



# The solution structure of the complement deregulator FHR5 reveals a compact dimer and provides new insights into CFHR5 nephropathy

Received for publication, July 6, 2020, and in revised form, September 6, 2020. Published, Papers in Press, September 14, 2020, DOI 10.1074/jbc.RA120.015132

Nilufar Kadkhodayi-Kholghi<sup>1</sup> , Jayesh S. Bhatt<sup>1</sup> , Jayesh Gor<sup>1</sup> , Lindsay C. McDermott<sup>2</sup> , Daniel P. Gale<sup>3</sup> , and Stephen J. Perkins<sup>1,\*</sup>

From the <sup>1</sup>Department of Structural and Molecular Biology, Division of Biosciences, University College London, London, United Kingdom, the <sup>2</sup>School of Life Sciences, University of Bedfordshire, Luton, United Kingdom, and <sup>3</sup>UCL Department of Renal Medicine, Royal Free Hospital, University College London, London, United Kingdom

Edited by Karen G. Fleming

The human complement Factor H-related 5 protein (FHR5) antagonizes the main circulating complement regulator Factor H, resulting in the deregulation of complement activation. FHR5 normally contains nine short complement regulator (SCR) domains, but a FHR5 mutant has been identified with a duplicated N-terminal SCR-1/2 domain pair that causes CFHR5 nephropathy. To understand how this duplication causes disease, we characterized the solution structure of native FHR5 by analytical ultracentrifugation and small-angle X-ray scattering. Sedimentation velocity and X-ray scattering indicated that FHR5 was dimeric, with a radius of gyration ( $R_g$ ) of  $5.5 \pm 0.2$  nm and a maximum protein length of 20 nm for its 18 domains. This result indicated that FHR5 was even more compact than the main regulator Factor H, which showed an overall length of 26–29 nm for its 20 SCR domains. Atomistic modeling for FHR5 generated a library of 250,000 physically realistic trial arrangements of SCR domains for scattering curve fits. Only compact domain structures in this library fit well to the scattering data, and these structures readily accommodated the extra SCR-1/2 domain pair present in CFHR5 nephropathy. This model indicated that mutant FHR5 can form oligomers that possess additional binding sites for C3b in FHR5. We conclude that the deregulation of complement regulation by the FHR5 mutant can be rationalized by the enhanced binding of FHR5 oligomers to C3b deposited on host cell surfaces. Our FHR5 structures thus explained key features of the mechanism and pathology of CFHR5 nephropathy.

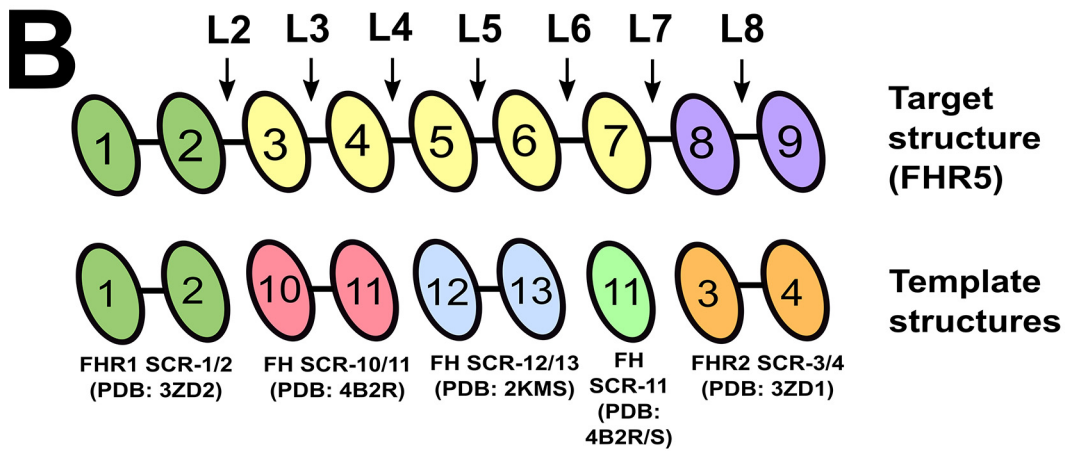
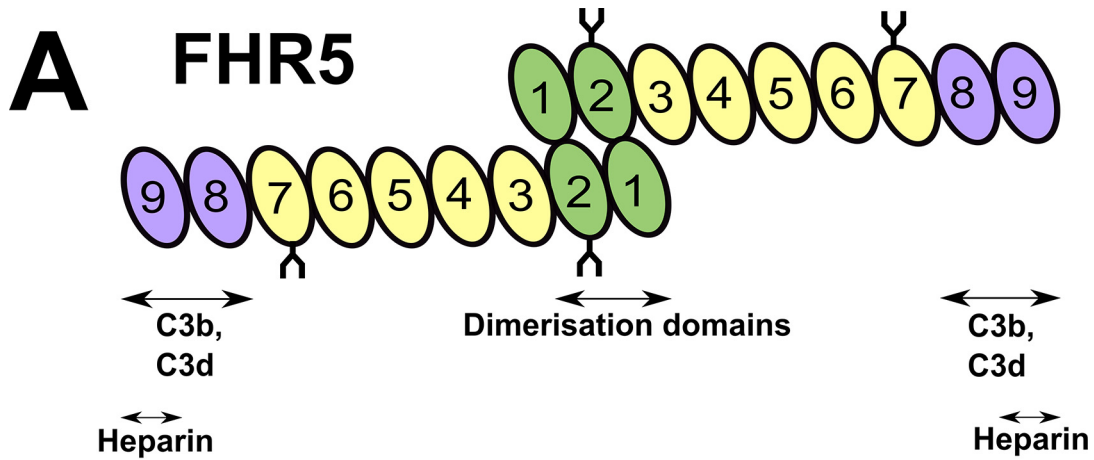
Complement activation and regulation is of major importance in enabling clearance of pathogens, while preventing complement-mediated host cell damage. Complement factor H-related 5 protein (FHR5) was first identified co-localized with C3 in glomerular immune deposits from patients with glomerulonephritis and is a member of a family of structurally related proteins comprising the major serum complement regulator Factor H and five complement Factor H-related proteins. Factor H, comprising 20 short complement regulator (SCR) domains, has been well characterized, both in terms of its structure and function, binding to activated C3b and its frag-

