Structural biochemistry of the pentraxins

Noone, D.P.

9 General Discussions

Dylan P. Noone[1] and Thomas H. Sharp[1,*]

[1]Department of Cell and Chemical Biology, Leiden University Medical Center, 2300 RC Leiden, The Netherlands

*To whom correspondence should be addressed: t.sharp@lumc.nl

9.1 Targeting the innate immune system

Throughout this thesis we have explored the structural and biochemical behavior of the pentraxins, to understand their structure-function relationships, with the aim of exploiting them for therapeutic benefit. Many treatments utilize the adaptive immune system such as mRNA vaccines(278), monoclonal antibodies (279) and CAR-T cells(280). However, the innate immune system is relatively underutilized. Moreover, the innate immune system appears to be more complex than previously thought, with innate immune cells having a limited immunological memory of events(6). For example, vaccination of children with bacillus Calmette-Guerin (BCG) provides protection against a range of other respiratory pathogens and sepsis in addition to *M. tuberculosis*, which was ascribed to innate immune mechanisms (281). This indicates that innate immune therapeutics may provide broader, though less specific, effects that could complement more conventional therapies such as monoclonal antibodies. Whilst this would not provide a specific induction of a response to a particular epitope like a vaccine, this could provide a general immune boost, which may prove useful at the early stages of a pandemic prior to vaccine development or in tandem with conventional treatments. Indeed, blockage of C5a signaling has been used concurrently with immune checkpoint therapies to such an effect against some cancers (282). Additionally, given the link between innate immune dysregulation and pathologies such as COVID-19(283) Alzheimer's disease (AD)(60) and autoimmune disorders such as systemic lupus erythematosus (SLE)(64), exploiting this large portion of the immune system may lead to novel and efficacious therapies.

9.2 A case study of pentraxins as a therapy: CRP in autoimmune disorders

SLE is an autoimmune disorder characterized by the immune system targeting healthy tissues and organs such as the skin, joints and the kidneys(284). SLE has been postulated to stem from the defective clearance of apoptotic cells, resulting in the presentation of autoantigens to the immune system(58). In support of this hypothesis, SLE patients often have macrophages and neutrophils that have impaired phagocytic activity(285, 286), with polymorphisms in phagocytic receptors resulting in an increased risk of SLE(287). Additionally, SLE patients often have deficiencies in the early components of the classical complement cascade such as C1q(65), which is itself an opsonin(20), that could prevent the formation of "eat-me" and "find-me" signals such as deposited C3b and soluble C3a. The correct disposal of dying cells is postulated to prevent dendritic cells being exposed to autoantigens, which could then present these autoantigens to autoreactive B or T -cells that stochastically arise(288). Interestingly, PTX3 has been reported to sequester autoantigens away from dendritic cells, suggesting pentraxins may play a crucial role in the silent clearance of cells to prevent autoreactivity in the immune system(61, 234). In support of this, auto antibodies against CRP and Serum amyloid P component (SAP) are found in SLE patients(289), with CRP having a documented role in phagocytosis(144-146).

Taken together, the evidence illustrates that CRP plays a scavenger role, mopping up autoantigens from apoptotic cells to facilitate their phagocytosis, via direct interactions with

Fcγ receptors(22, 24) and indirectly via deposited complement factors such as C3b(20). This model suggests that CRP could be used as a therapeutic for SLE and autoimmune disorders more generally, as more CRP could result in the more efficient phagocytosis of apoptotic cells. Indeed, intravenous injection of CRP showed therapeutic potential in a mouse model of SLE(66), which was shown to be mediated via interactions with macrophages and interleukin-10, with both complement and Fcγ mediated phagocytosis playing a role(126). In fact, this was shown to be beneficial if applied prior to disease onset but also during the acute phase of the disease. However, patients with SLE also show aberrant complement activation (67, 68), suggesting complement can also have a negative impact on autoimmune disorders. CRP can activate complement (***chapter 3***)(30) and so may cause unwanted side effects if left unchecked. In mouse models of arthritis, lowering serum CRP levels improved symptoms(187), though a later clinical trial led to inconclusive results(188). We suggest that these conflicting findings may represent differences in the biasing of CRP towards either complement activation or phagocytosis, with membrane attach complex (MAC) mediated lysis worsening autoimmune conditions via increased exposure of autoantigens and causing greater inflammation.

