Zinc binding to the Tyr402 and His402 allotypes of complement factor H: possible implications for age-related macular degeneration

Ruodan Nan,[1] Irene Farabella,[1] Felix F. Schumacher,[1] Ami Miller,[1] Jayesh Gor,[1] Andrew C. R. Martin,[1] David T. Jones,[1] Imre Lengyel[2] and Stephen J. Perkins [1*]

[1]Department of Structural and Molecular Biology,
Division of Biosciences
Darwin Building,
University College London,
Gower Street,
London WC1E 6BT, U. K.

[2] Department of Ocular Biology and Therapeutics,
UCL Institute of Ophthalmology,
University College London,
11-43 Bath Street,
London EC1V 9EL, UK

Running Title: Zinc binding to homozygous CFH

Keywords: X-ray scattering; ultracentrifugation; molecular modelling; age-related macular degeneration; retinal pigment epithelium

Abbreviations: AMD, age-related macular degeneration; FH, factor H; RPE, retinal pigment epithelium; SCR, short complement regulator.

* Author to whom correspondence and requests for reprints should be addressed (Tel: 020-7679-7048; Fax: 020-7679-7193; Email: s.perkins@medsch.ucl.ac.uk). The text consists of 48 pages (including 10 Figures and 2 Tables).

ABSTRACT

In the innate immune system, Factor H (FH) with 20 short complement regulator (SCR) domains is a major regulator of the alternative pathway of complement activation. High millimolar concentrations of zinc are found in drusen deposits, a key feature of age-related macular degeneration (AMD). Heterozygous FH is inhibited by zinc which causes FH to aggregate strongly. Here, analytical ultracentrifugation showed that large amounts of oligomers formed with both the native Tyr402 and the AMD-risk His402 homozygous allotypes of FH in the presence of zinc, while the functionally-important SCR-6/8 allotypes aggregated with zinc. X-ray scattering showed that both FH allotypes and both SCR-6/8 allotypes strongly aggregated at > 20 μM zinc. The other functionally-important SCR-1/5 and SCR-16/20 fragments were less likely to bind zinc. Starting from known zinc-binding sites, a total of 202 putative surface zinc binding sites were predicted in FH. Of the 20 SCR domains, the most sites were predicted in SCR-6. Metal site prediction web servers suggested that SCR-6 and several other domains may bind zinc. Predictions based on docked SCR-6/8 dimeric structures revealed potential zinc binding sites at the protein-protein interface that may lead to daisy-chained oligomers. We conclude that zinc binds weakly to FH at several surface locations, most probably within SCR-6/8, and this explains why zinc inhibits FH activity. Given that there are high pathophysiological levels of zinc present in subretinal deposits, we discuss how zinc binding to FH may contribute to deposit formation and inflammation associated with AMD.

INTRODUCTION

The complement system of the innate immune system is activated by the alternative pathway through the low level spontaneous hydrolysis of C3 to form C3u (also known as $C3_{H2O}$), which leads to a positive-feedback amplification of C3 cleavage to form activated C3b.[1,2] While complement is targeted against pathogenic bacteria, complement-mediated host cell damage is prevented by the complement regulator Factor H, which acts as a cofactor for factor I to cleave C3b, competes with factor B to inhibit the formation of the C3 convertase C3bBb, and accelerates the decay of C3bBb.[3] FH is constructed from 20 short complement regulator (SCR) domains, each of length about 61 residues.[4] The SCR domains in heterozygous and homozygous FH form a partially-folded back structure in solution.[5-7] FH binds to glycosaminoglycans on the host cell surface through its C-terminal region and to SCR-7, and this is followed by the decay accelerating and co-factor activity of FH against C3b through the N-terminal region.[8-10] FH binds to C3b at SCR-1/4 and SCR-19/20.[11] FH also binds to other ligands such as C-reactive protein, M-protein of Streptococci bacterium, and FH binding protein of Neisseria meningitides.[12-15] FH also self-associates to form dimer, trimer and higher oligomers.[7,16-18] FH oligomerisation is strongly promoted by zinc, followed by copper.[19,20] X-ray scattering and analytical ultracentrifugation (AUC) experiments showed that pooled native heterozygous FH aggregates strongly in the presence of ≥ 20 μM of zinc or copper, and this is matched by the decrease of fluid-phase FH activity.[20]

FH is genetically associated with age-related macular degeneration (AMD), the most common cause of blindness in the elderly in the Western population.[21-25] FH is also involved with atypical haemolytic uraemic syndrome (aHUS), membranoproliferative glomerulonephritis type II (MPGN), and Alzheimer's disease (AD).[26-28] Even with the analysis of over 100 genetic alterations, the molecular role of FH in these diseases is still unclear.[27] This lack of knowledge is attributed to the multivalent weak affinities of FH for its ligands, which makes experimental studies difficult.[29,30] A hallmark of early AMD is the appearance and growth of drusen deposits that develop within Bruch's membrane, an extracellular matrix layer interposed between the retinal pigment epithelium (RPE) and the choroidal vasculature.[31-33] Drusen contains oxidized lipids, carbohydrates, cellular materials and over 200 aggregated proteins including FH and other complement components. FH is secreted by RPE cells or is delivered by the choroidal blood circulation.[21,25,34,35] Complement activation-related inflammation has been linked with the formation of drusen.[21,25,36] A

Tyr402His polymorphism in FH SCR-7 is associated with 50% of AMD cases.[22-25] The SCR-6/8 His402 allotype shows slightly greater self-association than the Tyr402 allotype.[7,16] While the binding of C-reactive protein to FH was recently disputed, the most recent study confirmed that the two proteins interact with each other, and that the SCR-6/8 His402 allotype binds more weakly to C-reactive protein than the Tyr402 allotype.[3,12,29] The two FH allotypes showed similar cofactor activity in C3b degradation.[11,37] The effect of glycosaminoglycans on these two FH allotypes remains unclear, sometimes with conflicting results.[38]

Zinc is abundant in human retina.[39] It is vital for retinal cell survival and the functioning of antioxidant enzymes and the visual cycle, while an excess of zinc can be detrimental and exacerbate neuronal damage.[40] Under oxidative stress during light exposure, the secretion of zinc from the zinc-rich RPE cells is elevated.[40] High mM concentrations of zinc have been found in sRPEds and Bruch's membrane, suggesting that the pathological release of zinc from surrounding tissues like the RPE-choroid complex might be involved in sRPEd formation.[39,41] Studies of the effect of zinc supplements on reducing the development of AMD gave different results. Based on the premise that zinc is an antioxidant in retina, the large Age-Related Eye Disease Study (AREDS) indicated that supplements with zinc alone or with zinc together with vitamins C and E and β carotene reduces the risk for progression to advanced AMD.[42,43] Interestingly, a follow-up study found that the treatment response to the AREDS-recommended zinc supplement was influenced by the Tyr402His FH polymorphism. A smaller reduction in the progression to AMD to advanced stage disease was found in patients with the FH His402 allotype than in those with the FH Tyr402 allotype.[44]

Zinc binding sites show a tetrahedral coordination geometry at which the binding site is often formed from surface His, Asp, Glu and Cys residues.[45] When these zinc sites are found between two protein surfaces, between one to three residues at each surface contribute to a tetrahedral coordination involving four residues. The Tyr402His polymorphism potentially provides a new His zinc-binding ligand within the SCR-6/8 region that binds to heparin and C-reactive protein. To clarify whether the two allotypes exhibit different zinc binding properties, we studied the homozygous FH Tyr402 and His402 allotypes and the two allotypes of recombinant SCR-6/8 titrated with zinc. By analytical ultracentrifugation and synchrotron X-ray scattering, we showed that zinc interacts with both homozygous allotypes of FH. We also studied the effect of zinc on the other functionally important SCR-1/5 (binds

C3b) and SCR-16/20 (binds C3b, C3d, C-reactive protein and heparin) regions. Bioinformatics predictions supported the observation of zinc-binding to SCR-6/8 and less so to SCR-1/5 or SCR-16/20. We discuss the molecular implications of our results for complement regulation and sRPEd formation in AMD.

RESULTS AND DISCUSSION

Sedimentation velocity of the complexes of the two FH and SCR-6/8 allotypes with zinc

The homozygous full-length FH and recombinant SCR-6/8 allotypes (Tyr402 and His402) were purified (Materials and Methods). During this study, the two FH allotypes were studied at a concentration of 0.7-0.9 mg/ml (4.6-5.8 μM) in order to be comparable in concentration with the physiological FH range of 0.235-0.810 mg/ml in serum.[27] The SCR-6/8 allotypes were studied at 0.2-0.3 mg/ml (9.7-14.6 μM) which was the lowest concentration that produced analysable data. Both proteins were studied in Hepes buffer (Materials and Methods), in order to avoid the precipitation of zinc with phosphate. Both were titrated with 0-600 μM $ZnSO_4$ in ultracentrifugation and scattering experiments.

Analytical ultracentrifugation studies the sedimentation behaviour of macromolecules on subjecting these to a high centrifugal force in order to determine their sizes and shapes.[46] Sedimentation velocity experiments were performed using speeds up to 60,000 r.p.m on two freshly-prepared Tyr402 and two His402 FH allotypes at concentrations of 0.7-0.8 mg/ml, each being titrated with zinc from 0.2 μM to 200 μM. Similar sets of rapidly-moving sedimentation boundaries were observed for both allotypes when [Zn] increased. Good fits to the sedimentation boundaries were obtained in all cases using sedimentation coefficient distribution analyses $c(s)$ (Figures 1(a-d,f-i)). At all [Zn] values, the FH monomer was observed at a mean value of 5.67 ± 0.05 S for the Tyr402 allotype, and at 5.74 ± 0.07 S for the His402 allotype at 50,000 r.p.m. This indicated the presence of zinc-free FH even with a 40-fold excess of zinc, meaning that the binding of zinc is weak. For both allotypes with [Zn] between 0.2-20 μM, additional peaks that corresponded to similar small amounts of oligomers from dimers to nonamers were visible from 7 S to 30 S that resembled those for zinc-free FH.[6,7,18] For both allotypes with [Zn] between 60-200 μM, sizeable amounts of rapidly sedimenting species were observed with increased $s_{20,w}$ values up to 100 in the *c(s)* distributions (Figure 1(e,j)). The proportions of oligomers of the FH Tyr402 and His402 allotypes were derived by integration of the *c(s)* size-distribution analyses. No differences in the proportions of the FH monomer and oligomers at each zinc concentration were seen between the allotypes. Like heterozygous pooled native FH,[20] both the Tyr402 and His402 allotypes form heavy oligomers with zinc and share similar dependences on [Zn] (Figure 2(a)).

The sedimentation boundaries showed that differential ranges of S values for the zinc-induced FH oligomers were formed with the two allotypes. For the Tyr402 allotype, the oligomers ranged in S values in excess of 100 S, indicating a greater proportion of more compact or higher molecular mass species (Figure 1(e)). For the His402 allotype, the oligomers peaked in the region of 20-80 S, indicating a greater proportion of more extended or smaller sized oligomers (Figure 1(i)). These differences were reflected in Figures 1(a-d,f-i), where the upper part of the boundaries sedimented more rapidly for the Tyr402 allotype compared to the His402 allotype at all zinc concentrations. Further differential effects were observed for the other two FH samples, thus it was not possible to define differences between the two allotypes. In addition, using the homozygous FH proteins whose heterogeneity is much reduced compared to the pooled FH used in earlier studies, it was noteworthy that the *c(s)* peaks show no sequence of sharp peaks with well-resolved S values similar to the peaks 2-9 seen in zinc-free full-length FH.[6,7,18] Overall, these observations suggested that a range of zinc-induced FH oligomers of different sizes formed through the cross-linking of different weak zinc-binding sites at the surface of FH.

In order to clarify the molecular basis for zinc-induced FH oligomers, sedimentation velocity experiments were performed at 50,000 r.p.m on the Tyr402 and His402 allotypes of the SCR-6/8 fragment at 0.3 mg/ml (14.6 μM) titrated with zinc from 0.2 μM to 600 μM. Similar sedimentation boundaries were observed for the two allotypes of SCR-6/8 at each [Zn] (Figure 3(a-d)). In contrast to full-length FH, no rapidly sedimenting larger species were detected in the size distribution *c(s)* fits at any [Zn]. The *c*(*s*) analyses consistently showed a major monomer peak at 2.12 ± 0.09 S and a minor dimer peak at 4.03 ± 0.20 S for the Tyr402 allotype, and likewise two peaks at 2.19 ± 0.08 S and 3.95 ± 0.16 S for the His402 allotype (Figure 3(e)). This agrees with the previous observation of SCR-6/8 monomer and dimer by sedimentation velocity.[16] However, with [Zn] above 60 μM, the signal intensity of the monomer peak for both allotypes decreased significantly with increase in [Zn] (Figures 2(b) and 3(e)). This indicates that, at [Zn] above 60 μM, zinc induced the formation of aggregated SCR-6/8 which were sufficiently large to sediment to the bottom of the sample cell before the first boundary scan was recorded. While the His402 allotype showed slightly greater aggregation than the Tyr402 allotype, this may result from experimental variability in quantifying the different sedimenting species. Overall, it was noteworthy that SCR-6/8 aggregated at the same [Zn] value that caused large FH oligomers to form. This indicated that major binding sites for zinc in full-length FH were located within SCR-6/8.

The reversibility of the zinc-induced aggregation of the SCR-6/8 allotypes was studied by adding 2 mM EDTA to the ultracentrifuge cell contains 0.3 mg/ml SCR-6/8 and 300 μM of zinc. After mixing, the sedimentation velocity experiment was immediately re-run at 50,000 r.p.m on this cell again. Even though the size distribution *c(s)* analyses showed that the monomer peak for both allotypes of SCR-6/8 with 300 μM of zinc was much reduced, the addition of EDTA caused a significant increase of the monomer peak intensity for both allotypes (Figure 4). This showed that the zinc-induced aggregation of SCR-6/8 was reversible, in good agreement with the reversibility of the zinc-induced oligomers of heterozygous pooled FH by adding EDTA.[20]

Additional ultracentrifugation experiments with the SCR-1/5 and SCR-16/20 fragments in the presence of zinc were performed as controls of the SCR-6/8 experiments. SCR-1/5 is less soluble at 0.3 mg/ml when expressed, compared to SCR-6/8 and SCR-16/20. As [Zn] increased, SCR-1/5 aggregated in the same manner as SCR-6/8 at a similar [Zn] value to that for SCR-6/8, and this may reflect its lower solubility. This was visible from both the reduced intensities of the fringes as [Zn] increased and the marked reduction of the monomer peak intensity in the *c(s)* plot (Figure 5(a-d)). In contrast to this, SCR-16/20 is more soluble than SCR-1/5 and SCR-6/8, and SCR-16/20 remained in solution as [Zn] increased with a smaller reduction in the number of observed fringes at 600 μM. Here, the *c(s)* plot showed a smaller reduction in the sizes of the monomer and dimer peaks, while at the same time as many as three other species were observed (peaks 1, 2 and 3 in Figure 5(e-h)). These experiments showed that other different types of zinc-induced oligomers formed in two other regions of FH. Unlike SCR-6/8, both SCR-1/5 and SCR-16/20 possessed His_6-tags which may comprise potential zinc binding sites.[17] This was evaluated by further zinc titrations in which a 2:1 molar ratio of nickel (17 μM) was present to saturate its binding site at the His_6-tag, to which nickel binds more tightly with a dissociation constant of 1 μM than zinc.[47] The resulting *c(s)* plots showed that no SCR-1/5 or SCR-16/20 aggregates were seen, and that only the monomer or monomer/dimer peaks were observed that were similar to those observed for zinc-free protein. These final experiments indicate either that the His_6-tag represented an oligomeric zinc-binding site which is eliminated by the stronger binding of nickel to this, and/or that nickel competes with zinc for weak interaction sites in FH. The first explanation is preferred, because nickel has no strong effect on FH oligomer formation.[20] Control experiments with full-length FH His402 in the presence of 200 μM zinc and 20 μM

nickel, 200 μM zinc, or 20 μM nickel were performed to complete these studies. This showed that, even in the presence of nickel, zinc still causes full-length FH to aggregate in similar quantities. Even though these experiments are cautionary in that other possible weak zinc-binding sites in FH may exist, these experiments show that a major region for zinc-binding to FH lies in SCR-6/8.

