B]	ARG
1	B	128[A]	HIS
1	B	128[B]	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	CIT	A	301	-	12,12,12	0.90	0	17,17,17	1.66	1 (5%)
2	CIT	A	302	3	12,12,12	0.80	0	17,17,17	2.38	4 (23%)
2	CIT	A	303	-	12,12,12	0.89	0	17,17,17	1.78	1 (5%)
2	CIT	B	301	-	12,12,12	0.95	0	17,17,17	1.80	5 (29%)
2	CIT	B	302	3	12,12,12	0.91	0	17,17,17	2.09	5 (29%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CIT	A	301	-	-	0/16/16/16	0/0/0/0
2	CIT	A	302	3	-	0/16/16/16	0/0/0/0
2	CIT	A	303	-	-	0/16/16/16	0/0/0/0
2	CIT	B	301	-	-	0/16/16/16	0/0/0/0
2	CIT	B	302	3	-	0/16/16/16	0/0/0/0

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	302	CIT	O6-C6-C3	7.57	123.90	112.89
2	B	302	CIT	O6-C6-C3	6.09	121.75	112.89
2	A	303	CIT	O6-C6-C3	5.41	120.76	112.89

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	301	CIT	O6-C6-C3	5.06	120.24	112.89
2	B	301	CIT	O6-C6-C3	4.29	119.12	112.89
2	A	302	CIT	O5-C6-C3	3.16	117.85	122.20
2	A	302	CIT	C3-C2-C1	3.10	121.27	113.77
2	B	301	CIT	C4-C3-C6	2.75	103.75	110.12
2	B	301	CIT	C2-C3-C6	2.73	116.45	110.12
2	A	302	CIT	O7-C3-C4	2.69	103.80	109.22
2	B	302	CIT	O7-C3-C6	2.56	105.27	108.95
2	B	301	CIT	O2-C1-C2	2.40	123.09	114.63
2	B	302	CIT	O2-C1-O1	2.33	117.37	123.30
2	B	302	CIT	O4-C5-O3	2.25	117.56	123.30
2	B	301	CIT	O2-C1-O1	2.17	117.78	123.30
2	B	302	CIT	O5-C6-C3	2.15	119.23	122.20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	134/134 (100%)	0.16	3 (2%)	59	66	14, 20, 37, 52	0
1	B	133/134 (99%)	-0.04	2 (1%)	70	77	15, 22, 35, 45	0
All	All	267/268 (99%)	0.06	5 (1%)	65	71	14, 21, 36, 52	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	75[A]	ARG	2.7
1	A	0	MET	2.6
1	A	85	LEU	2.5
1	B	133	ILE	2.3
1	A	101	ARG	2.1

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å²)	Q<0.9
2	CIT	A	303	13/13	0.25	13.75	28,47,51,53	0
2	CIT	B	302	13/13	0.20	6.37	20,28,34,38	13
2	CIT	A	302	13/13	0.21	3.11	11,18,30,32	13
2	CIT	A	301	13/13	0.14	2.74	23,27,34,36	0
2	CIT	B	301	13/13	0.08	-0.68	20,24,30,32	0
3	FE	B	303	1/1	0.05	-1.62	24,24,24,24	1
3	FE	A	304	1/1	0.06	-1.70	17,17,17,17	1

6.5 Other polymers (i)

There are no such residues in this entry.

2.5.   Structure of ECP/H15A mutant at 1.47 Å (PDB ID: 4OWZ)

WORLDWIDE wwPDB
PROTEIN DATA BANK
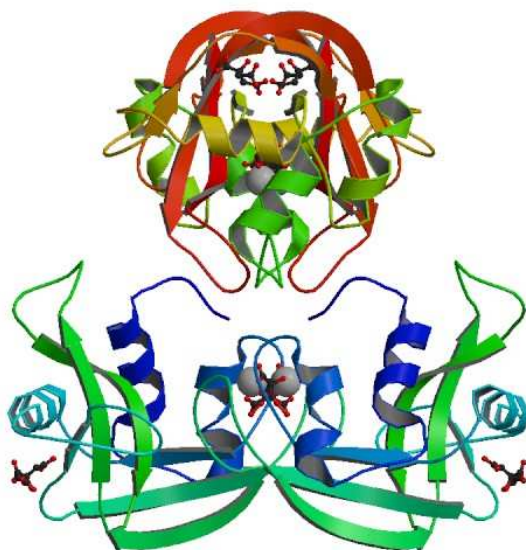
Full wwPDB X-ray Structure Validation Report (i)

