UAB
Universitat Autònoma de Barcelona

Doctoral thesis

EXPLORING THE MECHANISM OF ACTION OF HUMAN ANTIMICROBIAL RIBONUCLEASES

VIVIAN ANGÉLICA SALAZAR MONTOYA

Barcelona, 2015

PAPERS RELATED TO THE THESIS

Review

Structural determinants of the eosinophil cationic protein antimicrobial activity

Ester Boix*, Vivian A. Salazar, Marc Torrent, David Pulido, M. Victòria Nogués and Mohammed Moussaoui

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

*Corresponding author
e-mail: ester.boix@uab.cat

Abstract

Antimicrobial RNases are small cationic proteins belonging to the vertebrate RNase A superfamily and endowed with a wide range of antipathogen activities. Vertebrate RNases, while sharing the active site architecture, are found to display a variety of noncatalytical biological properties, providing an excellent example of multitask proteins. The antibacterial activity of distant related RNases suggested that the family evolved from an ancestral host-defence function. The review provides a structural insight into antimicrobial RNases, taking as a reference the human RNase 3, also named eosinophil cationic protein (ECP). A particular high binding affinity against bacterial wall structures mediates the protein action. In particular, the interaction with the lipopolysaccharides at the Gram-negative outer membrane correlates with the protein antimicrobial and specific cell agglutinating activity. Although a direct mechanical action at the bacteria wall seems to be sufficient to trigger bacterial death, a potential intracellular target cannot be discarded. Indeed, the cationic clusters at the protein surface may serve both to interact with nucleic acids and cell surface heterosaccharides. Sequence determinants for ECP activity were screened by prediction tools, proteolysis and peptide synthesis. Docking results are complementing the structural analysis to delineate the protein anchoring sites for anionic targets of biological significance.

Keywords: bactericidal; glycosaminoglycans; host defence; innate immunity; nucleotides; RNase.

Introduction: antimicrobial RNases

Antimicrobial RNases belong to the vertebrate-secreted RNase family, also called RNase A or pancreatic RNase superfamily (Beintema and Kleineidam, 1998). The family gathers all the vertebrate RNases homologous to RNase A, where RNase A refers to the bovine pancreatic RNase, probably the best studied enzyme during the 20th century (Cuchillo et al., 2011). The protein was one of the first enzymes discovered (1920) and crystallized (1939). The enzyme was easily obtained from bovine pancreas at slaughterhouses, and at the beginning of the 1940s, 1 kg of purified protein was obtained and made available for researchers. Its abundance, small molecular weight and high stability converted this protein as one of the favourites of the early 20th century biochemists. It was the first enzyme whose primary structure was known, and in 1972, three Nobel prizes were awarded for their work. Pioneering sequencing provided around 40 homologous sequences of ruminants and other mammals, setting the basis for evolution studies. It was proposed that the pancreatic RNase would digest the huge amount of RNA in the stomach of ruminants produced by the intestinal flora. Together with a primary role in RNA digestion ascribed for RNase A, a variety of non-catalytical biological properties were described for the other family members. The quest in trying to find out the family physiological role was just starting, and despite the joined efforts, the puzzle is not fully solved. The antibacterial activity of distantly related RNases suggested that ancestral RNases were in fact host-defence RNases (Cho and Zhang, 2007; Pizzo and D'Alessio, 2007; Rosenberg, 2008b).

The RNase A family is strictly restricted to vertebrates and includes eight human members that maintain the active site residues (Cho et al., 2005), referred to as the 'canonical RNases' (Sorrentino, 2010) (Figure 1, Table 1). Although the primary sequence identity in some cases is as low as 30%, they all share a common overall three dimensional structure with three to four conserved disulphide bonds and equivalent catalytic residues. The conserved catalytic triad is comprised of two histidines (H12 and H119 in RNase A) that participate by a general acid-basic mechanism in the cleavage of the RNA 3'5' phosphodiester bond and one lysine (K41 in RNase A) that stabilises the transition state intermediate. Together with the RNase catalytic activity, other biological activities have been described for some family members (Boix and Nogues, 2007; Rosenberg, 2008b). In particular, several human RNases display a toxicity for several pathogens, and a role in host defence is proposed (Table 1). However, trying to define structural determinants for the so-called 'antimicrobial RNases' is not straightforward. No unique structural identity patterns distinguish the 'antimicrobial' members from the others. As a general rule, the antimicrobial RNases would work as small cationic bullets, which are secreted locally and mostly exert a cytotoxic action on a wide range of targets.
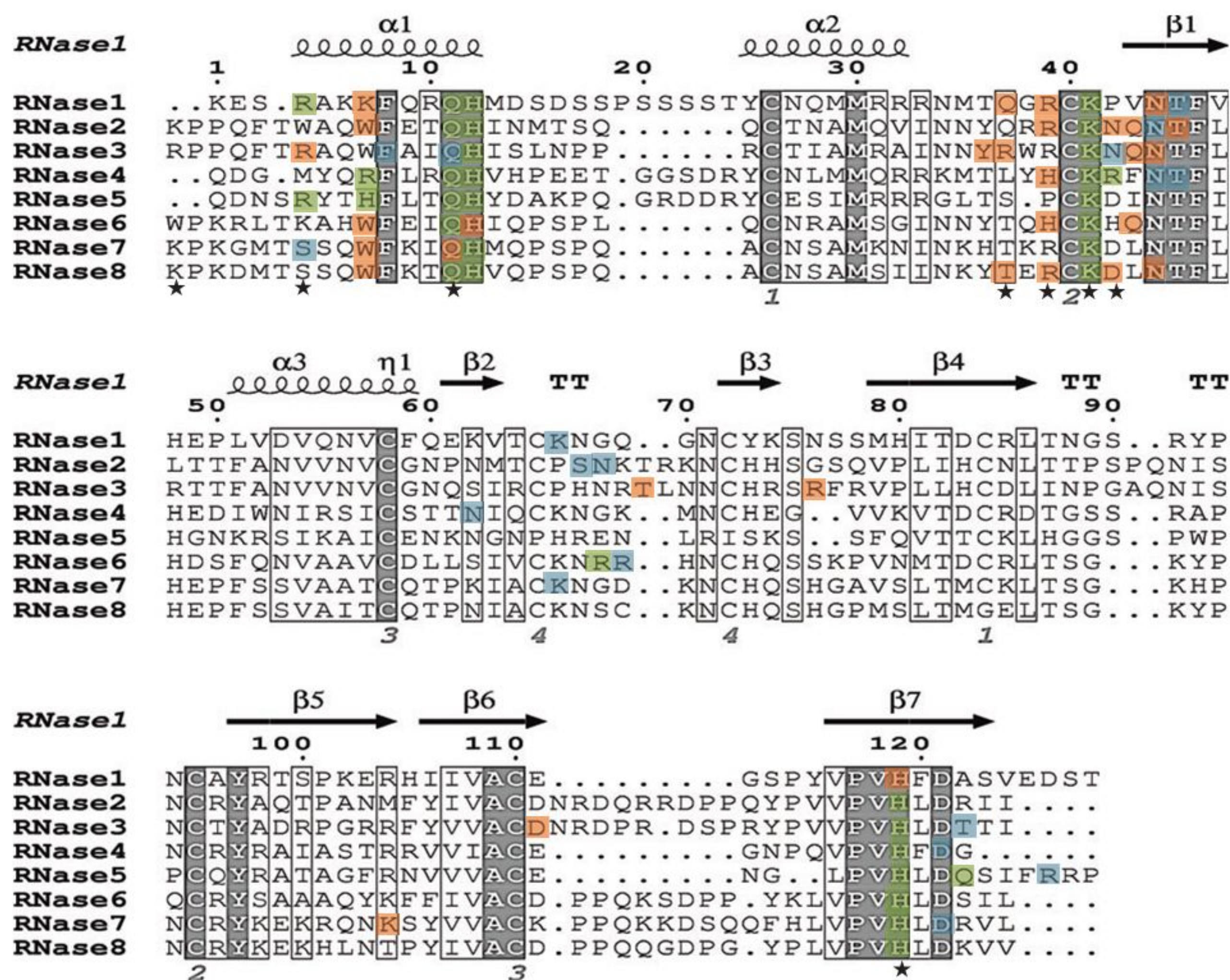
RNase1

α1   α2   β1

	1	10	20	30	40	
RNase1	..KES.RAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCKPVNTFV
RNase2	KPPQFTWAQWFETQHINMTSQ......QCTNAMQVINNYQRRCKNQNTFL
RNase3	RPPQFTRAQWFAIQHISLNPP......RCTIAMRAINNYRWRCKNQNTFL
RNase4	..QDG.MYQRFLRQHVHPEET.GGSDRYCNLMMQRRKMTLYHCKRFNTFI
RNase5	..QDNSRYTHFLTQHYDAKPQ.GRDDRYCESIMRRRGLTS.PCKDINTFI
RNase6	WPKRLTKAHWFEIQHIQPSPL......QCNRAMSGINNYTQHCKHQNTFL
RNase7	KPKGMTSSQWFKIQHMQPSPQ......ACNSAMKNINKHTKRCKDLNTFL
RNase8	KPKDMTSSQWFKTQHVQPSPQ......ACNSAMSIINKYTERCKDLNTFL

α3   η1   β2   TT   β3   β4   TT   TT

	50	60	70	80	90	
RNase1	HEPLVDVQNVCFQEKVTCKNGQ..GNCYKSNSSMHITDCRLTNGS..RYP
RNase2	LTTFANVVNVCGNPNMTCPSNKTRKNCHHSGSQVPLIHCNLTTPSPQNIS
RNase3	RTTFANVVNVCGNQSIRCPHNRTLNNCHRSRFRVPLLHCDLINPGAQNIS
RNase4	HEDIWNIRSICSTTNIQCKNGK..MNCHEG..VVKVTDCRDTGSS..RAP
RNase5	HGNKRSIKAICENKNGNPHREN..LRISKS..SFQVTTCKLHGGS..PWP
RNase6	HDSFQNVAAVCDLLSIVCKNRR..HNCHQSSKPVNMTDCRLTSG...KYP
RNase7	HEPFSSVAATCQTPKIACKNGD..KNCHQSHGAVSLTMCKLTSG...KHP
RNase8	HEPFSSVAITCQTPNIACKNSC..KNCHQSHGPMSLTMGELTSG...KYP

β5   β6   β7

	100	110	120	
RNase1	NCAYRTSPKERHIIVACE.........GSPYVPVHFDASVEDST
RNase2	NCRYAQTPANMFYIVACDNRDQRRDPPQYPVVPVHLDRII....
RNase3	NCTYADRPGRRFYVVACDNRDPR.DSPRYPVVPVHLDTTI....
RNase4	NCRYRAIASTRRVVIACE........GNPQVPVHFDG.......
RNase5	PCQYRATAGFRNVVVACE......NG..LPVHLDQSIFRRP
RNase6	QCRYSAAAQYKFFIVACD.PPQKSDPP.YKLVPVHLDSIL....
RNase7	NCRYKEKRQNKSYVVACK.PPQKKDSQQFHLVPVHLDRVL....
RNase8	NCRYKEKHLNTPYIVACD.PPQQGDPG.YPLVPVHLDKVV....

Figure 1   Primary sequence alignment of the human RNases belonging to the vertebrate-secreted RNase A family. Conserved amino acids in all sequences are boxed in black. The specific side chain-interacting residues identified in docking analysis (see Tables S1–S4) are highlighted by coloured boxes: residues interacting with nucleotide probes are in blue, residues interacting with heparin probes are in orange and common residues are in green. Residues that interact with sulphate anions in ECP crystal complex structure (4A2O.pdb) are labelled with black stars at the bottom. The alignment was performed using *Clustal W*, and the picture was drawn using the *ESPript* program (http://espript.ibcp.fr/ESPript/).

Their RNase activity participates in some of their cytotoxic power, but frequently, protein denaturation or ablation of active residues does not abolish their antipathogen activities (Pizzo and D'Alessio, 2007). Overall, the eight human RNases comply the main common properties of antimicrobial peptides, with a positive net charge that would enable them to interact electrostatically with the negatively charged phospholipids of the bacterial membrane and the lipolysaccharides or teichoic acids at the Gram-negative and -positive surface, having a high percentage of hydrophobic amino acids, and regions that can adopt amphipatic secondary structures upon membrane interaction (Boix and Nogues, 2007). However, although all the family members share common structural patterns, only some of them have been reported to display antimicrobial activity (see Table 1). Moreover, antimicrobial RNases, as do many other antimicrobial proteins, may combine membrane lytic and cytotoxic actions together with other immunomodulatory properties (Boix et al., 2008; Bystrom et al., 2011; Nakatsuji and Gallo, 2012).

A suggestive hypothesis regards antimicrobial RNases not as bactericidal proteins per se, but rather as a mere source of antimicrobial peptides. The proteins would act as a reservoir of active peptides (Pizzo and D'Alessio, 2007; Pizzo et al., 2008; Rosenberg, 2008b; Giancola et al., 2011; Sanchez et al., 2011). This proposal relies on the observation that some antimicrobial proteins retain their activity or are only active when denatured (During et al., 1999; Schroeder et al., 2011). The secreted RNases could have a primary role linked to its RNase activity and secondarily would constitute a source of antimicrobial peptides, which could intervene when secreted at the infectious focus. There are also many examples in the literature of antimicrobial proteins that are activated *in situ* by specific proteolysis (Levy, 2000). In fact, the proteolytic processing is used extensively in immunology cascade events. A local cleavage can release the active peptides in the infection area in the same manner as cecropin is released from cathelicidins expressed in neutrophils (Zanetti et al., 1995; Burton and Steel, 2009), lactoferricin is derived from the N-terminus

Table 1 Properties of human RNases from the RNase A family.

Protein name (alternative name)[a]	Main source	Reported bactericidal activity	Known natural variants[b]	Posttranslational modifications	Proposed role
RNase 1 (Human pancreatic RNase) [P07998] 128/28	Pancreas, human endothelial cells and other tissues and body fluids	None		3 N-linked glycosylation sites (Beintema et al., 1988)	RNA digestion in ruminant homologues. Unknown non-digestive physiological role for the human member (Sorrentino, 2010)
RNase 2 (EDN) [P10153] 134/27	Matrix of eosinophil large specific granules, liver, lung, spleen, neutrophils and monocytes	None	His73/Gln His129/Asn	5 potential N-glycosylation sites; C-linked mannose at Trp7 (Krieg et al., 1998; Clark et al., 2003); Tyr33 nitration (Ulrich et al., 2008)	Host defense against viral infections; chemotaxis for dendritic cells (Rosenberg, 2008a)
RNase 3 (ECP) [P12724] 133/27	Matrix of eosinophil large specific granules and also minoritary in neutrophils	High (nM–μM range) $MIC_{100}=0.3–1.5$ μM (tested against several Gram-negative and -positive species (Torrent et al., 2009b, 2011b)	Arg97/Thr (related to asthma propensity and disease-induced pathologies) (Eriksson et al., 2007a; Adu et al., 2011) Gly103Arg	3 potential N-glycosylation sites; 10 purified variants of N-linked glycosylated forms (Eriksson et al., 2007b); N-glycosylation processed upon secretion by activated eosinophils (Woschnagg et al., 2009); Tyr33 nitration (Ulrich et al., 2008)	Host defense against parasite, viral and bacterial infections; immunomodulation (Venge et al., 1999; Boix et al., 2008; Rosenberg, 2008b; Bystrom et al., 2011)
RNase 4 [P34096] 119/28	Ubiquitous, with predominance in liver and lung	None		Pyro-Glu (N-terminal Gln cyclation) (Hofsteenge et al., 1998)	Unknown. mRNA cleavage? (Hofsteenge et al., 1998)
RNase 5 (angiogenin) [P03950] 123/24	Expressed predominantly in the liver. Also detected in endothelial cells, neurons, intestinal epithelium and keratinocytes	High-moderate (low-μM range) Contradictory reports (Hooper et al., 2003; Abtin et al., 2009)	Multiple substitutions related to amyotrophic lateral sclerosis (ALS9) (Greenway et al., 2006)	Pyro-Glu (N-terminal Gln cyclation) (Strydom et al., 1985)	Angiogenesis; cell differentiation; host defense? (Moenner et al., 1994; Strydom, 1998; Hooper et al., 2003; Gao and Xu, 2008)
RNase 6 (RNase K6) [Q93091] 127/23	Strong expression in lung, followed by heart, placenta, kidney, pancreas, liver, brain and skeletal muscle. Also expressed in monocytes and neutrophils	None	Arg66/Gln	2 potential N-glycosylation sites (Rosenberg and Dyer, 1996)	Host defense?; RNA catabolism? (Dyer et al., 2004; Rosenberg and Dyer, 1996)
RNase 7 (Skin-derived antimicrobial protein 2, SAP-2) [Q9H1E1] 128/28	Expressed in various epithelial tissues including skin, respiratory tract, genito-urinary tract and, at a low level, in the gut. Expressed in liver, kidney, skeletal muscle and heart	High (nM–μM range) Broad-spectrum antimicrobial activity against many pathogenic microorganisms and remarkably potent activity (LD90<30 nM) against a vancomycin-resistant *Enterococcus faecium* (Harder and Schroder, 2002)	Ala75/Pro His88/Tyr (Clark et al., 2003)	1 potential N-glycosylation site	Host defense. Skin protection. Protection of respiratory, intestinal epithelium and urinary tract (Harder and Schroder, 2002; Zhang et al., 2003; Spencer et al., 2011; Simanski et al., 2012)

221

Table 1 (Continued)

Protein name (alternative name)[a]	Main source	Reported bactericidal activity[b]	Known natural variants[b]	Posttranslational modifications	Proposed role
RNase 8 (Placental RNase) [Q8TDE3] 127/27	Expressed in placenta	High-moderate (low-μM range) Contradictory reports. Tested against a wide bacteria spectra (LD90 from 0.1 to 1 μM) (Rudolph et al., 2006)			Placenta protection against infection? (Zhang et al., 2002; Rudolph et al., 2006) Unknown role not related to secreted RNases (Chan et al., 2012)

Information was taken, unless indicated, from UniProt Protein Knowledgebase data (UniProtKB; http://www.uniprot.org). Bactericidal activity reported in the literature is indicated and classified as low, moderate or high, when an appropriate reference is available.

[a][UniProtKB ID], Mature protein/ signal peptide lengths.
[b]Only natural variants in the mature protein reported in the UniProtKB are included.

of lactoferrin (Nibbering et al., 2001) by action of pepsin in the stomach or cryptdins are expressed in a precursor form and secreted by Paneth cells (Ayabe et al., 2002). Another example is lysozyme endowed with a specific muramidase activity together with a mechanical lytic action at the bacteria wall, which is retained when denatured (Masschalck et al., 2001). Lysozyme and RNases, as small cationic proteins, which are easy to purify, are indeed long travel companions recently recovered from the past as potential multitask effectors in innate immunity (Nakatsuji and Gallo, 2012).

In particular, in the RNase A family, we find a direct experimental evidence of a protein, the zebrafish RNase, that releases an antimicrobial peptide when processed *in vivo* by a bacterial protease (Zanfardino et al., 2010). Another family-related RNase, the leukocyte chicken RNase, can provide derived peptides that retain most of the parental protein activity (Nitto et al., 2006). RNase 3, the eosinophil cationic protein (ECP) is also offering an interesting example: the enzymatic protein proteolysis of ECP with a lysyl-endopeptidase produces a unique active segment corresponding to the N-terminus 1–38 fragment (Sanchez et al., 2011). Interestingly, human ECP shows a single Lys in all its coding sequence (Lys 38). All other Lys residues in ECP lineage diverging from a common eosinophil RNase ancestor have been lost during evolution in detriment of Arg residues, which ultimately would have conferred the protein with the highest pI amongst the family counterparts. However, lysyl-endopeptidases are only reported in bacteria species, and no evidence is available from any *in vivo* processing of the eosinophil-secreted ECP. In any case, further scanning analysis using chemical cleavage and synthetic peptides confirms that an N-terminus peptide can reproduce most of the protein antimicrobial properties (Torrent et al., 2009b; Sanchez et al., 2011). In the following section, we will try to address the structure-function relationship of antimicrobial RNases, taking ECP as a reference.

The eosinophil cationic protein (ECP) in host defence

The human RNase 3, or ECP, is an RNase specifically expressed in eosinophils and secreted during infection and inflammation processes (Venge et al., 1999; Bystrom et al., 2011). ECP levels can be detected in circulating blood (Venge et al., 1977), and its levels are routinely used to monitor asthma and other inflammatory diseases (Blanchard and Rothenberg, 2009; Bystrom et al., 2011). Eosinophil traffic and specific degranulation have been located in diverse pathological conditions, and a host-defence role for ECP at the respiratory and intestinal mucosa is attributed (Bystrom et al., 2011).

ECP, as an eosinophil granule secretion protein, is a key mediator against parasite infections, as helminths and protozoa (Tischendorf et al., 1996; Boix et al., 2008; Bystrom et al., 2011; Malik and Batra, 2012). Recent genomic studies revealed a polymorphism at the coding sequence (Arg97Thr), which correlates with asthma predisposition and parasitic infection susceptibility (Eriksson et al., 2007a). Indeed, eosinophilia and specific eosinophil tissue infiltration are directly associated

with parasite infections (Klion and Nutman, 2004; Bystrom et al., 2011; Makepeace et al., 2012). ECP toxicity was reported against helminths, as *Schistosoma mansoni* (Mclaren et al., 1984; Ackerman et al., 1985) and *Brugia malayi* (Hamann et al., 1990), and protozoa, as *Leishmania* (Elshafie et al., 2011; Singh and Batra, 2011), *Trypanosoma* (Kierszenbaum et al., 1986) and *Plasmodium* (Waters et al., 1987). Eosinophils are activated during *Plasmodium* infection (Kurtzhals et al., 1998), and the secreted protein participates in the parasite killing inside infected erythrocytes (Waters et al., 1987). *Plasmodium falciparum* infection induces eosinophilia, and ECP levels are increased during acute malaria (Kurtzhals et al., 1998). Unfortunately, the eosinophil protein is not only involved in the host defence against infection but also in the disease pathogenesis. A severe pathology is attributed to the protein neurotoxicity, which causes cerebral malaria, a major cause of children's death at an early stage. The natural variant (Arg97Thr), which reduces the protein cytotoxicity (Rubin et al., 2009), seems to have been selected in the Ghana population exposed to malaria infection, as a mechanism to avoid the protein severe neurotoxic side effects (Adu et al., 2011). A similar scenario was reported for *Schistosoma* infection, a common helminth parasite in tropical areas. Eosinophils participate not only in the host protection against schistosomes but also contribute to the derived pathogenesis, as fibrosis development and also the non-toxic variant would have been selected (Eriksson et al., 2007a).

Together with the well-documented involvement of eosinophils in parasitic infection, ECP levels are also increased in viral (Kristjansson et al., 2005) and bacterial infection (Venge et al., 1978; Karawajczyk et al., 1995; Niehaus et al., 2000; Ashitani et al., 2003), and a specific eosinophil activation is reported in bacterial infection (Kvarnhammar and Cardell, 2012). Although the direct contribution of eosinophils in bacterial infection is not so well established, many references in the literature indicate that this cell type complements significantly the neutrophil's role during host defence. Besides, ECP is also secondarily expressed in activated neutrophils (Monteseirin et al., 2007). In fact, the protein discovery was closely followed by the description of its antimicrobial properties (Gleich et al., 1986; Gabay et al., 1989). Preliminary work on eosinophil granules proteins already highlighted the high bactericidal activity of purified ECP on *Escherichia coli* and *Staphylococcus aureus* (Lehrer et al., 1989). Interestingly, the RNase catalytic site was found to be not required for the protein bactericidal activity (Rosenberg, 1995), and the protein toxicity is dependent on a direct action at the bacteria wall and cytoplasmic membrane levels (Torrent et al., 2007, 2008). Site-directed mutagenesis studies identified the main critical residues exposed at the protein surface involved in membrane lysis and cytotoxicity (Carreras et al., 2003; Torrent et al., 2007; Singh and Batra, 2011). Additionally, complementary studies on the protein native posttranslational forms and comparison of polymorphism variants highlighted the contribution of glycosylation on the protein cytotoxic power (Trulson et al., 2007; Rubin et al., 2009; Woschnagg et al., 2009). In the following section, we detail the current knowledge on ECP antimicrobial properties.