X-ray scattering of the two allotypes of FH and SCR-6/8 with zinc

X-ray scattering is a diffraction method used to study solution structures of macromolecules in random orientations.[48] The effect of zinc on the structures of freshly-purified two Tyr402 and two His402 allotypes of full-length FH and the two allotypes of FH fragment SCR-6/8 were investigated by synchrotron X-ray scattering. FH was studied at 0.9 mg/ml (5.8 μM) in order to be comparable with physiological FH concentrations, while SCR-6/8 was studied at 0.2 mg/ml (9.7 μM), this being the lowest concentration of SCR-6/8 that produced analysable data. Each FH or SCR-6/8 sample was titrated with 0 μM to 600 μM $ZnSO_4$. The scattering data $I(Q)$ showed excellent signal-noise ratios and no detectable effect from radiation damage.

Guinier fits at low Q values (where $Q = 4\pi \sin\theta/\lambda$; 2θ = scattering angle; λ = wavelength) gives the radius of gyration (R_G) which monitors the degree of elongation of the protein, and the $I(0)/c$ value which is proportional to the relative molecular mass (Figure 6(a,b)).[48,49] At larger Q values, Guinier fits give the cross-sectional radius of gyration R_{XS} which monitors the structural proximity relationships between non-neighbouring SCR domains (R_{XS-1}) and neighbouring SCR domains (R_{XS-2}) (Figure 6(c-f)).[5] In our previous study of heterozygous and homozygous FH, the presence of oligomers caused concentration-dependent increases in the values of R_G and R_{XS-1}.[7,18] Here, at the lowest zinc concentration of 0.2 μM, the mean R_G and R_{XS-1} values were 8.91 ± 0.17 nm and 3.03 ± 0.17 nm respectively for the Tyr402 FH allotype, and 9.23 ± 0.72 nm and 3.15 ± 0.39 nm respectively for the His402 FH allotype (Figure 7(b,c)). Both the R_G and R_{XS-1} values were similar to those for heterozygous FH at 0.9 mg/ml without zinc.[18] Both R_G and R_{XS-1} values were the same within error of the corresponding values for zinc-free homozygous FH at 1.1 mg/ml.[7] When both the Tyr402 and His402 FH allotypes were titrated with zinc up to 6 μM, the scattering curves $I(Q)$ remained unchanged (Figure 6(a-d)), and the Guinier values remained unchanged (Figure 7(a-c)). For [Zn] from 20 μM to 600 μM, the $I(Q)$ intensities increased significantly

at low Q and decreased at high Q as the result of oligomer formation (Figure 6(a-f)). The Guinier $I(0)/c$ and R_G values increased significantly (Figure 7(a,b)), although the R_{XS-1} and R_{XS-2} valued decreased at [Zn] above 200 μM as the result of the decrease in the intensity of $I(Q)$ in the Q ranges used for the R_{XS} determinations (Figures 6(c-f) and 7(c,d)). The changes for both homozygous FH allotypes are comparable with previous observations for heterozygous native FH,[18] indicating again that FH heterogeneity is not responsible for these observations. For [Zn] between 20 μM and 120 μM, the two His402 allotypes showed slightly higher Guinier $I(0)/c$, R_G and R_{XS-1} values than the two Tyr402 allotype (Figure 7(a-c)). If this is significant, this would indicate that the His402 allotype formed more oligomers than the Tyr402 allotype, however the difference is not considered to be large enough.

To test whether the Tyr402 and His402 allotypes of SCR-6/8 formed zinc-induced oligomers, both were titrated with [Zn] from 0.2 to 600 μM. For SCR-6/8 at 0.2 mg/ml with 0.2 μM zinc, the Guinier R_G and R_{XS} values were 2.63 ± 0.46 nm and 0.88 ± 0.29 nm respectively for the Tyr402 allotype, which were similar to values of 2.79 ± 0.47 nm and 1.00 ± 0.27 nm for the His402 allotype. Those Guinier values are slightly less but within error of the experimental R_G and R_{XS} values of 3.26-3.35 nm and 1.04 nm respectively for the zinc-free Tyr402 allotype, and 3.12-3.21 nm and 1.15 nm (± 0.06-0.20 nm) for the zinc-free His402 allotype.[16] For [Zn] between 0.2 μM to 6 μM, the scattering curves $I(Q)$ were unchanged for both allotypes (Figures 6(g-j) and 7(e-g)). At 20 μM [Zn], the $I(Q)$ curves of both allotypes underwent dramatic changes in which the intensities sharply increased at low Q and decreased rapidly at large Q (Figure 6(g-j)), and these changes resulted in much increased Guinier $I(0)/c$ and R_G values (Figure 7(e,f)). Above 20 μM [Zn], the $I(Q)$ intensities continued to rise at very low Q (Figure 6(g,h)). The Guinier $I(0)/c$ and R_G values peaked at [Zn] of 60 μM and declined at higher [Zn] values (Figure 7(e,f)). These changes showed that both allotypes of SCR-6/8 interacted with zinc starting from 20 μM, which is the same starting concentration as the FH interaction with zinc. Unlike the FH-zinc oligomers which remained in solution, the changes observed for SCR-6/8 caused this to aggregate, and this accounted for the apparent decline in the $I(0)/c$ and R_G above 60 μM. Like full-length FH, no significant difference was observed for the zinc-induced oligomerisation of the Tyr402 and His402 allotypes of SCR-6/8.

The distance distribution function $P(r)$ of the two FH allotypes provided shape information in real space on the zinc-induced FH oligomers. The $P(r)$ curve summarises the distances between all pairs of atoms within FH, and gives the most frequently occurring distance M from the position of the peak maximum, the maximum length L from the point at which *P(r)* becomes zero at large Q, and an independent calculation of the R_G and *I(0)* values for comparison with the Guinier determinations (not shown).[48] At [Zn] up to 6 μM, the shape of the $P(r)$ curves remained unchanged with increase in [Zn], and was indistinguishable between the Tyr402 and His402 FH allotypes (Figure 8). The length L of both allotypes was 32 nm, and the maximum M was at 5.36 ± 0.11 nm and 5.20 ± 0.11 nm for the Tyr402 and His402 allotypes. These values are in good agreement with previous observations for heterozygous FH with zinc.[20] Also similar to previous, the intensity of the $P(r)$ curves and the L and M values for both allotypes increased with increase in [Zn] above 6 μM, which indicates that zinc-induced oligomerisation occurred with both the Tyr402 and His402 allotypes and was not dependent on heterogeneity (Figure 8). For [Zn] at 20 μM, 60 μM and 120 μM, the shapes of the *P(r)* curves were similar between the two allotypes, indicating that similar-sized oligomers had formed under these conditions (Figure 8). The *P(r)* intensities and M values were slightly higher for the His402 allotype than the Tyr402 allotype, which is consistent with the Guinier analyses, however the difference is not considered to be significant (Figure 8). When the *P(r)* analyses were repeated for the SCR-6/8 allotypes with [Zn] up to 6 μM, L was determined to be 9 nm for both allotypes, and the curves were indistinguishable between the two allotypes (data not shown). The $P(r)$ curves were similar in appearance to those for SCR-6/8 without zinc.[16] Above 6 μM zinc, the aggregation of SCR-6/8 made further $P(r)$ analyses impossible. The lack of difference between the SCR-6/8 allotypes indicated that His402 had no effect on zinc-induced oligomerisation compared to Tyr402.

Fluid-phase activity assays of two allotypes of full-length FH complexes with zinc

In order to test the impact of zinc on the regulatory role of the Tyr402 and His402 FH allotypes, assays of fluid-phase ammonium-inactivated C3 (C3u) cleavage by factor I and two allotypes of FH were performed by zinc titration.[50] Previous fluid-phase activity assays had been performed at significantly lower FH concentrations that included 45 μg/ml,[50] 3 μg/ml,[51] 30 μg/ml[52] and 0.3 μg/ml[53] than the 0.235-0.81 mg/ml (1.5–5.3 μM) observed *in vivo*. Our assays with the FH allotypes were performed at 0.3 mg/ml (1.9 μM) which is

comparable with both its physiological range and the heterozygous FH concentrations used previously.[20] Thus the α-chain of C3u was cleaved by factor I in the presence of the FH Tyr402 allotype without zinc to produce two major degradation fragments at apparent sizes of 45 kDa and 75 kDa (Figure 9(a)), and the cleavage rate was significantly reduced in the presence of 200 μM zinc (Figure 9(b)). The influence of zinc on this reaction for both FH allotypes was investigated at concentrations of 0 μM, 120 μM and 200 μM (Figure 9(b)). This revealed that the C3u cleavage rate was decreased by zinc, in good agreement with the decrease seen previously for heterozygous FH.[20] No difference between the Tyr402 and His402 allotypes was seen (Figure 9(c)). If similar levels of FH oligomerisation with both allotypes accounted for the decrease of FH activity by blocking FH access to C3u and/or factor I,[20] this would account for the observed reduced rate of C3u cleavage.

Bioinformatics prediction of zinc-binding sites on full-length FH

Two bioinformatics strategies were applied in order to assess potential zinc-binding sites in FH. The first approach utilised a distance-dependent knowledge-based method that was based on known zinc-binding sites. A total of 3705 proteins with Zn-binding sites were identified in the Protein Data Bank, and these were screened to identify the residues that were within 0.3 nm of the metal site. Zinc binding sites at the interface between two protein surfaces generally contain Asp, His and Glu sidechains as zinc-coordination sites.[45] Cys residues were not included in this survey because there are no exposed Cys residues in full-length FH. Accordingly a total of 830 zinc-binding sites composed of His, Glu and /or Asp residues only were indentified and classified into 33 subgroups based on their compositions. A contact matrix was defined for each subgroup that is based on the mean separation ± one standard deviation between the α-carbon atoms and that between the β-carbon atoms (Materials & Methods). This matrix was used to screen the FH domains in order to predict potential interface zinc-binding sites. For this, crystal and NMR structures for 11 SCR domains and nine homology-modelled structures available at the time of this work were used (Table 1(a)).[27,54-57] A total of 202 possible zinc binding sites were identified as residue pairs or triplets that could act as one-half of a zinc-interface binding site. In some cases, a pair or triplet of residues could form a partial zinc binding site with more than one coordination geometry (Table 1(b)). In order of probability, 31% of these sites were identified in SCR-6 which had 62.5 sites, followed by 12% in SCR-3, 10% in SCR-16, 7% sites in SCR-2 and 7% sites in SCR-14. Two or fewer predicted zinc–binding sites were found in five SCR domains

(Table 1(a)). All the triplets were located in SCR-6 or between SCR-6 and SCR-7 (Table 1(b)). The predicted highest number of zinc binding sites in SCR-6 agrees with the similar effect of zinc on the aggregation of full-length FH and SCR-6/8. The predicted involvement of SCR-2, SCR-3 and SCR-16 concurs with the ultracentrifugation experiments with SCR-1/5 and SCR-16/20 that suggest that other weak binding sites may be present, assuming that the His-tag has no effect. Even though the disease-related His402 residue may be a potential zinc-binding residue, His402 occurred in only two out of the 202 zinc binding sites predicted in FH. The His402 polymorphism thus has a low probability of altering the interaction with zinc, and this agrees with the observed lack of difference between the Tyr402 and His402 allotypes of FH and SCR-6/8 by ultracentrifugation and scattering.

The second bioinformatics approach was based on submitting the 20 crystal, NMR and homology SCR structures to web servers that predict metal binding sites in proteins (Materials and Methods). METSITE was developed using relative residue positions and does not require side-chain atoms to be present, therefore this is applicable to low resolution structures.[58] It has a high sensitivity of 94.6% and a medium specifity of 47.8% for zinc. This method is based on recognizing residue environments in which a metal ion is likely to be found within 0.7 nm. These environments are represented by structural features such as residue beta-carbon distances and solvent accessibility. If a 0.4 nm threshold is applied to the METSITE results, METSITE predicted that a total of 26 out of 175 His, Asp and Glu residues in FH are candidates for 13 predicted interaction sites with zinc. SCR-6 with five residues gave the highest total among the 20 SCR domains, followed by SCR-15 and SCR-20 with three each. A ranking based on the number of residues per zinc binding site (between 1.3 to 1.4) showed that SCR-6, SCR-7, SCR-8, SCR-15 and SCR-20 were the most likely to contain zinc binding sites. Given this outcome, the crystal structure of SCR-6/8 (PDB code 2UWN) was used to create 195 possible homodimeric structures for SCR-6/8 using four different protein docking servers (Materials and Methods). All 195 dimeric structures were subjected to zinc binding site prediction using the web servers METSITE and CHED (Materials and Methods). METSITE identified possible zinc binding sites in 193 out of 195 structures, while CHED with no activated filters identified possible zinc binding sites in 91 structures. The difference is attributed to the higher sensitivity but lower selectivity of METSITE compared to CHED.[58,59] The ten most frequent possible zinc binding site residues in the 195 dimeric structures were Glu359, His360, His371, Asp370, Asp485, His332, His417, Glu462, Glu487 and Asp497 in order of frequency (Table 2). Interestingly, His402

was not included in those ten residues, which is consistent with the experimental observations of no difference seen between the two SCR-6/8 allotypes in the presence of zinc. In order to rationalise the outcome of these searches, the 195 putative dimeric structures were analysed with RMSDCLUST (Materials and Methods) which superposes every pair of structures and identifies clusters of similar structures. Nine different zinc binding sites were found in seven clusters that each contained 18, 15, 4, 2, 3, 4 and 5 SCR-6/8 dimer structures, together with a single outlier that did not belong to any cluster. Only two out of the nine possible zinc binding sites possessed all the residues identified as the ten most frequent possible zinc binding residues. One of these binding sites involved Glu359 and His360 on each of two SCR-6/8 monomers (Figure 10(a)), and the other involved His371 on one SCR-6/8 monomer and Asp485 and Glu487 on the other monomer (Figure 10(b)). If zinc cross-links pairs of SCR domains, the predictions of Figures 10(a,b) leads to a schematic cross-linking cartoon in which adjacent pairs of two SCR-6 domains or SCR-6 and SCR-8 domains form contacts with each other. The indefinite self-association suggested in Figure 10(c) shows how zinc may cross link SCR-6/8 molecules to form large aggregates. As a control of this predictive approach, the same procedure was performed with three protein structures with known zinc binding sites (Table 2). These were human interferon-β, the complex of human growth hormone and prolactin, and an archael cytochrome P450. All the known zinc binding residues in the three known protein structures were included in the top ten most frequently occurring residues, and this outcome supports the credibility of this predictive method for FH SCR-6/8.

