UNIVERSITÀ DEGLI STUDI DI NAPOLI "FEDERICO II"

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

TESI DI DOTTORATO IN

SCIENZE CHIMICHE

XXIII CICLO

STRUCTURAL STUDIES OF NUCLEIC ACIDS AND PROTEINS INVOLVED IN NUCLEIC ACID RECOGNITION

Dottoranda

Irene Russo Krauss

TUTORE
Prof.ssa Filomena Sica

RELATORE
Prof.ssa Concetta Giancola

COORDINATORE DEL XXIII CICLO
Prof. Lucio Previtera

PREFACE

Macromolecular recognition is based on the requirement of dual geometric and chemical complementarity, eventually leading to the formation of a thermodynamically stable and specific complex between interacting molecules. These aspects are key elements for understanding the function of biological systems: enzymes that bind substrates and effectors, proteins that mediate signal transduction *via* networks of alternative or specific protein-protein pair, and nucleic acids that, *via* the binding of transcription factors, repressors, co-activators, regulate protein expression. In particular, the site-specific associations between DNA and proteins regulate most biological events, with key involvement in transcription, replication and recombination. In order to deeply understand the chemical and physical mechanisms by which the cellular components exploit their functions, a prerequisite is a description of their structure at atomic level. The best technique currently available for visualizing the structures of macromolecules at atomic scale is X-ray diffraction analysis of single crystals. X-ray crystallography has transformed our understanding of biological processes. For example it provided the first clues to the structure of the DNA double helix 50 years ago, giving profound insights into how DNA is replicated. Since the pioneering work of Perutz and Kendrew on the structure of haemoglobin and myoglobin in the late 1950s, about 60000 protein structures have been solved to date by X-ray crystallography. Nowadays, crystallographers and structural biologists attention has moved on macromolecular complexes, which control the complex signaling and molecular recognition processes that occur in virology, cell signaling, DNA hybridization, immunology. Indeed, protein-protein or protein-nucleic acids cross-talk is in many cases even more important than the structure of an individual molecule.

A field where the detailed knowledge of macromolecular structures and interactions plays a pivotal role is drug discovery and development. The process of drug discovery involves the identification of target molecules (usually but not always proteins), which serve as points of attack for future medicines, the design and synthesis of potential lead compounds, and further characterization, screening, and assays for therapeutic efficacy and toxicity. Target structures do not lead directly to a drug, but the information of potential binding sites and substrate geometries enable a focused approach to screening for potential therapeutic agents. On the other hand, structures of targets in complexes with potential drugs may be much more useful: they suggest synthetic variation strategies, depending on optimization needs. For examples potency can be increased by optimizing interactions, pharmacokinetic properties may be improved with targeted variation of moieties not essential for binding, selectivity may be enhanced (or decreased) by comparing binding modes in related targets and new binding modes may be discovered by targeted variation of essential binding moieties. The structural information obtained

(where does the molecule bind; what are the ligand/protein/solvent interactions; how does the protein/ligand structure change upon binding) can be invaluable in the generation of novel molecules or in the re-design of existing molecules whose drug properties are not optimal. Thus crystallography has became a standard technique used by many pharmaceutical and biotechnology companies. However, a deeper characterization and knowledge of macromolecules dynamics, protein-protein and protein-nucleic acid interactions is reached when crystallography is combined with computational and biophysical studies, including different spectroscopic techniques.

The present study focuses on the structural analysis of two different kind of proteins that interacts with nucleic acids and their ligands. The first part concerns the structural characterization of complexes between human α-thrombin, a protein of physiological and pathological relevance, and two oligonucleotidic aptamers, which adopt a quadruplex fold. This knowledge opens the door to the development of novel aptamers with improved antithrombotic activity, preserved stability and increased binding efficacy and represent a further affirmation of the crucial role of macromolecular crystallography for drug development.

The second part of the thesis is subdivided into more issues. It globally deals with the study of a protein family, whose cellular target is RNA: the ribonucleases or RNases. This is a class of enzymes whose functions are much wider and fall outside the limits of simple destruction of RNA molecules. A particular attention is devoted to ribonucleases that, thanks to their structural features (domain swapping), can exert an antitumor activity. The results obtained through a combination of crystallographic and docking analyses allowed to identify the structural determinants of antitumor dimeric RNases and have been successfully utilized to convert non-cytotoxic RNase A in a dimeric cytotoxic mutant. These results are very useful to further improve biological properties of these macromolecules that can be used as a new generation of antitumor agents.

Another topic of the thesis is the characterization of the whole ribonuclease system in zebrafish organism. This is an ideal vertebrate model system for phylogenetic and evolutionary analyses and for biomedical studies. In particular, zebrafish is considered a new powerful tool in research and drug discovery that will accelerate screening processes and the finding of new lifesaving drugs for humans. Although assay development to model human diseases in zebrafish is still in its infancy, a variety of diseases, among which cancer, have been successfully modelled in zebrafish. Structural determination by crystallography and homology modelling of these ribonucleases, endowed with angiogenic activity, allowed to understand the molecular bases of their different properties. Furthermore it represents a good starting point to study evolution of a protein scaffold from ancestral organisms to contemporary mammalians.

Finally this study has regarded chimeric molecules, with specific cell-type cytotoxicity, containing

RNase moieties. A computational approach was used to identify the epitope recognized by a powerful antitumor construct, immuno-RNase. This fully human fusion protein is highly selective towards tumor cell that overexpress the receptor ErbB2, and does not present immunogenicity nor cardiotoxicity as side effects, whereas approved antitumor antibodies do. Since immuno-RNase properties depend on the ErbB2 region that it binds, a detailed knowledge of interactions between immuno-RNase and its epitope is useful to implement immuno-RNase pharmacological properties, and also to design specific antidotes to control its effect. The study has been successful: a small ErbB2 region (14 aminoacid residues) bound by immuno-RNase was identified and the results obtained by our computational approach were also validated by experimental data. Remarkably immuno-RNase binds a different epitope with respect to immunoagents approved for cancer therapy. This finding on one hand could justify immuno-RNase lack of cardiotoxicity, on the other hand suggests that this protein could be used in combination with other drugs to obtain more efficient antitumoral therapies.

PART I

NUCLEIC ACIDS-PROTEINS INTERACTIONS: THE THROMBIN-APTAMERS CASE

INTRODUCTION

First discovered in 1869 DNA continues to amaze and intrigue the researchers devoted to its study. The Watson and Crick model for B-form duplex DNA dramatically improved the mechanistic understanding of the replication and flow of genetic information, termed "the central dogma of molecular biology" by Francis Crick in 1958. Since then, the B-form has usually been regarded as the biologically relevant structure of DNA. However, in the context of nucleosome core particles and other nucleoprotein complexes, DNA can adopt a wide variety of conformations, including highly distorted A-, B-, Z-form double helices [1]. In addition to various duplex structures, single stranded DNAs can fold into a wide variety of hairpin, triplex, G-quadruplex, and i-motif structures containing non-canonical base pairs [2]. Single stranded DNA can exhibit structural diversity comparable, in some ways, to that of RNA. Polypurine–polypyrimidine tracts and other repetitive sequences can form non-duplex and/or higher-order chromatin structures, possibly related to many different biological activities. One family of structures originating from guanine-rich tracts is referred to as G-quadruplex (or G-tetraplex)[3].

G-quadruplex

Since from the fifties it is known that guanine-rich oligonucleotidic sequences can dispose in a quadruple helix structure, named G-quadruplex. The main component of the G-quadruplex is the G-tetrad or G-quartet, a planar arrangement of four guanine bases associated through a cyclic array of Hoogsteen-like hydrogen bonds, in which each guanine base both accepts and donates two hydrogen bonds [4] (Figure 1). A quadruplex is composed of at least two stacked G-quartets, which can be connected by residues not involved in the quartet itself that define the loops [5].

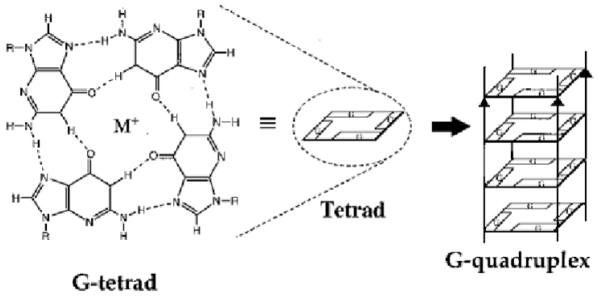

Figure 1 Top view of a G-quartet and its arrangement in a G-quadruplex

G-quadruplex can be formed by one, two or four strands and are named uni-, bi-, ore tetra-molecular, respectively. Furthermore they are differentiated for several structural features:

- the orientation of the strands, which could be parallel or antiparallel, where the last term is used for a quadruplex with at least one strand opposite to the other ones,
- the conformation of guanines glycosidic bonds angles,
- the size of grooves spanning contiguous backbones,
- the conformation of connecting loops.

In particular three types of grooves exist, narrow, wide and medium, depending on the arrangement of glycosidic bonds angles of hydrogen bonded guanines. Narrow and wide grooves are the result of hydrogen bonded bases with different glycosidic bond angles, whereas contiguous bases with the same conformation form a medium groove [6]. Concerning the loop conformation, three types of loops are observed in quadruplexes: lateral, or edgewise, diagonal and propeller.

The most simple quadruplex structure is the tetra-molecular one. It is composed by four strands, typically with the same sequence, which are parallel to each other. There is no connecting loop and all guanines have *anti* glycosidic bonds angles, resulting in grooves with the same width. The topologies allowed for bi-molecular and uni-molecular quadruplexes are more complex and variegate. In principle each arrangement of the strands is possible, parallel or antiparallel. Furthermore antiparallel quadruplexes are formed by guanines with both *anti* and *syn* glycosidic torsion angles, thus different types of grooves are observed between contiguous backbones. A great variability of structures derives from loop orientations and loops with different conformations can be present in a single quadruplex [7, 8] (Figure 2).

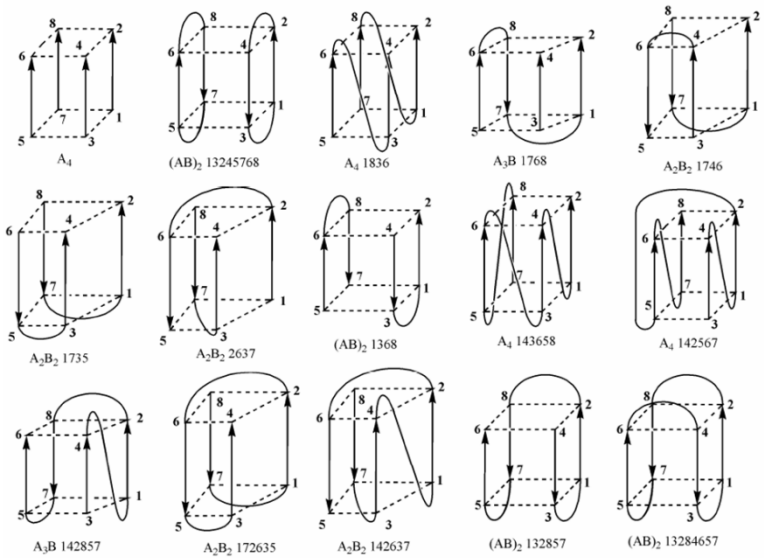

Figure 2 Overview of different G-quadruplex topologies

Furthermore no severe rule establishes a unique quadruplex fold for a given oligonucleotidic sequence, and examples of oligonucleotides that fold in different structures exist. In particular, it has been observed that a crucial role in determining three-dimensional structure is played by metal ions or molecular crowding.

Stacked quartets line with their O6 atoms a central cavity with a strong negative electrostatic potential, where cations are well accommodated. In particular the coordination of potassium [9], sodium [10] and strontium [11] provides thermodynamic stability to the G-quadruplex structure. It has been proved that G-quadruplexes are stabilized by the alkaline series in the following order: $K^+ >> Na^+ > Rb^+ > Cs^+ >> Li^+$, and by the alkaline earth series in the order: $Sr^{2+} >> Ba^{2+} > Ca^{2+} > Mg^{2+}$, indicating that the atomic radii of potassium and strontium fit best in the coordination site between adjacent G-tetrads [12, 13]. Moreover the binding of heavy metals, such as tallium [14], or of positive molecular groups, such as ammonium [15], has been also observed. The location of the cation with respect to the quartet plane strongly depends on the metal. Potassium ions are always found sandwiched between stacked tetrads, almost equidistant from the eight oxygen atoms [16], whereas a variety of coordination geometries is possible for the smaller sodium ion. Indeed, sodium ions have been observed on the quartet plane (coordination number four), very close to the plane (coordination number five, the fifth ligand being water or another metal ion) and even between the two tetrads with coordination number eight [17-19]. In some cases metal ion binding is determinant for the quadruplex conformation, and there are several well-established examples where the change from Na^+ to K^+ induces a profound structural alteration in the quadruplex fold [20-22].

G-quadruplex biological roles

The first direct evidence for the presence of G-quadruplex DNA structures *in vivo* came from studies using G-quadruplex DNA-specific antibodies to detect intermolecular structures at ciliate telomeres, where their formation and dissolution are cell cycle regulated [23-25]. In vertebrates, the telomeres that cap chromosome ends are composed of tandem repeats of the sequence d(TTAGGG)n, with a 39 single stranded extension (G-overhang) associated with six protein factors (TRF1, TRF2, RAP1, TIN2, TPP1 and POT1) that form a protecting complex (*shelterin*) essential for genome stability [26]. Telomeres protect the ends of chromosomes from recombination, fusion, or being recognized as damaged DNA [27, 28] and need to be maintained above a critical length. Furthermore they are involved in the natural process of cellular senescence: at each cellular replication telomere shortening occurs, thus defining life of healthy cell [29]. Cellular mechanisms that counteract this loss confer indefinite proliferation potential, characteristic of cancerous cells [30]. The finding that telomeric DNAs can associate to form quadruplex structures in physiological ionic conditions is particularly intriguing because quadruplex structures are not recognized by telomerase. The high activity of telomerase in the majority of tumors combined with its absence in most adult normal tissues has generated considerable interest in targeting the enzyme and associated telomeres in a cancer therapeutic context [31-34]. In particular it has been suggested that stabilization of quadruplexes in telomeric ends of cancerous cell can inhibit telomerase activity and induce cellular death [35, 36].

Furthermore quadruplex-forming guanine-rich sequences are also found in a number of cancer-related genes, such as c-myc [37-40], c-kit [41, 42], HIF-1α [43], VEGF [44], BCL-2 [45, 46], KRAS [47], Rb [48, 49] and RET protooncogene [50]. For example, it has been shown that G-quadruplexes are involved in transcriptional control of the c-myc oncogene [38], whose overexpression is associated with a variety of human malignancies, including breast, colon, cervix, small-cell lung carcinomas, osteosarcomas, glioblastomas, and myeloid leukemias [51-54]. Thus G-quadruplex interacting agents with specificity for binding to silencer element can be good anticancer agents.

Aptamers

Aptamers are short DNA- [55] or RNA- [56]based oligonucleotides selected from large combinatorial pools of sequences for their capacity to efficiently recognize targets . More important, the structure and conformation of aptamers can be rationally designed and tailor-made to be selectively responsive to various kinds of targets, including small molecules [57], biological macromolecules [58-60] and even cells [61, 62]. Like antibodies, aptamers exhibit high specificity and affinity for target binding. As a result, they may display effective interference in biological processes, which renders them not only valuable diagnostic tools, but also promising therapeutic agents [63]. Aptamers display a large number of structural arrangements, which accounts for their binding efficiency and selectivity for unrelated

targets. The G-quadruplex architecture is adopted by several of them, the most popular of which shows inhibitory properties against α-thrombin, a pharmacologically relevant protein. G-quadruplex conformational plasticity consents fine tuning of target recognition as obtained by aptamer selection [63].

As therapeutic agents, aptamers may have some advantages over monoclonal antibodies and other oligonucleotide-based approaches. Unlike monoclonal antibodies, they are non-immunogenic and present higher heat stability. Thanks to their high affinity and specificity, it should be allowed a therapeutic dosing at submicromolar levels, reducing potential non-specific effects. Moreover antidotes can be designed to control their pharmacologic activity [64]. G-rich aptamers that adopt a G-quadruplex structure have also the additional benefit of an enhanced cellular uptake and a certain resistance to serum nucleases, when compared to unstructured nucleotide sequences used in therapeutic approaches [65]. This latter finding is particularly significant: indeed, with the aim to obtain more resistant non-quadruplex aptamers, modified oligonucleotides are used. They present several drawbacks, including an increased toxicity and reduced specificity. On the contrary the resistant G-quadruplex arrangement allows the use of unmodified or minimally modified oligonucleotides [65].

Tests performed with some quadruplex aptamers have reported antiproliferative activity correlated with protein-binding properties [66], although specific targets have been identified for only few of them, such as nucleolin [67, 68] and interferon-γ [69-71]. However the most known quadruplex aptamer is the DNA 15-mer with anti-thrombotic activity.

Thrombin

α-thrombin is a trypsin-like serine protease that plays a key role in the coagulation cascade [72]. The coagulation is a complex process by which blood forms clots upon the damage of a blood vessel wall. In mammals this process involves both a cellular component, the platelet, and a protein component, which includes several coagulation factors. Immediately upon the injury, platelets start to aggregate: this process is known as primary haemostasis. Simultaneously the secondary haemostasis occurs: the coagulation factors circulating in the plasma as inactive zymogens trigger a complex succession of reactions that results in the final production of fibrin strands, which strengthen the platelet plug [73].

The most important constituent of the coagulation cascade is α-thrombin, since it is the only enzyme capable of catalyzing the cleavage of soluble fibrinogen to produce insoluble fibrin clot. It is produced by the enzymatic cleavage of two sites on the zymogen prothrombin, or coagulation factor II, by the activated coagulation factor X. Apart from the processing of fibrinogen, α-thrombin also activates blood coagulation factors V and VIII, amplifying the coagulation cascade, because their activation results in a huge amplification of thrombin generation, in a positive feedback action. Furthermore it

activates coagulation factor XIII, which cross-links the fibrin monomers, thus reinforcing the initial fibrin clot (Figure 3). In addition to its activity in coagulation cascade, α-thrombin promotes platelet aggregation [72, 74, 75].

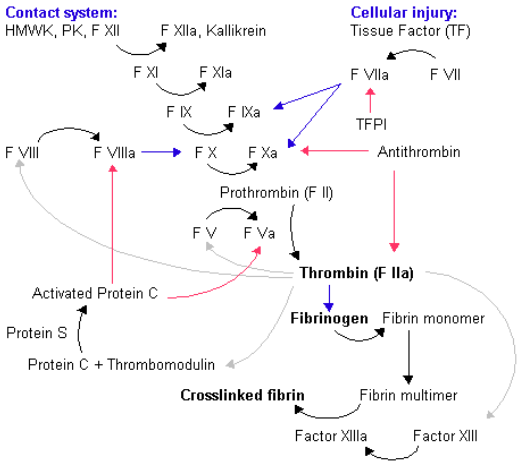

Figure 3 Schematic representation of the coagulation cascade process

α-thrombin has the surprising property to switch its activity from a procoagulant to an anticoagulant one, through the binding to different macromolecules. Indeed, in complex with thrombomodulin α-thrombin undergoes to a considerable conformational variation and loses its platelet activating and protease activities. Furthermore it becomes able to activate protein C, which in turn inhibits the formation of factors V and VIII. This process is very important in the normal physiological state, because normal endothelial cells produce thrombomodulin that binds any circulating thrombin, thus preventing clot formation in undamaged vessels [74].

The thrombin capability to perform many different functions depends on its ability to recognize a large variety of substrates, inhibitors and cofactors. This ability is finely regulated by the binding of several ligands that allosterically modulate the substrate specificity of the enzyme, thereby rendering it capable of selecting its distinct substrates and cofactors in the opposite procoagulant and anticoagulant pathways[72, 74, 76].

The crystal structure of α-thrombin indicated the typical serine protease fold, composed of two adjacent β-barrels, which line a deep cleft where the characteristic serine-histidine-aspartate catalytic triad (Ser195, His57 and Asp102, in α-thrombin) is placed. However, with respect to the other enzymes of the family, α-thrombin has long insertion loops that are crucial for the recognition of its different ligands and make its active site less accessible than those of other proteases [77, 78]. In ligand recognition an essential role is played by two positive patches on thrombin surface, placed on the opposite sides of the active site, named exosites I and II (Figure 4).

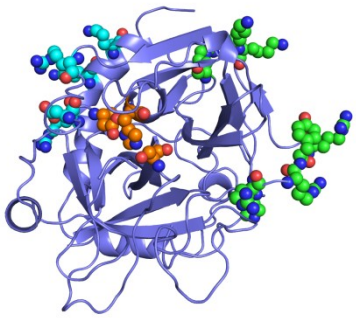

Figure 4 Cartoon representation of thrombin molecule (PDB code 1PPB). Residues of active site, exosite I and exosite II, shown as spheres, are coloured orange, green and cyan respectively

Exosite I is the binding site of thrombin physiological substrate fibrinogen, therefore it is also known as the fibrinogen recognition site (FRE), and contains also hydrophobic patches, whereas exosite II contains only a very large number of positive residues and is the binding site of thrombin inhibitor heparin [72, 77-79].

Thrombin generation is closely regulated to locally achieve rapid haemostasis after injury, without causing uncontrolled systemic thrombosis. In the absences of efficient and timely thrombin generation, stable blood clots cannot form, resulting in haemorrhage [80]. Conversely, unregulated thrombin activity produces an excessive coagulation function, resulting in dissemination of the clot beyond the site of vessel damage. It can obstruct normal blood flow thus causing thrombosis.

The capability of inhibiting and regulating thrombin activity *in vivo* by synthetic compounds is an important goal in prevention and treatment of clotting abnormalities. Thrombin inhibition can be achieved by different mechanisms [81]. An option is blocking its active site, through the binding of substrate analogues that form covalent bonds with thrombin catalytic residues, in this way the hydrolysis of both macromolecular and small substrates is prevented. Another mechanism is blocking the binding of macromolecular substrates to exosite I or even favouring the binding of thrombin to SERPINs (SERine Protease INhibitors), which irreversibly suppress thrombin activity. The last mechanism is that by which heparin acts: it binds thrombin exosite II and drives the formation of the complex between thrombin and a SERPIN, antithrombin [80, 82].

Thrombin Binding Aptamer

Many different thrombin inhibitors exist, including natural and modified peptides, glycosaminoglycans and even oligonucleotides. Amid these inhibitors, oligonucleotidic aptamers are particularly interesting. In 1992 a DNA 15-mer consensus sequence, $^{5'}$GGTTGGTGTGGTTGG$^{3'}$, which inhibits thrombin clotting activity at nanomolar concentration, was identified [55]. It was named thrombin binding aptamer, or TBA. TBA is a strong anticoagulant *in vitro*, acting on the two procoagulant function of thrombin, since it inhibits thrombin-catalyzed activation of fibrinogen and thrombin induced platelet

aggregation. Moreover it is able to bind both free and clot-bound thrombin [83, 84]. Nonetheless its promising properties as drug in anticoagulant therapies, TBA presents the trouble of a rather rapid degradation by nucleases *in vivo*.

NMR solution structure of TBA revealed that it adopts a uni-molecular quadruplex structure, composed of two G-quartets and three loops, two TT loops on one side and a TGT loop on the opposite side (Figure 5). TT loops span narrow grooves, whereas TGT loop spans a wide groove [85, 86].

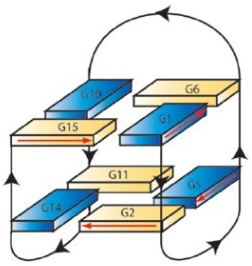

Figure 5 Schematic representation of the structure adopted by TBA. Black arrowheads indicate 5'→3' polarity of the strands. *Anti* and *syn* guanines are depicted yellow and blue, respectively.

The structure of thrombin-TBA complex was solved by X-ray crystallography at 2.9 Å resolution [87]. Although the central core of TBA is the same in the NMR and crystallographic models, some structural differences exist in the way these central bases are connected. In particular, because of the poor quality of diffraction data, the disposition of aptamer loops with respect to grooves was not clear and in the crystallographic model TGT loop spans a narrow groove, whereas TT loops span a wide groove (Figure 6).

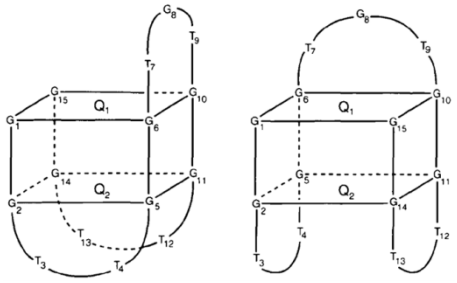

Figure 6 Schematic representation of the two TBA models, NMR model (left) and crystallographic model (right): in the first one TT loops span narrow grooves, whereas TGT loop spans a wide groove, the opposite occurs in the second one

In a second time the structure of the complex was re-determined with better diffraction data and by using the coordinates of the NMR structure of the 15-mer in the fitting procedure (NMR derived model) [88]. The two models refined to similar values of crystallographic R factor and R_{free} and seemed to be equally correct. They differ for the orientation of TBA with respect to thrombin: the aptamer is rotated 180° about an axis between the two G-quartets. Thus in the pure crystallographic model TBA

interacts with thrombin exosite I by its TGT loop, whereas in the NMR-derived model by its TT loops (Figure 7). Thus, apart for establishing that the aptamer binds the enzyme through the loops, it remained fully uncertain the specific loop that is involved and the nature of the interactions that stabilize the complex.

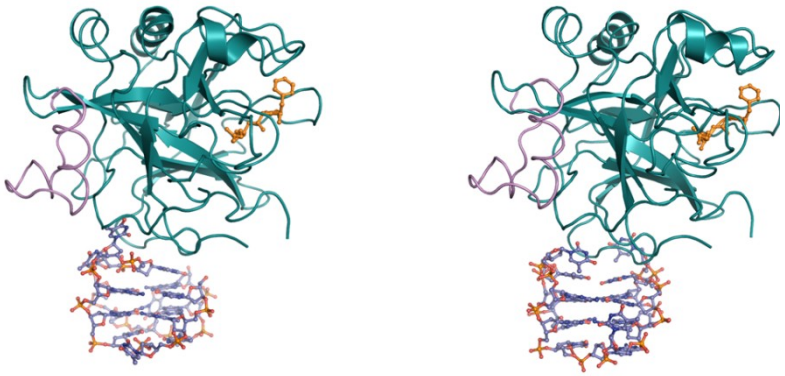

Figure 7 Representation of the two proposed models of thrombin-TBA complex: on the left the pure crystallographic one (PDB code 1HAP), and on the right the NMR-derived one (PDB code 1HAP). Thrombin molecule is shown as cartoon with heavy chain in green and light chain in violet. PPACK (orange) and TBA are shown in ball and stick.

In the crystal, TBA also interacts with the exosite II of a second symmetry-related thrombin molecule (Figure 8) [88].

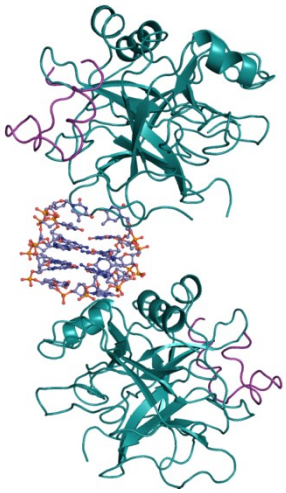

Figure 8 TBA sandwiched between two thrombin molecules in the NMR-derived model

These additional packing contacts raise the question of whether the recognition site of the aptamers on the thrombin molecule is the fibrinogen binding site or the heparin binding site, or both of them. The latter possibility has found credit on the basis of isothermal titration calorimetric data that were interpreted in terms of a complex in solution formed by two thrombin molecules and one aptamer [89]. However, more recent data support the 1:1 stoichiometry [90].

Modified Thrombin Binding Aptamer

A great effort has been devoted in the last years to development of new thrombin binding aptamers with improved pharmacological properties. As mentioned before, unmodified TBA presents some drawbacks, in particular a rather rapid digestion by endogenous nucleases. Thus several modified aptamers have been produced and characterized, including aptamers with modified residues [91], LNA (locked nucleic acid) [92, 93], 2'F-ANA (2'-deoxy-2'-fluoro-arabino nucleic acid) [94], isoguanine [90], acyclic thymidine nucleosides [95], or non-canonical linkages between contiguous residues, formacetal bonds [96] or 5'-5' polarity inversion sites [97]. The latter strategy is particularly promising because it has been reported that polarity inversion sites confer an increased resistance to endogenous nucleases [98], allowing the use of natural monomers, whose metabolites should not be toxic for organism. At the same time it has been proved that oligonucleotides with polarity inversion sites are still able to adopt the quadruplex structure required for thrombin binding.

A modified thrombin binding aptamer, with a 5'-5' polarity inversion site between thymine 3 and 4, named mTBA (Figure 9), was produced and amply characterized [89, 97].

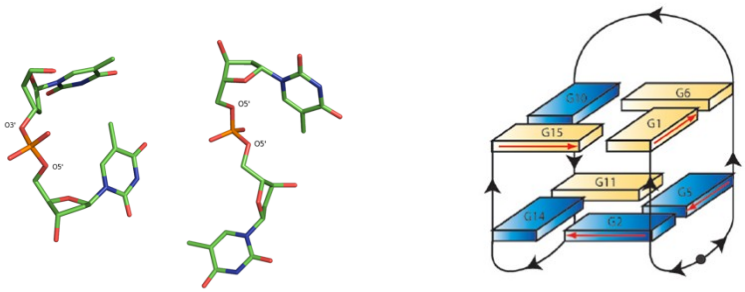

Figure 9 Comparison between a canonical 3'-5' and a 5'- 5' bond (on the left). On the right schematic representation of the structure adopted by mTBA. Black arrowheads indicate 5'→3' polarity of the strands, black circle represents the 5'–5' polarity inversion site. *Anti* and *syn* guanines are depicted yellow and blue, respectively

mTBA presents some improved properties with respect to its unmodified counterpart: it is slightly more stable (T_m is 53°C and 58°C for TBA and mTBA, respectively) and has a higher thrombin affinity (K_b is $3x10^6$ M and $4x10^7$ M for TBA and mTBA, respectively). However its inhibitor activity is less than that of TBA, as assessed by prothrombin time tests [89]. mTBA structure was solved in solution by NMR: it folds as a uni-molecular G-quadruplex with a chair-like conformation, similar to that of unmodified TBA. Few differences are observed between the two molecules [97], but they do not provide a clear justification of their different properties.

Despite its reduced inhibitory activity, mTBA represents a good starting point for the design of improved therapeutics.

Thesis purpose

So far the rational design of aptamer based anti-thrombotic agents has been prevented by the absence of a clear knowledge of interactions between thrombin and TBA, due to the lack of an unambiguous and detailed structure of their complex. In this context the purpose of the present research work is the structural characterization by X-ray diffraction of the complex of human α-thrombin with mTBA and the rivisitation at higher resolution of the structure of α-thrombin with TBA.

The initial part of this thesis job consisted in screening and optimization of crystallization conditions, in order to obtain large, well diffracting crystals of thrombin-mTBA and thrombin-TBA complexes. These crystals were then used for diffraction data collection. Complexes structures, solved and refined at high resolution, were analyzed and compared to each other, with the aim to clarify the doubts existing about the recognition between the molecules and the complex assembly and to justify the different properties deriving from the polarity inversion site. Moreover, in order to investigate the effect of ions on aptamer folding and binding to target enzyme, thrombin-TBA complex was characterized in two different conditions, in the presence either of potassium or sodium ions.

RESULTS AND DISCUSSION

Aptamers folding

Nucleotides conformational homogeneity was checked by circular dichroism measurements. The aptamers were dissolved in potassium or sodium phosphate buffer at pH 7.1, heated 10 minutes at 85°C, then they were slowly cooled down to room temperature and stored at 20°C. After one night a small aliquot of each sample was diluted at 20 μM in order to record circular dichroism spectra.

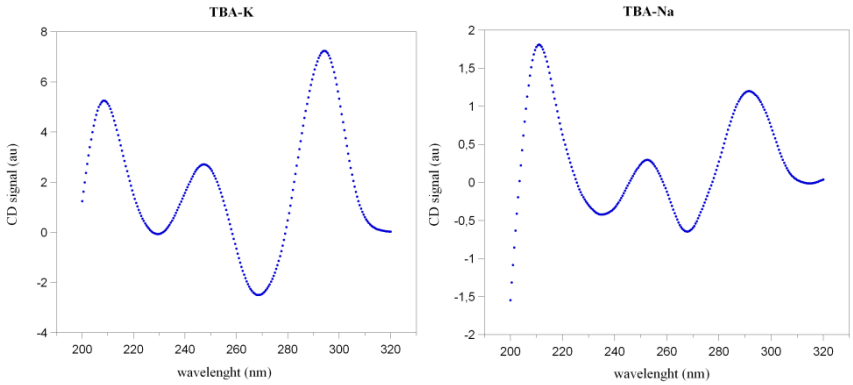

Figure 10 CD spectra of TBA-K (left) and TBA-Na (right)

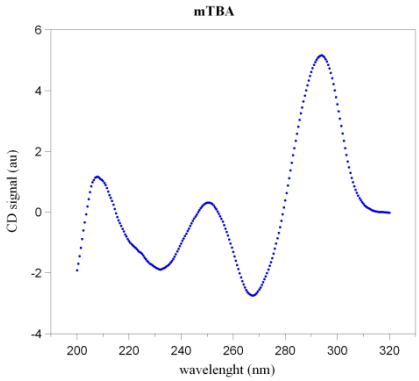

Figure 11 CD spectrum of mTBA

The spectra are very similar to each other, despite the polarity inversion site of mTBA: they present two positive bands at 248 and 294 nm and a negative one at 267 nm (see Figure 10 and Figure 11). These characteristic signals indicate that each aptamer consists of at least two G-tetrads with an alternating *anti* and *syn* glycosidic conformation along each strand, they are typical for anti-parallel quadruplex structures [99, 100]. Slight differences exist between CD spectra of the two TBA samples: the spectrum of TBA-K is more resolved, suggesting that the potassium ion stabilizes the quadruplex fold better than sodium ion.

Crystal structures

Overall structure and thrombin molecule

After an extensive screening of crystallization conditions, good quality crystals of thrombin-mTBA and thrombin-TBA complexes have been obtained (Figure 12). The formers belong to the space group I222 and diffract up to a resolution of 2.15 Å. Crystals of the complex between thrombin and TBA-K and TBA-Na are isomorphous, space group P1, and diffract up to a resolution of 2.05 Å and 1.80 Å, respectively.

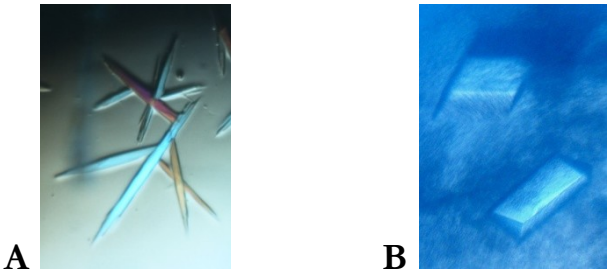

Figure 12 Best crystals of thrombin-mTBA complex (A) and thrombin-TBA complex (B)

In all cases the asymmetric unit contains one thrombin and one aptamer molecule (Figure 13).