ment C3d and regulating excess C3 activation (1). However, the principal physiological function of FHR5 is poorly understood. FHR5 circulates in plasma in extremely low concentrations of 3–6  $\mu\text{g/ml}$  (2), which is  $\sim 100$ -fold lower than Factor H. It is also the least abundant of the FHR proteins, yet its structure is the longest of these proteins, with a linear sequence of nine SCR domains (Fig. 1). The SCR domain (3) is the major domain type found in the complement regulators. An SCR domain is characterized by a consensus sequence of  $\sim 61$  amino acids, with four invariant cysteine residues that form two disulfide bridges (I–III and II–IV) and a conserved tryptophan residue. It folds compactly, with a hydrophobic core, in a  $\beta$ -sandwich arrangement of six hydrogen-bonded  $\beta$ -strands. The key C-terminal C3b/C3d recognition sites are conserved between SCR-19/20 of Factor H and SCR-8/9 of FHR5 (Fig. 1). FHR5 also interacts with heparin (2); however, FHR5 has no complement regulatory domains equivalent to SCR-1/4 of Factor H. FHR5 forms native homodimers via its two N-terminal domains SCR-1/2 that exhibit increased avidity for C3b/C3d compared with the monovalent Factor H, and, although early studies using supraphysiological concentrations of FHR5 showed evidence of weak (compared with Factor H) complement regulating activity (2), more recent work has shown that, at physiological concentrations, FHR5 competitively antagonizes Factor H, thus deregulating complement (4, 5). Conflicting data exist on whether FHR5 forms heterodimers with other FHRs *in vivo* (6, 7).

CFHR5 nephropathy, a monogenic cause of kidney failure endemic in Cypriots (individuals residing in or with ancestry from the island of Cyprus), is characterized in almost all affected individuals by persistent microscopic hematuria and, in a proportion of patients, episodes of kidney damage and visible blood in the urine that occur at times of otherwise trivial mucosal infections, with repeated episodes typically resulting in progressive kidney damage and eventually end-stage kidney failure occurring in  $>80\%$  of affected males and  $<20\%$  of affected females by the age of 55 years. Kidney biopsy shows predominantly mesangial-based glomerular inflammation with deposition of C3 but not Igs in the mesangium and, under electron microscopic examination, the subendothelial part of the glomerular basement membrane—appearances termed C3 glomerulopathy that suggest defective regulation of the complement system. The disease is a highly penetrant autosomal

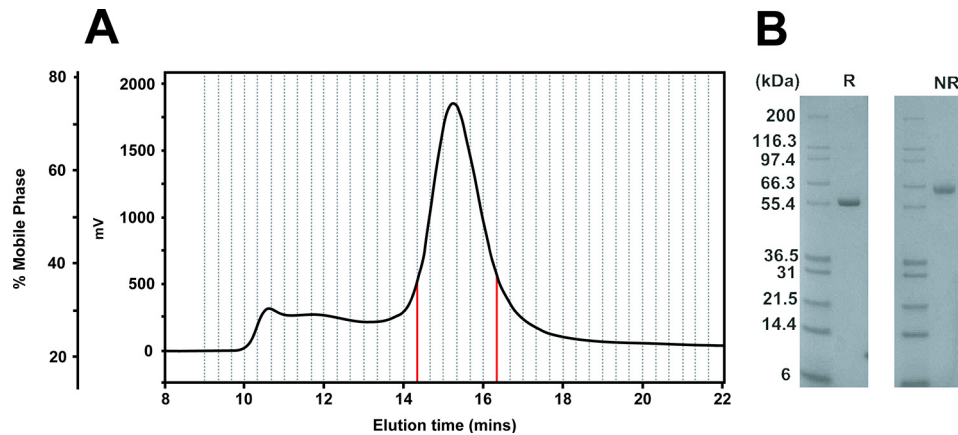
This article contains supporting information.