Dissecting ECP antimicrobial mechanism of action

Currently, amongst the diverse reported antipathogen ECP activities, the bactericidal activity is one of the most extensively documented. Table 2 tries to compile all the experimental data that have contributed to define the main protein active regions, and the location of the protein key motives is illustrated in Figure 2. The Figure highlights the residues involved in the protein antimicrobial activity (Carreras et al., 2003; Torrent et al., 2009b, 2011b; Singh and Batra, 2011) together with the defined RNase catalytic site (Rosenberg, 1995; Mohan et al., 2002). In particular, the N-terminus domain (1–45) includes the main determinants for the protein antimicrobial activity (Torrent et al., 2009b, 2011b), and this region encompasses a stretch required for the protein self-aggregation and the bacteria agglutinating activity (Torrent et al., 2010b, 2011b; Pulido et al., 2012). Taking into consideration the amino acid conservation in the context of the RNase A superfamily pylogenetic tree (Figure S1), the family-conserved catalytic groove stands out among ECP-specific residues related to the antimicrobial activity, as the colouring conservation score shows off (Figure 3).

By scanning the protein sequence using a prediction software based on an antimicrobial propensity scale (Torrent et al., 2012), an active domain at the N-terminus was identified (Torrent et al., 2009b). To further delimitate the protein key active regions and with a more biotechnological perspective, a structure-minimisation strategy was applied, to select the best pharmacophore candidate. The parental 1–45 peptide was prone at both the N- and C-terminus ends, and the two key helical stretches were further linked with a connector, which provides flexibility to the construct. The best final analogue (6–17)–(23–36) represents a 40% size reduction and can mostly reproduce the antimicrobial activity of the parental protein (Torrent et al., 2011b) (Table 3). NMR spectroscopy and dynamic light scattering confirmed that the final construct adopted the same helical conformation with a higher structuration degree upon incubation on a membrane-like environment (Torrent et al., 2011b).

Another interesting feature of ECP antimicrobial action relies on its particular bacteria-agglutinating activity (Torrent et al., 2010a), which was observed to be Gram-negative specific by testing several representative species (Torrent et al., 2011b) (see Table 3). Interestingly, a hydrophobic patch required for the protein self-aggregation mediates the protein ability to aggregate liposomes (Torrent et al., 2010b) and agglutinate bacterial cells. When scanning the eight human RNases with an aggregation prediction software (Conchillo-Sole et al., 2007), a unique hydrophobic patch (8–16) is identified in ECP (Figure 2), which is not present in the other homologue sequences. The Ile13Ala substitution breaks the aggregation-prone sequence, and the 1–19 peptide, corresponding to the first α helix, can form amyloid-like aggregates (Torrent et al., 2010b). Indeed, switching off the protein self-aggregation propensity by a single amino acid substitution abolishes the protein cell agglutination activity and impairs the antimicrobial activity (Torrent et al., unpublished results).

Table 2 ECP regions reported to be involved in the protein biological activities.

Protein region	Biological activities	References
K38, H128	RNase activity	(Rosenberg, 1995)
Q14, H15, H128, K38,R34	Nucleotide binding	(Mohan et al., 2002)
W10, W35R36, R101R104	Membrane lysis	(Carreras et al., 2003)
W35R36, R75F76, R101R104	Bactericidal against *E. coli*	(Carreras et al., 2003)
W10, W35R36, 115–122	Bactericidal against *S. aureus*	(Carreras et al., 2003)
W35R36	Cytotoxicity for HL60	(Carreras et al., 2005)
W35	Membrane interaction, membrane lysis	(Torrent et al., 2007; García-Mayoral et al., 2010)
R97	Cytotoxicity to host cells; antiparasitic activity	(Trulson et al., 2007; Rubin et al., 2009)
33–38	Heparin binding	(Fan et al., 2008)
24–45	LPS binding, bactericidal	(Torrent et al., 2009b)
1–19	Membrane destabilisation	(Torrent et al., 2009b)
33–36	Bactericidal against *E. coli*	(Torrent et al., 2009a)
1–45	Bactericidal, bacteria agglutination, membrane lysis, LPS binding	(Torrent et al., 2009b; Torrent et al., 2011b)
1–45	Membrane binding, heparin binding	(García-Mayoral et al., 2010)
1–19	Protein aggregation	(Torrent et al., 2010b)
I13	Protein aggregation	(Torrent et al., 2010b)
8–14, 33–36, 40–42, H64, R66, H128, D130	Heparin disaccharide binding	(García-Mayoral et al., 2010)
1–38	Bactericidal	(Sanchez et al., 2011)
23–36	Bacteria agglutination, LPS binding	(Torrent et al., 2011b)
R22, R34	Bactericidal Antiparasitic activity	(Singh and Batra, 2011)
R1, R7, Q14, R34	Heparin disaccharide binding	(Torrent et al., 2011a)
R1,W10, Q14, K38, Q40	LPS binding	(Pulido et al., 2012)
R1,R7,R34,R36,N39, Q14,K38,H128	Sulphate binding	(Boix et al., 2012)
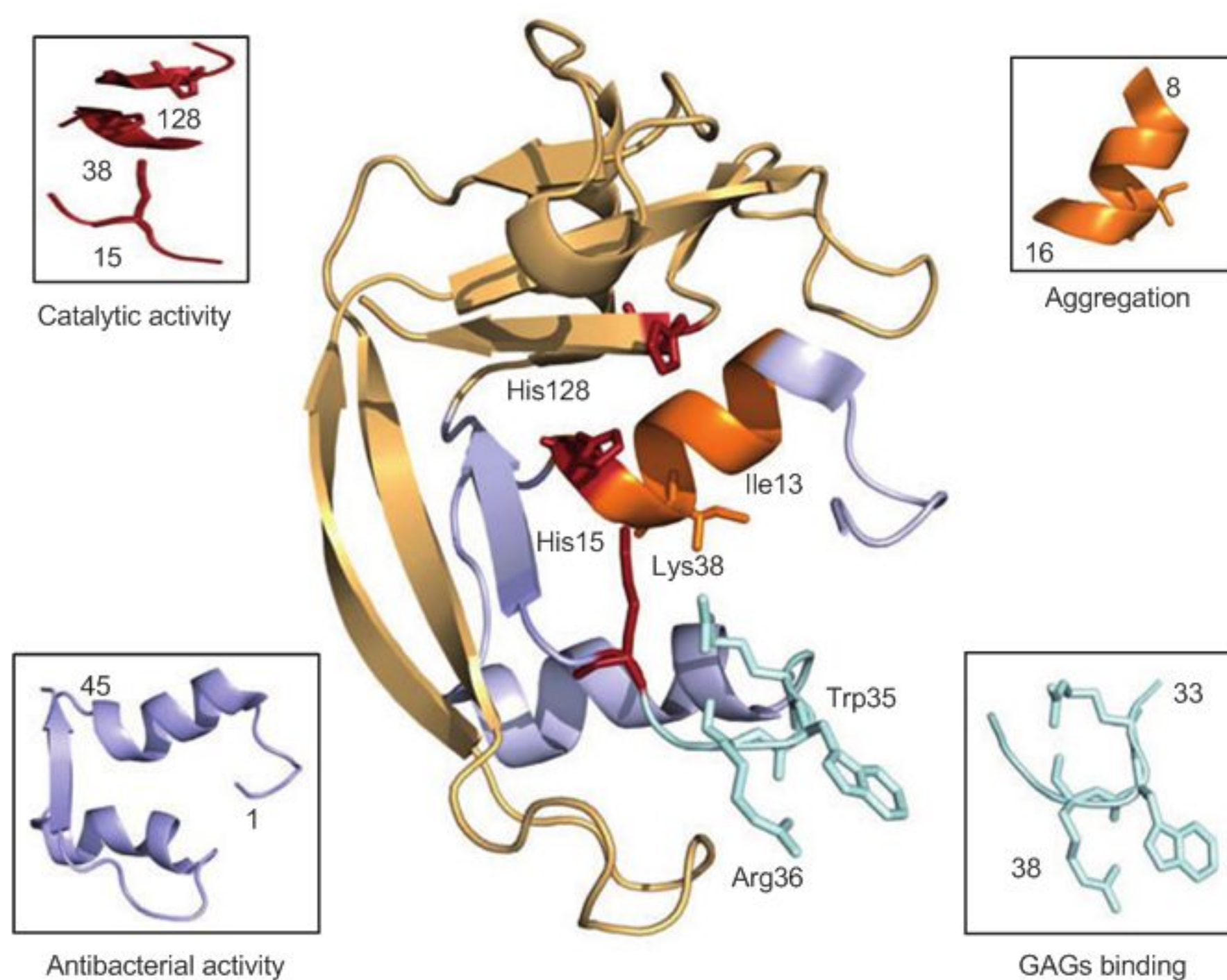
Figure 2 Representation of ECP three-dimensional structure showing the key regions and main residues identified to be involved in the protein biological properties.
Picture was drawn with *PYMOL* (DeLano Scientific).
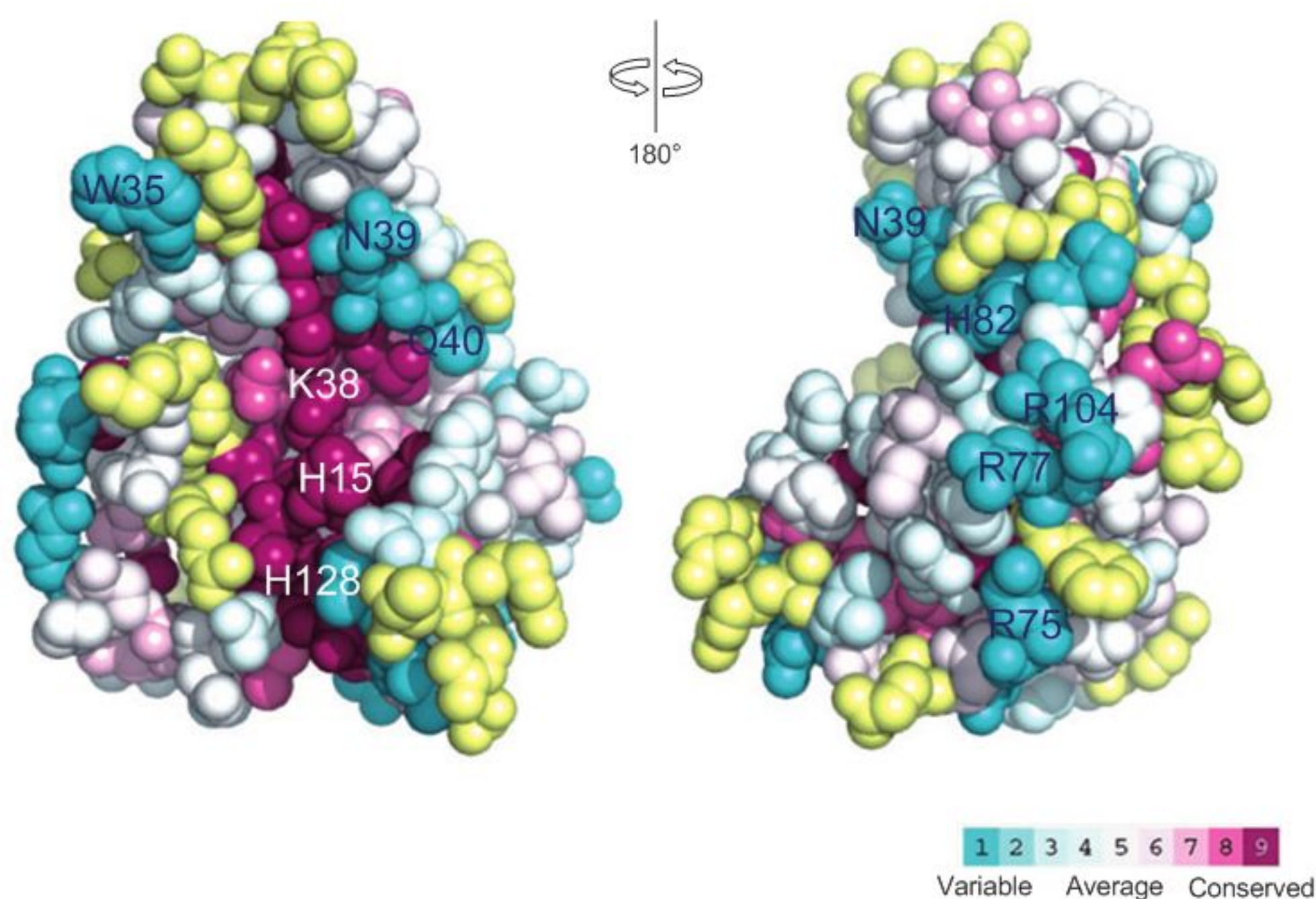
Figure 3 ECP three-dimensional structure surface representation using the CONSURF web server (http://consurf.tau.ac.il/) featuring the relationships among the evolutionary conservation of amino acid positions in the RNase A family.
Amino acid sequences were aligned using *Clustal W* taking human ECP as a reference (sequence information and their phylogenetic relationship is included in Figure S1). The three-dimensional structure shows the residues coloured by their conservation score using the colour-coding bar at the bottom image. Residues were coloured in yellow when not enough information was available. Conserved residues belonging to the RNase catalytic site are labelled in pink, and the residues most specific to ECP are labelled in blue.

ECP bacteria-agglutinating activity was also correlated with the protein ability to interact with the Gram-negative outer membrane lipopolysaccharide (LPS) structure by using a battery of *E. coli* cells with progressively truncated LPS (Pulido et al., 2012). Results indicated that the protein requires for a full antimicrobial activity both the LPS sugar core and the anionic phosphate groups, the agglutination activity being hampered by LPS truncation, mostly when the full core is removed. LPS-binding determinants were proposed by docking simulation, and residues Arg 1, Trp 10, Gln 14, Lys 38 and Gln 40 were spotted at hydrogen-bonding distance from the phosphorylated saccharide moiety anchored to the Lipid A portion. The data would correspond to the previous region identified to be involved in LPS binding using synthetic peptides (Torrent et al., 2009b, 2011b). Moreover, the YRWR (33–36) tag alone attached to the protein C-terminus provides the human pancreatic RNase 1 with Gram-negative specific bactericidal activity (Torrent et al.,

Table 3 Bactericidal and agglutination activities of ECP and peptide analogues calculated as MIC_{100} (μM) and minimal agglutinating concentration (MAC) values.

Peptide	E. coli	Pseudomonas sp	Acinetobacter baumannii	S. aureus	Micrococcus luteus	Enterococcus faecium	E. coli agglutination activity
ECP	0.4±0.1	0.62±0.07	0.6±0.1	0.40±0.06	1.5±0.3	0.87±0.07	+++[a]
1–45	0.62±0.07	0.6±0.1	0.6±0.1	0.62±0.07	1.5±0.3	0.87±0.07	++[b]
24–45	7±1	7±1	1.1±0.2	1.5±0.3	7±1	7±1	-
1–19	>10	>10	>10	>10	>10	>10	-
8–45	0.45±0.09	1.5±0.3	0.88±0.07	0.7±0.1	1.5±0.3	1.5±0.3	+[c]
8–36	1.5±0.3	3.5±0.9	1.5±0.3	0.7±0.1	3.5±0.9	1.5±0.3	+[c]
16–45	1.5±0.3	1.5±0.3	3.5±0.9	1.5±0.3	3.5±0.9	3.5±0.9	-
(6–17)~(23–36)	0.6±0.1	1.5±0.3	0.88±0.07	0.87±0.07	1.5±0.3	1.5±0.3	+[c]
(8–15)~(23–36)	1.5±0.3	3.5±0.9	7±1	1.1±0.2	>10	>10	-
(8–15)~(23–31)	>10	>10	>10	>10	>10	>10	-

MAC values were classified as: [a]below 0.2 μM, [b]below 0.5 μM and [c]below 1 μM. Modified from Torrent et al. (2011b). 6-Aminohexanoic acid (Ahx) connectors between α1 and α2 segments are represented by a tilde symbol.

2009a). The results' interest scope goes beyond the RNase A family. Finding the structural determinants for LPS recognition is a cherished goal for the pharmaceutical industry (Frecer et al., 2004; Kaconis et al., 2011). LPS released by bacteria during infection can induce septic shock, and LPS binders are indeed excellent candidates to neutralise the induction of proinflammatory cytokines and modulate the host immune response (Gutsmann et al., 2010; Brandenburg et al., 2011; Pulido et al., 2012).

The last but not the least, ECP cytotoxicity can also target the human host tissues (Navarro et al., 2008, 2010), contributing to undesired pathological side effects concomitant to eosinophil degranulation at the infection focus (Venge et al., 1999; Bystrom et al., 2011). Cell surface glycosaminoglycan (GAGs) recognition is one of the first steps that would mediate the protein toxicity to host cells (Fuchs and Raines, 2006). In particular, heparin and heparan sulphate are the main heterosaccharides exposed at the extracellular matrix (Capila and Linhardt, 2002). ECP shows a high binding affinity to heparin and inhibits the enzyme RNase activity (Torrent et al., 2011a). Critical residues involved in the interaction were identified by molecular docking (Torrent et al., 2011a) and NMR spectroscopy (García-Mayoral et al., 2010) and corroborated by mutagenesis studies (Fan et al., 2008). In particular, the 33–38 stretch is involved in ECP binding to GAGs (Figure 2) (Fan et al., 2008; Torrent et al., 2011a). The protein interaction with heparan sulphate at the matrix of epithelial cells would facilitate the protein endocytosis (Fan et al., 2007). Further insight on the protein residues potentially involved in heparin binding has been deduced by molecular docking as detailed below.

Exploring ECP binding sites

Molecular docking simulations using sulphated heterosaccharides and nucleotide templates were performed to correlate the diverse biological properties of ECP with its binding site structure. Molecular modelling was carried out previously (Torrent et al., 2011a) using the *Autodock* software (The Scripps Research Institute, La Jolla, CA, USA) as detailed in the supplemental material. Protein complexes for all eight human RNases were predicted in the presence of dimeric and tetrameric structures for nucleotide and heparin targets. CpA and ApTpApA nucleotides where chosen as representative putative RNA substrates, and the IdoA (2S) GlcNS (6S) (5S) disaccharide and the corresponding tetramer were selected as heparin main components. Tables S1–S4 show each probe molecule details and collect all the protein and ligand atoms at a hydrogen bond distance for the eight human RNases. Overall comparison of the RNase binding mode confirmed the contribution of equivalent protein recognition regions for most of the studied templates. Ligand binding mostly relies on interactions between protein polar and cationic residues and anionic groups, as nucleotide phosphates and heparin sulphates and carboxylates do. Overall, all tested ligands are located at the main protein catalytic groove. As an example, Figure 4 depicts the surface electrostatic representation of ECP structure and

illustrates the overlapping of both nucleotide and saccharide tetramers. We have highlighted the residues potentially involved in hydrogen bond interactions in the RNase sequence alignment (Figure 1). The data reveal that many nucleotide and sugar-binding residues are overlapping in most studied RNases, although a more tight binding is predicted for the heparin target. Most of the spotted residues are related either to the RNase catalytic groove or to the protein N-terminus region. To analyse the RNase binding site architecture, we have followed the reference nomenclature for RNase A where pn, Bn and Rn designates the subsites assigned to phosphate, base and ribose units, p1 being the position where the RNA phosphodiester bond is cleaved (Raines, 1998; Cuchillo et al., 2011). Docking results confirm that protein regions involved in both nucleotide and sugar binding (labelled in blue and orange, respectively, in the figure alignment) are mostly overlapping. In particular, the residues belonging to the main p1 nucleotide substrate binding site, that is, Gln 11, His 12, Lys 41 and His 119 for the pancreatic RNase, are conserved in all the predicted complexes (common residues shared for both targets are labelled in green). Besides, complementary interactions would facilitate the sugar binding at α1 and the connecting loop between α2
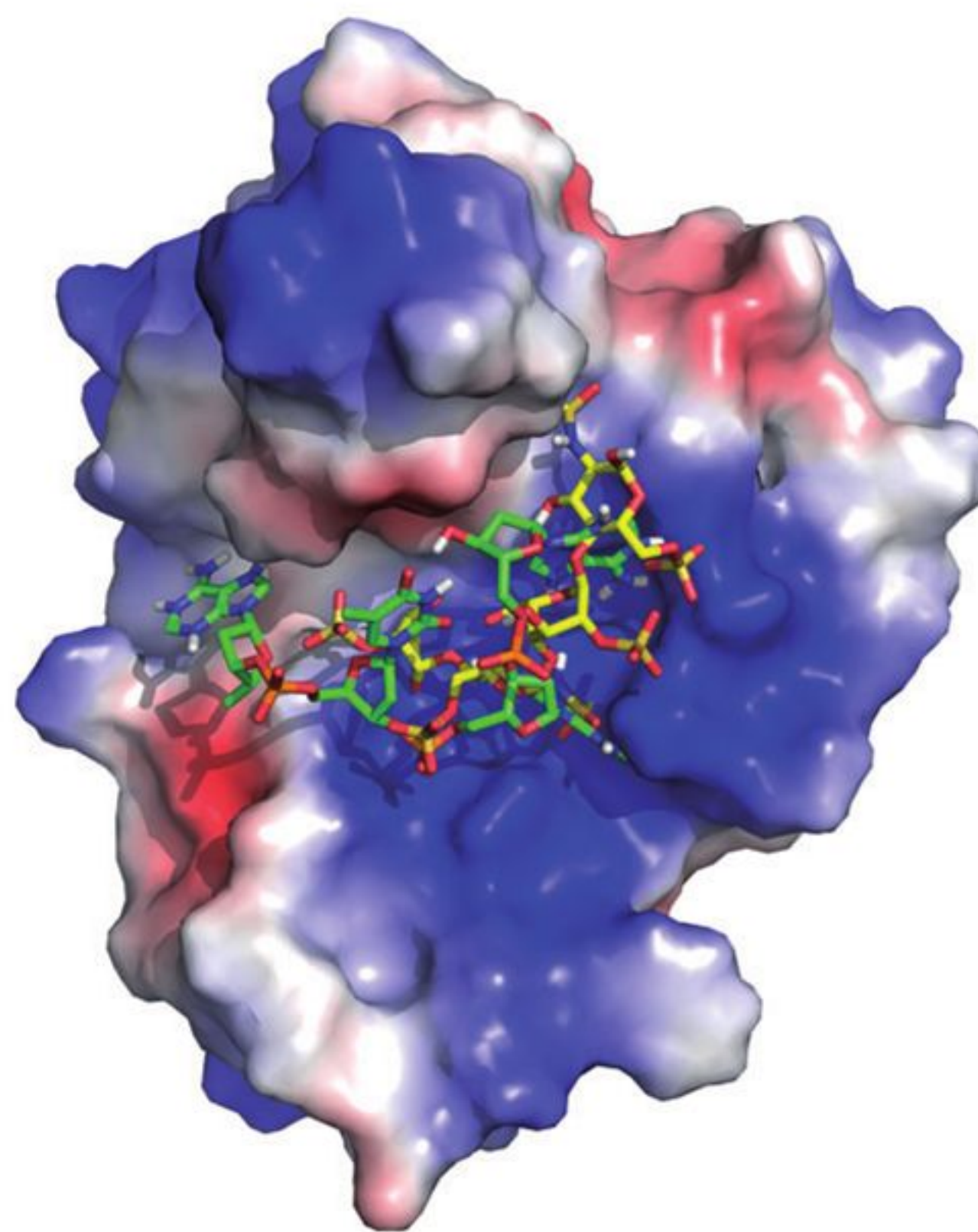

Figure 4 Electrostatic surface of ECP structure showing the modelled ligand location.
Atoms corresponding to the ApTpApA nucleotide are coloured in green, red, blue, white and orange for carbon, oxygen, nitrogen, hydrogen, and phosphate atoms, respectively. IdoA(2S) GlcNS (6S)-IdoA(2S) GlcNS (6S) heparin tretramer atoms are coloured in yellow, red, blue, white and orange for carbon, oxygen, nitrogen, hydrogen, and sulphate atoms, respectively.

and β1 (labelled in orange). These regions correspond to the previously assigned p-1 and p2 sites, which are flanking the main p1 (Nogues et al., 1998).