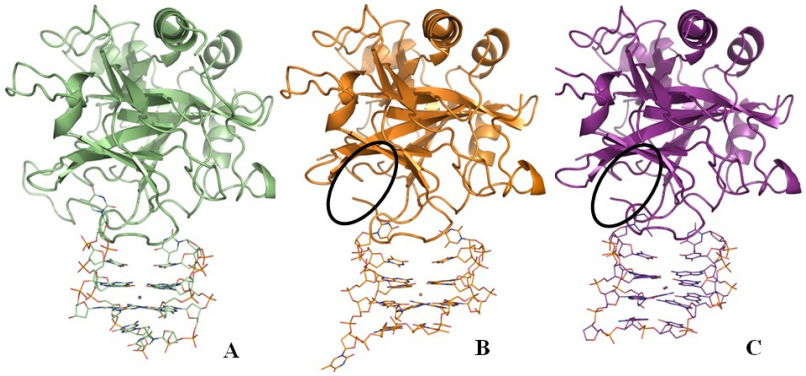

Figure 13 Overall structure of the thrombin-mTBA (A), thrombin-TBA-K (B) and thrombin-TBA-Na (C) complexes. The missing γ autolysis loop in B and C is marked with a black circle

The final models of the complexes were refined to R-factor/R_{free} values of 0.196/0.252, 0.182/0.234 and 0.178/0.224, for thrombin-mTBA, thrombin-TBA-K and thrombin-TBA-Na, respectively. A full list of refinement statistics is reported in Table 1.

Table 1 Summary of refinement statistics for thrombin-aptamers complexes

	Thrombin-mTBA	Thrombin-TBA-K	Thrombin-TBA-Na
Refinement Results			
Resolution limits (Å)	50.00-2.15	50.00-2.05	50.00-1.80
Number of reflections used in the refinement (F>2σ (F))	23598	19447	26555
No. of reflections in working set	22464	18499	25273
No. of reflections in test set	1134	948	1282
R/R_{free}	0.196/0.252	0.182/0.234	0.178/0.224
No. of protein atoms	2374	2266	2265
No. of aptamer atoms	306	315	306
No. of ligand atoms	44	44	44
No. of water molecules	250	191	222
No. of ions	2	6	7
RMSD from ideal values			
Bond lengths (Å)	0.02	0.02	0.02
Bond angles (°)	2.11	1.92	2.10
Average B-factors (Å²)			
Protein, overall	18.29	30.55	25.74
Main chains	17.03	28.84	23.98
Side chains	19.52	32.18	27.41
Ions, solvent and ligand atoms	28.49	39.83	36.78
Aptamer atoms	25.90	40.79	38.12
Ramachandran plot statistics			
Residues in the most favored regions (%)	86.3	97.0	98.1
Residues in the additionally allowed regions (%)	13.7	2.2	1.9
Residues in the generously allowed regions (%)	-	0.7	-

The polypeptide chain of thrombin is very well defined for the whole heavy chain (residues 16-247) and for residues 1H–14L of the light chain, in the case of thrombin-mTBA complex, whereas in both the thrombin-TBA structures it is well defined for the heavy chain (residues 16-246), with the exception of the γ-autolysis loop (residues 148-150), and for residues 1B-14K of the light chain. The light chain and the γ-autolysis loop are very flexible regions, whose conformation is usually found to be strongly affected by crystal packing: when intermolecular crystal contacts are missing these region are generally disordered and not visible in electron density maps, as happens in the present triclinic crystal form of thrombin.

All residues in the refined structures, with the exception of two residues of a flexible loop in thrombin-TBA-K, lie within "allowed" regions of the Ramachandran plot. These results are remarkable when compared with the quality of the old thrombin-TBA structures, PDB codes 1HAO and 1HAP [87, 88], for which only the 79.6% of residues are in favoured region, whereas the 17.9/16.0% lie in allowed region and the 2.6/4.4% in forbidden regions, for the structures 1HAO/1HAP respectively.

The quality of electron density maps is very high for all the models. As an example in Figure 14 omit Fo-Fc map in the N-terminal region of light chain calculated for each of the two dataset of thrombin-mTBA complex is represented.

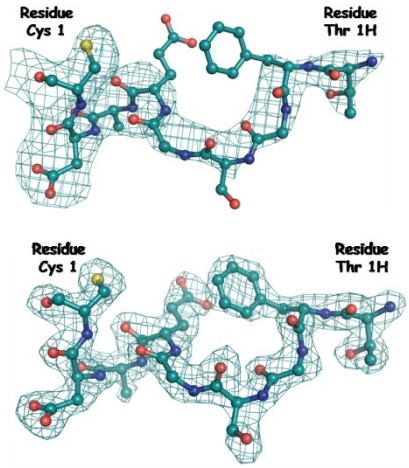

Figure 14 Comparison of electron density maps in the N-terminal region of thrombin light chain (thrombin-mTBA complex) calculated with data collected from thin crystals at 2.99Å (up) and from best crystals at 2.15 Å (down). See methods section and Table 7 for further details

As expected, the final model maintains all the structural features of the uncomplexed molecule. The thrombin structure is very similar in all three models (after superposition of all Cα atoms the RMSD is in the range 0.32-0.56 Å) and to the PPACK-inhibited enzyme (PDB code 1PPB [77]) (the RMSD is 0.67 Å, 0.81 Å and 0.59 Å for the complex with mTBA, TBA-K and TBA-Na, respectively). The largest deviations are localized in more flexible regions, such as the light chain, and in particular its N- and C-terminal tails. With respect to thrombin-TBA structures, thrombin in complex with mTBA is more ordered, thanks to crystal packing interactions. In particular the thrombin light chain is placed between two symmetry related complexes and interacts at the N-terminal tail with phosphates from residues 14 and 15 of the mTBA and at the C-terminal end with the light chain of a second thrombin molecule, whereas the γ autolysis loop (residues 149A-149L) is involved in hydrophobic interactions with residues 89 -92 of a symmetry related molecule. Moreover several hydrogen bonds, direct or mediated by water molecules, connect the two thrombin molecules. This pattern of interactions is comparable to that observed in PPACK-thrombin crystal structure (PDB code 1PPB [77]).

The architecture of the active site is strictly conserved among all the PPACK inhibited thrombin structures and the density corresponding to the inhibitor is very well defined in the present models (Figure 15).

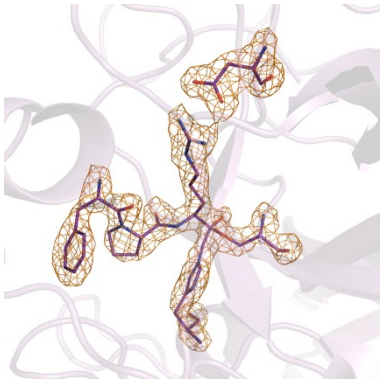

Figure 15 Omit Fo-Fc electron density map (4.5 σ level) of the PPACK inhibitor and active site residues of thrombin-TBA-Na.

PPACK (D-Phe-Pro-Arg-chloromethylketone) forms two covalent bonds with Oγ of Ser195 and Nε of His57, Pro is encapsulated in a hydrophobic pocket formed by several thrombin residues and Phe of PPACK itself and finally Arg engages salt bridges with Asp189.

Another important region is the sodium binding site, located in the 217-225 loop, since this ion plays a key role in modulating thrombin allostery and strongly affects thrombin structure. The analysis of the positive difference peaks generated by an Fo-Fc map and of the coordination geometry clearly indicates the presence of sodium ion near the segment 220-225, in agreement with other thrombin models deposited in the PDB. The ion is hexa-coordinated with octahedral geometry, the six ligands are four water molecules (placed in a range from 2.2 Å to 2.9 Å apart from Na^+) and the carbonilic oxygen atoms of Lys224 and Asp221A (Figure 16). It is notable that the sodium ion is present in all the structures: thrombin affinity for sodium is so high that the ion, already present in the protein sample, either cannot be removed by dialysis against potassium buffers or thrombin binds Na^+ from crystallization solution.

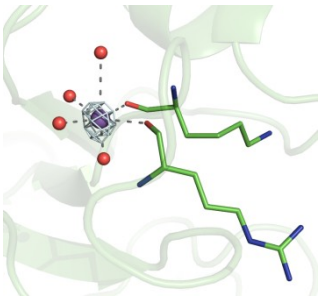

Figure 16 Omit Fo-Fc electron density map (4.5 σ level) of the sodium ion in thrombin-mTBA. Water molecules and thrombin residues that form its coordination sphere are also shown

Furthermore in both the thrombin-TBA structures, two zinc ions from the crystallization solution bind residues on thrombin surface and are involved in packing contacts. Both the ions have tetrahedral coordination, typical for Zn^{2+}: the first one binds His119, Asp1A, Asp222 of a symmetry related molecule and a water molecule, the second one binds Asp60E, Asp125 of a symmetry related thrombin

molecule and two chloride ions, in turn bound to a bridging water molecule (Figure 17). The tetrahedral coordination of the first zinc site is regular (metal-ligand distances in the range 1.8-2.1 Å), whereas it is more distorted in the second one. Indeed zinc-chloride distances are longer than zinc-oxigen ones.

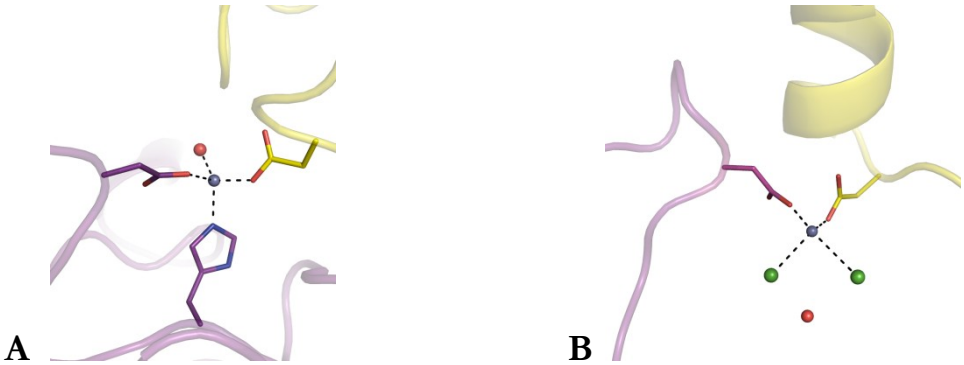

Figure 17 Representation of the two zinc sites on thrombin surface. The two symmetry related thrombin molecules are depicted purple and yellow, respectively. Chloride ions in second zinc site (B) are shown in green

Modified and unmodified thrombin binding aptamers

Electron density maps corresponding to aptamer molecules are clear and continuous, and have allowed the building of the whole TBA and mTBA in the three structures, with the exception of Thy9 base of TBA-Na and Thy7 base of mTBA.

In particular in the case of mTBA electron density corresponding to the polarity inversion site is perfectly defined (Figure 18).

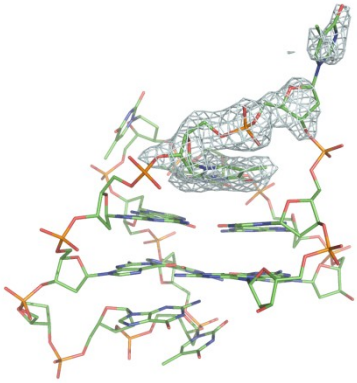

Figure 18 Omit Fo-Fc electron density map (4.0 σ level) of the two mTBA residues (Thy3 and Thy4) involved in the polarity inversion site

X-ray analysis of aptamers in complex with thrombin reveals that all the molecules fold as chair-like quadruplexes with two stacked G-quartets, the TGT loop on one side and the two TT loops on the other side. It must be underlined that TBA adopts almost the same structure both in the presence of sodium or potassium ion, as also suggested by CD spectra. TBA possesses alternated antiparallel strands, and the disposition of glycosidic bond angles along the strands determines an alternation in

each G-tetrad of *anti* and *syn* guanines and in the quadruplex of wide grooves (between strands Gua1-Gua2/Gua14-Gua15 and Gua5-Gua6/Gua10-Gua11) and narrow grooves (between strands Gua1-Gua2/Gua5-Gua6 and Gua10-Gua11/Gua14-Gua15). TT loops span narrow grooves, whereas TGT loop spans a wide groove. This result indicates that between the previously proposed TBA models, the correct one is the NMR model [85, 86]. Differences between modified and unmodified aptamers are observed in strands orientation and grooves dimension. Because of the presence of the polarity inversion site, mTBA possesses three strands parallel to each other and the fourth one oriented in the opposite sense. Along the chain, the bases forming the G-quartets have alternating *syn*/*anti* conformations, as happens in TBA, but, because of the inversion site, within the two tetrads the arrangement of the guanines is *anti-anti-syn-anti* and *syn-syn-anti-syn*, respectively. These sets of glycosidic bond angles give rise to two medium grooves (between Gua1-Gua2/Gua5-Gua6 and Gua1-Gua2/Gua14-Gua15), one narrow groove (between Gua10-Gua11/Gua14-Gua15) and a wide groove (between Gua5-Gua6/Gua10-Gua11). These features are also found in the solution structure of free mTBA [97].

A deeper insight into the quadruplex structures was obtained by defining average G-quartet planes and calculating with respect to them the RMSD of guanine atoms and the angle that each guanine residue describes. Furthermore distance and angle between the two average G-tetrad planes were also calculated. Results obtained are reported in Table 2.

Table 2 Comparison among the core structures of the different models of mTBA and TBA

			RMSD of Gua atoms with respect to the average tetrad plane (Å)	Angle between the two average tetrad planes (°)	Distance between the two average tetrad planes (Å)	Angle between guanine plane and the average tetrad plane (°)			
mTBA	**X-ray model**	Tetrad 1	0.25	0.8	3.40	Gua1 2.3	Gua6 11.4	Gua10 11.3	Gua15 8.9
		Tetrad 2	0.22			Gua2 8.0	Gua5 2.7	Gua11 10.6	Gua14 6.3
	NMR model	Tetrad 1	0.37	1.5	3.70	Gua1 14.7	Gua6 10.6	Gua10 10.3	Gua15 -17.1
		Tetrad 2	0.38			Gua2 -8.0	Gua5 10.4	Gua11 -20.9	Gua14 11.6
TBA	**K+**	Tetrad 1	0.35	0.6	3.45	Gua1 8.9	Gua6 12.2	Gua10 8.0	Gua15 13.5
		Tetrad 2	0.39			Gua2 9.2	Gua5 4.5	Gua11 12.8	Gua14 7.8
	Na+	Tetrad 1	0.39	0.8	3.40	Gua1 10.0	Gua6 12.0	Gua10 11.5	Gua15 15.8
		Tetrad 2	0.44			Gua2 10.1	Gua5 6.9	Gua11 13.7	Gua14 7.6
	1HAO	Tetrad 1	0.35	1.6	3.66	Gua1 3.3	Gua6 12.6	Gua10 2.0	Gua15 17.6
		Tetrad 2	0.36			Gua2 14.2	Gua5 10.9	Gua11 11.8	Gua14 4.6

These results clearly show that in our three models the two G-tetrad planes are more parallel to each other. Apart from this, all the TBA structures are comparable as far as a single G-tetrad is concerned. In the case of mTBA more marked differences exist between our crystallographic model and the NMR one: in the latter structure quartet planes are more distorted.

Larger differences can be observed in the loops of the two mTBA structures (Figure 19). In particular, the TGT loop adopts a different conformation: Gua8 and Thy9, which in the NMR model stack on the quartet formed by Gua1, Gua6, Gua10 and Gua15, loose the interaction with the tetrad, the bases become more parallel to each other and make direct stabilizing interactions. This conformational variation determines the expulsion of Thy7 from the wide groove, so that the base becomes more exposed to the solvent and disordered. With respect to free mTBA, the two TT loops move away from each other, the direct hydrogen bond between Thy4 and Thy13 is lost with the insertion of a bridging water molecule. These two bases lie on a plane that stacks on the second G-quartet formed by Gua2, Gua5, Gua11 and Gua14. Finally, Thy3 adopts a conformation that is completely different in the two models.

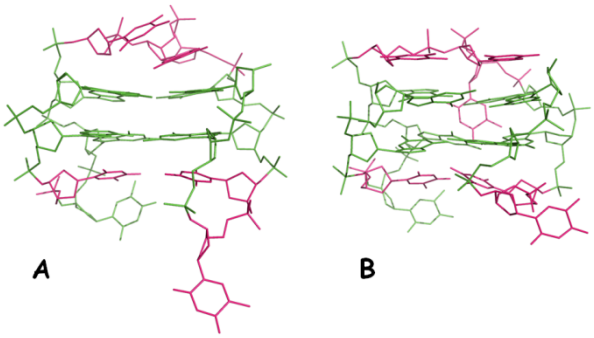

Figure 19 Comparison between the X-Ray (A) and the NMR (B) structures of mTBA. Residues that adopt different conformations in the two models are depicted in magenta.

The movement of TT loops upon thrombin binding observed for mTBA is found in both the structures of TBA with respect to the unbound NMR model, furthermore the Thy4-Thy13 pair described above is also present in our crystal structures of TBA.

The detailed analysis of oligonucleotide structures highlights some interesting features, never stressed before. Whereas in the structures of nucleic acids the C8 of purine in *anti* conformation has been reported to form weak intra-residue C-H•••O interactions with the O5' [101-103], in our models inter-residue contacts between C8 and O4' are observed. A brief analysis of other quadruplex structures (PDB codes: 2AVH, 1JPQ, 1I34) shows that this type of interaction is not unusual in antiparallel quadruplex, where guanine residues with *syn* and *anti* conformations are alternated along the strand. On the contrary it is absent in parallel structures with all *anti* guanines. This interaction is likely to contribute to some extent to the stability of the G-quadruplex.

Examination of the Fo-Fc difference map clearly shows the presence of an ion sandwiched between the two G-quartets of the three aptamer structures (Figure 20). Alkaline ions (usually K^+ or Na^+) coordinated to the buried carbonyl oxygens of the nucleotide bases are crucial for the stabilization of quadruplex structures, since they balance the strong negative electrostatic potential created by the guanines O6 oxygen atoms [104]. In the case of mTBA and TBA-K, the residual electron density is spherical and almost equidistant from the eight O6 atoms of the guanines of the tetrads. The ion, placed between the two quartets, has a coordination number of eight, at distances from the O6 atoms ranging between 2.68 and 2.90 Å, in a distorted anti-prism geometry. This coordination geometry is typical for potassium ion [16], whereas sodium ion often lies on or close to the quartet plane [17, 19]. Moreover, it is well known that potassium, because of its higher ionic radius, is more efficient than sodium ion in stabilizing quadruplex structures [105]. Altogether these findings strongly suggest that the density seen in the residual map is produced by a potassium ion bound to mTBA and TBA-K. Unexpectedly, in the case of the thrombin-TBA-Na structure the residual electron density is placed between the G-quartets, and not on or close to them as expected for sodium ions. However it has an

elliptical and not spherical shape. Therefore two distinct conformations of sodium ion, with the same atomic occupancy, were modelled.

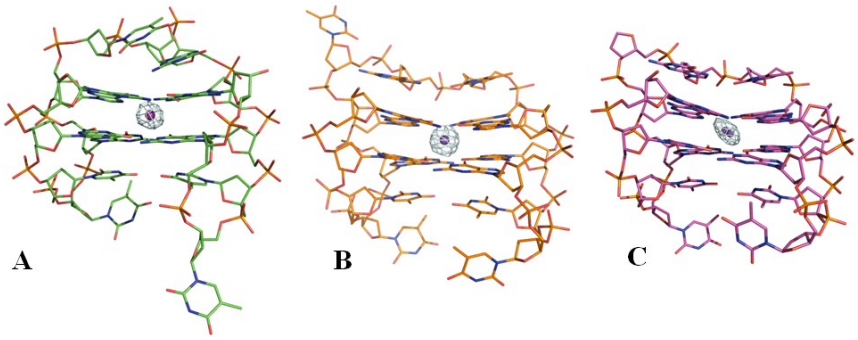

Figure 20 Omit Fo-Fc electron density map (6.0 σ level) of the ion stacked between aptamer quartets of mTBA (A), TBA-K (B) and TBA-Na (C)

This difference could be due to the different ionic radius of Na^+ and K^+. The larger potassium ion fits better in the cavity lined by the guanine residues and occupies a unique site that is equidistant from the two planes, whereas for the smaller sodium ion two different positions are possible.

These results underline the importance of good quality high resolution diffraction data: indeed in the previous crystallographic model of TBA no ion was identified in the structure. Furthermore interactions between quadruplex and positive ions are not directly visible in NMR structures.

Special attention has been devoted to investigate the nucleotides water interactions that are decisive for stability, dynamics, and recognition, since this is another structural feature not directly derivable from NMR. The high resolution of the diffraction data gives the opportunity to study in details the organization of the water molecules bound to the aptamers, particularly in the four grooves. In these regions water molecules arrange themselves in a continuous network of hydrogen bonds, linking the backbone and the bases of the quadruplex.

It has been previously observed that water molecules are preferentially placed close to N2 and C8 atoms of guanines in G-quadruplex [106, 107]. This feature is confirmed in the three aptamers structures. Moreover in both TBA and mTBA a conserved network of sequential N2-water-water-C8 hydrogen bonds connecting adjacent strands is present, in particular in both wide and narrow grooves of TBA and in the narrow groove and in one of the two medium grooves of mTBA. New and less regular patterns of hydration, involving N2, N3 or backbone oxygens and a greater number of water molecules between adjacent strands, are observed in TBA and in the wide groove and in the other medium groove of mTBA. Thymine bases, being primarily involved in the binding with thrombin, are less hydrated than guanines in all the models. The non-conventional binding of waters to the C6 atom, observed, for example, for thymines of the quadruplex of *Oxytricha nova* telomeric sequence [17, 19] or

for all thymines of B-DNA [108], in the case of TBA and mTBA is marginal and involves only Thy9 of TBA-K and Thy4 of mTBA. Non conventional hydrogen bonding of water to methyl group of thymines is preferred in both the aptamers. This rich pattern of hydration involves both the four grooves and the three loops and it is likely to play a key role in maintaining the loop conformation.

Some similarities are observed between the two TBA models and differences with respect to mTBA. For example, concerning the TGT loop, it adopts the same conformation in TBA-K and TBA-Na, with Thy7 and Gua8 stacked on the tetrad formed by Gua1, Gua6, Gua10, Gua15. This conformation is further stabilized by a water molecule interacting to the O6 atoms of Gua6 and Gua10, the N3 atom of Thy7 and the N7 atom of Gua8 (Figure 21).

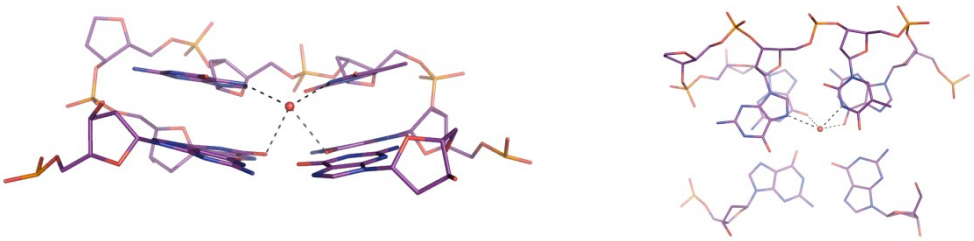

Figure 21 Side view (left) and top view (right) of the TBA-Na TGT loop stacking on G-tetrad formed by Gua1, Gua6, Gua10, Gua15 and the water molecule interacting with Thy7, Gua8, Gua6 and Gua10

This water molecule is absent in the mTBA structure, whose TGT loop adopts a different conformation (described above).

Furthermore in both the TBA models a water molecule links the phosphate groups of Thy13-Gua11 and Thy4-Gua2, blocking the position of the two thymines. Residues Thy3 and Thy12 are more mobile (B factor values are higher with respect to other TBA residues) and indeed they assume two slight different positions in TBA-Na and TBA-K.

Architecture of the complex

The crystallographic structures of the complexes definitely show that the aptamers interact with thrombin exosite I by their TT loops, with a further contribution of Gua2 backbone in the case of mTBA. This feature is in line with the NMR derived model, proposed by Padmanabhan *et al.* in 1996 [88]. It must be underlined, however, that in our structures different binding modes are observed for TBA and mTBA that differ by a 180° rotation of the aptamer with respect to thrombin (Figure 22).

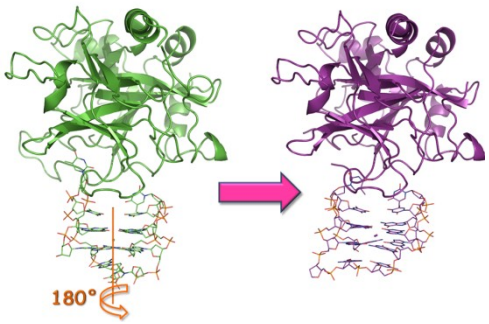

Figure 22 Comparison between the two binding modes found for mTBA (on the left) and TBA (on the right)

Notably, in the old model (1HAO) TBA interacts with thrombin similarly to mTBA.

It should be stressed that, apart from the TGT loop, TBA has an approximate twofold symmetry around the helix axis, whereas this is not true for mTBA, because of the inversion site. As the TGT loop is far from the protein binding site, two practically equivalent binding modes are allowed to TBA. Indeed, a deeper analysis of aptamers-protein interactions shows that most of them are preserved in both binding modes, with the only difference that they involve aptamer residues related by the twofold symmetry.

Thrombin-aptamers interactions

Upon binding the aptamer undergoes to a slight structural rearrangement, whereas the thrombin molecule strictly retains its original structure (as observed in the PPACK-inhibited molecule, PDB code 1PPB [77]), except for the flipping of the peptide bond between Arg77A and Asn78. In the free enzyme residues 77A-79 are involved in a type I β-turn, which upon flipping converts in a type II β-turn. In this new arrangement the residue in position i+3 of the turn (Asn78) adopts a conformation in the left-handed α-helical (Lα) region of Ramachandran map that is not a favoured region for branched residues, although asparagine is often found to adopt the Lα conformation [109]. This modification of the protein appears to have a functional role, as it optimizes the contacts with the bound aptamer: the carbonyl oxygen of Arg77A is hydrogen bonded to the N1 and N3 atoms of TBA Thy4 or mTBA Thy13 (Figure 23).

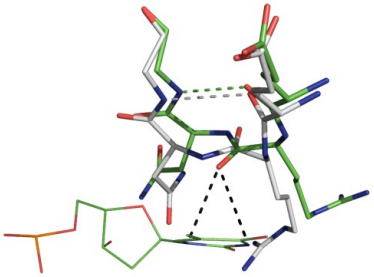

Figure 23 The 77-79 β-turn of thrombin-mTBA (green) and PPACK-thrombin (grey). Interactions between O atom of Arg77A and N1 and N3 atoms of Thy13 in thrombin-mTBA complex are also shown

Interactions between aptamers and thrombin are both hydrophobic and hydrophilic. As an example in Figure 24 and Figure 25 are represented some of them for TBA-K and mTBA, respectively.

Figure 24 Examples of thrombin-TBA-K interactions. On the left disposition of Thy3 with respect to hydrophobic thrombin residues, on the right hydrogen bonds involving Thy4

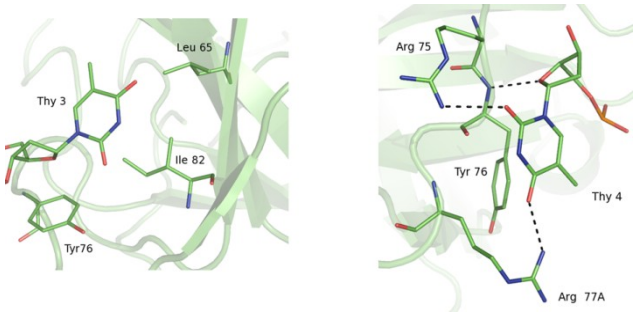

Figure 25 Examples of thrombin-mTBA interactions. On the left disposition of Thy3 with respect to hydrophobic thrombin residues, on the right hydrogen bonds involving Thy4

A detailed summary of interactions between thrombin and aptamers is reported in Table 3 and Table 4. The analysis of thrombin-aptamer interfaces shows that some protein residues are involved in the aptamer binding in all the complexes (Arg75, Tyr76, Arg77A and Ile79), other residues bind specifically mTBA (Leu65 and Ile82) or TBA (His71 and Tyr117) (Figure 26).

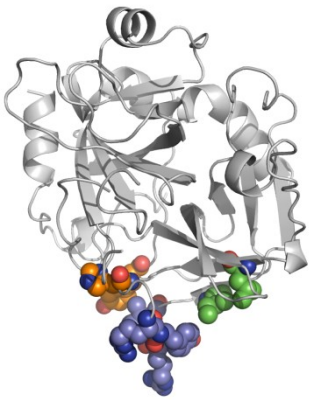

Figure 26 Cartoon representation of thrombin with residues involved in aptamer binding represented as spheres. Residues that bind both aptamers are coloured blue, those that bind only TBA are coloured orange, those that bind only mTBA are coloured green

Finally some slight differences exist between the thrombin-TBA-Na and thrombin-TBA-K interfaces. For example TBA-Na, but not TBA-K, is in contact with Ile24; on the contrary hydrophobic interaction with Asn78 observed in the case of thrombin-TBA-K is missing in thrombin-TBA-Na.

Table 3 Polar interactions (hydrogen bonds and salt bridges) between thrombin and aptamers

Thrombin-mTBA		**Thrombin-TBA-K**		**Thrombin-TBA-Na**	
mTBA residue	**Thrombin residue**	**TBA-K residue**	**Thrombin residue**	**TBA-Na residue**	**Thrombin residue**
Gua2 (O2P)	Arg77 (NH1, NH2)	-	-	-	-
-	-	Thy3 (O4)	His71 (ND1)	Thy3 (N3)	Glu77 (OE2)
Thy4 (O4')	Tyr76 (N)	Thy4 (O4)	Arg75 (NH2)	Thy4 (O4)	Arg75 (NH1, NH2)
Thy4 (O2)	Arg75 (NH2)	Thy4 (O2)	Arg77A (NH1)	Thy4 (O2)	Arg77A (NH1)
Thy4 (O4)	Arg77A (NH1)	Thy4 (N1)	Arg77A (O)	Thy4 (N1, N3)	Arg77A (O)
-	-	Gua5 (O4')	Arg77A (NH2)	Gua5 (O4')	Arg77A (NH2)
Thy12 (N3)	Glu77 (OE2)	-	-	-	-
Thy13 (O4)	Arg75 (NH1, NH2)	Thy13 (O2)	Arg75 (NE, NH2)	Thy13 (O2)	Arg75 (NH2)
Thy13 (O3')	Asn78 (ND2)	Thy13 (O4')	Tyr76 (N)	Thy13 (O4')	Tyr76 (N)
Thy13 (N1)	Arg77A (O)	Thy13 (N3)	Tyr76 (O)	Thy13 (N3)	Tyr76 (O)
Thy13 (N3)	Arg77A (O)	-	-	-	-

Table 4 Hydrophobic interactions between thrombin and aptamers

Aptamer residue	Thrombin residue		
	Thrombin-mTBA	**Thrombin-TBA-K**	**Thrombin-TBA-Na**
Thy3	Tyr76, Leu65, Ile82	His71, Ile79, Tyr117	Ile24, Ile79, Tyr117
Thy4	Tyr76	Asn78, Ile79	Ile79
Thy12	Ile79	-	Tyr76
Thy13	Asn78, Ile79	Tyr76	Tyr76

Furthermore, thanks to the good quality of the electron density maps, several water-mediated interactions between aptamers and thrombin were also identified. Some contacts that are direct in the case of TBA are water mediated for mTBA (for example those with His 71 or Tyr117), in other cases water mediated interactions are similar for both the modified and unmodified aptamers (such as those with Gly25, Thr74 and Gly69) and finally some of them are observed only for one of the two aptamers (Lys81 and Ser115 for mTBA, Ser27 and Lys70 for TBA).

Comparison among thrombin-aptamer and other protein-DNA complexes

In order to obtain a deeper characterization of thrombin-mTBA and thrombin-TBA complexes, several interface parameters, including the buried surface area (ΔASA), the "gap volume index", the number and type of residues at the interface, the contribution to the interface by polar, non-polar and neutral atoms, were calculated and compared with values obtained for the old thrombin-TBA complex (PDB code 1HAO) and with mean values of these parameters obtained using a non-redundant dataset of 25 protein-DNA complexes (Table 5).

Notably, the thrombin-aptamers interface is much smaller than that generally found in protein-DNA complexes, it contains an unusual large number of non-polar atoms and it is constituted also by residues with a low propensity to bind nucleic acids, such as leucine, isoleucine and glutamic acid residues [110]. These features are observed for both mTBA and TBA. A remarkable feature of thrombin-aptamers interface is tight packing. The surface complementarity index (Sc) value obtained is one of the highest observed for protein-DNA complexes. On the contrary a particularly small value is obtained for the old thrombin-TBA complex. This could indicate a worse binding mode or could depend on the low accuracy of the old structure. It should be underlined that in thrombin-aptamers complexes the typical features of aptamer-target interfaces are observed: a particularly small interface area characterized by an excellent shape complementarity and a high number of intermolecular

interactions. Indeed it is peculiar that aptamers realize high affinity binding to their target thanks to a tertiary structure, which optimizes shape complementarity and intermolecular contacts [64].

It is interesting to note that some differences exist between thrombin-TBA-Na and thrombin-TBA-K at the protein-aptamer interface. The first structure is characterized by a larger buried surface area, a higher number of interface residues and intermolecular interactions. The binding affinity of mTBA and TBA to thrombin has been characterized by ITC measurements in a potassium buffer (10 mM potassium phosphate pH 7, 70 mM KCl, 0.1 mM EDTA) [89], whereas a successive ITC study has investigated the thrombin-TBA recognition in a sodium buffer (50 mM sodium citrate pH 6.5, 200 mM NaCl, 0.1% PEG 8K) [90]. Dissociation constant values obtained for TBA are comparable (K_d is ≈300 nM in potassium buffer and ≈100 nM in sodium buffer). To verify whether the observed differences between the two new thrombin-TBA structures correspond to a different binding affinity, a new calorimetric parallel analysis, in which the only variable is the type of cation present in solution, is necessary.

Table 5 Protein interface properties

	Thrombin-mTBA	Thrombin-TBA-K	Thrombin-TBA-Na	Thrombin-TBA (1HAO)	Protein-DNA complexes *
Sequence segmentation	4	5	5	5	7 (4)
Δ ASA (Å^2)	570	445	502	462	1320 (495)
Δ %ASA	4.8	3.8	4.4	4.0	13 (7)
Atoms in Interface	60	39	48	40	122 (51)
% Polar Atoms Contribution to Interface	24.3	26.3	23.7	25.0	30 (7)
% Non-Polar Atoms Contribution to Interface	54.2	33.1	41.9	38.0	30 (11)
% Neutral Atoms Contribution to Interface	45.5	40.1	34.0	37.7	40 (9)
Residues in Interface	15	10	13	12	43 (20)
% Polar Residues in Interface	46.7	50.0	53.8	58.3	37 (10)
% Non-Polar Residues in Interface	33.3	20.0	30.8	25.0	28 (9)
% Charged Residues in Interface	20.0	30.0	15.4	16.7	35 (9)
Hydrogen Bonds	11	10	11	2	21 (9)
Hydrogen Bonds (/100 Å^2 ΔASA)	1.9	2.2	2.2	0.5	1.6 (0.4)
Bridging Water Molecules	4	0	3	0	17 (10)
Bridging Water Molecules (/100 Å^2 ΔASA)	0.7	0	0.6	0	1.3 (0.7)
Gap Volume (Å^3)	2710	2281	2808	2997	7408 (3050)
Gap Volume Index (Å)	2.38	2.6	2.8	3.49	3.0 (1.0)
Sc	0.73	0.72	0.76	0.56	0.65 (0.05)

*properties were calculated from a non-redundant dataset of 25 protein-DNA complexes belonging to different protein-structure families with resolution better than 2.5 Å.

Stoichiometry of the thrombin-aptamer complexes

Once defined that protein-aptamer interactions involve exoxite I of thrombin and the TT loops of the aptamers, we analysed the crystal packing of the new and the old thrombin-aptamers complexes.