\* For correspondence: Stephen J. Perkins, [s.perkins@ucl.ac.uk](mailto:s.perkins@ucl.ac.uk).



**C**

	20	30	40	50	60	70	80
SCR-1 3ZD2	EGTLCD	FPKIHGFLYDEEDYNPFSQVPT	GEVFIYS	CEYNFVSPSKSFWTRIT	CTEEGWSP	TPKCLRM	
	EATLCD	FPKINHGILYDEEKYPFSQVPT	GEVFIYS	CEYNFVSPSKSFWTRIT	CTEEGWSP	TPKCLRL	
SCR-2 3ZD2	CS	PPFVKNHSESS	GLIHLE	GDTVQIICNTGYSLQNN	KNISQ	VERGWST	PPICSF
	CF	FPFVENHSESS	GQTHLE	GDTVQIICNTGYRLQNN	NNISQ	VERGWST	PPKCRST
SCR-3 4B2R	CH	VPILANVDAQPK	KESYKV	GDVLKFSCKRNLIRVGS	DSVQC	YQFGWSPNFP	TCRQGVRS
	CE	LPKIDVHLVPDRK	KDQYKV	GEVLKFSCKPGFTIVGP	NSVQC	YHFGLS	PDLPICKEQVQS
SCR-4 4B2R	CG	PPQLSNGEVKEIR	KEEYGH	NEVVEYDCNPNFIINGP	KKIQ	VDGEWTT	LPTC
	CG	PPPELLNGNVKEKT	KEEYGH	SEVVEYCNPRFLMKGP	NKIQC	VDGEWTT	LPVCI
SCR-5 2KMS	CG	YIPELEYGVQPS	VPPYQH	GVSVEVNCRNEYAMIGN	NMITC	INGI	WTE
	CG	DIPELEHGWAQLS	SPPYYY	GDSVEFN	SESFTMIGH	RSITC	IHG
SCR-6 2KMS	CK	IAGVNIKTLKLS	GKEFNH	NSRIRYRSDIFRYRHS	VCING	KWNP	EVD
	CK	SSNLIILEEHLKN	KKEFDH	NSNIRYRCRGEKGIHT	VCING	RWDP	EVNCS
SCR-7 4B2R 4B2S	CG	PPPQIPNAQ	NMTT	TVNYQD	GEKVALCKENYLLPEA	KEIVC	KDGRWQS
	CG	PPPELLNGNVKEKT	KEEYGH	SEVVEYCNPRFLMKGP	NKIQC	VDGEWTT	LPVCI
	CG	PPPELLNGNVKEKT	KEEYGH	SEVVEYCNPRFLMKGP	NKIQC	VDGEWTT	LPVCI
SCR-8 3ZD1	CG	PPPSINNGDITTSFP	LSVYPP	GSTVTYRCQSFYKLGQS	VTVTC	RNKQWSE	PPRCLDP
	CG	PPPIDNGDITTSFL	LSVYAP	GSSVEYCNQNYLQLEGN	NQITC	RNGWSE	PPKCLDP
SCR-9 3ZD1	CV	SEENMKNKIQLKWRNDGKLYAKT	DAVEFQ	CKFPHKAMISSPPFRAI	CE	QEGKFEY	PI
	CV	SQEI MEKYNIKLKWNTNQKLYSR	TGDI	VEFV	CKSGYHP	TKSHSFRAMC	QNGKLVY



**Figure 2. Purification of FHR5.** *A*, gel-filtration profile showing the removal of aggregates from the FHR5 sample. The sample was loaded onto a Superdex™ 200 gel-filtration column. Fractions containing homogeneous FHR5 from the large peak as indicated were pooled. Mass determinations are shown in *Fig. 3, B*, nonreducing (*NR*) and reducing (*R*) SDS-PAGE analyses of FHR5 (~65 kDa monomeric). The molecular masses of the Mark 12™ protein standard (Invitrogen) in kDa are labeled.

dominant disorder that is caused by heterozygosity for an in-frame duplication of exons 2 and 3 of the *CFHR5* gene that results in production of an elongated FHR5 protein with an extra two N-terminal SCR-1/2 domains in tandem. No extrarenal features of the disease have been reported, despite the review of clinical data from over 100 affected individuals of all ages (8, 9). The molecular mechanisms that make the kidney susceptible to complement-mediated damage in *CFHR5* nephropathy and other common causes of glomerulonephritis (*e.g.* lupus nephritis and IgA nephropathy, in which flares of disease triggered by mucosal infections also occur) are not well-understood.