Feb 26, 2014 – 02:04 PM EST

PDB ID : 4OWZ
Title : Structure of ECP/H15A mutant.
Authors : Blanco, J.A.; Salazar, V.A.; Boix, E.; Moussaoui, M.
Deposited on : 2014-02-04
Resolution : 1.47 Å(reported)

1 Overall quality at a glance (i)

The reported resolution of this entry is 1.47 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
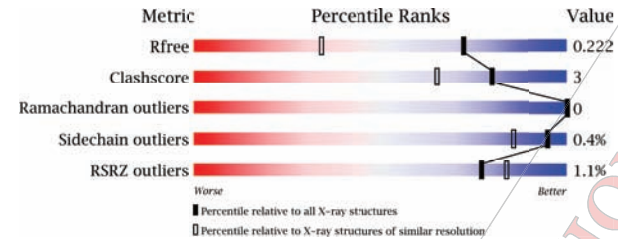
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	2222 (1.50-1.46)
Clashscore	79885	2555 (1.50-1.46)
Ramachandran outliers	78287	2496 (1.50-1.46)
Sidechain outliers	78261	2494 (1.50-1.46)
RSRZ outliers	66119	2223 (1.50-1.46)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	134	
1	B	134	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	CIT	A	301	-	X
2	CIT	A	302	-	X
2	CIT	B	302	-	X

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 2798 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Eosinophil cationic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	134	Total 1185	C 733	N 251	O 189	S 12	0	15	0
1	B	134	Total 1148	C 710	N 240	O 188	S 10	1	8	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P12724
A	15	ALA	HIS	engineered mutation	UNP P12724
A	97	ARG	THR	variant	UNP P12724
B	0	MET	-	initiating methionine	UNP P12724
B	15	ALA	HIS	engineered mutation	UNP P12724
B	97	ARG	THR	variant	UNP P12724

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 13	C 6	O 7	0	0
2	A	1	Total 13	C 6	O 7	0	0
2	B	1	Total 13	C 6	O 7	0	0
2	B	1	Total 13	C 6	O 7	0	0

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Fe 1	0	0
3	A	1	Total 1	Fe 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	229	Total 230	O 230	0	1
4	B	181	Total 181	O 181	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Eosinophil cationic protein

Chain A:

- Molecule 1: Eosinophil cationic protein

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, α, β, γ	62.54Å 62.54Å 175.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.45 – 1.47 62.54 – 1.47	Depositor EDS
% Data completeness (in resolution range)	100.0 (25.45-1.47) 96.7 (62.54-1.47)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.77 (at 1.47Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.8.4_1496)	Depositor
R, R_{free}	0.206 , 0.223 0.204 , 0.222	Depositor DCC
R_{free} test set	2933 reflections (5.04%)	DCC
Wilson B-factor (Å^2)	20.9	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 43.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< \lvert L \rvert > = 0.48, < L^2 > = 0.31$	Xtriage
Outliers	0 of 60217 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2798	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.69% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: FE, CIT

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z| >2$	RMSZ	#$|Z| >2$
1	A	0.40	1/870 (0.1%)	0.60	6/985 (0.6%)
1	B	0.34	0/925	0.55	3/1046 (0.3%)
All	All	0.37	1/1795 (0.1%)	0.58	9/2031 (0.4%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	10	TRP	CE3-CZ3	2.07	1.42	1.38

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	34	ARG	NE-CZ-NH2	-5.61	117.49	120.30
1	A	34	ARG	NE-CZ-NH1	4.38	122.49	120.30
1	A	107	TYR	CA-CB-CG	2.87	118.84	113.40
1	B	34	ARG	NE-CZ-NH2	-2.73	118.94	120.30
1	B	107	TYR	CA-CB-CG	2.41	117.98	113.40
1	A	61	ARG	NE-CZ-NH2	-2.41	119.10	120.30
1	A	61	ARG	NE-CZ-NH1	2.26	121.43	120.30
1	B	115	ASP	CB-CG-OD2	2.10	120.19	118.30
1	A	112	ASP	CB-CG-OD1	2.05	120.14	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2   Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens

added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1185	0	1174	9	0
1	B	1148	0	1138	4	0
2	A	26	0	9	3	1
2	B	26	0	9	4	2
3	A	1	0	0	0	1
3	B	1	0	0	0	2
4	A	230	0	0	7	1
4	B	181	0	0	3	0
All	All	2798	0	2330	16	4