Although these complexes have different crystal packing, the aptamers in all cases interact with symmetry related thrombin molecules by their TGT loop. However, these packing contacts differ from one complex to the other (Figure 27).

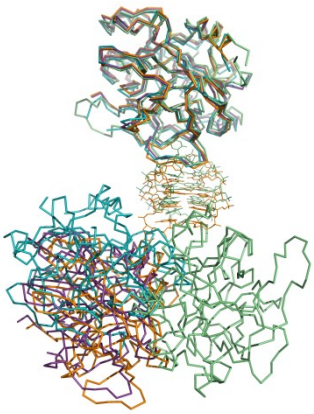

Figure 27 Comparison between the putative 2:1 thrombin-aptamers complexes (mTBA: green, TBA-Na: purple, TBA-K: orange and TBA-1HAO: cyan)

In the old thrombin-TBA structure the TGT loop interacts extensively with the exosite II of the second molecule. In the case of mTBA the additional packing interactions of the TGT loop are far away from the exosite II and are limited to Thy9 that makes a cation-π interaction with Arg97, two hydrogen bonds with Glu97A, which in turn is hydrogen bonded to Gua15. In our structures of the TBA complexes, the TGT loop loosely interacts with two symmetry related thrombin molecules; only for one of them one residue of the exosite II is marginally involved . In all cases, these additional packing contacts are weaker and bury a much smaller surface area with respect to that involving the loop TT and exosite I. Details are reported in (Table 6). These observations suggest that the interactions involving thrombin regions distinct from exosite I are aspecific ionic contacts, that result from crystal packing, and do not imply a second binding site for a putative 2:1 complex in solution. These findings are in agreement with the more recent ITC data ([90]). In conclusion it is reasonable to say that the complex in solution has a 1:1 stoichiometry maintained by a specific recognition of the aptamer TT loops by the thrombin exosite I.

Table 6 Comparison among thrombin-aptamer interfaces in crystal structures of thrombin-TBA-K, thrombin-TBA-Na, thrombin-mTBA, thrombin-TBA-1HAO

		Buried Surface Area (Å^2)	% Buried Surface Area
Thrombin-TBA-K	Exosite I	445	3.8
	TBA-sym1	280	2.4
	TBA-sym2	204	1.7
Thrombin-TBA-Na	Exosite I	502	4.4
	TBA-sym1	270	2.3
	TBA-sym2	148	1.3
Thrombin-mTBA	Exosite I	570	4.8
	mTBA-sym	248	2.1
Thrombin-TBA (1HAO)	Exosite I	462	4.0
	TBA-sym (exosite II)	468	4.1

Furthermore, whereas in the old thrombin-TBA complex model interactions at the second interface extensively involve thrombin exosite II, in the case of mTBA-sym and TBA-sym1 the binding site is placed far away from this exosite, and in the case of TBA-sym2 it marginally involves just one exosite II residue. These observations suggest that the interactions involving thrombin regions distinct from exosite I are aspecific ionic contacts, that result from crystal packing, and do not imply a second binding site for a putative 2:1 complex in solution. These findings are in agreement with the more recent ITC data [90]. In conclusion it is reasonable to say that the complex in solution has a 1:1 stoichiometry, maintained by a specific recognition of the aptamer TT loops by the thrombin exosite I.

About the different biological properties of modified and unmodified aptamers

An extensive comparison between thrombin-mTBA and thrombin-TBA structures has been performed with the aim to investigate the molecular bases of the different biological properties of the two aptamers. With respect to the old crystallographic model proposed by Padmanabhan *et al.*[88], the peculiar structure of mTBA determines the formation of a greater number of contacts with the enzyme, hides a larger buried surface area (570 *vs* 462 Å^2 for mTBA and TBA respectively) and leads to an increased shape complementarity (the surface complementarity indexes Sc are 0.72 and 0.56 for mTBA and TBA respectively as reported in Table 5). The comparison of the three new crystal structures highlights new features: although number of interactions and Sc values are comparable, in all the cases a smallest buried surface area is present in complexes of thrombin with unmodified aptamer.

Altogether these findings suggest that the polarity inversion site in mTBA produces a better fit of the aptamer at the thrombin exosite I and accounts for the higher thrombin affinity of mTBA as compared

to TBA. Furthermore these results agree with some thermodynamic features of the thrombin-aptamers recognition phenomenon: the complex formation is enthalpically driven and the higher thrombin affinity of mTBA derives from its larger favourable enthalpy. On the contrary a less adverse entropy contribute has been found for TBA [89]. This could be rationalized by noting that, thanks to its symmetry, two different binding modes are available to TBA and probably not to mTBA.

Analysis of the crystal structures sheds lights also on the lower inhibitory activity of mTBA with respect to TBA. An important interaction found in the case of the unmodified aptamer involves His71, a residue that plays a key role in recognition of fibrinogen and inhibition of thrombin clotting activity by TBA [111]. This contact between His71 and Thy12 (or Thy3, depending from the two possible binding modes of TBA) is lost in the case of mTBA. The position of this histidine is the same in all the complexes, but the characteristic structure of mTBA causes a subtle reorganization of the region close to the thymine that determines the lost of the His71-Thy12 (Thy3) contact and the formation of a salt bridge between the side chain of Arg77A and the phosphate oxygen atom of Gua2, a base in close proximity due to the reversed polarity of the strand. In the thrombin-TBA complexes the aptamer backbone is far away from Arg77A, whose side chain is flipped in the opposite direction. This causes the slight conformational change of the 77-79 segment mentioned before. In all the complexes Ile79 is in hydrophobic contact with Thy12 (Thy3), however to preserve this interaction mTBA moves away from His71 (Figure 28). The absence of this direct interaction between aptamer and His71 could determine the lower anticoagulant activity of mTBA.

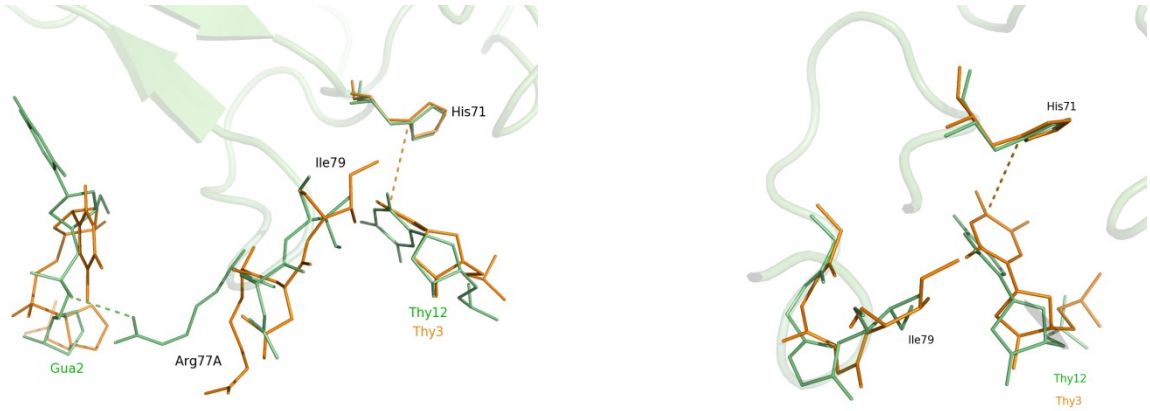

Figure 28 Comparison between thrombin-mTBA (green) and thrombin-TBA-Na (orange) interactions at exosite I region. Interactions between TBA Thy12 and His71 and between mTBA Gua2 and Arg77A are explicitly shown. On the right a zoom on the His71-Thy3 contact is reported

CONCLUSIONS AND FUTURE PERSPECTIVES

Aptamers are considered very promising anticoagulant agents. They typically exhibit high affinity and specific binding to their target protein, which should allow for therapeutic dosing at submicromolar levels reducing potential non-specific effects, they are supposed to be non-immunogenic and antidotes can be designed to control their pharmacologic activity [64]. Since α-thrombin is the key enzyme of the coagulation cascade, it represents an ideal target for anticoagulant therapeutics and, indeed, a few years ago the Phase I clinical trial of TBA was started [112]. Despite this encouraging scenario, up to now the rational design of new aptamer based anti-thrombotic agents has been prevented by the absence of a clear knowledge of interactions between thrombin and TBA, due to the lack of an unambiguous and detailed structure of their complex. In this context the complexes structures here presented are of particular interest, because they not only clarify unresolved questions about thrombin-aptamers interaction, but also suggest the molecular bases of the different properties of TBA and mTBA and furnish a good starting point for the rational design of new modified aptamers for use in anticoagulant therapies. Moving on from these observations it is possible to design modified aptamers that preserve both the interactions essential for inhibition and the stability against nucleases. In particular, the polarity inversion site could be inserted in the second TT loop, between Thy12 and Thy13. In fact a deeper analysis of thrombin structure shows that a hydrophobic pocket, composed by Ile24, His71 and Tyr117 is close to Thy12. The insertion of an inversion polarity site could cause a conformational variation in this TT loop, as observed for Thy3 of mTBA, that could bring the Thy12 base in hydrophobic contact with the three residues, Ile24, His71 and Tyr117. This could improve thrombin-aptamer interactions with respect to unmodified TBA, thus preserving the high thrombin affinity of mTBA. At the same time a better interaction between Thy12 and His71 could be achieved, which should enhance aptamer inhibitory activity. Finally, in order to obtain the best compromise between thrombin affinity and inhibitory activity, the two polarity inversion types, 3'-3' and 5'-5, should be tested, because they should determine different loop conformations.

MATERIAL AND METHODS

Thrombin-aptamers complexes preparation

A similar protocol was utilized to prepare complexes of human α-thrombin with modified and unmodified aptamers. The human D-Phe-Pro-Arg-chloromethylketone (PPACK) inhibited α-thrombin was purchased from Haemtech and two different samples were prepared, changing the initial buffer to KCl or NaCl 0.75 M, using Centricon mini-concentrator and a refrigerated centrifuge. The protein concentration was determined spectrophotometrically, by UV absorption at 280 nm (ε_{280} = 66612 M^{-1} cm^{-1}). Samples at 0.9 mg/ml (25 μM) protein concentration were quickly frozen at -20°C.

TBA was purchased from Genosys, whereas mTBA, synthesized and purified as previously described [97], was kindly provided by prof. Antonio Randazzo. Stock solutions of aptamers were prepared dissolving lyophilized oligonucleotides in sodium (TBA-Na) or potassium (TBA-K and mTBA) phosphate buffer 10 mM pH 7.1 at a concentration of 2 mM. Oligonucleotides concentration was verified spectrophotometrically, measuring absorption at 260 nm and using the same molar extinction coefficient, ε_{260} = 147300 $M^{-1}cm^{-1}$, for TBA and mTBA. The measurements have to be performed on single strand oligonucleotides, so all the samples were previously heated at 90°C and maintained at this temperature during the scan, in order to disrupt any folded structure. To induce the folding into quadruplex structure, nucleotide samples were heated for 10 minutes at 85°C and then slowly cooled down and stored one night at 20°C. Then aptamers folding was verified by circular dichroism measurements. CD spectra, in the 200-350 nm range, were recorded at a constant temperature of 20°C on a Jasco J-710 spectropolarimeter equipped with a Peltier type temperature control system (Model PTC-348WI). Once stated that aptamer assumed an antiparallel quadruplex fold, the complexes with thrombin were prepared by placing a twofold molar excess of aptamer on the frozen sample of inhibited thrombin and leaving the sample three hours at4°C. Then the samples were diluted and the buffer changed to potassium phosphate 50 mM pH 7.1 and KCl 0.1 M, for thrombin in complex with mTBA and TBA-K, or sodium phosphate 50 mM pH 7.1 and NaCl 0.1 M, for thrombin-TBA-Na. The thrombin-aptamers complexes were extensively washed to take off aptamer excess and finally concentrated to about 8 mg/ml using Centricon mini-concentrator and a refrigerated centrifuge. It was not possible to determine spectrophotometrically the complex concentration because thrombin and aptamer absorption bands are superposed and no well resolved peak is observed in the UV spectra.

Crystallization

Thrombin-mTBA complex

Initial crystallization trials of thrombin-mTBA complex were performed at 20°C replying the crystallization conditions found for thrombin-TBA complex [87] or other thrombin complexes [113, 114], but none of these trial was successful. Therefore other crystallization experiments were done using the sparse-matrix sampling method. The screening was carried out using commercially available kits (Crystal Screen kits I and II and Index kit from Hampton Research) and performing the trials both by hand and by a Nanodispersion Robotic System, "Microlab Starlet Nanojet 8+1" by *Hamilton Life Science Robotics*, at the Istituto di Biostrutture e Bioimmagini of CNR. Bushes of thin and damaged crystals (Figure 29) grew from a crystallization condition consisting of polyethylene glycol 4000 25% w/v, ammonium sulphate 0.2 M, sodium acetate 0.1 M at pH 4.5.

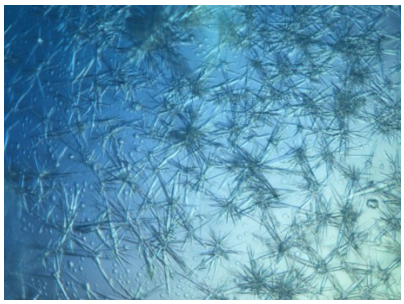

Figure 29 Crystal of thrombin-mTBA before optimization of crystallization conditions

In order to obtain crystals suitable for diffraction data collection it was necessary to perform a minute optimization of the crystallization conditions. Initially the concentration of protein and of precipitant, the pH values and the nature of the buffer were fine-tuned, also performing the crystallization trials at different temperatures (4 and 20°C). Since none of these trials brought to a substantial improvement of crystal quality, a screen of polyethylene glycols and methyl substituted poly(ethylene) glycols with different molecular weights was also tried. Good results were obtained using PEG 20000. It was possible to collect diffraction data from these crystals, but the resolution was low (2.99 Å). Indeed crystals were still too thin and fragile and most of them broke down when touched with the loop. Thus, in order to avoid this problem, the effect of several additives was tested. In particular alcohols (ethanol, 2-methyl-2,4-pentandiol, isopropanol, n-propanol, n-butanol, tert-butanol, glicerol), sugars (treaolose, sucrose and xylitol), and organic liquids (dimethylformamide, acetonytril, dimethylsolfoxide) were added to the crystallization mixture at concentrations ranging from 0.5% to 10% v/v.

Eventually good quality crystals of the complex were obtained from a crystallization solution that contained polyethylene glycol 20000 20% w/v, ammonium sulfate 0.2 M, n-propanol 3%, sodium

acetate 0.1 M at pH 5.8. Crystals grew in few days by sitting drop vapor diffusion method at 20°C. The protein-to-reservoir solution ratio was 1:1 in 1 µl drops and protein concentration was 8 mg mL^{-1}.

Thrombin-TBA complexes

The crystallization condition found for thrombin-mTBA complex was tested to grow thrombin-TBA complexes too, but only amorphous precipitates were obtained. Thus, also in this case, it was necessary to perform an extensive screening of crystallization conditions, using commercial kits and the automatic crystallization system. Two different promising conditions were identified, in which small and thin crystals, morphologically distinct, grew (Figure 30). The first solution contained polyethylene glycol 8000 20% w/v, zinc acetate 0.2 M, sodium cacodilate 0.1 M pH 6.5, the other one consisted of MPD 50% v/v, ammonium phosphate 0.2 M, Tris/HCl 0.1 M pH 7.5.

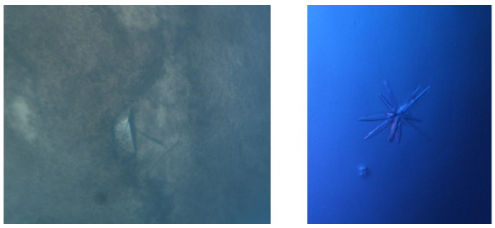

Figure 30 The two crystal forms (on the left crystals grown from PEG, on the right those grown from MPD) of thrombin-TBA complex before optimization of crystallization conditions

In order to improve crystal quality a minute optimization of these crystallization conditions was performed. The concentration of precipitants and salts and the pH values were fine-tuned. Furthermore crystallization trials were performed using both the hanging drop and the sitting drop techniques. Although crystals obtained by the first technique seemed to be better, they were particularly fragile and disordered. Eventually crystals of both the complexes were obtained by sitting drop vapour diffusion method, at a protein-to-reservoir solution ratio of 1:1 in 1 µl drops. The crystallization solution containing polyethylene glycol 8000 10-15% w/v, zinc acetate 50 mM, sodium cacodilate 0.1 M pH 7.4 produced good crystals for both the samples of thrombin-TBA complex (thrombin-TBA-Na and -TBA-K). Furthermore, crystals of only thrombin-TBA-Na were obtained in MPD 53% v/v, ammonium phosphate 50 mM, Tris/HCl 0.1 M pH 7 (thrombin-TBA-Na_2) (Figure 31).

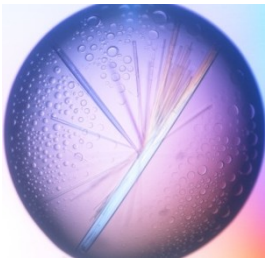

Figure 31 Best crystals of thrombin-TBA-Na_2 grown from MPD

Data collection

Diffraction data from thrombin-mTBA complex and thrombin-TBA complexes grown in PEG were collected in-house on a Saturn944 CCD detector. The X-ray radiation used was Cu *Ka* radiation from a Rigaku Micromax 007 HF generator. Diffraction data from thrombin-TBA complex grown in MPD were collected by using synchrotron light at ELETTRA, Trieste. After the addition of 25% glycerol to the harvesting solution, all crystals were flash-cooled at 100K in supercooled N_2 gas produced by an Oxford Cryosystem and maintained at 100 K during the data collection.

All the data were indexed, processed and scaled with HKL2000 [115]. Thrombin-aptamers crystals belong to three different space groups and diffract at different resolutions. Detailed statistics on data collection are reported in Table 7.

Matthews coefficient calculation suggested the presence of a single thrombin molecule in the asymmetric unit of thrombin-mTBA and thrombin-TBA crystals grown from PEG, whereas three thrombin molecules should be in the asymmetric unit of the thrombin-TBA-Na_2 crystals. Due to their low resolution (2.8 Å), the diffraction data collected from these crystals have not yet been used for structure determination.

Table 7 X-ray data collection statistics for thrombin-aptamers complexes

	Thrombin-TBA-K	**Thrombin-TBA-Na**	**Thrombin-TBA-Na_2**	**Thrombin-mTBA (low resolution)**	**Thrombin-mTBA (high resolution)**
Space group	P1	P1	P3(1)	I222	I222
Unit-cell parameters					
a (Å)	43.369	43.018	94.253	68.211	68.208
b (Å)	45.299	44.716	94.253	111.171	110.574
c (Å)	51.519	52.689	124.180	120.566	120.916
α (°)	68.975	65.222	90.000	90.000	90.000
β (°)	86.114	85.327	90.000	90.000	90.000
γ (°)	69.662	69.970	120.00	90.000	90.000
Resolution limits (Å)	50.00-2.05 (2.12-2.05)	50.00-1.80 (1.86-1.80)	50.00-2.80 (2.90-2.80)	50.00-2.99 (3.10-2.99)	50.00-2.15 (2.23-2.15)
No. of observations	537199	645846	615127	541181	817402
No. of unique reflections	20850	28160	30359		25096
Completeness (%)	94.5 (84.5)	90.9 (75.6)	99.2 (92.8)	92.0 (85.5)	99.1 (93.2)
I/σ(I)	16.8 (2.6)	19.4 (2.8)	12.3 (2.1)	7.6 (3.7)	9 (3.0)
Average multiplicity	5.0	4.2	4.6	4.5	5.2
Rmerge (%)	8.8 (36.4)	6.8 (40.2)	12.1(35.4)	13.4 (27.4)	13.7(42.7)
Mosaicity	0.8	1.2	0.8	0.9	0.6
V_M ($Å^3Da^{-1}$)	2.26	2.21	2.72		2.92
Solvent content (%)	45.7	44.3	54.8		57.9

Rmerge = $\sum_{hkl}\sum_i |I_i(hkl)-\langle I_i(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is ith intensity measurement of the reflection hkl, including symmetry related reflections, and $\langle I_i(hkl)\rangle$ is its average.
Values in brackets are for the highest resolution shell.

Structure determination and refinement

All the structures, solved by molecular replacement or Fourier analysis, were refined with CNS. At the end of refinement the geometry of all the protein structures were monitored using PROCHECK [116] and WHATCHECK [117] programs.

Thrombin-mTBA complex

The structure of thrombin-mTBA was solved using the program AMoRe [118] and the PPACK-inhibited thrombin structures (PDB code 1PPB) as a search model. To avoid bias, inhibitor and water molecules were removed from the model. The correlation for the highest solution was 0.59, with an R factor of 0.36. The starting model was subjected to few cycles of rigid body refinement with CNS [119] in the resolution range 10.0-3.00 Å. This resulted in a model with R=0.327 and R_{free}=0.338. Fourier difference maps, calculated with (Fo-Fc) and (2Fo-Fc) coefficients, showed continuous electron density in proximity of thrombin exosite I. The presence in these maps of two parallel planes allowed the initial fitting of the eight guanine residues of the mTBA core. The structure was subjected to several cycles of coordinate minimization and B-factor refinement. Each run was alternated with manual model building using O program [110]. Analysis of Fourier difference maps at following stages of refinement allowed the fitting of PPACK in the active site, the stepwise rebuilding of the whole aptamer molecule, the positioning of two ions, one stacked between aptamer quartets and the second one interacting with thrombin residues, of several water molecules, and finally the identification of an N-acetyl glucosamine residue linked to Asn60G of the heavy chain and absent in the initial search model. During these cycles the R and R_{free} reduced to 0.250 and 0.295, respectively. Finally the resolution range was extended to include low resolution data and the bulk solvent correction was applied. The crystallographic R factor and R_{free} for the final model in the resolution range 50.0-2.15 Å are 0.196 and 0.252, respectively. Definition of topologies and parameters for mTBA and PPACK molecules to be used in CNS refinement was particularly challenging. Concerning mTBA, CNS parameters for nucleic acids were modified in order to include the polarity inversion site and the reverse polarity of the segment Gua1-Thy3. CNS topology and parameter files for the segment Gua1-Thy4 were calculated using the Prodrg (http://davapc1.bioch.dundee.ac.uk/prodrg/) server, then they were merged with default parameters and the correct connectivity between residues was obtained modifying by hand the input files. For what concern the inhibitor molecule, parameters for a fictitious ligand formed by PPACK and thrombin residues Ser195 and His57 to whom inhibitor is covalently bound were calculated using the Prodrg (http://davapc1.bioch.dundee.ac.uk/prodrg/) server. These files were used for PPACK molecule and to modify the default protein topology and parameter files for the side chain of amino acid residues involved in PPACK binding. Furthermore it was necessary to insert in the refinement input files a

specific patch to include the non-conventional covalent bonds between the side chains of thrombin residues and PPACK.

Thrombin-TBA complexes

The structure of thrombin-TBA-K was solved by the molecular replacement method using the program Phaser [120] of the CCP4 package and coordinates of inhibited thrombin derived from the structure of its complex with mTBA and coordinates of TBA derived from the structure of the thrombin complex (PDB code 1HAO) as search models. To avoid bias, PPACK, ions and water molecules were removed from model. A clear solution was obtained with a Z-score of 12.5 and an LLG of 1206. In a second moment diffraction data from thrombin-TBA-Na crystals were collected. Since crystals of the two complexes were isomorphous, it was not necessary to perform molecular replacement: the partially refined structure of thrombin-TBA-K was utilized as starting model for the other complex. Slight rearrangement of the complex assembly were taken in count performing several cycles of rigid body refinement of the initial model, treating TBA and thrombin as distinct rigid groups.

A preliminary comparison between these structures and the other crystal form of the thrombin-TBA complex (PDB code 1HAO) showed that the orientation of the aptamer with respect to thrombin is different in the two cases. The analysis of Fourier difference maps, calculated with (Fo-Fc) and (2Fo-Fc) coefficients, seemed to confirm the novel complex architecture for both thrombin-TBA-K and thrombin-TBA-Na. However, in order to be sure of the correctness of our structures, the whole TBA molecule was removed from the models and rebuilt stepwise: first of all the two G-quartets, then Thy 4 and Thy 13, followed by Thy 3 and Thy 12 and finally residues of the TGT loop. A protocol similar to that described for thrombin-mTBA complex was used for refinement of thrombin-TBA complexes. Also in these cases, electron density maps were particularly clear and allowed the fitting of PPACK in the active site, the positioning of several ions, one stacked between aptamers quartets, one interacting with the same thrombin residues identified for thrombin-mTBA complex, and the others involved in packing interactions, of several water molecules, and the identification of the glycosilation site. After inclusion of low resolution data and bulk solvent correction, the crystallographic R factor and R_{free} for the thrombin-TBA-K final model in the resolution range 50.0-2.05 Å are 0.182 and 0.234 respectively, whereas for the thrombin-TBA-Na final model, in the resolution range 50.0-1.80 Å, are 0.178 and 0.224, respectively.

Structural analysis of aptamers and thrombin-aptamers complexes

A deep characterization of aptamers structures was obtained compiling a program that defines the average G-quartet plane and calculates the RMSD of guanine atoms from this plane, the angle that each guanine residue describes with respect to the plane, the distance and the angle between the two average G-quartet planes. The characterization of hydration has been obtained identifying the networks of

water molecules with HBPLUS [121] and by visual inspection of the structure, whereas unconventional CH•••O hydrogen bonds have been identified using HBAT [122].

Features of the protein-aptamer interfaces, including buried surface area, number and types of residues at the interface, and gap volume index, were calculated using the Protein-Protein Interaction Server (http://www.bioinformatics.sussex.ac.uk/protorp/). For comparison, the same parameters were calculated as average values from a non-redundant dataset of 25 protein-double stranded DNA complexes solved at a resolution better than 2.4 Å. The shape complementarity score, Sc, as defined by Lawrence and Colman [123], was calculated using the CCP4 package [124], opportunely modified to include nucleic acid parameters. For atomic groups in DNA, standard radii values as defined by Nadassy *et al.* were used [125]. Finally the molecular interactions between aptamers and thrombin molecules in the same or in other asymmetric units were found using the CCP4 package [124] (in particular the Contact program) and by visual inspection of the structure with O program.

All the figures were generated using Pymol (http://pymol.org).

PART II

PROTEINS INVOLVED IN NUCLEIC ACIDS RECOGNITION: THE RIBONUCLEASE SYSTEM

INTRODUCTION

The Vertebrate Ribonuclease Superfamily

One of the best known and most studied classes of nucleic acid binding protein is that of ribonucleases. Ribonucleases (RNases), which are enzymes catalyzing the hydrolysis of the phosphodiesteric bond of RNA, play a key role in metabolism of ribonucleic acid [126], operating between transcription and translation. In particular they are involved in RNA degradation, transformation of precursors into mature forms of RNA and production of short segments of regulator RNA. Other important biological functions ascribed to RNases are related to the control of gene expression, growth and cell differentiation, cell protection from the attack of pathogens and induction of apoptosis [127].

A well characterized subfamily is that of pancreatic like ribonucleases, whose prototype is the bovine pancreatic ribonuclease or RNase A. Pancreatic like ribonucleases are endoribonucleases that specifically cleave phosphodiesteric bonds with a pyrimidine base at the 3' end [126]. They usually act on single stranded RNA [128], but some members of the family are able to hydrolyze also double stranded RNA and DNA-RNA hybrid structures [129-131]. Among the signature features of this gene family, all active pancreatic-like ribonuclease coding sequences are initiated with a signal sequence, all mature proteins include six to eight appropriately spaced cysteines that form distinct disulfide bonds, and all include two catalytic histidines and a single lysine, the latter within an invariant sequence motif (CKXXNTF) [132]. As a rule, their coding sequences are found on a single exon [133].

Although the members of the family present primary sequences very different to each other, they share a common three-dimensional structure and catalytic site. In particular the "RNase fold" consists of three α-helices and two β-sheets, each composed of three antiparallel β-strands, arranged in a typical V shape with a deep cleft, where the active site is placed [134].

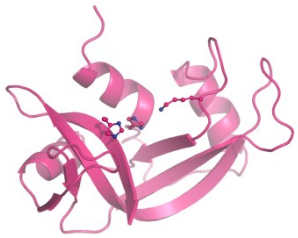

Figure 32 Cartoon representation of RNase A. Catalytic residues (His12, Lys41, His119) are also shown as ball and sticks.

The catalytic centre, as mentioned before, is composed by two histidines, one located in helix I and one located in the C-terminal strand (respectively His12 and His119 in the RNase A numeration) and a lysine (Lys41 in RNase A). The catalytic mechanism consists of two steps: in the first one the imidazolium side chain of His12 acts as a base to extract a proton from the 2'-hydroxyl of the substrate, thus facilitating its attack on phosphorus atom. The acidic His119 side chain protonates the 5'-oxygen, facilitating the departure of the leaving group. In the second step, the enzyme catalyzes the attack on the 2'-3'-cyclic phosphodiester by a water molecule, producing a phosphate monoester on the C3' of the ribose sugar unit of RNA, through the same pathway in reverse. Thus the mechanism of action involves a transphosphorilation reaction followed by an hydrolysis to form a 3' monophosphate nucleotide [126].

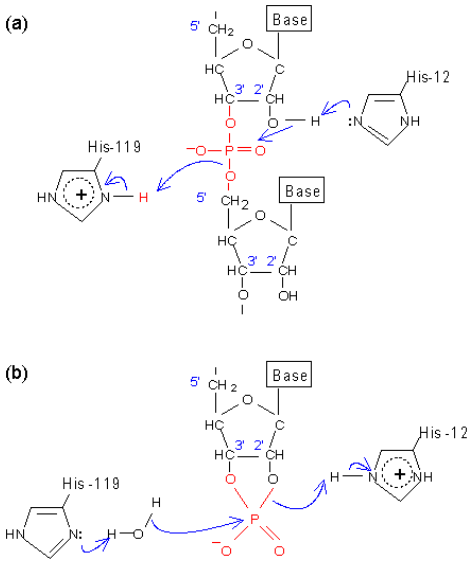

Figure 33 Schematic representation of RNase A catalytic mechanism

The third active site residue, Lys41, also interacts with the phosphate group, stabilizes the trigonal bipyramidal intermediate and can potentially act as a general base in RNA cleavage.

Thanks to crystallographic studies on RNases, especially bovine pancreatic ribonuclease, and their complexes with mononucleotides and oligonucleotides, other sites, apart from the active one, have been identified on RNases surface. These sites play a key role in determining activity and specificity of the enzyme [135]. Thus usually it is preferable to talk about RNases subsites [136-138]. In particular

these subsites are divided into P subsites, in which RNA phosphate groups are bound, and B subsite, which interact specifically with bases. Referring to RNase A numeration, P1 subsite, lined by catalytic residues His12, Lys41 and His119, binds the phosphate group to be hydrolyzed, whereas the P0 (residue Lys66) and the P2 (Lys7 and Arg10) subsites bind the upstream and downstream phosphate groups, respectively. The pyrimidine specificity of enzymes results from the binding of the base in a narrow pocket on the surface of the enzyme, named B1, formed by Val43, Thr45, Phe120, and Ser123, whereas the downstream base is located in the less specific B2 subsite, constituted by the loop 65–72 containing Asn71, Asn67 and Gln69. These sites, well defined in the case of RNase A, can be identified also in other members of the family [139, 140].

The RNase A superfamily has been known for a long time also as tetrapods RNases superfamily, because members have been found in amphibians, reptiles, birds and mammals [141]. Recently, a large number of sequences of fish DNAs and proteins has become available and it makes possible to recognize in fish proteins the unmistakable signature of an RNase sequence (CKXXNTF) [142, 143]. This has led to the proposal to rename the pancreatic-like superfamily more generally as the vertebrate RNase superfamily.

Ribonucleases with special biological functions

Many members of the vertebrate RNase superfamily are endowed with special biological functions, distinct from their catalytic action of degrading RNA, but strictly dependent on it. In particular these functions include angiogenic [144], antifertility [145], anti-pathogen [146], cytotoxic [147] and immunosuppressive [148] activities.

Angiogenins

Angiogenesis describes a process of endothelial cell stimulation to form new blood vessels and is critical for maintenance of tissues and support of growth. This process is essential for embryonic development, wound healing, proliferation of the endometrium and several other physiological processes [149], but plays an important role also in a variety of pathological conditions, such as diabetic retinopathy, psoriasis and neoplasia [150, 151]. Furthermore angiogenesis is crucial for tumour development, through the vascularisation of the tumor mass.

The RNases with angiogenic activity, usually called angiogenins, form a distinct family within the superfamily, and were firstly identified for their ability to stimulate the growth of blood vessels [152] and then recognized as ribonucleases. With respect to the superfamily prototype, RNase A, they have a sequence homology of about 30% and only three disulphide bridges instead of four [153]. They lack the disulphide bond between Cys65 and Cys72, which has been found to be important in determining RNase A activity and stability [154]. Angiogenins are characterized by a very low ribonucleolitic activity (10^3-10^4 lower than RNase A one), which is essential for angiogenic activity and is probably caused by

the obstruction of pyrimidine binding subsite B1 by the side chain of a C-terminal residue (Gln117 in human angiogenin) [139]. Interestingly, it has been proved that an increasing of ribonucleolitic activity determines the lost of angiogenic function. However, it is not yet known how these proteins stimulate blood vessels growth, which are their physiological substrates and why angiogenic function is so strictly dependent on a low catalytic activity.

Angiogenins from fishes

Several fish RNases have been identified and studied in the last years: five recombinant RNases from the tropical bony fish zebrafish (*Danio rerio*)[142, 143, 155, 156], two from the atlantic salmon (*Salmo salar*) [157], two from the rainbow trout (*Oncorhynchus mikiss*) and ten from medaka (*Oryzias latipes*) [143]. These proteins share about 30% sequence homology with RNase A and about 35% with human angiogenin (hAng). Their RNA degrading activity is comparable to that of mammalian angiogenins and, as in the case of these latter enzymes, their angiogenic activity is strictly dependent on the integrity of the catalytic activity. On the other hand, all fish RNases studied so far have been found to be endowed with bactericidal activity. Surprisingly, this latter activity has been found to be maintained when the RNases are either catalytically inactivated or fully denatured.

A phylogenetic analysis (Figure 34) indicates that fish RNases form two distinct, related clusters. One of them comprises the RNases from trout, salmon, and zebrafish, excluding ZF-RNase-4, the other, comprising zebrafish RNase 4 and some medaka RNases, includes also the amphibian RNases. It has been proposed that this indicates the presence of two distinct RNase genes in the last common ancestor of fish and amphibians.