A closer detailed inspection was carried out for ECP-binding pattern. A side by side comparison of RNase A-binding mode to the ApTpApA tetranucleotide, as reported in the crystal structure by Fontecilla-Camps and co-workers (Fontecilla-Camps et al., 1994), with the predicted tetranucleotide complex for ECP (Figure 5), indicates that while most phosphate groups are located at equivalent positions, the orientations of the nucleotide bases differ significantly. Moreover, phosphate-conserved binding sites are indeed matching the sulphate sites as defined in the ECP-sulphate crystal complex (4A2O.pdb; Boix et al., 2012) (see residues labelled by stars in the sequence alignment). Besides, the sulphate-binding sites identified in ECP-sulphate crystal structure are shown to match the position of the sulphate groups of the heparin tetramer, as depicted in Figure 6. Three main sulphate-binding sites have been defined for ECP (S1–S3) (Boix et al., 2012), where S1 and S2 overlap with the previously ascribed phosphate sites deduced from the crystal complex with 2′5′ADP (Mohan et al., 2002), while Arg 1 and Arg 7 would correspond to an exposed anchoring region at the protein N-terminus defining the sulphate S3 site. In fact, side by side comparison of ECP-2′5′ ADP crystal structure and heparin-disaccharide predicted complex shows equivalent key residues at S1 and S2 sites, together with the S3 site that would enhance the probe binding (Figure 7). Interestingly, S1 and S2 sites were also reported for the other eosinophil RNase [RNase 2 or eosinophil-derived neurotoxin (EDN)] from protein-sulphate crystal complexes (Mosimann et al., 1996; Swaminathan et al., 2002) and nucleotide analogue
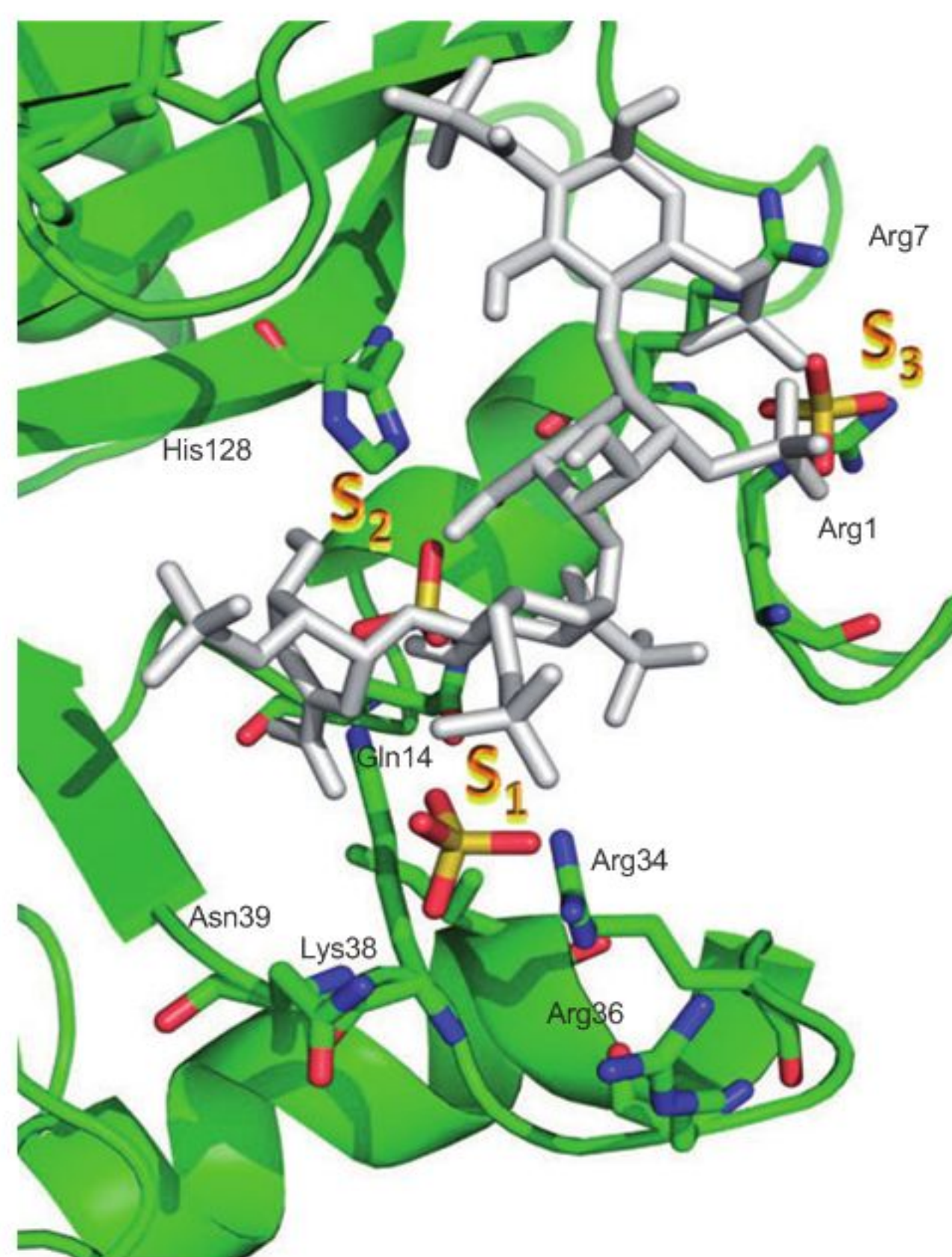
Figure 6   Representation of ECP-sulphate crystal complex (4A2O. pdb; Boix et al., 2012).
Protein-interacting residues and sulphate anions are shown in ball and sticks. The heparin tetramer IdoA(2S) GlcNS (6S)-IdoA(2S) GlcNS (6S) from the superposed modelled complex is shown coloured in white. Atoms types are coloured in green, red, blue and yellow for carbon, oxygen, nitrogen and sulphur atoms, respectively.
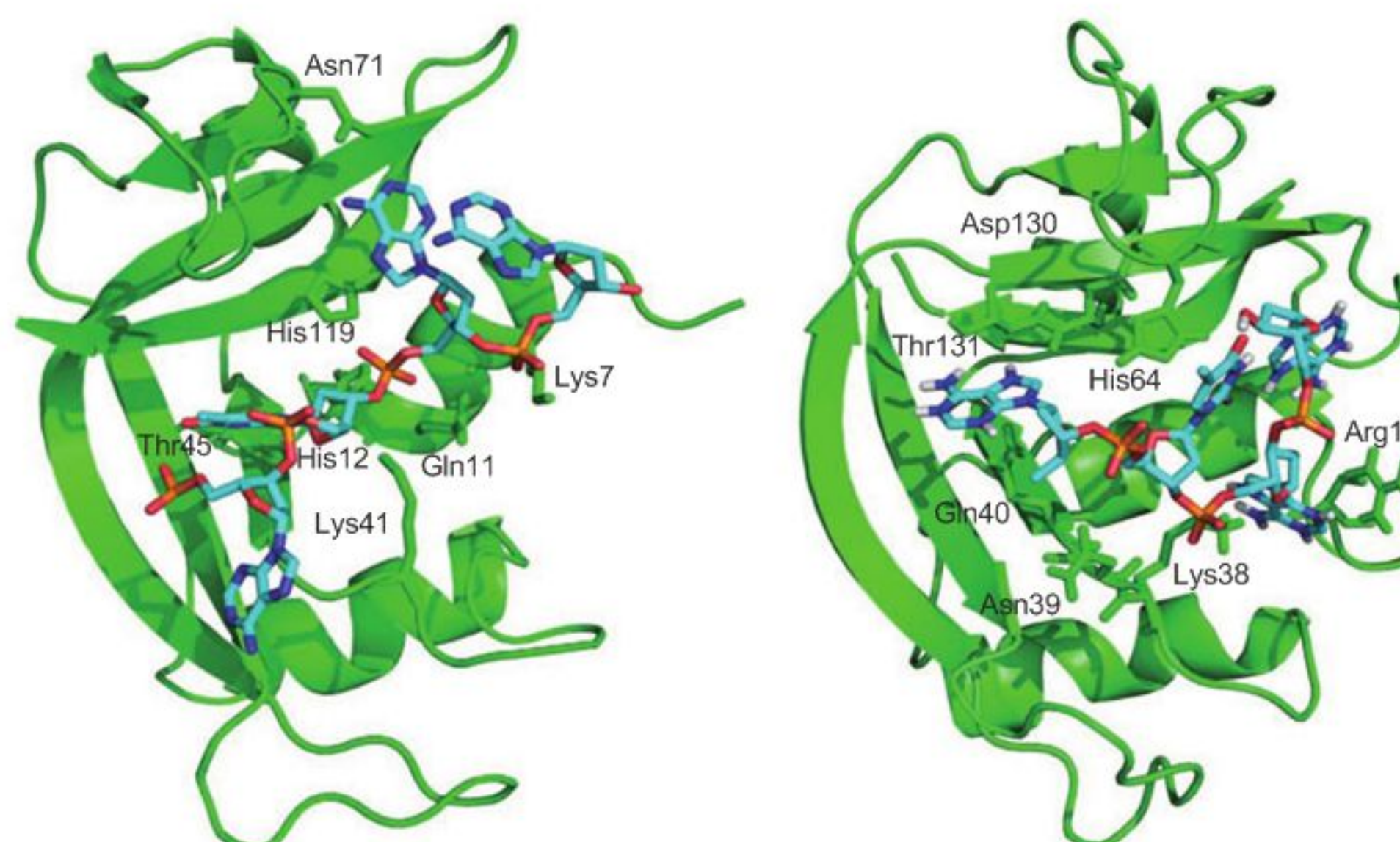
Figure 5   On the left, representation of RNase A with the d(ApTpApA) nucleotide taken from the complex crystal structure (1RCN.pdb; Fontecilla-Camps et al., 1994). On the right, representation of the modelled complex of ECP with ApTpApA. Amino acid side chains at hydrogen bond distance are shown in ball and sticks. Atom types of the ligand are coloured in cyan, red, blue, white, and yellow for carbon, oxygen, nitrogen, hydrogen and phosphate atoms, respectively. Docking was performed using *Autodock* v.4.2 as detailed in the supplemental material section.
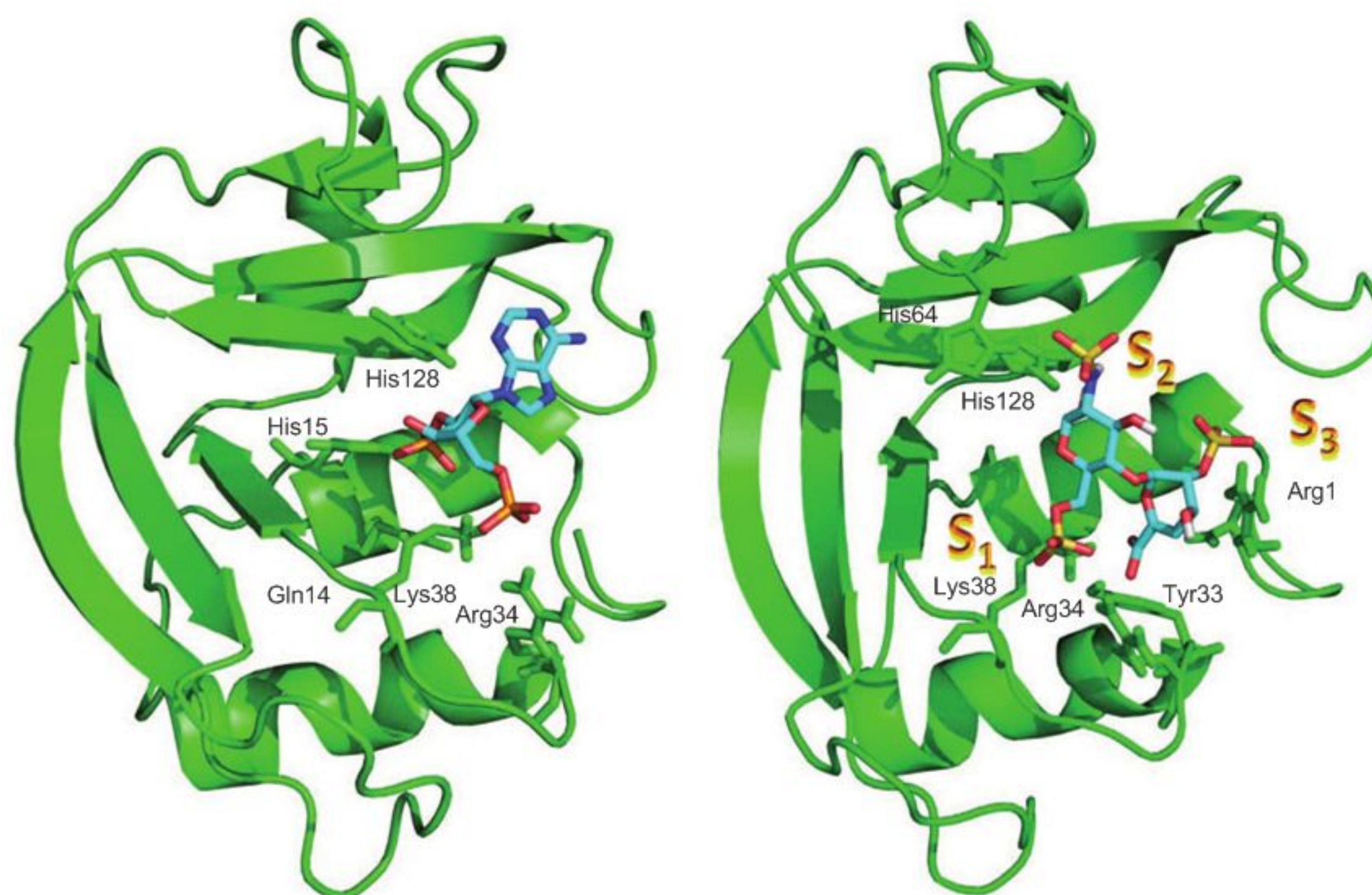
Figure 7   On the left, representation of ECP with the 2'5'ADP nucleotide taken from the crystal structure complex 1H1H.pdb (Mohan et al., 2002).
On the right, representation of the modelled complex for ECP with the heparin disaccharide IdoA(2S) GlcNS (6S) (5S). Amino acids side chains at hydrogen bond distance are shown in ball and sticks. Atom types are coloured in cyan, red, blue, white, orange and yellow for carbon, oxygen, nitrogen, hydrogen, phosphate and sulphate atoms, respectively. Docking was performed using *Autodock* v.4.2 as detailed in the supplemental material section.

complexes (Leonidas et al., 2001; Baker et al., 2006). Some of the residue contributions were confirmed experimentally by site-directed mutagenesis and functional studies for both eosinophil RNases (Carreras et al., 2003, 2005; Sikriwal et al., 2007, 2009; Singh and Batra, 2011).

In any case, although common traits are shared by the two eosinophil RNases, ECP diverged from its counterpart, acquiring an unusual cationicity, with a pI close to 11. Indeed ECP shows 19 Arg on its coding sequence, most of them recently incorporated during its divergent evolution from the other eosinophil RNase (Rosenberg et al., 1995; Zhang et al., 1998). Close inspection of the additional Arg residues, unique to ECP (Figure 3), suggests that their strategic location may enhance the binding of longer anionic polymers of biological significance, as LPS and heparan sulphate, which could mediate the protein action in host defence. Additionally, the protein-sulphate complex (4A2O.pdb) identifies other bound anions at unique Arg residues, namely, Arg 75, Arg 77 and Arg 104, also involved in the protein antipathogen activities (Carreras et al., 2003; Singh and Batra, 2011). To conclude, ECP structural analysis has contributed to define the overall traits that assist the protein-binding mode to glycosaminoglycans and mediate the protein-cell attachment and its immunomodulating properties (Lortat-Jacob et al., 2002; Imberty et al., 2007; Boix et al., 2011) and, hence, provides infor-

mation of general interest for structure-based drug design applications.

Conclusions

Structural analysis on ECP/RNase 3 defines the most relevant traits involved in the protein antimicrobial mechanism of action. The sequence determinants for ECP antimicrobial activity were mapped by applying predictive tools, proteolysis and experimental testing of synthetic peptides. Bactericidal activity was tested using representative Gram-positive and Gram-negative species, and a high binding affinity for LPS was observed. Indeed, a bacteria-agglutinating activity is specific for Gram-negative strains, and mutant strains with progressively truncated LPS were recently screened to assess the protein binding to the LPS polysaccharide moiety as a priming step in the bacteria-agglutination process.

Complementary, comparative docking analysis using both nucleotide and heterosaccharide probes was applied to define ECP-binding site architecture, in the context of the pancreatic RNase A family. The eight human members, with a low primary sequence identity, share a common three-dimensional fold. Together with the conserved catalytic triad, some variability is found for the residues involved in nucleotide secondary binding sites. Docking results and the analysis of ECP

crystal structure in complex with sulphate anions indicate that ECP would expose additional surface cationic clusters that would assist the protein anchoring to anionic heterosaccharide molecules, as LPS at the Gram-negative outer membrane and GAGs at the host cell extracellular matrix. The structural analysis unravels the protein key-binding motives that can modulate its role in innate immunity. The data provide the structural basis for ECP biological properties and can assist the design of inhibitors of the protein cytotoxicity to host tissues during inflammation processes, together with the development, in a much wider context, of therapeutic agents, such as LPS binders or alternative antibiotics.

Acknowledgments

References

Abtin, A., Eckhart, L., Mildner, M., Ghannadan, M., Harder, J., Schroder, J.M., and Tschachler, E. (2009). Degradation by stratum corneum proteases prevents endogenous rnase inhibitor from blocking antimicrobial activities of RNase 5 and RNase 7. J. Invest. Dermatol. *129*, 2193–2201.

Ackerman, S.J., Gleich, G.J., Loegering, D.A., Richardson, B.A., and Butterworth, A.E. (1985). Comparative toxicity of purified human eosinophil granule cationic proteins for schistosomula of Schistosoma mansoni. Am. J. Trop. Med. Hyg. *34*, 735–745.

Adu, B., Dodoo, D., Adukpo, S., Gyan, B.A., Hedley, P.L., Goka, B., Adjei, G.O., Larsen, S.O., Christiansen, M., and Theisen, M. (2011). Polymorphisms in the RNASE3 gene are associated with susceptibility to cerebral malaria in Ghanaian children. PLoS One *6*, e29465.

Ashitani, J., Yanagi, S., Arimura, Y., Sano, A., and Mukae, H. (2003). Acute respiratory distress syndrome induced by rifampicin with high levels of neutrophil and eosinophil products in bronchoalveolar lavage fluid. Respiration *70*, 541–543.

Ayabe, T., Satchell, D.P., Pesendorfer, P., Tanabe, H., Wilson, C.L., Hagen, S.J., and Ouellette, A.J. (2002). Activation of Paneth cell α-defensins in mouse small intestine. J. Biol. Chem. *277*, 5219–5228.

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

Beintema, J.J. and Kleineidam, R.G. (1998). The ribonuclease A superfamily: general discussion. Cell. Mol. Life Sci. *54*, 825–832.

Beintema, J.J., Blank, A., Schieven, G.L., Dekker, C.A., Sorrentino, S., and Libonati, M. (1988). Differences in glycosylation pattern of human secretory ribonucleases. Biochem. J. *255*, 501–505.

Blanchard, C. and Rothenberg, M.E. (2009). Biology of the Eosinophil. Adv. Immunol. *101*, 81–121.

Boix, E. and Nogues, M.V. (2007). Mammalian antimicrobial proteins and peptides: overview on the RNase A superfamily members involved in innate host defence. Mol. Biosyst. *3*, 317–335.

Boix, E., Torrent, M., Sánchez, D., and Nogués, M.V. (2008). The antipathogen activities of eosinophil cationic protein. Curr. Pharm. Biotechnol. *9*, 141–152.

Boix, E., Torrent, M., Nogués, M., and Salazar, V. (2011). Searching for heparin binding patterns. In: Heparin: Properties, Uses and Side Effects. E.D. Piyathilake and R. Liang, ed. (NY, USA: Nova Sciences Publishers, Inc).

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

Brandenburg, K., Andra, J., Garidel, P., and Gutsmann, T. (2011). Peptide-based treatment of sepsis. Appl. Microbiol. Biotechnol. *90*, 799–808.

Burton, M.F. and Steel, P.G. (2009). The chemistry and biology of LL-37. Nat. Prod. Rep. *26*, 1572–1584.

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

Capila, I. and Linhardt, R.J. (2002). Heparin-protein interactions. Angew. Chem. Int. Ed. *41*, 391–412.

Carreras, E., Boix, E., Rosenberg, H.F., Cuchillo, C.M., and Nogues, M.V. (2003). Both aromatic and cationic residues contribute to the membrane-lytic and bactericidal activity of eosinophil cationic protein. Biochemistry *42*, 6636–6644.

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

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

Cho, S., Beintema, J.J., and Zhang, J. (2005). The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. Genomics *85*, 208–220.

Cho, S. and Zhang, J. (2007). Zebrafish ribonucleases are bactericidal: implications for the origin of the vertebrate RNase A superfamily. Mol. Biol. Evol. *24*, 1259–1268.

Clark, H.F., Gurney, A.L., Abaya, E., Baker, K., Baldwin, D., Brush, J., Chen, J., Chow, B., Chui, C., Crowley, C., et al. (2003). The Secreted Protein Discovery Initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment. Genome Res. *13*, 2265–2270.

Conchillo-Sole, O., de Groot, N.S., Aviles, F.X., Vendrell, J., Daura, X., and Ventura, S. (2007). AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. BMC Bioinformatics *8*, 65.

Cuchillo, C.M., Nogues, M.V., and Raines, R.T. (2011). Bovine pancreatic ribonuclease: fifty years of the first enzymatic reaction mechanism. Biochemistry *50*, 7835–7841.

During, K., Porsch, P., Mahn, A., Brinkmann, O., and Gieffers, W. (1999). The non-enzymatic microbicidal activity of lysozymes. FEBS Lett. *449*, 93–100.

Dyer, K.D., Rosenberg, H.F., and Zhang, J.Z. (2004). Isolation, characterization, and evolutionary divergence of mouse RNase 6: evidence for unusual evolution in rodents. J. Mol. Evol. *59*, 657–665.

Elshafie, A.I., Ahlin, E., Hakansson, L.D., Elghazali, G., El Safi, S.H., Ronnelid, J., and Venge, P. (2011). Activity and turnover of eosinophil and neutrophil granulocytes are altered in visceral leishmaniasis. Int. J. Parasitol. *41*, 463–469.

Eriksson, J., Reimert, C.M., Kabatereine, N.B., Kazibwe, F., Ireri, E., Kadzo, H., Eltahir, H.B., Mohamed, A.O., Vennervald, B.J., and Venge, P. (2007a). The 434(G>C) polymorphism within the coding sequence of Eosinophil Cationic Protein (ECP) correlates with the natural course of Schistosoma mansoni infection. Int. J. Parasitol. *37*, 1359–1366.

Eriksson, J., Woschnagg, C., Fernvik, E., and Venge, P. (2007b). A SELDI-TOF MS study of the genetic and post-translational molecular heterogeneity of eosinophil cationic protein. J. Leukoc. Biol. *82*, 1491–1500.

Fan, T.C., Chang, H.T., Chen, I.W., Wang, H.Y., and Chang, M.D. (2007). A heparan sulfate-facilitated and raft-dependent macropinocytosis of eosinophil cationic protein. Traffic *8*, 1778–1795.

Fan, T.C., Fang, S.L., Hwang, C.S., Hsu, C.Y., Lu, X.A., Hung, S.C., Lin, S.C., and Chang, D.T. (2008). Characterization of molecular interactions between eosinophil cationic protein and heparin. J. Biol. Chem. *283*, 25468–25474.

Fontecilla-Camps, J.C., de Llorens, R., le Du, M.H., and Cuchillo, C.M. (1994). Crystal structure of ribonuclease A.d(ApTpApApG) complex. Direct evidence for extended substrate recognition. J. Biol. Chem. *269*, 21526–21531.