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (16) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:A:302:CIT:O7	4:A:401:HOH:O	2.02	0.76
1:A:39[A]:ASN:OD1	4:A:402:HOH:O	2.06	0.73
1:A:101[B]:ARG:NH1	4:A:405:HOH:O	2.35	0.58
1:B:34:ARG:HH22	2:B:302:CIT:H21	1.70	0.55
1:A:34:ARG:HH22	2:A:302:CIT:C5	2.20	0.54
2:B:302:CIT:O4	4:B:401:HOH:O	2.17	0.54
1:A:53:ASN:ND2	4:A:404:HOH:O	2.35	0.53
1:A:101[B]:ARG:NH2	4:A:534:HOH:O	2.37	0.53
1:B:34:ARG:HH22	2:B:302:CIT:C2	2.22	0.52
1:A:117:ARG:NH2	4:A:613:HOH:O	2.44	0.49
1:B:123:PRO:HG2	4:B:520:HOH:O	2.15	0.47
1:A:34:ARG:HH22	2:A:302:CIT:C4	2.31	0.43
1:A:32[A]:ASN:ND2	4:A:565:HOH:O	2.39	0.42
1:B:75:ARG:HD3	4:B:532:HOH:O	2.19	0.42
1:A:79:PRO:HG3	1:A:104[A]:ARG:NH1	2.35	0.41
2:B:302:CIT:O7	2:B:302:CIT:O3	2.39	0.41

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:A:302:CIT:O7	3:A:303:FE:FE[5_555]	1.66	0.54
2:B:302:CIT:C6	3:B:303:FE:FE[5_555]	1.82	0.38

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
4:A:403:HOH:O	4:A:443:HOH:O[5_555]	1.95	0.25
2:B:302:CIT:C3	3:B:303:FE:FE[5_555]	2.06	0.14

5.3   Torsion angles

5.3.1   Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	149/134 (111%)	146 (98%)	3 (2%)	0	100	100
1	B	141/134 (105%)	138 (98%)	3 (2%)	0	100	100
All	All	290/268 (108%)	284 (98%)	6 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2   Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	139/122 (114%)	137 (99%)	2 (1%)	78	50
1	B	131/122 (107%)	131 (100%)	0	100	100
All	All	270/244 (111%)	268 (99%)	2 (1%)	95	75

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	121[A]	ARG
1	A	121[B]	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3   RNA ⓘ

There are no RNA chains in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6   Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	CIT	A	301	-	12,12,12	0.88	0	17,17,17	1.64	2 (11%)
2	CIT	A	302	3	12,12,12	1.01	0	17,17,17	2.24	9 (52%)
2	CIT	B	301	-	12,12,12	1.10	0	17,17,17	2.00	6 (35%)
2	CIT	B	302	3	12,12,12	1.21	1 (8%)	17,17,17	2.10	4 (23%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CIT	A	301	-	-	0/16/16/16	0/0/0/0
2	CIT	A	302	3	-	0/16/16/16	0/0/0/0
2	CIT	B	301	-	-	0/16/16/16	0/0/0/0
2	CIT	B	302	3	-	0/16/16/16	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	302	CIT	C2-C3	2.06	1.50	1.53

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	302	CIT	O6-C6-C3	5.53	120.93	112.89
2	A	301	CIT	O6-C6-C3	5.07	120.27	112.89
2	A	302	CIT	O6-C6-C3	4.07	118.81	112.89
2	B	302	CIT	O7-C3-C6	3.76	114.37	108.95
2	B	301	CIT	O6-C6-C3	3.59	118.12	112.89
2	B	302	CIT	C3-C4-C5	3.55	105.16	113.77
2	A	302	CIT	O4-C5-C4	3.40	126.61	114.63
2	B	301	CIT	O4-C5-O3	3.37	114.71	123.30
2	B	301	CIT	C3-C4-C5	3.33	121.84	113.77
2	A	302	CIT	O2-C1-O1	3.15	115.30	123.30
2	A	302	CIT	O2-C1-C2	3.14	125.70	114.63
2	B	301	CIT	C3-C2-C1	3.09	106.29	113.77
2	A	302	CIT	O3-C5-C4	2.73	114.11	122.74
2	A	302	CIT	C4-C3-C6	2.59	104.13	110.12
2	B	302	CIT	O6-C6-O5	2.22	116.77	123.76
2	A	302	CIT	O6-C6-O5	2.22	116.77	123.76
2	B	301	CIT	C2-C3-C6	2.16	105.13	110.12
2	A	302	CIT	O7-C3-C4	2.11	113.47	109.22
2	A	302	CIT	O7-C3-C6	2.10	105.93	108.95
2	B	301	CIT	O6-C6-O5	2.08	117.23	123.76
2	A	301	CIT	O2-C1-O1	2.02	118.16	123.30

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7   Other polymers ⓘ

There are no such residues in this entry.