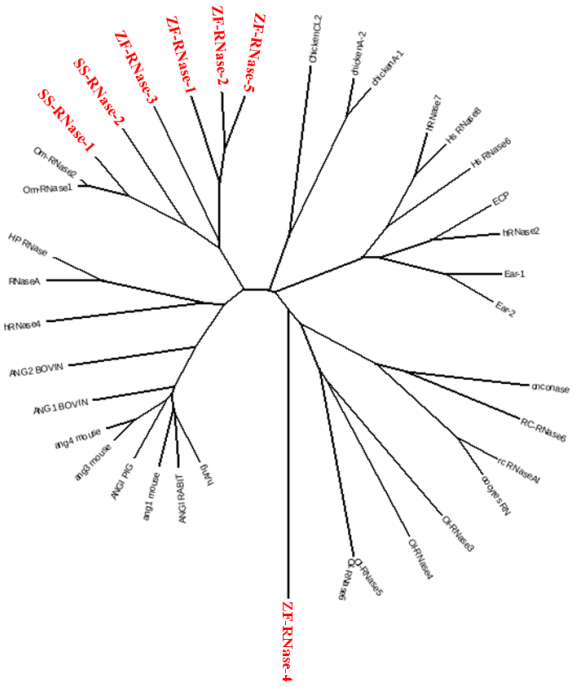

Figure 34 Rootless phylogenetic tree of the vertebrate RNase superfamily. Zebrafish and salmon RNases are marked in red

On the basis of the findings of angiogenic and antibacterial activities of zebrafish and salmon RNases, different hypotheses have been proposed on the physiological and evolutive significance of the early-diverged fish RNases. The relationship of fish RNases to mammalian angiogenins prompts the hypothesis that the whole superfamily of vertebrate RNases evolved from early angiogenic RNases. In the course of evolution, RNases endowed with strong RNase activity were selected for, and used for digestive functions, or diverse non-catalytic bioactions. Separate clades of RNases instead, poorly active as RNA degrading enzymes, conserved the angiogenic activity, all through the divergence of mammals. On the contrary the finding that fish RNases have a strong bactericidal activity suggests that the ancestral RNases were more likely proteins engaged in host defence [158]. Further studies and a deeper characterization of fish ribonucleases may show that the two hypotheses are indeed only one: that the ancestral RNases were engaged both in angiogenic and host defence activities.
A detailed analysis of structural and functional characteristics of these ancestral ribonucleases could provide the opportunity to study evolutionary forces at work on a conserved protein scaffold.

Antitumor RNases

As it happened in the case of angiogenins, also other proteins have been discovered thanks to their biological functions and then recognized as ribonucleases. This is the case of Onconase, a homolog of RNase A present in the oocytes and early embryos of *Rana pipiens*, which shows a high cytotoxic activity towards cancer cells and thus it has been early identified as a promising cancer chemotherapeutic agent [159-161]. Numerous cancer lines are sensitive to Onconase; their treatment with 10-100 nM enzyme leads to suppression of cell cycle progression, followed by apoptosis or cell senescence. Onconase has advanced to a phase IIIb confirmatory clinical trial for the treatment of unresectable malignant mesothelioma [162]. Other amphibian RNases exhibit a specific toxicity to cancer cells, such as ribonucleases extracted from *Rana catesbaiana* (RC-RNase) and *Rana japonica* (RJ-RNase) [163, 164]. All these enzymes are monomeric, share about 30% sequence homology with RNase A and present four disulphide bridges, three of which common to bovine pancreatic ribonuclease. Furthermore they have a particularly high thermal stability compared with other members of the family [161].
Among cytotoxic RNases peculiar is the case of bovine seminal ribonuclease, or BS-RNase, since it is the only natural dimeric protein of the superfamily. In particular BS-RNase is a homodimer that shares more than 80% amino acid homology with the family prototype RNase A. The dimeric structure is stabilized by two intermolecular disulphide bridges (between Cys31 of one chain and Cys32 of the other one). Moreover the enzyme exists as two conformational isoforms [165], one of which (MxM) presents the reciprocal exchange between the two chains of the N-terminal helix [166], a phenomenon absent in the other one (M=M) [167]. The special structural feature observed in MxM BS-RNase is

known as 3D domain swapping.

Although the research concerning antitumor RNases is extremely rich, many aspects of their properties are still unclear. It is evident that their molecular target is RNA, but these proteins have distinct preference to cellular RNAs [168]. For example it has been demonstrated that Onconase *in vivo* acts on tRNA, but not on mRNA and rRNA [169-172], whereas BS-RNase preferentially hydrolyze ribosomial RNA [173]. These results are not valuable *in vitro*, where the enzymes cleave indistinguishably rRNA and mRNA. It has been suggested that the *in vivo* specificity may be caused by a different cellular localization, by protein modifications after internalization or by the different accessibility of RNA when complexed with cellular proteins [168]. The internalization is a crucial step for exploiting the antitumor activity, but it is not fully understood. RNases are basic proteins that well interact with the negative cellular membrane and it has been proved that an increasing of positive charge on the RNases surface rises their cytotoxic activity [174, 175]. Specific receptors have been identified on cellular membrane that bind RC-RNase and RJ-RNase [163], an Onconase receptor has been discovered only on glyoma9 cells [176], whereas no cellular receptor for BS-RNase has been found [168]. Thus it has been proposed that RNases are internalized by endocytosis, and indeed they have been localized in endosoms [177-179]. In contrast to Onconase, which enters the cytosol from the endosome recycling compartment, mammalian RNases are translocated into the cytosol from the endosomal/lysosomal pathway [177, 179]. In the endosomes (as well as in the lysosomes) and in the cytosol the RNases have to resist the attack by proteases. Furthermore, to exploit their cytotoxic function, RNases have to avoid the binding by the proteic ribonuclease inhibitor (RI), which represents the 0.01-0.1% of the whole protein component of mammalian cells cytosol [180]. RI is a 50 kDa protein with the typical leucine rich fold, composed by 15 β-α motifs [181] arranged in a characteristic horseshoe shape [182].

Figure 35 Cartoon representation of the proteic ribonuclease inhibitor, RI (coordinates derived from the structure of RI-RNase A complex, PDB code 1DFJ)

RI contains many reduced cysteine residues, essential for its activity: when one single residues is oxidized all the others are quickly oxidized and RI loses its capability of binding and inhibiting RNases and is digested by cellular proteases [183]. The biological role of this protein is still unclear, probably it exerts a protective function to the cell against secreted ribonucleases. Furthermore it could also regulate

RNases biological activities [184].

RI forms with RNase A a 1:1 complex, which is one of the strongest known proteic complex. It is characterized by a dissociation constant of 10^{-14} M [185] and by an incredible shape complementarity between the two components [186, 187]. Mammalian ribonuclease inhibitor is able to bind many RNases, such as angiogenins [188], but it does not recognize amphibian ribonucleases, which lack several aminoacid residues essential for the binding [189], and the dimeric BS-RNase, which is sterically inaccessible [187, 190], all proteins endowed with cytotoxic and antitumor activities. On the basis of these findings it has been proposed that monomeric RNases can be converted into antitumor agents if modified to evade the binding of RI [191].

BS-RNase

As mentioned before, BS-RNase represents a unique case in the pancreatic ribonuclease superfamily because of its quaternary structure. It is a covalent homodimeric enzyme, isolated as an equilibrium mixture of swapped (MxM-BS) and unswapped (M=M-BS) forms, at 7:3 ratio [165]. The interchanging N-terminus (residues 1–15) is linked to the main body of the enzyme through the residues 16-22 [166]. This peptidic segment, which is called hinge peptide and is very important in determining the swapping phenomenon, adopts different conformations in the two isomers. In both isoforms, the two chains are linked by two antiparallel disulfide bridges, involving Cys31 and Cys32 (Figure 36), and the two Leu residues in position 28 form stabilizing hydrophobic interactions across the molecular twofold axis [166, 167].

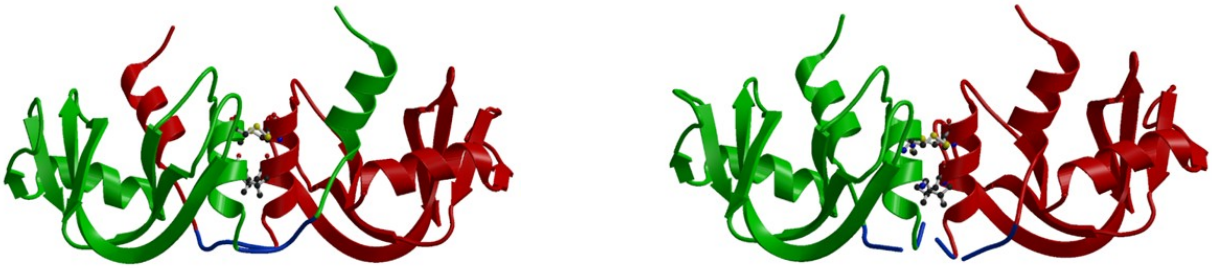

Figure 36 Cartoon representation of the BS-RNase isoforms, MxM on the left (PDB code 1BSR), M=M on the right (PDB code 1R3M). The two subunits are coloured red and green respectively, the hinge peptides are in blue, cysteines involved in interchain bridges are represented in ball and sticks

Although the two dimers present an almost undistinguishable external shape, the cytotoxic action is a peculiar property of MxM-BS [192]. It is believed that this feature is linked to the fate of the enzyme in the cytosol. Due to the reducing property of this cellular compartment, M=M-BS dissociates into monomers (Figure 37 A), which are strongly inhibited by RI [193], whereas MxM-BS survives as a metastable non-covalent swapped dimer (NCD-BS), whose overall shape hinders the formation of the complex with RI [190]. Furthermore, in NCD-BS the cysteines are close to each other (Figure 37 B), and this feature may explain the higher stability of its disulfide bridges with respect to those of M=M.

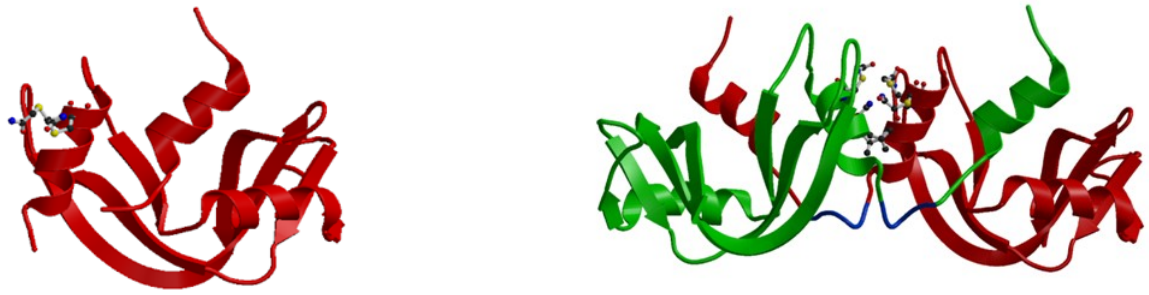

Figure 37 Cartoon representation of the monomer (PDB code 1N1X), on the left, and of the non covalent dimeric form (NCD) (PDB code 1TQ9), on the right, of BS-RNase. In NCD-BS the two subunits are coloured red and green respectively, in both models cysteines involved in interchain bridges are represented in ball and sticks

The 3D domain swapping, first observed in the case of MxM-BS-RNase in 1993, it has been found in the last years in an increasing number of proteins. It has been suggested that 3D domain swapping could play a role in important biological functions, such as up- and down-regulation of enzymatic activities [194], immunosuppressive and antitumoral properties [195, 196], regulation of molecular assembly [197, 198], structural organization of viruses [199]. It has been also proposed that it is a structural feature in the formation of amyloids fibers [200], which are associated with several degenerative diseases such as Alzheimer, type II diabetes, and many others[201]. It is a powerful process to create dimers, since the interactions that pre-exist within the monomer are re-used in the dimer (Closed-Interface) and at the same time novel chain-chain interactions are formed (Open-interface) [197]. Therefore in the present context it represents a promising mechanism to transform non-cytotoxic monomeric ribonucleases into antitumor dimeric or oligomeric enzymes. Indeed many efforts have been devoted to deeply understand the phenomenon and to identify the structural requirements of the BS-RNase swapped structure. This especially because the seminal enzyme presents some risks as therapeutic due to its immunogenicity [202]. This problem could be overcome inducing antitumor activity in pancreatic ribonucleases, in particular in the bovine and human proteins [203]. Transforming these proteins in a NCD-BS-like dimeric form seemed to be an easy task, because of the high sequence homology (more than 80% in both cases). Through a minute analysis of MxM-BS crystal structure a few amino acid residues that are likely to play a crucial role in the swapping phenomenon and in the stabilization of the quaternary structure have been identified [166, 191, 204]. In particular the two cysteine residues in position 31 and 32, requested for the dimerization, the proline in position 19 and the leucine in position 28. Pro19 determines a characteristic conformation of the hinge peptide and stabilizes the quaternary structure occupying a hydrophobic patch on the other subunit lined by residues Tyr25 and Gln101, whereas Leu28, as said before, stabilizes the dimer through symmetric hydrophobic interactions [166, 167, 190]. The importance of these two residues has been proved by the production and characterization of the BS-RNase double mutant, P19→A/L28→Q, with the two residues substituted by the corresponding RNase A ones. This mutant (PALQ) presents a reduced

antitumor activity with respect to the wild type enzyme, and its NCD form has an open structure that allows the binding of RI (Figure 38) [205].

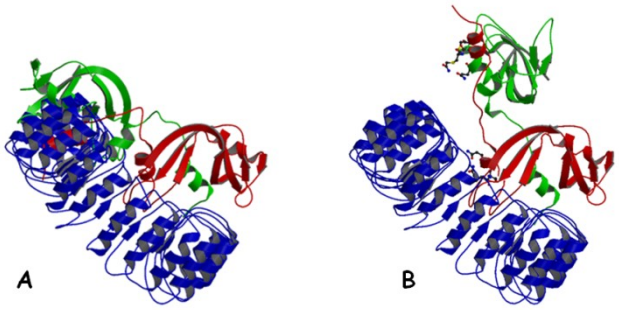

Figure 38 Cartoon representation of the putative complexes between RI and NCD-BS (A) and NCD-PALQ (B). RI is coloured blue, whereas the two subunits of NCD molecules are coloured red and green respectively

Starting from these findings several human pancreatic RNase (HP-RNase) and RNase A mutants with different ability to evade RI and to exert cytotoxic activity have been obtained [191, 204, 206].

New strategies to obtain antitumor RNases: immuno-RNases

A very promising strategy to obtain efficient and selective antitumor therapeutics is based on the use of antibodies, which specifically bind antigens on the external surface of sick cells [207]. During the past two decades, immunotoxins, made up of antibodies or miniantibodies (single chain variable fragment, scFv), fused to toxins, have been proposed as anticancer drugs [208]. These chimeric proteins combine the potent toxicity of toxins with the antigen specificity of antibodies. However, in many cases the murine nature of the immunomoieties, the plant or bacterial nature of the toxin moieties, and their high toxicity have greatly limited the therapeutic potential of these constructs [209]. A very recent strategy in anticancer immunotherapy is that based on immunoconjugates, in which an RNase substitutes for the toxin [210, 211]. Since mammalian RNases are not toxic to cells unless internalized, these fusion proteins are not immunotoxins but rather immunoprotoxins [212]. Furthermore antibodies can deliver a RNase into tumor cells leading to a high RNase concentration in the cytosol. In this way, although some enzymes are inhibited by RI, proteins able to exert their catalytic activity are still available in the cell, through a saturation effect [213].

Because of its preferential expression in tumor cells and its extracellular accessibility, an attractive target for immunotherapy is ErbB2, a trans-membrane tyrosine kinase receptor that is overexpressed on many carcinoma cells of different origin, and is associated with poor prognosis and with a more aggressive clinical behaviour [214, 215]. ErbB2 consists of a cytoplasmatic region, a trans-membrane tract and finally an extracellular domain (ECD) [216], in turn composed by four sub domain (I-IV) [217, 218].

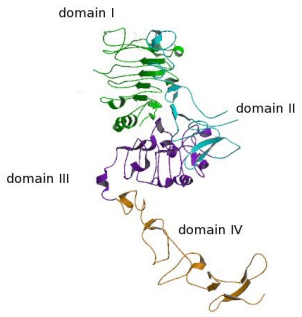

Figure 39 Cartoon representation of the ErbB2 extra-cellular domain (ECD). The four domains are coloured green, cyan, purple and orange, respectively

Several humanized immunoagents against the extracellular domain of ErbB2 have been produced and have been approved for clinical use or tests, such as Trastuzumab [219, 220] and Pertuzumab [221]. They are humanized antibodies of murine origin and present immunogenicity and cardiotoxicity as side effects. To overcome these problems the production of several human immunoagents against ErbB2 has been attempted by prof. D'Alessio and his colleagues [212, 222-224]. In particular, they produced and characterized a human anti-ErbB2 scFv, named Erbicin [223], which *in vitro* is active against tumors that are resistant to other antibodies and both *in vitro* and *in vivo* is not cardiotoxic. This protein has been fused to HP- RNase to obtain an immuno-RNase, which retains the enzymatic activity of the wild-type enzyme and specifically binds to ErbB2-positive cells with the high affinity of the parental scFv [212]. It has been proved that immuno-RNase action is based on its RNase activity, exerted in the cytosol of internalized cells, reached directly from the endosomal compartment. After internalization by target cells, immuno-RNase becomes selectively cytotoxic in a dose-dependent manner at nanomolar concentrations [213]. Remarkably it does not present Trastuzumab drawbacks: it have no adverse effects on cardiac cells *in vitro*, it does not alter cardiac function *in vivo*, when tested on mice [225], and it is active on Trastuzumab resistant tumors. Is has been suggested that these immuno-RNase properties depend on a different binding on ECD, which in turn triggers a different signaling mechanism [213].

Thesis purposes

Concerning the ribonuclease system, the thesis purposes have been multiple.

First of all, I focused my attention on the structural features of fish ribonucleases. During the first year of my PhD, the crystal structures of two zebrafish RNases, ZF-RNase-1 and -3, determined at alkaline pH, were published [155]; my work has thus been devoted to complete the structural characterization of ZF-RNases, through a combination of crystallographic analysis and computational techniques, principally homology modelling. The main purpose was to explain on molecular basis the different stability and activity properties of these highly homologues proteins. Furthermore, the study of a ribonucleases from *Salmo salar* (SS-RNase-2), a fish emerged about 70 million years before zebrafish

genome divergence, has been also attempted, because structural comparison among these fish enzymes could shed light on the evolution of the RNase scaffold from an ancestral protein.

As a second point, domain swapped RNases endowed with cytotoxic activity, but characterized by a low immunogenicity, have been studied. Inspired by the dimeric nature of the cytotoxic BS-RNase, RNase A dimeric variants were designed to disturb the affinity of RI to the RNase A molecule. The structural determination at high resolution of the non covalent swapped dimeric forms (NCD) of RNase A mutants allowed to understand the molecular causes of the specific biological properties of these proteins and represents a crucial step toward the design of an RNase with low immunogenic response and improved antitumoral activity.

Finally, this study has been focused on immunotoxins with specified cell-type cytotoxicity, constructed by fusing human pancreatic RNase to antibodies. In particular, a potent antitumor immuno-RNase was produced by combining HP-RNase with an human anti-ErbB2 scFv, Erbicin [212]. In order to define and implement the antitumor potential of immuno-RNase and to rationalize its properties with respect to other known anti-ErbB2 antibodies, the immuno-RNase epitope on ECD should be identified. The most straightforward method to characterize the intermolecular recognition is the structural determination of the complex using X-ray diffraction. However, when good quality crystals are not available, a powerful alternative is *in silico* analysis, based on homology modelling, rigid docking and molecular dynamic. Since the search for crystallization conditions was unsuccessful we decided to use a computational approach. To simplify the problem it was decided to first characterize the epitope that is recognized by the antibody. The results obtained will be used to clarify the role of ribonuclease in the immunoconjugate action.

RESULTS AND DISCUSSION

Fish angiogenins

In the last years RNases have been identified in genomes of several fishes: in particular five recombinant RNases from the tropical bony fish zebrafish (*Danio rerio*) [142, 143, 155, 156] and two from the atlantic salmon (*Salmo salar*) [157]. These proteins were expressed, purified and characterized by the group directed by prof. D'Alessio. They are endowed with bactericidal and angiogenic activities, the last function being strictly dependent on ribonucleolitic activity. Among zebrafish ribonuclease subfamily the sequence identity ranges from 26% to 73%. The enzymes present small differences in the thermodynamic properties, with ZF-RNase-2, -4 and -5 showing a slight higher stability with respect to ZF-RNase-1 and -3. A more variable catalytic activity is observed among ZF-RNases: a difference of three order of magnitude is found between the most active member, ZF-RNase-4, and the less active ones, ZF-RNase-1 and -2 [156]. Concerning salmon RNases, SS-RNase-1 is similar to zebrafish RNases for what concern the stability, whereas SS-RNase-2 is much more stable. The latter protein has also an elevated catalytic activity [157], comparable to that of ZF-RNase-5, but still lower than that of ZF-RNase-4. A report of catalytic efficiency and melting temperature of ZF-RNases and SS-RNases compared to hAng and RNase A is presented in Table 8.

Table 8 Catalytic efficiency and melting temperatures of ZF-RNases and SS-RNases compared to hAng and RNase A

Enzyme	K_{cat}/K_M ($M^{-1}s^{-1}$)	T_m (°C)
ZF-RNase-1	$5.3\pm0.39 \times 10^2$	54.9
ZF-RNase-2	$1.9\pm0.13 \times 10^2$	63.8
ZF-RNase-3	$2.2\pm0.15 \times 10^3$	55.5
ZF-RNase-4	$0.9\pm0.05 \times 10^5$	64.8
ZF-RNase-5	$1.2\pm0.09 \times 10^4$	64.3
SS-RNase-1	$2.3\pm0.22 \times 10^2$	62.4
SS-RNase-2	$7.0\pm0.69 \times 10^4$	71.1
hAng	$1.7\pm0.12 \times 10^3$	62.6
RNase A	$8.4\pm0.46 \times 10^6$	64.0

ZF-RNase-5

The search for crystallization conditions produced ZF-RNase-5 crystals which diffracted to 1.8 Å resolution.

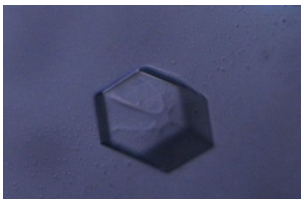

Figure 40 Typical ZF-RNase-5 crystal

The final protein model (residues 1-122), obtained by molecular replacement and refined to R factor=0.178 and R_{free}=0.210, contains 969 protein atoms, 114 water molecules, 1 acetate and 2 sulphate ions. A summary of the refinement statistics is presented in Table 9.

Table 9 Refinement statistics for ZF-RNase-5

Refinement Results	
Resolution limits (Å)	50.00-1.80
Number of reflections used in the refinement (F>2σ(F))	15282
No. of reflections in working set	13762
No. of reflections in test set	1520
R/R_{free}	0.178/0.210
No. of protein atoms	969
Ligand atoms and water molecules	127
RMSD from ideal values	
Bond lengths (Å)	0.031
Bond angles (°)	2.5
Average B-factors (Å²)	
Protein, overall	24.96
Main chains	23.61
Side chains	26.30
Solvent atoms	38.29
Ion atoms	41.36
Ramachandran plot statistics	
Residues in the most favored regions (%)	95.0
Residues in the additionally allowed regions (%)	3.4
Residues in the generously allowed regions (%)	1.7

ZF-RNase-5 is very similar to the other zebrafish RNases, whose crystal structures have been solved [155] (in comparison to ZF-RNase-1 and ZF-RNase-3 the RMSDs of the corresponding Cα atoms are 0.70 Å and 0.91 Å respectively), whereas it is slightly different with respect to hAng [139] (RMSD=1.28 Å).

Figure 41 Cartoon representation of ZF-RNase-5

Remarkably ZF-RNase-5 is the first structure of a zebrafish RNase that presents a sulphate ion in the P1 subsite; its interactions with the surrounding catalytic residues (His17, His117 and Lys45) and structured water molecules are similar to those found in RNase A [226] and other angiogenins [227, 228]. The similarity of the active site architecture with that of RNase A extends to the subsite P2 subsite, where Lys12 and Arg15 are the analogues of Lys7 and Arg10 in the bovine pancreatic enzyme. However, in ZF-RNase-5 the B1 subsite, delimited by Val47, Thr49, and Tyr118, is obstructed by that part of the C-terminal tail which is visible in the electron density map (residues 119-122). In particular, the side chain of Glu120 makes two hydrogen bonds with Thr49 and partly mimics the binding interactions of the pyrimidine base.

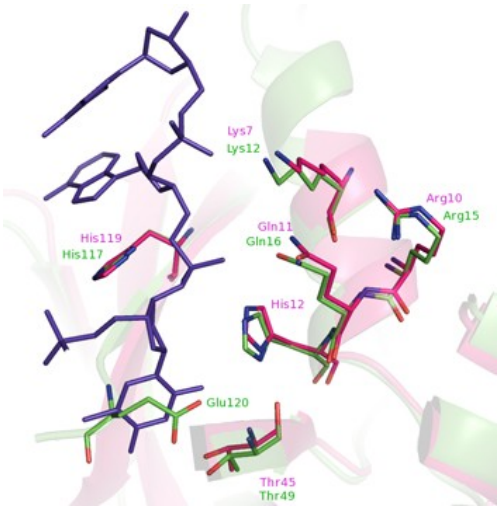

Figure 42 Active site region of ZF-RNase-5. The structure of ZF-RNase-5 (green) is superimposed to that of RNase A (pink) in complex with the dATAA tetranucleotide (violet)

In the present case, the orientation of the Glu side chain is enforced by the involvement of this residue in a pseudo type II' β-bend, stabilized by a C_{10} hydrogen bond between the side chain of the preceding Asp119 and Gly121.

In order to investigate whether the different behavior in the sulphate binding between ZF-RNase-5 and ZF-RNase-1 and -3 was not merely a result of differences in the environmental parameters used in the crystallization trials, structural analysis has been extended to crystals of ZF-RNase-1 grown at acidic pH and in the presence of an elevated concentration of SO_4^{2-} ions.

ZF-RNase-1

After an extensive screening of crystallization conditions, good quality crystals of ZF-RNase-1 were obtained at acid pH (ZF1_pH4.5), from a crystallization solution that contained 2.0 M ammonium sulphate and 0.1 M sodium acetate pH 4.5. These new crystals belong to the space group $P2_12_12_1$, contain two molecules in the independent unit and diffract up to a resolution of 1.38 Å.

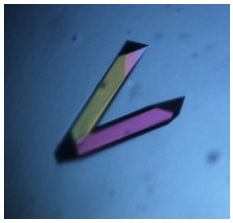

Figure 43 Best crystals of ZF1_pH4.5

The structure was solved by molecular replacement. The final model contains two protein molecules (A and B), very similar to each other (RMSD=0.42 Å when superimposing Cα atoms), 419 water molecules, 7 phosphate groups and 1 acetate one (Figure 44), and it has been refined to R-factor/R_{free} values of 0.162/0.192.

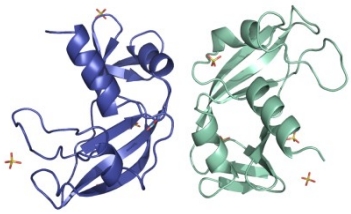

Figure 44 Cartoon representation of the two ZF-RNase-1 molecules in the asymmetric unit. The acetate ion and sulphate ions are shown as stick

In order to investigate the dependence of the structural features of this protein on the pH, the structure at higher pH (7.3) was also solved (ZF1_pH7.3). It is very similar to ZF1_pH4.5: the Cα atomic positions for the two molecules in the asymmetric unit at the two pH do not differ significantly and their RMSD are in the range 0.14-0.47 Å. All non-glycine residues in the refined structures lie within "allowed" regions of the Ramachandran plot and the analysis of the electron density maps indicates that they are extremely well defined for both the protein chains at both pH. Detailed refinement statistics are reported in Table 10. The recombinant protein here analyzed lacks the first three residues in the sequence with respect to the crystal structure of ZF-RNase-1 at basic pH (ZF1_pH8, PDB code 2VQ8 [155]). To simplify the comparison, the numbering of the latter is maintained in this work.

Table 10 Refinement statistics for ZF1_pH4.5 and ZF1_pH7.3

	ZF-RNase-1 PH 4.5	**ZF-RNase-1 PH 7.3**
Refinement Results		
Resolution limits (Å)	30.00-1.38	30.00-1.61
Number of reflections used in the refinement (F>2σ(F))	40180	27488
No. of reflections in working set	36128	24765
No. of reflections in test set	4052	2723
R/R_{free}	0.162/0.192	0.152/0.205
No. of protein atoms	2075	2084
Ligand atoms and water molecules	458	434
RMSD from ideal values		
Bond lengths (Å)	0.027	0.028
Bond angles (°)	2.4	2.4
Average B-factors (Å²)		
Protein, overall	15.63	13.55
Main chains	13.67	11.53
Side chains	17.50	15.43
Solvent atoms	36.06	30.42
Ion atoms	33.06	36.15
Ramachandran plot statistics		
Residues in the most favored regions (%)	96.4	96.4
Residues in the additionally allowed regions (%)	2.8	2.8
Residues in the generously allowed regions (%)	0.8	0.8

The inspection of the electron density maps of ZF1_pH4.5 reveals that chain A binds five sulphate ions and chain B binds three sulphates ions and one acetate ion. Details of the interactions of these ions with the protein are reported in Table 11.

Table 11 Hydrogen bonds of sulphate and acetate ions in the structure of Zf1_pH4.5

Chain A			Chain B		
Sulphate	**Atom name (protein residue and residue number)**	**Distance (Å)**	**Sulphate and acetate**	**Atom name (protein residue and residue number)**	**Distance (Å)**
SO_4^{2-} 3001			SO_4^{2-} 3003		
O1	N (Tyr 107)	3.00	O1	N (Tyr 107)	2.85
	NE (Arg 106)	3.02		NE (Arg 106)	2.98
O2	NH2 (Arg 106)	3.17	O3	NE (Arg 106)	3.03
				NH2 (Arg 106)	2.93
SO_4^{2-} 3002			CH_3COO^-		
O1	NH2 (Arg 25)	2.88	O	NH2 (Arg 25)	2.89
	NH2 (Arg 35)	3.05			
O2	N (Met 21)	3.20			
	N (Asp 20)	2.86			
	N (Pro 19)	3.16			
O4	NH1 (Arg 25)	3.18			
SO_4^{2-} 3000			SO_4^{2-} 3004		
O3	ND1 (His 119)	2.82	O4	NH2 (Arg 93)	2.23
	NE2 (His 121)	2.53			
SO_4^{2-} 3006			SO_4^{2-} 3005		
O2	ND2 (Asn 73)	3.13	O3	NE (Arg 100 AC1)	2.59
	NZ (Lys111 AC1)	2.85	O4	NH2 (Arg 100 AC1)	2.69
			SO_4^{2-} 3007		
			O4	N (Glu 114)	2.89

The binding site of a sulphate (acetate in molecule B) and that of a second sulphate site are well conserved in both chains. The remaining non-equivalent sites are involved in packing contacts. The map of the anionic binding sites is strictly conserved in the structure at neutral pH, with the only exception of the acetate ion that is replaced by an additional sulphate anion. The surprising finding is that, in contrast to numerous crystallographic reports on RNase A [140, 226] and other members of the family [166, 229, 230], and angiogenins [227, 228], and despite the relatively high concentration in the crystallization mixture, none of the sulphate ions is located in the active site, as the anion binds preferentially to other regions of ZF-RNase-1, both at acid and neutral pH.

The putative binding subsite B1 of the pyrimidine base is partially obstructed by the side chain of Glu122, located in the C-terminal segment of the protein. The position of this residue, fixed by two hydrogen bonds to Thr52, is common to all the angiogenins that have a Glu/Gln residue in the equivalent position, and to ZF1_pH8 [155].

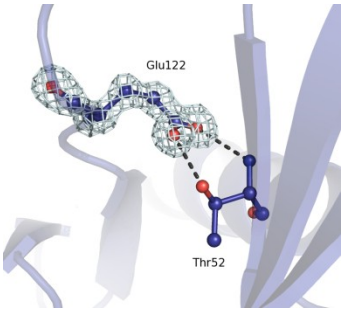

Figure 45 Omit Fo-Fc electron density map (5.0 σ level) of the obstructive Glu122 of ZF1_pH4.5

Interestingly, the C-terminal segment, to which the obstructive glutamic residue belongs, is also the region where the largest differences are observed between the structure at basic pH [155] and those at lower pH. In ZF1_pH4.5 and ZF1_pH7.3, this region is better defined, probably due to packing interactions involving residues 125-127. In particular the acidic/neutral conformation is characterized by hydrophobic interactions between residues Val125, Val127, Pro48, Val49 and Val87 and by the formation of a hydrogen bond between N atom of Gly128 and O atom of Leu88. On the overall, the C-terminus causes an obstruction of B1 that is even greater than that produced in ZF1_pH8.

Comparison among ZF-RNases, hAng and RNase A

An extensive comparison among the known zebrafish RNases crystal structures (ZF-RNase-1, -3, -5) and homology models of the other two proteins (ZF-RNase-2 and -4) has been performed, with the aim to explain the different properties of these enzymes. Interestingly some general trends can be extracted from the structural features of the different members of the subfamily.

Although the geometry of the P1 subsite is strictly conserved among ZF-RNases, it should be stressed that only ZF-RNase-5 binds a sulphate ion in this subsite. This result is enforced by the fact that ZF-RNase-5 crystals were grown from solutions with a low ammonium sulphate concentration (0.2 M), as mentioned before. This feature is probably related to a cluster of positive charges (Arg9, Lys12 and Arg15) in the proximity of the ZF-RNase-5 active site. In particular, Lys12 and Arg15 correspond to Lys7 and Arg10 of RNase A, two residues which line the P2 subsite and are considered to be important for substrate binding [135]. An important role in determining the low catalytic activity of angiogenins in comparison to RNase A has been ascribed to the obstruction of the B1 subsite that occurs in most angiogenins [139, 227, 231]. In these molecules a Glu/Gln residue on the C-terminal tail protrudes in the B1 subsite and forms two hydrogen bonds with the threonine residue involved in the pyrimidine base recognition. This feature is also observed both in ZF-RNase-1 and ZF-RNase-5. In the latter, as for hAng, the Glu residue is involved in a pseudo type II' β-bend, stabilized by a hydrogen bond between the amide nitrogen of Gly121 and the side chain of Asp119. This structural motif, which should further restrict the conformational space available to the glutamic residue, is not found in ZF-RNase-1, where the Asp residue is replaced by a His. In this respect, it should be recalled that the

corresponding aspartate of RNase A (Asp121) forms a sort of "catalytic dyad" with His119 [232]. In the bovine pancreatic enzyme, the hydrogen bond between the side chains of these two residues stabilizes the histidine conformation that is supposed to be active during the transphosphorylation step. In ZF-RNase-5, just like in hAng, no interaction between the corresponding Asp and His is observed, and Asp is engaged in the two enzymes in interactions with Ser118 and Gly121, respectively (Figure 46).

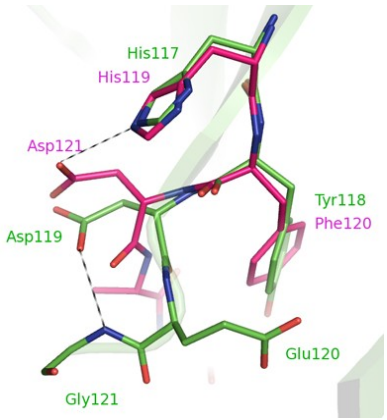

Figure 46 Comparison between the C-terminal tails of ZF-RNase-5 (green) and RNase A (pink)

However, it should be noted that a rotation of the Asp side chain from the *t* to *g*⁻ conformation would bring the carboxylate group in a position to form the hydrogen bond with the catalytic histidine. Moreover, the rotation of the aspartic side chain also results in the disruption of the hydrogen bond with the second residue next along the chain (Gly121 in ZF-RNase-5). This would also destabilize the position of the C-terminal residue and favor the expulsion of Glu120 from the B1 subsite. This structural feature, together with the greater ability of ZF-RNase-5 to bind in the active site a sulphate anion, may account for the enhanced enzymatic activity of the latter with respect to ZF-RNase-1.