Protein structural studies of full-length FHR5 are complicated by its large size and its eight potentially flexible interdomain linkers of lengths between three and eight residues (*Fig. 1*), both of which make it difficult to crystallize to determine its three-dimensional appearance. To date, atomic-level structures have not been determined for any small FHR5 fragments. However, alternative methods can be used for structural studies. Previously for full-length factor H, EM, small-angle X-ray scattering (SAXS), analytical ultracentrifugation (AUC), and molecular modeling showed that full-length factor H has a partially folded-back structure that is relevant to its regulatory function (10–12). This combination of analytical ultracentrifugation, X-ray solution scattering, and atomistic modeling has been effective in determining many macromolecular structures in solution (13–15). Many of the first structural explanations for factor H-associated diseases, such as atypical hemolytic uremic syndrome, were based on homology models for the SCR domains (16–18). Here, these solution structural and modeling approaches were applied to determine the solution conformation of full-length FHR5 to explain its role in healthy individuals and how *CFHR5* nephropathy may arise through the SCR-1/2

duplication. Following SAXS and AUC data collection, full-length FHR5 was modeled using molecular dynamics, followed by Monte Carlo simulations to generate a large library of physically realistic trial atomistic structures for the FHR5 dimer (14, 19). The theoretical scattering profiles of this library were compared with the experimental SAXS curves to determine best-fit FHR5 structures. We thus defined a small subset of compact folded-back solution structures. The extra SCR-1/2 domain pair in mutant FHR5 was readily added to these structures, their presence leading to the formation of multivalent oligomers of FHR5. Our work explains how FHR5 regulates complement activation in the kidney and how *CFHR5* nephropathy arises.

**Results**

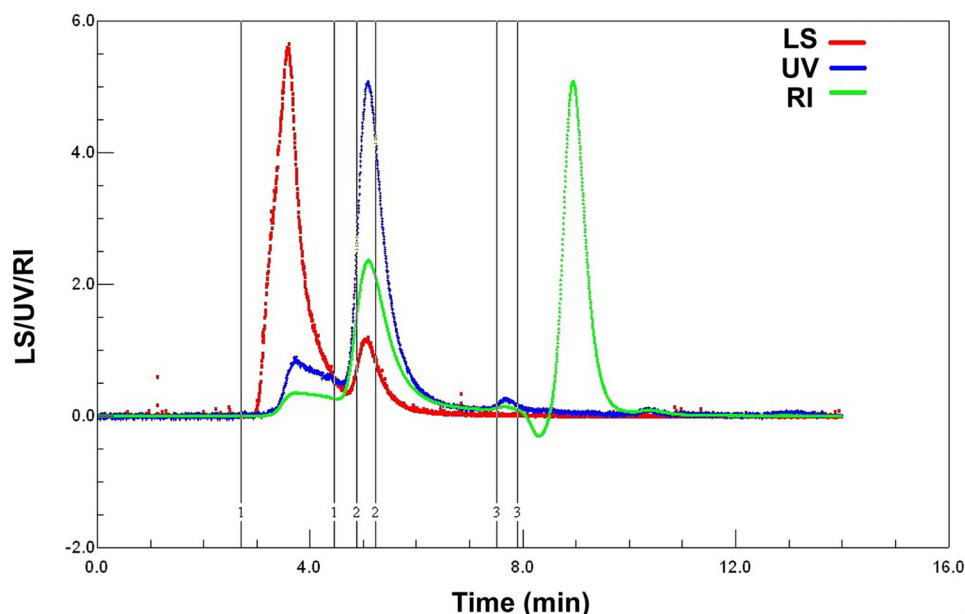
**Purification of full-length FHR5**

Human FHR5 SCR-1/9 purchased from Creative Biolabs was subjected to gel-filtration chromatography to ensure monodispersity and removal of aggregates prior to SAXS experiments. The protein eluted as a single symmetrical peak at ~15-ml elution volume (*Fig. 2A*). This was preceded by a broader peak that was eluted between 10 and 14 ml, which was attributed to protein aggregates. Only the protein fractions between 14.3 and 16.3 ml (*red in Fig. 2A*) were retained. By SDS-PAGE (*Fig. 2B*), a single band was seen at 60–66 kDa (nonreduced) that corresponds well to the expected monomer molecular mass of 62.4 kDa. Reducing conditions resulted in another single band but at a slightly lower mass, this difference being attributed to the presence of glycan chains on FHR5.

SEC-MALLS was used to determine the mass and self-association of FHR5 in our Tris-150 purification buffer, as in previous work (4). FHR5 from a size-exclusion column was detected by

**Figure 1. The human FHR5 SCR-1/9 domain structure.** *A*, schematic representation of the domain structure of the human FHR5 dimer and its ligand binding sites. The two putative N-linked glycosylation sites (not modeled) on SCR-2 and SCR-7 are depicted as Y-shaped symbols. Domains SCR-1/2 (green) form a head-to-tail dimer. Domains SCR-3/7 (yellow) have high sequence identities to human Factor H SCR-10/14. The C3b/C3d- and heparin-binding sites are indicated as arrows on SCR-8 and SCR-9 (purple). *B*, summary of the templates used for the homology modeling of FHR5. Each FHR5 domain is aligned with the corresponding crystal/NMR structure denoted by its PDB code. *C*, the sequences of the nine FHR5 SCR domains are aligned with the template SCR sequences identified by their PDB codes. The conserved cysteine and tryptophan residues are highlighted in pink and yellow respectively, and the linkers in cyan. The residue numbering corresponds to the full FHR5 protein sequence (SWISSPROT accession code Q9BXR6) including the signal peptide, with +1 corresponding to the methionine start codon.