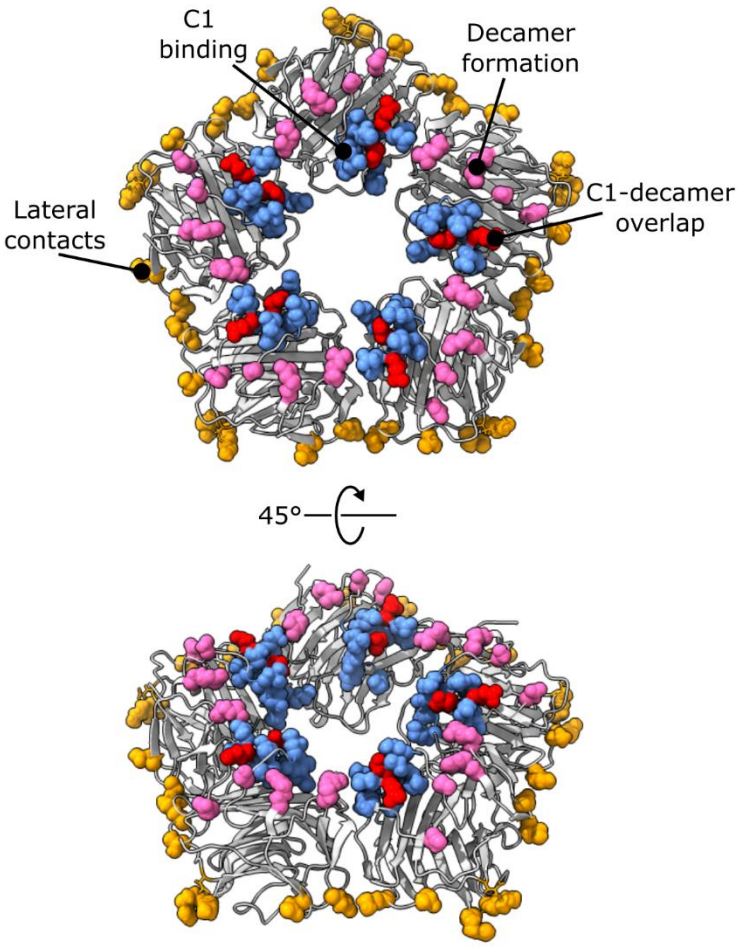

Figure 1: Hot spots for manipulation of CRP function. *A-face view of CRP with residues mediating C1 binding residues (P168-G177 and D112-K114), decamer formation (R6, D163, D169, E170, N172, S181 and R188) and lateral contacts (K57, R58, T126, E130, K123, Q195 and E197) in blue, pink and orange respectively. Residues that are linked to both decamer formation and C1 binding are shown in red.*

To remedy this, the next generation of CRP based therapies should be more precise, by biasing CRP towards particular beneficial effector functions and away from other deleterious processes. In an autoimmune disease context this may involve the inhibition of CRP based complement activation, whilst maintaining potentially helpful pathways such as phagocytosis. In fact, all therapies targeting CRP thus far have either targeted the B-face of CRP (123, 186, 290) or induced reduced serum CRP concentrations(187-190), therefore inhibiting all CRP effector functions. We propose the rational design of recombinant CRP or the development of molecules such as nanobodies that could be intravenously administered (***chapter 5***), to bias CRP to

or away from phagocytosis or complement activation. This could allow the enhancement of phagocytosis and clearance of apoptotic cells in SLE, whilst preventing excessive complement activation that could worsen the disease.