It should be stressed that the most relevant differences between ZF-RNases and hAng are localized at the C-terminal tail. In hAng, this region adopts a well defined 3_{10} helix conformation, which is anchored to the protein core by hydrophobic interactions involving residues downstream to the obstructive Gln117 (Ile119 and Phe120). In ZF-RNases the C-terminal segment seems to be much more mobile. Indeed, in ZF1_pH8 it is partially disordered and the segment visible in the electron density map assumes a different conformation when compared to that observed in the structures of ZF1_pH4.5 and ZF1_pH7.3 here reported. In ZF-RNase-5 the C-terminus appears even more disordered: the last visible residue in the electron density map is Gly121. This could be reflected in an increased capability of ZF-RNase-5 to adopt an alternative conformation of the C-terminal tail that allows the binding of substrate. Interestingly, despite ZF-RNase-2 shares a 73% of sequence identity with ZF-RNase-5, its catalytic activity is 100-fold lower. This result could be ascribed to a lower number of basic residues in the active site region of ZF-RNase-2 in comparison to ZF-RNase-5. Moreover, the putative B2 subsite

of the former enzyme appears to be less accessible, as a result of the presence of Arg107 replacing Gly109 in the latter (Figure 47).

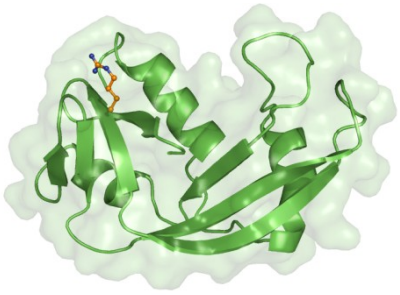

Figure 47 Homology model of ZF-RNase-2 with Arg107 represented as balls and sticks

Concerning ZF-RNase-4, the most active member of the family of zebrafish RNases, its high activity could be due to the concomitant occurrence of an unobstructed B1 subsite, a feature also found in ZF-RNase-3, and of a cluster of positive charges at the N-terminal helix, in particular the Lys9, just adjacent to catalytic His10 (Figure 48). In the neighborhood of the catalytic histidine (His10) this cluster could help substrate recognition.

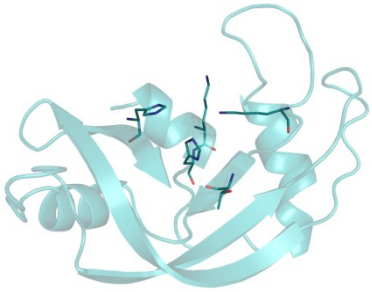

Figure 48 Homology model of ZF-RNase-4 with active site residues (Lys9, His10, Lys38, Thr42 and His106) represented as sticks

From these observations what emerges is that, at least in the zebrafish subfamily, the B1 obstruction is not sufficient to justify the low ribonucleolitic activity of angiogenins; a key role is also played by electrostatic properties of P1 subsite.

Furthermore, the structural differences noted among the different members of the zebrafish RNases subfamily, in particular at the C-terminal region, roughly correlate with their thermodynamic behavior. The presence of the pseudo type II' β-bend at the C-terminal tail in ZF-RNase-2 and ZF-RNase-5, but not in ZF-RNase-3 and ZF-RNase-1, may be the source of their slightly different thermal stability. This feature is not present in the modeled structure of ZF-RNase-4, which also has a shorter N-terminal helix (Figure 48). Its stability, comparable to that of ZF-RNase-2 and ZF-RNase-5, may well be due to the compactness of the protein and in particular to the shorter length of the loop connecting helix II to the β-strand next in the sequence, a feature that the enzyme shares with hAng and RNase A.

Finally, structural analysis highlights slight differences and common features in helix II of the zebrafish

proteins. In ZF-RNase-5 this helix, which encompasses residues 24-36, begins with a turn in a 3_{10} conformation and terminates with a turn in a π helix conformation, with main chain hydrogen bonds Met31-Ile36 and Ser32-Lys35 and Lys35 in the Lα conformation. In the corresponding helix of ZF-RNase-1 (residues 23-37) the network of hydrogen bonds (Ile30-Ile37 and Gly31-Lys36) is spatially conserved, with Lys36 in the Lα conformation, while the intervening extra residues Pro32 and Asn33 bulges out of the helix, forming a type I β-bend (Figure 49).

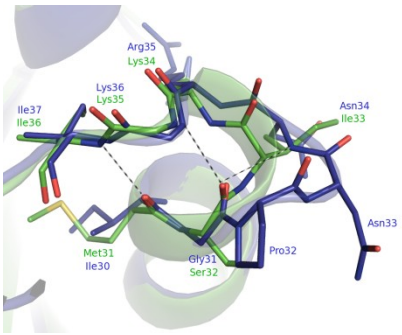

Figure 49 Superimposition of the C-terminal region of helix II of ZF-RNase-5 (green) and ZF-RNase-1 (blue). In the latter, residues Pro32 and Asn33 bulge out of the helix

It is remarkable that despite the presence of the two extra residues, the C-terminal region of helix II of ZF-RNase-1 adopts the same conformation found in many other vertebrate RNases, such as RNase A [233], BS-RNase [166], hAng [139] and ZF-RNase-5. Since ZF-RNase-1 is more ancient than other members of the subfamily it may be surmised that RNase sequence has evolved removing the two extra residues to optimize this peculiar conformation involving a residue in a unfavorable Lα conformation.

SS-RNase-2

An extensive investigation of protein crystal growth conditions produced crystals of SS-RNase-2 diffracting at 1.89 Å resolution.

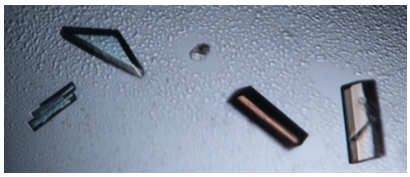

Figure 50 Typical SS-RNase-2 crystals

Crystals contain two molecules in the asymmetric unit, rather similar to each other: their Cα atoms overlay with an RMSD of 0.8 Å. The Ramachandran plot reveals that in the partially refined model 91.8% of residues lies in the most favored regions and 6.6% in the other allowed regions with the remainder 1.6% in outlier regions.

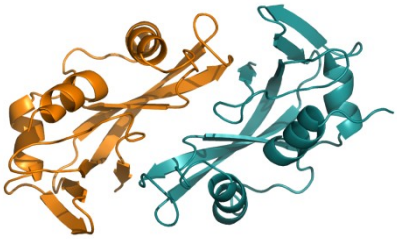

Figure 51 Cartoon representation of the two SS-RNase-2 molecules in the asymmetric unit

Globally SS-RNase-2 structure results very flexible: in both molecules no electron density is observed for the last two C-terminal residues and fragmental density is found corresponding to 113-117 segment, an important region to which a putative catalytic residue, His113, belongs. It should be underlined that thermal factors values for both molecules are quite high, also in usually rigid regions, such as helix II and helix III (Figure 52), or in correspondence to residues such as Gln12 and Lys40 of the active site.

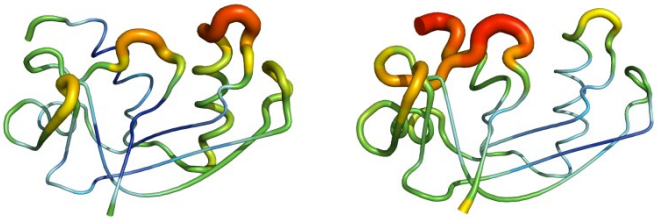

Figure 52 Sausage representation of the two SS-RNase-2 molecules in the asymmetric unit. The color and tube size are based on the B factor, where the color scale is from dark blue to red, with red corresponding to the highest B factor and dark blue to the lowest B factor.

SS-RNase-2 globally adopts the typical RNase fold, with three α-helices and six β-strands and its structure is stabilized by three disulphide bridges, typical of angiogenins. Furthermore SS-RNase-2 shares another common feature with angiogenins, the B1 subsite obstruction by the side chain of a C-terminal residue [139, 155, 231], in this case Asp119. However slight differences are observed with respect to other RNases: SS-RNase-2 has a longer β2 and a shorter β5 strands and substantial dissimilarities are found in helix II. In SS-RNase-2 this region is longer and more regular than in the other proteins: it starts with a turn in 3_{10} conformation, a conformational feature conserved in ZF-RNase-1 and ZF-RNase-5 structures, but continues with an α-helix (residues 23-33), with the canonical i-i+4 hydrogen bonds pattern. As previously discussed, this helix in the other RNases encompasses only six residues, whereas the following four describe a π-helix, involving a residue with a unfavorable Lα conformation. On the contrary in SS-RNase-2 all the residues adopt allowed conformations.

The most relevant and interesting structural features of SS-RNase-2 are localized at the C-terminal tail. A surprising finding is that segment 112-118 does not adopt a β-strand conformation, as happens in all the other RNases, but it is unstructured and very flexible, whereas the 118-124 segment, which adopts the expected β-strand conformation, is anchored to the protein main body thanks to interactions

between Asp119 and Thr44 (a residue analogue to RNase A Thr45). Furthermore this peptide fragment is also involved in intermolecular contacts between the two protein molecules in the asymmetric unit. The unusual conformation of the 112-118 fragment results in a collapsed P1 subsite: the side chain of His113 occupies the substrate binding site. However it should be stressed that electron density maps corresponding to this residue are not very clear. Activity tests have shown that this protein has a consistent ribonucleolitic activity [157], which seems in apparent contrast with the structural results. However, the high flexibility of SS-RNase-2 could facilitate a conformational variation of all the C-terminal segment (residues 112-126) toward a well defined structure, with unobstructed subsites, which favours the substrates recognition.

It should be also recalled that atlantic salmon lives at 10°C and therefore it is a cold adapted organism. Psycrophilic proteins have been evolved in order to exert an efficient catalytic activity also at low temperature. It has been demonstrated that cold adapted enzyme could achieve this result thanks to a higher flexibility of the whole structure [234] or of active site residues [235]. In this context, SS-RNase-2 could represent an extreme case: its active site is so mobile that it is only partially structured.

SS-RNase-2: a new tile in the RNase mosaic

Sequence alignment of SS-RNase-2, zebrafish RNases, human angiogenin and RNase A indicates a variability of the C-terminal tail, as is evident in Figure 53.

```
Ss-RNase2       IACDGK--FPVHYDGDVDIGITDGK 126
zf-RNase1       LKCEEG--WPVHYHEDEVNVG---- 125
zf-RNase3       VACEGE--WPTHYEKGVI------- 123
zf-RNase5       LGCDKG--WPVHYDEGIIDVNRSG- 129
hAng            VACENG--LPVHLDQSIFRRP---- 123
RNaseA          VACEGNPYVPVHFDASV-------- 124
```

Figure 53 Sequence alignment among SS-RNase-2, zebrafish RNases of known structure, hAng and RNase A. Residues belonging to the C-terminal β-strand are coloured in red

However sequence alignment does not correspond to the structural one. Structural analysis allows to cluster these enzymes into two groups: RNases with an obstructive C-terminal segment (hAng, ZF-RNase-1 and -5), which have a shorter β-strand, and RNases with a free B1 subsite (RNase A and ZF-RNase-3), which have a longer β-strand. Remarkably SS-RNase-2 does not correspond to either of these two groups: it has an obstructed B1 subsite, but its β-strand better aligns with that of RNases with a free B1 subsite. Thus a different alignment, based on structural features including the obstructive residue, could be proposed (Figure 54)

```
Ss-RNase2       IACDGK--FPVHYDGDVDIGITDGK----126
zf-RNase1       LKCEEG--WPVHYH---EDEVNVG---- 125
zf-RNase3       VACEGE--WPTHYE---KGVI------- 123
zf-RNase5       LGCDKG--WPVHYD---EGIIDVNRSG- 129
hAng            VACENG--LPVHLD---QSIFRRP---- 123
RNaseA          VACEGNPYVPVHFD---ASV-------- 124
```

Figure 54 Structural alignment among SS-RNase-2, zebrafish RNases of known structure, hAng and RNase A. Residues belonging to the C-terminal β-strand are coloured in red

From this it emerges that in the salmon protein a three residue insertion is present downstream the putative catalytic histidines (His113), which could determine the greater mobility of the C-terminal tail observed in the crystal structure. Furthermore, it should be noted that, although in SS-RNase-2 the obstructive residue (Asp119) has a shorter side chain with respect to that of other angiogenins (Glu or Gln), it is still able to interact with Thr44, thanks to the peculiar structure of the C-terminal segment, discussed before.

In the case of SS-RNases we thought better do not apply the homology modelling approach used to obtain the structures of ZF-RNase-2 and -4. The SS-RNase-2 model is so unusual and different from the other RNases that we preferred not use it for homology modelling of SS-RNase-1. In the future we will further search crystallization conditions for the latter protein, moreover we will try to obtained new crystal forms of SS-RNase-2, in order to confirm its peculiar structure.

Non covalent swapped dimeric form of RNase A mutants

Many efforts have been devoted in the last decades to convert monomeric non-cytotoxic RNases, which though present a low immunogenicity, such as human and bovine pancreatic RNases, into non-covalent swapped dimeric forms, endowed with antitumor activity. The availability of a well known model, BS-RNase, has allowed the production, through a homology mutation approach, of several dimeric RNase A variants. In particular I focused my attention on three RNase A mutants produced by prof. Picone and her coworkers: LCC, with substitution Q28→L, K31→C, S32→C, PLCC, with the additional substitution A19→P, and GNPSCC, which embodies the whole sequence of the seminal enzyme hinge peptide plus the two cysteines (S16→G, T17→N, A19→P, A20→S, K31→C, S32→C). All these proteins do exhibit a cytotoxic activity, though at a much lower level than BS-RNase [191, 204, 206, 236]. Structural determination by X-ray crystallography of their NCD forms provides a powerful tool to understand the molecular basis of the different biological activity of these proteins with respect to the model enzyme, BS-RNase, and to design new mutants with improved biological actions. And indeed, thanks to the results achieved, a new RNase A mutant has been proposed to prof. Picone during my PhD, namely GPLCC (PLCC with the additional substitution S16→G), which has been found to have an antitumor activity similar to that of the seminal enzyme. Thus NCD-GPLCC has been crystallized as well, and its structure determined from two different crystal forms.

NCD-PLCC, NCD-LCC, NCD-GNPSCC

The non covalent dimeric form of the three RNase A mutants, LCC, PLCC and GNPSCC, were crystallized and diffraction data were collected in the range 1.60-1.99 Å. The structure of NCD-LCC was determined in the presence of the substrate deoxycytidylyl (3',5')-2' -deoxyguanosine (dCpdG), and that of NCD-PLCC was obtained both in the ligand-free (lf) and in the dCpdG legated form (lb) (structural results obtained during the development of my graduated thesis). Matthews coefficients calculations show that in all the cases the asymmetric unit contains one dimer molecule and the inspection of the residual electron density maps corresponding to hinge peptide regions confirmed the somewhat expected result that the three dimers do swap their N-terminal α-helices (Figure 55). However, it should be underlined that in NCD-PLCC and in NCD-LCC, one of the two hinge peptides is disordered, electron density corresponding to residues 20-21 and 19-21 of the second chain of NCD-LCC and NCD-PLCC respectively is missing.

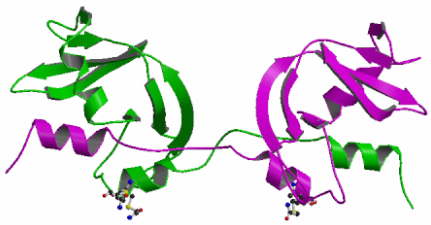

Figure 55 Cartoon representation of the overall structure of lf-NCD-PLCC. Alchilated cystein residues are also shown in ball and stick representation

The basic structural units, formed by residues 23-124 of one chain and the swapped residues 1-15 of the partner chain, are very similar in the three dimers (RMSDs after superimposition of Cα atoms are in the range 0.5-0.6 Á), and, for each dimer, are related by a local twofold symmetry, as observed for most of the NCD swapped dimers of RNases. In the refined structures all non-glycine residues, except Ser23 in both forms of NCD-PLCC and in NCD-GNSPCC, lie within allowed regions of the Ramachandran plot. Detailed refinement statistics are reported in Table 12.

Table 12 Refinement statistics for the NCD forms of the three RNase A mutants, LCC, PLCC and GNSPCC

	Lb-NCD-LCC	**NCD-GNSPCC**	**Lf-NCD-PLCC**	**Lb-NCD-PLCC (previous results)**
Refinement Results				
Resolution limits (Å)	30.0-1.94	30.0-2.00	40.00-1.60	40.00-1.90
Number of reflections used in the refinement (F>2σ(F))	18262	14995	29314	18705
No. of reflections in working set	16459	13467	26371	16879
No. of reflections in test set	1803	1528	2943	1826
R/R_{free}	0.181/0.227	0.188/0.246	0.172/0.217	0.175/0.223
No. of protein atoms	1915	1915	1931	1893
Ligand atoms	20	0	23	50
Water molecules	233	123	397	210
RMSD from ideal values				
Bond lengths (Å)	0.021	0.021	0.029	0.028
Bond angles (°)	2.04	2.01	2.34	2.29
Average B-factors (Å²)				
Protein, overall	22.24	25.83	27.80	22.26
Main chains	21.01	25.00	26.28	21.06
Side chains	23.52	26.72	29.37	23.52
Solvent atoms	31.03	30.88	42.92	32.32
Ligand atoms	40.04	-	30.93	47.44
Ramachandran plot statistics				
Residues in the most favored regions (%)	97.0	96.6	97.0	96.5
Residues in the additionally allowed regions (%)	3.0	3.0	2.6	3.5
Residues in the generously allowed regions (%)	-	0.4	0.4	-

Diversified features are observed at the active site region of the dimers, in terms of ligands and their location. It should be stressed that the active sites are composite, as they are lined by residues of different chains. At one active site of lf- and lb- NCD-PLCC, a well-defined electron density clearly reproduces a trehalose molecule (Figure 56), a compound used for the cryo-stabilization of the crystals. The sugar is hydrogen bonded to Thr45, Lys66 and Asp121 and two water molecules, which, in turn, are linked to Asp83 and Ser123. The binding of trehalose does not affect significantly the active site.

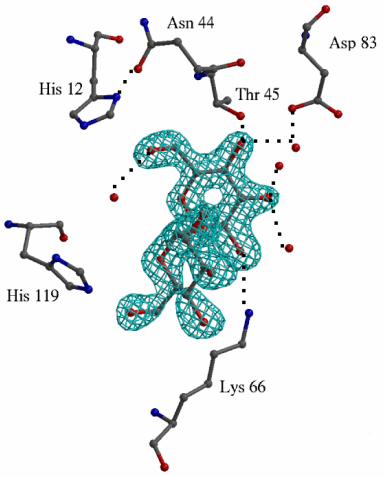

Figure 56 Omit Fo-Fc electron density map (3.8σ level) at the active site region of the first structural unit of NCD-PLCC

In the ligand-free form, water molecules are located in the active site (Figure 57 A) similarly to other RNases [194, 226, 230, 233], whereas in lb-NCD-PLCC, dCpdG is positioned in the retrobinding mode, similar to that observed for guanine containing dinucleotide complexed to RNase A [140, 237, 238] and BS-RNase [194]. The purine base is located in B1 subsite and is hydrogen bonded to Thr45, the guanine ribose moiety is exposed to the solvent and does not interact with the protein and the phosphate group is only partially visible (Figure 57 B).

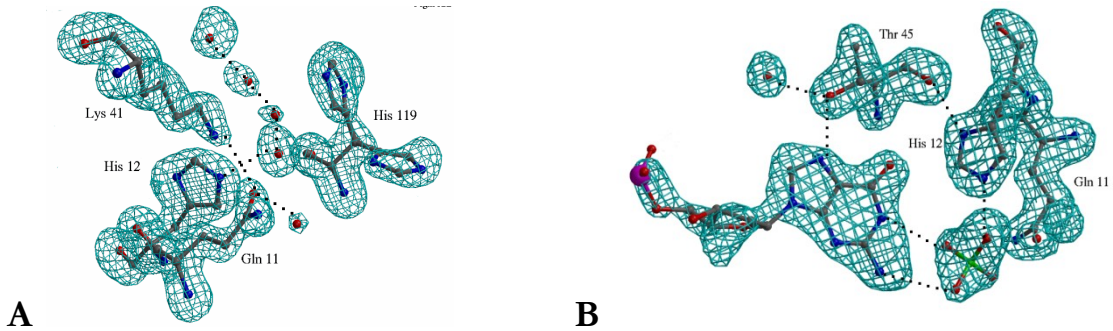

Figure 57 A) Omit Fo-Fc electron density map (3.0 σ level) at the active site region of the second structural unit of lf-NCD-PLCC. Catalytic residue His119 adopts two different conformations. B) Omit Fo-Fc electron density map (3.5 σ level) at the active site region of the second structural unit of lb-NCD-PLCC.

Differences in the active site organization between the two structural units are also apparent for lb-NCD-LCC. Water molecules are located in one site, whereas at the second site the dinucleotide binds in the productive mode (Figure 58), with the pirimidine base in B1 and the phosphate in P1. The ribose and the downstream guanine base are not visible, as found also in the productive binding of UpG to RNase A [140].

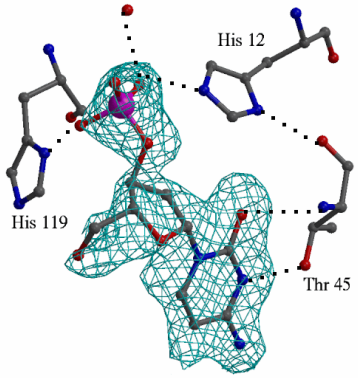

Figure 58 Omit Fo-Fc electron density map (3.5 σ level) at the active site region of the second structural unit of lb-NCD-LCC

In contrast with the first two variants, the active sites of lf-NCD-GNPSCC present a very similar organization: well-defined water molecules occupy the B1 and P1 subsites at positions similar to those found for the ligand-free structures of other pancreatic-like RNases [140, 226, 230, 233].

Dimers structures have been compared to each other and with the model protein NCD-BS [190]. In particular the subunit assembly between two dimers can be monitored by the value of the angle ψ needed to superimpose the second structural unit after the superposition of the first one. NCD-LCC and NCD-GNPSCC are very similar to each other (ψ=4°), whereas PLCC is slightly different (ψ≈30° with respect to the other two dimers). Remarkably a particularly high ψ value is obtained when the three dimeric mutants are compared with the closed quaternary structure of NCD-BS (ψ≈105°). It should be stressed that in NCD-BS interface between subunits (Open interface) is formed by the two hinge peptides (residues 16-22) and by two helices II (residues 23-34), which comprise the S-carboxyamidomethyl-cysteines and the two Leu28 [190]; on the contrary in the three structures here reported, the helices 23-34 are far away from each other and the corresponding residues 28 are fully exposed to the solvent, despite the fact that this position is occupied by a Leu in LCC and PLCC (Figure 59).

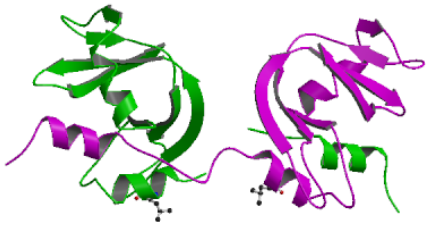

Figure 59 Cartoon representation of the overall structure of lf-NCD-PLCC. Leu28 of the two subunits are also shown in ball and stick representation

This structural results provide an explanation for cytotoxic assays performed on these proteins. The insertion of the two Cys residues, Pro19 and Leu28 is not sufficient to induce the RNase A dimeric variants to adopt the quaternary structure of NCD-BS. In the open quaternary structure of these mutants one subunit can be easily modeled into RI horseshoe cavity without steric interference

produced by the partner chain (Figure 60), as previously shown for other open dimers [205] as well as for the more compact ND-RNase A molecule [190].

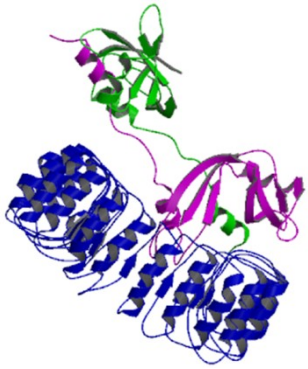

Figure 60 Cartoon representation of the putative complex between NCD-GNSPCC (green and magenta) and RI (blue)

The finding that NCD-PLCC adopts a quaternary structure completely different with respect to NCD-BS has been somewhat surprising, since it was predicted that the insertion of Pro19 and Leu28 in the RNase A sequence would strongly push the quaternary organization toward that observed for NCD-BS. However, interesting novel features emerges from a comparison of the hinge peptide structures of various dimers. For the present compounds, the electron density in this region reveals significant differences. In lf-NCD-GNPSCC both hinge peptides are well ordered and adopt two different conformations starting from Ser20. In one chain, residues 21-24 and 22-25 form two nested type IV β-turns, whereas in the second chain, a type I β-turn is formed by residues 18-21. For the structures of NCD-PLCC and NCD-LCC one of the two hinge peptides is disordered, whereas the second one can be fully traced in a well defined electron density map. The ordered conformation in the two structures is similar (RMSD=0.37 Å), irrespective of the presence of a proline residue in PLCC, and is characterized by a type I β-turn encompassing residues 18-21, followed by a 3_{10} helix (residues 22-24).
When comparing hinge peptides with those of other swapped dimers of RNases remarkable results are obtained. In Figure 61 the hinge peptides are shown together with those of various swapped dimers of RNases: the overlay of the peptides was obtained by imposing for all dimers a common orientation of the partner chain.

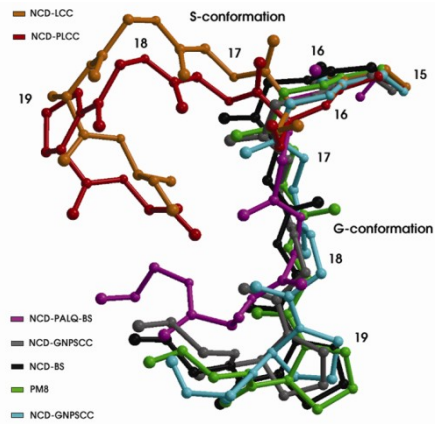

Figure 61 Superimposition of the hinge peptide regions of lf-NCD-PLCC, lb-NCD-LCC, lf-NCD-GNPSCC, NCD-BS (PDB code 1TQ9 [190]), PM8 (PDB code 1H8X [239]) and NCD-PALQ (PDB code 3BCP [205]). For sake of clarity, the side chains of some residues have been omitted

As it clearly emerges from Figure 61, the observed conformations cluster in two broad families, which part at the level of residue 16, according to whether this residue is Gly (G-conformation) or Ser (S-conformation). In the first case (G-conformation), the stretch 16-19 is tightly bound on the surface of the partner chain and, in particular, has the pirrolydine ring of Pro19 in the pocket formed by the side chains of Tyr25 and Gln101 of the other subunit (Figure 62).

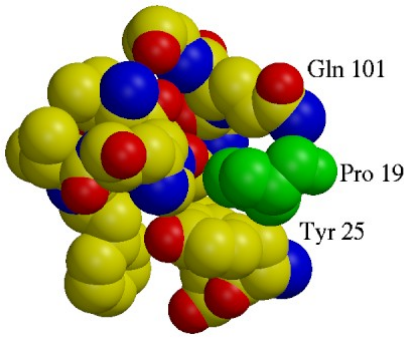

Figure 62 Van der Waals representation of the *proline pocket*: Pro19 side chain (green) is inserted into a cavity of the partner subunit lined by Tyr25 and Gln101

Interestingly, in all cases, the conformation of Gly16 falls in a region of the Ramachandran map forbidden for a non-Gly residue. When Pro is replaced by an Ala (as in NCD-PALQ [205]), this residue is shifted away from the *proline pocket*. On the other hand, when Gly is replaced by Ser, there is a dramatic modification of the hinge structure at the level of Ser16, which favours more opened and flexible conformations of the whole peptide, irrespective of the presence of proline in position 19 (S-conformation in Figure 61). In one subunit of NCD-LCC and NCD-PLCC, residues 20-21 and 19-21, respectively, cannot be traced in the electron density maps. However, residues 16-18 adopt a structure that is very close to the S-conformation.

Thus the analysis of the conformational preferences of the hinge peptide clearly suggests a strong synergic effect of the residues in the 16 and 19 position of the hinge peptide sequence. In order the

proline side chain be located in the *proline pocket* built on the surface of the partner chain, the residue in position 16 must assume a conformation not allowed for a non-glycine residue. With Pro19 firmly bound in the pocket, residues 16-19 stick on the protein surface and the flexibility of the hinge peptide is limited to the conformational variability of the segment 20-22. This significantly reduces the quaternary space accessible to the dimer and, in the presence of a Leu residue in position 28, should favor the acquisition of a compact structure similar to that of NCD-BS.

NCD-GPLCC

Moving on from structural results obtained for the three RNase A mutants, the new mutant GPLCC-RNase A (with substitution S16G, A19P, Q28L, R31C and S32C) has been produced and tested for cytotoxic activity by prof. Picone group. The obtained results are particularly encouraging: GPLCC exerts a cytotoxic activity on SVT2 cells comparable to that of seminal enzyme (Figure 63).

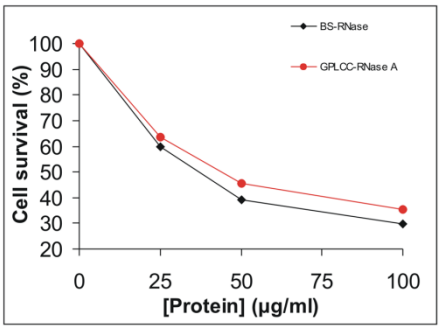

Figure 63 Cytotoxicity of BS-RNase and GPLCC on SVT2 cells

In order to verify whether the biological properties of the new mutant depend on a quaternary structure of the NCD form similar to NCD-BS, the crystallization of NCD-GPLCC was attempted.

Good quality crystals were obtained from a solution containing 30% PEG 20000 and 0.2 M ammonium sulphate at protein concentration 10 mg/ml and in presence of 7.4 mM dCpdG (lb-NCD-GPLCC) or 7.4 mM dATAA (lf-NCD-GPLCC). Diffraction data were collected for both crystals at 2.61 Á and 1.60 Á. Matthews coefficients calculations show that in both the cases asymmetric unit contains one dimer molecule, a clear solution was obtained by molecular replacement performed with AMoRe [118] and finally the inspection of the residual electron density maps corresponding to hinge peptide regions confirmed the expected result that the dimer swaps its N-terminal α-helices in both the crystal forms, although one of the two hinge peptides is disordered in the structure at 2.61 Å resolution. Furthermore, electron density analysis show that both active sites of the structure at 1.60 Å resolution bind a sulphate ion and a dCpdG molecule in retrobinding mode [140] (lb-NCD-GPLCC), but no electron density corresponding to the tetranucleotide dATAA is observed in the active sites of the other structure (lf-NCD-GPLCC).

Lb-NCD-GPLCC structure has been completely refined to R/R_{free} values of 0.179/0.228 respectively, whereas refinement of lf-NCD-GPLCC is in progress. Refinement statistics for the first model are reported in Table 13.

.

Table 13 Refinement statistics for lb-NCD-GPLCC

Refinement Results	
Resolution limits (Å)	50.00-1.60
Number of reflections used in the refinement (F>2σ(F))	27792
No. of reflections in working set	25014
No. of reflections in test set	2778
R/R_{free}	0.179/0.228
No. of protein atoms	1949
Ligand atoms and water molecules	283
RMSD from ideal values	
Bond lengths (Å)	0.026
Bond angles (°)	2.14
Average B-factors (Å²)	
Protein, overall	24.50
Main chains	23.44
Side chains	25.60
Solvent atoms	39.19
Ligand atoms	31.43
Ramachandran plot statistics	
Residues in the most favored regions (%)	94.9
Residues in the additionally allowed regions (%)	4.7
Residues in the generously allowed regions (%)	0.4

The surprising result is that quaternary structures of the dimer in the two crystal forms are completely different to each other (the value of the angle ψ needed to superimpose the second structural unit of the two dimers, after the superposition of the first one, is 88°).

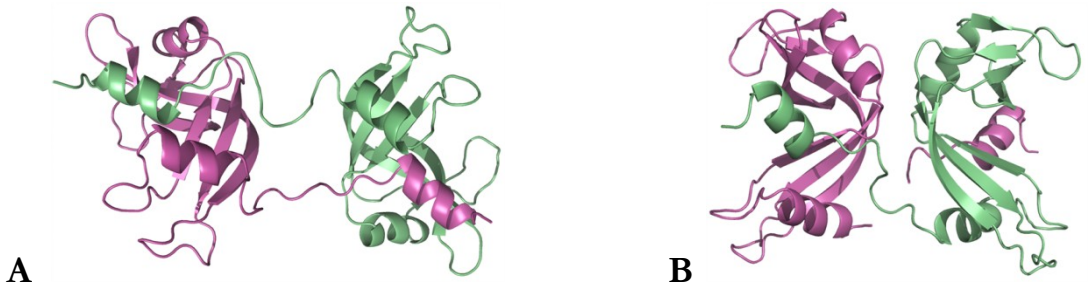

Figure 64 Cartoon representation of lb-NCD-GPLCC (A) and lf-NCD-GPLCC (B). In each structure the two subunits are coloured green and magenta, respectively

Lb-NCD-GPLCC adopts an open quaternary structure, whose subunits are related by a 165° rotation around a pseudo symmetry rotation axis (Figure 64 A), very similar to the NCD forms of the other mutants analyzed before (the value of the angle ψ needed to superimpose the second structural unit of lb-NCD-GPLCC and NCD-LCC, NCD-PLCC and NDC-GNPSCC, after the superposition of the first one, is 3° in the three cases). On the contrary lf-NCD-GPLCC is more compact (Figure 64 B) (although it lacks one of the hinge peptides its buried surface area is 1780 $Å^2$ versus the 1460 $Å^2$ of lb-NCD-GPLCC). It is characterized by a non-crystallographic twofold symmetry and a tight Open interface, mainly composed of residue 48-50 and 78-80 from the two subunits. However its assembly is different with respect to other closed non covalent dimers, especially with respect to NCD-BS [190] (ψ=144°) and ND-RNase A (PDB code 1A2W [240]) (ψ=156°).

A deeper analysis of lf-NCD-GPLCC was performed to rationalize its open structure. The two hinge peptides are clearly defined in electron density maps and they adopt two different conformations. The first one is characterized by two nested β-turns, a type IV encompassing residues 21-24 and a type I encompassing residues 22-25. In the second one residues 18-21 are involved in a type IV β-turn and residues 21-24 in a type I β-turn. A comparison with the two hinge peptide conformations (G- or S-conformation) described above shows that in NCD-GPLCC the first hinge peptide adopts a G-conformation and the other one an S-conformation. In particular in lb-NCD-GPLCC the first one has a structure identical to that of NCD-BS for residues 15-19 (Figure 65 A) and, as expected, the Pro19 occupies the *proline pocket* on the other subunit. However marked differences are observed for residue 20-23 with respect to NCD-BS. This diversity could be ascribed to the presence of a sulphate ion in the present structure, which interacts with N atoms of Ser22 and Asn24 (Figure 65 B), and, directly or by water mediated contacts, with side chains of residue Ser21, Ser22, Ser23 and Asn24.