Frecer, V., Ho, B., and Ding, J.L. (2004). De novo design of potent antimicrobial peptides. Antimicrob. Agents Chemother. *48*, 3349–3357.

Fuchs, S.M. and Raines, R.T. (2006). Internalization of cationic peptides: the road less (or more?) traveled. Cell. Mol. Life Sci. *63*, 1819–1822.

Gabay, J.E., Scott, R.W., Campanelli, D., Griffith, J., Wilde, C., Marra, M.N., Seeger, M., and Nathan, C.F. (1989). Antibiotic proteins of human polymorphonuclear leukocytes. Proc. Natl. Acad. Sci. USA *86*, 5610–5614.

Gao, X.W. and Xu, Z.P. (2008). Mechanisms of action of angiogenin. Acta Biochim. Biophys. Sin. *40*, 619–624.

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

Giancola, C., Ercole, C., Fotticchia, I., Spadaccini, R., Pizzo, E., D'Alessio, G., and Picone, D. (2011). Structure-cytotoxicity relationships in bovine seminal ribonuclease: new insights from heat and chemical denaturation studies on variants. FEBS J. *278*, 111–122.

Gleich, G.J., Loegering, D.A., Bell, M.P., Checkel, J.L., Ackerman, S.J., and McKean, D.J. (1986). Biochemical and functional similarities between human eosinophil-derived neurotoxin and eosinophil cationic protein: homology with ribonuclease. Proc. Natl. Acad. Sci. USA *83*, 3146–3150.

Greenway, M.J., Andersen, P.M., Russ, C., Ennis, S., Cashman, S., Donaghy, C., Patterson, V., Swingler, R., Kieran, D., Prehn, J., et al. (2006). ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. Nat. Genet. *38*, 411–413.

Gutsmann, T., Razquin-Olazaran, I., Kowalski, I., Kaconis, Y., Howe, J., Bartels, R., Hornef, M., Schurholz, T., Rossle, M., Sanchez-Gomez, S., et al. (2010). New antiseptic peptides to protect against endotoxin-mediated shock. Antimicrob. Agents Chemother. *54*, 3817–3824.

Hamann, K.J., Gleich, G.J., Checkel, J.L., Loegering, D.A., McCall, J.W., and Barker, R.L. (1990). In vitro killing of microfilariae of Brugia pahangi and Brugia malayi by eosinophil granule proteins. J. Immunol. *144*, 3166–3173.

Harder, J. and Schroder, J.M. (2002). RNase 7, a novel innate immune defense antimicrobial protein of healthy human skin. J. Biol. Chem. *277*, 46779–46784.

Hofsteenge, J., Vicentini, A., and Zelenko, O. (1998). Ribonuclease 4, an evolutionarily highly conserved member of the superfamily. Cell. Mol. Life Sci. *54*, 804–810.

Hooper, L.V., Stappenbeck, T.S., Hong, C.V., and Gordon, J.I. (2003). Angiogenins: a new class of microbicidal proteins involved in innate immunity. Nat. Immunol. *4*, 269–273.

Imberty, A., Lortat-Jacob, H., and Perez, S. (2007). Structural view of glycosaminoglycan-protein interactions. Carbohydr. Res. *342*, 430–439.

Kaconis, Y., Kowalski, I., Howe, J., Brauser, A., Richter, W., Razquin-Olazaran, I., Inigo-Pestana, M., Garidel, P., Rossle, M., de Tejada, G.M., et al. (2011). Biophysical mechanisms of endotoxin neutralization by cationic amphiphilic peptides. Biophys. J. *100*, 2652–2661.

Karawajczyk, M., Pauksen, K., Peterson, C.G.B., Eklund, E., and Venge, P. (1995). The differential release of eosinophil granule proteins-studies on patients with acute bacterial and viral infections. Clin. Exp. Allergy *25*, 713–719.

Kierszenbaum, F., Villalta, F., and Tai, P.C. (1986). Role of inflammatory cells in Chagas' disease .3. Kinetics of human eosinophil activation upon interaction with parasites (*Trypanosoma cruzi*). J. Immunol. *136*, 662–666.

Klion, A.D. and Nutman, T.B. (2004). The role of eosinophils in host defense against helminth parasites. J. Allergy Clin. Immunol. *113*, 30–37.

Krieg, J., Hartmann, S., Vicentini, A., Glasner, W., Hess, D., and Hofsteenge, J. (1998). Recognition signal for C-mannosylation of Trp-7 in RNase 2 consists of sequence Trp-x-x-Trp. Mol. Biol. Cell *9*, 301–309.

Kristjansson, S., Bjarnarson, S.P., Wennergren, G.E., Palsdottir, A.H., Arnadottir, T., Haraidsson, A., and Jonsdottir, I. (2005). Respiratory syncytial virus and other respiratory viruses during the first 3 months of life promote a local T(H)2-like response. J. Allergy Clin. Immunol. *116*, 805–811.

Kurtzhals, J.A.L., Reimert, C.M., Tette, E., Dunyo, S.K., Koram, K.A., Akanmori, B.D., Nkrumah, F.K., and Hviid, L. (1998). Increased eosinophil activity in acute Plasmodium falciparum infection - association with cerebral malaria. Clin. Exp. Immunol. *112*, 303–307.

Kvarnhammar, A.M. and Cardell, L.O. (2012). Pattern-recognition receptors in human eosinophils. Immunology *136*, 11–20.

Lehrer, R.I., Szklarek, D., Barton, A., Ganz, T., Hamann, K.J., and Gleich, G.J. (1989). Antibacterial properties of eosinophil major basic protein and eosinophil cationic protein. J. Immunol. *142*, 4428–4434.

Leonidas, D.D., Boix, E., Prill, R., Suzuki, M., Turton, R., Minson, K., Swaminathan, G.J., Youle, R.J., and Acharya, K.R. (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*, 15009–15017.

Levy, O. (2000). Antimicrobial proteins and peptides of blood: templates for novel antimicrobial agents. Blood *96*, 2664–2672.

Lortat-Jacob, H., Grosdidier, A., and Imberty, A. (2002). Structural diversity of heparan sulfate binding domains in chemokines. Proc. Natl. Acad. Sci. USA *99*, 1229–1234.

Makepeace, B.L., Martin, C., Turner, J.D., and Specht, S. (2012). Granulocytes in helminth infection-who is calling the shots? Curr. Med. Chem. *19*, 1567–1586.

Malik, A. and Batra, J.K. (2012). Antimicrobial activity of human eosinophil granule proteins: involvement in host defence against pathogens. Crit. Rev. Microbiol. *38*, 168–181.

Masschalck, B., Van Houdt, R., Van Haver, E.G.R., and Michiels, C.W. (2001). Inactivation of gram-negative bacteria by lysozyme, denatured lysozyme, and lysozyme-derived peptides under high hydrostatic pressure. Appl. Environ. Microbiol. *67*, 339–344.

Mclaren, D.J., Peterson, C.G.B., and Venge, P. (1984). *Schistosoma mansoni*-further studies of the interaction between schistosomula and granulocyte-derived cationic proteins in vitro. Parasitology *88*, 491–503.

Moenner, M., Gusse, M., Hatzi, E., and Badet, J. (1994). The widespread expression of angiogenin in different human cells suggests a biological function not only related to angiogenesis. Eur. J. Biochem. *226*, 483–490.

Mohan, C.G., Boix, E., Evans, H.R., Nikolovski, Z., Nogues, M.V., Cuchillo, C.M., and Acharya, K.R. (2002). 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*, 12100–12106.

Monteseirin, J., Vega, A., Chacon, P., Camacho, M.J., El Bekay, R., Asturias, J.A., Martinez, A., Guardia, P., Perez-Cano, R., and Conde, J. (2007). Neutrophils as a novel source of eosinophil cationic protein in IgE-mediated processes. J. Immunol. *179*, 2634–2641.

Mosimann, S.C., Newton, D.L., Youle, R.J., and James, M.N. (1996). X-ray crystallographic structure of recombinant eosinophil-derived neurotoxin at 1.83 Å resolution. J. Mol. Biol. *260*, 540–552.

Nakatsuji, T. and Gallo, R.L. (2012). Antimicrobial peptides: old molecules with new ideas. J. Invest. Dermatol. *132*, 887–895.

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

Navarro, S., Boix, E., Cuchillo, C.M., and Nogues, M.V. (2010). Eosinophil-induced neurotoxicity: the role of eosinophil cationic protein/RNase 3. J. Neuroimmunol. *227*, 60–70.

Nibbering, P.H., Ravensbergen, E., Welling, M.M., van Berkel, L.A., van Berkel, P.H., Pauwels, E.K., and Nuijens, J.H. (2001). Human lactoferrin and peptides derived from its N terminus are highly effective against infections with antibiotic-resistant bacteria. Infect. Immun. *69*, 1469–1476.

Niehaus, M.D., Gwaltney, J.M., Hendley, J.O., Newman, M.J., Heymann, P.W., Rakes, G.P., Platts-Mills, T.A.E., and Guerrant, R.L. (2000). Lactoferrin and eosinophilic cationic protein in nasal secretions of patients with experimental rhinovirus colds, natural colds, and presumed acute community-acquired bacterial sinusitis. J. Clin. Microbiol. *38*, 3100–3102.

Nitto, T., Dyer, K.D., Czapiga, M., and Rosenberg, H.F. (2006). Evolution and function of leukocyte RNase A ribonucleases of the avian species, *Gallus gallus*. J. Biol. Chem. *281*, 25622–25634.

Nogues, M.V., Moussaoui, M., Boix, E., Vilanova, M., Ribo, M., and Cuchillo, C.M. (1998). The contribution of noncatalytic phosphate-binding subsites to the mechanism of bovine pancreatic ribonuclease A. Cell. Mol. Life Sci. *54*, 766–774.

Pizzo, E. and D'Alessio, G. (2007). The success of the RNase scaffold in the advance of biosciences and in evolution. Gene *406*, 8–12.

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

Pulido, D., Moussaoui, M., Andreu, D., Nogués, M., Torrent, M., and Boix, E. (2012). Antimicrobial action and cell agglutination

by eosinophil cationic protein is modulated by the cell wall lipopolysaccharide structure. Antimicrob. Agents Chemother. *56*, 2378–2385.

Raines, R.T. (1998). Ribonuclease A. Chem. Rev. *98*, 1045–1066.

Rosenberg, H.F. (1995). Recombinant human eosinophil cationic protein. Ribonuclease activity is not essential for cytotoxicity. J. Biol. Chem. *270*, 7876–7881.

Rosenberg, H.F. (2008a). Eosinophil-derived neurotoxin/RNase 2: connecting the past, the present and the future. Curr. Pharm. Biotechnol. *9*, 135–140.

Rosenberg, H.F. (2008b). RNase A ribonucleases and host defense: an evolving story. J. Leukoc. Biol. *83*, 1079–1087.

Rosenberg, H.F. and Dyer, K.D. (1996). Molecular cloning and characterization of a novel human ribonuclease (RNase k6): increasing diversity in the enlarging ribonuclease gene family. Nucleic Acids Res. *24*, 3507–3513.

Rosenberg, H.F., Dyer, K.D., Tiffany, H.L., and Gonzalez, M. (1995). Rapid evolution of a unique family of primate ribonuclease genes. Nat. Genet. *10*, 219–223.

Rubin, J., Zagai, U., Blom, K., Trulson, A., Engstrom, A., and Venge, P. (2009). The coding ECP 434(G>C) gene polymorphism determines the cytotoxicity of ECP but has minor effects on fibroblast-mediated gel contraction and no effect on RNase activity. J. Immunol. *183*, 445–451.

Rudolph, B., Podschun, R., Sahly, H., Schubert, S., Schroder, J.M., and Harder, J. (2006). Identification of RNase 8 as a novel human antimicrobial protein. Antimicrob. Agents Chemother. *50*, 3194–3196.

Sanchez, D., Moussaoui, M., Carreras, E., Torrent, M., Nogues, V., and Boix, E. (2011). Mapping the eosinophil cationic protein antimicrobial activity by chemical and enzymatic cleavage. Biochimie *93*, 331–338.

Schroeder, B.O., Wu, Z.H., Nuding, S., Groscurth, S., Marcinowski, M., Beisner, J., Buchner, J., Schaller, M., Stange, E.F., and Wehkamp, J. (2011). Reduction of disulphide bonds unmasks potent antimicrobial activity of human β-defensin 1. Nature *469*, 419–423.

Sikriwal, D., Seth, D., Dey, P., and Batra, J.K. (2007). Human eosinophil-derived neurotoxin: involvement of a putative non-catalytic phosphate-binding subsite in its catalysis. Mol. Cell. Biochem. *303*, 175–181.

Sikriwal, D., Seth, D., and Batra, J.K. (2009). Role of catalytic and non-catalytic subsite residues in ribonuclease activity of human eosinophil-derived neurotoxin. Biol. Chem. *390*, 225–234.

Simanski, M., Köten, B., Schröder, J.M., Gläser, R., and Harder, J. (2012). Antimicrobial RNases in cutaneous defense. J. Innate Immun. *4*, 241–247.

Singh, A. and Batra, J.K. (2011). Role of unique basic residues in cytotoxic, antibacterial and antiparasitic activities of human eosinophil cationic protein. Biol. Chem. *392*, 337–346.

Sorrentino, S. (2010). The eight human "canonical" ribonucleases: molecular diversity, catalytic properties, and special biological actions of the enzyme proteins. FEBS Lett. *584*, 2194–2200.

Spencer, J.D., Schwaderer, A.L., DiRosario, J.D., McHugh, K.M., McGillivary, G., Justice, S.S., Carpenter, A.R., Baker, P.B., Harder, J., and Hains, D.S. (2011). Ribonuclease 7 is a potent antimicrobial peptide within the human urinary tract. Kidney Int. *80*, 175–181.

Strydom, D.J. (1998). The angiogenins. Cell. Mol. Life Sci. *54*, 811–824.

Strydom, D.J., Fett, J.W., Lobb, R.R., Alderman, E.M., Bethune, J.L., Riordan, J.F., and Vallee, B.L. (1985). Amino-acid sequence of human-tumor derived angiogenin. Biochemistry *24*, 5486–5494.

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

Tischendorf, F.W., Brattig, N.W., Buttner, D.W., Pieper, A., and Lintzel, M. (1996). Serum levels of eosinophil cationic protein, eosinophil-derived neurotoxin and myeloperoxidase in infections with filariae and schistosomes. Acta Trop. *62*, 171–182.

Torrent, M., Cuyas, E., Carreras, E., Navarro, S., Lopez, O., de la Maza, A., Nogues, M.V., Reshetnyak, Y.K., and Boix, E. (2007). Topography studies on the membrane interaction mechanism of the eosinophil cationic protein. Biochemistry *46*, 720–733.

Torrent, M., Navarro, S., Moussaoui, M., Nogues, M.V., and Boix, E. (2008). Eosinophil cationic protein high-affinity binding to bacteria-wall lipopolysaccharides and peptidoglycans. Biochemistry *47*, 3544–3555.

Torrent, G., Ribo, M., Benito, A., and Vilanova, M. (2009a). Bactericidal activity engineered on human pancreatic ribonuclease and onconase. Mol. Pharm. *6*, 531–542.

Torrent, M., de la Torre, B.G., Nogues, V.M., Andreu, D., and Boix, E. (2009b). Bactericidal and membrane disruption activities of the eosinophil cationic protein are largely retained in an N-terminal fragment. Biochem. J. *421*, 425–434.

Torrent, M., Badia, M., Moussaoui, M., Sanchez, D., Nogues, M.V., and Boix, E. (2010a). Comparison of human RNase 3 and RNase 7 bactericidal action at the Gram-negative and Gram-positive bacterial cell wall. FEBS J. *277*, 1713–1725.

Torrent, M., Odorizzi, F., Nogues, M.V., and Boix, E. (2010b). Eosinophil cationic protein aggregation: identification of an N-terminus amyloid prone region. Biomacromolecules *11*, 1983–1990.

Torrent, M., Nogues, M.V., and Boix, E. (2011a). Eosinophil cationic protein (ECP) can bind heparin and other glycosaminoglycans through its RNase active site. J. Mol. Recognit. *24*, 90–100.

Torrent, M., Pulido, D., De la Torre, B.G., Garcia de la Torre, J., Nogues, M.V., Bruix, M., Andreu, D., and Boix, E. (2011b). Refining the eosinophil cationic protein antibacterial pharmacophore by rational structure minimization. J. Med. Chem. *54*, 7.

Torrent, M., Di Tommaso, P., Pulido, D., Nogués, M., Notredame, C., Boix, E., and Andreu, D. (2012). AMPA: an automated web server for prediction of protein antimicrobial regions. Bioinformatics *28*, 130–131.

Trulson, A., Bystrom, J., Engstrom, A., Larsson, R., and Venge, P. (2007). The functional heterogeneity of eosinophil cationic protein is determined by a gene polymorphism and post-translational modifications. Clin. Exp. Allergy *37*, 208–218.

Ulrich, M., Petre, A., Youhnovski, N., Promm, F., Schirle, M., Schumm, M., Pero, R.S., Doyle, A., Checkel, J., Kita, H., et al. (2008). Post-translational tyrosine nitration of eosinophil granule toxins mediated by eosinophil peroxidase. J. Biol. Chem. *283*, 28629–28640.

Venge, P., Roxin, L.E., and Olsson, I. (1977). Radioimmunoassay of human eosinophil cationic protein. Br. J. Haematol. *37*, 331–335.

Venge, P., Stromberg, A., Braconier, J.H., Roxin, L.E., and Olsson, I. (1978). Neutrophil and eosinophil granulocytes in bacterial infection: sequential studies of cellular and serum levels of granule proteins. Br. J. Haematol. *38*, 475–483.

Venge, P., Bystrom, J., Carlson, M., Hakansson, L., Karawacjzyk, M., Peterson, C., Seveus, L., and Trulson, A. (1999). Eosinophil cationic protein (ECP): molecular and biological properties and the use of ECP as a marker of eosinophil activation in disease. Clin. Exp. Allergy *29*, 1172–1186.

Waters, L.S., Taverne, J., Tai, P.C., Spry, C.J.F., Targett, G.A.T., and Playfair, J.H.L. (1987). Killing of *Plasmodium falciparum* by eosinophil secretory products. Infect. Immun. *55*, 877–881.

Woschnagg, C., Rubin, J., and Venge, P. (2009). Eosinophil cationic protein (ECP) is processed during secretion. J. Immunol. *183*, 3949–3954.

Zanetti, M., Gennaro, R., and Romeo, D. (1995). Cathelicidins: a novel protein family with a common proregion and a variable C-terminal antimicrobial domain. FEBS Lett. *374*, 1–5.

Zanfardino, A., Pizzo, E., Di Maro, A., Varcamonti, M., and D'Alessio, G. (2010). The bactericidal action on *Escherichia coli* of ZF-RNase-3 is triggered by the suicidal action of the bacterium OmpT protease. FEBS J. *277*, 1921–1928.

Zhang, J., Rosenberg, H.F., and Nei, M. (1998). Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proc. Natl. Acad. Sci. USA *95*, 3708–3713.

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

Zhang, J., Dyer, K.D., and Rosenberg, H.F. (2003). Human RNase 7: a new cationic ribonuclease of the RNase A superfamily. Nucleic Acids Res. *31*, 602–607.

Ester Boix graduated in Biological Sciences in 1986 at the Universitat de Barcelona and in 1994 obtained her PhD in Biochemistry at the Universitat Autònoma de Barcelona (UAB). During a first postdoctoral stage at the National Institutes of Health, Bethesda, that she enrolled on cytotoxic RNases and following that she enrolled in the Structural Biology Group at the University of Bath. In 2001 she moved to the Biochemistry and Molecular Biology Dpt., UAB, first as a lecturer, then as a "Ramón y Cajal" research fellow and since 2004 as an Associate Professor. Her current research project focuses on the antipathogen mechanism of action of human RNases involved in host defense.

Vivian Salazar obtained her BSc in Bacteriology from Pontificia Universidad Javeriana (Colombia) in 2002, and MSc in Biological Sciences at Andes University (Colombia) in 2008. She then moved to Universitat Autònoma de Barcelona where she obtained her MSc in Biochemistry, Molecular Biology and Biomedicine in 2011. Currently she is a PhD student working on the structure–function characterization of human ribonucleases.

Marc Torrent obtained his PhD in Biochemistry at the Universitat Autònoma de Barcelona in 2009. From 2009 to 2011, he worked as a postdoctoral fellow at the Universitat Pompeu Fabra and currently he is a Career Development Fellow at the Laboratory of Molecular Biology at Cambridge, UK. His main research interests include the study of the structural activity relationships on antimicrobial proteins and peptides and how evolution has shaped their sequence to develop and improve their antimicrobial activity.
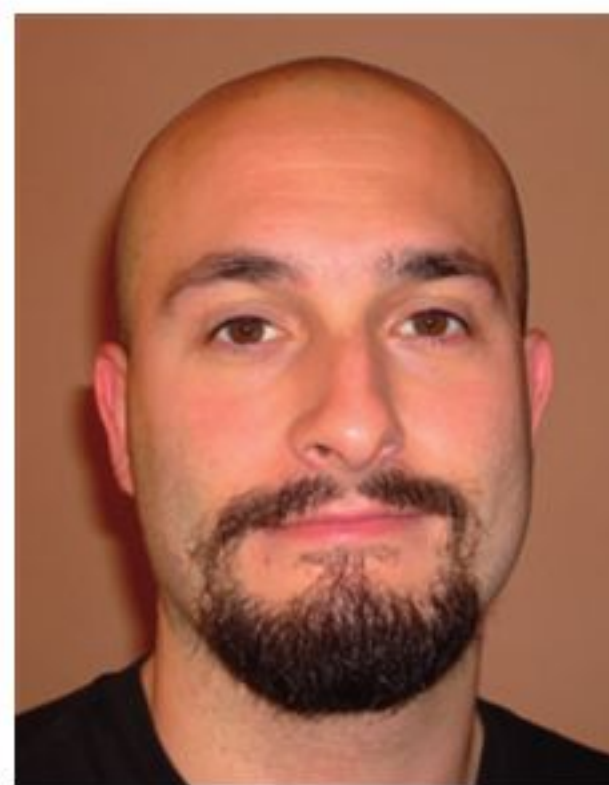
David Pulido graduated in Microbiology at the Universitat Autònoma de Barcelona (UAB) in 2009. One year later he obtained a MSc in Biochemistry and Molecular Biology at the same university. Currently he is a PhD student. His work in the Biochemistry department of UAB is centered on the study of human RNases structure, their antimicrobial mechanism of action and role in the host immune system.

M. Victòria Nogués (1954, Barcelona, Spain) graduated in Biological Sciences and obtained her PhD in Biological Sciences in 1982 at Universitat Autònoma de Barcelona (Spain). Presently she is Full Professor at the Department of Biochemistry and Molecular Biology at the Faculty of Biosciences of the Universitat Autònoma de Barcelona. Her main scientific interests have been focused in the field of the proteins of the ribonuclease A family. Her contributions include kinetics and structural studies directed to analyze catalytic mechanism and the contribution of non-catalytic binding sites in the cleavage of polymeric substrates and the structural bases and mechanisms associated to the biological functions of ribonucleases, especially the cytotoxicity directed to bacteria and eukaryotic cells.

Mohammed Moussaoui received his PhD in Biochemistry and Molecular Biology in 1998 at the Universitat Autònoma de Barcelona (UAB), Spain. Since 1999, he works at the Biochemistry and Molecular Biology Department (UAB) as a Researcher and Lecturer. His main research interests are in the catalytic and antipathogen mechanism of ribonucleases.