5.8   Polymer linkage issues

There are no chain breaks in this entry.

6   Fit of model and data ⓘ

6.1   Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	134/134 (100%)	0.05	2 (1%)	70	77	15, 21, 34, 45	1 (0%)
1	B	134/134 (100%)	-0.03	1 (0%)	84	90	16, 25, 39, 47	3 (2%)
All	All	268/268 (100%)	0.01	3 (1%)	77	84	15, 23, 37, 47	4 (1%)

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	0	MET	2.4
1	A	128[A]	HIS	2.3
1	A	85	LEU	2.1

6.2   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3   Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4   Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å²)	Q<0.9
2	CIT	B	302	13/13	0.22	5.46	20,28,32,33	13
2	CIT	A	301	13/13	0.12	3.49	22,26,33,35	0
2	CIT	A	302	13/13	0.26	3.04	15,18,28,29	13
2	CIT	B	301	13/13	0.12	1.10	21,28,35,41	0
3	FE	A	303	1/1	0.07	-2.19	16,16,16,16	1
3	FE	B	303	1/1	0.05	-2.70	26,26,26,26	1

6.5 Other polymers ⓘ

There are no such residues in this entry.

2.6. Structure of ECP/H128N mutant with sulphate anions at 1.34 Å (PDB ID: 4X08)

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Nov 24, 2014 – 05:22 AM EST

PDB ID : 4X08
Title : Structure of H128N/ECP mutant in complex with sulphate anions at 1.34 Angstroms.
Authors : Blanco, J.A.; Garcia, J.M.; Salazar, V.A.; Sanchez, D.; Moussauoi, M.
Deposited on : 2014-11-21
Resolution : 1.34 Å(reported)

The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable24103
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable24103

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 1.34 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
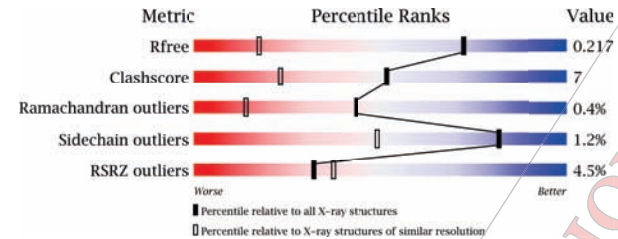
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1196 (1.38-1.30)
Clashscore	79885	1305 (1.38-1.30)
Ramachandran outliers	78287	1262 (1.38-1.30)
Sidechain outliers	78261	1262 (1.38-1.30)
RSRZ outliers	66119	1196 (1.38-1.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	134	
1	B	134	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	A	201	-	X
2	SO4	A	202[A]	-	X
2	SO4	A	202[B]	-	X
2	SO4	A	204	-	X
2	SO4	A	205	-	X
2	SO4	A	206	-	X

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Type	Chain	Res	Geometry	Electron density
2	SO4	A	207	-	X
2	SO4	B	201	-	X
2	SO4	B	202	-	X
2	SO4	B	208	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2837 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Eosinophil cationic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	133	Total 1156	C 717	N 236	O 194	S 9	0	13	0
1	B	134	Total 1169	C 728	N 238	O 190	S 13	0	14	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P12724
A	97	ARG	THR	variant	UNP P12724
A	128	ASN	HIS	engineered mutation	UNP P12724
B	0	MET	-	initiating methionine	UNP P12724
B	97	ARG	THR	variant	UNP P12724
B	128	ASN	HIS	engineered mutation	UNP P12724

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 10	O 8	S 2	0	1
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 10	O 8	S 2	0	1
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0
2	B	1	Total 5	O 4	S 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	210	Total 210	O 210	0	0
3	B	212	Total 212	O 212	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Eosinophil cationic protein

Chain A:

- Molecule 1: Eosinophil cationic protein

Chain B:

XCII

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	92.87Å 51.10Å 55.60Å 90.00° 111.09° 90.00°	Depositor
Resolution (Å)	41.37 – 1.34 41.37 – 1.34	Depositor EDS
% Data completeness (in resolution range)	98.8 (41.37-1.34) 98.8 (41.37-1.34)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.90 (at 1.34Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.184 , 0.218 0.184 , 0.217	Depositor DCC
R_{free} test set	2623 reflections (4.86%)	DCC
Wilson B-factor (Å^2)	12.3	Xtriage
Anisotropy	0.059	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 38.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Outliers	0 of 54033 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2837	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.09% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.34	0/1239	0.53	0/1688
1	B	0.35	0/1250	0.55	0/1698
All	All	0.34	0/2489	0.54	0/3386

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1156	0	1124	22	1
1	B	1169	0	1156	11	0
2	A	40	0	0	6	1
2	B	50	0	0	4	0
3	A	210	0	0	12	1
3	B	212	0	0	6	0
All	All	2837	0	2280	34	2

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 7.