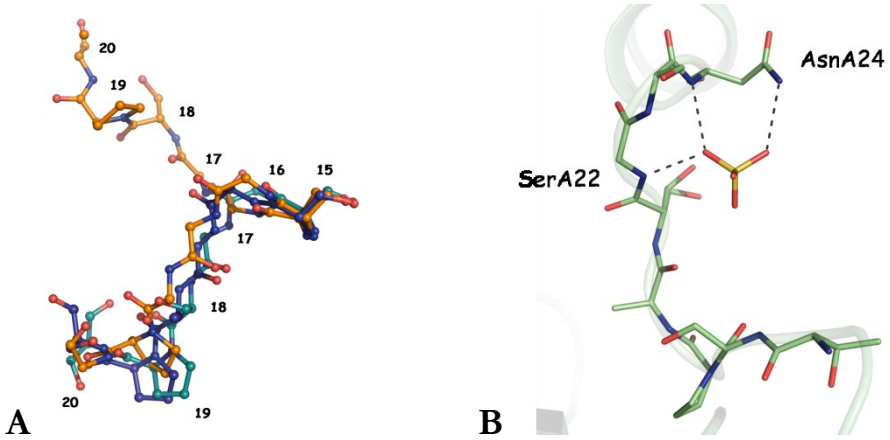

Figure 65 A) Superimposition of the hinge peptide regions of lb-NCD-GPLCC (orange), lf-NCD-GPLCC (cyan) and NCD-BS (blue). For sake of clarity, the side chains of some residues have been omitted. B) Interactions between a sulphate ion and one of the hinge peptides of lb-NCD-GPLCC

Thus it can be suggested that sulphate binding strongly affects hinge peptide conformation. As mentioned before the second hinge peptide adopts the S-conformation, with a solvent exposed Pro19, but its structure is different with respect to that observed in NCD-PLCC and NCD-LCC. Hinge peptides affect each other and this could determine the specific conformation found in NCD-GPLCC: the sulphate binding prevents the first hinge peptide to adopt a BS-like structure, in turn it pushes the other hinge peptide towards a different conformation and altogether this results in a open quaternary structure. Thus the observed dimer assembly could be influenced by crystallization conditions. Concerning lf-NCD-GPLCC, at a first glance (it should be recalled that we are talking of a preliminary model and not of a fully refined structure) the only visible hinge peptide adopts the expected G-conformation.

Both the NCD-GPLCC structures, even the most compact one, are still compatible with RI binding, in apparent contrast with its observed cytotoxic activity. However the high variability of quaternary structure found for this dimer should be stressed. This feature has been amply discussed for swapped dimers: they do not need to conserve interface interactions in order to preserve a dimeric structure, thanks to the presence of the swapping phenomenon. This results in a marked variability of the relative arrangement of the subunits within the dimers. Nevertheless, the case of NCD-GPLCC seems to be peculiar. The two quaternary assemblies are obtained from identical crystallization conditions, the only observed difference being the presence of the substrate analogue. This does not happen for other RNase dimeric mutants, such as PLCC, whose NCD structure is the same irrespective to substrate binding.

Immuno-RNase: ErbB2 epitope identification using a computational approach

One of the most promising strategy in anticancer therapy is that based on immuno-conjugates in which an RNase is fused to an antibody or scFv specific for tumor cell. In this context prof. D'Alessio, prof. De Lorenzo and their co-workers have produced an immuno-RNase [212], composed of human pancreatic RNase and Erbicin, a human scFv direct against ErbB2 [223]. Both Erbicin and immuno-RNase are selectively cytotoxic for ErbB2-positive cancer cells *in vitro* and *in vivo* and remarkably do not show cardiotoxicity, the most negative side effect of Trastuzumab, the only humanized anti-ErbB2 antibody used in the therapy of breast carcinoma. Furthermore they act also on Trastuzumab resistant tumors. These attractive properties of Erbicin and immuno-RNase are likely due to a different ErbB2 epitope recognized by Erbicin with respect to Trastuzumab. The structural characterization of the antibody-antigen complexes is required to obtain several insights into the molecular mechanism of action. In the present case, it has not been possible to perform crystallization experiments for these complexes, because the procedure of expression and purification of ECD has not yet been optimized by the group of prof. De Lorenzo. On the contrary, a careful search for crystallization conditions of Erbicin and immuno-RNase has been carried out without, unfortunately, a positive outcome. It seems clear that the modular structure of these immuno-agents induces a wide conformational variability that could interfere with the crystallization process. Thus, in order to identify the epitope and to characterize the immuno-RNase-ECD interactions, a computational approach has been employed. It should be underlined that HP-RNase and Erbicin are linked together through a 11 residues peptide, whose sequence is AAASGGPEGGS. The linker is long and flexible enough to assure that the two molecules can be considered separately. Thus, for simplicity, the analysis has been carried out on the antibody moiety of the fusion protein.

To perform the calculation a three-dimensional model of Erbicin was built, using the canonical structures method for the hypervariable loops [241, 242] and standard homology modelling techniques for the framework regions. The model has a Prosa Z-score of -6.53, a value in the range of scores typically found in proteins of similar sequence length, and shows the 96.4% of residues in most favoured or in allowed regions of the Ramachandran map. The modelled protein

(Figure 66) is characterized by a predominantly canonical structure with the peculiar feature of a short H3 CDR loop. Complementarity determining regions (CDR) are the antibody regions where the molecule complements an antigen's conformation, they determine the molecule specificity. Therefore, this Erbicin structural feature is particularly important to determine its properties with respect to other antibodies, as will be discussed below.

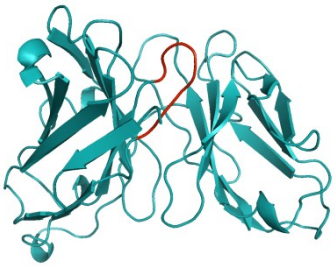

Figure 66 Cartoon representation of Erbicin. The H3 CDR loop is coloured red

The model was used to identify the structural origins of the different binding properties between Erbicin and the two immuno-agents of known structures, Trastuzumab and Pertuzumab. The crystal structures of ECD in complex with the two humanized antibodies are known (PDB codes 1N8Z and 1S78 for complexes with Trastuzumab [218] and Pertuzumab [243], respectively), thus putative complexes between Erbicin and ECD were obtained aligning Erbicin to Trastuzumab (RMSD after Cα superimposition is 0.94 Å) in the first complex (trastuzumab-like) and to Pertuzumab in the second complex (pertuzumab-like) (RMSD after Cα superimposition is 1.00 Å).

Interactions between Trastuzumab and ECD are localized in ECD domain IV and mainly involve the antibody H3 CDR loop [218]. A similar binding mode is not allowed to Erbicin because, as mentioned before, it has an unusually short H3 loop (6 *vs* 11 residues) (Figure 67).

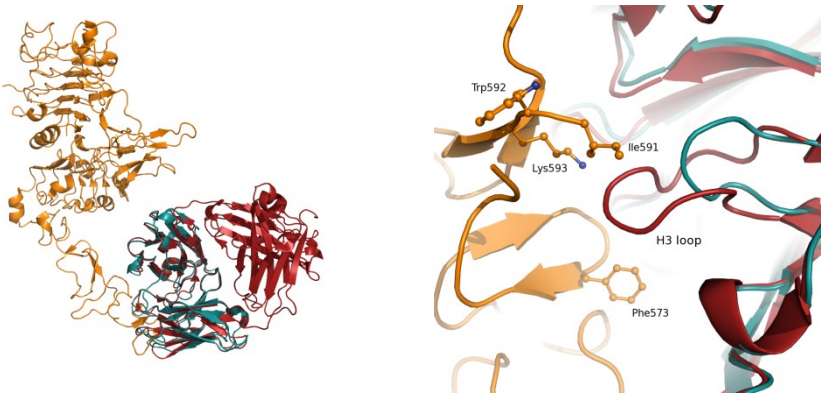

Figure 67 On the left, trastuzumab-like putative complex of Erbicin (cyan) with ECD (orange). The structure of Trastuzumab (red) has been also reported for comparison. On the right a zoom on the interface region

For what concern Pertuzumab, significant differences are found in primary structures of the two molecules. A few amino acid substitution (Asp31 with Ser31, Asn52 with Tyr52 and Asn54 with Gly54) in Erbicin prevents a pertuzumab-like binding. Asp31 of Pertuzumab forms a strong hydrogen bond with the side chain of Ser288 of ECD and participates to hydrophobic interactions with the C atoms of Val286 and Thr290 of ECD. Furthermore, ND2 atom of Asn52 forms a hydrogen bond with the backbone oxygen of Val286 of ECD, whereas OD1 and ND2 atoms of Asn54 interact with ECD backbone atoms belonging to Cys246 and Val286 and with side chain atoms of Thr268 [243]. In the putative pertuzumab-like complex these interactions are missing due to the presence of a serine instead

of Asp31 and a glicine instead of Asn54. An even more dramatic effect can be related to Erbicin Tyr52: the phenol ring is spatially too close to the backbone atoms of ECD Val286, thus the binding could be prevented due to an hindrance effect (Figure 68).

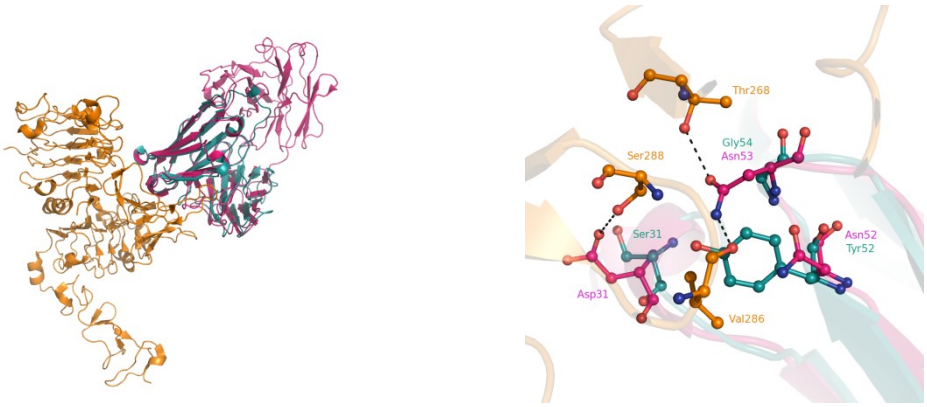

Figure 68 On the left, pertuzumab-like putative complex of Erbicin (cyan) with ECD (orange). The structure of Pertuzumab (pink) has been also reported for comparison. On the right a zoom on the interface region

In order to reveal the region of ECD involved in the interaction with Erbicin, a computational docking was performed by FTDock program. First of all the program performance was checked verifying if it was able to predict the ECD-Trastuzumab and ECD-Pertuzumab complexes. Then the calculations were carried on using the model of Erbicin, here reported. The solutions were evaluated with respect to several experimental and theoretical criterions. In particular, the model should have a high surface complementarity at the interface and should bury a surface area > 600 $Å^2$ per molecule, it should have low energy and should be reproduced when docking calculations were repeated using different programs and/or input parameters. Upon clustering the 30 solutions with the lowest energy values, three potential models were identified and visually examined (Figure 69).

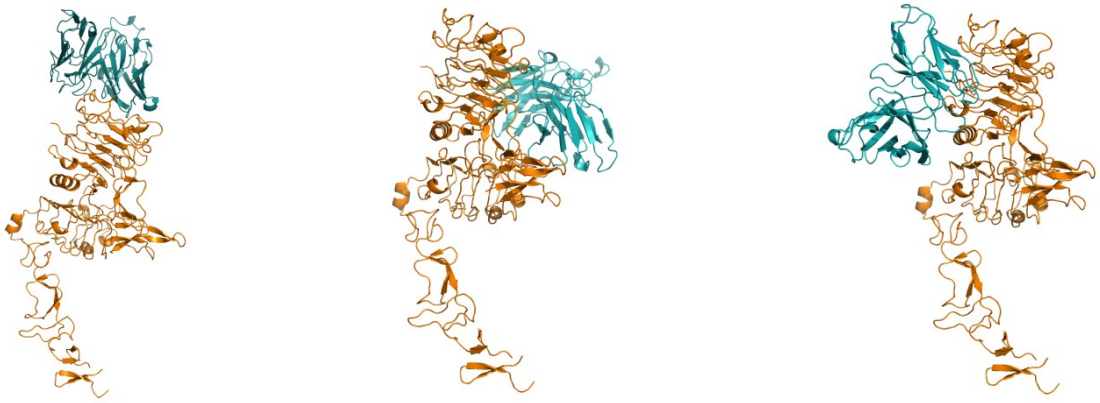

Figure 69 Cartoon representation of the potential Erbicin-ECD models (model 1-3, from the left to the right). Erbicin is coloured cyan, ECD orange.

Among the three solutions, model 1 fulfils the previous criteria. Indeed a total of 23 residues forms the interface, which is characterized by a good surface complementarity (Sc 0.55) and by an extensive buried surface area (about 750 $Å^2$) over a long groove. In this model Erbicin binds ECD in the cleft

between the light and the heavy chain variable domains. As expected CDR loops are mainly involved in antigen recognition. The antibody peptide regions participating in direct contacts with ECD include the CDR H3 loop (Arg100, Asp101, Ser 102), the CDR H1 loop (Thr30, Ser31, Tyr32) and residues Tyr181, Ser182, Gly225, Ser226, Pro227. On the other side, the ErbB2 residues at the interface primarily involve a cysteine rich fragment, corresponding to region 162-190. In particular, Erbicin tightly binds the ECD fragment $_{167}$SRACHPCSPMCKGS$_{180}$, in which Cys173 forms a disulfide bridge with Cys182 and Cys177 forms a disulfide bridge with Cys190. These findings well agree with the experimental evidences obtained by prof. D'Alessio and prof. De Lorenzo, which have shown that the epitope involves ECD residues 122-195, on domain I of ECD. A deep analysis of the model shows that a central role in the ECD-Erbicin interaction is played by His171 of ECD. This residue fills a hydrophobic cavity on Erbicin surface, lined by the side chains of tyrosines 223, 163, 164 and 181 and by Ser182 (Figure 70), where it may be involved in stacking interactions with one of the aromatic residues and in a hydrogen bond with the OG of the serine.

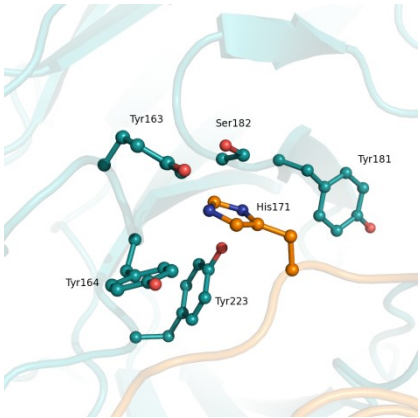

Figure 70 A zoom on the Erbicin-ECD interface

The results obtained by computational docking were validated by fluorescence studies. A peptide with the aminoacid sequence of ECD 167-180, and cysteines substituted by serines (SRASHPSSPMSKGS), was synthesized. Furthermore the peptide with the substitution His171→Glu was also produced, with the aim to investigate the role of His171. The fluorescence emission of Erbicin alone and in presence of each of the two peptides was monitored. A variation of the signal intensity is observed when the first peptide is added to Erbicin, whereas no marked difference with respect to Erbicin alone exists in spectra registered in presence of the mutant peptide. Finally the two peptides were used for competition ELISA assays with Erbicin and ECD by prof. De Lorenzo group. In according to predictions based on the docking model and to fluorescence results, it is found out that the not-mutated peptide inhibits binding between the two molecules, whereas the other one exerts no inhibitory activity.

To further refine the structure of the ECD-Erbicin complex, four 10ns MD trajectories were

performed using the coordinates of the rigid-docking obtained model as starting structure. Each of the produced trajectories represents a distinct exploration of the phase space accessible from the common starting structure. The generated structures are possibly representative of actual conformations of ECD-Erbicin complex in solution.

The time evolutions of the secondary structure elements and of the total number of hydrogen bonds indicate that the two molecules maintain their correct globular folding in all the simulations. Similar results are obtained analyzing the root mean square deviations of the coordinates during the simulation, which indicate that the complex structure is quite stable in three trajectories and diverges in the last 2 ns of the last one.

With the aim to characterize the relative orientation of the two molecules in the complex during the simulations, an average structure from the equilibrated portion of the four trajectories was calculated. The resulting structure is characterized by a slight larger buried surface area with respect to the starting model (790 $Å^2$ versus 750 $Å^2$). Although interface between Erbicin and ECD remains essentially the same, the antibody performs a rigid body accommodation with respect to the receptor (Figure 71). In this new model, Erbicin shows a translation of 3.8 Å of its centre of mass and a shift of the principal axis of inertia of the molecule of about 14° with respect to its initial position. The RMSDs per residue show that largest deviations are mainly localized at the ECD molecule, whereas very small deviations (RMSDs less than 3 Å) are detected for almost all the Erbicin structure. Notably, the largest deviations are far away from the receptor-ligand interface, thus confirming the reliability of the overall structure of the docked model.

Figure 71 Comparison between the starting model of the ECD-Erbicin complex, as obtained from rigid docking (coloured blue) and the average model from MD simulations (Erbicin coloured cyan, ECD orange)

Analysis of hydrogen bonds lifetime reveals that only two among the five hydrogen bonds that characterize the starting structure are well conserved (>50% lifetime) during the simulations. However, it is worth noting that new additional hydrogen bonds are observed. In particular, rather stable hydrogen bonds are formed between Asp143 and Cys170 of ECD and Arg100 of Erbicin (Figure 72),

between Ala169 of ECD and Gly225 of Erbicin and between His171 of ECD and Arg100 of the antibody.

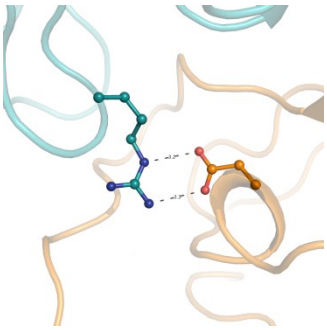

Figure 72 Hydrogen bond interactions between Erbicin Arg100 (cyan) and ECD Asp143 (orange)

These results confirm the role of ECD residues Ala169, Cys170 and His171, three residues already identified as important for the ECD-Erbicin recognition from the analysis of rigid docking complex. Furthermore these data point up that an important role in the ECD-Erbicin recognition should be played by Erbicin Arg100, placed in the H3 loop.

MATERIAL AND METHODS

Samples preparation

Fish angiogenins

All the fish angiogenins, ZF-RNase-1-5 and SS-RNase-1-2, were cloned, expressed and purified by prof. D'Alessio group, as previously reported [142, 143, 156, 157]. The proteins were provided as lyophilized samples from water unbuffered solutions. Several buffers were tested in order to identify the conditions ensuring the correct protein folding and the minimal interference with crystallization screening. All the lyophilized proteins were dissolved in 20 mM Tris/HCl buffer, pH 7.8. The accurate protein concentration was determined spectrophotometrically for all the ribonucleases, measuring their absorption at 280 nm. A small aliquot from each sample was used for circular dichroism analysis to check the conformational homogeneity before crystallization trials. CD spectra of fish angiogenins are typical of proteins with a α/β structure: they present a deep minimum at 208 nm, a shoulder at about 220 nm and a high positive peak at 194 nm. As example in Figure 73 CD spectrum of ZF-RNase-5 is reported.

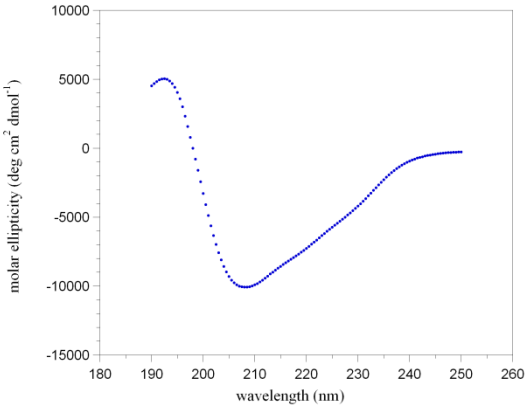

Figure 73 CD spectrum of ZF-RNase-5

Non covalent swapped dimeric form of RNase A mutants

The four RNase A mutants, LCC, PLCC, GNPSCC and GPLCC, were cloned, expressed and purified in the monomeric form, with Cys31 and Cys32 linked to two glutathione molecules, by prof. Picone and her group, as described elsewhere [191, 244]. These protecting moieties were removed by mild reduction and the resulting free monomers readily formed swapped and unswapped dimers. The non-covalent dimers were obtained from the equilibrium mixture (MxM and M=M) through selective reduction of interchain disulfide bridges, followed by alkylation of free thiol groups and by separation of monomeric and dimeric species by using gel permeation chromatography. All the samples were stored in 25 mM phosphate buffer at pH 6.7, because phosphate groups is known to stabilize non covalent swapped dimers by mimicking substrate phosphodiesteric group. This because the NCD form

contains two composite active sites with catalytic residues belonging to the different chains. In a few cases complexes between the ribonucleases and the substrate analog 2'-deoxycytidylyl(3',5')-2'-deoxyguanosine (dCpdG) were prepared, by adding to the protein solution a 1:10 molar excess of the dinucleotide.

Crystallization

Fish angiogenins

Initial crystallization trials for all the fish ribonucleases, ZF-RNase-1-5 and SS-RNase-1-2, were performed at 20°C using the sparse-matrix sampling method. The screening was carried out using commercially available kits (Crystal Screen kits I and II and Index kit from Hampton Research) and performing the trials both by hand and by a Nanodispersion Robotic System, "Microlab Starlet Nanojet 8+1" by *Hamilton Life Science Robotics*, at the Istituto di Biostrutture e Bioimmagini of CNR. Promising conditions were identified for ZF-RNase-1, -3, -5 and SS-RNase-2. After optimization process, well diffracting crystals of ZF-RNase1 and -5 and SS-RNase-2 were grown.

Crystals of all the three proteins were obtained by hanging drop vapour diffusion method at 20°C. In the case of ZF-RNase-1 the drops were prepared by mixing equal volumes of protein solution (15 mg/ml) and reservoir solution containing 1.8-2 M ammonium sulphate and 0.1 M sodium acetate pH 4.2-4.8 (ZF1_pH4.5). ZF-RNase-1 crystals were particularly big and resistant. Thus to analyze the structure of this enzyme at higher pH, the crystals, grown at pH 4.5, were transferred to harvesting solutions (2.3 M ammonium sulphate) buffered at pH values in the range 4.5-7.3. A stepwise procedure in which the pH was adjusted in steps of 0.2 was necessary to avoid deterioration of crystal quality. Crystals were stored under the final pH condition (ZF1_pH7.3) for a week with daily changes of the mother liquor. The best crystals of ZF-RNase-5 were obtained by mixing equal volumes of protein (15 mg/ml) and precipitant solution containing 32 % w/v MPEG 2000, 0.1 M sodium acetate (pH 4.5) and 0.2 M ammonium sulphate. Finally the best crystals of SS-RNase-2 grew at 15 mg/ml protein concentration from a solution containing 32 % w/v PEG 4000, 0.1 M Tris/HCl pH 7.4, 0.2 M sodium acetate and 1% v/v acetonytril at a protein-to-reservoir solution ratio of 1:1 in 1 μl drops.

Preparation and crystallization of fish angiogenins in complex with substrate analogues

With the aim to investigate the angiogenins/oligonucleotide interactions, many attempts to obtain the complexes were done. Because it was found that thermal unfolding of RNases, including fish RNases, is a reversible process, we tried to facilitate the substrate analogues binding by heating solutions of the proteins in presence of an excess of oligonucleotide at 40°C or 60°C, two temperatures corresponding to starting and ending points of the unfolding process, respectively, for ten minutes. Then the samples were slowly cooled down to room temperature and left one night at 20°C. The correct refolding of the

RNases was monitored by circular dichroism measurements. It was found that a small aliquot of the sample precipitates, but the supernatant is perfectly folded. Thus the samples were used for crystallization trials. Good quality crystals were obtained in the case of ZF-RNase-1 and ZF-RNase-3 in presence of dCpdG or 3'UMP, and in the case of SS-RNase-2 in presence of 3'UMP or of the tetra deoxynucleotide dACAA. Unfortunately diffraction data analysis showed that in none of these cases the substrate analogue binds the angiogenin.

Non covalent swapped dimeric form of RNase A mutants

Crystals of all the NCD mutants were obtained by hanging drop vapour diffusion method at 20°C using a protein concentration of 10 mg/ml. NCD-GNPSCC and NCD-PLCC were crystallized in a ligand free form, lf-NCD-GNPSCC and lf-NCD-PLCC (the NCD-PLCC-dCpdG complex structure, lb-NCD-PLCC, was previously solved during the degree thesis), NCD-LCC was crystallized in complex with the dinucleotide, lb-NCD-LCC, whereas NCD-GPLCC in both the forms (lf- and lb-NCD-GPLCC). Crystallization conditions are reported in Table 14.

Table 14 Crystallization conditions for NCD forms of RNase A mutants

Lb-NCD-LCC	Lf-NCD-PLCC	Lf-NCD-GNPSCC	Lb-/ lf-NCD-GPLCC
21% w/v PEG 35000, 0.2 M sodium acetate, 0.1 M Tris/HCl H 8.5	26% w/v PEG 4000, 0.2 M ammonium sulphate, 0.1 M Tris/HCl H 8.5	30% w/v PEG 4000, 0.2 M lithium sulphate, 0.1 M Tris/HCl H 8.5	30% w/v PEG 20000, 0.2 M ammonium sulphate
7.4 mM dCpdG	-	-	7.4 mM dCpdG/7.4 mM dATAA

Data collection

Diffraction data from all the NCD-RNase A mutants, with the exception of lf-NCD-PLCC and lf-NCD-GPLCC, and from all the fish angiogenins were collected in-house on a Saturn944 CCD detector. The X-ray radiation used was Cu *Ka* radiation from a Rigaku Micromax 007 HF generator. Diffraction data from lf-NCD-PLCC and lf-NCD-GPLCC and from one crystal of ZF1_pH4.5 were collected by using synchrotron light at ELETTRA, Trieste. All crystals were flash-cooled at 100K in supercooled N_2 gas produced by an Oxford Cryosystem and maintained at 100K during the data collection. Crystals were flash-cooled after the addition of 25% v/v glycerol to the harvesting solution. All the data were indexed, processed and scaled with HKL2000 program [115]. Detailed statistics on data collection are reported in Table 15 and Table 16. Matthews coefficient calculation suggested the presence of one dimer in the asymmetric unit of all the NCD-mutants crystals, one protein molecule in the asymmetric unit of ZF-RNase-5 crystals and two protein molecules in the asymmetric unit of ZF-RNase-1 and SS-RNase-2.

Table 15 X-ray data collection statistics for fish angiogenins

	ZF-1_pH4.5 (low resolution)	ZF-1_pH4.5 (high resolution)	ZF-1_pH7.3	ZF-RNase-5	SS-RNase-2
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$I4_132$	$P2_1$
Unit-cell parameters					
a (Å)	41.43	41.34	41.57	126.49	28.97
b (Å)	49.01	48.97	49.18	126.49	70.75
c (Å)	110.05	109.91	110.28	126.49	53.54
α (°)	90.00	90.00	90.00	90.000	90.00
β (°)	90.00	90.00	90.00	90.000	90.92
γ (°)	90.00	90.00	90.00	90.000	90.00
Resolution limits (Å)	30.00-1.65 (1.71-1.65)	30.00-1.38 (1.42-1.38)	30.00-1.61 (1.67-1.61)	50.00-1.80 (1.86-1.80)	50 - 1.89 (1.96 - 1.89)
No. of observations	746245	430152	451397	913538	
No. of unique reflections	27166	41877	28093	16327	17350
Completeness (%)	97.8 (81.9)	89.7 (54.1)	93.3 (51.4)	99.9 (99.9)	99.4 (96.3)
$I/\sigma(I)$	21 (3)	29 (2)	24 (5)	24 (3)	15 (2)
Average multiplicity	7.1	3.2	4.9	10.0	3
Rmerge (%)	5.9 (23.8)	4.9 (38.5)	5.4 (11.5)	9.5 (64.8)	7.2 (31)
Mosaicity	0.48	0.43	0.65	0.52	
V_M ($Å^3Da^{-1}$)	1.95	1.94	1.97	2.92	1.97
Solvent content (%)	37.0	36.7	37.54	57.90	37.67

Rmerge = $\sum_{hkl}\sum_i |I_i(hkl)-\langle I_i(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is ith intensity measurement of the reflection hkl, including symmetry related reflections, and $\langle I_i(hkl)\rangle$ is its average.
Values in brackets are for the highest resolution shell

Table 16 X-ray data collection statistics for non covalent swapped dimeric form of RNase A mutants

	lb-NCD-LCC	lf-NCD-PLCC	lf-NCD-GNPSCC	lf-NCD-GPLCC	lb-NCD-GPLCC
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters					
a (Å)	37.09	37.67	38.14	46.76	37.39
b (Å)	62.35	64.68	57.46	55.90	59.10
c (Å)	56.64	55.12	58.50	46.72	56.51
α (°)	90.00	90.00	90.00	90.00	90.00
β (°)	93.6	97.2	90.2	106.18	91.98
γ (°)	90.00	90.00	90.00	90.00	90.00
Resolution limits (Å)	30.00-1.94 (2.01-1.94)	50.00-1.60 (1.66-1.60)	50.00-1.99 (2.06-1.99)	50.00-2.61 (2.70-2.61)	50.00-1.60 (1.66-1.60)
No. of observations	252202	875737	505371	63530	650134
No. of unique reflections	19072	33531	15901	6945	29332
Completeness (%)	99.5 (95.5)	97.0 (93.8)	90.7 (56.5)	96.7 (74.5)	89.8 (74.8)
$I/\sigma(I)$	19 (7)	18 (3)	21 (5)	11.9 (4.4)	13.3 (2.8)
Average multiplicity	3	7	6	3.7	3.8
Rmerge (%)	4.9 (14.3)	7.0 (37.0)	8.0 (28.5)	12.2 (24.6)	8.4 (48.7)
Mosaicity	0.8	0.8	0.9	0.60	0.64
V_M ($Å^3Da^{-1}$)	2.41	2.38	2.29	2.10	2.23
Solvent content (%)	48.98	48.30	46.27	41.29	44.81

Rmerge = $\sum_{hkl}\sum_i |I_i(hkl)-\langle I_i(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is ith intensity measurement of the reflection hkl, including symmetry related reflections, and $\langle I_i(hkl)\rangle$ is its average.
Values in brackets are for the highest resolution shell

Structure determination and refinement

All the structures, solved by molecular replacement, were refined with CNS program [119]. At the end of the refinement the geometry of all the refined protein structures were monitored using PROCHECK [116] and WHATCHECK [117] programs.

Fish angiogenins

The structure of Zf1_pH4.5 was solved by molecular replacement, using the program Phaser [120] and the model of the homologous bovine angiogenin (PDB code 1AGI) [231], from whom loop regions and peptide segments with very different amino acid sequence were removed, as the structure of the protein determined at pH 8.0 (Zf1_pH8) was not yet available. Molecular replacement was carried out using lower resolution data at 1.65 Å whereas higher resolution data (1.38 Å), collected later, were used for refinement. A clear solution was obtained with a Z-score of 5.3 and an LLG of 43.

Phaser [120] program was also employed to solve the structure of ZF-RNase-5 using Zf1_pH4.5 as starting model, whereas the structure of SS-RNase-2 was solved using the program AMoRe [118] and Zf1_pH4.5 as search model (the correlation for the highest solution was 67.2, with an R factor of 33.0). Missing peptide segments were rebuilt, when possible, using the program arp/warp of the CCP4 package [124] and by manual interventions using the program O [110] and Fourier difference maps calculated with CNS [119]. Each refinement step was followed by manual interventions using the program O to correct minor errors in the positions of some side chains and to identify solvent and ion sites. At the end of the refinement process, after inclusion of low resolution data and bulk solvent correction, good values of the crystallographic factors R and R_{free} were obtained (see Table 9 and Table 10).

Non covalent swapped dimeric form of RNase A mutants

All the structures were solved by the molecular replacement method, using the program AMoRe [118] and a search model formed by the region 1-14 and 25-124 of the RNase A structure (PDB code 1KF3) [233]. In all the cases the cross-rotation function, calculated using data in the 10–3.5 Å resolution range, clearly showed two prominent peaks, which were ascribed to the two protein chains. Then the translation function was calculated and the solution with the highest correlation and lowest R factor was fixed. A new translation function search, performed on the remaining rotation solutions, allowed to find the peaks corresponding to the second subunits. Eventually, by visual inspection of solutions with the program O, it was possible to build all the dimers, whose subunits correspond to the first molecular replacement solution and to a molecule that is symmetry related to the second solution. Hinge peptides were rebuilt manually, when possible, visualizing Fourier difference maps with the program O. The structures were refined using a similar protocol: each refinement run performed with the CNS program

was followed by manual interventions using the program O to correct minor errors in the positions of some side chains. Visual inspections were also performed to identify water molecule sites and to locate the ligand molecules. At the end of the refinement process, after inclusion of low resolution data and bulk solvent correction, the crystallographic factors R and R_{free} reached values indicating good quality structures (see Table 12 and Table 13). The partially refined structure of lf-GPLCC, in the 10.0-2.61 Å resolution range, has a R factor of 0.23 and a R_{free} of 0.32 .

Homology modelling

Fish angiogenins

Since no crystal was obtained for ZF-RNase-2 and ZF-RNase-4, the structures of these enzymes were determined *in silico* by homology modelling. In particular they were modelled with the SWISS-MODEL server [245], using as reference the crystal structure of ZF-RNase-5 and ZF-RNase-3 [155], respectively. The templates were chosen on the basis of the highest score and the lowest E-value found with a BLAST search against the PDB database. Positional sequence identity between ZF-RNase-2 and ZF-RNase-5 is 73%, whereas that between ZF-RNase-4 and ZF-RNase-3 is 33%. The models produced by the server were then manually adjusted to built missing residues on the N-terminal α-helical region and energy-minimized *in vacuo* by means of the GROMOS96 force-field, following a procedure previously reported [246, 247].

Immuno-RNase

Also in the case of the immuno-RNase the extensive and detailed research of the crystallization conditions did not produce positive results. It was, therefore, decided to obtain a model of the antibody, Erbicin, by homology modelling. Erbicin structure was built by using the canonical structures method for the hypervariable loops [241, 242, 248] and standard homology modelling techniques for the framework regions. Briefly, the framework structure of the light and heavy chain variable domains (V_L and V_H) from the PDB code 1DZB [249] was used as the scaffolding on which the six complementarity determining regions (CDR) loops are built. The CDR loops were assigned according to the definitions proposed by Chothia and coworkers [241, 242] with the exception of the H3 CDR loop that is predicted *de novo*. This is a short (6 residues) loop, which should have a reduced conformational accessible space and only few conformations compatible with the rest of the protein structure. The Erbicin model was validated by the PROCHECK [116] program ProsaII [250] and CCP4 [124] programs.

Rigid docking and molecular dynamics

Rigid docking [251] of the Erbicin model onto ErbB2 ECD was performed using FTDock [252]. Given two molecules, FTDock computes the three-dimensional transformations of one of the molecules with

respect to the other with the goal of maximizing surface shape complementarity while minimizing the number of steric clashes. The scoring method of FTDock also includes electrostatic filters. The candidate models were then scored according to an energy function. The solutions were visually examined, clusterized and evaluated with respect to experimental and theoretical criterions. The final complex structure was then studied to analyze the intermolecular contacts and identify specific residue interactions between the proteins. This protocol allowed a successful prediction of the structure of ECD-Pertuzumab (PDB code:1S78 [243]) and ECD-Trastuzumab (PDB code 1N8Z [218]) complexes. A protein-protein interaction server was used to identify the residues at the interface in the complex and to evaluate the interface features. The presence of putative hydrogen bonds and salt bridges were calculated with HBPLUS program [121]. Assessment of the complex model using the PROCHECK [116], ProsaII [250] and CCP4 [124] programs suggests that it has a low energy, a good stereochemical quality, and structural features of the interface (including the surface complementarity Sc value comparable with those observed in ECD-Tratuzumab and ECD-Pertuzumab complexes.

The model obtained by the rigid docking calculation was further refined using molecular dynamics (MD) techniques. MD simulations were performed with the GROMACS 3.2 package [253] using a procedure described elsewhere [229, 254, 255]. Briefly, four distinct simulations of 10ns, differing in initial velocity coupled to protein atoms, were performed at a constant temperature of 300K, using periodic boundary conditions. In all cases a non bonded cut-off of 14 Å for the Lennard Jones potential and a dielectric constant of 1 were used. The treatment of coulombic interactions was carried out by using the Part Mesh Ewald method (PME) [256], with a grid size of 0.9 Å. All bond lengths were constrained by Lincs [257]. The Newton's equation was integrated within a time step of 2 fs and atomic coordinates were saved for analyses every 0.5 ps.