Supplemental data

Ester Boix*, Vivian A. Salazar, Marc Torrent, David Pulido, M. Victòria Nogués and Mohammed Moussaoui

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

Materials and methods

Molecular docking modelling was performed as previously described (Torrent et al., 2011). Briefly, the docking simulations were performed using *Autodock 4.2* (The Scripps Research Institute, La Jolla, CA, USA). The protein coordinates of the eight human RNases were taken from X-ray crystallography or NMR data. When no three-dimensional structure was available, the protein structures were predicted using the Swiss Modeller server (http://swissmodel.expasy.org/). For protein coordinates, the following crystal structures were used: RNase 1 (2K11.pdb), RNase2 (1H1H.pdb), RNase3 (1DYT.pdb), RNase 5 (2ANG.pdb) and models from nuclear magnetic resonance of RNase 4 (1RNF.pdb) and RNase 7 (2HKY.pdb), the RNases 6 and 7 three dimensional models were designed using Swiss Model software. The ligands d(ApTpApA) and heparin tetramer fragment were taken from protein data bank coordinates 1RCN.pdb and 1BFB.pdb, respectively; CpA and heparin IdoA(2S) GlcNS (6S) (5S) were designed using PRODRG server. As first step, all ribonucleases were superposed using Swiss-PdbViewer to the RNase A model (1RCN.pdb). In all crystal structures used, water molecules were removed from the structure. Hydrogen atoms were added using Autodock Tools. This software was also used to add atomic partial charges, using the Gasteiger method, and non-polar hydrogen merging. The protein was kept rigid during the docking simulation. The docking was accomplished using
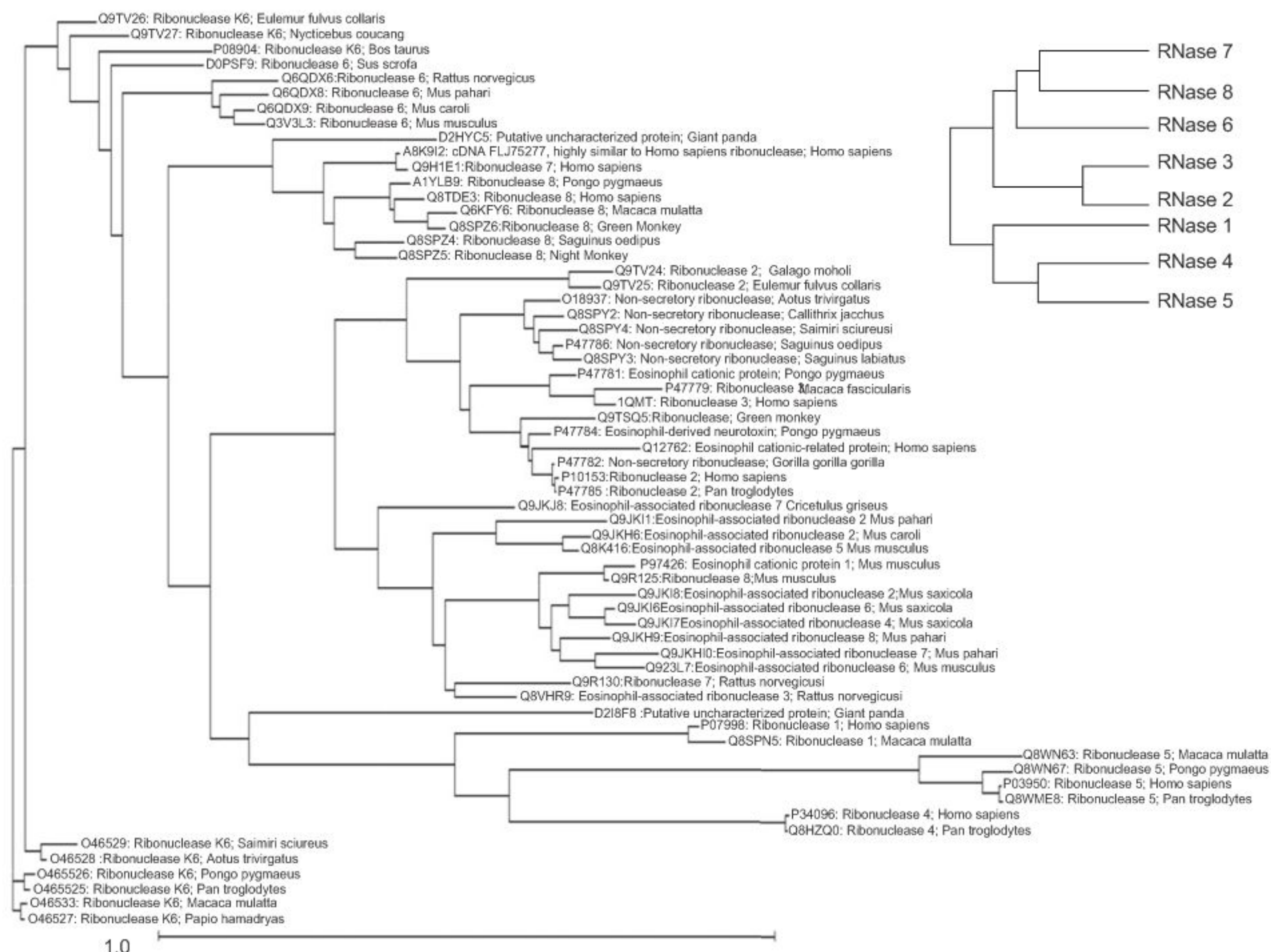
Figure S1 Phylogenetic tree among the RNase A family members used to prepare the evolutionary conservation score colouring, as provided in the CONSURF web server (http://consurf.tau.ac.il/). RNase 3 sequence was taken as a reference. The Protein database used for the Multiple Sequence Alignment (MSA) was Clean Uniprot (UniProt Knowledgebase). The tree was created using Rate4Site program from the CONSURF Server. Top right box includes the phylogenetic tree for the eight human RNases, as depicted in Zhang et al. (2002).

Table S1 Human RNases docking to CpA. The results include binding energy and potential hydrogen bond interactions between each RNase and CpA. The picture shows the free selected rotable bonds, out of 10. For protein coordinates, the following crystal structures were used: RNase 1, 2K11.pdb; RNase 2, 1H12.pdb; RNase 3, 1DYT.pdb; RNase 5, 2ANG.pdb; and models from nuclear magnetic resonance of RNase 4, 1RNF.pdb and RNase 7, 2HKY.pdb. The RNases 6 and 8 three dimensional models were designed using the Swiss modeller server (http://swissmodel.expasy.org), and CpA was built with the PRODRG software tool. All measures were calculated using the CONTACT program from the CCP4 interface (CCP4i, 1994) and checked using COOT 0.6.2 program (Emsley and Cowtan, 2004), selecting a maximum cutoff distance of 3.4 Å.
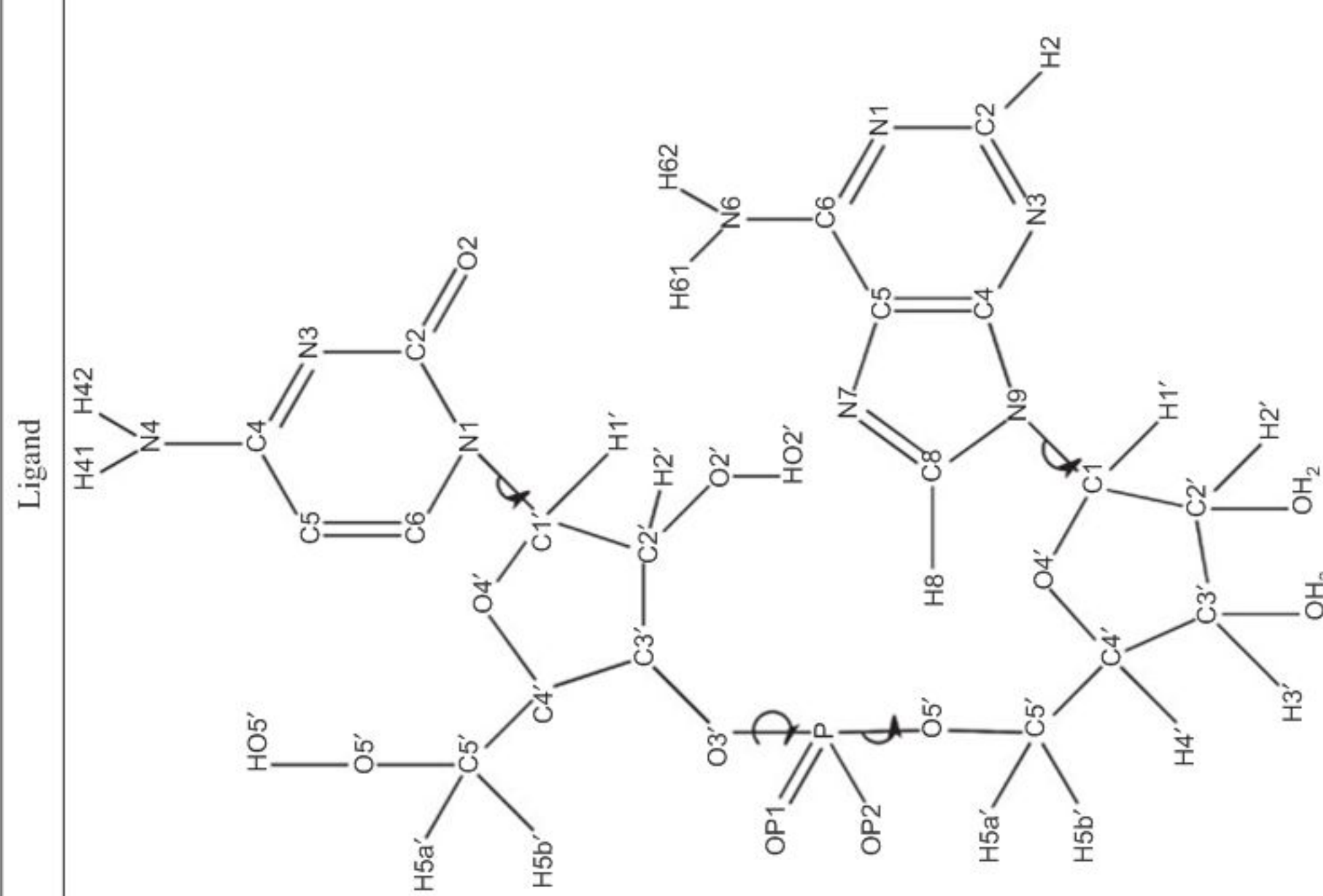
Protein	Binding energy	Interactions			Ligand
RNase 1	−21.58kJ/mol	Gln11 NE22:O3' Gln11 NE2:OP1	His12 NE2:O2' Lys41 HZ1:OP1	Asn44 ND2:O3' Phe120 O:N1 Asp121 O:N6	
RNase 2	−22.90 kJ/mol	Gln14 NE2:O4' His15 HE2:O4'	Ser64 OG:N7 His129 HE2:O3'	His129 ND1:N3 Leu130 O:N3 Asp131 O:N4	
RNase 3	−20.29 kJ/mol	Arg7 O:O2' Phe11 N:N4 Gln14 OE1:O2 Gln14 NE2:N3 Gln14 NE2: HN3	Arg7 O:H42 His15 HE2:O4' His15 HE2:O5' Lys38 HZ2:OP2 Val127 O:N1 Val 127 O:HN1	His128 HD1:N3 His128 HD1:O5' His128 ND1:OH5 Leu129 O:O5' Leu129 N:O'	
RNase 4	−24.97 kJ/mol	Gln11 OE:OP2 His12 NE2:O2 Lys40 NZ:O4' Asn43 OD:O2 Thr44 OG1:N4	Thr44 OG1:H42 Thr44 NH:O2 Asn66 ND2: N6 Asn66 HD22:H62	His116 ND1:O5' His116 ND:O4' His116 ND:N3 Phe117 N:OP1	
RNase 5	−16.44 kJ/mol	His8 NE2:O2' Gln12 HE22:OP1 His13 NE3:OP1 Lys40 NZ:P	Lys40 HZ1:OP1 Val113 O:N7 His114 HD1:N3	His114 ND1:N1 Aspl16 OD1:N3 Gln117 N:O2	
RNase 6	−13.30 kJ/mol	Lys7 O:O2' Gln14 HE22:P Lys38 HZ1:OP1 Asn64 ND2:H62 Asn64 ND2:N7	Asn64 OD1:N7 Arg65 NH2:N3 Arg65 NH2:HN3 Arg66 NH2:H41 Arg66 NH2:H42	His122 HD1:O5' His122 HD1:P His122 HD1:OP2 Val121 O:O5' Leu123 NH:OP2	
RNase 7	−23.47 kJ/mol	Ser7 OG:N6 Lys38 HZ3:OP2	His123 HE2:O5' Asp125 OD1:HN7	Asp125 OD1:N7	
RNase 8	−22.38 kJ/mol	Gln14 OE:O3' His15 NE2:O3' His15 NE2:OP1 Lys38 NZ:O4'	Thr42 OG1:N3 Thr42 HN:N3 Thr42 HN:O2 Val121 O:O4'	His122 ND:N7 His122 HD1:N7 Leu123 N:OP1 Leu123 HN:OP1	

Table S2   Human RNases docking to the tetranucleotide ApTpApA. The results include binding energy and potential hydrogen bond interactions between each human RNase and the ligand. The picture shows the three free selected rotable bonds, out of 21, from the ligand; coordinates were taken from 1RCN.pdb. For protein coordinates, the following crystal structures were used: RNase 1, 2K11.pdb; RNase 2, 1HI2.pdb; RNase 3, 1DYT.pdb; RNase 5, 2ANG.pdb; and models from nuclear magnetic resonance of RNase 4, 1RNF.pdb and RNase7, 2HKY.pdb. The RNases 6 and 8 three dimensional models were modelled using the Swiss Modeller. All measures were calculated using the CONTACT program from the CCP4 interface and checked using COOT 0.6.2 program, selecting a maximum cutoff distance of 3.4 Å.

Protein	Binding energy	Interactions						Ligand
		Protein	Nucleic acid	Protein	Nucleic acid	Protein	Nucleic acid	
RNase 1	-27.01 kJ/mol	Lys66 NZ	5'-A¹ N6	Val43 O	O3'	Arg4 HN2	5'-A³ H3'	[chemical structure of ligand ApTpApA]
		Thr45 N	5'-T² N3	His12 NE2	5'-A³ H3	Arg4 NE2	H3'	
		Thr45 H	O2	His12 HE2	N3	Gln11 OE1	O4'	
		His12 ND1	O2	Lys41 NZ	OP2			
		Thr45 OG1	O4	Lys41 HZ3	OP2			
				Lys41 NZ	O5'			
RNase 2	-6.39 kJ/mol	Glu14 OE1	5'-T² O4'	Asp41 OD1	OP2	Asn65 ND2	5'-A³ N3	
		His15 NE2	OP1	Thr42 N	OP2			
		Lys38 NZ	O4'	Val 128 O	O4'			
				Leu130 O	OP1			
RNase 3	-11.46 kJ/mol	Gln40 OE	5'-A¹ N3	His64 ND1	N3	Lys38 NZ2	O5'	
		Asp130 O	H3	His64 ND1	H3	Lys38 NZ2	OP	
		Thr131 HG1	N7	Arg1 N	5'-A³ N1	His64 HD	3'-A⁴ H3'	
		Asn39 OD1	5'-T² H7	Arg1 N	N6			
			OP	Lys38 NZ3	O4'			
RNase 4	-23.90 kJ/mol	Arg41 HN	5'-A¹ O3'	Gly3 O	N62	Asn66 ND2	3'-A⁴ O3'	
		Arg41 HN	5'-T² OP1	Arg7 HH21	OP2	Asn66 HD2	O3	
		Phe42 O	H3	Arg7 NH2	OP2	Asp118 OD1	H3	
		Gly3 O	5'-A³ H62	Gln11 HE11	P	Asp118 OD1	N3	
				Gln11 HE11	OP2			
				Lys40 HZ1	OP1			
				His116 HD1	O3'			
RNase 5	-3.52 kJ/mol	His114 HE2	5'-T² O2	His13 HE2	P	His114 HD1	O4'	
		Gln117 N	O2	Lys40 HZ1	OP2	Leu115 O	OP1	
		Arg121 HH22	O4'	Lys40 HZ3	O3'	Arg5 HH11	3'-A⁴ N3	
		His13 NE2	5'-A³ OP2	His114 ND1	N7	Gln12 NE2	P	
				His114 HD1	N7	Gln12 HE22	OP1	
				His114 HD1	N9			

Table S2 (Continued)

Protein	Binding energy	Interactions						Ligand
		Protein	Nucleic acid	Protein	Nucleic acid	Protein	Nucleic acid	
RNase 6	−30.10 kJ/mol		5'-A^1		5'-A^3	Leu123 N	OP2	
		Arg65 HH12	O3'	Pro2 N	N6		3'-A^4	
			5'-T^2	Gln14 NE2	OP1	Gln40 NE2	N7	
		Lys7 O	O2	Gln14 NE2	O4'	Thr42 N	H62	
		Arg65 NH2	OP2	Gln14 OE1	O4'	Thr42 OG	N6	
		Arg65 NH1	OP2	His15 NE2	OP1	Thr42 OG1	N1	
		Val121 O	O3'	His15 NE2	OP2	Arg65 NH2	O3'	
		His122 ND1	O3'	Lys38 NZ	O4'	Leu123 O	H3	
		His122 HD1	O3'	His122 ND1	OP2	Leu123 O	N3	
				His122 N	O3'	Asp124 O	N3	
RNase 7	−15.10 kJ/mol		5'-T^2	His123 NE2	O2	Trp10 HE1	3'-A^4	
		Lys38 NZ	O4'	His123 NE2	H3	Trp10 HE	H62	
		Lys38 HZ3	O4'		5'-A^3	His12 NE	N6	
		His15 NE2	O2	Trp10 NE1	H7	His123 NE2	OP2	
		His15 NE2	N3	His123 HE2	O3'	His123 HE2	OP2	
							OP2	
RNase 8	−6.73 kJ/mol		5'-A^1	Lys38 NZ	O3'	Leu123 O	O4'	
		Arg36 NE	N7		5'-A^3		3'-A^4	
		Arg36 NE	N9	His15 HE2	OP2	His122 HE2	OP1	
		Arg36 NE	O4'	Lys38 NZ	OP1	Leu123 O	O4'	
			5'-T^2	Asn41 OD1	OP1			
		Lys1 N	OP1	His122 NE2	OP2			
		Trp10 NE1	OP2					

237

Table S3 Human RNases docking to heparin IdoA(2S) GlcNS (6S). The results include binding energy and potential hydrogen bond interactions between each human RNase and heparin disaccharide. The picture shows the ligand-free 4 rotable bonds out of 13. For protein coordinates, the following crystal structures were used: RNase 1, 2K11.pdb; RNase 2, 1HI2.pdb; RNase 3, 1DYT.pdb; RNase 5, 2ANG.pdb and models from nuclear magnetic resonance of RNase 4, 1RNF.pdb and RNase 7, 2HKY.pdb, the RNases 6 and 8 three dimensional models were designed using the Swiss Modeller server, and the ligand was built using the PRODRG server. All measures were calculated using the CONTACT program from the CCP4 interface and checked using COOT 0.6.2 program, selecting a maximum cutoff distance of 3.4 Å.

Protein	Binding energy	Interactions	Ligand
RNase 1	-36.65 kJ/mol	Arg4 NH2:C6(SO$_3^-$); Lys7 HZ3:O6; Lys7 HZ1:O6; Gln11 OE1:C5'(CO$_2^-$); Gln11 HE22:C5'(CO$_2^-$); Gln11 HE22:O'; Asn34 O:HN; Met35 SD:C1(OH); Gln37 H:C2(SO$_3^-$); Lys41 HZ1:C3(OH); Lys41 HZ3:C3(OH)	
RNase 2	-38.26 kJ/mol	His15 NE2:C6(SO$_3^-$); His15 NE2:C5'(CO$_2^-$); Lys38 NZ:C5'(CO$_2^-$); Lys38 NZ:O'; Asn39 N:C2'(SO$_3^-$); Gln40 OE:C2'(SO$_3^-$); Thr42 HN:C5'(CO$_2^-$); His129 ND1:C6(SO$_3^-$)	
RNase 3	-39.24 kJ/mol	Arg1 NE:C2'(SO$_3^-$); Arg1 NH2:C2'(SO$_3^-$); Tyr33 OH:C5'(CO$_2^-$); Arg34 NH1:C6(SO$_3^-$); Lys38 HZ2:C6(SO$_3^-$); His64 ND1:C2(SO$_3^-$); Lys38 HZ1:C6(SO$_3^-$); His128 ND1: C1(OH)	
RNase 4	-33.34 kJ/mol	Arg7 HN2:C2(SO$_3^-$); Arg7 HN2:C3'(OH); Gln11 NE2:C2'(SO$_3^-$); Gln11 OE1:C2(HN); Gln11 OEI:C1(OH); His12 HEI:C1(OH); Lys40 NZ1:C1(OH); Arg41 N:C5'(CO$_2^-$); Phe42 NH:C5'(CO$_2^-$); His116 ND1:C2'(SO$_3^-$)	
RNase 5	-26.19 kJ/mol	His8 NE2:C2(HN); Gln12 HE22:C2'(SO$_3^-$); Gln12 OE1:C1(OH); Gln12 OE1:C2(HN); His13 NE2:C1(OH); Ile42 O:C5'(CO$_2^-$); Val113 O: C2(SO$_3^-$); His114 HE2:C5(SO$_3^-$); Leu115 O:C5(SO$_3^-$); Leu115 O:C6(O); Gln117 N:C5(SO$_3^-$); Gln117 HN:C5(SO$_3^-$)	
RNase 6	-33.76 kJ/mol	Pro2 N:C2'(SO$_3^-$); Trp10 NE1:O'; Trp10 HE1:O'; Gln14 OE1:C5'(CO$_2^-$); Lys38 NZ:C5(SO$_3^-$); Lys38 NZ:C5'(CO$_2^-$); Lys38 HZ3:C5'(CO$_2^-$); Arg65 HH12:C2(SO$_3^-$) Val121 O:C(HN)	
RNase 7	-33.17 kJ/mol	Ser7 OG:C2(HN); Gln14 OE:C5'(CO$_2^-$); Trp10 NE1:O'; Lys38 HZ1:C5'(CO$_2^-$); Lys38 NZ:C5'(CO$_2^-$); Lys107 HZ1:C2(SO$_3^-$); His122 ND: O; His123 HE2:C5(SO$_3^-$); His123 HE2:C5(SO$_3^-$)	
RNase 8	-40.48 kJ/mol	Trp10 NE1:C2(HN); Gln14 HE21:C2(SO$_3^-$); His15 NE:C2'(SO$_3^-$); Arg36 HH22:C6(O); Arg36 HH22:O; Arg36 NH22:O; Lys38 HZ1: C3'(OH); Lys38 NZ: C3'(OH); Leu40 O:C2(SO$_3^-$); Asn41 OD1:C3'(OH); Asp39 N:C5(SO$_3^-$); His122 ND1:C2'(SO$_3^-$); Leu123 N:C2'(SO$_3^-$)	

Table S4 Human RNases docking to heparin oligosaccharide. The results include binding energy and interactions between each human ribonuclease and heparin oligosaccharide. The ligand shows 3 rotable bonds of 27. Ligand coordinates were taken from 1BFB.pdb. For protein coordinates, the following crystal structures were used: RNase 1, 2K11.pdb; RNase 2, 1HI2.pdb; RNase 3, 1DYT.pdb; RNase 5, 2ANG.pdb and models from nuclear magnetic resonance of RNase 4, 1RNF.pdb and RNase 7, 2HKY.pdb. The RNases 6 and 8 three dimensional models were modelled using the Swiss Modeller Server. All measures were calculated using the CONTACT program from the CCP4 interface and checked using COOT 0.6.2 program, selecting a maximum cutoff distance of 3.4 Å.