CONCLUSIONS

Overall, we have located a major FH zinc binding site to lie within SCR-6/8, and this accounts for the inhibition of FH activity by zinc. Our earlier studies showed that heterozygous FH aggregated in the presence of five transition metals at physiological FH concentrations, in particular at zinc and copper concentrations of ≥ 20 μM.[19,20] Zinc has the strongest effect on FH, and inhibits its function. Apart from FH, high zinc concentrations have been reported in retinal deposits, suggesting a potential role for zinc-induced FH aggregation in age-related macular degeneration.[41] In this study, using ultracentrifugation and X-ray scattering, we have shown that similar zinc-induced aggregation occurs for both homozygous forms of FH with either the Tyr402 or disease-related His402 allotypes. The FH Val62 or Ile62 polymorphism (which is also disease-related) also has no influence on zinc binding effects. Therefore zinc-induced aggregation is independent of FH heterogeneity. Based on ultracentrifugation and scattering and bioinformatics predictions, we showed that SCR-6/8 is a major locus for zinc-induced aggregation through the presence of partial zinc binding sites at the FH surface. This region can be cross-linked in the presence of zinc to form large daisy-chained oligomers (Figure 10). Control experiments with the two other functionally-active regions of FH at SCR-1/5 and SCR-16/20 showed that both regions may aggregate in the presence of zinc, but this is less significant than that at SCR-6/8. This work now explains the zinc-induced formation of FH oligomers for the first time in terms of multiple weak surface zinc-binding sites located primarily within the SCR-6/8 region. The compact nature of these FH oligomers (Figure 8) is explained by zinc binding at a central location in FH within SCR-6/8. The proximity of SCR-6/8 to the C3b-binding region of FH at SCR-1/5 explains why zinc binding inhibits FH. The Tyr402 or His402 allotypes have no significant influence on zinc-induced aggregation, even though His402 might be considered as a potential surface zinc-binding residue.

FH is a major complement regulator in blood, and is expressed and secreted by many different cell types including the RPE.[60] The major physiological ligands of FH include C3b and its C3d fragment, heparin and other glycososaminoglycans, C-reactive protein and FH self-association, in addition to zinc binding.[29,30] All these ligands bind weakly to FH with μM affinities, and this is consistent with the relative μM abundance of FH in serum. The FH-zinc interaction has already been discussed previously.[20] Since that 2008 study, more recent studies show that C3b and C3d bind with K_D values of 3.5-14 μM and 2.6 μM

respectively,[11,61] C-reactive protein binds with a K_D value of 4-15 μM,[12] and FH dimerises with a K_D value of 28 μM.[18] If FH is found at 0.6 mg/ml in serum, this means that 11% of FH exists as dimers. If C3b occurs at 1 mg/ml, 29% of FH will be bound to C3b if the K_D value is 3.5 μM. If C-reactive protein occurs at the acute-phase level of 0.4 mg/ml, 32% of FH will be bound to C-reactive protein if the K_D value is 4.2 μM. These calculations assume that no other factors in serum influence these equilibria, which have been considered separately from each other for purpose of illustration here. In comparison to these physiological interactions, the pathophysiological interaction between FH and zinc is characterised by comparable K_D values at approximately 6-20 μM, implying that FH will be 50% bound to zinc if bioavailable zinc is present at 13 μM. However most of the zinc in plasma is bound to proteins such as human serum albumin (HSA), and only 2-8% of this zinc is bioavailable.[62] Potentially higher bioavailable zinc levels may exist in drusen (below). Because two of the most frequently predicted zinc binding residues His360 and His371 (Table 2) have been implicated in glycosaminoglycan binding to FH,[56] it is possible that zinc binding to FH can perturb the binding of glycosaminoglycans to FH.

The comparison of zinc binding to FH and HSA is of interest. HSA is a three-domain protein that circulates in serum at reference levels between 30-50 mg/ml (450-750 μM). A zinc-binding site was identified at the protein surface interface between Domain I at His67 and Asn99 and Domain II at His247 and Asp249 (PDB code 1AO6) that is conserved in almost all mammalian species.[62,64] Here, asparagine is a rare zinc-binding ligand. Being stereochemically well-defined, the K_D for zinc binding to HSA is reduced to 1 μM compared to that of approximately 20 μM for zinc binding to FH. The larger K_D value for the FH-zinc interaction is explained by the lack of proximity in monomeric FH of appropriate pairs of the SCR domains that possess potential zinc-binding residues. Given the 450-750 μM concentration of HSA compared to the serum zinc concentration of 12.5 μM, where HSA typically binds about 80% of plasma zinc, the presence of HSA will reduce the level of bioavailable zinc, and accordingly there is little risk of FH aggregation with zinc in blood.

Human ocular tissues contain unusually high concentrations of zinc, where the zinc concentration is 472 μg/g of dry tissue in the RPE-choroid complex, and 464 μg/g in the neurosensory retina.[39] In contrast, the zinc level in plasma remains low at 14.7 μM even after a daily diet supplement with 80 mg zinc in the AREDS trials.[42,63] Therefore, normal zinc

levels in plasma are too low to induce FH oligomerisation. Zinc in the RPE-choroidal complex is expected to be mostly bound to melanin, metallothionein and other zinc binding proteins, and it has been proposed that less than 10% of zinc in retina is free or chelatable.[39,40,65] Zinc appears to be essential for the normal function of the retina, but its exact role is not clear. RPE cells can accumulate zinc following oral supplementation and retain it longer than any other tissues in the body.[65] Zinc might be released from the RPE through photo-oxidative damage. The released zinc could then be entrapped in the Bruch's membrane and contribute to the several hundred ppm of zinc (mM) observed in sub-RPE deposits.[40,41] Even a small percentage of the RPE-released zinc might be sufficient to induce a localised aggregation of FH that has been recruited to the location because of the process of inflammation associated with RPE damage. The above comparison with the K_D values for FH interactions with its other ligands shows that the regulatory balance of FH activity in terms of interactions with C3b, heparin and C-reactive protein can be disturbed by zinc binding. The main interaction of zinc with the SCR-6/8 regions, which is a locus for interactions with heparin and C-reactive protein indicates that FH function with these two ligands will be perturbed if sufficient excess zinc is available. In addition, the indefinite self-association of FH at SCR-6/8 is expected to block the binding of the neighbouring FH SCR-1/4 domains to C3b. Such a mechanism can contribute to both the formation of drusen deposits and the inhibition of the regulatory function of FH to cause inflammation and host cell damage.

Daily supplements with 80 mg zinc, 2 mg copper, vitamins C and E, and β carotene are recommended by AREDS as a means of reducing the risk for progression to advanced AMD.[42,43] Interestingly, the AREDS-recommended zinc supplements with antioxidants was recently correlated with a greater reduction in progression to advanced AMD in patients who are homozygous for the FH Tyr402 allotype (34% of cases) compared to patients who are homozygous for the FH His402 allotype (11% of cases).[44] Our current study offers no direct explanation for this observation because no significant difference in zinc-induced FH aggregation was seen between the Tyr402 and His402 allotypes of FH and SCR-6/8.

The most common coordination geometry of zinc binding sites is tetrahedral.[45] Here, our bioinformatics analyses indicated the availability of 3705 zinc binding co-ordinations in the Protein Data Bank for evaluation. Protein interface zinc binding sites are primarily supplied by His, Glu and Asp residues, but sites containing Cys are also found (not relevant

to FH because it contains no free Cys residues), and β-sheet secondary structures predominately contribute to zinc binding sites. Zinc binding is able to induce protein self-association or link two different proteins.[45] Our previous discussion of potential FH zinc binding sites[20] is now clarified by the abundance of potential zinc sites that were predicted using a distance-based algorithm, suggesting that multiple weak zinc binding sites exist in FH. This outcome predicts that a range of different zinc-induced oligomers will form at several sites. This multiplicity of oligomers is in good agreement with the observation of many overlapping peaks in the ultracentrifugation *c(s)* size distribution analyses (i.e. the opposite of the formation of fewer well-resolved signals such as the nine peaks observed for FH alone[7,18] and also with the formation of indefinite oligomers at high zinc concentrations. The prediction of multiple sites also explains better why different types of FH oligomers were observed with different metals such as copper.[20] Finally the predictions were also informative in that they explained the lack of difference seen between the SCR-6/8 Tyr402 and His402 allotypes, because His402 was not predicted to be a significant zinc-binding ligand. These analyses provide insight into the weak binding of zinc to other plasma proteins. It is possible that the outcome of this FH-zinc study will be applicable to other plasma proteins. For example, we have observed that zinc will induce oligomer formation in C3 (K. Li & S. J. Perkins, unpublished data).

In summary, these results contribute to a better understanding of how zinc might be involved in two distinctively different ways that contribute to AMD.[66] In early AMD disease, zinc release may trigger drusen formation and contribute to the growth of these deposits through oligomerisation of proteins including FH. This will not only lead to deposit formation but also to uncontrolled inflammation, the two hallmarks for the development of this blindness condition. In late AMD, however, zinc supplementation could slow disease progression through a currently unidentified mechanism. This however is unlikely to involve FH oligomerisation.

MATERIALS AND METHODS

Protein purification of FH and FH SCR-6/8

Human blood was obtained from 48 anonymous genotyped healthy volunteers following ethical approval. Direct DNA sequencing of the PCR product of the FH gene was used to identify the Tyr402His and Val62Ile FH polymorphisms in each volunteer. Four homozygous FH Tyr402 allotypes (coded 026II, 022LM, 018DK and 016CC), and four FH His402 allotypes (coded 027BA, 032KO, 030TJ and 033NV) were purified using monoclonal affinity chromatography with an MRC-OX23 Sepharose column.[18,67] The four Tyr402 allotypes were homozygous for Ile62, homozygous for Val62, heterozygous for Val62/Ile62, and homozygous for Val62 in that order, while all four His402 allotypes were homozygous for Val62. Native factor I was purified from a pool of just-outdated anonymised human plasma using another monoclonal affinity MRC-OX21 Sepharose column for FI.[18,67] Bound homozygous FH and native factor I were each eluted from the columns using 3 M $MgCl_2$, pH 6.9, then each was dialysed into HEPES buffer (10 mM HEPES, 137 mM NaCl, pH 7.4) in the presence of 0.5 mM EDTA to remove Mg^{2+}. The MRC-OX23 column was washed with guanidine (0.2 M Tris, 4 M guanidine-HCl, pH 8.0) after each FH preparation to avoid allotype cross-contamination. IgG contaminants were removed by passing the FH and FI purifications through a Hitrap™ Protein G HP. Non-specific aggregates and human serum albumin were removed by gel filtration on a Superose™ 6 prep grade XK 16/60 size-exclusive column. Native C3 was purified from fresh plasma by ion-exchange chromatography, and haemolytically inactive C3 in which the reactive thioester is hydrolysed (C3u) was prepared by incubating native C3 in 200 mM hydrazine at 37 °C for 1 h.[68] The recombinant FH fragment SCR-6/8 Tyr402 was expressed in BL21 (DE3) *E. coli* cells using a pET21ab vector system, while SCR-6/8 His402 was expressed in BL21 (DE3)pLysS *E. coli* cells using a pET14b vector system. SCR-6/8 was extracted from inclusion bodies, solubilised and refolded. The refolded SCR-6/8 was purified by using ion-exchange chromatography with a Hitrap™ 5 ml Heparin HP column (A. Miller and S. J. Perkins, unpublished data).[16] The purifications of SCR-1/5 and SCR-16/20 are described elsewhere.[17] Proteins were dialysed into HEPES buffer for zinc experiments because HEPES does not interact with zinc. Protein concentrations were determined from absorption coefficients of 16.2 for full-length FH Tyr402, and 16.1 for full-length FH His402, 12 for FI, 9.4 for C3u, 17.2 for SCR-1/5, 22.7 for SCR-6/8 Tyr402, 22.1 for SCR-6/8 His402, and 15.7 for SCR-16/20 (1%, 280 nm, 1 cm path length).[5,6,17,69,70] The integrity of protein samples were

routinely checked by SDS-PAGE before and after scattering and ultracentrifugation experiments.

Sedimentation velocity data collection and analyses

Analytical ultracentrifugation data were obtained on two Beckman XL-I instruments equipped with AnTi50 or AnTi60 rotors, using two-sector cells with column heights of 12 mm at rotor speeds of 50,000 r.p.m and 60,000 r.p.m. Sedimentation velocity experiments at 20^{o}C were performed for the two FH allotypes Tyr402 (026II and 022LM) and the two His402 allotypes (027BA and 032KO) at 0.7 – 0.8 mg/ml which were titrated with $ZnSO_4$ at concentrations of 0.2 μM, 2 μM, 6 μM, 20 μM, 60 μM, 120 μM and 200 μM. SCR-1/5 at 0.3 mg/ml was titrated with $ZnSO_4$ at concentrations of 0 μM, 2 μM, 6 μM, 20 μM, 60 μM, 200 μM and 600 μM. The Tyr402 and His402 allotypes of SCR-6/8 at 0.3 mg/ml were titrated with $ZnSO_4$ at concentrations of 0.2 μM, 2 μM, 6 μM, 20 μM, 60 μM, 200 μM, 300 μM and 600 μM. SCR-16/20 at 0.1 mg/ml was titrated with $ZnSO_4$ at concentrations of 0 μM, 2 μM, 20 μM, 200 μM and 600 μM. Data analyses were performed using SEDFIT software (version 11.71).[71,72] The distribution analyses *c(s)* provided size and shape data by directly fitting the boundary Lamn equation to 350 scans for FH, 480 scans for SCR-1/5, 640 scans for SCR-6/8, and 440 scans for SCR-16/20. The *c(s)* analyses were based on a fixed resolution of 200, and floated the meniscus, the bottom of the cell, the baseline and the average frictional ratio *f/fo* starting from 1.781 for FH and from 1.200 for SCR-6/8, until the overall root mean square deviation and fits between the observed and calculated sedimentation boundaries were satisfactory. The *f/fo* values were fixed at 1.45 for SCR-1/5, 1.35 for the SCR-16/20 monomer, and 1.58 for the SCR-16/20 dimer.[17] Monomer and oligomers were quantitated using the integration function in the *c(s)* analyses. The percentage fraction of monomer and oligomers of full-length FH was derived by assuming that the sum of the signal intensities of monomer and oligomer was 100%. Other details are described elsewhere.[17,18]

X-ray scattering data collection and analysis

X-ray scattering data was acquired for the FH allotypes in one beam session, and for the SCR-6/8 allotypes in a second beam session on Beamline ID02 at the European Synchrotron Radiation Facility (Grenoble, France) operating with a ring energy of 6.0 GeV in 4-bunch mode and 16-bunch mode in the two beam session respectively to reduce the incident flux.[73] Storage ring currents were 29 - 43 mA in the first beam session, and 56 - 91

mA in the second beam session. The sample-detector distance was 3 m for the first beam session, and 1.5 m for the second beam session. Potential radiation damage was eliminated by the continuous movement of the sample in a flow cell during beam exposure, the use of 10 time frames of duration between 0.1 sec and 0.25 sec each during each acquisition, and on-line checks for the absence of radiation damage at low Q. In the first beam session, the FH Tyr402 allotypes 018DK and 016CC and the His402 allotypes 030TJ and 033NV were studied at 0.9 mg/ml (5.8 μM) with $ZnSO_4$ at concentrations of 0.2 μM, 2 μM, 6 μM, 20 μM, 60 μM, 120 μM, 200 μM and 600 μM. In the second beam session, the Tyr402 and His402 allotypes of SCR-6/8 were studied at 0.2 mg/ml (9.7 μM) with $ZnSO_4$ at concentrations of 0.2 μM, 2 μM, 6 μM, 20 μM, 60 μM, 120 μM, 200 μM, 300 μM and 600 μM. All the measurements were done in HEPES buffer (10mM HEPES, 137mM NaCl, pH 7.4). Other details including data reduction are described elsewhere.[16,17]

In a given solute-solvent contrast, the radius of gyration R_G corresponds to the mean square distance of scattering elements from their centre of gravity, and is a measure of structural elongation. Guinier analyses at low Q gives the R_G value and the forward scattering at zero angle $I(0)$ from the expression:[49]

$$\ln I(Q) = \ln I(0) - R_G^2 Q^2/3.$$

This expression is valid in a $Q.R_G$ range up to 1.5. The $I(0)/c$ value (c is the protein concentration in mg/ml) is proportional is the relative molecular mass Mr. If the structure is elongated, the mean cross-sectional radius of gyration R_{XS} and the cross-sectional intensity at zero angle $[I(Q)Q]_{Q\to 0}$ are determined from Guinier analyses in a Q range larger than that used for the R_G determination:[49]

$$\ln [I(Q)Q] = [I(Q)Q]_{Q\to 0} - R_{XS}^2 Q^2/2.$$

The Guinier analyses were performed using an interactive PERL script program SCTPL7 (J. T. Eaton and S. J. Perkins, unpublished software) on Silicon Graphics OCTANE Workstations.