All (34) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:203:SO4:O3	3:B:301:HOH:O	1.96	0.83
1:A:39:ASN:O	3:A:301:HOH:O	1.95	0.83
1:B:36:ARG:NH1	3:B:302:HOH:O	2.10	0.82
1:A:66:ARG:NH2	2:A:205:SO4:O3	2.12	0.79
1:A:85:LEU:N	3:A:301:HOH:O	2.15	0.79
2:A:206:SO4:O2	3:A:448:HOH:O	2.04	0.74
1:B:82[B]:HIS:ND1	3:B:450:HOH:O	2.19	0.71
1:A:122[A]:TYR:OH	3:A:441:HOH:O	2.09	0.70
1:B:128:ASN:OD1	3:B:493:HOH:O	2.12	0.67
1:A:97:ARG:NH2	3:A:306:HOH:O	2.27	0.67
1:B:37[A]:CYS:SG	3:B:505:HOH:O	2.52	0.66
1:A:101[B]:ARG:NH2	3:A:308:HOH:O	2.31	0.63
1:B:42:THR:OG1	1:B:82[B]:HIS:HD2	1.86	0.59
1:A:86[A]:ILE:HG12	1:A:97:ARG:O	2.03	0.58
2:B:202:SO4:O4	3:B:303:HOH:O	2.17	0.58
1:A:19[B]:ASN:OD1	1:B:33:TYR:OH	2.16	0.57
1:A:122[A]:TYR:OH	1:B:49:ALA:HB2	2.10	0.52
1:B:86[A]:ILE:HD11	1:B:99:ALA:HB2	1.94	0.49
1:A:104:ARG:NH1	2:A:203:SO4:O3	2.46	0.48
1:A:77:ARG:NE	2:A:203:SO4:O4	2.27	0.48
1:A:101[B]:ARG:HG3	3:A:463:HOH:O	2.13	0.47
1:A:28:ARG:HG3	3:A:321:HOH:O	2.15	0.47
1:A:75:ARG:HB2	2:A:206:SO4:O1	2.17	0.45
1:A:68:LEU:HD13	3:A:485:HOH:O	2.17	0.43
1:A:98:TYR:HE2	3:A:301:HOH:O	2.01	0.43
1:B:77:ARG:NH1	2:B:205:SO4:O2	2.48	0.42
1:A:53[B]:ASN:OD1	3:A:302:HOH:O	2.22	0.42
1:A:82[A]:HIS:NE2	1:A:84:ASP:OD1	2.53	0.42
1:A:92:ASN:HB3	3:A:359:HOH:O	2.20	0.42
1:A:66:ARG:HB3	2:A:205:SO4:O4	2.20	0.41
1:B:57:ASN:HB3	2:B:206[A]:SO4:O3	2.20	0.41
1:A:18:LEU:HD12	1:A:18:LEU:HA	1.91	0.41
1:B:82[A]:HIS:NE2	1:B:84:ASP:OD1	2.52	0.41
1:A:2:PRO:HA	1:A:3:PRO:HD3	1.86	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:32:ASN:ND2	2:A:205:SO4:O1[4_556]	2.04	0.16
3:A:321:HOH:O	3:A:413:HOH:O[4_556]	2.17	0.03

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	144/134 (108%)	141 (98%)	2 (1%)	1 (1%)	30	5
1	B	146/134 (109%)	145 (99%)	1 (1%)	0	100	100
All	All	290/268 (108%)	286 (99%)	3 (1%)	1 (0%)	43	18

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	91	GLN

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	135/123 (110%)	133 (98%)	2 (2%)	76	40
1	B	137/123 (111%)	133 (97%)	4 (3%)	55	14
All	All	272/246 (111%)	266 (98%)	6 (2%)	82	23

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	122[A]	TYR
1	A	122[B]	TYR
1	B	37[A]	CYS
1	B	37[B]	CYS
1	B	76[A]	PHE
1	B	76[B]	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3   RNA (i)

There are no RNA chains in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates (i)

There are no carbohydrates in this entry.