Protein	Binding energy	Interactions		Ligand
RNase 1	-31.36 kJ/mol	Lys7 NZ:C5'(CO_2^-)I	Gln37 N:C2'(SO_3^-)I	Val43 O:C6(SO_3^-)II
		His12 HE21:C2'(SO_3^-)III	Arg39 NH2:C5(SO_3^-)II	Asn44 ND2:C6(SO_3^-)III
		Asn34 O:C2'(SO_3^-)I	Arg39 NE:C5(SO_3^-)II	His119 ND1:C2'(SO_3^-)III
		Met 35 O:C2'(SO_3^-)I	Lys41 HZ3:C3'(OH)I	Phe120 N:C2'(SO_3^-)III
		Gln37 O:C2'(SO_3^-)I	Lys41 HZ2:C5'(CO_2^-)III	Asp121 O:C3'(OH)III
		Gln37 H:C2'(SO_3^-)I		
RNase 2	$-21/76$ kJ/mol	Trp10 NE1: C5(SO_3^-)II	Arg36 NH1:C2'(SO_3^-)III	Thr42 OG1:C2 (SO_3^-) IV
		Gln14 NE2:C2'(SO_3^-)I	Arg36 NH2:C2'(O)III	His129 ND1:C2'(SO_3^-)I
		His15 NE2: C2'(SO_3^-)I	Asn39 N:C2'(SO_3^-)III	Leu130 N:C2'(SO_3^-)I
			Gln40 OE1:C3(OH)IV	Asp131:C5'(CO_2^-)I
RNase 3	-17.49 kJ/mol	Arg1NH2:C2'(SO_3^-)III	Gln40 O:C6' I	Asp112 OD:C2 (SO_3^-) IV
		Arg1 N:C2(SO_3^-) II	Asn41 OD:C6(CO_2^-)I	Val127 O:C6(CO_2^-)III
		Arg7 HH1:C6(SO_3^-)II	Asn41 ND2:C6'(CO_2^-)I	His128 ND1:C3'(OH)I
		His15 NE2:C3'(OH)I	His64 NE2:C2'(SO_3^-)I	Leu129 H1:C3'(OH)I
		His15 NE2:C3'(OH)I	Ala110 O:C2 (SO_3^-) IV	Asp130 O:C2'(SO_3^-)I
RNase 4	-25.84 kJ/mol	Arg7 NH1:C2(SO_3^-)IV	Met34 O: C5(SO_3^-)IV	Phe42 N:C2(SO_3^-)II
		Arg7 NH1:C3(OH)IV	His38 O:C5(SO_3^-)IV	His16 ND2:C2'(O)I
		Arg7 HH22:C3(OH)IV	His38 HD:C5(SO_3^-)IV	His16 ND:O1
		Arg7 NH2:C3(OH)IV	Lys40 HZ1:C3(OH)II	Phe117 HN:C5'(CO_2^-)I
		Arg7 NH2:C5'(CO_2^-)III	Phe42 HN:C2(SO_3^-)II	Phe117 N: C5'(CO_2^-)I
RNase 5	-15.80 kJ/mol	Arg5 HH1:C2'(SO_3^-)I	Pro38 O:C5(SO_3^-)IV	Val13 O:C2'(O)I
		His8 NE2:C4'(OH)I	Lys40 N:C5(SO_3^-)IV	His14 ND1:C2'OI
		Gln12 OE1:C3'(OH)I	Lys40 HZ3:C5'(CO_2^-)III	His14 HD1:C2'OI
		His13 HE2:C5' (CO_2^-)I	Lys40 NZ:C5'(CO_2^-)III	Leu15 N:C5'(CO_2^-)I
		Leu35 O:C5(SO_3^-)IV		
RNase 6	-15.47 kJ/mol	Pro2 HN1:C2(SO_3^-)II	Gln40 NE2:C2(SO_3^-)IV	Arg65 NH2:C2'(SO_3^-)I
		Lys7 O:C5'(CO_2^-)I	Gln40 HE2:C2(SO_3^-)IV	Val121 O:C2'(SO_3^-)III
		Trp10 NE1:C2(SO_3^-)II	His15 NE2:C5(SO_3^-)II	His122 HD1:C2'(SO_3^-)I
		Gln14 NE2:C5(SO_3^-)II	His15 HE2:C5(SO_3^-)II	Leu123 HN:C5(SO_3^-)II
		Gln14 NE2:C5(SO_3^-)I	His36 NE2:C2'(SO_3^-)III	
RNase 7	-14.72 kJ/mol	Trp10 HE1:C5'(CO_2)III	Lys38 HZ2:C3'(OH)I	His123 NE2:(O)II
		His15 HE2:C5(SO_3^-)II	Leu40 H:C2'(SO_3^-)I	His123 HE2:(O)II
		Lys38 HZ1:C3'(OH)I	His123 NE2:C5(SO_3^-)II	Leu124 O:C5(SO_3^-)II
RNase 8	-5.31 kJ/mol	Trp10 HE1:C3'(OH)I	Arg36 HH12:C5(SO_3^-)IV	Val121 O:C5'(CO_2)I
		His15 NE2:C5'(CO_2^-)I	Arg36 HH12:(O)IV	His122 HD1:(O)I
		His15 HE2:C5'(CO_2^-)I	Arg36 HN1:(O)IV	His122 HD1:C5(SO_3^-)II
		Thr34 HG1:C2(SO_3^-)IV	Asp39 N:C2(SO_3^-)IV	His122 ND1:C5(SO_3^-)II
				Leu123 N:C5'(CO_2)I

150 Lamarkian genetic algorithms (LGA) runs, and the initial position of the ligand was random. Binding energies are estimated for a cluster population within 2.0 Å of RMSD tolerance.

References

CCP4i. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Torrent, M., Nogues, M.V., and Boix, E. (2011). Eosinophil cationic protein (ECP) can bind heparin and other glycosaminoglycans through its RNase active site. J. Mol. Recognit. *24*, 90–100.

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

Chapter VI

Searching for Heparin Binding Partners

Ester Boix[1,], Marc Torrent[1,2], M. Victòria Nogués[1],*
Vivian A. Salazar[1]

[1]Department of Biochemistry and Molecular Biology, Universitat Autònoma de Barcelona, Biosciences Faculty, 08193, Cerdanyola del Vallès, Spain [2]Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Biomedical Research Park of Barcelona, Aiguader 80, 08003, Barcelona, Spain

Abstract

Heparin modulates diverse key physiological processes, from coagulation to angiogenesis, also contributing to the host –pathogen recognition process. New potential roles are still emerging, such as the inhibition of amyloidogenesis. To fully understand heparin function, and eventually modulate its action, we must first find out who are the travel companions.

The chapter presents an overview on the heparin binding proteins known up to date. Following, the review offers useful tools to assist the discovering of new binding molecules, and identify the structural determinants involved in the interaction.

Heparin complexes can be studied by experimental and *in silico* approaches, including docking and molecular dynamics simulations. Next, protein-ligand interfaces can be further analyzed by chemical-based computational tools to redesign even discontinuous binding epitopes in order to develop new active drugs.

New *in silico* developed leads can be efficiently synthesized by high-throughput and combinatorial chemical synthesis allowing the screening of thousands of compounds. Thenceforth, they can be tested by experimental screening analysis to refine the leads designed, which once validated can undergo first clinical trials.

* Corresponding Author: Ester Boix. Departament de Bioquímica i Biologia Molecular, Facultat de Biociències, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Spain. E –mail: ester.boix@uab.cat; Phone: 34-93-5814147. Fax: 34-93-5811264.

Introduction

Glycosaminoglycans (GAGs) are complex carbohydrates that participate in key cell biological processes by regulation of their binding protein partners [1]. They are linear negatively charged polysaccharides with a molecular weight varying from 10 to 100 kDa and a high heterogeneity arises from all the enzymatic reactions involved in their biosynthesis pathways. In this chapter we will focus on the sulfated polysaccharides: heparin and heparan sulfate (HS). Heparin is synthesized by mast cells, whereas HS can be produced by virtually all type of cells. On its turn, heparin structure is highly conserved with similar composition in a broad range of vertebrate and invertebrate organisms, while HS can be found as components of the heparin sulfate proteoglycans (HSPGs) in the extracellular matrix and in the surface of animal cells.

Heparin and HS are constituted by repeating units of GlcNAcα1–4GlcAβ1–4, which can undergo subsequent modifications (Figure 1). HS chain can vary in the extent of sulfation and epimerization [2] and heparin can in fact be described as a highly sulfated type of HS. On the other hand, HS contains a higher level of acetylated glucosamine. Each heparin disaccharide contain between 2 and 3 sulfate groups, giving together with carboxyl groups a very anionic character. Indeed, heparin is considered the biological molecule with the highest negatively charged density [3].

Heparin is a linear molecule that tends to have an extended conformation in solution because of its highly hydrophilic nature. The three dimensional structure of heparin was deduced from NMR studies. The structure is not unique, as sugars tend to adopt several conformations. NMR data indicates the presence of three possible conformers: the $^4C_1$ and the $^1C_4$ chairs and the $^2S_0$ skew boat. Mostly, the IdoA2S form shows a predominance over the chair ($^1C_4$) and the skew-boat ($^2S_0$) [4]. The molecule adopts an helical conformation, where clusters of sulfate groups align at both sides at regular intervals. The NMR structure of an heparin dodecasaccharide has often been used as a reference for modeling and docking analysis (1HPN [5]). However, crystal structures of protein complexes with short oligomers suggest that when heparin binds to a protein, a conformational change is induced for a better fit [1].

A Bit of History

Heparin was discovered in 1916 at Johns Hopkins University by a medical student, Jay McLean [3]. In 1928 one of the sugars was identified as uronic acid and in 1935 glucosamine was found as the second sugar component. Later on, it was establish that heparin was a polysaccharide containing a very high percentage of covalently linked sulfates, making it one of more acidic known biological macromolecule. Heparin name derives from the first isolation source: canine liver cells (*hepar* is liver in greek) [6, 7]. By 1935 pure heparin was available for clinical tests and the compound was found effective in preventing postoperative thrombosis. The observed anti-coagulation activity was in fact due to its interaction with the thrombin inhibitor (antithrombin III, AT-III). Pharmaceutical heparin is now usually derived from bovine lung or porcine intestinal mucosa. The development of better heparin derived anticoagulant drugs is very active and most drug development research is based on the

sulfated glycosaminoglycans (GAGs): heparin and heparan sulfate [8]. The undesirable side effects of heparin, such as bleeding complications, and the better understanding of the coagulation cascade, led to the isolation and analysis of low molecular weight heparin derivates with better defined properties [3]. Several heparin –mimetic oligosaccharides are currently under clinical trials [8].

Heparin
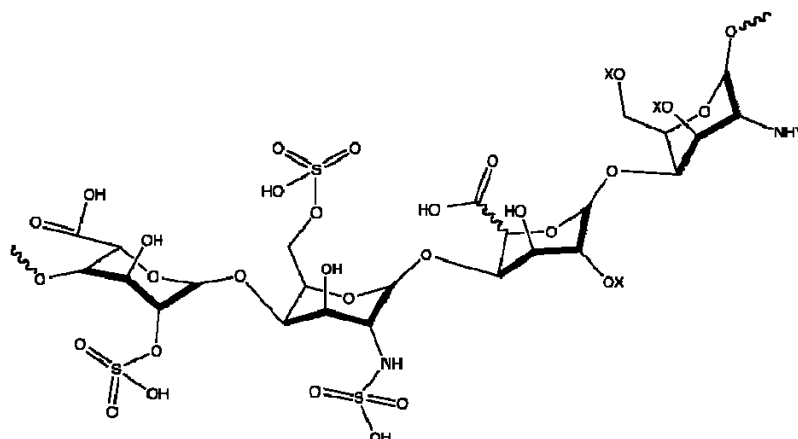
Heparan sulfate
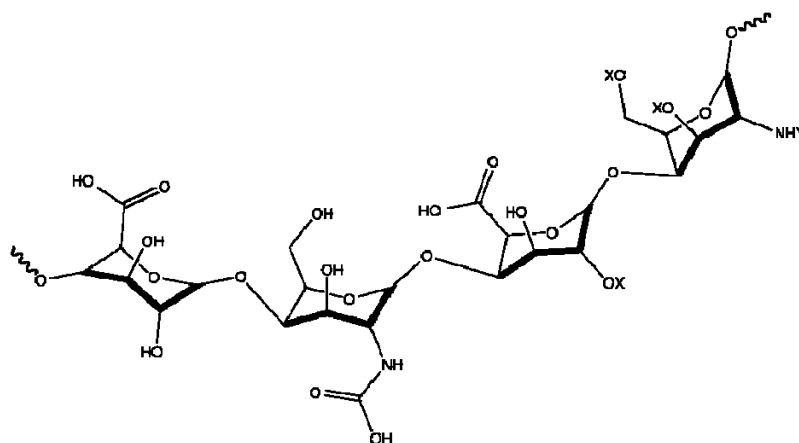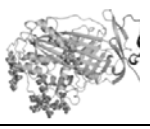
Figure 1. Disaccharide repeating units in heparin and heparan sulfate (X= H or SO_3^-, Y = Ac, SO_3, or H). Scheme was modified from [3].

Looking for Identity Patterns

As heparin properties depend on its binding partners, we must first identify which are the structural determinants that guide their selective affinity. Heparin protein interaction is mainly driven by ionic interactions, dependant on the specific arrangement of sulfated residues. Among glycosaminoglycans, heparin and HS have a greater chemical diversity and capacity to interact with proteins through their varied arrangements of sulfate groups and glucuronic acid/iduronic acid residues [3], as depicted in figure 1.

A database of glycosaminoglycan binding proteins (GAGPROT) that collected all the available protein- ligand complexes coordinates deposited at the Protein Data Bank was created [9]. The database was part of a global glycomics project (Glyco3D) which also included an oligosaccharide database. We include here in this chapter an updated and exhaustive list of the currently available protein structures in complex with heparin and heparan derivates (Table 1). One of the first and best-studied example is antithrombin complex with an heparin pentamer (Figure 2).

Table 1. Three-dimensional structures of protein complexes with heparin and heparan sulfate oligosaccharides

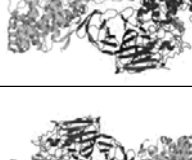| Protein name	Protein type	Ligand	pdb code[1]	Complex structure[2]	Ref.
Antithrombin	Serpin (Serine protease inhibitor)	Heparin pentasaccharide	1AZX (2.9 Å)		[43]
Antithrombin	Serpin (Serine protease inhibitor)	Heparin pentasaccharide	1E03 (2.9 Å)		[60]
Antithrombin	Serpin	Heparin pentasaccharide	1NQ9 (2.6 Å)		[61]
Antithrombin	Serpin (Serine protease inhibitor)	Trisulfoamino heparin pentasaccharide	3EVJ (3.0 Å)		Unpubl.
Antithrombin-prothrombin	Serpin/ protease	Heparin heptasaccharide	1SR5 (3.1 Å)		[62]

244
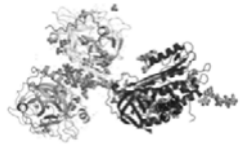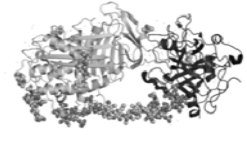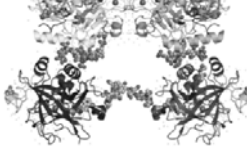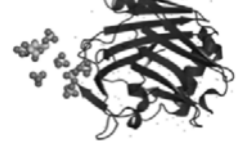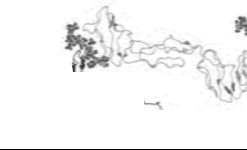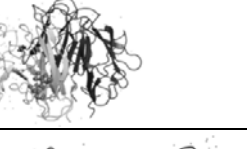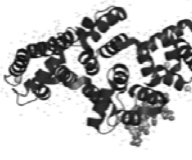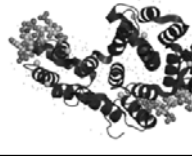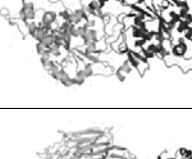
Antithrombin-thrombin	Serpin/protease	Heparin mimetic SR123781	2B5T (2.1 Å)		[63]
Antithrombin-thrombin	Serpin/protease	Heparin hexadecasaccharide	1TB6 (2.5 Å)		[44]
Antithrombin-Factor X	Serpin/protease	Synthetic heparin pentasaccharide (Fondaparinux)	2GD4 (3.3 Å)		[64]
Antithrombin III- Factor IXa	Serpin/protease	Heparin pentasaccharide	3KCG (1.7 Å)		[65]
Thrombin	Protease	Heparin disaccharide	1XMN (1.85 Å)		[66]
Thrombospondin-1	Membrane glycoprotein	Synthetic heparin pentasaccharide Arixtra® (Idraparinux)	1ZA4 (1.9 Å)		[67]
Vaccinia Complement Protein	Inhibition of the host inflammation response	Heparin octasaccharide	1RID (2.1 Å)		[68]
Complement C1Q subunit	Subunit of the C1 enzyme complex	Heparan sulfate tetrasaccharide	2WNU (2.3 Å)		[69]
RANTES	Chemokine	Heparin disaccharides (I-S and III-S)	1U4L (2.0 Å) 1U4M (2.0 Å)		[70]

Table 1. Continued

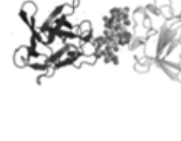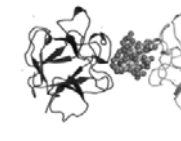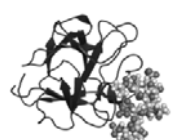| Protein name	Protein type	Ligand	pdb code[1]	Complex structure[2]	Ref.
Stromal Cell-derived Factor-1α (CXCL12)	Chemokine	Heparin disaccharide	2NWG (2.07Å)		[71]
Annexin A2	Extracellular protein	Heparin tetrasaccharide	2HYU (1.86 Å)		[49]
Annexin V	Extracellular protein	Heparin tetrasaccharide	1G5N (1.9 Å)		[72]
Sulfotransferase	Enzyme	Heparin, tetrasaccharide	1T8U (1.95 Å)		[73]
Heparinase II	Enzyme	Heparin disaccharide	2FUT (2.3 Å)		[74]
Heparinase II mutant	Enzyme	Heparan sulfate tetrasaccharide	3E7J (2.1 Å)		[75]
Heparinase II	Enzyme	Heparan sulfate disaccharide	3E80 (2.35 Å)		[75]
Heparin lyase I	Enzyme	Heparin disaccharide	3IN9 (2.0 Å)		[76]

246
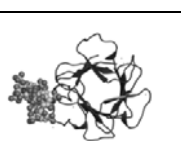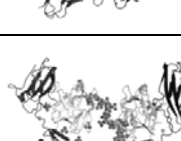
Heparin lyase I mutant	Enzyme	Heparin dodesaccharide	3INA (1.9 Å)		[76]
Human Acidic Fibroblast Growth Factor	Growth factor	Heparin pentasaccharide	1AXM (3.0 Å)		[77]
Human Acidic Fibroblast Growth Factor	Growth factor	Heparin hexasaccharide	2AXM (3.0 Å)		[77]
Human fibroblast growth factor (FGF-1)	Growth factor	Heparin synthetic hexasaccharide	2ERM (NMR)		[78]
Human Basic Fibroblast Growth Factor (bFGF)	Growth factor	Heparin tetrasaccharide	1BFB (1.9 Å)		[79]
Human Basic Fibroblast Growth Factor (bFGF)	Growth factor	Heparin hexasaccharide	1BFC (2.2 Å)		[79]
bFGF2/ecto-domain of FGF receptor 1	Growth factor/ receptor	Heparin hexasaccharide octasaccharide	1FQ9 (3.0 Å)		[80]
aFGF1/ecto-domain of FGF receptor 2	Growth factor/ receptor	Heparin decasaccharide	1E0O (2.8 Å)		[45]
Nk1 hepatocyte growth factor	Growth factor	Heparin pentasaccharide	1GMN (2.3 Å)		[81]

247

Table 1. Continued

Protein name	Protein type	Ligand	pdb code[1]	Complex structure[2]	Ref.
Nk1 hepatocyte growth factor	Growth factor	Heparin hexamer	1GMO (3.0 Å)		[81]
Eosinophil-granule major basic protein	Lectin C type	Heparin disaccharide	2BRS (2.2 Å)		[82]
Cardiotoxin A3	Snake venom toxin	Heparin hexasaccharide	1XT3 (2.4 Å)		[83]
Foot-and-Mouth Disease Virus-Oligosa-ccharide Receptor	Picornavirus capsid protein	Heparin pentasaccharide	1QQP (1.9 Å)		[84]
Human papilloma virus18 (HPV18) capsid L1	Virus capsid protein	Heparin trisaccharide	3OFL (3.4 Å)		[85]
Peptidoglycan recognition protein	Receptor for bacteria recognition	Heparin disaccharide	3OGX (2.8 Å)		Unpubl.

[1] Data resolution is indicated below.

[2] Assemblies of complex structures are depicted as defined by the authors or predicted by the *PDBe* PISA server [86]. Images were taken from the *PDBe* server.

In 1989 Cardin and Weintraub published the first study aimed at identifying the structural requirements for GAGs interaction [10]. The analysis of the available complexes at that time lead to the proposal of two binding consensus sequences: XBBXBX and XBBBXXBX [B is a basic and X a hydropathic (neutral and hydrophobic) amino acid residue], which would align respectively the basic group in beta and alpha-helix structures respectively for interaction with sulfated groups. Cardin-Weintraub consensus sequences is indeed present in many heparin –binding proteins, as Annexin, Vitronectin, AT-III, ApoE or the Protein C inhibitor [1]. A third consensus sequence was later proposed, XBBBXXBBBXXBBX, serving as initial sequence probe to determine whether a protein can possibly bind heparin [3]. However, the model was not always fitting and spacial orientation of basic residues should also be considered. By inspection of heparin binding sites with known secondary structure, a spacing

248

of 20 Å between cationic residues was proposed. By using peptide libraries Fromm and coworkers [11] assessed the affinity of randomly synthesized heptamers for both heparin and HS, concluding that Lys and mostly Arg, together with polar residues, are preferred. In particular, the guanidinium group of Arg can provide a specific tight interaction with the sulfo groups, as observed in the antithrombin-pentamer interactions (Figure 2). In fact, the Cardin and Weintraub motif shows a predominant preference for SO_4 binding, as shown by statistics analysis upon all currently protein-ligand complexes, when performed using the *PDBe motif* software [12]. Although main interaction between protein and heparin would be of ionic type with sulfo and carboxyl groups, the contribution of hydrogen bonds with hydroxyl groups should also be considered [13]. A second ligand binding hit for the heparin binding motif, as revealed by *PDBe motif* analysis is the phosphate group. Heparin resembles nucleic acids as both are highly charged linear polymers, and not surprisingly several heparin binding proteins were reported to bind to nucleotides and the other way round, nucleic acid binding proteins show a high affinity for oligosaccharides.
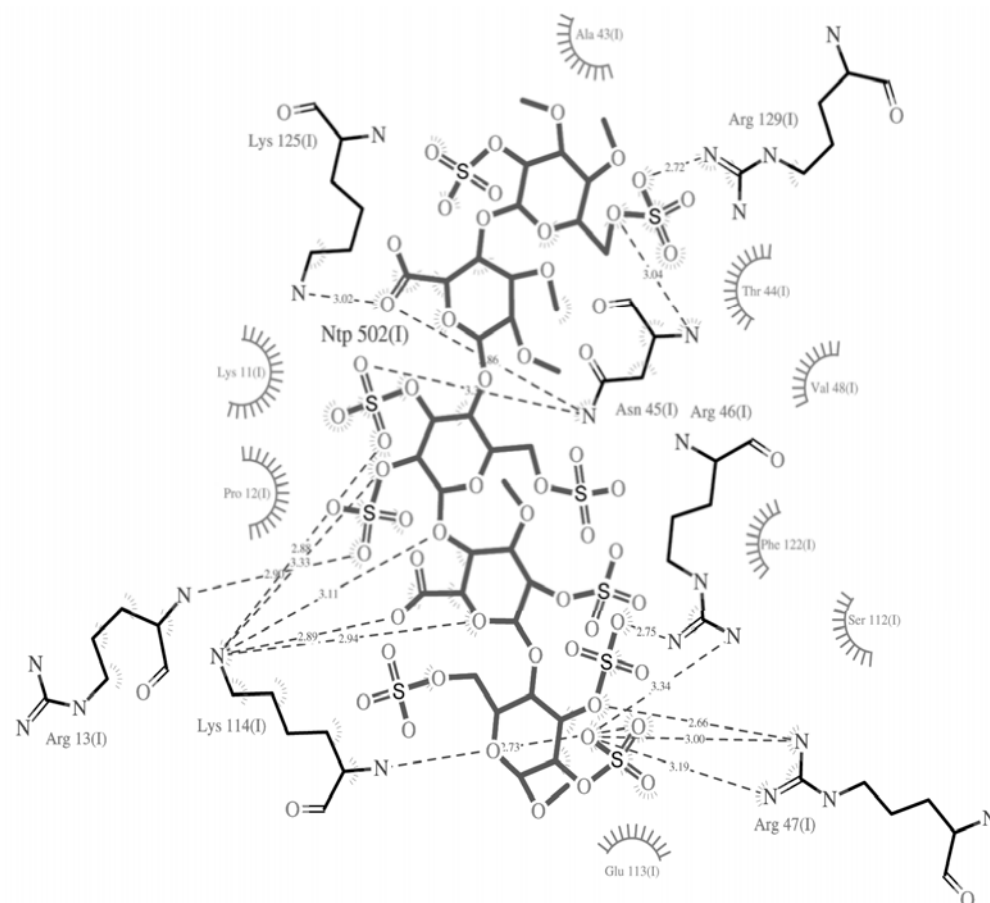

Figure 2. Protein- ligand interactions at the binding site of the antithrombin- heparin pentamer complex (1azx.pdb, [43]). The figure was drawn using *LigPlot* and provided by the *PDBe* server.
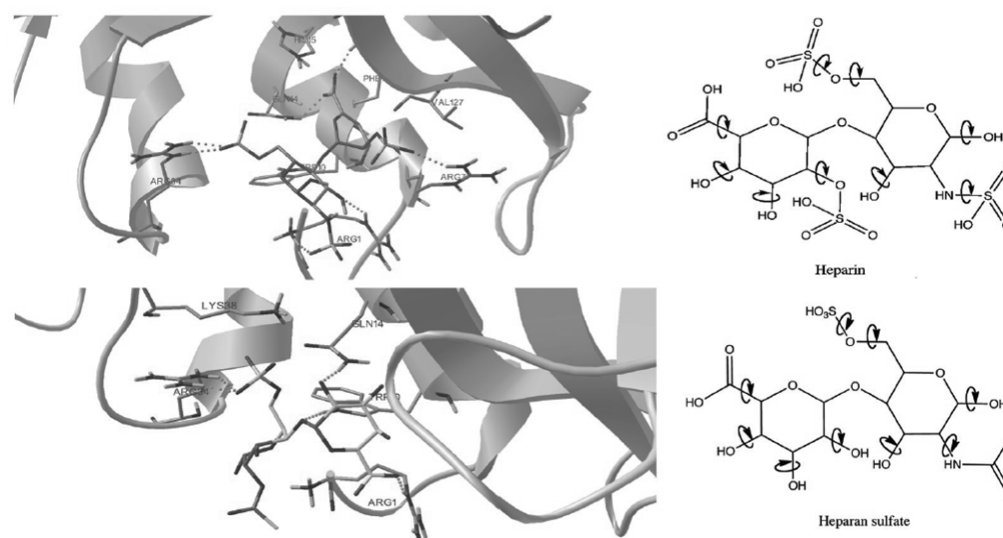
249

Figure 3. Docking results for heparin [IdoA(2S)-GlcNS(6S)] and heparan sulfate [IdoA-GlcNAc(6S)] [18]. The best ligand conformation of each cluster is selected. Rotable bonds are depicted and hydrogen bonds are indicated with dashed lines.