The figures were generated using Pymol (http://pymol.org) and Molscript [258].

APPENDIX: SIDE PROJECTS

During the PhD work, I have also been involved in the structural analysis, performed by X-ray diffraction and spectroscopic techniques, of superoxide dismutases (SODs) extracted from bacteria: *Ph*SOD, from the psycrophilic eubacterium *Pseudoalteromonas haloplanktis*, and *Sm*SOD, from *Streptococcus mutans*.

The characterization of *Ph*SOD falls in the more general study of psycrophilic proteins, voted to understand the molecular basis of cold-adaptation. The structures of this protein from two different crystal forms and of its complex with sodium azide, that inhibits the enzyme catalysis, were delucidated at a resolution in the range 2.10-1.50 Å. The crystallographic information was combined with a characterization of the chemical and thermal stability, performed by circular dichroism and fluorescence measurements, of both *Ph*SOD and its mesophilic counterpart from *Escherichia coli*, *Ec*SOD. The extensive comparison of structural and biophysical data, in particular of the conformational flexibility that was obtained by analysis of crystallographic B factors, for *PhSOD* and *EcSOD* leads to remarkable results. These dimeric enzymes share a large structural similarity, such as a conserved tertiary structure and arrangement of the two monomers, an almost identical total number of inter- and intra-molecular hydrogen bonds and salt bridges. However, the cold-adapted *Ph*SOD shows an increased flexibility of the active site residues with respect to its mesophilic homologue. The denaturing temperature of *Ph*SOD is comparable with that of *Ec*SOD, on the contrary, the values of the denaturant (urea and guanidinium chloride) concentration at the transition midpoint, lower for *Ph*SOD with respect to *Ec*SOD, are in line with the psychrophilic/mesophilic origin of the proteins. Altogether, these data provide additional support to the hypothesis that cold-adapted enzymes achieve efficient catalysis at low temperature by increasing the flexibility of their active site; moreover, our results underline how fine structural modifications can alter enzyme flexibility and/or stability without compromising the overall structure of a typical rigid enzyme, such as SODs.

Two *Ph*SOD mutants, with a fully exposed free cysteine (Cys57) mutated either to Arg (C57R) or to Ser (C57S), were also characterized by X-ray diffraction and biophysical techniques. The structure of C57R was determined from two different crystal forms at 2.06 and 1.66 Å, respectively, whereas that of C57S was solved at 1.97 Å. The X-ray structures of the mutants show that the introduction of the arginine allows the formation of additional salt bridges on the protein surface, whereas the Cys→Ser substitution leaves the protein structure essentially unaltered. In agreement with crystallographic results, thermal and denaturant induced unfolding followed by circular dichroism and fluorescence reveals that the mutant C57R significantly alters the protein thermal and denaturant resistance, whereas the C57S

mutant does not. These findings are useful to understand the role of electrostatic interactions in the unusual thermal tolerance of this psychrophilic enzyme: they support the conclusion that the thermal resistance of *Ph*SOD strongly depends on the distribution of charged atoms on the surface.

The case of *Sm*SOD is particularly interesting because it is a cambialistic enzyme. Unlike most metal-enzymes, it is able to bind and/or exchange two different metals, Mn or Fe, in its active site. The specific activity, as well as the sensitivity to inhibitors, depends on the metal nature. Temperature and guanidinium chloride induced unfolding was investigated by means of circular dichroism, fourth derivative UV spectroscopy and fluorescence measurements. Our data show that both unfolding processes occur through a three-state model, characterized by functional dimeric intermediates species, with different three dimensional structures. It is interesting that *Sm*SOD is the smallest dimeric protein, known so far, that populates dimeric rather than monomeric intermediates during denaturation processes. Very recently diffracting crystals of *Sm*SOD have been obtained and the structure at 2.15 Å resolution is under refinement. The comparison of *Sm*SOD structure with those of other Fe-SOD and Mn-SOD, will help in the identification of the molecular bases of the enzyme ability to bind either Fe and Mn.

REFERENCES

[1] D. Svozil, J. Kalina, M. Omelka, B. Schneider, DNA conformations and their sequence preferences, Nucleic Acids Res 36 (2008) 3690-3706.

[2] J. Suhnel, Beyond nucleic acid base pairs: from triads to heptads, Biopolymers 61 (2001) 32-51.

[3] J. Alzeer, B.R. Vummidi, P.J. Roth, N.W. Luedtke, Guanidinium-modified phthalocyanines as high-affinity G-quadruplex fluorescent probes and transcriptional regulators, Angew Chem Int Ed Engl 48 (2009) 9362-9365.

[4] M. Gellert, M.N. Lipsett, D.R. Davies, Helix formation by guanylic acid, Proc Natl Acad Sci U S A 48 (1962) 2013-2018.

[5] J.R. Williamson, G-quartet structures in telomeric DNA, Annu Rev Biophys Biomol Struct 23 (1994) 703-730.

[6] S.M. Kerwin, G-Quadruplex DNA as a target for drug design, Curr Pharm Des 6 (2000) 441-478.

[7] S. Burge, G.N. Parkinson, P. Hazel, A.K. Todd, S. Neidle, Quadruplex DNA: sequence, topology and structure, Nucleic Acids Res 34 (2006) 5402-5415.

[8] T.M. Ou, Y.J. Lu, J.H. Tan, Z.S. Huang, K.Y. Wong, L.Q. Gu, G-quadruplexes: targets in anticancer drug design, ChemMedChem 3 (2008) 690-713.

[9] T. Simonsson, R. Sjoback, DNA tetraplex formation studied with fluorescence resonance energy transfer, J Biol Chem 274 (1999) 17379-17383.

[10] G. Laughlan, A.I. Murchie, D.G. Norman, M.H. Moore, P.C. Moody, D.M. Lilley, B. Luisi, The high-resolution crystal structure of a parallel-stranded guanine tetraplex, Science 265 (1994) 520-524.

[11] F.M. Chen, Sr2+ facilitates intermolecular G-quadruplex formation of telomeric sequences, Biochemistry 31 (1992) 3769-3776.

[12] W. Guschlbauer, J.F. Chantot, D. Thiele, Four-stranded nucleic acid structures 25 years later: from guanosine gels to telomer DNA, J Biomol Struct Dyn 8 (1990) 491-511.

[13] C.C. Hardin, T. Watson, M. Corregan, C. Bailey, Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG), Biochemistry 31 (1992) 833-841.

[14] C. Caceres, G. Wright, C. Gouyette, G. Parkinson, J.A. Subirana, A thymine tetrad in d(TGGGGT) quadruplexes stabilized with Tl+/Na+ ions, Nucleic Acids Res 32 (2004) 1097-1102.

[15] P. Schultze, N.V. Hud, F.W. Smith, J. Feigon, The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the Oxytricha nova telomere repeat oligonucleotide d(G(4)T(4)G(4)), Nucleic Acids Res 27 (1999) 3018-3028.

[16] S. Haider, G.N. Parkinson, S. Neidle, Crystal structure of the potassium form of an Oxytricha nova G-quadruplex, J Mol Biol 320 (2002) 189-200.

[17] K. Phillips, Z. Dauter, A.I. Murchie, D.M. Lilley, B. Luisi, The crystal structure of a parallel-stranded guanine tetraplex at 0.95 A resolution, J Mol Biol 273 (1997) 171-182.

[18] P. Schultze, F.W. Smith, J. Feigon, Refined solution structure of the dimeric quadruplex formed from the Oxytricha telomeric oligonucleotide d(GGGGTTTTGGGG), Structure 2 (1994) 221-233.

[19] M.P. Horvath, S.C. Schultz, DNA G-quartets in a 1.86 A resolution structure of an Oxytricha nova telomeric protein-DNA complex, J Mol Biol 310 (2001) 367-377.

[20] Y. Xu, Y. Noguchi, H. Sugiyama, The new models of the human telomere d[AGGG(TTAGGG)3] in K+ solution, Bioorg Med Chem 14 (2006) 5584-5591.

[21] A. Ambrus, D. Chen, J. Dai, T. Bialis, R.A. Jones, D. Yang, Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in

potassium solution, Nucleic Acids Res 34 (2006) 2723-2735.
[22] K.N. Luu, A.T. Phan, V. Kuryavyi, L. Lacroix, D.J. Patel, Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold, J Am Chem Soc 128 (2006) 9963-9970.
[23] K. Paeschke, S. Juranek, T. Simonsson, A. Hempel, D. Rhodes, H.J. Lipps, Telomerase recruitment by the telomere end binding protein-beta facilitates G-quadruplex DNA unfolding in ciliates, Nat Struct Mol Biol 15 (2008) 598-604.
[24] K. Paeschke, T. Simonsson, J. Postberg, D. Rhodes, H.J. Lipps, Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo, Nat Struct Mol Biol 12 (2005) 847-854.
[25] C. Schaffitzel, I. Berger, J. Postberg, J. Hanes, H.J. Lipps, A. Pluckthun, In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei, Proc Natl Acad Sci U S A 98 (2001) 8572-8577.
[26] W. Palm, T. de Lange, How shelterin protects mammalian telomeres, Annu Rev Genet 42 (2008) 301-334.
[27] E.H. Blackburn, Telomere states and cell fates, Nature 408 (2000) 53-56.
[28] E.H. Blackburn, Switching and signaling at the telomere, Cell 106 (2001) 661-673.
[29] M.Z. Levy, R.C. Allsopp, A.B. Futcher, C.W. Greider, C.B. Harley, Telomere end-replication problem and cell aging, J Mol Biol 225 (1992) 951-960.
[30] L.K. Wai, Telomeres, telomerase, and tumorigenesis--a review, MedGenMed 6 (2004) 19.
[31] S. Neidle, L.R. Kelland, Telomerase as an anti-cancer target: current status and future prospects, Anticancer Drug Des 14 (1999) 341-347.
[32] A.E. Pitts, D.R. Corey, The telomerase challenge - an unusual problem in drug discovery, Drug Discov Today 4 (1999) 155-161.
[33] D.J. Bearss, L.H. Hurley, D.D. Von Hoff, Telomere maintenance mechanisms as a target for drug development, Oncogene 19 (2000) 6632-6641.
[34] L.R. Kelland, Telomerase inhibitors: targeting the vulnerable end of cancer?, Anticancer Drugs 11 (2000) 503-513.
[35] A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J.F. Riou, J.L. Mergny, Targeting telomeres and telomerase, Biochimie 90 (2008) 131-155.
[36] F. Pendino, I. Tarkanyi, C. Dudognon, J. Hillion, M. Lanotte, J. Aradi, E. Segal-Bendirdjian, Telomeres and telomerase: Pharmacological targets for new anticancer strategies?, Curr Cancer Drug Targets 6 (2006) 147-180.
[37] T. Simonsson, P. Pecinka, M. Kubista, DNA tetraplex formation in the control region of c-myc, Nucleic Acids Res 26 (1998) 1167-1172.
[38] A. Siddiqui-Jain, C.L. Grand, D.J. Bearss, L.H. Hurley, Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription, Proc Natl Acad Sci U S A 99 (2002) 11593-11598.
[39] A.T. Phan, Y.S. Modi, D.J. Patel, Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter, J Am Chem Soc 126 (2004) 8710-8716.
[40] A. Ambrus, D. Chen, J. Dai, R.A. Jones, D. Yang, Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization, Biochemistry 44 (2005) 2048-2058.
[41] S. Rankin, A.P. Reszka, J. Huppert, M. Zloh, G.N. Parkinson, A.K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, Putative DNA quadruplex formation within the human c-kit oncogene, J Am Chem Soc 127 (2005) 10584-10589.
[42] H. Fernando, A.P. Reszka, J. Huppert, S. Ladame, S. Rankin, A.R. Venkitaraman, S. Neidle, S. Balasubramanian, A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene, Biochemistry 45 (2006) 7854-7860.
[43] R. De Armond, S. Wood, D. Sun, L.H. Hurley, S.W. Ebbinghaus, Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor

1alpha promoter, Biochemistry 44 (2005) 16341-16350.
[44] D. Sun, K. Guo, J.J. Rusche, L.H. Hurley, Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents, Nucleic Acids Res 33 (2005) 6070-6080.
[45] J. Dai, D. Chen, R.A. Jones, L.H. Hurley, D. Yang, NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region, Nucleic Acids Res 34 (2006) 5133-5144.
[46] J. Dai, T.S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R.A. Jones, D. Yang, An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution, J Am Chem Soc 128 (2006) 1096-1098.
[47] S. Cogoi, L.E. Xodo, G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription, Nucleic Acids Res 34 (2006) 2536-2549.
[48] A.I. Murchie, D.M. Lilley, Retinoblastoma susceptibility genes contain 5' sequences with a high propensity to form guanine-tetrad structures, Nucleic Acids Res 20 (1992) 49-53.
[49] Y. Xu, H. Sugiyama, Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb), Nucleic Acids Res 34 (2006) 949-954.
[50] K. Guo, A. Pourpak, K. Beetz-Rogers, V. Gokhale, D. Sun, L.H. Hurley, Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene, J Am Chem Soc 129 (2007) 10220-10228.
[51] C.A. Spencer, M. Groudine, Control of c-myc regulation in normal and neoplastic cells, Adv Cancer Res 56 (1991) 1-48.
[52] K.B. Marcu, S.A. Bossone, A.J. Patel, myc function and regulation, Annu Rev Biochem 61 (1992) 809-860.
[53] L.M. Facchini, L.Z. Penn, The molecular role of Myc in growth and transformation: recent discoveries lead to new insights, FASEB J 12 (1998) 633-651.
[54] S. Pelengaris, B. Rudolph, T. Littlewood, Action of Myc in vivo - proliferation and apoptosis, Curr Opin Genet Dev 10 (2000) 100-105.
[55] L.C. Bock, L.C. Griffin, J.A. Latham, E.H. Vermaas, J.J. Toole, Selection of single-stranded DNA molecules that bind and inhibit human thrombin, Nature 355 (1992) 564-566.
[56] J.M. Burke, A. Berzal-Herranz, In vitro selection and evolution of RNA: applications for catalytic RNA, molecular recognition, and drug discovery, FASEB J 7 (1993) 106-112.
[57] B. Chakraborty, Z. Jiang, Y. Li, H.-Z. Yu, Rational design and performance testing of aptamer-based electrochemical biosensors for adenosine, Journal of Electroanalytical Chemistry 365 (2009) 75-82.
[58] M. del Toro, R. Gargallo, R. Eritja, J. Jaumot, Study of the interaction between the G-quadruplex-forming thrombin-binding aptamer and the porphyrin 5,10,15,20-tetrakis-(N-methyl-4-pyridyl)-21,23H-porphyrin tetratosylate, Anal Biochem 379 (2008) 8-15.
[59] E. Torres-Chavolla, E.C. Alocilja, Aptasensors for detection of microbial and viral pathogens, Biosens Bioelectron 24 (2009) 3175-3182.
[60] A. Numnuam, K.Y. Chumbimuni-Torres, Y. Xiang, R. Bash, P. Thavarungkul, P. Kanatharana, E. Pretsch, J. Wang, E. Bakker, Aptamer-based potentiometric measurements of proteins using ion-selective microelectrodes, Anal Chem 80 (2008) 707-712.
[61] C. Wang, M. Zhang, G. Yang, D. Zhang, H. Ding, H. Wang, M. Fan, B. Shen, N. Shao, Single-stranded DNA aptamers that bind differentiated but not parental cells: subtractive systematic evolution of ligands by exponential enrichment, J Biotechnol 102 (2003) 15-22.
[62] K.N. Morris, K.B. Jensen, C.M. Julin, M. Weil, L. Gold, High affinity ligands from in vitro selection: complex targets, Proc Natl Acad Sci U S A 95 (1998) 2902-2907.
[63] B. Gatto, M. Palumbo, C. Sissi, Nucleic acid aptamers based on the G-quadruplex structure: therapeutic and diagnostic potential, Curr Med Chem 16 (2009) 1248-1265.
[64] S.M. Nimjee, C.P. Rusconi, R.A. Harrington, B.A. Sullenger, The potential of aptamers as

anticoagulants, Trends Cardiovasc Med 15 (2005) 41-45.
[65] P.J. Bates, D.A. Laber, D.M. Miller, S.D. Thomas, J.O. Trent, Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer, Exp Mol Pathol 86 (2009) 151-164.
[66] E.W. Choi, L.V. Nayak, P.J. Bates, Cancer-selective antiproliferative activity is a general property of some G-rich oligodeoxynucleotides, Nucleic Acids Res 38 1623-1635.
[67] P.J. Bates, J.B. Kahlon, S.D. Thomas, J.O. Trent, D.M. Miller, Antiproliferative activity of G-rich oligonucleotides correlates with protein binding, J Biol Chem 274 (1999) 26369-26377.
[68] X. Xu, F. Hamhouyia, S.D. Thomas, T.J. Burke, A.C. Girvan, W.G. McGregor, J.O. Trent, D.M. Miller, P.J. Bates, Inhibition of DNA replication and induction of S phase cell cycle arrest by G-rich oligonucleotides, J Biol Chem 276 (2001) 43221-43230.
[69] M. Ramanathan, M. Lantz, R.D. MacGregor, M.R. Garovoy, C.A. Hunt, Characterization of the oligodeoxynucleotide-mediated inhibition of interferon-gamma-induced major histocompatibility complex class I and intercellular adhesion molecule-1, J Biol Chem 269 (1994) 24564-24574.
[70] M. Ramanathan, M. Lantz, R.D. MacGregor, B. Huey, S. Tam, Y. Li, M.R. Garovoy, C.A. Hunt, Inhibition of interferon-gamma-induced major histocompatibility complex class I expression by certain oligodeoxynucleotides, Transplantation 57 (1994) 612-615.
[71] V. Balasubramanian, L.T. Nguyen, S.V. Balasubramanian, M. Ramanathan, Interferon-gamma-inhibitory oligodeoxynucleotides alter the conformation of interferon-gamma, Mol Pharmacol 53 (1998) 926-932.
[72] J.A. Huntington, Molecular recognition mechanisms of thrombin, J Thromb Haemost 3 (2005) 1861-1872.
[73] J.T. Crawley, S. Zanardelli, C.K. Chion, D.A. Lane, The central role of thrombin in hemostasis, J Thromb Haemost 5 Suppl 1 (2007) 95-101.
[74] E. Di Cera, Thrombin as procoagulant and anticoagulant, J Thromb Haemost 5 Suppl 1 (2007) 196-202.
[75] E. Di Cera, Q.D. Dang, Y.M. Ayala, Molecular mechanisms of thrombin function, Cell Mol Life Sci 53 (1997) 701-730.
[76] P.E. Bock, P. Panizzi, I.M. Verhamme, Exosites in the substrate specificity of blood coagulation reactions, J Thromb Haemost 5 Suppl 1 (2007) 81-94.
[77] W. Bode, I. Mayr, U. Baumann, R. Huber, S.R. Stone, J. Hofsteenge, The refined 1.9 A crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment, EMBO J 8 (1989) 3467-3475.
[78] W. Bode, D. Turk, A. Karshikov, The refined 1.9-A X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human alpha-thrombin: structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure-function relationships, Protein Sci 1 (1992) 426-471.
[79] W. Bode, Structure and interaction modes of thrombin, Blood Cells Mol Dis 36 (2006) 122-130.
[80] J.C. Rau, L.M. Beaulieu, J.A. Huntington, F.C. Church, Serpins in thrombosis, hemostasis and fibrinolysis, J Thromb Haemost 5 Suppl 1 (2007) 102-115.
[81] A. Lombardi, G. De Simone, S. Galdiero, N. Staiano, F. Nastri, V. Pavone, From natural to synthetic multisite thrombin inhibitors, Biopolymers 51 (1999) 19-39.
[82] E.A. Nutescu, N.L. Shapiro, A. Chevalier, New anticoagulant agents: direct thrombin inhibitors, Cardiol Clin 26 (2008) 169-187, v-vi.
[83] L.C. Griffin, G.F. Tidmarsh, L.C. Bock, J.J. Toole, L.L. Leung, In vivo anticoagulant properties of a novel nucleotide-based thrombin inhibitor and demonstration of regional anticoagulation in extracorporeal circuits, Blood 81 (1993) 3271-3276.
[84] W.X. Li, A.V. Kaplan, G.W. Grant, J.J. Toole, L.L. Leung, A novel nucleotide-based thrombin inhibitor inhibits clot-bound thrombin and reduces arterial platelet thrombus formation, Blood

83 (1994) 677-682.
[85] R.F. Macaya, P. Schultze, F.W. Smith, J.A. Roe, J. Feigon, Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution, Proc Natl Acad Sci U S A 90 (1993) 3745-3749.
[86] K.Y. Wang, S. McCurdy, R.G. Shea, S. Swaminathan, P.H. Bolton, A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA, Biochemistry 32 (1993) 1899-1904.
[87] K. Padmanabhan, K.P. Padmanabhan, J.D. Ferrara, J.E. Sadler, A. Tulinsky, The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer, J Biol Chem 268 (1993) 17651-17654.
[88] K. Padmanabhan, A. Tulinsky, An ambiguous structure of a DNA 15-mer thrombin complex, Acta Crystallogr D Biol Crystallogr 52 (1996) 272-282.
[89] B. Pagano, L. Martino, A. Randazzo, C. Giancola, Stability and binding properties of a modified thrombin binding aptamer, Biophys J 94 (2008) 562-569.
[90] S.R. Nallagatla, B. Heuberger, A. Haque, C. Switzer, Combinatorial synthesis of thrombin-binding aptamers containing iso-guanine, J Comb Chem 11 (2009) 364-369.
[91] G.X. He, S.H. Krawczyk, S. Swaminathan, R.G. Shea, J.P. Dougherty, T. Terhorst, V.S. Law, L.C. Griffin, S. Coutre, N. Bischofberger, N2- and C8-substituted oligodeoxynucleotides with enhanced thrombin inhibitory activity in vitro and in vivo, J Med Chem 41 (1998) 2234-2242.
[92] L. Bonifacio, F.C. Church, M.B. Jarstfer, Effect of locked-nucleic acid on a biologically active g-quadruplex. A structure-activity relationship of the thrombin aptamer, Int J Mol Sci 9 (2008) 422-433.
[93] A. Virno, A. Randazzo, C. Giancola, M. Bucci, G. Cirino, L. Mayol, A novel thrombin binding aptamer containing a G-LNA residue, Bioorg Med Chem 15 (2007) 5710-5718.
[94] C.G. Peng, M.J. Damha, G-quadruplex induced stabilization by 2'-deoxy-2'-fluoro-D-arabinonucleic acids (2'F-ANA), Nucleic Acids Res 35 (2007) 4977-4988.
[95] A. Pasternak, F.J. Hernandez, L.M. Rasmussen, B. Vester, J. Wengel, Improved thrombin binding aptamer by incorporation of a single unlocked nucleic acid monomer, Nucleic Acids Res.
[96] G.X. He, J.P. Williams, M.J. Postich, S. Swaminathan, R.G. Shea, T. Terhorst, V.S. Law, C.T. Mao, C. Sueoka, S. Coutre, N. Bischofberger, In vitro and in vivo activities of oligodeoxynucleotide-based thrombin inhibitors containing neutral formacetal linkages, J Med Chem 41 (1998) 4224-4231.
[97] L. Martino, A. Virno, A. Randazzo, A. Virgilio, V. Esposito, C. Giancola, M. Bucci, G. Cirino, L. Mayol, A new modified thrombin binding aptamer containing a 5'-5' inversion of polarity site, Nucleic Acids Res 34 (2006) 6653-6662.
[98] E.M. McGuffie, C.V. Catapano, Design of a novel triple helix-forming oligodeoxyribonucleotide directed to the major promoter of the c-myc gene, Nucleic Acids Res 30 (2002) 2701-2709.
[99] M. Lu, Q. Guo, N.R. Kallenbach, Thermodynamics of G-tetraplex formation by telomeric DNAs, Biochemistry 32 (1993) 598-601.
[100] I. Smirnov, R.H. Shafer, Effect of loop sequence and size on DNA aptamer stability, Biochemistry 39 (2000) 1462-1468.
[101] J. Rubin, T. Brennan, M. Sundaralingam, Crystal and molecular structure of a naturally occurring dinucleoside monophosphate. Uridylyl-(3'-5')-adenosine hemihydrate. Conformational "rigidity" of the nucleotide unit and models for polynucleotide chain folding, Biochemistry 11 (1972) 3112-3128.
[102] E. Shefter, K.N. Trueblood, The Crystal and Molecular Structure of D(+)-Barium Uridine-5'-Phosphate, Acta Crystallogr 18 (1965) 1067-1077.
[103] M.C. Wahl, S.T. Rao, M. Sundaralingam, Crystal structure of the B-DNA hexamer d(CTCGAG): model for an A-to-B transition, Biophys J 70 (1996) 2857-2866.

[104] E.H. Clay, I.R. Gould, A combined QM and MM investigation into guanine quadruplexes, J Mol Graph Model 24 (2005) 138-146.
[105] N.V. Hud, F.W. Smith, F.A. Anet, J. Feigon, The selectivity for K+ versus Na+ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by 1H NMR, Biochemistry 35 (1996) 15383-15390.
[106] B. Schneider, H.M. Berman, Hydration of the DNA bases is local, Biophys J 69 (1995) 2661-2669.
[107] B. Schneider, K. Patel, H.M. Berman, Hydration of the phosphate group in double-helical DNA, Biophys J 75 (1998) 2422-2434.
[108] C.L. Kielkopf, S. Ding, P. Kuhn, D.C. Rees, Conformational flexibility of B-DNA at 0.74 A resolution: d(CCAGTACTGG)(2), J Mol Biol 296 (2000) 787-801.
[109] C.M. Deane, F.H. Allen, R. Taylor, T.L. Blundell, Carbonyl-carbonyl interactions stabilize the partially allowed Ramachandran conformations of asparagine and aspartic acid, Protein Eng 12 (1999) 1025-1028.
[110] T.A. Jones, J.Y. Zou, S.W. Cowan, M. Kjeldgaard, Improved methods for building protein models in electron density maps and the location of errors in these models, Acta Crystallogr A 47 (Pt 2) (1991) 110-119.
[111] M. Tsiang, A.K. Jain, K.E. Dunn, M.E. Rojas, L.L. Leung, C.S. Gibbs, Functional mapping of the surface residues of human thrombin, J Biol Chem 270 (1995) 16854-16863.
[112] S. Lancellotti, R. De Cristofaro, Nucleotide-derived thrombin inhibitors: a new tool for an old issue, Cardiovasc Hematol Agents Med Chem 7 (2009) 19-28.
[113] A. Lombardi, G. De Simone, F. Nastri, S. Galdiero, R. Della Morte, N. Staiano, C. Pedone, M. Bolognesi, V. Pavone, The crystal structure of alpha-thrombin-hirunorm IV complex reveals a novel specificity site recognition mode, Protein Sci 8 (1999) 91-95.
[114] G. De Simone, A. Lombardi, S. Galdiero, F. Nastri, R. Della Morte, N. Staiano, C. Pedone, M. Bolognesi, V. Pavone, Hirunorms are true hirudin mimetics. The crystal structure of human alpha-thrombin-hirunorm V complex, Protein Sci 7 (1998) 243-253.
[115] Z. Otwinowsky, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, Methods Enzymol., 1997, pp. 307-326.
[116] R.A. Laskowski, M.W. MacArthur, M.D. Moss, J.M. Thorton, PROCHECK: A program to check the stereochemical quality of protein structure., J Appl Crystallogr 26 (1993) 283-291.
[117] R.W. Hooft, G. Vriend, C. Sander, E.E. Abola, Errors in protein structures, Nature 381 (1996) 272.
[118] J. Navaza, P. Saludjian, AMoRe: an automated molecular replacement program package, Methods Enzymol. 276 (1997) 581-594.
[119] A.T. Brunger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, Crystallography & NMR system: A new software suite for macromolecular structure determination, Acta Crystallogr D Biol Crystallogr 54 (Pt 5) (1998) 905-921.
[120] A.J. McCoy, R.W. Grosse-Kunstleve, L.C. Storoni, R.J. Read, Likelihood-enhanced fast translation functions, Acta Crystallogr D Biol Crystallogr 61 (2005) 458-464.
[121] I.K. McDonald, J.M. Thornton, Satisfying hydrogen bonding potential in proteins, J Mol Biol 238 (1994) 777-793.
[122] A. Tiwari, S.K. Panigrahi, HBAT: a complete package for analysing strong and weak hydrogen bonds in macromolecular crystal structures, In Silico Biol 7 (2007) 651-661.
[123] M.C. Lawrence, P.M. Colman, Shape complementarity at protein/protein interfaces, J Mol Biol 234 (1993) 946-950.
[124] The CCP4 suite: programs for protein crystallography, Acta Crystallogr D Biol Crystallogr 50 (1994) 760-763.
[125] K. Nadassy, I. Tomas-Oliveira, I. Alberts, J. Janin, S.J. Wodak, Standard atomic volumes in double-stranded DNA and packing in protein--DNA interfaces, Nucleic Acids Res 29 (2001)

3362-3376.
[126] R.T. Raines, Ribonuclease A, Chem. Rev. 98 (1998) 1045-1065.
[127] O.N. Il'inskaia, A.A. Makarov, [Why ribonucleases cause death of cancer cells], Mol Biol (Mosk) 39 (2005) 3-13.
[128] D.A. Usher, E.S. Erenrich, F. Eckstein, Geometry of the first step in the action of ribonuclease-A (in-line geometry-uridine2',3'-cyclic thiophosphate- 31 P NMR), Proc Natl Acad Sci U S A 69 (1972) 115-118.
[129] M. Libonati, A. Floridi, Breakdown of double-stranded RNA by bull semen ribonuclease, Eur J Biochem 8 (1969) 81-87.
[130] G. D'Alessio, S. Zofra, M. Libonati, Action of dimeric ribonucleases on double-stranded RNA, FEBS Lett 24 (1972) 355-358.
[131] G. D'Alessio, J. Doskocil, M. Libonati, Action of dimeric hybrids of native and selectively alkylated ribonuclease A on double-stranded ribonucleic acid, Biochem J 141 (1974) 317-320.
[132] H.F. Rosenberg, RNase A ribonucleases and host defense: an evolving story, J Leukoc Biol 83 (2008) 1079-1087.
[133] S. Cho, J.J. Beintema, J. Zhang, The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories, Genomics 85 (2005) 208-220.
[134] G. Kartha, J. Bello, D. Harker, Tertiary structure of ribonuclease, Nature 213 (1967) 862-865.
[135] M. Moussaoui, A. Guasch, E. Boix, C. Cuchillo, M. Nogues, The role of non-catalytic binding subsites in the endonuclease activity of bovine pancreatic ribonuclease A, The Journal of biological chemistry 271 (1996) 4687-4692.
[136] X. Pares, M.V. Nogues, R. de Llorens, C.M. Cuchillo, Structure and function of ribonuclease A binding subsites, Essays Biochem 26 (1991) 89-103.
[137] 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.
[138] M.V. Nogues, M. Vilanova, C.M. Cuchillo, Bovine pancreatic ribonuclease A as a model of an enzyme with multiple substrate binding sites, Biochim Biophys Acta 1253 (1995) 16-24.
[139] K.R. Acharya, R. Shapiro, S.C. Allen, J.F. Riordan, B.L. Vallee, Crystal structure of human angiogenin reveals the structural basis for its functional divergence from ribonuclease, Proceedings of the National Academy of Sciences of the United States of America 91 (1994) 2915-2919.
[140] L. Vitagliano, A. Merlino, A. Zagari, L. Mazzarella, Productive and nonproductive binding to ribonuclease A: X-ray structure of two complexes with uridylyl(2',5')guanosine, Protein Sci 9 (2000) 1217-1225.
[141] J.J. Beintema, H.J. Breukelman, A. Carsana, A. Furia, Evolution of vertebrate ribonucleases: ribonuclease A superfamily, in: G. D'Alessio, J.F. Riordan (Eds.), Ribonucleases: structures and functions, Academic press, New York, 1997, pp. 245-269.
[142] E. Pizzo, P. Buonanno, A. Di Maro, S. Ponticelli, S. De Falco, N. Quarto, M.V. Cubellis, G. D'Alessio, Ribonucleases and angiogenins from fish, J Biol Chem 281 (2006) 27454-27460.
[143] S. Cho, J. Zhang, Zebrafish ribonucleases are bactericidal: implications for the origin of the vertebrate RNase A superfamily, Mol Biol Evol 24 (2007) 1259-1268.
[144] J.F. Riordan, Structure and function of angiogenin, in: G. D'Alessio, J.F. Riordan (Eds.), Ribonucleases: structures and functions, Academic press, New York, 1997, pp. 426-444.
[145] J. Matousek, Aspermatogenic effect of the bull seminal ribonuclease (BS RNase) in the presence of anti-BS RNase antibodies in mice, Anim Genet 25 Suppl 1 (1994) 45-50.
[146] J. Harder, J.M. Schroder, RNase 7, a novel innate immune defense antimicrobial protein of healthy human skin, J Biol Chem 277 (2002) 46779-46784.
[147] R.J. Youle, G. D'Alessio, Antitumor RNases, in: G. D'Alessio, J.F. Riordan (Eds.), Ribonucleases: structures and functions, Academic press, New York, 1997, pp. 491-514.
[148] M. Tamburrini, G. Scala, C. Verde, M.R. Ruocco, A. Parente, S. Venuta, G. D'Alessio,