5.6   Ligand geometry (i)

18 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	SO4	A	201	-	4,4,4	0.19	0	6,6,6	0.11	0
2	SO4	A	202[A]	-	4,4,4	0.19	0	6,6,6	0.07	0
2	SO4	A	202[B]	-	4,4,4	0.22	0	6,6,6	0.09	0
2	SO4	A	203	-	4,4,4	0.22	0	6,6,6	0.09	0
2	SO4	A	204	-	4,4,4	0.21	0	6,6,6	0.07	0
2	SO4	A	205	-	4,4,4	0.45	0	6,6,6	0.29	0
2	SO4	A	206	-	4,4,4	0.34	0	6,6,6	0.26	0
2	SO4	A	207	-	4,4,4	0.21	0	6,6,6	0.12	0
2	SO4	B	201	-	4,4,4	0.20	0	6,6,6	0.09	0
2	SO4	B	202	-	4,4,4	0.16	0	6,6,6	0.23	0
2	SO4	B	203	-	4,4,4	0.28	0	6,6,6	0.13	0
2	SO4	B	204	-	4,4,4	0.31	0	6,6,6	0.40	0
2	SO4	B	205	-	4,4,4	0.17	0	6,6,6	0.07	0
2	SO4	B	206[A]	-	4,4,4	0.14	0	6,6,6	0.09	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	SO4	B	206[B]	-	4,4,4	0.11	0	6,6,6	0.11	0
2	SO4	B	207	-	4,4,4	0.18	0	6,6,6	0.14	0
2	SO4	B	208	-	4,4,4	0.30	0	6,6,6	0.19	0
2	SO4	B	209	-	4,4,4	0.24	0	6,6,6	0.15	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SO4	A	201	-	-	0/0/0/0	0/0/0/0
2	SO4	A	202[A]	-	-	0/0/0/0	0/0/0/0
2	SO4	A	202[B]	-	-	0/0/0/0	0/0/0/0
2	SO4	A	203	-	-	0/0/0/0	0/0/0/0
2	SO4	A	204	-	-	0/0/0/0	0/0/0/0
2	SO4	A	205	-	-	0/0/0/0	0/0/0/0
2	SO4	A	206	-	-	0/0/0/0	0/0/0/0
2	SO4	A	207	-	-	0/0/0/0	0/0/0/0
2	SO4	B	201	-	-	0/0/0/0	0/0/0/0
2	SO4	B	202	-	-	0/0/0/0	0/0/0/0
2	SO4	B	203	-	-	0/0/0/0	0/0/0/0
2	SO4	B	204	-	-	0/0/0/0	0/0/0/0
2	SO4	B	205	-	-	0/0/0/0	0/0/0/0
2	SO4	B	206[A]	-	-	0/0/0/0	0/0/0/0
2	SO4	B	206[B]	-	-	0/0/0/0	0/0/0/0
2	SO4	B	207	-	-	0/0/0/0	0/0/0/0
2	SO4	B	208	-	-	0/0/0/0	0/0/0/0
2	SO4	B	209	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7   Other polymers (i)

There are no such residues in this entry.

5.8　Polymer linkage issues

There are no chain breaks in this entry.

6　Fit of model and data ⓘ

6.1　Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	133/134 (99%)	0.28	6 (4%)	32	37	8, 13, 25, 48	2 (1%)
1	B	134/134 (100%)	0.26	6 (4%)	32	37	8, 13, 23, 45	0
All	All	267/268 (99%)	0.27	12 (4%)	32	37	8, 13, 24, 48	2 (0%)

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	35	TRP	5.3
1	B	0	MET	5.1
1	A	91	GLN	4.4
1	A	133	ILE	3.7
1	B	76[A]	PHE	3.7
1	A	89	GLY	3.4
1	A	86[A]	ILE	2.6
1	A	90	ALA	2.5
1	B	133	ILE	2.4
1	B	91	GLN	2.4
1	B	86[A]	ILE	2.2
1	B	95[A]	ASN	2.1

6.2　Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3　Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4　Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors($Å^2$)	Q<0.9
2	SO4	A	202[B]	5/5	0.19	12.41	30,33,36,38	5
2	SO4	A	202[A]	5/5	0.19	9.92	29,31,34,38	5
2	SO4	B	202	5/5	0.20	8.62	20,37,55,56	0
2	SO4	A	205	5/5	0.21	8.55	11,20,28,30	5
2	SO4	A	204	5/5	0.20	7.41	14,16,19,21	5
2	SO4	B	208	5/5	0.19	5.82	18,24,26,30	5
2	SO4	A	206	5/5	0.18	3.27	16,23,30,35	5
2	SO4	A	201	5/5	0.20	3.22	43,45,52,56	0
2	SO4	A	207	5/5	0.12	2.45	19,26,29,33	5
2	SO4	B	201	5/5	0.16	2.15	31,36,39,40	5
2	SO4	B	205	5/5	0.21	1.17	42,47,52,53	0
2	SO4	B	206[B]	5/5	0.15	0.81	12,14,17,19	5
2	SO4	B	203	5/5	0.13	0.74	17,21,29,32	5
2	SO4	B	206[A]	5/5	0.15	0.71	12,13,15,19	5
2	SO4	B	207	5/5	0.10	0.54	18,19,24,25	5
2	SO4	B	204	5/5	0.12	-0.05	16,17,23,25	0
2	SO4	B	209	5/5	0.12	-0.34	16,20,24,25	5
2	SO4	A	203	5/5	0.10	-0.71	21,25,28,31	5