A mutant version of CRP that can selectively activate complement whilst not agglutinating lipids has already been presented in this thesis (***chapter 3***). However, CRP variants that can form decamers but not activate C1 could also be produced. These variants should focus on two hotspots of CRP (**Fig. 1**); the putative C1 binding site P168-G177 and D112-K114 (***chapter 3***)(90-92), or lateral contact points between CRP pentamers associating on a surface at K57 and R58(92), as well as T126, E130, K123, Q195 and E197 (***chapter 3***)(87). The first hotspot would inhibit interactions with the C1 complex directly, though care must be taken given certain residues overlap with the inter-pentamer interface in decameric CRP such as D169, E170 and N172 (**Fig. 1, *chapter 1***)(35). However, the crystal structure of the short pentraxin SAP complexed with an Fcγ receptor revealed that Fcγ binds to homologous regions to those involved in C1 binding in CRP (P168-G177)(24). As such, any mutants targeting this hotspot should be extensively tested for their ability to agglutinate liposomes together (***chapter 5***), but also their ability to bind to Fcγ receptors and stimulate phagocytosis in a similar manner as has been reported for SAP(24).

The second hotspot containing the lateral regions, could prevent formation of higher order oligomers required for C1 binding (***chapter 3***) and would have the benefit of not interfering with decamer formation or Fcγ receptor binding. This could be achieved via two routes; the elimination of ionic bonds or electrostatic repulsion between CRP pentamers. Simply mutating residues at these lateral contacts to alanine, could prevent salt bridges that are needed to form higher order oligomers prior to C1 binding (***chapter 3***). However, a stronger effect may be achieved via mutating these charged residues to have the opposite charge, e.g. mutating D to K. This would change pre-existing salt bridges to contain amino acids of the same charge, repelling neighbouring CRP pentamers from each other on a lipid membrane. Nevertheless, given the large size of C1 and flexibility in its stalk region(16, 291), C1 may be able to bridge any gaps between CRP pentamers this may cause. In this case, bulkier molecules such as monoclonal antibodies that create a larger physical distance between CRP pentamers may be required.

Using the pipeline presented in ***chapter 5*** would allow a proof of concept of these ideas, before scaling up to phagocytic and complement mediated cell killing assays and ultimately testing in mouse models. Indeed, direct comparisons of complement and phagocytic deficient CRP mutants may reveal fundamental relationships between the two processes and the onset of autoimmune diseases such as SLE.

9.3 CRP and the inflammatory microenvironment

Perhaps the most underdeveloped part of this thesis, though potentially most relevant to the clinic, were the chapters on the characteristics of CRP in an acidic microenvironment. The inflammatory microenvironment has been shown to be acidic(75, 77, 80, 81), with complement activation(181) and leukocyte metabolism(80) causing localized acidification.