As an illustrative example we include here the comparative analysis of RNase 3 binding mode by both structural and modeling studies (Figs. 3 and 4). RNase 3, also called the Eosinophil Cationic Protein (ECP) is an eosinophil RNase of the RNase A superfamily involved in host defense [14]. It is a very cationic protein, with an unusual high pI (~ 11). RNase 3 was already reported to bind heparin when first purified from eosinophils [15]. Binding to heparin oligosaccharides was also observed later on when using the recombinant protein and mutagenesis studies identified a binding stretch required for interaction [16, 17]. Interestingly, the protein RNase catalytic activity is inhibited by the presence of heparin [18]. NMR analysis [19] and docking studies (Figure 3, [18]) confirmed subsequently that protein residues at the active site cleft [20, 21] were involved in the sugar binding. Figure 4 illustrates the overlapping of residues involved in the RNA substrate anchoring, taking the RNase A- tetranucleotide complex as a model [22], and RNase 3 docking to an heparin tetramer.

Cationic RNases binding to heparan sufate by Arg residues was suggested to promote the protein internalization in host cells, and may facilitate their cytotoxic side effects [23]. The virulence of pathogens such as protozoa also depends on HS binding at the mammalian cell surface and HS are used as a receptor for virus entry.

Critical Arg residues are also contributing to RNase 3 antimicrobial activity [24]. An Arg cluster was also identified at lactoferrin N-terminus which could anchor not only heparin and DNA, but also the lipopolysaccharides (LPS) present at the Gram –negative species outer membrane, accounting for its antimicrobial activity [25]. Screening the RNase 3 primary sequence for an antimicrobial activity motif [26] identified the region 33 to 38, which was also involved in LPS binding [27, 28]. Patterns for LPS have been proposed, bearing a striking similarity with the ones related to heparin binding, as referred above. In particular, the following stretches were reported by Frecer and coworkers [29]: BH(P)HB; HBHPHBH and HBHBHBH, where B, H and P stand for basic, hydrophobic and polar residues respectively.
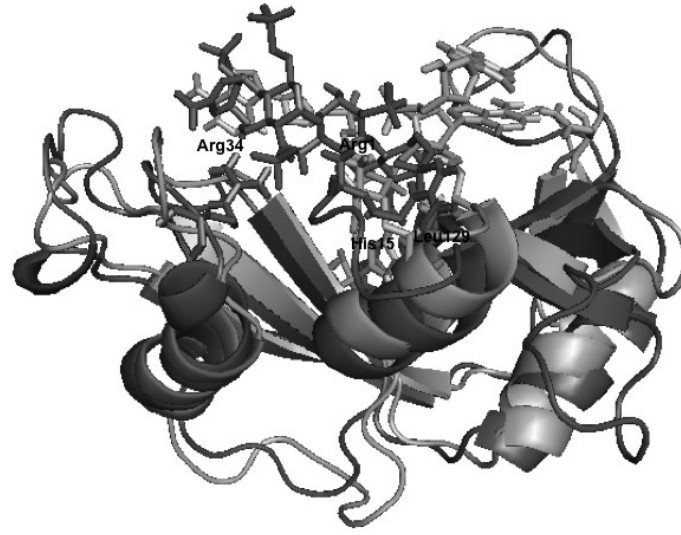
Figure 4. Structure of RNase A-d(ApTpApApG) Complex [22] in grey and Eosinophil Cationic Protein (ECP)[20] docking to tetrasaccharide (IdoA-GlcN-IdoA-GlcN) in black (1BFB.pdb). ECP interacting residues are labeled.

Searching for Methodological Tools

Many approaches can be used in order to discover, identify and characterize heparin-binding sites in proteins [30]. The techniques employed can be classified depending on the information provided (e.g. the resolution achieved) as summarized in Table 2.

Table 2. Summary of selected methods to study the interaction of heparin with proteins grouped in experimental and theoretical methods

	Atomic resolution	Epitope characterization	Binding affinity*
Experimental	Nuclear magnetic resonance (NMR) spectroscopy Protein crystallography	Proteolytic footprinting	Heparin affinity chromatography Equilibrium competition binding Isothermal fluorescence titration
Theoretical	Molecular docking simulation Molecular dynamics simulation	Shape complementary algorithms	Molecular docking simulation Molecular dynamics simulation

* Binding affinity assays can be used together with protein mutants to scan for heparin epitopes.

251

The experimental data can provide information on the binding affinity for heparin and even offer a kinetic picture of the association/dissociation process and, in more complex experimental setups, give atomic resolution details of the protein-heparin complex. In its turn, computational methods can easily provide atomic resolution details. Though inexpensive, computational results are not as reliable as those provided by experimental methods.

Experimental Methods for Characterizing Heparin-protein Binding Interactions

Heparin Affinity Chromatography

The main purpose of heparin affinity chromatography is to provide a rough estimation of heparin-protein affinity using a fast and easy experimental setup. First, the protein is applied to an heparin sepharose column and is eluted using a NaCl gradient from 0 to 1M NaCl. The NaCl concentration at which the protein is eluted denotes the protein affinity for heparin.

The protocol can also be applied to characterize the binding of engineered protein mutants with respect to the wild type. In that case, the measure of the NaCl concentration that promotes protein elution is a measure of the impact of the mutated residues on heparin affinity. Using these values the specificity index ($\Delta\Delta[NaCl]$) can be calculated as:

$$\Delta\Delta[NaCl] = \Delta[NaCl]^H - \Delta[NaCl]^S$$

The superscripts H and S refer to the chromatography type, heparin-sepharose and sepharose column respectively and:

$$\Delta[NaCl]^H = \Delta[NaCl]^H_{wt} - \Delta[NaCl]^H_{mut}$$

$$\Delta[NaCl]^S = \Delta[NaCl]^S_{wt} - \Delta[NaCl]^S_{mut}$$

Low amounts of protein are required to calculate the specificity index. However, sepharose chromatography assays are biased towards electrostatic interactions and do not give a real value for the affinity constant [30].

This method has been extensively used in the characterization of chemokines, to study the importance of cationic residues (e.g. lysine and arginine) for heparin interaction. As an example, Severin et al. [31] studied the heparin specificity index for several mutants of the CXCL11 chemokine. In addition, they observed that the results correlate well with a model developed using molecular docking simulations (see Molecular Docking Simulations section). Moreover, the results obtained together with equilibrium competition assays (see Equilibrium competition assays) and other related tests allowed the authors to propose a cationic binding patch on the surface of CXCL11.

Equilibrium Competition Binding and Tritiated Heparin Binding Assay

Although the heparin affinity chromatography gives a rough idea on the protein binding affinity for heparin, more detailed studies have to be conducted in order to calculate real binding constants that cannot be inferred from sepharose chromatography. Equilibrium competition binding is a classical assay that can provide accurate binding affinities for protein-ligand complexes. In this test, immobilized heparin in sepharose beads are incubated with radioactively labeled protein and the binding competition is followed using an increasing concentration of heparin or unlabeled protein. In both cases, the competitor will displace the radiolabeled protein bound to heparin. As sepharose beads can be easily isolated by filtration, the radioactivity recovered together with heparin-sepharose beads can be measured in a scintillation counter. As a result, affinity constants can be determined by this method and, when protein mutants are available, it can be used to infer protein-binding epitopes.

On the other hand, the tritiated heparin-binding assay can be used instead of the equilibrium competition assay; although it is less reliable and larger amounts of protein are required [30]. In this method, radioactive heparin is incubated with increasing concentrations of protein in 96-well plates fitted with cellulose phosphate paper. After washing the papers with an appropriate buffer to remove unbound heparin, the total amount of protein can be quantified in a scintillation counter.

Tritiated heparin binding assays have been performed by Lau et al. [32] to determine the amino acids defining the heparin-binding region of CC chemokine MCP-1 and found that His66 and the initiating Met had an active role in heparin binding. These residues could not be detected using heparin affinity chromatography.

Isothermal Fluorescence Titration and Surface Plasmon Resonance

Isothermal fluorescence is a biophysical technique that can be used to characterize heparin binding by monitoring the fluorescence of aromatic residues in the protein. Usually, the tryptophan fluorescence signal, rather than tyrosine or phenylalanine, is used because of its higher fluorescence quantum yield. The binding isotherms built with this method can be used to calculate apparent binding constants. For example, Lau et al. calculated the dissociation constants for the interaction of MCP-1 mutants with unfractionated heparin [32]. When no aromatic residue is present in the sample, the protein can be engineered to replace a residue for a Trp if that does not affect to the protein biological function.

Surface plasmon resonance (SRP) is a powerful system to study protein-ligand interactions and, particularly heparin-protein interactions. To describe the method briefly, the absorption or desorption of proteins in the sensor surface induces a change in the refractive index that is proportional to the protein concentration. The heparin substrate is immobilized on the chip and the solution containing the protein is constantly flowed over the chip. Heparin immobilization can be achieved by biotinylation using streptavidin sensor chip surfaces. The associated kinetics profile is followed to obtain the binding constant, detected as a change in protein concentration as a function of time. Finally, buffer is flowed over the chip to follow the dissociation.

SPR sensor chips SA, C1 and CM5 have been employed to determine the kinetic parameters of lymphotactin binding to immobilized heparin and used to show that mutation of residues Arg23 and Arg43 to Ala decrease the protein affinity for heparin 300-fold [33].

Proteolytic Footprinting Coupled to Mass Spectrometry

In order to identify protein regions involved in heparin binding, a combination of protein digestion with proteases and mass spectrometry (MS) can be used. The protein of interest is incubated with or without heparin and subsequently digested by trypsin. The peptide MS pattern is recorded and compared for both incubated and non-incubated protein. Regions that interact with heparin are protected from digestion and are identified. Furthermore, the resolution of this method can be improved by digestion with complementary proteases such as chymotrypsin [30].

Proteolytic footprinting can be used as a complementary assay to confirm heparin- binding regions in proteins. As a matter of example, Severin et al. have used proteolytic digesting combined with MS spectrometry to identify the region [53]CLNPKSKQAR[62] as a putative heparin-binding region for chemokine CXCL11 [31].

Computational Methods for Characterizing Heparin-protein Binding Interactions

Molecular Docking Simulations

Molecular docking is a computational approach designed to predict the binding sites of interacting molecules, from ligand-to-receptor binding to protein-protein complex interactions, with the aim to elucidate drug candidates or study interaction networks.

Up to date, only few molecular docking methods have been explicitly developed to predict protein-heparin complexes [5, 34, 35]. Molecular docking of reasonable heparin fragments ($\geq$ 4 monosaccharide units) to a protein-binding site is still difficult due to the high electrostatic charge carried in both the ligand and the protein. Moreover, the flexibility displayed by the glycosidic rings and saccharide linkages also contributes to difficult the procedure. In order to low the computation cost of docking heparin oligosaccharides in proteins, ligands can be treated as semi-rigid bodies, with no rotations allowed in the glycosidic rings and linkages. Although this procedure has been found not to affect the precision of the models obtained, it may not be adequate (e.g. precision may not be enough) for some models. In fact, the highly sulfated iduronic acid ring of heparin may induce some flexibility to the oligosaccharide chain. Also the plasticity on protein-heparin complexes due to NA-domains is important in multivalent interactions [34]. Thus, taking into account all degrees of freedom on protein-heparin complexes must be considered to obtain a reliable model.

One of the most used docking algorithms is *Autodock* but other algorithms such as DOCK have been successfully used to predict heparin-binding sites. Here, we briefly describe the docking procedure of heparin ligands in *Autodock* [18].

First, water molecules are removed from the structure, hydrogen atoms and atomic partial charges assignment and non-polar hydrogen merging can be added using *Autodock Tools*. The assignment of atomic partial charges can also be done using the AMBER force field or the REST procedure (available at the RED web server; http://q4md-forcefieldtools.org/RED/ [36]).

Then the grid is set depending on the docking strategy. For blind-docking strategies, the entire molecule has to be embedded on the grid to search for the potential heparin-binding

site. However, when the binding site is already known or expected, restricting the grid volume can lower the docking computation cost.

An acceptable docking run can be accomplished using 100 to 150 Lamarckian genetic algorithms (LGA) runs with 150 individuals in populations. The maximum number of energy evaluations that the genetic algorithm should make is usually set to 2500000 with 27000 maximum generations. The number of top individuals that are guaranteed to survive into the next generation was 1. Rates of gene mutation and crossover are usually 0.02 and 0.80, respectively. Following docking, all structures generated for the same compound are subjected to cluster analysis, cluster families being based on a tolerance of 2 Å for an all-atom root mean square (RMS) deviation from a lower energy structure. This protocol has been successfully applied to identify the heparin-binding site of the Eosinophil Cationic Protein (Figure 3) [18].

For the second stage, the global minimum structure can be subjected to redocking under the same conditions and to cluster analysis, using a grid box centered in the ligand with a 0.25 Å of grid spacing. Multiple docking runs can increase the performance of docking programs, as was shown specifically in the case of *Autodock*. This two-stage analysis allows a more complete exploration of conformational space in the first stage with optimization of only the successful docking modes in the second. Also, molecular dynamics simulations can be run at this stage to get more details about the interactions and flexibility of the binding site.

Molecular Dynamics Simulation

Molecular dynamics simulation (MD) is one of the main tools in the theoretical study of protein-ligand interactions. The main aim of MD is to calculate the time dependent behavior of a molecular system. MD simulations can provide detailed information on the fluctuations and conformational changes of biomacromolecules and complexes, e.g. protein-ligand complexes, and also be used to examine the structure, dynamics and thermodynamics of biological molecules and their complexes (See [37] for a guide on computational tools for protein-glycan interactions).

In MD, the time-dependent behavior of a system is obtained by integrating Newton's equations by numerical integration with a potential energy function. Simulation results are time series of conformations namely the path followed by each atom according to Newton's laws. Most MD simulations are performed under conditions of constant N,V,E (the microcanonical ensemble) though latest methods implement simulations at constant N, T and P to mimic experimental conditions.

The main drawback in running MD simulations is due to the high computational resources needed. A typical run about 100 ns time consumes between 5500 to 21000 CPU hours [38] and is commonly accomplished on supercomputers or large CPU clusters. However, the details obtained are, usually, more reliable and provide more information that those obtained by molecular docking simulations.

The method reproduced here was successfully applied by N. Sapai et al. [34] to study the interactions of fibroblast growth factor 2 (FGF2) and chemokine CXCL12 and can be useful to study protein-heparin binding interactions. Protein-heparin systems are neutralized with chlorine or sodium ions placed in a TIP3P water box and equilibrated over 1 ns with a 2 fs time step. The temperature is set to 298 K and a pressure of 1 bar by Langevin dynamics. Non-bonded interactions are cut off at 10 Å and updated every 5 steps. A switch function is applied beyond 9 Å. Long-range electrostatics is treated by the particle mesh Ewald method

255

using 1 $\mathring{A}$ grid spacing. Simulations are conducted for 20 ns with coordinates recorded every 5 ps to give 4000 snapshots per trajectory. Visualization and analysis of the obtained trajectories can be done with VMD.

The free energy of binding is calculated with the MM/PBSA method (as implemented in the AMBER 11 package). Ligand and protein coordinates are extracted every 50 ps giving 400 snapshots per trajectory. The dielectric constant is set to 1 for solutes and 80 for solvent and the ion concentration set to 0. Surface calculation is performed with the MSMS software using a probe radius of 1.4 $\mathring{A}$. The gamma and beta constants are set to 0.00542 kcal $\mathring{A}^{-2}$ and -1.008 kcal mol^{-1}. The entropy is assessed by running a normal modes calculation on 10 snapshots extracted every 2 ns, after minimization using a convergence criterion of 0.0001 kcal mol^{-1}.

Structural Analysis for Characterizing Heparin-protein Binding Interactions

The inherent difficulties in solving three-dimensional structures of protein- carbohydrate complexes have delayed the understanding of the structural determinants of heparin recognition. Most 3D-complexes have been characterized by the X-ray crystallography methodology and NMR has provided a complementary approach when the protein complexes size was not a limiting factor. High resolution structure of protein complexes with heparin oligosaccharides are now used as a reference for modeling studies and as a template for the development of novel glycotherapeutics.

NMR Spectroscopy

NMR has been applied to characterize the conformational changes that can take place on both protein and ligand upon interaction. The flexibility of certain glycosidic linkage produces multiple co-existing conformations. In particular NMR reveal the three conformers present in equilibrium of the flexible iduronate ring in the free state, and how each conformer was favored upon binding in the distinct complexes analyzed [9]. NMR was applied for example to study the conformational fold heterogeneity that modifies the heparin binding site for the lymphotactin chemokine [39, 40].

Few three-dimensional complexes solved by NMR are now available. The small size of some GAG binding proteins enable the NMR analysis, and titration experiments also identified the interacting groups [41]. The methodology has also proven appropriate for the analysis of peptide – oligomer complexes [42].

X-ray Crystallography

Crystallization of protein complexes with heparin derived ligands is particularly challenging due to the heterogeneity of the heparin fragments and the multiple binding modes. Table 1 provides an overview on the great diversity of heparin binding proteins. The table includes all the available crystal structure complexes, with the corresponding coordinate codes together with an illustrative picture of their assembly mode. The heparin pentasaccharide –antithrombin III complex in 1997 [43] was a reference starting step to unravel the protein interactions to sugar oligomers (Figure 2). Another emblematic structure

is the ternary complex where an heparin hexadecasaccharide binds to both thrombin and anti-thrombin [44].

Heparin binding can also promote dimerization, as suggested by the crystal structures of growth factor receptor complexes [45]. Chemokines bind GAGs also in a dimeric or tetrameric state, regulating thereby their role in inflammation [46, 47].

Oligomerization can be induced by sugar interaction, as observed for IFNγ association [48] and corroborated by other complementary methodologies [1].

The X-ray crystallography methodology has also provided some examples on how metal ion can modulate the GAG binding, as annexin A2 dependence of Ca^{2+} [49]. Selective divalent cation, as Zn^{2+} preference for heparin, suggests that the interaction is not merely electrostatic [50].

Table 3. List of selected reference binding proteins for the main physiological processes regulated by heparin, together with their respective known biological roles

Blood coagulation	Cell-cell interactions	Inflammation	Lipid metabolism	Cell growth and morpho-genesis	Host-pathogen interaction
Antithrombin AT III (Serine protease inhibitor)	Selectin (adhesion protein)	Lymphotactin (chemokine)	Annexin II (lipid binding protein)	FGF-1 (fibroblast growth factor)	HIV-1-gp120 (viral surface glycoprotein that mediates cell entry)
Thrombin (protease)	Fribronectin (adhesion protein)	Interleukine IL-8 (cytokine)	Annexin V (lipid binding protein, blood coagulation)	FGF-2 (fibroblast growth factor)	Malaria CS protein (sporozoite attachment to hepatocytes)
Protein C inhibitor (Serine protease inhibitor)	Vitronectin (adhesion protein)	Platelet factor 4 (PF4) (Chemokine)	ApoE (lipid transport)	VEGF (Vascular endothelial growth factor)	Foot and mouth disease virus (FMDV) receptor
Factor Xa (blood coagulation)	Laminin (adhesion protein)	CXCL12 (Chemokine)	LPL (lipoprotein lipase) lipolysis	TGFβ-1 (Transforming growth factor)	Vaccinia virus complement protein (VCP) (inhibition of the host inflammatory response)

257

Searching for Therapeutic Agents

Heparin and GAGs in general are involved in many biological processes, providing a unique field for novel drug development. Table 3 illustrates the diversity of related biological functions. More detailed information is provided by exhaustive review work [1, 3, 51, 52]. To assess *in vivo* the involvement of heparan sulfate in mammalian physiology numerous knockout and transgenic mice have also been generated, as reviewed by Bishop and coworkers [53].

Clinical significance of GAGs derives from their diversity of roles in cell function, from cell signaling to pathogenesis, as tumor growth and inflammation. HS at the cell surface contribute to many cell pathways, as lipid metabolism, by binding to apolipoproteins and lipases and promoting their endocytosis, participating in cellular crosstalk; cell signaling and tissue repair processes and under certain conditions, induce some pathological state.

Heparin/ heparan GAGs (HSGAG) are involved in several path of the tumor growth. We find HSGAG proteoglycans in tumor cell surface, and neighboring cells, such as endothelial cells. Exposed HSGAG at tumor cell surface bind to growth factors and endothelial HSGAG bind to angiogenic factors. HSGAG at the cell –extracellular matrix (ECM) regulate both the tumor development and the metastasis process. Heparin derivates have been applied for the inhibition of both tumor growth and metastasis. New heparin derivates are being engineered to remove the anticoagulant activity while preserving the required properties for cancer treatment.

Heparin and HS are also found in Alzheimer's tissue deposits [54]. In fact, HS have been reported to bind amyloidogenic peptides *in vitro* and *in vivo*, promoting fibril formation, and derived drugs have been considered to slow down or reverse the formation of amyloid deposition [3].