6.5　Other polymers ⓘ

There are no such residues in this entry.

XCVII

2.7. Structure of human RNase 6 in complex with sulphate anions at 1.72 Å (PDB ID: 4X09)

WORLDWIDE

wwPDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Dec 11, 2014 – 10:00 AM EST

PDB ID   :   4X09
Title   :   Structure of human RNase 6 in complex with sulphate anions
Authors   :   Arranz, J.; Blanco, J.A.; Pulido, D.; Moussaoui, M.; Boix, E.
Deposited on   :   2014-11-21
Resolution   :   1.72 Å(reported)

The following versions of software and data (see references) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable24103
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable24103

1   Overall quality at a glance ⓘ

The reported resolution of this entry is 1.72 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
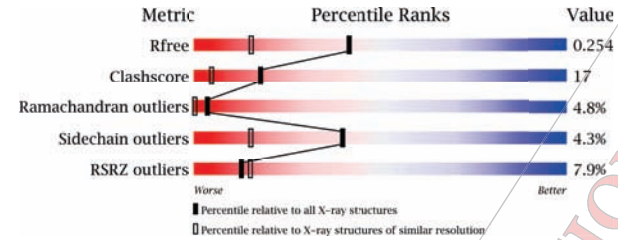
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	2979 (1.74-1.70)
Clashscore	79885	3506 (1.74-1.70)
Ramachandran outliers	78287	3449 (1.74-1.70)
Sidechain outliers	78261	3449 (1.74-1.70)
RSRZ outliers	66119	2979 (1.74-1.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

| Mol	Chain	Length	Quality of chain
1	A	127	

2　Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 1223 atoms, of which 0 are hydrogen and 0 are deuterium.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Ribonuclease K6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	127	Total	C	N	O	S	0	4	0
			1049	661	197	181	10			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O$_4$S).
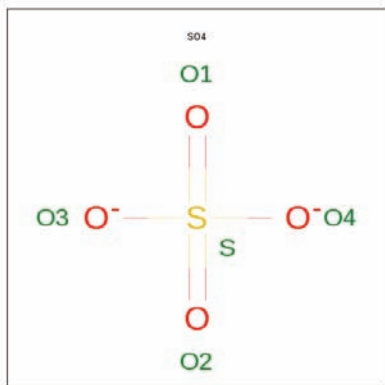
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	164	Total	O	0	0
			164	164		

3　Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Ribonuclease K6

Chain A:
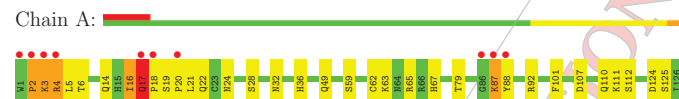
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	27.73Å 38.86Å 97.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.98 – 1.72 48.98 – 1.72	Depositor EDS
% Data completeness (in resolution range)	99.2 (48.98-1.72) 99.2 (48.98-1.72)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.16 (at 1.72Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.208 , 0.255 0.209 , 0.254	Depositor DCC
R_{free} test set	579 reflections (4.94%)	DCC
Wilson B-factor (Å^2)	24.7	Xtriage
Anisotropy	0.081	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 47.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< \|L\| > = 0.45, < L^2 > = 0.28$	Xtriage
Outliers	1 of 11717 reflections (0.009%)	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	1223	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.27% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#|Z|>5$	RMSZ	$\#|Z|>5$
1	A	0.41	0/1096	0.62	0/1484

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	16	ILE	Peptide
1	A	17	GLN	Peptide

5.2 Close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

CI

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1049	0	1028	34	0
2	A	10	0	0	2	0
3	A	164	0	0	21	0
All	All	1223	0	1028	36	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 17.