Immunosuppressive activity of bovine seminal RNase on T-cell proliferation, Eur J Biochem 190 (1990) 145-148.
[149] P.J. Polverini, The pathophysiology of angiogenesis, Crit Rev Oral Biol Med 6 (1995) 230-247.
[150] J. Folkman, M. Klagsbrun, Angiogenic factors, Science 235 (1987) 442-447.
[151] J. Folkman, R. Cotran, Relation of vascular proliferation to tumor growth, Int Rev Exp Pathol 16 (1976) 207-248.
[152] J.W. Fett, D.J. Strydom, R.R. Lobb, E.M. Alderman, J.L. Bethune, J.F. Riordan, B.L. Vallee, Isolation and characterization of angiogenin, an angiogenic protein from human carcinoma cells, Biochemistry 24 (1985) 5480-5486.
[153] D.J. Strydom, J.W. Fett, R.R. Lobb, E.M. Alderman, J.L. Bethune, J.F. Riordan, B.L. Vallee, Amino acid sequence of human tumor derived angiogenin, Biochemistry 24 (1985) 5486-5494.
[154] J.J. Beintema, C. Schuller, M. Irie, A. Carsana, Molecular evolution of the ribonuclease superfamily, Prog Biophys Mol Biol 51 (1988) 165-192.
[155] K. Kazakou, D.E. Holloway, S.H. Prior, V. Subramanian, K.R. Acharya, Ribonuclease A homologues of the zebrafish: polymorphism, crystal structures of two representatives and their evolutionary implications, J Mol Biol 380 (2008) 206-222.
[156] E. Pizzo, A. Merlino, M. Turano, I. Russo Krauss, F. Coscia, A. Zanfardino, M. Varcamonti, A. Furia, C. Giancola, L. Mazzarella, F. Sica, G. D'Alessio, A new RNase sheds light on the RNase/angiogenin subfamily from zebrafish, Biochem J. (2010).
[157] E. Pizzo, M. Varcamonti, A. Di Maro, A. Zanfardino, C. Giancola, G. D'Alessio, Ribonucleases with angiogenic and bactericidal activities from the Atlantic salmon, FEBS J 275 (2008) 1283-1295.
[158] E. Pizzo, G. D'Alessio, The success of the RNase scaffold in the advance of biosciences and in evolution, Gene 406 (2007) 8-12.
[159] Z. Darzynkiewicz, S.P. Carter, S.M. Mikulski, W.J. Ardelt, K. Shogen, Cytostatic and cytotoxic effects of Pannon (P-30 Protein), a novel anticancer agent, Cell Tissue Kinet 21 (1988) 169-182.
[160] W. Ardelt, S.M. Mikulski, K. Shogen, Amino acid sequence of an anti-tumor protein from Rana pipiens oocytes and early embryos. Homology to pancreatic ribonucleases, J Biol Chem 266 (1991) 245-251.
[161] W. Ardelt, K. Shogen, Z. Darzynkiewicz, Onconase and amphinase, the antitumor ribonucleases from Rana pipiens oocytes, Curr Pharm Biotechnol 9 (2008) 215-225.
[162] S.M. Mikulski, J.J. Costanzi, N.J. Vogelzang, S. McCachren, R.N. Taub, H. Chun, A. Mittelman, T. Panella, C. Puccio, R. Fine, K. Shogen, Phase II trial of a single weekly intravenous dose of ranpirnase in patients with unresectable malignant mesothelioma, J Clin Oncol 20 (2002) 274-281.
[163] K. Nitta, K. Ozaki, M. Ishikawa, S. Furusawa, M. Hosono, H. Kawauchi, K. Sasaki, Y. Takayanagi, S. Tsuiki, S. Hakomori, Inhibition of cell proliferation by Rana catesbeiana and Rana japonica lectins belonging to the ribonuclease superfamily, Cancer Res 54 (1994) 920-927.
[164] W. Ardelt, B. Ardelt, Z. Darzynkiewicz, Ribonucleases as potential modalities in anticancer therapy, Eur J Pharmacol 625 (2009) 181-189.
[165] R. Piccoli, M. Tamburrini, G. Piccialli, A. Di Donato, A. Parente, G. D'Alessio, The dual-mode quaternary structure of seminal RNase, Proc Natl Acad Sci U S A 89 (1992) 1870-1874.
[166] L. Mazzarella, S. Capasso, D. Demasi, G. Di Lorenzo, C.A. Mattia, A. Zagari, Bovine seminal ribonuclease: structure at 1.9 Å resolution, Acta Cryst. D 49 (1993) 389-402.
[167] R. Berisio, F. Sica, C. De Lorenzo, A. Di Fiore, R. Piccoli, A. Zagari, L. Mazzarella, Crystal structure of the dimeric unswapped form of bovine seminal ribonuclease, FEBS Lett 554 (2003) 105-110.
[168] A. Benito, M. Ribo, M. Vilanova, On the track of antitumour ribonucleases, Mol Biosyst 1 (2005) 294-302.
[169] J.J. Lin, D.L. Newton, S.M. Mikulski, H.F. Kung, R.J. Youle, S.M. Rybak, Characterization of the mechanism of cellular and cell free protein synthesis inhibition by an anti-tumor

ribonuclease, Biochem Biophys Res Commun 204 (1994) 156-162.
[170] S.K. Saxena, M. Gravell, Y.N. Wu, S.M. Mikulski, K. Shogen, W. Ardelt, R.J. Youle, Inhibition of HIV-1 production and selective degradation of viral RNA by an amphibian ribonuclease, J Biol Chem 271 (1996) 20783-20788.
[171] M.S. Iordanov, O.P. Ryabinina, J. Wong, T.H. Dinh, D.L. Newton, S.M. Rybak, B.E. Magun, Molecular determinants of apoptosis induced by the cytotoxic ribonuclease onconase: evidence for cytotoxic mechanisms different from inhibition of protein synthesis, Cancer Res 60 (2000) 1983-1994.
[172] S.K. Saxena, R. Sirdeshmukh, W. Ardelt, S.M. Mikulski, K. Shogen, R.J. Youle, Entry into cells and selective degradation of tRNAs by a cytotoxic member of the RNase A family, J Biol Chem 277 (2002) 15142-15146.
[173] M.R. Mastronicola, R. Piccoli, G. D'Alessio, Key extracellular and intracellular steps in the antitumor action of seminal ribonuclease, Eur J Biochem 230 (1995) 242-249.
[174] J. Futami, T. Maeda, M. Kitazoe, E. Nukui, H. Tada, M. Seno, M. Kosaka, H. Yamada, Preparation of potent cytotoxic ribonucleases by cationization: enhanced cellular uptake and decreased interaction with ribonuclease inhibitor by chemical modification of carboxyl groups, Biochemistry 40 (2001) 7518-7524.
[175] J. Futami, E. Nukui, T. Maeda, M. Kosaka, H. Tada, M. Seno, H. Yamada, Optimum modification for the highest cytotoxicity of cationized ribonuclease, J Biochem 132 (2002) 223-228.
[176] Y. Wu, S.M. Mikulski, W. Ardelt, S.M. Rybak, R.J. Youle, A cytotoxic ribonuclease. Study of the mechanism of onconase cytotoxicity, J Biol Chem 268 (1993) 10686-10693.
[177] A. Bracale, D. Spalletti-Cernia, M. Mastronicola, F. Castaldi, R. Mannucci, L. Nitsch, G. D'Alessio, Essential stations in the intracellular pathway of cytotoxic bovine seminal ribonuclease, Biochem J 362 (2002) 553-560.
[178] M.C. Haigis, R.T. Raines, Secretory ribonucleases are internalized by a dynamin-independent endocytic pathway, J Cell Sci 116 (2003) 313-324.
[179] M. Bosch, A. Benito, M. Ribo, T. Puig, B. Beaumelle, M. Vilanova, A nuclear localization sequence endows human pancreatic ribonuclease with cytotoxic activity, Biochemistry 43 (2004) 2167-2177.
[180] F.S. Lee, B.L. Vallee, Structure and action of mammalian ribonuclease (angiogenin) inhibitor, Prog Nucleic Acid Res Mol Biol 44 (1993) 1-30.
[181] J. Hofsteenge, B. Kieffer, R. Matthies, B.A. Hemmings, S.R. Stone, Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals the presence of leucine-rich repeats, Biochemistry 27 (1988) 8537-8544.
[182] B. Kobe, J. Deisenhofer, Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats, Nature 366 (1993) 751-756.
[183] M. Blazquez, J.M. Fominaya, J. Hofsteenge, Oxidation of sulfhydryl groups of ribonuclease inhibitor in epithelial cells is sufficient for its intracellular degradation, J Biol Chem 271 (1996) 18638-18642.
[184] P.A. Leland, R.T. Raines, Cancer chemotherapy--ribonucleases to the rescue, Chem Biol 8 (2001) 405-413.
[185] A.M. Vicentini, B. Kieffer, R. Matthies, B. Meyhack, B.A. Hemmings, S.R. Stone, J. Hofsteenge, Protein chemical and kinetic characterization of recombinant porcine ribonuclease inhibitor expressed in Saccharomyces cerevisiae, Biochemistry 29 (1990) 8827-8834.
[186] B. Kobe, Z. Ma, J. Deisenhofer, Complex between bovine ribonuclease A and porcine ribonuclease inhibitor crystallizes in a similar unit cell as free ribonuclease inhibitor, J Mol Biol 241 (1994) 288-291.
[187] B. Kobe, J. Deisenhofer, Mechanism of ribonuclease inhibition by ribonuclease inhibitor protein based on the crystal structure of its complex with ribonuclease A, J Mol Biol 264 (1996) 1028-1043.

[188] A.C. Papageorgiou, R. Shapiro, K.R. Acharya, Molecular recognition of human angiogenin by placental ribonuclease inhibitor--an X-ray crystallographic study at 2.0 A resolution, EMBO J 16 (1997) 5162-5177.
[189] E. Boix, Y. Wu, V.M. Vasandani, S.K. Saxena, W. Ardelt, J. Ladner, R.J. Youle, Role of the N terminus in RNase A homologues: differences in catalytic activity, ribonuclease inhibitor interaction and cytotoxicity, J Mol Biol 257 (1996) 992-1007.
[190] F. Sica, A. Di Fiore, A. Merlino, L. Mazzarella, Structure and stability of the non-covalent swapped dimer of bovine seminal ribonuclease: an enzyme tailored to evade ribonuclease protein inhibitor, J Biol Chem 279 (2004) 36753-36760.
[191] A. Di Donato, V. Cafaro, G. D'Alessio, Ribonuclease A can be transformed into a dimeric ribonuclease with antitumor activity, J Biol Chem 269 (1994) 17394-17396.
[192] G. D'Alessio, A. Di Donato, L. Mazzarella, R. Piccoli, Seminal ribonuclease: the importance of diversity, in: G. D'Alessio, J.F. Riordan (Eds.), Ribonucleases: structures and functions, Academic press, New York, 1997, pp. 383-423.
[193] F. Sica, A. Di Fiore, A. Zagari, L. Mazzarella, The unswapped chain of bovine seminal ribonuclease: Crystal structure of the free and liganded monomeric derivative, Proteins 52 (2003) 263-271.
[194] L. Vitagliano, S. Adinolfi, F. Sica, A. Merlino, A. Zagari, L. Mazzarella, A potential allosteric subsite generated by domain swapping in bovine seminal ribonuclease, J Mol Biol 293 (1999) 569-577.
[195] S. Vescia, D. Tramontano, Antitumoral action of bovine seminal ribonuclease, Mol Cell Biochem 36 (1981) 125-128.
[196] V. Cafaro, C. De Lorenzo, R. Piccoli, A. Bracale, M.R. Mastronicola, A. Di Donato, G. D'Alessio, The antitumor action of seminal ribonuclease and its quaternary conformations, FEBS Lett 359 (1995) 31-34.
[197] M.P. Schlunegger, M.J. Bennett, D. Eisenberg, Oligomer formation by 3D domain swapping: a model for protein assembly and misassembly, Adv Protein Chem 50 (1997) 61-122.
[198] M.J. Bennett, M.P. Schlunegger, D. Eisenberg, 3D domain swapping: a mechanism for oligomer assembly, Protein Sci 4 (1995) 2455-2468.
[199] C. Qu, L. Liljas, N. Opalka, C. Brugidou, M. Yeager, R.N. Beachy, C.M. Fauquet, J.E. Johnson, T. Lin, 3D domain swapping modulates the stability of members of an icosahedral virus group, Structure 8 (2000) 1095-1103.
[200] Y. Liu, G. Gotte, M. Libonati, D. Eisenberg, A domain-swapped RNase A dimer with implications for amyloid formation, Nat Struct Biol 8 (2001) 211-214.
[201] M.J. Bennett, M.R. Sawaya, D. Eisenberg, Deposition diseases and 3D domain swapping, Structure 14 (2006) 811-824.
[202] D. Spalletti-Cernia, R. Sorrentino, S. Di Gaetano, R. Piccoli, M. Santoro, G. D'Alessio, P. Laccetti, G. Vecchio, Highly selective toxic and proapoptotic effects of two dimeric ribonucleases on thyroid cancer cells compared to the effects of doxorubicin, Br J Cancer 90 (2004) 270-277.
[203] M. Hursey, D.L. Newton, H.J. Hansen, D. Ruby, D.M. Goldenberg, S.M. Rybak, Specifically targeting the CD22 receptor of human B-cell lymphomas with RNA damaging agents: a new generation of therapeutics, Leuk Lymphoma 43 (2002) 953-959.
[204] A. Di Donato, V. Cafaro, I. Romeo, G. D'Alessio, Hints on the evolutionary design of a dimeric RNase with special bioactions, Protein Sci 4 (1995) 1470-1477.
[205] A. Merlino, C. Ercole, D. Picone, E. Pizzo, L. Mazzarella, F. Sica, The Buried Diversity of Bovine Seminal Ribonuclease:Shape and Cytotoxicity of the Swapped Noncovalent Form of the Enzyme, J Mol Biol 376 (2008) 427-437.
[206] C. Ercole, R.A. Colamarino, E. Pizzo, F. Fogolari, R. Spadaccini, D. Picone, Comparison of the structural and functional properties of RNase A and BS-RNase: a stepwise mutagenesis approach, Biopolymers 91 (2009) 1009-1017.

[207] J. Ross, K. Gray, D. Schenkein, B. Greene, G.S. Gray, J. Shulok, P.J. Worland, A. Celniker, M. Rolfe, Antibody-based therapeutics in oncology, Expert Rev Anticancer Ther 3 (2003) 107-121.
[208] I. Pastan, R. Hassan, D.J. FitzGerald, R.J. Kreitman, Immunotoxin treatment of cancer, Annu Rev Med 58 (2007) 221-237.
[209] C. De Lorenzo, G. D'Alessio, From immunotoxins to immunoRNases, Current pharmaceutical biotechnology 9 (2008) 210-214.
[210] S.M. Rybak, H.R. Hoogenboom, D.L. Newton, J.C. Raus, R.J. Youle, Rational immunotherapy with ribonuclease chimeras. An approach toward humanizing immunotoxins, Cell Biophys 21 (1992) 121-138.
[211] S.M. Rybak, M.A. Arndt, T. Schirrmann, S. Dubel, J. Krauss, Ribonucleases and immunoRNases as anticancer drugs, Curr Pharm Des 15 (2009) 2665-2675.
[212] C. De Lorenzo, A. Arciello, R. Cozzolino, D.B. Palmer, P. Laccetti, R. Piccoli, G. D'Alessio, A fully human antitumor immunoRNase selective for ErbB-2-positive carcinomas, Cancer Res 64 (2004) 4870-4874.
[213] C. De Lorenzo, C. Di Malta, G. Cali, F. Troise, L. Nitsch, G. D'Alessio, Intracellular route and mechanism of action of ERB-hRNase, a human anti-ErbB2 anticancer immunoagent, FEBS Lett 581 (2007) 296-300.
[214] D.J. Slamon, G.M. Clark, S.G. Wong, W.J. Levin, A. Ullrich, W.L. McGuire, Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene, Science 235 (1987) 177-182.
[215] J. Baselga, J. Albanell, Mechanism of action of anti-HER2 monoclonal antibodies, Ann Oncol 12 Suppl 1 (2001) S35-41.
[216] T. Yamamoto, S. Ikawa, T. Akiyama, K. Semba, N. Nomura, N. Miyajima, T. Saito, K. Toyoshima, Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor, Nature 319 (1986) 230-234.
[217] H.S. Cho, D.J. Leahy, Structure of the extracellular region of HER3 reveals an interdomain tether, Science 297 (2002) 1330-1333.
[218] H.S. Cho, K. Mason, K.X. Ramyar, A.M. Stanley, S.B. Gabelli, D.W. Denney, Jr., D.J. Leahy, Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab, Nature 421 (2003) 756-760.
[219] J. Baselga, D. Tripathy, J. Mendelsohn, S. Baughman, C.C. Benz, L. Dantis, N.T. Sklarin, A.D. Seidman, C.A. Hudis, J. Moore, P.P. Rosen, T. Twaddell, I.C. Henderson, L. Norton, Phase II study of weekly intravenous trastuzumab (Herceptin) in patients with HER2/neu-overexpressing metastatic breast cancer, Semin Oncol 26 (1999) 78-83.
[220] J. Stebbing, E. Copson, S. O'Reilly, Herceptin (trastuzamab) in advanced breast cancer, Cancer Treat Rev 26 (2000) 287-290.
[221] A. Badache, N.E. Hynes, A new therapeutic antibody masks ErbB2 to its partners, Cancer cell 5 (2004) 299-301.
[222] C. De Lorenzo, R. Cozzolino, A. Carpentieri, P. Pucci, P. Laccetti, G. D'Alessio, Biological properties of a human compact anti-ErbB2 antibody, Carcinogenesis 26 (2005) 1890-1895.
[223] C. De Lorenzo, D.B. Palmer, R. Piccoli, M.A. Ritter, G. D'Alessio, A new human antitumor immunoreagent specific for ErbB2, Clin Cancer Res 8 (2002) 1710-1719.
[224] C. De Lorenzo, A. Tedesco, G. Terrazzano, R. Cozzolino, P. Laccetti, R. Piccoli, G. D'Alessio, A human, compact, fully functional anti-ErbB2 antibody as a novel antitumour agent, Br J Cancer 91 (2004) 1200-1204.
[225] G. Riccio, G. Esposito, E. Leoncini, R. Contu, G. Condorelli, M. Chiariello, P. Laccetti, S. Hrelia, G. D'Alessio, C. De Lorenzo, Cardiotoxic effects, or lack thereof, of anti-ErbB2 immunoagents, Faseb J 23 (2009) 3171-3178.
[226] R. Berisio, V.S. Lamzin, F. Sica, K.S. Wilson, A. Zagari, L. Mazzarella, Protein titration in the crystal state, Journal of molecular biology 292 (1999) 845-854.
[227] D.E. Holloway, G.B. Chavali, M.C. Hares, V. Subramanian, K.R. Acharya, Structure of murine

angiogenin: features of the substrate- and cell-binding regions and prospects for inhibitor-binding studies, Acta crystallographica 61 (2005) 1568-1578.
[228] D.D. Leonidas, G.B. Chavali, A.M. Jardine, S. Li, R. Shapiro, K.R. Acharya, Binding of phosphate and pyrophosphate ions at the active site of human angiogenin as revealed by X-ray crystallography, Protein Sci 10 (2001) 1669-1676.
[229] A. Merlino, L. Mazzarella, A. Carannante, A. Di Fiore, A. Di Donato, E. Notomista, F. Sica, The importance of dynamic effects on the enzyme activity: X-ray structure and molecular dynamics of onconase mutants, The Journal of biological chemistry 280 (2005) 17953-17960.
[230] A. Merlino, L. Vitagliano, F. Sica, A. Zagari, L. Mazzarella, Population shift vs induced fit: the case of bovine seminal ribonuclease swapping dimer, Biopolymers 73 (2004) 689-695.
[231] K.R. Acharya, R. Shapiro, J.F. Riordan, B.L. Vallee, Crystal structure of bovine angiogenin at 1.5-A resolution, Proc Natl Acad Sci U S A 92 (1995) 2949-2953.
[232] L.W. Schultz, D.J. Quirk, R.T. Raines, His...Asp catalytic dyad of ribonuclease A: structure and function of the wild-type, D121N, and D121A enzymes, Biochemistry 37 (1998) 8886-8898.
[233] R. Berisio, F. Sica, V.S. Lamzin, K.S. Wilson, A. Zagari, L. Mazzarella, Atomic resolution structures of ribonuclease A at six pH values, Acta Crystallogr D Biol Crystallogr 58 (2002) 441-450.
[234] K.S. Siddiqui, R. Cavicchioli, Cold-adapted enzymes, Annu Rev Biochem 75 (2006) 403-433.
[235] A. Merlino, I. Russo Krauss, I. Castellano, E. De Vendittis, B. Rossi, M. Conte, A. Vergara, F. Sica, Structure and flexibility in cold-adapted iron superoxide dismutases: the case of the enzyme isolated from *Pseudoalteromonas haloplanktis*, Journal of Structural Biology 172 (2010) 343-352.
[236] V. Cafaro, A. Bracale, A. Di Maro, S. Sorrentino, G. D'Alessio, A. Di Donato, New muteins of RNase A with enhanced antitumor action, FEBS Lett 437 (1998) 149-152.
[237] C.F. Aguilar, P.J. Thomas, A. Mills, D.S. Moss, R.A. Palmer, Newly observed binding mode in pancreatic ribonuclease, J Mol Biol 224 (1992) 265-267.
[238] C.F. Aguilar, P.J. Thomas, D.S. Moss, A. Mills, R.A. Palmer, Novel non-productively bound ribonuclease inhibitor complexes--high resolution X-ray refinement studies on the binding of RNase-A to cytidylyl-2',5'-guanosine (2',5'CpG) and deoxycytidylyl-3',5'-guanosine (3',5'dCpdG), Biochim Biophys Acta 1118 (1991) 6-20.
[239] A. Canals, J. Pous, A. Guasch, A. Benito, M. Ribo, M. Vilanova, M. Coll, The structure of an engineered domain-swapped ribonuclease dimer and its implications for the evolution of proteins toward oligomerization, Structure (Camb) 9 (2001) 967-976.
[240] Y. Liu, P.J. Hart, M.P. Schlunegger, D. Eisenberg, The crystal structure of a 3D domain-swapped dimer of RNase A at a 2.1-A resolution, Proc Natl Acad Sci U S A 95 (1998) 3437-3442.
[241] C. Chothia, A.M. Lesk, Canonical structures for the hypervariable regions of immunoglobulins, Journal of molecular biology 196 (1987) 901-917.
[242] C. Chothia, A.M. Lesk, A. Tramontano, M. Levitt, S.J. Smith-Gill, G. Air, S. Sheriff, E.A. Padlan, D. Davies, W.R. Tulip, et al., Conformations of immunoglobulin hypervariable regions, Nature 342 (1989) 877-883.
[243] M.C. Franklin, K.D. Carey, F.F. Vajdos, D.J. Leahy, A.M. de Vos, M.X. Sliwkowski, Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex, Cancer cell 5 (2004) 317-328.
[244] C. Ercole, R. Colamarino, E. Pizzo, F. Fogolari, R. Spadaccini, D. Picone, Comparison of the structural and functional properties of RNase A and BS-RNase: a stepwise mutagenesis approach, submitted for publication to the same issue of Biopolymers (2009).
[245] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: An automated protein homology-modeling server, Nucleic acids research 31 (2003) 3381-3385.
[246] A. Merlino, G. Graziano, L. Mazzarella, Structural and dynamic effects of alpha-helix deletion in Sso7d: implications for protein thermal stability, Proteins 57 (2004) 692-701.

[247] M. Porcelli, M.A. Moretti, L. Concilio, S. Forte, A. Merlino, G. Graziano, G. Cacciapuoti, S-adenosylhomocysteine hydrolase from the archaeon Pyrococcus furiosus: biochemical characterization and analysis of protein structure by comparative molecular modeling, Proteins 58 (2005) 815-825.
[248] B. Al-Lazikani, A.M. Lesk, C. Chothia, Standard conformations for the canonical structures of immunoglobulins, Journal of molecular biology 273 (1997) 927-948.
[249] J. Ay, T. Keitel, G. Kuttner, H. Wessner, C. Scholz, M. Hahn, W. Hohne, Crystal structure of a phage library-derived single-chain Fv fragment complexed with turkey egg-white lysozyme at 2.0 A resolution, Journal of molecular biology 301 (2000) 239-246.
[250] M.J. Sippl, Recognition of errors in three-dimensional structures of proteins, Proteins 17 (1993) 355-362.
[251] A.D. van Dijk, R. Boelens, A.M. Bonvin, Data-driven docking for the study of biomolecular complexes, The FEBS journal 272 (2005) 293-312.
[252] H.A. Gabb, R.M. Jackson, M.J. Sternberg, Modelling protein docking using shape complementarity, electrostatics and biochemical information, Journal of molecular biology 272 (1997) 106-120.
[253] D. van der Spoel, R. van Druner, H.J.C. Berendsen, GROningen MAchine for CHemical Simulation, Department of Biophysical Chemistry. Bioson Research Institute, Groningen, 1994.
[254] A. Merlino, L. Esposito, L. Vitagliano, Polyglutamine repeats and beta-helix structure: molecular dynamics study, Proteins 63 (2006) 918-927.
[255] A. Merlino, S. Varriale, M.R. Coscia, L. Mazzarella, U. Oreste, Structure and dimerization of the teleost transmembrane immunoglobulin region, J. Mol. Graph. Model 27 (2008) 401-407.
[256] T. Darden, D. York, L. Pedersen, Part Mesh Ewald: an log (N) method for Ewald sums in large systems., J Chem Phys 98 (1993) 10089-10092.
[257] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Freaje, LINCS: a linear constraint solver for molecular simulations, J Comp Chem 18 (1997) 1463-1472.
[258] P.J. Kraulis, MOLSCRIPT:a program to produce both detailed and schematic plots of protein structures., J. Appl. Crystallogr. 24 (1991) 946-950.

PUBLICATIONS

- A. Merlino, **I. Russo Krauss**, I. Castellano, E. De Vendittis, A. Vergara, F. Sica

Crystallization and preliminary X-ray diffraction studies of a psychrophilic iron superoxide dismutase from *Pseudoalteromonas haloplanktis*

Protein & Peptide Letters, 15 (4): 415-418 (2008)

Abstract

The Antarctic eubacterium *Pseudoalteromonas haloplanktis* (*Ph*) produces a cold-active iron superoxide dismutase (SOD). *Ph*SOD is a homodimeric enzyme, that displays a high catalytic activity even at low temperature. Using hanging-drop vapour-diffusion technique, *Ph*SOD has been successfully crystallized in two different crystal forms. Both crystal forms are monoclinic with space group P21 and diffract to 2.1 Å resolution. Form I has unit-cell parameters a=45.49Å b=103.63Å c=50.37Å β=108.2° and contains a homodimer in the asymmetric unit. Form II has unit-cell parameters a=50.48Å b=103.78Å c=90.25Å β=103.8° and an asymmetric unit containing two *Ph*SOD homodimers. Structure determination has been achieved using molecular replacement. The crystallographic study of this cold-adapted enzyme could contribute to the understanding of the molecular mechanisms of cold-adaptation and of the high catalytic efficiency at low temperature.

- A. Merlino**, I. Russo Krauss**, M. Perillo, C.A. Mattia, C. Ercole, D. Picone, A. Vergara, F. Sica

Towards an antitumor form of bovine pancreatic ribonuclease: the crystal structure of three non-covalent dimeric mutants

Biopolymers 91(12): 1029-37 (2009)

Abstract

The cytotoxic action of bovine seminal ribonuclease (BS-RNase) depends on its non-covalent swapped dimeric form (NCD-BS), which presents a compact structure that allows the molecule to escape ribonuclease inhibitor (RI). A key role in the acquisition of this structure has been attributed to the concomitant presence of a proline in position 19 and a leucine in position 28. The introduction of Leu28, Cys31 and Cys32 and, in addition, of Pro19 in the sequence of bovine pancreatic ribonuclease (RNase A) has produced two dimeric variants LCC and PLCC, which do exhibit a cytotoxic activity, though at a much lower level than BS-RNase. The crystal structure analysis of the non-covalent swapped form (NCD) of LCC and PLCC, complexed with the substrate analogue 2'-deoxycytidylyl(3',5')-2'-deoxyguanosine, has revealed that, differently from NCD-BS, the dimers adopt

an opened quaternary structure, with the two Leu residues fully exposed to the solvent, that does not hinder the binding of RI. Similar results have been obtained for a third mutant of the pancreatic enzyme, engineered with the hinge peptide sequence of the seminal enzyme (residues 16-22) and the two cysteines in position 31 and 32, but lacking the hydrophobic Leu residue in position 28. The comparison of these three structures with those previously reported for other ribonuclease swapped dimers strongly suggests that, in addition to Pro19 and Leu28, the presence of a glycine at the N-terminal end of the hinge peptide is also important to push the swapped form of RNase A dimer into the compact quaternary organization observed for NCD-BS.

- **I. Russo Krauss,** A. Merlino, A. Randazzo, L. Mazzarella, F. Sica

Crystallization and preliminary X-ray analysis of the complex of human alpha-thrombin with a modified thrombin binding aptamer

Acta Crystallogr Sect F Struct Biol Cryst Commun. 66(Pt 8): 961-3 (2010)

Abstract

The thrombin-binding aptamer (TBA) is a consensus DNA 15-mer that binds specifically to human α-thrombin at nanomolar concentrations and inhibits its procoagulant functions. Recently, a modified TBA (mTBA) containing a 5'-5' inversion-of-polarity site has been shown to be more stable and to possess a higher thrombin affinity than its unmodified counterpart. The structure of the thrombin-TBA complex has previously been determined at low resolution, but did not provide a detailed picture of the aptamer conformation or of the protein-DNA assembly, while that of the complex with mTBA is unknown. Crystallographic analysis of the thrombin-mTBA complex has been attempted. The crystals diffracted to 2.15 Å resolution and belonged to space group I222.

- A. Merlino, **I. Russo Krauss**, I. Castellano, E. De Vendittis, B. Rossi, M. Conte, A. Vergara, F. Sica

Structure and flexibility in cold-adapted iron superoxide dismutases: the case of the enzyme isolated from *Pseudoalteromonas haloplanktis*

Journal of Structural Biology 172: 343-352 (2010)

Abstract

Superoxide dismutases (SODs) are metalloenzymes catalysing the dismutation of superoxide anion radicals into molecular oxygen and hydrogen peroxide. Here, we present the crystal structure of a cold-adapted Fe-SOD from the Antarctic eubacterium Pseudoalteromonas haloplanktis (PhSOD), and that

of its complex with sodium azide. The structures were compared with those of the corresponding homologues having a high sequence identity with PhSOD, such as the mesophilic SOD from Escherichia coli (EcSOD) or Pseudomonas ovalis, and the psychrophilic SOD from Aliivibrio salmonicida (AsSOD). These enzymes shared a large structural similarity, such as a conserved tertiary structure and arrangement of the two monomers, an almost identical total number of inter- and intramolecular hydrogen bonds and salt bridges. However, the two cold-adapted SODs showed an increased flexibility of the active site residues with respect to their mesophilic homologues. Structural information was combined with a characterization of the chemical and thermal stability performed by CD and fluorescence measurements. Despite of its psychrophilic origin, the denaturation temperature of PhSOD was comparable with that of the mesophilic EcSOD, whereas AsSOD showed a lower denaturation temperature. On the contrary, the values of the denaturant concentration at the transition midpoint were in line with the psychrophilic/ mesophilic origin of the proteins.

These data provide additional support to the hypothesis that cold-adapted enzymes achieve efficient catalysis at low temperature, by increasing the flexibility of their active site; moreover, our results underline how fine structural modifications can alter enzyme flexibility and/or stability without compromising the overall structure of typical rigid enzymes, such as SODs.

- E. Pizzo, A. Merlino, M. Turano, **I. Russo Krauss**, F. Coscia, A. Zanfardino, M. Varcamonti, A. Furia, C. Giancola, L. Mazzarella, F. Sica, G. D'Alessio,

A new RNase sheds light on the RNase/angiogenin subfamily from zebrafish

Biochem J. (2010) DOI: 10.1042/BJ20100892

Abstract

Only recently extracellular RNases of the RNase A superfamily, with the characteristic CKxxNTF sequence signature, have been identified in fish. This has led to the recognition that these RNases are present in the whole vertebrate sub-phylum. In fact, they comprise the only enzyme family unique to vertebrates. Four RNases from zebrafish (*Danio rerio*) have been previously reported, with a very low RNase activity, some endowed, like human angiogenin, with powerful angiogenic and bactericidal activities. Here we report the three-dimensional structure, the thermodynamic behaviour, and the biological properties of a novel zebrafish RNase, ZF-RNase-5. The investigation of its structural and functional properties, extended to all other sub-family members.

- F. Troise, M. Monti, A. Merlino, F. Cozzolino, C. Fedele, **I. Russo Krauss**, F. Sica, P. Pucci, G. D'Alessio, C. De Lorenzo

A novel ErbB2 epitope targeted by human antitumor immunoagents *Submitted*

Abstract

Two novel human antitumor immunoconjugates, engineered by fusion of a human anti-ErbB2 scFv, termed Erbicin, with either a human RNase or the Fc region of a human IgG1, are selectively cytotoxic for ErbB2-positive cancer cells in vitro and in vivo and do not show the most negative properties of Herceptin, the only humanized anti-ErbB2 mAb used in the therapy of breast carcinoma: cardiotoxicity and inability to act on resistant tumors.

These differences are likely due to the different ErbB2 epitopes recognized by EDIA and Herceptin as it was reported that they induce different signaling mechanisms which control tumor and cardiac cell viability.

Thus, to accurately identify the novel epitope recognized by EDIA, which could also become a potential therapeutic target to mitigate anti-ErbB2-associated toxicity and eventually overcome resistance, three independent methodologies were used which gave coherent results, here reported: EDIA bind to a different ErbB2 epitope, if compared to Herceptin and the other human/humanized anti-ErbB2 antibodies reported so far, and this epitope is located in the region 122-195 of extracellular domain I of ErbB2.

- **I. Russo Krauss,** A. Merlino, C. Giancola, A. Randazzo, L. Mazzarella, F. Sica

A revealed ambiguity: the crystal structure of α-thrombin in complex with modified thrombin binding aptamer *In preparation*

- A. Merlino, **I. Russo Krauss**, I. Castellano, E. De Vendittis, A. Vergara, F. Sica,

Structural and denaturation studies of two mutants of a cold adapted superoxide dismutase point to the importance of electrostatic interactions in protein stability *In preparation*

- A. Merlino, **I. Russo Krauss**, I. Castellano, E. De Vendittis, B. Rossi, A. Vergara, F. Sica

Spectroscopic analysis of temperature and guanidinium chloride-induced unfolding of the superoxide dismutase from the denthal pathogenic microorganism *Streptococcus mutans* supports a three state mechanism *In preparation*

Structures published in Protein Data Bank

PDB code

3FKZ X-ray structure of the non covalent swapped form of the S16G/T17N/A19P/A20S/K31C/S32C mutant of bovine pancreatic ribonuclease
A. Merlino, **I. Russo Krauss**, M. Perillo, C. Ercole, D. Picone, A. Vergara, F. Sica.
(24-03-2009) DOI: 10.2210/pdb3fkz/pdb

3FL0 X-ray structure of the non covalent swapped form of the Q28L/K31C/S32C mutant of bovine pancreatic ribonuclease in complex with 2'-DEOXYCYTIDINE--2'-DEOXYGUANOSINE-3',5'-MONOPHOSPHATE
A. Merlino, **I. Russo Krauss**, M. Perillo, C. Ercole, D. Picone, A. Vergara, F. Sica.
(24-03-2009) DOI: 10.2210/pdb3fl0/pdb

3FL1 X-ray structure of the non covalent swapped form of the A19P/Q28L/K31C/S32C mutant of bovine pancreatic ribonuclease in complex with 2'-DEOXYCYTIDINE--2'-DEOXYGUANOSINE-3',5'-MONOPHOSPHATE
A. Merlino, **I. Russo Krauss**, M. Perillo, C. Ercole, D. Picone, A. Vergara, F. Sica.
(24-03-2009) DOI: 10.2210/pdb3fl1/pdb

3FL3 X-ray structure of the ligand free non covalent swapped form of the A19P/Q28L/K31C/S32C mutant of bovine pancreatic ribonuclease
A. Merlino, **I. Russo Krauss**, M. Perillo, C. Ercole, D. Picone, A. Vergara, F. Sica.
(24-03-2009) DOI: 10.2210/pdb3fl3/pdb

3LIO The X-ray structure of iron superoxide dismutase from Pseudoalteromonas haloplanktis (crystal form I)
A. Merlino, **I. Russo Krauss**, B. Rossi, M. Conte, A. Vergara, F. Sica
(08-09-2010) DOI: 10.2210/pdb3lio/pdb

3LJF The X-ray structure of iron superoxide dismutase from Pseudoalteromonas haloplanktis (crystal form II)
A. Merlino, **I. Russo Krauss**, B. Rossi, M. Conte, A. Vergara, F. Sica
(08-09-2010) DOI: 10.2210/pdb3ljf/pdb

3LJ9 The X-ray structure of iron superoxide dismutase from Pseudoalteromonas haloplanktis in complex with sodium azide
A. Merlino, **I. Russo Krauss**, B. Rossi, M. Conte, A. Vergara, F. Sica
(08-09-2010) DOI: 10.2210/pdb3lj9/pdb

3LJE The X-ray structure of zebrafish RNase5
I. Russo Krauss, A. Merlino, F. Coscia, L. Mazzarella, F. Sica
(24-11-2010) DOI: 10.2210/pdb3lje/pdb

3LJD The X-ray structure of zebrafish RNase1 from a new crystal form at pH 4.5

I. Russo Krauss, A. Merlino, L. Mazzarella, F. Sica

DOI: 10.2210/pdb3ljd/pdb

3LN8 The X-ray structure of zebrafish RNase1 from a new crystal form at pH 7.3

I. Russo Krauss, A. Merlino, L. Mazzarella, F. Sica

DOI: 10.2210/pdb3ln8/pdb