Indirect transformation of the $I(Q)$ curve measured in reciprocal space into real space gives the distance distribution function $P(r)$ and was carried out using the program GNOM:[74]

$$P(r) = \frac{1}{2\pi^2}\int_o^\infty I(Q)\, Qr \sin(Qr)\, dQ$$

$P(r)$ corresponds to the distribution of distances r between volume elements. This offers an alternative calculation of the R_G and *I(0)* values that is based on the full scattering curve *I(Q)*, and not that at low Q. It also gives the most frequently occurring distance M and the maximum dimension of the macromolecule L. For FH at 0.9 mg/ml titrated with zinc, the X-ray curves utilised up to 344 data points for Q between 0.08 nm^{-1} and 1.60 nm^{-1}. Other details are described elsewhere.[5,16-18]

Fluid-phase activity assays

To determine activities, the reaction mixtures containing 0.3 mg/ml FH, 0.3 mg/ml C3u and 0.003 mg/ml factor I were incubated in a water bath at 37 °C in three concentrations of $ZnSO_4$ (0 μM, 120 μM and 200 μM). The functions of two pairs of the two allotypes of FH were tested. The FH Tyr402 allotype 013HH and the FH His402 allotype 015LT were tested with 0 μM and 200 μM of zinc, and the FH Tyr402 allotype 022LM and the His402 allotype 032KO were tested with 120 μM of zinc. At timed intervals, 5 μl aliquots were removed for reducing SDS-PAGE analyses. Three blank controls were used, namely 0.3 mg/ml C3u; 0.3 mg/ml C3u and 0.3 mg/ml FH; and 0.3 mg/ml C3u, 0.3 mg/ml FH and zinc. The C3 α-chain band densities from SDS-PAGE were measured using the gel analysis system SYNGENE (Synoptics Ltd., Cambridge, UK). The C3 α-chain cleavage was referenced to the averaged density of the uncleaved bands in the controls.

Molecular prediction of surface zinc binding sites

Structural database analyses were used to predict protein-interface-type zinc binding sites on protein surfaces. A dataset of 3705 Zn-binding protein crystal structures with resolutions better than 0.3 nm was identified from the Protein Data Bank. Residues that were located within 0.3 nm distance of the zinc atom were assigned as putative zinc coordinating residues using a PERL script to scan the Protein Data Bank. Of the 4883 putative zinc binding sites that were identified, a total of 830 zinc-binding sites composed of His, Glu and /or Asp as zinc coordination residues were analysed and classified into 33 subgroups based on the amino acid composition of each site. In each subgroup, the distance x between the Cα atoms i and that between the Cβ atoms j of each zinc-coordinating residue were calculated as a residue pair type x(i,j), and stored in a contact matrix CM. For each residue pair in CM, the average distance $\bar{x}$ and its standard deviation σ was calculated. An algorithm was implemented to screen each protein structure of interest for residues that form putative partial

Zn binding sites. For a given protein structure, solvent-exposed Asp, His and Glu side-chains were identified. A residue is considered exposed if its side-chain has a relative surface accessibility greater than 20%. The geometric distances between pairs and triplets of Cα atoms and Cβ atoms of the putative zinc-coordinating residues were calculated. If these distances fall within the distance cut-off $\bar{x} \pm \sigma$, then the pair or triplet of residues became considered as possible zinc coordinating residues. Using this algorithm, 11 experimental SCR crystal or NMR structures for SCR-1/2, SCR-2/3, SCR-5, SCR-6/8, SCR-15/16 and SCR-19/20,[27,54-56] and nine homology-modelled SCR structures for SCR-4, SCR-9, SCR-10, SCR-11, SCR-12, SCR-13, SCR-14, SCR-17 and SCR-18[27] were screened.

Further predictions of metal binding sites in FH were carried out using web servers. First, the 20 experimental and homology SCR structures in FH were evaluated using METSITE (http://bioinf.cs.ucl.ac.uk/metsite)[58] which uses a set of neural network classifiers trained to identify potential cation sites. Second, to form FH dimer structures, the SCR-6/8 His402 crystal structure (PDB code 2UWN) was docked using four different docking approaches available as web servers including CLUSPRO (http://cluspro.bu.edu/login.php), HEX (http://www.csd.abdn.ac.uk/hex_server/), PATCHDOCK (http://bioinfo3d.cs.tau.ac.il/PatchDock/) and ROSETTADOCK (http://rosettadock.graylab.jhu.edu/).[75-79] These four docking programs were selected according to their performance in the Critical Assessment of Predicted Interactions surveys.[80,812]Third, the possible biological relevant dimeric structures were tested again by METSITE and another web server CHED [59] to detect the possible zinc binding sites at the protein interface of the dimeric structures. The CHED web server searches for Cys, His, Glu and Asp residues with atoms within several specified distances to each other and a 0.27 nm distance to the putative zinc atom position, and if the distances do not match, a independent rotamer library is used to find an appropriate fit configuration. The sites found in CHED are filtered by a decision tree and a support vector machine to verify their suitability. Fourth, the dimeric structures that possess zinc binding sites predicted by METSITE and CHED were clustered by using the program RMSDCLUST (Sean Ward and D. T. J, unpublished software) to group similar structures. RMSDCLUST carries out rigid body superposition between each pair of dimeric structures, computes the alpha-carbon root mean square deviation (RMSD), and then clusters all pairs of structures within an adjustable RMSD threshold (0.6 nm in this case).. To validate the methods, the same predictions were performed with three zinc binding structures as benchmarks, namely human interferon-β

(PDB code 1AU1), the complex of human growth hormone and prolactin (PDB code 1BP3) and thermophilic cytochrome P450 from *S. solfataricus* (PDB code 1F4T).[82-84]

ACKNOWLEDGEMENTS

We are very grateful to Prof. A. C. Bird for useful discussions, Dr A. Shukla and Dr T. Narayanan (European Synchrotron Radiation Facility, Grenoble, France) for excellent instrumental support, and Prof R. B. Sim for the MRC affinity columns. We thank the Biotechnology and Biological Sciences Research Council (R.N., A.M.), the Mercer Fund of the Fight for Sight Charity (R.N.) and the Wellcome Trust (I.F., F.S.) for graduate studentship and equipment grant support. I.L. thanks the Mercer Fund from Fight for Sight, the Special Trustees of Moorfields Eye Hospital and the Bill Brown Charitable Trust for support.

REFERENCES

1. Law, S. K. A. & Reid, K. B. M. (1995). Complement (Second edition). IRL Press, Oxford.
2. Walport, M. J. (2001). Complement. *N. Engl. J. Med.* **344**, 1058-1066 and 1140-1144.
3. Ferreira, V. P., Pangburn, M. K. & Cortés, C. (2010). Complement control protein factor H: The good, the bad, and the inadequate. *Molec. Immunol.* **47**, 2187-2197.
4. Soares, D. & Barlow, P. N. (2005). Complement control protein modules in the regulators of complement activators. In "Structural Biology of the Complement System" (Eds. Morikis, D & Lambris, J. D.) pp 19-62. Taylor & Francis, Boca Raton, USA.
5. Aslam, M. & Perkins, S. J. (2001). Folded-back solution structure of monomeric Factor H of human complement by synchrotron X-ray and neutron scattering, analytical ultracentrifugation and constrained molecular modelling. *J. Mol. Biol.* **309**, 1117-1138.
6. Okemefuna, A. I., Nan, R., Gor, J. & Perkins, S. J. (2009). Electrostatic interactions contribute to the folded-back conformation of wild-type human Factor H. *J. Mol. Biol*. **391**, 98-118.
7. Nan, R., Ward, G., Gavigan, L., Miller, A., Gor, J., McKay, A. R., Lengyel, I. & Perkins, S. J. (2010). The His402 allotype of complement Factor H show similar self-association to the Tyr402 allotype. *Molec. Immunol.* **47**, 2263-2263 (abstract).
8. Sharma, A. K. & Pangburn, M. K. (1996). Identification of three physically and functionally distinct binding sites for C3b in human complement factor H by deletion mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10996-11001.
9. Pangburn, M. K. (2000). Host recognition and target differentiation by factor H, a regulator of the alternative pathway of complement. *Immunopharmacology.* **49**,149-157.
10. Oppermann, M., Manuelian, T., Józsi, M., Brandt, E., Jokiranta, T. S., Heinen, S., Meri, S., Skerka, C., Götze, O. & Zipfel, P. F. (2006). The C-terminus of complement regulator Factor H mediates target recognition: evidence for a compact conformation of the native protein. *Clin. Exp. Immunol.* **144**, 342-352.
11. Schmidt, C. Q., Herbert, A. P., Kavanagh, D., Gandy, C., Fenton, C. J., Blaum, B. S., Lyon, M., Uhrín, D. & Barlow, P. N. (2008b). A new map of glycosaminoglycan and C3b binding sites on factor H. *J. Immunol.* **181**, 2610-2619.

12. Okemefuna, A. I., Nan, R., Miller, A. Gor, J. & Perkins, S. J. (2010). Complement Factor H binds at two independent sites to C-reactive protein in acute-phase concentrations. *J. Biol. Chem.* **285**, 1053-1065.

13. Horstmann, R. D., Sievertsen, H. J., Knobloch, J. & Fischetti, V. A. (1988). Antiphagocytic activity of streptococcal M protein: selective binding of complement control protein factor H. *Proc. Natl. Acad. Sci. U.S.A.*. **88**, 1657–1661.

14. Giannikis, E., Jokiranta, T. S., Male, D. A., Ranganathan, S., Ormsby, R. J., Fischetti, V. A., Mold, C. & Gordon, D. L. (2003). A common site within factor H SCR7 responsible for binding heparin, c-reactive protein and streptococcal M protein. *Eur. J. Immunol.* **33**, 962-969.

15. Schneider, M. C., Prosser, B. E., Caesar, J. J. E., Kugelberg, E., Li, S., Zhang, Q., Quoraishi, S., Lovett, J. E., Deane, J. E., Sim, R. B., Roversi, P., Johnson, S., Tang, C. M. & Lea, S. M. (2009). *Neisseria meningitidis* recruits factor H using protein mimicry of host carbohydrates. *Nature* **458**, 890-893.

16. Fernando, A. N., Furtado, P. B., Clark, S. J., Gilbert, H. E., Day, A. J., Sim, R. B. & Perkins, S. J. (2007). Associative and structural properties of the region of complement Factor H encompassing the Tyr402His disease-related polymorphism and its interactions with heparin. *J. Mol. Biol.* **368**, 564-581.

17. Okemefuna, A. I., Gilbert, H. E., Griggs, K. M., Ormsby, R. J., Gordon, D. L. & Perkins, S. J. (2008). The regulatory SCR-1/5 and cell-surface-binding SCR-16/20 fragments of Factor H reveal partially folded-back solution structures and different self-associative properties. *J. Mol. Biol.* **375**, 80-101.

18. Nan, R., Gor, J. & Perkins, S. J. (2008). Implications of the progressive self-association of wild-type human Factor H for complement regulation and disease. *J. Mol. Biol.* **375**, 891-900.

19. Perkins, S. J., Nealis, A. S. & Sim, R. B. (1991). Oligomeric domain structure of human complement factor H by X-ray and neutron solution scattering. *Biochemsitry* **30**, 2847-2857.

20. Nan. R., Gor, J., Lengyel, I., Perkins, S. J. (2008). Uncontrolled zinc- and copper-induced oligomerisation of the human complement regulator factor H and its possible implications for function and disease. *J. Mol. Biol.* **384**,1341-1352.

21. Hageman, G. S., Luthert, P. J., Victor Chong, N. H., Johnson, L. V., Anderson, D. H. & Mullins, R. F. (2001). An integrated hypothesis that considers drusen as biomarkers

of immune-mediated processes at the RPE-Bruch's membrane interface in aging and age-related macular degeneration. *Progr. Retin. Eye Research,* **20**, 705-732.

22. Klein, R. J., Zeiss, C., Chew, E. Y., Tsai, J. Y., Sackler, R. S., Haynes, C., Henning, A. K., Sangiovanni, J. P., Mane, S. M., Mayne, S. T., Bracken, M. B., Ferris, F. L., Ott, J., Barnstable, C. & Hoh, J. (2005). Complement factor H polymorphism in age-related macular degeneration. *Science,* **308**, 385-389.
23. Haines, J. L., Hauser, M. A., Schmidt, S., Scott, W. K., Olson, L. M., Gallins, P., Spencer, K. L., Kwan, S. Y., Noureddine, M., Gilbert, J. R., Schnetz-Boutaud, N., Agarwal, A., Postel, E. A. & Pericak-Vance, M. A. (2005). Complement factor H variant increases the risk of age-related macular degeneration. *Science,* **308**, 419-421.
24. Edwards, A. O., Ritter, R. III, Abel, K. J., Manning, A., Panhuysen, C. & Farrer, L. A. (2005). Complement factor H polymorphism and age-related macular degeneration. *Science,* **308**, 421-424.
25. Hageman, G. S., Anderson, D. H., Johnson, L. V., Hancox, L. S., Taiber, A. J., Hardisty, L. I., Hageman, J. L., Stockman, H. A., Borchardt, J. D., Gehrs, K. M., Smith, R. J., Silvestri, G., Russell, S. R., Klaver, C. C., Barbazetto, I., Chang, S., Yannuzzi, L. A., Barile, G. R., Merriam, J. C., Smith, R. T., Olsh, A. K., Bergeron, J., Zernant, J., Merriam, J. E., Gold, B., Dean, M. & Allikmets, R. (2005). A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7227-7232.
26. Strohmeyer, R., Ramirez, M., Cole, G. J., Mueller, K. & Rogers, J. (2002). Association of factor H of the alternative pathway of complement with agrin and complement receptor 3 in the Alzheimer's disease brain. *J. Neuroimmunol.* **131**, 135-146.
27. Saunders, R. E. Abarrategui-Garrido, C., Frémeaux-Bacchi, V., Goicoechea de Jorge, E., Goodship, T. H. J., López Trascasa, M., Noris, M., Ponce Castro, I. M., Remuzzi, G., Rodríguez de Córdoba, S., Sánchez-Corral, P., Zipfel, P. F. & Perkins, S. J. (2007). The interactive factor H-atypical haemolytic uraemic syndrome mutation database and website: Update and integration of membrane cofactor protein and factor I mutations with structural models. *Hum. Mutat.* **28**, 222-234.
28. Zetterberg, M., Landgren, S., Andersson, M. E., Palmér, M. S., Gustafson, D. R., Skoog, I., Minthon, L., Thelle, D. S., Wallin, A., Bogdanovic, N., Andreasen, N., Blennow, K. & Zetterberg, H. (2008). Association of complement factor H Y402H

gene polymorphism with Alzheimer's disease. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **147**, 720-726.