**Figure 3. SEC-MALLS analysis of FHR5.** The elution profile (chromatogram) for FHR5 was analyzed using UV detection (blue), MALLS (light scattering), detector (red), and refractive index detector (green). Three successive prominent peaks (1–3 as indicated by the pairs of numbers below) were examined for their molecular mass. The calculated molecular masses were >5,400, 162, and 27 kDa for peaks 1, 2, and 3, respectively. The refractive index peak above 8.0 min is attributed to the end of the gel-filtration step.

UV (blue, Fig. 3) and refractive index (green) measurements, in parallel with multi-angle light scattering (red) to analyze size distributions. Three peaks were observed in the elution profile. Peak 1 at 2.7–4.2 min was assigned as aggregated material, because this had a lower UV and refractive index, but high light scattering intensities that indicated very large sizes. Its molecular mass was calculated to be above 5,400 kDa. Peak 2 at 4.9–5.2 min was the FHR5 dimer that eluted with higher UV and refractive index values but with lower light scattering. Its molecular mass was estimated as 162 kDa, this being consistent with FHR5 dimer formation, given that the mass of the monomer was 62.4 kDa from its composition (20). Despite a large inherent error associated with light scattering, no evidence of an FHR5 monomer peak was detectable. A small peak 3 at 7.6–7.9 min was assigned to fragments below 30 kDa.

#### AUC analyses of FHR5

AUC sedimentation velocity experiments on FHR5 studied its oligomerization and shape using size distribution  $c(s)$  analyses to determine its molecular mass and sedimentation coefficient  $s_{20,w}$ . Absorbance data for FHR5 at 0.16 mg/ml in PBS were collected for five different salt concentrations between 20 and 250 mM NaCl. SEDFIT analyses involved as many as 500 absorbance scans. The experimental sedimentation boundaries (left, Fig. 4) gave good fits to the Lamm equation to give the size-distribution  $c(s)$  profiles (right, Fig. 4), despite the low concentrations in use. These fits were obtained by floating the meniscus, the bottom of the cell, the baseline, and the frictional ratio  $f/f_0$  of around 1.5.

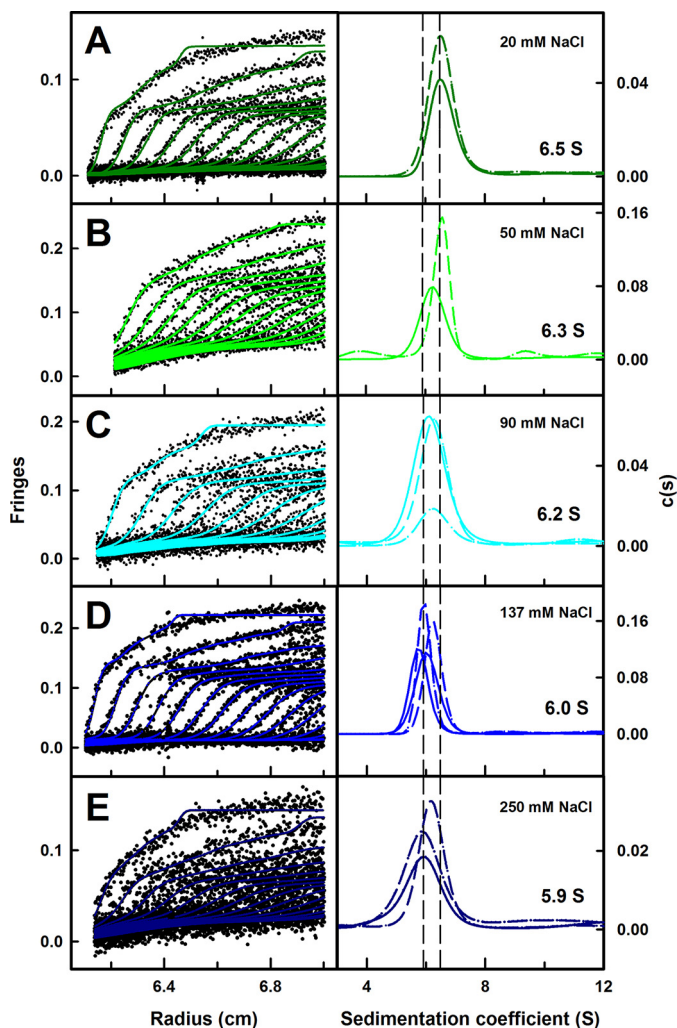
Protein aggregation was visible in the earliest boundaries that sedimented rapidly at the start of the runs, to leave behind the FHR5 dimer that sedimented more slowly (Fig. 4). This agreed with SEC-MALLS. A major  $c(s)$  peak at 6.0 S was ob-