Localized acidosis is not a new idea in biology, as the endocytic system comprises a series of gradually acidifying membranous compartments(292). Furthermore, extracellular acidosis is also a recognized feature of the cancer cell niche(77, 293). CRP has been shown to have different properties at an acidic pH. For instance, CRP can interact with gC1q preferentially at an acidic pH(85, 136) (***chapter 2***) and has been shown to induce more complement activation(30). CRP has been linked to atherosclerosis(171), and in this thesis it has been indicated that CRP binds to apolipoproteins in a pH dependent manner (***chapter 4***). CRP has also been shown to bind to cholesterol crystals(294), crystalline cholesterol structures characteristic of atherosclerosis, which are known to activate the complement system to stimulate inflammation(295, 296). How this interaction is affected by acidosis is unknown, though it has been observed that acidosis and cholesterol crystals have a synergistic effect on inflammasome activation in macrophages(297). More data on CRP interactions with cholesterol crystals and macrophages under acidic conditions may help inform if this enhancement of CRP at an acidic pH is physiologically relevant. Furthermore, a cryoET derived structure of the C1-CRP-cholesterol crystal complex would complement the C1-CRP-phosphocholine structures shown in ***chapter 3*** of this thesis. The C1-CRP-choelsterol crystal complex is linked to a pathological disease setting rather than homeostatic clearance of cellular debris and represents misregulation of the innate immune response opposed to its normal function. Whilst CRP mediated complement activation often results in minimal MAC deposition(27, 59), C1-CRP bound to cholesterol crystals appears to result in colocalization with the terminal pathway(294), suggestive of full complement activation. Comparisons between the two may therefore shed light on structural factors, in addition to the binding of regulatory proteins such as factor H, that affect CRP based complement activation. This may also have direct relevance to the clinic given that extracellular acidosis also appears to lead increase macrophage mediated phagocytosis(298), to defective cholesterol metabolism in macrophages(299) and induce inflammasome activation(297), providing a possible mechanism for foam cells formation after phagocytosis of opsonized cholesterol containing complexes, such as C1-CRP coated cholesterol crystals. However, the milieu of atherosclerotic plaques are reported to have a pH of 6.0-6.5(298), 1.0-1.5 pH units higher than those used in ***chapters 2*** and ***4***. As such, repeating these experiments at pH 6.0-6.5 will validate the biological relevance of the acidic behavior of CRP and discount the possibility that these results are simply an artefact of such a low pH. Nonetheless, hints at the continued relevance of the pH dependent behaviour of CRPs interactions with complement at pH 6.0-6.5(30, 136) (***chapters 2 and 4***) are present in the literature. For example, the optimal pH for CRP mediated complement activation has been shown to be pH 6.3(30), putting it right in the middle of the pH range found in atherosclerotic plaques(298). In conclusion whilst the data presented in this thesis is underdeveloped and should be interpreted with caution, future studies focusing on a more physiological pH range and the effect of acidosis on CRP interactions with cholesterol crystals coupled with an accompanying cryoET derived structures would bolster our understanding of this phenomenon.

9.4 The long pentraxins

The long pentraxins such as PTX3 are linked to medically relevant processes such as viral infections(51, 201, 202, 215, 245) and autoimmunity (61, 234). However, the long pentraxin field is not as developed as the body of literature around the short pentraxins. For example, the first long pentraxin domain was resolved in 2020 (38) and the first complete long pentraxin structure in 2022 (***chapter 6***)(36), compared to the first structures of CRP being published in the 1990s(34, 86). As such, in order to unlock the potential of these proteins as therapies more basic research should be carried out to understand their structural and biochemical mechanisms of action.

Long pentraxins have an extra N-terminal domain(39), which enables a different protein architecture in PTX3 compared to the short pentraxins(36). In PTX3 the N-terminal region was revealed to be a tetrameric coiled coil, which was stapled together via a network of disulfide bonds (***chapter 6***)(36, 40), supporting previous biochemical studies(41). This is also reported to be the case for the long pentraxins NPTX1 and NPTX2, where N-terminal protein-protein interactions and disulfide bonding between NPTX monomers is shown to mediate oligomerisation(42). Thus it would appear that N-terminal domains regulate the oligomeric states of the long pentraxins via the formation of coiled coil domains and are stabilized by disulfide bonding.

Coiled coils consist of amphipathic helices that contain a heptad repeat (***abcdefg***) characterized by hydrophobic residues at the ***a*** and ***d*** positions(300), that cause the helices to wrap around each other forming a supercoil. The residues at ***a*** and ***d*** form a central hydrophobic seam with side chains packing between helices with characteristic knobs into holes interactions (**Fig. 2A**)(300). Additionally, residues at the ***e*** and ***g*** positions on opposing alpha helices may also participate in ionic or hydrogen bonding further stabilizing the quaternary structure of the coiled coil(**Fig. 2A**)(300). Coiled coils can produce a range of oligomeric states, such as trimers and tetramers (**Fig. 2A&B**), which can be predicted following established structure to function relationships between the nature of the heptad repeat and the quaternary structure (301). Both trimers and tetramers would form useful structures in terms of complement and pentraxins, as tetramers could form PTX3 like scaffolding proteins, whereas trimers could adopt a similar geometry to C1q that is well suited to the hexameric arrays seen in complement activators such as IgG1(8), IgM(9) and IgG3(138) (**Fig. 2C**).