29. Perkins, S. J., Okemefuna, A. I. & Nan, R. (2010). Unravelling protein-protein interactions between complement factor H and C-reactive protein by a multidisciplinary strategy. *Biochem. Soc. Transact*. **38**, 894-900.
30. Perkins, S. J., Nan, R., Okemefuna, A. I., Li, K., Khan, S. & Miller, A. (2010). Multiple interactions of complement factor H with its ligands in solution: a progress report. Current Topics on Complement and Eye Diseases (Ed. J.D. Lambris and A. Adamis). *Adv. Exp. Med. Biol*. **703**, 25-47.
31. Bird, A. C. (1992). Bruch's membrane changes with age. *Brit. J. Ophthalmol.* **76**, 166-168.
32. Bird, A. C., Bressler, N. M., Bressler, S. B., Chisholm, I. H., Coscas, G., Davis, M. D., de Jong, P. T., Klaver, C. C., Klein, B. E., Klein, R., *et al.* (1995). An international classification and grading system for age-related maculopathy and age-related macular degeneration. *Surv. Ophthalmol.* **39**, 367-374.
33. Guymer, R. & Bird, A. C. (1998). Bruch's membrane, drusen, and age-related macular degeneration. In: Marmor M, Wolfensberger T (eds) The retinal pigment epithelium. Oxford University Press, Oxford, 693-705.
34. Crabb, J. W., Miyagi, M., Gu, X., Shadrach, K., West, K. A., Sakaguchi, H., Kamei, M., Hasan, A., Yan,_L., Rayborn, M. E., Salomon, R. G. & Hollyfield, J. G. (2002). Drusen proteome analysis: An approach to the etiology of age-related macular degeneration. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14682-14687.
35. Bok, D. (2005). Evidence for an inflammatory process in age-related macular degeneration gains new support. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7053-7054.
36. Anderson, D. H., Mullins, R. F., Hageman, G. S. & Johnson, L. V. (2002). A role for local inflammation in the formation of drusen in the aging eye. *Am. J. Ophthalmol.* **134**, 411-431.
37. Skerka, C., Lauer, N., Weinberger, A. A., Keilhauer, C. N., Sühnel, J., Smith, R., Schlötzer-Schrehardt, U., Fritsche, L., Heinen, S., Hartmann, A., Weber, B. H. & Zipfel, P. F. (2007). Defective complement control of factor H (Y402H) and FHL-1 in age-related macular degeneration. *Mol. Immunol.* **44**, 3398-3406.
38. Clark, S. J., Higman, V. A., Mulloy, B., Perkins, S. J., Lea, S. M., Sim, R. B. & Day, A. J. (2006). H384 allotypic variant of factor H associated with age-related macular

degeneration has different heparin-binding properties from the non-disease associated form. *J. Biol. Chem.* **281**, 24713-24720.

39. Galin, M. A., Nano, H. D. & Hall, T. (1962). Ocular zinc concentration. *Investig. Ophthalmol.* **1**, 142-148.

40. Ugarte, M. & Osborne, N. N. (2001). Zinc in the retina. *Prog. Neurobiol.* **64**, 219-249.

41. Lengyel, I., Flinn, J. M., Peto, T., Linkous, D. H., Cano, K., Bird, A. C., Lanzirotti, A., Frederickson, C. J. & van Kuijk, F. J. G. M. (2007). High concentration of zinc in sub-retinal pigment epithelial deposits. *Exp. Eye Res.* **84**, 772-780.

42. Age-Related Eye Disease Study Research Group. (2001). A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. *Arch Ophthalmol.* **119**, 1417-1436.

43. Age-Related Eye Disease Study Research Group. (2001). A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E and beta carotene for age-related cataract and vision loss: AREDS report no. 9. *Arch Ophthalmol.* **119**, 1439-1452.

44. Klein, R. J., Zeiss, C., Chew, E. Y., Tsai, J. Y., Sackler, R. S., Haynes, C., Henning, A. K., Sangiovanni, J. P., Mane, S. M., Mayne, S. T., Bracken, M. B., Ferris, F. L., Ott, J., Barnstable, C. & Hoh, J. (2005). Complement factor H polymorphism in age-related macular degeneration. *Science,* **308**, 385-389.

45. Auld, D. S. (2001). Zinc coordination sphere in biochemical zinc sites. *Biometals.* **14**, 271-313.

46. Cole, J. L., Lary, J. W., Moody, T. P. & Laue, T. M. (2008). Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Meth. Cell Biol.* **84**, 143-211.

47. Nieba, L., Nieba-Axmann, S. E., Persson, A., Hämäläinen, M., Edebratt, F., Hansson, A., Lidholm, J., Magnusson, K., Karlsson, Å. F. & Plückthun, A. (1997). BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. *Anal. Biochem.* **252**, 217-228.

48. Perkins, S. J., Okemefuna, A. I., Fernando, A. N., Bonner, A., Gilbert, H. E. & Furtado, P. B. (2008). X-ray and neutron scattering data and their constrained molecular modelling. *Meth. Cell Biol.* **84**, 375-423.

49. Glatter, O. & Kratky, O. (Editors) (1982). *Small angle X-ray scattering.* Academic Press, New York.

50. Crossley, L. G. & Porter, R. R. (1980). Purification of the human complement control protein C3b inactivator. *Biochem. J.* **191**, 173-182.

51. Day, A. J. & Sim, R. B. (1986). Inhibitory effect of Zn^{2+} ion on the degradation of the complement activation fragment C3b. *Biochem. Soc. Transact.* **14**, 73-74.

52. Blom, A. M., Kask, L., Ramesh, B. & Hillarp, A. (2003). Effects of zinc on factor I cofactor activity of C4b-binding protein and factor H. *Arch. Biochem. Biophys.* **418**, 108-118.

53. Tsiftsoglou, S. A. & Sim, R. B. (2004). Human complement factor I does not require cofactors for cleavage of synthetic substrates. *J. Immunol.* **173**, 367-375.

54. Barlow, P. N., Steinkasserer, A., Norman, D. G. , Kieffer, B., Wiles, A. P., Sim, R. B. & Campbell, I. D. (1993). Solution structure of a pair of complement modules by nuclear magnetic resonance. *J. Mol. Biol.* **232**, 268-284.

55. Jokiranta, T. S., Jaakola, V. P., Lehtinen, M. J., Parepalo, M., Meri, S. & Goldman, A. (2006). Structure of complement factor H carboxyl-terminus reveals molecular basis of atypical haemolytic uremic syndrome. *EMBO J.* **25**, 1784-1794.

56. Prosser, B. E., Johnson, S., Roversi, P., Herbert, A. P., Blaum, B. S., Tyrrell, J., Jowitt, T. A., Clark, S. J., Tarelli, E., Uhrin, D., Barlow, P. N., Sim, R. B., Day, A. J. & Lea, S. M. (2007). Structural basis for complement factor H – linked age-related macular degeneration. *J. Exp. Med.* **204**, 2277-2283.

57. Hocking, H. G., Herbert, A. P., Kavanagh, D., Soares, D. C., Ferreira, V. P., Pangburn, M. K., Uhrín, D. & Barlow, P. N. (2008). Structure of the N-terminal region of complement factor H and conformational implications of disease-linked sequence variations. *J. Biol. Chem.* **283,** 9475-9487.

58. Sodhi J.S., Bryson K., McGuffin L.J., Ward J.J., Wernisch L. & Jones D.T. (2004). Predicting metal-binding site residues in low-resolution structural models. *J. Mol. Biol.* **342**, 307-320.

59. Babor M., Gerzon S., Raveh B., Sobolev V. & Edelman M. (2008). Prediction of transition metal-binding sites from apo protein structures. *Proteins,* **70**, 208-217.

60. An, E., Lu, X., Flippin, J., Devaney, J. M., Halligan, B., Hoffman, E. P., Strunnikova, N., Csaky, K. & Hathout, Y. (2006). Secreted proteome profiling in human RPE cell cultures derived from donors with age related macular degeneration and age matched healthy donors. *J. Proteome Res.* **5**, 2599-2610.

61. Okemefuna, A. I., Li, K., Nan, R., Ormsby, R. J., Sadlon, T., Gordon, D. L. & Perkins, S. J. (2009). Multimeric interactions between complement Factor H and its C3d ligand provide new insight on complement regulation. *J. Mol. Biol.* **391**, 119-135.

62. Lu, J., Stewart, A. J., Sadler, P. J., Pinheiro, T. J. & Blindauer, C. A. (2008). Albumin as a zinc carrier: properties of its high-affinity zinc-binding site. *Biochem. Soc. Trans.* **36**, 1317-1321.

63. Age-Related Eye Disease Study Research Group. (2002). The effect of five-year zinc supplementation on serum zinc, serum cholesterol and hematocrit in persons randomly assigned to treatment group in the Age-Related Eye Disease Study. *J. Nutr.* **132**, 687–702.

64. Sugio, S., Kashima, A., Mochizuki, S., Noda, M. & Kobayashi, K. (1999). Crystal structure of human serum albumin at 2.5 A resolution. *Protein Eng.* **12,** 439-446.

65. Newsome, D. A., Oliver, P. D., Deupree, D. M., Miceli, M. V. & Diamond, J. G. (1992). Zinc uptake by primate retinal pigment epithelium and choroid. *Curr. Eye Res.* **11**, 213-217.

66. Lengyel, I. & Peto, T. (2008). Cure or cause: the opposing roles for zinc in age-related macular degeneration. *Expert Rev. Ophthalmol.* **3**, 1-4.

67. Sim, R. B., Day, A. J., Moffatt, B. E. & Fontaine, M. (1993). Complement factor I and cofactors in control of complement system convertase enzymes. *Meth. Enzymol.* **223**, 13-35.

68. Dodds, A. W. (1993). Small-scale preparation of complement components C3 and C4. *Meth. Enzymol.* **223**, 46-61.

69. Perkins, S. J. & Sim, R. B. (1986). Molecular modelling of human complement component C3 and its fragments by solution scattering. *Eur. J. Biochem.* **157**, 155-168

70. Ullman, C. G., Chamberlain, D., Emery, V. C., Haris, P. I., Sim, R. B. & Perkins, S. J. (1998). Human complement factor I: its expression by insect cells and its biochemical and structural characterisation. *Mol. Immunol.* **35**, 503-512.

71. Schuck, P. (1998). Sedimentation analysis of non-interacting and self-associating solutes using numerical solutions to the Lamm equation. *Biophys. J.* **75**, 1503-1512.

72. Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606-1619.

73. Narayanan, T., Diat, O. & Bosecke, P. (2001). SAXS and USAXS on the high brilliance beamline at the ESRF. *Nucl. Instrum. Methods Phys. Res. A*, **467-468**, 1005-1009

74. Semenyuk, A. V. & Svergun, D. I. (1991). GNOM - a program package for small-angle scattering data-processing. *J. Appl. Crystallogr.* **24,** 537-540.

75. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. (2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics,* **20**, 45-50.

76. Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A. & Baker, D. (2003). Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol.* **331**, 281-299.

77. Ritchie, D.W. (2003). Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. *Proteins,* **52**, 98-106.

78. Schneidman-Duhovny D., Inbar Y., Nussinov R. & Wolfson H. J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**, W363-367.

79. Gray, J. J. (2006). High-resolution protein-protein docking. *Curr. Opin. Struct. Biol*. **16**, 183-193.

80. Mendez, R., Leplae, R., Lensink, M. F. & Wodak, S. J. (2005). Assessment of CAPRI predictions in rounds 3-5 shows progress in docking procedures. *Proteins,* **60**, 150-169.

81. Lyskov, S. & Gray, J. J. (2008). The RosettaDock server for local protein-protein docking. *Nucleic Acids Res.* **36**, W233-238.

82. Karpusas, M., Nolte, M., Benton, C. B., Meier, W., Lipscomb, W. N. & Goelz, S. (1997). The crystal structure of human interferon beta at 2.2-A resolution. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 11813-11818.

83. Somers. W., Ultsch. M., De Vos. A. M. & Kossiakoff. A. A. (1994). The X-ray structure of a growth hormone-prolactin receptor complex. *Nature,* **372**, 478-481.

84. Yano, J. K., Koo, L. S., Schuller, D. J., Li, H., Ortiz de Montellano, P. R. & Poulos, T. L. (2000). Crystal structure of a thermophilic cytochrome P450 from the archaeon Sulfolobus solfataricus. *J. Biol. Chem.* **275**, 31086-31092.

Table 1. Prediction of partial zinc binding sites in FH.

(a) Distribution of predicted zinc coordination sites in 20 SCR domains. [a]

Domain	SCR-1	SCR-2	SCR-3	SCR-4	SCR-5
Structure	2RLP	2RLP, 2RLQ	2RLQ	homology	NMR structure
Sites	6	14.5	23.5	2	0

Domain	SCR-6	SCR-7	SCR-8	SCR-9	SCR 10
Structure	2UWN	2UWN	2UWN	homology	homology
Sites	62.5	8.5	1	5	4

Domain	SCR 11	SCR 12	SCR 13	SCR 14	SCR-15
Structure	homology	homology	homology	homology	1HFH
Sites	14	9	8	4	7.5

Domain	SCR-16	SCR 17	SCR 18	SCR-19	SCR-20
Structure	1HFH	homology	homology	2G7I	2G7I
Sites	19.5	0	1	7.5	4.5

(b) Possible predicted zinc binding sites in the FH fragment SCR-6/8. [b]

Residues	**Location**	**Possible zinc coordination**
D326-D329	SCR-6	DDH, DDHH
D329-H332	SCR-6	DDHH, DHHH
H332-D358	SCR-6	DDDH, DDHH, DEHH, DH, DHH, DHHH
H332-D358-E359	SCR-6	DEHH
H332-D358-H360	SCR-6	DDHH, DEHH, DHH, DHHH
H332-D358-E362	SCR-6	DDHH, DEHH
H332-E359	SCR-6	EH, EHH, EHHH
H332-H360	SCR-6	DDHH, DEHH, DHH, DHHH, EHH, HH, HHH
H332-H360-E362	SCR-6	DEHH
H332-H360-E433	SCR-6/7	DEHH
H332-E362	SCR-6	DEHH, EHH, EHHH
H332-E433	SCR-6/7	EHHH, DEH, DEHH
H337-E338	SCR-6	EH, EHHH
D358-E359	SCR-6	DE
D358-H360	SCR-6	DDHH, DH, DHH, DHHH
D358-H360-E362	SCR-6	DEHH
H360-E362	SCR-6	DEHH, EEEH, EEHH, EH, EHHH
H360-E433	SCR-6/7	EHH, DEHH, EHHH
D370-H371	SCR-6	DDHH, DEHH, DHH
D370-H371-H373	SCR-6	DDHH, DHH, DHHH
D370-H373	SCR-6	DDHH, DHH, DHHH
H371-H373	SCR-6	DHHH, EEHH, EHH, HHH, HHHH
H373-D377	SCR-6	DHHH, DDHH
E395-H402	SCR-7	EHH, EHHH
E395-H417	SCR-7	DEHH, EHH,EHHH
D454-E456	SCR-8	DDH, DDE

[a] Zinc coordination sites that are shared between two adjacent SCR domains are counted as 0.5 per domain.

[b] His402 occurred in 2 of the 76 possible zinc coordinations in SCR-6/8. Zinc coordinations that contain two or three residues also contain two or one H_2O molecules respectively to form the required tetrahedral geometry with zinc.

Table 2. The most likely zinc-interaction ligands in FH SCR-6/8 and three protein benchmarks.

	FH SCR-6/8 (PDB code 2UWN)	Human interferon β (PDB code 1AU1)	Complex of human growth hormone and prolactin (PDB code 1BP3)	Archael cytochrome P450 (PDB code 1F4T)
Top 3	Glu359	Glu42	His18 *	Asp52
	His360	His93 *	Glu174 *	His340
	His371	His121 *	Asp387 *	Glu342
Top 5	Asp370	Glu85	His21	Asp42
	Asp485	His131	His388 *	His178 *
Top 10	His332	Glu81	Glu56	Glu48
	His417	His97 *	Asp171	Asp52
	Glu462	Glu103	Glu245	Glu139 *
	Glu487		Glu292	Glu274
	Asp497		Glu345	Asp294

* The original ligands in the three benchmarked structures 1AU1, 1BP3 and 1F4T are asterisked.

FIGURE LEGENDS

Figure 1. Sedimentation velocity analyses of the two FH allotypes titrated with zinc. Both allotypes were studied at 0.7 mg/ml. In the sedimentation boundary fits, only every tenth scan is shown for clarity.

(a-d) The fits corresponding to the FH Tyr402 allotype (026II; blue) was titrated with zinc concentrations of 20 μM, 60 μM, 120 μM and 200 μM in that order.