served for FHR5 in PBS-137 that corresponded to an average molecular mass of 134 kDa. This mass confirmed the presence of dimer in solution. The aggregates made little contribution to the  $c(s)$  analyses between 3 and 12 S, even though they contributed as much as half of the protein present. The molecular masses for the five buffers were between 133 and 139 kDa (Table 1), showing that the FHR5 dimer was stable between 20 and 350 mM NaCl. The  $c(s)$  analyses did not reveal any FHR5 monomer at lower  $s$  values. The reproducibility of these data was tested at two different rotor speeds of 40,000 and 50,000 rpm, to show no difference.

The solution structure of FHR5 between 20 and 350 mM NaCl was monitored using the mean  $s_{20,w}$  values (Table 1). A significant decrease of 0.9 S from 6.48 to 5.35 S was seen on going from 20 mM NaCl to 350 mM NaCl. This shift in the FHR5 dimer peak was visible in the  $c(s)$  distribution plots (vertical dashed lines, Fig. 4). This result indicated a conformational change in FHR5, where the smaller  $s_{20,w}$  values at high NaCl concentration indicated a more elongated FHR5 domain structure that formed as the ionic strength was increased (Fig. 5).

#### SAXS analyses of FHR5

SAXS was used to study the solution structure of the FHR5 dimer in concentration series in three different buffers, two being physiological (PBS-137 and Tris-150) and one being low-salt (PBS-50). The FHR5 samples were purified by gel filtration (Fig. 2). In Tris-150, data were collected using 0.04–0.5 mg/ml FHR5. In PBS-137 and PBS-50, data were collected using 0.04–0.17 mg/ml FHR5. Guinier analyses of the solution structure gave high-quality linear plots in two distinct regions of the  $I(Q)$  curves that corresponded to the radius of gyration ( $R_g$ ) and the cross-sectional radius of gyration ( $R_{cs}$ ) from two distinct  $Q$  ranges (Fig. 6). These values are measures of the overall and the shorter



**Figure 4. Size-distribution  $c(s)$  sedimentation velocity analysis of FHR5.** Left, the experimentally observed sedimentation boundaries (black) by interference optics are shown for 0.16 mg/ml FHR5 SCR-1/9 based on the following PBS buffers: 20 mM NaCl (A), 50 mM NaCl (B), 90 mM NaCl (C), 137 mM NaCl (D), and 250 mM NaCl (E). In each case, 40–50 boundaries (colored lines) were fitted as shown. Right, the data fits corresponded to rotor speeds of 40,000 rpm (solid lines) and 50,000 rpm (dashed lines). The resulting size-distribution analyses  $c(s)$  revealed a major peak that shifted from 6.5 S in 20 mM NaCl to 5.9 S in 250 mM NaCl. The vertical dashed lines illustrate the shift in  $s_{20,w}$  values as the salt concentrations increased.

**Table 1**  
Sedimentation velocity data for FHR5 SCR-1/9

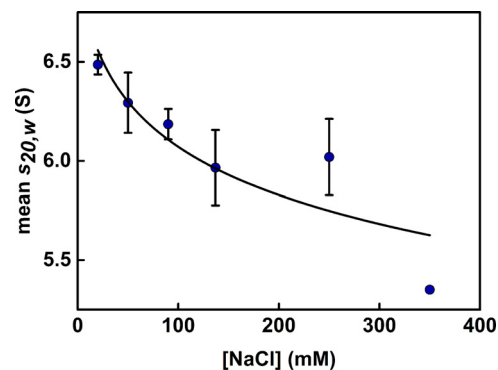
Sample <sup>a</sup>	$s_{20,w}$	Molecular mass	Frictional ratio ( $ff_0$ )
FHR5 in PBS-20	$6.48 \pm 0.05$	$139 \pm 2$	$1.49 \pm 0.0003$
FHR5 in PBS-50	$6.29 \pm 0.1$	$133 \pm 4$	$1.48 \pm 0.02$
FHR5 in PBS-90	$6.19 \pm 0.1$	$134 \pm 6$	$1.53 \pm 0.05$
FHR5 in PBS-137	$5.97 \pm 0.2$	$134 \pm 4$	$1.59 \pm 0.03$
FHR5 in PBS-250	$5.91 \pm 0.02$	$136 \pm 8$	$1.58 \pm 0.02$
FHR5 in PBS-350	$5.35^b$	138	1.80

<sup>a</sup>All samples were measured at 0.16 mg/ml. The values represent the mean  $\pm$  S.D. The data presented include those collected both at 40,000 and 50,000 rpm.