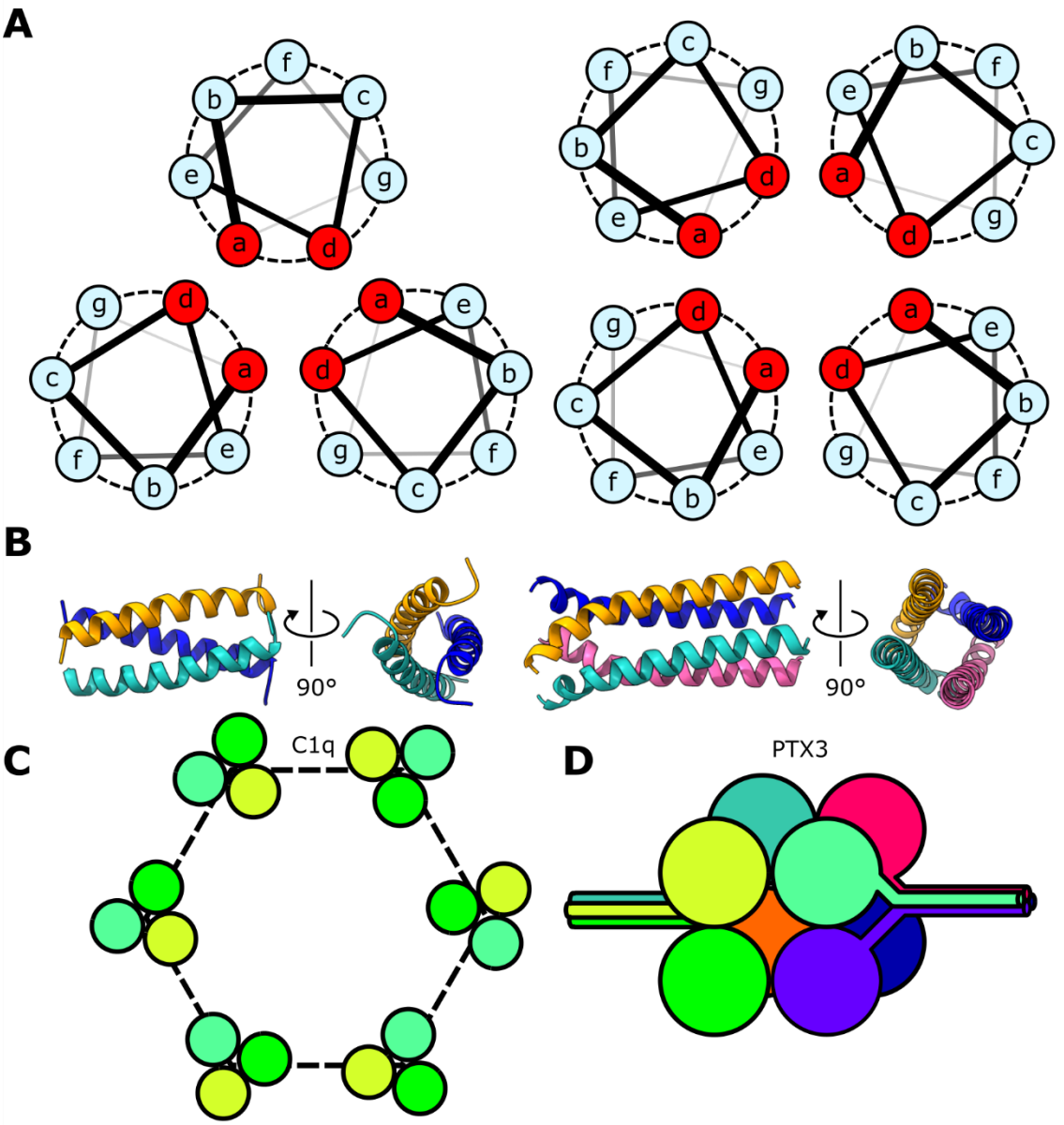

Figure 2: Schematics of trimeric and tetrameric coiled coil assemblies. *(**A**) Helical-wheel diagrams of trimer and tetrameric coiled coil. Hydrophobic residues at **a** and **d** form the hydrophobic core of the coil, which is complemented by ionic or hydrogen bonds between opposing residues at **e** and **g**. (**B**) Example trimeric and tetrameric coiled coils from the PDB (trimer PDB code; 1COI , tetramer PDB code; 7ZL1). (**C**) Schematic of the trimeric globular head domains of C1q forming hexameric arrays for binding to antibody Fc regions during complement activation. (**D**) Schematic of PTX3 showing that it forms a dimer of tetramers.*

LOGICOIL is an online tool that can be used to predict the presence of coiled-coil domains and their most probable and second most probable oligomeric states(207), and has been utilized in this theses previously (***chapter 6 & 8***). The sequences of each long pentraxin were submitted to LOGICOIL, producing a series of predicted coiled coil domains in the N-terminal for each long pentraxin (**Table 1**). PTX3 was correctly predicted as a tetrameric coiled coil, matching the literature(36, 41). PTX4(302) was also predicted as a tetramer similar to PTX3. However, little is known about PTX4 apart from sparse reports showing it is present in the gingival tissues of non-human primates(303) and is overexpressed, along with PTX3, NPTX1

and NPTX2, during brain trauma(304). PTX4 also appears to be prevalent in the lymphoid tissues such as the spleen and lymph nodes(305), though further work needs to be done to help reveal the structure and role of PTX4 in the body.

The neuronal pentraxins (NPTXs) were discussed in detail previously, so will not be covered further in great detail (***chapter 8***). However, it is noteworthy that NPTX2 is reported to be better at cluster AMPA receptors (AMPAR), a characteristic attributed to the N-terminal region of NPTX2(42). This is also the only NPTX to contain a region of tetrameric nature in the N-terminal coiled coil domain (**Table 1**). This may indicate that the enhanced ability of NPTX2 to cluster AMPAR may be mediated by the oligomeric state of the N-terminal region i.e., more pentraxins present to bind and cluster AMPAR. As the oligomeric state of coiled-coils follow predictable rules based on their primary sequence(300, 301), this raises the intriguing possibility that synthetic NPTX derivatives with N-terminal coiled-coil domains with the propensity to form different oligomeric states (monomer, dimer, trimer, tetramer, etc.) could be used to form synthetic proteins that are able to tune glutamatergic synaptic signaling. Indeed the production of synthetic synaptic organizers appears to be an exciting front in NPTX based synthetic biology(38).