(f-i) The fits corresponding to the FH His402 allotype (027BA; red) was titrated with the same zinc concentrations.

(e,j) The *c(s)* sedimentation coefficient distribution analyses from (a-d) and (f-i) are shown for the FH Tyr402 and FH His402 allotypes respectively. The height of the FH monomer peak at 5.7 S is normalised to 100 for clarity. The zinc concentrations (μM) are denoted numerically.

Figure 2. Dependence of the monomer and oligomers of FH and SCR-6/8 on zinc concentrations.

(a) The values from *c(s)* integrations correspond to the averages obtained at 50,000 r.p.m. and 60,000 r.p.m. The oligomer data were fitted to a two-parameter power function $y=a*(1 + x)^b$, and the monomer data were fitted to the function $f=y0+a*x^b$. Oligomers are denoted by filled symbols and monomers by open symbols for four experiments with two FH Tyr402 (026II, 022LM, blue, ●, ○, ▲, Δ) and two FH His402 (027BA, 032KO, red, ●, ○, ▲, Δ) allotypes.

(b) The proportion of monomer of the Tyr402 (blue) and His402 (red) allotypes of SCR-6/8 are shown.

Figure 3. Sedimentation velocity analyses of the two SCR-6/8 allotypes titrated with zinc. Both allotypes were studied at 0.3 mg/ml. In the sedimentation boundary fits, only every tenth scan is shown for clarity.

(a,b) The fits for the SCR-6/8 Tyr402 allotype (blue) titrated with zinc at 6 μM and 200 μM.

(c,d) The fits for the SCR-6/8 His402 allotype (red) titrated with zinc at 6 μM and 200 μM.

(e) The *c(s)* sedimentation coefficient distribution analyses for the Tyr402 allotype with 6 μM (blue) and 200 μM (cyan) zinc, and the His402 allotype with 6 μM (red) and 200 μM (pink) zinc. The *c(s)* distributions are labelled with their zinc concentration values

(μM). The peak at about 2.1 S corresponds to the monomer, and that at about 4.0 S is the dimer.

Figure 4. The effect of EDTA on the interactions between the SCR-6/8 allotypes and zinc using sedimentation velocity.

(a) The *c(s)* distribution analyses are shown for the Tyr402 allotype in the presence of 0.2 μM zinc (blue solid line), 300 μM zinc (cyan solid line), and 300 μM zinc and 2 mM EDTA (cyan dashed line).

(b) The *c(s)* distribution analyses are shown for the His402 allotype in the presence of 0.2 μM zinc (red solid line), 300 μM zinc (pink solid line), and 300 μM zinc and 2 mM EDTA (pink dashed line).

Figure 5. Sedimentation velocity analysis of SCR-1/5 and SCR-16/20 titrated with zinc. Red numbers indicate zinc concentrations (μM). (a-c) For SCR-1/5, the boundary fits at three zinc concentrations are shown for only every 24^{th} scan out of 480 for reason of clarity. (d) Size-distribution *c(s)* analyses for SCR-1/5 at seven zinc concentrations are shown. M indicates the monomer position, and two weak emerging zinc-induced signals are denoted by 1 and 2. The graphs were displaced vertically in units of 0.5 for reason of clarity. (e-g) For SCR-16/20, the boundary fits at three zinc concentrations are shown for only every 22^{nd} scan out of 440. (h) Size-distribution *c(s)* analyses for SCR-16/20 at seven zinc concentrations are shown. M and D indicates the monomer and dimer peaks. Weak emerging signals are denoted by 1, 2 and 3. Graphs are displaced vertically in units of 0.5 for reason of clarity.

Figure 6. X-ray Guinier analyses of the full-length FH and SCR-6/8 allotypes titrated with zinc. In all panels, the open symbols correspond to the experimental data and the filled symbols correspond to those used for the R_G and R_{XS} straight line fits. Data correspond to one FH Tyr402 allotype (018DK) in (a,c,e) and one FH His402 allotype (030TJ) in (b,d,f), both at 0.9 mg/ml, and SCR-6/8 Tyr402 (g,i) and SCR-6/8 His402 (h,j), both at 0.2 mg/ml. Data for the Tyr402 allotype of FH and SCR-6/8 are shown in blue from 0.2 μM to 120 μM of $ZnSO_4$, and in cyan from 200 μM to 600 μM of $ZnSO_4$. Data for the His402 allotype of FH and SCR-6/8 are shown in red from 0.2 μM to 120 μM of $ZnSO_4$, and in pink from 200 μM to 600 μM of $ZnSO_4$.

(a,b) Guinier R_G plots of ln $I(Q)$ vs. Q^2 for FH Tyr402 and FH His402 allotypes titrated with $ZnSO_4$ concentrations of 0.2 μM (O), 2 μM (□), 6 μM (Δ), 20 μM (▽), 60 μM (◇), 120 μM (⎔), 200 μM (O) and 600 μM (Δ). The Q fit range was 0.08 - 0.13 nm^{-1} for 0.2 μM to 120 μM $ZnSO_4$, and 0.03 - 0.16 nm^{-1} for 200 μM to 600 μM $ZnSO_4$. The zinc concentrations are numerically labelled as shown.

(c,d) The corresponding Guinier cross-sectional R_{XS-1} fits of ln $I(Q).Q$ vs. Q^2 for the two FH allotypes with zinc using a Q range of 0.16 - 0.26 nm^{-1}.

(d,e) The corresponding Guinier cross-sectional R_{XS-2} fits for the two FH allotypes with zinc using a Q range of 0.4 - 0.8 nm^{-1}.

(g,h) Guinier R_G plots for the SCR-6/8 Tyr402 and His402 allotypes at 0.2 mg/ml titrated with $ZnSO_4$ at concentrations of 0 μM (O), 2 μM (□), 6 μM (Δ), 20 μM (▽), 60 μM (◇), 120 μM (⎔), 200 μM (O), 300 μM (□) and 600 μM (Δ). The Q fit range was 0.16 - 0.5 nm^{-1} for 0.2 μM to 120 μM $ZnSO_4$, and 0.06 - 0.11 nm^{-1} for 200 μM to 600 μM $ZnSO_4$.

(i,j) The corresponding Guinier cross-sectional R_{XS} fits for the SCR-6/8 allotypes using a Q range of 0.55 – 1.0 nm^{-1}.

Figure 7. Dependence of the R_G, $I(0)/c$ and R_{XS} values of FH and SCR-6/8 on the zinc concentration. In all cases, the Tyr402 allotype is shown in blue and the His402 allotype is shown in red. Each value was measured in quadruplicate and averaged, and statistical error bars are shown where visible. In (a-d), the four titrations correspond to two pairs of FH Tyr402 allotypes (018DK, 016CC) and FH His402 allotypes (030TJ, 033NV) at 0.9 mg/ml titrated with zinc from 0.2 to 600 μM (each pair shown as dashed and continuous lines). In (e-g), the titrations correspond to the two SCR-6/8 allotypes titrated with the same zinc concentration range from 0.2 to 600 μM.

Figure 8. Dependence of the distance distribution function $P(r)$ of the FH allotypes on zinc concentration. The $P(r)$ curves were calculated from the scattering curves of the two FH Tyr402 allotypes (018DK, 016CC; blue) and two FH His402 allotypes (030TJ, 033NV; red). From bottom to top, the zinc concentrations were 6 μM, 20 μM, 60 μM and 120 μM (labelled). The average value of the most frequently occurring distance M for each allotype at each zinc concentration is marked by vertical straight lines.

Figure 9. Cleavage of fluid phase C3u by FI and the two FH allotypes in the presence of zinc.

(a) Reducing SDS-PAGE analysis of C3u cleavage by FI and the FH Tyr402 allotype (013HH) in the absence of zinc. Lane 1, High Mark™ Prestained High Molecular Weight Standard; Lane 2, 0.3 mg/ml C3u; Lane 3, 0.3 mg/ml C3u and 0.3 mg/ml FH; Lanes 4-15 correspond to Lane 3 with 0.003 mg/ml factor I added with the reaction times in min as labelled. FH and the α-chain and β-chain of C3u are arrowed on the left, and the α-chain cleavage products are arrowed on the right.

(b) Reducing SDS-PAGE analysis of C3u cleavage by FI and the FH Tyr402 allotype (013HH) with 200 μM zinc. Lane 1, 0.3 mg/ml C3u; Lane 2, 0.3 mg/ml C3u and 0.3 mg/ml FH; Lane 3, 0.3 mg/ml C3u, 0.3 mg/ml FH and 200 μM of zinc; Lanes 4 – 15 correspond to Lane 3 with the reaction times in min as labelled.

(c) The percentage cleavage of the C3 α-chain by FI and the FH Tyr402 and His402 allotypes are shown as a function of reaction time. The [Zn] values are labelled as shown. The Tyr402 allotypes are in blue (013HH, ●; 022LM, ▲) and the His402 allotypes are in red (015LT, ○; 032KO, Δ).

Figure 10. Predicted structural models for zinc binding to SCR-6/8. The two most likely predicted SCR-6/8 zinc sites from docking simulations are shown in (a,b), and a cartoon of the resulting putative zinc-induced SCR-6/8 aggregation is shown in (c).

(a) A predicted zinc binding site is composed of two Glu359 and His360 sidechains (Table 2) at the surface interface between two SCR-6 domains in the docked dimer. The left panel shows a close-up view of the zinc binding residues, whilst the right panel shows the overall structure of the putative dimer. The distances between each zinc ligand atom (nitrogen or oxygen) are displayed in nm.

(b) A second predicted zinc binding site is composed of a His371 sidechain in SCR-6 of the first SCR-6/8 monomer and the Asp485 and Glu487 sidechains in SCR-8 of the second SCR-6/8 monomer (Table 2). Other details follow (a).

(c) A mechanistic model for the zinc-induced aggregation of SCR-6/8 based on the two possible zinc binding sites shown in (a) and (b) shows how the two zinc sites may form a daisy-chain of SCR-6/8 that leads to aggregation and precipitation.

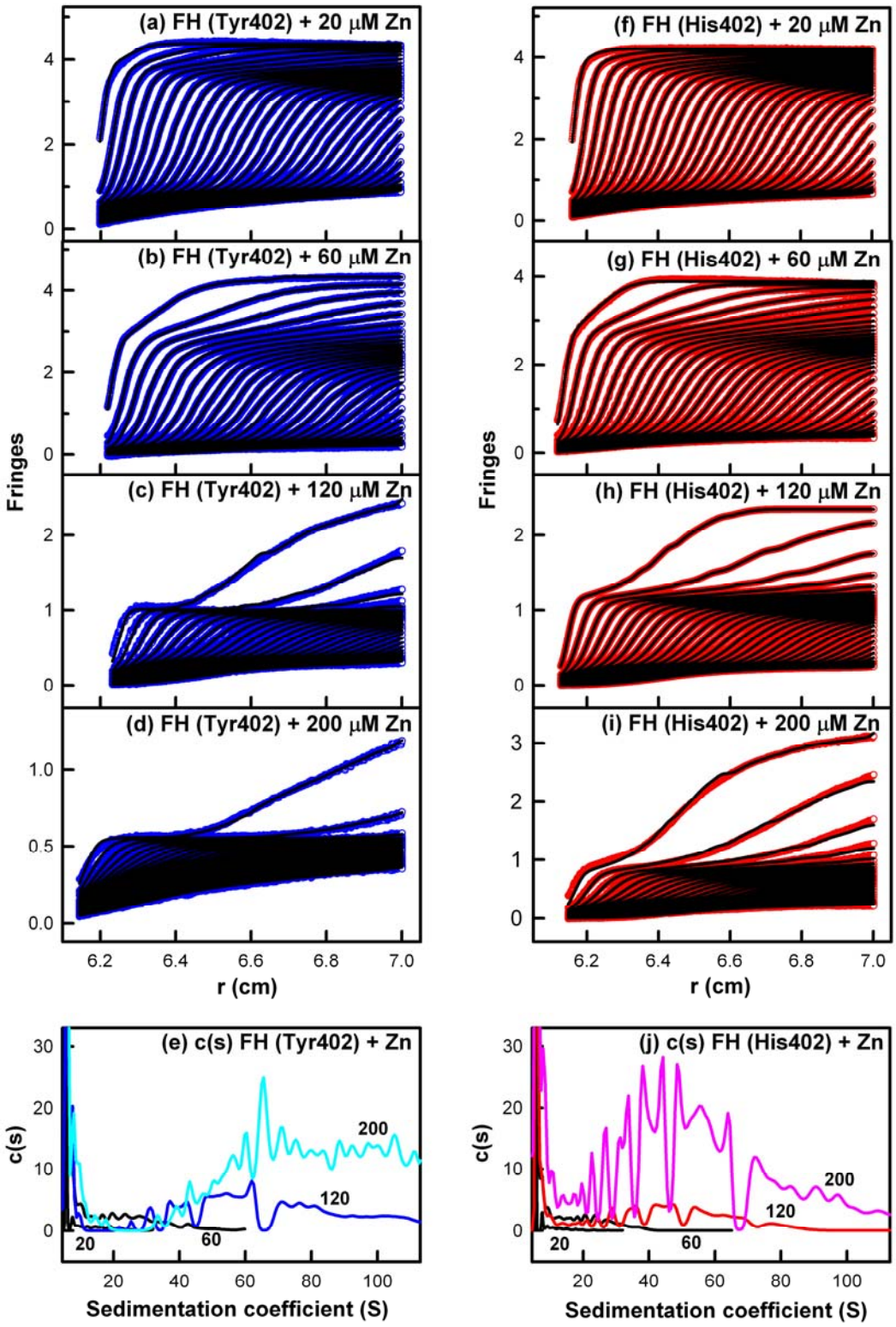
(a) FH (Tyr402) + 20 μM Zn
(b) FH (Tyr402) + 60 μM Zn
(c) FH (Tyr402) + 120 μM Zn
(d) FH (Tyr402) + 200 μM Zn
(f) FH (His402) + 20 μM Zn
(g) FH (His402) + 60 μM Zn
(h) FH (His402) + 120 μM Zn
(i) FH (His402) + 200 μM Zn
Fringes
r (cm)
(e) c(s) FH (Tyr402) + Zn
(j) c(s) FH (His402) + Zn
c(s)
Sedimentation coefficient (S)
20
60
120
200