<sup>b</sup>Only a single run was carried out for FHR5 SCR-1/9 in PBS-350.

dimensions of macromolecular elongation, respectively. Their values were deduced according to Equations 1 and 2, respectively, within satisfactory  $Q.R_g$  and  $Q.R_{xs}$  limits close to 1.0 as follows.

(i) In the overall structural Guinier  $R_g$  analyses in a low  $Q$  range of 0.1–0.27  $\text{nm}^{-1}$  (Fig. 7A), in Tris-150 and PBS-137



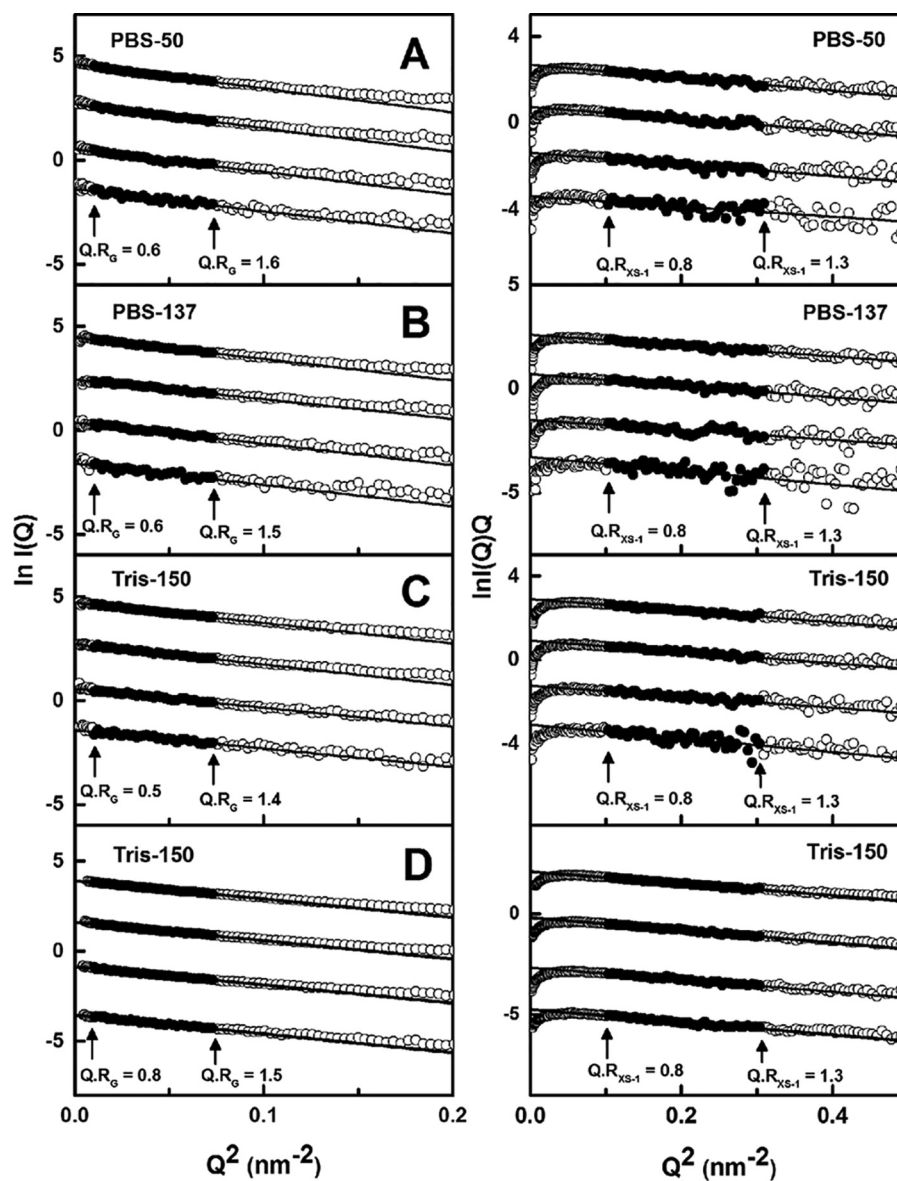
**Figure 5. Relationship between sedimentation coefficient and the buffer salt concentration.** The mean  $\pm$  S.D. of the sedimentation coefficient ( $s_{20,w}$ ) data are shown as a function of the NaCl concentration in each buffer. For PBS-350, only a single data point was available.

buffers with similar NaCl concentrations, the mean  $R_g$  values were  $5.36 \pm 0.14$  and  $5.48 \pm 0.17$  nm, respectively. However, in the PBS-50 buffer with lower NaCl, the mean  $R_g$  value increased slightly to  $5.91 \pm 0.13$  nm. This increase was attributed to trace aggregation in FHR5 that affected the lowest  $Q$  values (Fig. 7A). No concentration dependence was observed for the  $R_g$  values between 0.04 and 0.17 mg/ml; however, a slightly increased  $R_g$  value of up to 0.2 nm was seen at 0.2–0.5 mg/ml FHR5. (ii) In the cross-sectional Guinier  $R_{xs}$  analyses, using a  $Q$  range of 0.32–0.55  $\text{nm}^{-1}$  (Fig. 7B), the mean  $R_{xs}$  values in each buffer were  $2.41 \pm 0.06$ ,  $2.29 \pm 0.09$ , and  $2.46 \pm 0.14$  nm for Tris-150, PBS-137, and PBS-50, respectively (Table 2). No significant changes in the  $R_{xs}$  values were seen between the data sets for these NaCl and protein concentrations, indicating that the cross-sectional structure of FHR5 was unchanged in conformation.