Protein		N-terminal coiled coil region	Most likely	Second most likely
PTX3	Sequence	FIMLENSQMRERMLLQATDDVLRGELQRLREELGRLAE	Tetramer	Dimer
	Register	*defgabcdefgabcdefgabcdefgabcdefgabcdef*		
	Sequence	SLARPCAPGAPAEARLTSALDELLQATRDAGRRLARME	Tetramer	Dimer
	Register	*gabcdefgabcdefgabcdefgabcdefgabcdefgab*		
	Sequence	GAEAQRPEEAGRALAAVLEELRQTRADLHAVQGWAAR	Tetramer	Dimer
	Register	*cdefggabcdefgabcdefgabcdefgabcdefgabc*		
PTX4	Sequence	NVDVRFRSLAEESQAVAQAVNRSQASVQGELAQL	Tetramer	Trimer
	Register	*fgabcdefgabcdefgabcdefgabcdefgabcd*		
	Sequence	KAWVRKLQRRGRKVDTRLRALDLTLGERSQQ	Tetramer	Trimer
	Register	*efgabcdefgabcdefgabcdefgabcdefg*		
	Sequence	RARERKAHKAQRDALQDSLARLEGLVH	Tetramer	Trimer
	Register	*abcdefgabcdefgabcdefgabcdab*		
	Sequence	SQGARLAALEGRLPVAH	Tetramer	Trimer
	Register	*cdefgabcdefgabcde*		
NPTX1	Sequence	DMCAASVAAGGAEELRSSVLQLRETVL	Trimer	Tetramer
	Register	defgabcdefgabcdefgabcdefgab		
	Sequence	QQKETILSQKETIRELTAKLGRCESQSTL	Trimer	Tetramer
	Register	cdefgabcdefgabcdefgabcdefgabc		
	Sequence	LSRTPAAETLSQLGQTLQSLKTRLENLEQYSRLNSSS	Trimer	Tetramer
	Register	fgabcdefgabcdefgabcdefgabcdefgabefgab		
	Sequence	QTNSLKDLLQSKIDELERQVLSRVNTLEEGKGGPRNDT	Trimer	Tetramer
	Register	cdefgabcdefgabcdefgabcdefgabcdefgabcde		
	Sequence	EERVKIETALTSLHQRISELEKGQKDNRP	Trimer	Tetramer
	Register	fgabcdefgabcdefgabcdefgabcdef		
NPTX2	Sequence	QGGAQSPEEELRAAVLQLRETVVQQKETL	Trimer	Dimer
	Register	abcdefgabcdefgabcdefgabcdefga		
	Sequence	GAQREAIRELTGKLARCEGLAGGKAR	Trimer	Dimer
	Register	bcdefgabcdefgabcdefgabcdef		
	Sequence	DPGHVVEQLSRSLQTLKDRLESLEHQLRANVSNAGLP	Tetramer	Dimer
	Register	cdefgabcdefgabcdefgabcdefgabcdefgabcd		
	Sequence	GDFREVLQQRLGELERQLLRKVAELEDEKSLLHNETS	Tetramer	Dimer
	Register	efgabcdefgabcdefcdefgabcdefgabcdefgab		
	Sequence	AHRQKTESTLNALLQRVTELERGNSAFKS	Tetramer	Dimer
	Register	cdefgabcdefgabcgabcdefgabcdef		
NPTXR	Sequence	AQQGDAAGAAPGEREELLLLQSTAEQLRQTAL	Trimer	Tetramer
	Register	fgabcdefgabcdefgabcdefgabcdefgab		
	Sequence	QQEARIRADQDTIRELTGKLGRCESGLPR	Trimer	Tetramer
	Register	cdefgabcdefgabcdefgabcdefgabc		
	Sequence	DSPALILELEDAVRALRDRIDRLEQELPARVNLSA	Trimer	Dimer
	Register	cdefgabcdefgabcdefgabcdefgabcdefgab		
	Sequence	GLHSKMDQLEGQLLAQVLALEKERVALSH	Trimer	Tetramer
	Register	cdefgabcdabcgabcdefgabcdefgab		
	Sequence	SSRRQRQEVEKELDVLQGRVAELEHGSSAYS	Trimer	Tetramer
	Register	cdefgabcdefgabcdefgabcdefgabcde		

Table 1: Long pentraxins and their LOGICOIL predictions for N-terminal coiled coil regions and their oligomeric states. *Sequences of predicted coiled coil-regions and assigned heptad register shown, with the most likely and second most likely oligomeric state also shown. Sequences are shown from N-terminal to C-terminal.*

Concluding remarks

This thesis has covered the pentraxin family from a biochemical and structural point of view, of which I argue are not fully separate perspectives, from which the title "structural biochemistry of the pentraxins" was derived. Further structural and biochemical characterization of the pentraxins in physiologically relevant situations, such as bound to C1 will reveal more about some of the hypotheses postulated above. In addition, I hope to have highlighted the need for more therapies targeting the innate immune system, which appears relatively untapped compared to aspects of adaptive immunity such antibodies. Given the wide range of functions throughout the body of the pentraxins and innate pattern recognition receptors in general, their potential to enable us to overcome previously insurmountable pathologies such as Alzheimer's or atherosclerosis may just be around the corner.