Figure 1

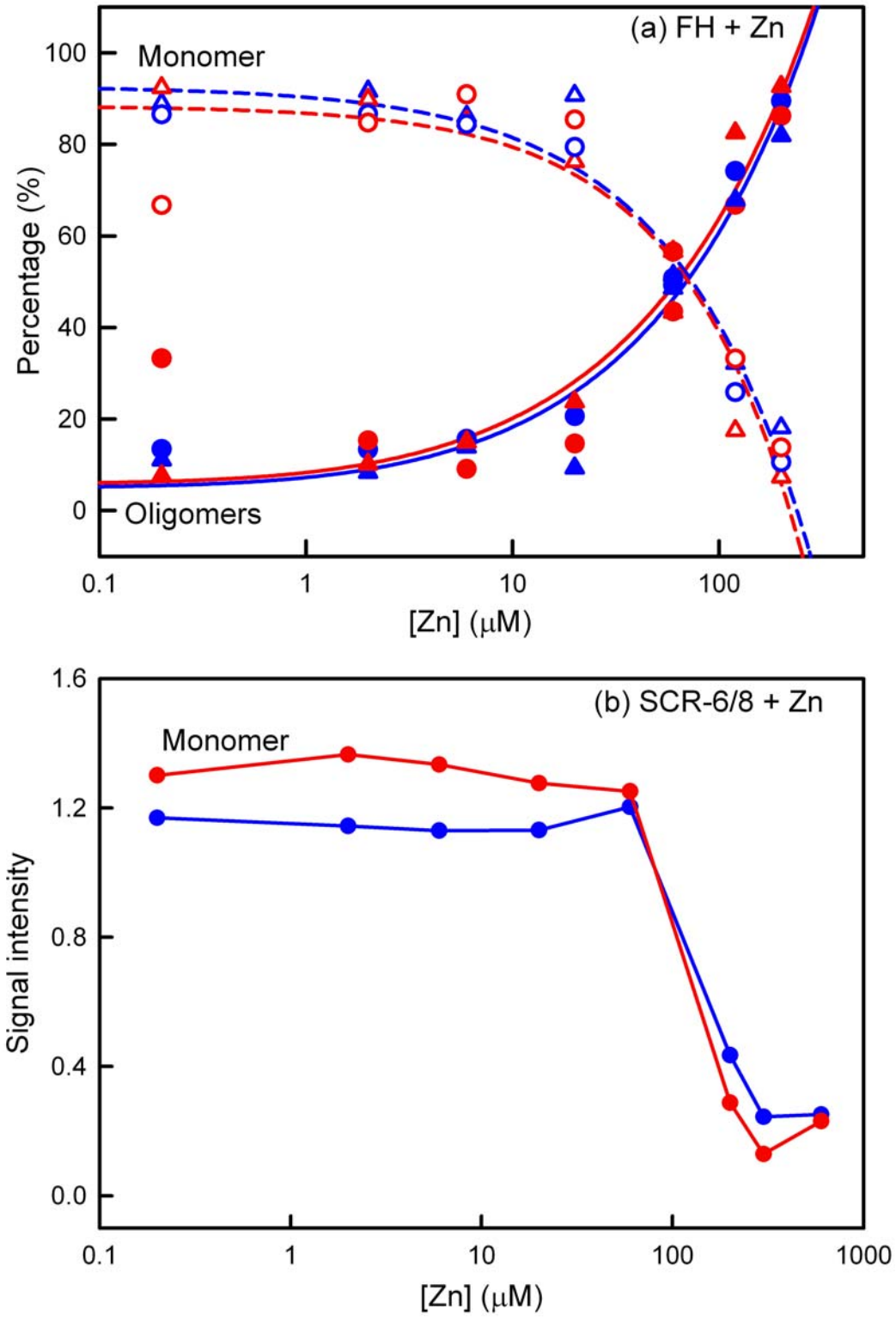
(a) FH + Zn
Monomer
Oligomers
Percentage (%)
100
80
60
40
20
0
0.1
1
10
100
[Zn] (μM)
(b) SCR-6/8 + Zn
Monomer
Signal intensity
1.6
1.2
0.8
0.4
0.0
0.1
1
10
100
1000
[Zn] (μM)

Figure 2

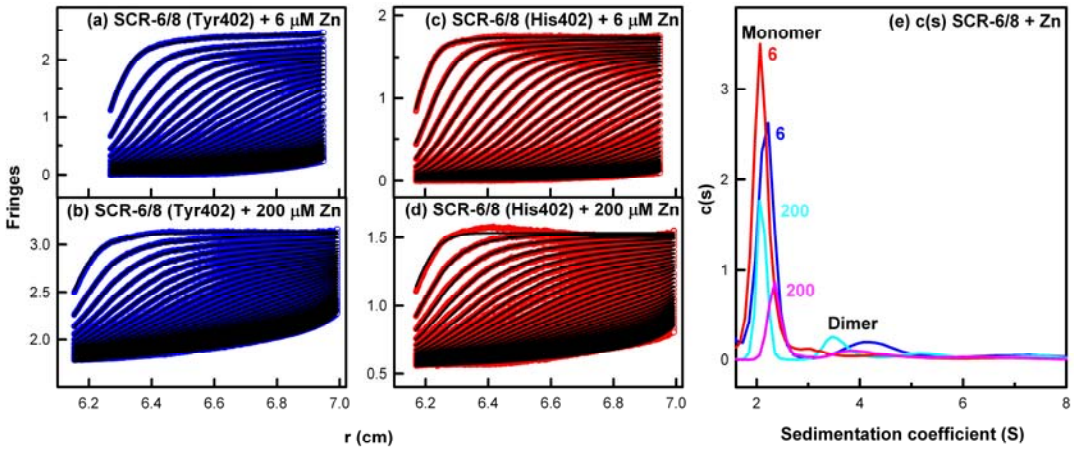
(a) SCR-6/8 (Tyr402) + 6 μM Zn
(b) SCR-6/8 (Tyr402) + 200 μM Zn
(c) SCR-6/8 (His402) + 6 μM Zn
(d) SCR-6/8 (His402) + 200 μM Zn
(e) c(s) SCR-6/8 + Zn
Monomer
6
6
200
200
Dimer
Fringes
r (cm)
c(s)
Sedimentation coefficient (S)

Figure 3

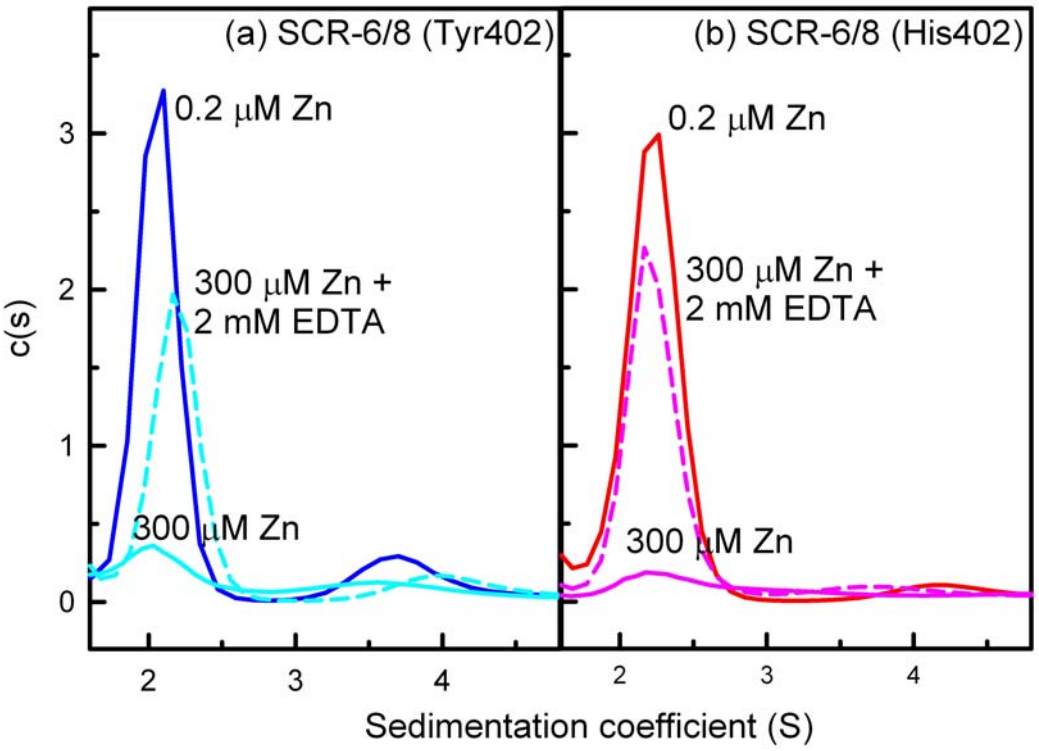
(a) SCR-6/8 (Tyr402)
(b) SCR-6/8 (His402)
0.2 μM Zn
300 μM Zn +
2 mM EDTA
300 μM Zn
c(s)
Sedimentation coefficient (S)

Figure 4

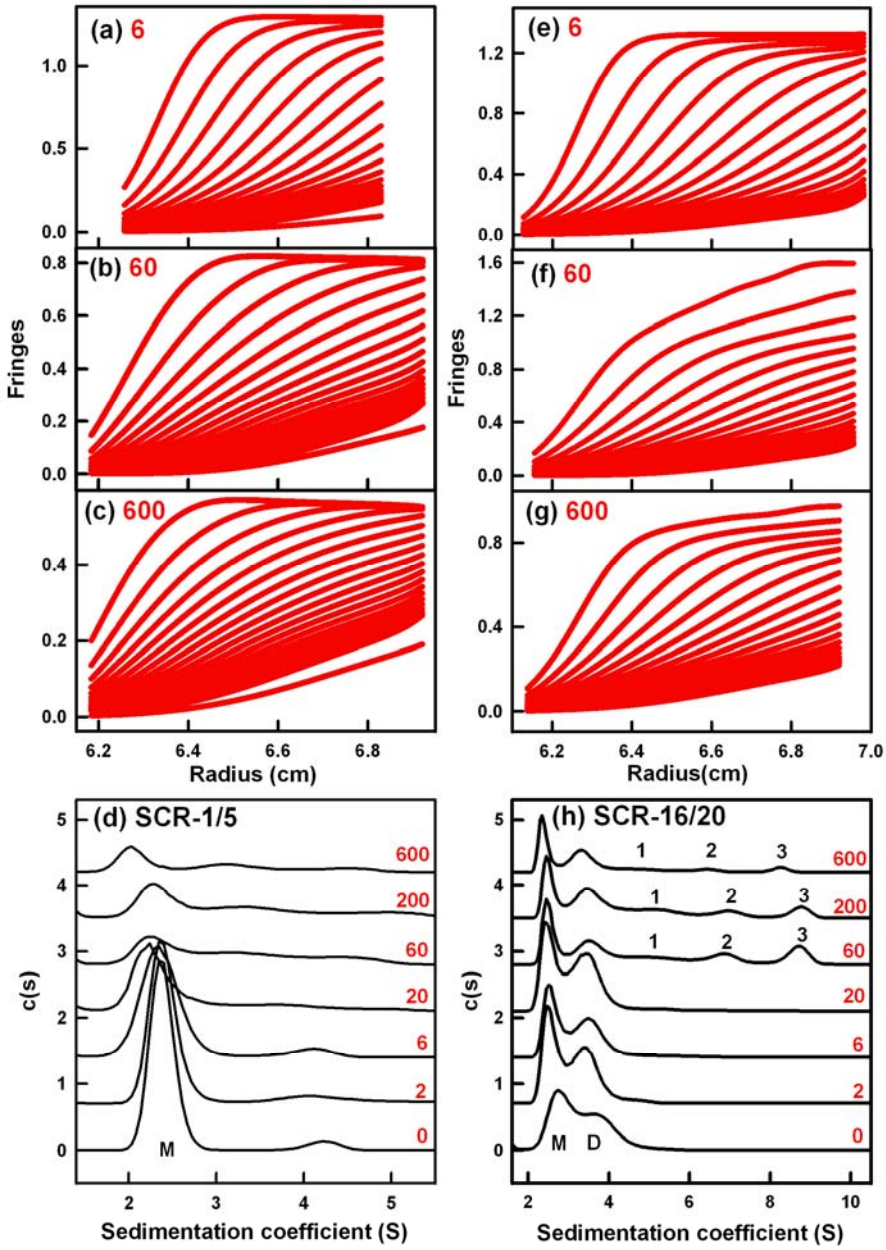
(a) 6
(b) 60
(c) 600
(e) 6
(f) 60
(g) 600
Fringes
Radius (cm)
Radius(cm)
(d) SCR-1/5
(h) SCR-16/20
c(s)
Sedimentation coefficient (S)
600
200
60
20
6
2
0
M
D
1
2
3

Figure 5

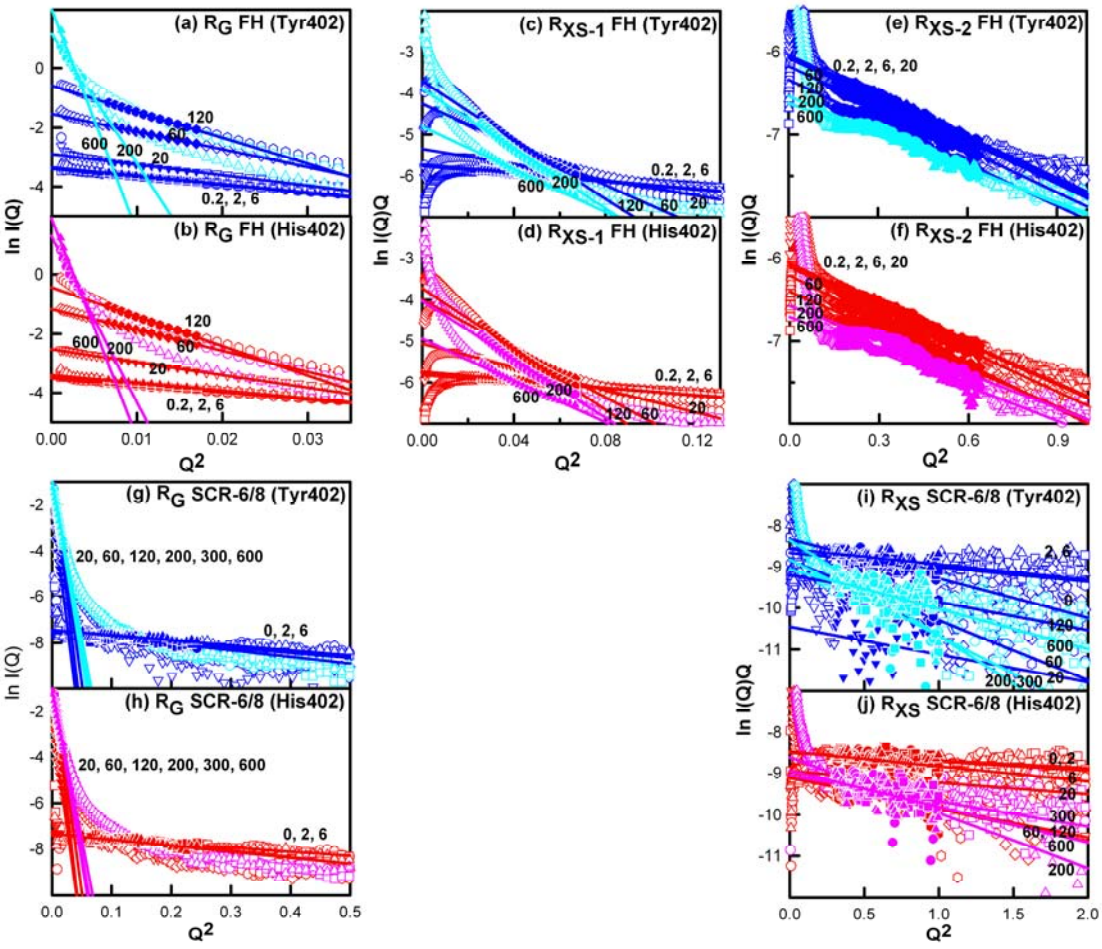
(a) R_G FH (Tyr402)
(b) R_G FH (His402)
(c) R_{XS-1} FH (Tyr402)
(d) R_{XS-1} FH (His402)
(e) R_{XS-2} FH (Tyr402)
(f) R_{XS-2} FH (His402)
(g) R_G SCR-6/8 (Tyr402)
(h) R_G SCR-6/8 (His402)
(i) R_{XS} SCR-6/8 (Tyr402)
(j) R_{XS} SCR-6/8 (His402)
ln I(Q)
ln I(Q)Q
Q2
0.2, 2, 6
0.2, 2, 6, 20
20, 60, 120, 200, 300, 600
0, 2, 6