All (36) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:17:GLN:HG3	1:A:18:PRO:HD3	1.52	0.92
1:A:87:LYS:HD2	1:A:88:TYR:H	1.43	0.82
1:A:65:ARG:NH2	3:A:301:HOH:O	2.14	0.80
1:A:36:HIS:NE2	3:A:303:HOH:O	2.16	0.77
1:A:4[B]:ARG:NH1	3:A:306:HOH:O	2.26	0.68
1:A:127:LEU:O	3:A:458:HOH:O	2.11	0.68
1:A:2:PRO:HD2	1:A:6:THR:HA	1.76	0.67
1:A:107:ASP:OD1	3:A:302:HOH:O	2.14	0.65
1:A:101:PHE:HB2	1:A:127:LEU:HD12	1.79	0.64
1:A:62:CYS:O	1:A:65:ARG:NH1	2.31	0.63
1:A:87:LYS:HD2	1:A:88:TYR:N	2.13	0.61
1:A:32:ASN:HB3	3:A:382:HOH:O	1.99	0.61
1:A:92[A]:ARG:NH2	3:A:456:HOH:O	2.35	0.59
1:A:17:GLN:HG3	1:A:18:PRO:CD	2.31	0.59
1:A:14:GLN:NE2	3:A:310:HOH:O	2.36	0.58
1:A:22:GLN:NE2	1:A:24:ASN:HB2	2.20	0.56
1:A:16:ILE:HG22	1:A:17:GLN:HB3	1.87	0.56
1:A:4[B]:ARG:HG2	3:A:381:HOH:O	2.05	0.56
1:A:4[B]:ARG:NH2	3:A:311:HOH:O	2.39	0.56
1:A:19:SER:N	3:A:308:HOH:O	2.33	0.55
1:A:79:THR:HB	3:A:454:HOH:O	2.06	0.55
1:A:111[B]:LYS:HD2	1:A:112:SER:N	2.23	0.54
1:A:2:PRO:HB2	1:A:5:LEU:O	2.09	0.53
1:A:110:GLN:NE2	3:A:448:HOH:O	2.41	0.52
1:A:4[A]:ARG:NH2	3:A:313:HOH:O	2.43	0.51
1:A:3:LYS:O	1:A:5:LEU:N	2.44	0.50
1:A:24:ASN:ND2	3:A:456:HOH:O	2.47	0.48
1:A:18:PRO:HD2	3:A:442:HOH:O	2.14	0.46
2:A:201:SO4:O4	3:A:304:HOH:O	2.21	0.46
1:A:3:LYS:C	1:A:5:LEU:H	2.19	0.45
1:A:49:GLN:HG3	3:A:395:HOH:O	2.16	0.45
1:A:87:LYS:NZ	3:A:437:HOH:O	2.50	0.44
1:A:63:LYS:HE2	1:A:124:ASP:O	2.20	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:16:ILE:O	1:A:17:GLN:HG2	2.20	0.41
1:A:101:PHE:HD1	3:A:458:HOH:O	2.03	0.41
2:A:202:SO4:O1	3:A:305:HOH:O	2.21	0.41

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	129/127 (102%)	119 (92%)	3 (2%)	7 (5%)	3	0

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	17	GLN
1	A	20	PRO
1	A	87	LYS
1	A	3	LYS
1	A	2	PRO
1	A	4[A]	ARG
1	A	4[B]	ARG

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	121/117 (103%)	115 (95%)	6 (5%)	34	11

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	21	LEU
1	A	28	SER
1	A	59[A]	SER
1	A	59[B]	SER
1	A	67	HIS
1	A	125	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	35	GLN
1	A	49	GLN

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	SO4	A	201	-	4,4,4	0.28	0	6,6,6	0.17	0
2	SO4	A	202	-	4,4,4	0.19	0	6,6,6	0.14	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SO4	A	201	-	-	0/0/0/0	0/0/0/0
2	SO4	A	202	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6　Fit of model and data ⓘ

6.1　Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	127/127 (100%)	0.29	10 (7%)	13	15	16, 26, 42, 57	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	TRP	9.7
1	A	2	PRO	5.9
1	A	87	LYS	3.9
1	A	18	PRO	3.6
1	A	86	GLY	3.6
1	A	20	PRO	3.2
1	A	3	LYS	2.7
1	A	17	GLN	2.3
1	A	4[A]	ARG	2.1
1	A	88	TYR	2.1

6.2　Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3　Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4　Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å²)	Q<0.9
2	SO4	A	202	5/5	0.15	1.41	36,39,51,55	0
2	SO4	A	201	5/5	0.08	-0.76	32,34,35,46	0

6.5　Other polymers ⓘ

There are no such residues in this entry.