GAGs interaction also mediates the microbial infection process. GAGs are known to promote microbial invasion by interacting with structures at the pathogen cell surface, as adhesion proteins, such as lectins and adhesins [55]. Many pathogens use HSPGs as adhesion receptors for infection [56] and glycocompounds are designed to work as binding competitors [57, 58]. As an example, sulfated polysaccharides have a potential for the therapy and prevention of HIV infection [59].

The potential of heparin derivates is already exploited by many companies, such as Intellihep, Zacharon Pharmaceuticals, GlycoMimetics, Endotis Pharma, Polymedix, Progen, OTR3 and Momenta. Table 4 includes representative heparin derived agents that have proven of great therapeutic [8]. The design of optimized low-molecular-weight heparin molecules is based on the knowledge of the spacial distribution and properties of oligosaccharide binding motives [52]. Heparin pentasaccharides were taken as a reference for the design of heparin mimetics as anticoagulant agents. The crystal structure of the AT-III pentasaccharide complex (Figure 2) mostly contributed to the understanding of the structural basis of heparin anticoagulation properties. The GlcNAc/NS6S -->GlcA -->GlcNS3S6S --> IdoA2S --> GlcNS6S comprises the anti-thrombin binding domain. Natural and synthetic oligosaccharides mostly differ in their substitution pattern [8]. The critical charged groups were identified and the N-sulfate groups were replaced by O-sulfates and alkylated hydroxyl groups. By modulating the oligosaccharide structures and the sulfation pattern distinct

properties could be achieved and an integrated functional glycomics approach would set the basis for the development of new effective drugs.

Table 4. Representative heparin and heparan sulfate based drugs

Commercial drug	Molecule	Mechanism	Function
Fondaparinux (Arixtra®)	Synthetic pentasaccharide analogue	Factor Xa inhibitor	anti-thrombotic
Idraparinux (SanOrg 34006)	Synthetic pentasaccharide analogue with methylated hydroxyl groups and O-sulfate groups	Factor Xa inhibitor	anti-thrombotic
Idrabiotaparinux (SSR126517)	Idraparinux linked to biotin	Factor Xa inhibitor	anti-thrombotic
Rivaroxaban, Apixaban and Dabigatran		Factor Xa inhibitor	anti-thrombotic
PI-88 (Muparfostat)	Sulfated phosphomannopentose and phosphomannotetraose oligosaccharide	Heparinase inhibitor	anti-angiogenic, anti-tumor and anti-metastatic
Tramiprosate (alzhemed™)	3-amino-1-propanesulfonic acid	Amyloid β-peptide binding. inhibition of Ab aggregation	Alzheimer's disease treatment
Eprodisate (Kiacts®, Fibrillex™)	1,3 propanedisulfonic acid	Inhibition of fibril formation and tissue amyloid deposition	amyloidosis treatment
M402	Mixture of oligosaccharide chains	Interaction with growth factors	anti-metastatic
OTR4120	Dextran derivatives with carboxymetyl and sulfonate groups	Replacement of HS bound to matrix proteins and growth factors destroyed by chronic tissue injury	regenerating agents

Conclusion

The understanding of the structural basis for heparin protein interaction is setting the path to the design of novel drugs that work as anti-thrombotic, anticoagulant, anti-inflammatory, anti-metastatic and anti-angiogenic agents. Past research has highlighted the drawbacks of using native heparin oligosaccharides as drugs. Heparin oligosaccharides due to their anionic

259

nature interact with multiple proteins, leading to many side-effects. Besides, the oligosaccharides have low tissue permeability, short serum half-life and poor stability. On the other hand, the multi-step synthesis of oligosaccharides poses serious challenges for chemists. The determination of X-ray crystal complex-structures has exponentially increased in recent   years contributing to the development of GAG-mimetics. Many pharmaceutical companies are now working on GAG-based drugs and some have been successful on clinical trials.

References

[1]   N.S. Gandhi, R.L. Mancera, The structure of glycosaminoglycans and their interactions with proteins, *Chem Biol Drug Des* 72 (2008) 455-482.

[2] C.J. Malavaki, A.D. Theocharis, F.N. Lamari, I. Kanakis, T. Tsegenidis, G.N. Tzanakakis, N.K. Karamanos, Heparan sulfate: biological significance, tools for biochemical analysis and structural characterization, *Biomed Chromatogr* 25 11-20.

[3] I. Capila, R.J. Linhardt, Heparin-protein interactions, *Angew Chem Int Ed Engl* 41 (2002) 391-412.

[4] D.R. Ferro, A. Provasoli, M. Ragazzi, B. Casu, G. Torri, V. Bossennec, B. Perly, P. Sinay, M. Petitou, J. Choay, Conformer populations of L-iduronic acid residues in glycosaminoglycan sequences, *Carbohydr Res* 195 (1990) 157-167.

[5]   B. Mulloy, M.J. Forster, C. Jones, A.F. Drake, E.A. Johnson, D.B. Davies, The effect of variation of substitution on the solution conformation of heparin: a spectroscopic and molecular modelling study, *Carbohydr Res* 255 (1994) 1-26.

[6]   D. Wardrop, D. Keeling, The story of the discovery of heparin and warfarin, *Br J Haematol* 141 (2008) 757-763.

[7]   J.A. Marcum, The origin of the dispute over the discovery of heparin, *J Hist Med Allied Sci* 55 (2000) 37-66.

[8]   N.S. Gandhi, R.L. Mancera, Heparin/heparan sulphate-based drugs, *Drug Discov Today* 15 (2010) 1058-1069.

[9]   A. Imberty, H. Lortat-Jacob, S. Perez, Structural view of glycosaminoglycan-protein interactions, *Carbohydr Res* 342 (2007) 430-439.

[10] A.D. Cardin, H.J. Weintraub, Molecular modeling of protein-glycosaminoglycan interactions, *Arteriosclerosis* 9 (1989) 21-32.

[11] J.R. Fromm, R.E. Hileman, E.E. Cadwell, J.M. Weiler, R.J. Linhardt, Differences in the interaction of heparin with arginine and lysine and the importance of these basic amino acids in the binding of heparin to acidic fibroblast growth factor. *Arch Biochem Biophys* 323 (1995) 279-287.

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

[13] M.A. Fath, X. Wu, R.E. Hileman, R.J. Linhardt, M.A. Kashem, R.M. Nelson, C.D. Wright, W.M. Abraham, Interaction of secretory leukocyte protease inhibitor with heparin inhibits proteases involved in asthma, *J Biol Chem* 273 (1998) 13563-13569.

[14] E. Boix, M. Torrent, D. Sánchez, M.V. Nogués, *The Antipathogen Activities of Eosinophil Cationic Protein, Current Pharm Biotec.* 9 (2008) 141-152.

[15] P. Venge, J. Bystrom, M. Carlson, L. Hakansson, M. Karawacjzyk, C. Peterson, L. Seveus, A. Trulson, Eosinophil cationic protein (ECP): molecular and biological properties and the use of ECP as a marker of eosinophil activation in disease, *Clin Exp Allergy* 29 (1999) 1172-1186.

[16] T.C. Fan, H.T. Chang, I.W. Chen, H.Y. Wang, M.D. Chang, A heparan sulfate- facilitated and raft-dependent macropinocytosis of eosinophil cationic protein, *Traffic 8* (2007) 1778-1795.

[17] 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, *J Biol Chem* 283 (2008) 25468-25474.

[18] M. Torrent, M.V. Nogues, E. Boix, Eosinophil cationic protein (ECP) can bind heparin and other glycosaminoglycans through its RNase active site, *J Mol Recognit* 24 (2011) 90-100.

[19] M.F.M. García-Mayoral, M.; de la Torre,B.G.; Andreu, D.; Boix, E.; Nogués, M.V.; Rico, M. Laurents, D.V.and Bruix,M., NMR structural determinants of eosinophil cationic proteinbinding to membrane and heparin mimetics, *Biohys J* 98 (2010) 2702- 2711.

[20] E. Boix, D.D. Leonidas, Z. Nikolovski, M.V. Nogues, C.M. Cuchillo, K.R. Acharya, Crystal structure of eosinophil cationic protein at 2.4 A resolution, *Biochemistry* 38 (1999) 16794-16801.

[21] 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.

[22] 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, *J Biol Chem* 269 (1994) 21526-21531.

[23] T.Y. Chao, L.D. Lavis, R.T. Raines, Cellular uptake of ribonuclease A relies on anionic glycans, *Biochemistry* 49 **(2010)** 10666-10673.

[24] E. Boix, M.V. Nogues, Mammalian antimicrobial proteins and peptides: overview on the RNase A superfamily members involved in innate host defence, *Mol Biosyst* 3 (2007) 317-335.

[25] P.H. van Berkel, M.E. Geerts, H.A. van Veen, M. Mericskay, H.A. de Boer, J.H. Nuijens, N-terminal stretch Arg2, Arg3, Arg4 and Arg5 of human lactoferrin is essential for binding to heparin, bacterial lipopolysaccharide, human lysozyme and DNA, *Biochem J* 328 (Pt 1) (1997) 145-151.

[26] M. Torrent, V.M. Nogues, E. Boix, A theoretical approach to spot active regions in antimicrobial proteins, *BMC Bioinformatics* 10 (2009) 373.

[27] M. Torrent, S. Navarro, M. Moussaoui, M.V. Nogues, E. Boix, Eosinophil cationic protein high-affinity binding to bacteria-wall lipopolysaccharides and peptidoglycans, *Biochemistry* 47 (2008) 3544-3555.

[28] M. Torrent, B.G. de la Torre, V.M. Nogues, D. Andreu, E. Boix, Bactericidal and membrane disruption activities of the eosinophil cationic protein are largely retained in an N-terminal fragment, *Biochem J* 421 (2009) 425-434.

[29] V. Frecer, B. Ho, J.L. Ding, De novo design of potent antimicrobial peptides, *Antimicrob Agents Chemother* 48 (2004) 3349-3357.

[30] D.J. Hamel, I. Sielaff, A.E. Proudfoot, T.M. Handel, Chapter 4. Interactions of chemokines with glycosaminoglycans, *Methods Enzymol* 461 (2009) 71-102.

[31] I.C. Severin, J.P. Gaudry, Z. Johnson, A. Kungl, A. Jansma, B. Gesslbauer, B. Mulloy, C. Power, A.E. Proudfoot, T. Handel, Characterization of the chemokine CXCL11-heparin interaction suggests two different affinities for glycosaminoglycans, *J Biol Chem* 285 (2010) 17713-17724.

[32] E.K. Lau, C.D. Paavola, Z. Johnson, J.P. Gaudry, E. Geretti, F. Borlat, A.J. Kungl, A.E. Proudfoot, T.M. Handel, Identification of the glycosaminoglycan binding site of the CC chemokine, MCP-1: implications for structure and function in vivo, *J Biol Chem* 279 (2004) 22294-22305.

[33] F.C. Peterson, E.S. Elgin, T.J. Nelson, F. Zhang, T.J. Hoeger, R.J. Linhardt, B.F. Volkman, Identification and characterization of a glycosaminoglycan recognition element of the C chemokine lymphotactin, *J Biol Chem* 279 (2004) 12598-12604.

[34] N. Sapay, E. Cabannes, M. Petitou, A. Imberty, Molecular modeling of the interaction between heparan sulfate and cellular growth factors: bringing pieces together, *Glycobiology* (2011).

[35] W.W. Bitomsky, R., Docking of Glycosaminoglycans to Heparin-Binding Proteins: Validation for aFGF, bFGF, and Antithrombin and Application to IL-8, *J Am Chem Soc* 121 (1999) 3004-3013.

[36] F.Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, P. Cieplak, The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building, *Phys Chem Chem Phys* 12 (2010) 7821- 7839.

[37] R.J. Woods, M.B. Tessier, Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes, *Curr Opin Struct Biol* 20 (2010) 575-583.

[38] F. Haberl, O. Othersen, U. Seidel, H. Lanig, T. Clark, Investigating Protein-Protein and Protein-Ligand Interactions by Molecular Dynamics Simulations, in: S. Wagner, Steinmetz, M., Bode, A., Brehm, M. (Eds.), *High Performance Computing in Science and Engineering,* Garching/Munich 2007, Springer Berlin Heidelberg, 2009, pp. 153- 164.

[39] E.S. Kuloglu, D.R. McCaslin, M. Kitabwalla, C.D. Pauza, J.L. Markley, B.F. Volkman, Monomeric solution structure of the prototypical 'C' chemokine lymphotactin, *Biochemistry* 40 (2001) 12486-12496.

[40] R.L. Tuinstra, F.C. Peterson, S. Kutlesa, E.S. Elgin, M.A. Kron, B.F. Volkman, Interconversion between two unrelated protein folds in the lymphotactin native state, *Proc Natl Acad Sci U S A 105* (2008) 5057-5062.

[41] C.D. Blundell, A. Almond, D.J. Mahoney, P.L. DeAngelis, I.D. Campbell, A.J. Day, Towards a structure for a TSG-6.hyaluronan complex by modeling and NMR spectroscopy: insights into other members of the link module superfamily, *J Biol Chem* 280 (2005) 18189-18201.

[42] C. Vanhaverbeke, J.P. Simorre, R. Sadir, P. Gans, H. Lortat-Jacob, NMR characterization of the interaction between the C-terminal domain of interferon-gamma and heparin-derived oligosaccharides, *Biochem J* 384 (2004) 93-99.

[43] L. Jin, J.P. Abrahams, R. Skinner, M. Petitou, R.N. Pike, R.W. Carrell, The anticoagulant activation of antithrombin by heparin, *Proc Natl Acad Sci* U S A 94 (1997) 14683-14688.

[44] W. Li, D.J. Johnson, C.T. Esmon, J.A. Huntington, Structure of the antithrombin- thrombin- heparin ternary complex reveals the antithrombotic mechanism of heparin, *Nat Struct Mol Biol* 11 (2004) 857-862.

[45] L. Pellegrini, D.F. Burke, F. von Delft, B. Mulloy, T.L. Blundell, Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin, *Nature* 407 (2000) 1029-1034.

[46] T.M. Handel, Z. Johnson, S.E. Crown, E.K. Lau, A.E. Proudfoot, Regulation of protein function by glycosaminoglycans--as exemplified by chemokines, *Annu Rev Biochem* 74 (2005) 385-410.

[47] C. Weber, R.R. Koenen, Fine-tuning leukocyte responses: towards a chemokine 'interactome', *Trends Immunol* 27 (2006) 268-273.

[48] H. Perez Sanchez, K. Tatarenko, M. Nigen, G. Pavlov, A. Imberty, H. Lortat-Jacob, J. Garcia de la Torre, C. Ebel, Organization of human interferon gamma-heparin complexes from solution properties and hydrodynamics, *Biochemistry* 45 (2006) 13227-13238.

[49] C. Shao, F. Zhang, M.M. Kemp, R.J. Linhardt, D.M. Waisman, J.F. Head, B.A. Seaton, Crystallographic analysis of calcium-dependent heparin binding to annexin A2, *J Biol Chem* 281 (2006) 31689-31695.

[50] D.M. Whitfield, B. Sarkar, Heavy metal binding to heparin disaccharides. II. First evidence for zinc chelation, *Biopolymers* 32 (1992) 597-619.

[51] J.D. Esko, R.J. Linhardt, *Proteins that Bind Sulfated Glycosaminoglycans*, in: V. A., ED, C., JD, E. (Eds.), Essential Glycobiology, Cold Spring Harbor, NY, 2009.

[52] R. Sasisekharan, R. Raman, V. Prabhakar, Glycomics approach to structure-function relationships of glycosaminoglycans, *Annu Rev Biomed Eng* 8 (2006) 181-231.

[53] J.R. Bishop, M. Schuksz, J.D. Esko, Heparan sulphate proteoglycans fine-tune mammalian physiology, *Nature* 446 (2007) 1030-1037.

[54] A.D. Snow, J.P. Willmer, R. Kisilevsky, Sulfated glycosaminoglycans in Alzheimer's disease, *Hum Pathol* 18 (1987) 506-510.

[55] A. Imberty, A. Varrot, Microbial recognition of human cell surface glycoconjugates, *Curr Opin Struct Biol* 18 (2008) 567-576.

[56] R.R. Dinglasan, J.G. Valenzuela, A.F. Azad, Sugar epitopes as potential universal disease transmission blocking targets, *Insect Biochem Mol Biol* 35 (2005) 1-10.

[57] M. Touaibia, R. Roy, Glycodendrimers as anti-adhesion drugs against type 1 fimbriated E. coli uropathogenic infections, *Mini Rev Med Chem* 7 (2007) 1270-1283.

[58] A. Imberty, Y.M. Chabre, R. Roy, Glycomimetics and glycodendrimers as high affinity microbial anti-adhesins, *Chemistry* 14 (2008) 7490-7499.

[59] C.C. Rider, The potential for heparin and its derivatives in the therapy and prevention of HIV-1 infection, *Glycoconj J* 14 (1997) 639-642.

[60] A.J. McCoy, X.Y. Pei, R. Skinner, J.P. Abrahams, R.W. Carrell, Structure of beta-antithrombin and the effect of glycosylation on antithrombin's heparin affinity and activity, *J Mol Biol* 326 (2003) 823-833.

[61] D.J. Johnson, J.A. Huntington, Crystal structure of antithrombin in a heparin-bound intermediate state, *Biochemistry* 42 (2003) 8712-8719.

[62] A. Dementiev, M. Petitou, J.M. Herbert, P.G. Gettins, The ternary complex of antithrombin-anhydrothrombin-heparin reveals the basis of inhibitor specificity, *Nat Struct Mol Biol* 11 (2004) 863-867.

[63] D.J. Johnson, J. Langdown, W. Li, S.A. Luis, T.P. Baglin, J.A. Huntington, Crystal structure of monomeric native antithrombin reveals a novel reactive center loop conformation, *J Biol Chem* 281 (2006) 35478-35486.

[64] D.J. Johnson, W. Li, T.E. Adams, J.A. Huntington, Antithrombin-S195A factor Xa- heparin structure reveals the allosteric mechanism of antithrombin activation, *Embo J* 25 (2006) 2029-2037.

[65] D.J. Johnson, J. Langdown, J.A. Huntington, Molecular basis of factor IXa recognition by heparin-activated antithrombin revealed by a 1.7-A structure of the ternary complex, *Proc Natl Acad Sci* U S A 107 (2010) 645-650.

[66] W.J. Carter, E. Cama, J.A. Huntington, Crystal structure of thrombin bound to heparin, *J Biol Chem* 280 (2005) 2745-2749.

[67] K. Tan, M. Duquette, J.H. Liu, R. Zhang, A. Joachimiak, J.H. Wang, J. Lawler, The structures of the thrombospondin-1 N-terminal domain and its complex with a synthetic pentameric heparin, *Structure* 14 (2006) 33-42.

[68] V.K. Ganesh, S.A. Smith, G.J. Kotwal, K.H. Murthy, Structure of vaccinia complement protein in complex with heparin and potential implications for complement regulation, *Proc Natl Acad Sci* U S A 101 (2004) 8924-8929.

[69] V. Garlatti, A. Chouquet, T. Lunardi, R. Vives, H. Paidassi, H. Lortat-Jacob, N.M. Thielens, G.J. Arlaud, C. Gaboriaud, Cutting edge: C1q binds deoxyribose and heparan sulfate through neighboring sites of its recognition domain, *J Immunol* 185 808-812.

[70] J.P. Shaw, Z. Johnson, F. Borlat, C. Zwahlen, A. Kungl, K. Roulin, A. Harrenga, T.N. Wells, A.E. Proudfoot, The X-ray structure of RANTES: heparin-derived disaccharides allows the rational design of chemokine inhibitors, *Structure* 12 (2004) 2081-2093.

[71] J.W. Murphy, Y. Cho, A. Sachpatzidis, C. Fan, M.E. Hodsdon, E. Lolis, Structural and functional basis of CXCL12 (stromal cell-derived factor-1 alpha) binding to heparin, *J Biol Chem* 282 (2007) 10018-10027.

[72] I. Capila, M.J. Hernaiz, Y.D. Mo, T.R. Mealy, B. Campos, J.R. Dedman, R.J. Linhardt, B.A. Seaton, Annexin V--heparin oligosaccharide complex suggests heparan sulfate--mediated assembly on cell surfaces, *Structure* 9 (2001) 57-64.

[73] A.F. Moon, S.C. Edavettal, J.M. Krahn, E.M. Munoz, M. Negishi, R.J. Linhardt, J. Liu, L.C. Pedersen, Structural analysis of the sulfotransferase (3-o-sulfotransferase isoform 3) involved in the biosynthesis of an entry receptor for herpes simplex virus 1, *J Biol Chem* 279 (2004) 45185-45193.

[74] D. Shaya, A. Tocilj, Y. Li, J. Myette, G. Venkataraman, R. Sasisekharan, M. Cygler, Crystal structure of heparinase II from Pedobacter heparinus and its complex with a disaccharide product, *J Biol Chem* 281 (2006) 15525-15535.

[75] D. Shaya, W. Zhao, M.L. Garron, Z. Xiao, Q. Cui, Z. Zhang, T. Sulea, R.J. Linhardt, M. Cygler, Catalytic mechanism of heparinase II investigated by site-directed mutagenesis and the crystal structure with its substrate, *J Biol Chem* 285 20051-20061.

[76] Y.H. Han, M.L. Garron, H.Y. Kim, W.S. Kim, Z. Zhang, K.S. Ryu, D. Shaya, Z. Xiao, C. Cheong, Y.S. Kim, R.J. Linhardt, Y.H. Jeon, M. Cygler, Structural snapshots of heparin depolymerization by heparin lyase I, *J Biol Chem* 284 (2009) 34019-34027.

[77] A.D. DiGabriele, I. Lax, D.I. Chen, C.M. Svahn, M. Jaye, J. Schlessinger, W.A. Hendrickson, Structure of a heparin-linked biologically active dimer of fibroblast growth factor, *Nature* 393 (1998) 812-817.

[78] A. Canales, R. Lozano, B. Lopez-Mendez, J. Angulo, R. Ojeda, P.M. Nieto, M. Martin- Lomas, G. Gimenez-Gallego, J. Jimenez-Barbero, Solution NMR structure of a human FGF-1 monomer, activated by a hexasaccharide heparin-analogue, *Febs J* 273 (2006) 4716-4727.

[79] S. Faham, R.E. Hileman, J.R. Fromm, R.J. Linhardt, D.C. Rees, Heparin structure and interactions with basic fibroblast growth factor, *Science* 271 (1996) 1116-1120.

[80] J. Schlessinger, A.N. Plotnikov, O.A. Ibrahimi, A.V. Eliseenkova, B.K. Yeh, A. Yayon, R.J. Linhardt, M. Mohammadi, Crystal structure of a ternary FGF-FGFR- heparin complex reveals a dual role for heparin in FGFR binding and dimerization, *Mol Cell* 6 (2000) 743-750.

[81] D. Lietha, D.Y. Chirgadze, B. Mulloy, T.L. Blundell, E. Gherardi, Crystal structures of NK1-heparin complexes reveal the basis for NK1 activity and enable engineering of potent agonists of the MET receptor, *Embo J* 20 (2001) 5543-5555.

[82] 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, *J Biol Chem* 276 (2001) 15009-15017.

[83] S.C. Lee, H.H. Guan, C.H. Wang, W.N. Huang, S.C. Tjong, C.J. Chen, W.G. Wu, Structural basis of citrate-dependent and heparan sulfate-mediated cell surface retention of cobra cardiotoxin A3, *J Biol Chem* 280 (2005) 9567-9577.

[84] E.E. Fry, S.M. Lea, T. Jackson, J.W. Newman, F.M. Ellard, W.E. Blakemore, R. Abu-Ghazaleh, A. Samuel, A.M. King, D.I. Stuart, The structure and function of a foot-and- mouth disease virus-oligosaccharide receptor complex, *Embo J* 18 (1999) 543-554.

[85] J. Dasgupta, M. Bienkowska-Haba, M.E. Ortega, H.D. Patel, S. Bodevin, D. Spillmann, B. Bishop, M. Sapp, X.S. Chen, Structural basis of oligosaccharide receptor recognition by human papillomavirus, *J Biol Chem* 286 (2011) 2617-2624.

[86] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, *J Mol Biol* 372 (2007) 774-797.