Figure 6

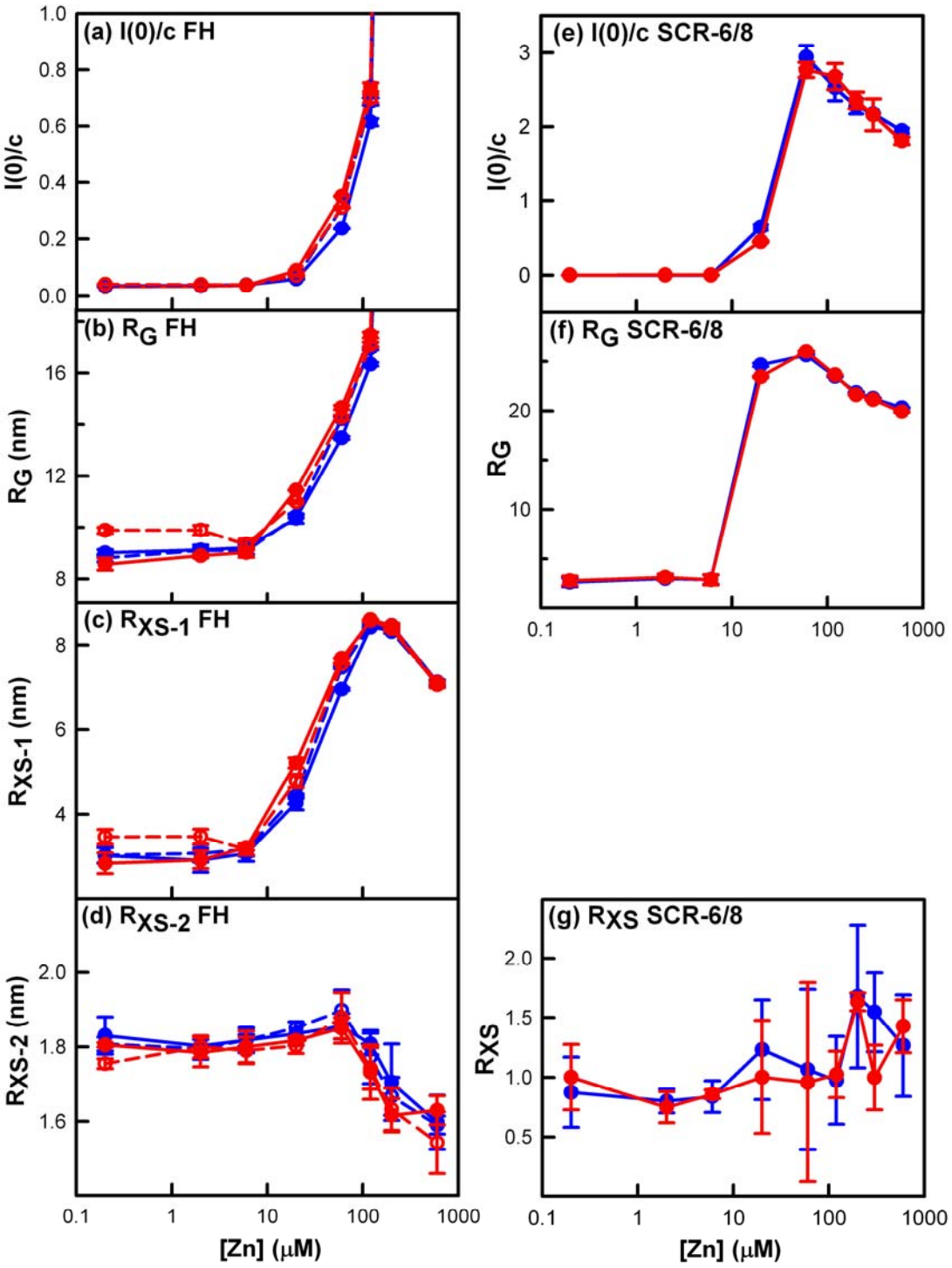
(a) I(0)/c FH
I(0)/c
(b) R_G FH
R_G (nm)
(c) R_{XS-1} FH
R_{XS-1} (nm)
(d) R_{XS-2} FH
R_{XS-2} (nm)
[Zn] (μM)
(e) I(0)/c SCR-6/8
I(0)/c
(f) R_G SCR-6/8
R_G
(g) R_{XS} SCR-6/8
R_{XS}
[Zn] (μM)

Figure 7

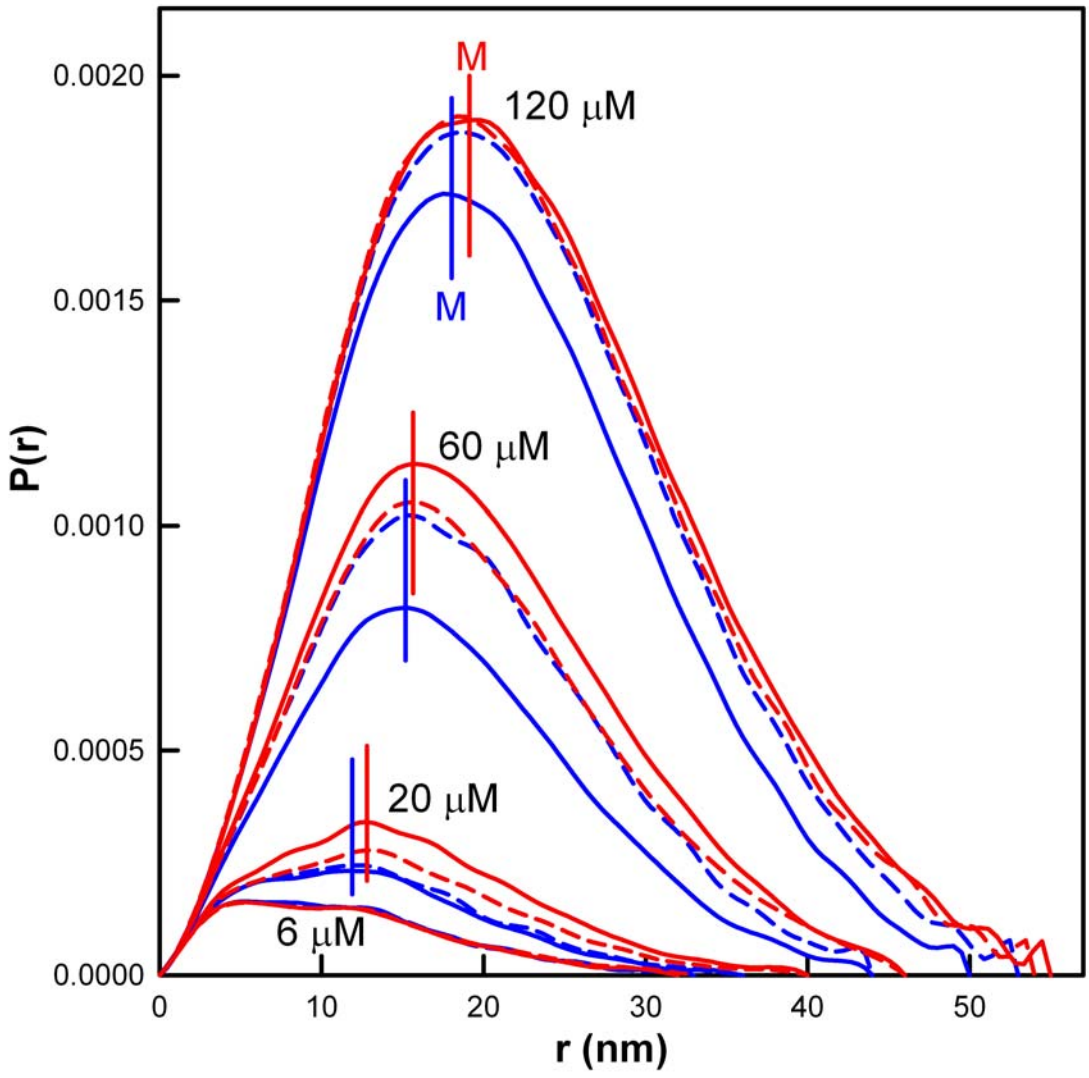
M
120 μM
M
60 μM
20 μM
6 μM
P(r)
0.0020
0.0015
0.0010
0.0005
0.0000
0
10
20
30
40
50
r (nm)

Figure 8

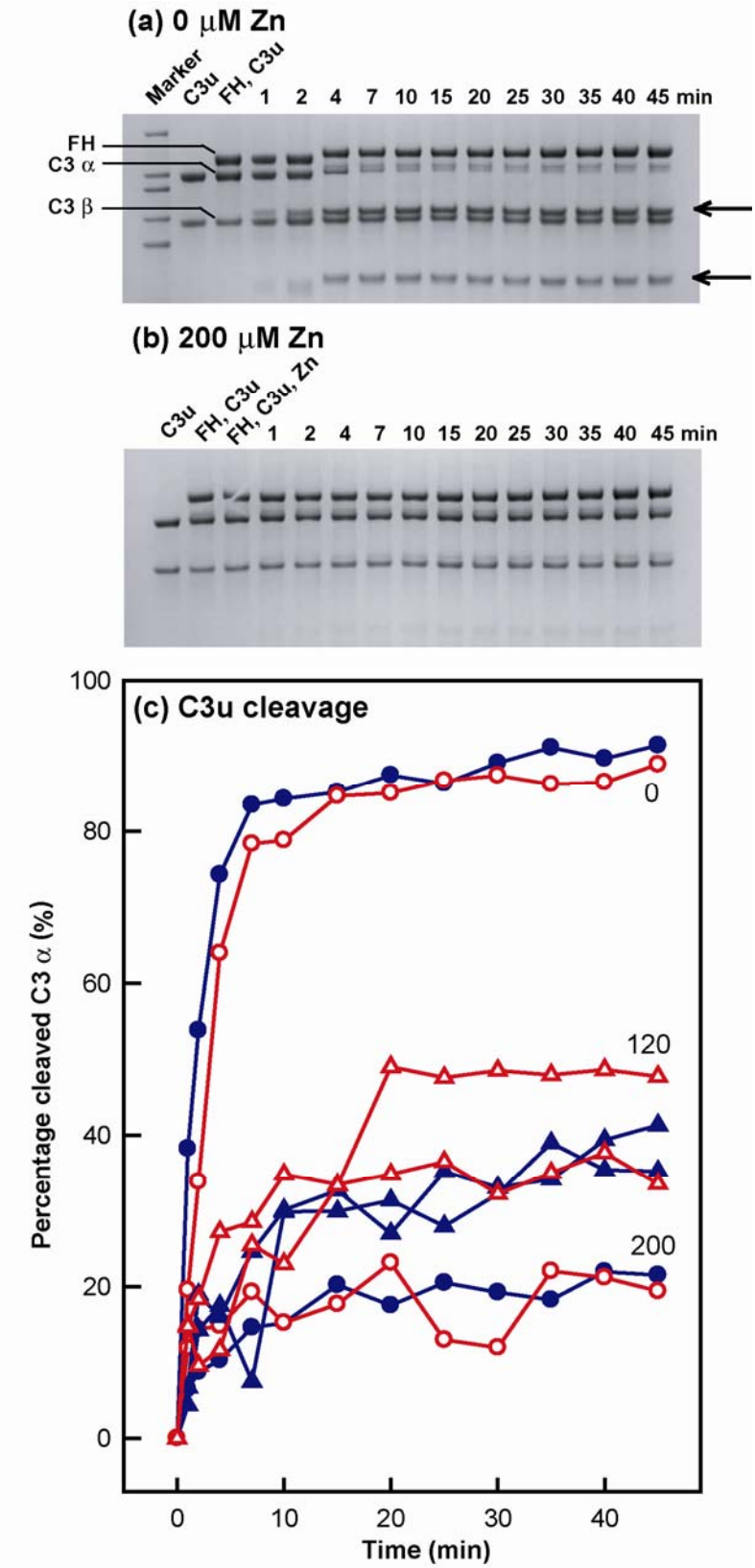
(a) 0 μM Zn
Marker
C3u
FH, C3u
1 2 4 7 10 15 20 25 30 35 40 45 min
FH
C3 α
C3 β
(b) 200 μM Zn
C3u
FH, C3u
FH, C3u, Zn
1 2 4 7 10 15 20 25 30 35 40 45 min
(c) C3u cleavage
Percentage cleaved C3 α (%)
Time (min)
0
120
200
0
20
40
60
80
100
0
10
20
30
40

Figure 9

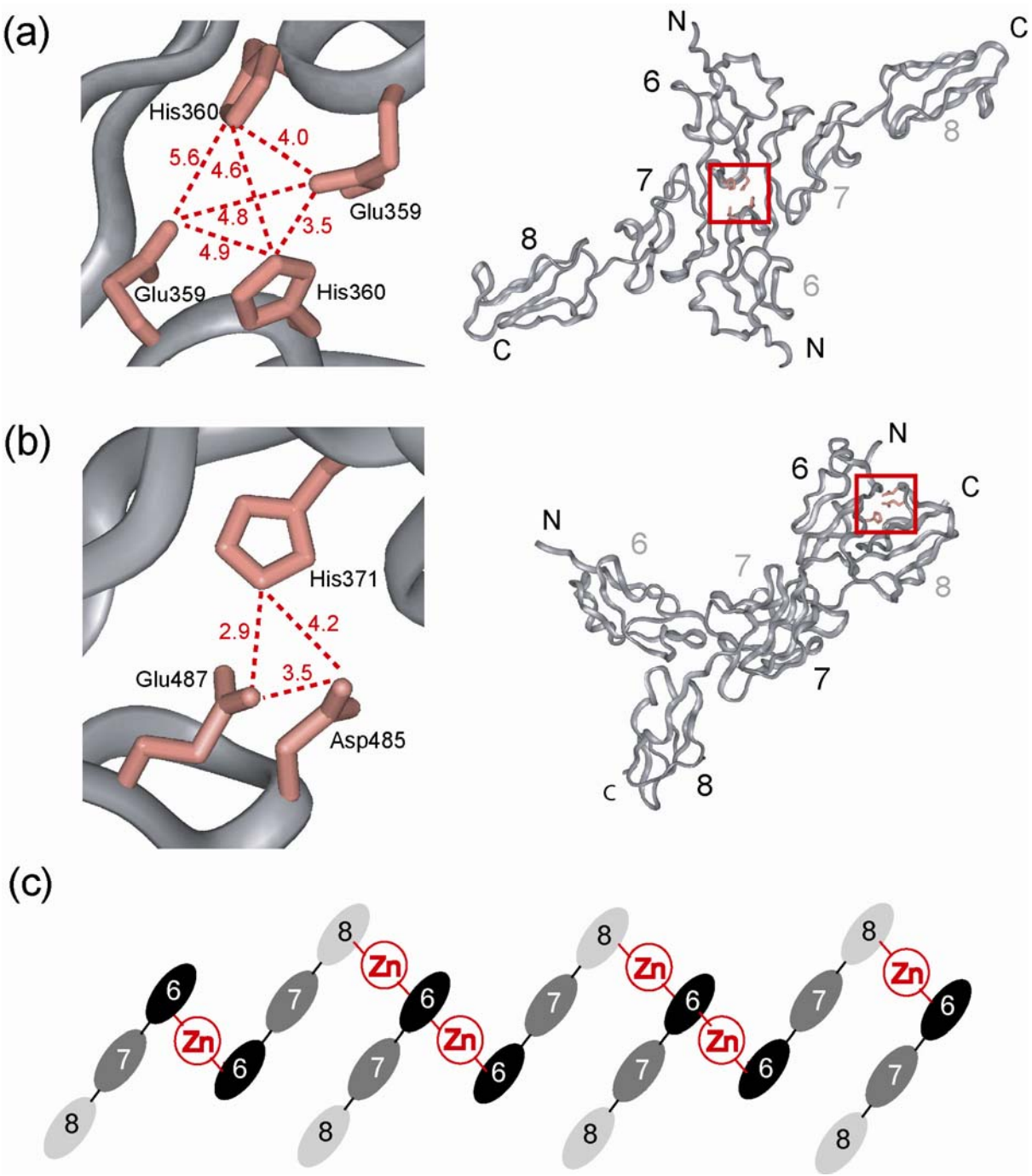
(a)
His360
4.0
5.6
4.6
Glu359
4.8
3.5
4.9
Glu359
His360
N
C
6
8
7
7
8
6
C
N
(b)
His371
2.9
4.2
Glu487
3.5
Asp485
N
6
C
N
6
7
8
7
C
8
(c)
8
Zn
6
7
6
Zn
6
7
8
8
Zn
6
7
6
Zn
6
7
8
8
Zn
6
7
8

Figure 10