The distance distribution function  $P(r)$  in real space represents all of the distances between pairs of atoms in FHR5. This was calculated from Fourier transformation of the full  $I(Q)$  scattering curve following the specification of the maximum dimension  $D_{\text{max}}$  (Equation 3; Fig. 8). The  $P(r)$  curve provided an independent  $R_g$  value for FHR5 for comparison with the Guinier value (Table 2). The  $R_g$  values from the  $P(r)$  analyses were in good agreement with those from the Guinier analyses (Table 2). The  $P(r)$  curve also gave the maximum length  $L$  of FHR5 from the value of  $r$  when  $P(r) = 0$ . The mean  $L$  values were  $19.5 \pm 0.4$  nm in Tris-150 (Fig. 8C),  $19.6 \pm 0.5$  nm in PBS-137 (Fig. 8B), and  $21.0$  nm in PBS-50 (Fig. 8A). The  $L$  value for PBS-50 was slightly higher than those in Tris-150 and PBS-137, most likely due to trace aggregation that resulted from the lower ionic strength used (see above). A single maximum  $M$  was observed in all of the  $P(r)$  curves. This corresponded to the most frequent interatomic distance within the FHR5 structure (Table 2). The mean  $M$  values were  $4.9 \pm 0.3$ ,  $4.9 \pm 0.1$ , and  $5.4 \pm 0.3$  nm for Tris-150, PBS-137, and PBS-50, respectively. The  $M$  values were relatively stable, although slightly higher for PBS-50 as the result of trace aggregates.

#### Initial model for the FHR5 dimer

Currently, there is no atomic level structural information on FHR5. To determine an atomistic-level solution structure for



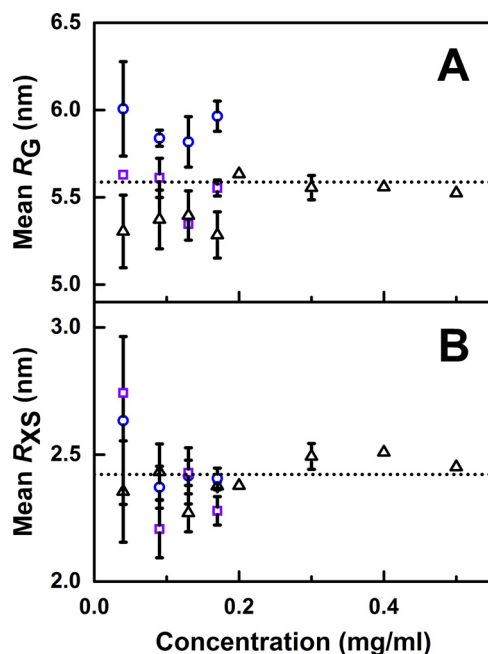
**Figure 6.** X-ray Guinier  $R_g$  and  $R_{xs}$  analyses for FHR5. In the Guinier analyses, plots of  $\ln I(Q)$  against  $Q^2$  for the  $R_g$  analyses (left), and  $\ln I(Q).Q$  against  $Q^2$  for the  $R_{xs}$  analyses (right) are shown. The straight lines indicate the slopes of each fit. The filled circles correspond to the  $Q.R_g$  and  $Q.R_{xs}$  ranges used for each fit, with the  $Q$  range used for the  $R_g$  values being 0.10–0.27  $\text{nm}^{-1}$ , and that for the  $R_{xs}$  values being 0.32–0.55  $\text{nm}^{-1}$  (arrowed). In rows A, B, and C, the fits correspond to FHR5 concentrations of 0.17, 0.13, 0.09, and 0.04 mg/ml from top to bottom in each panel. A, the FHR5 fits correspond to PBS-50; B, the FHR5 fits correspond to PBS-137; C, the FHR5 fits correspond to Tris-150. D, the FHR5 fits correspond to 0.5, 0.4, 0.3, and 0.2 mg/ml from top to bottom in Tris-150.

the FHR5 dimer, a starting model for the monomer was required. This was created by comparative modeling based on four known SCR crystal structures as structural templates (Fig. 1, B and C). Two used related crystal structures of the N-terminal FHR1 SCR-1/2 domains and the C-terminal FHR2 SCR-3/4 domains with high sequence identities of 85.2 and 61.7%, respectively, with SCR-1/2 and SCR-8/9 of FHR5. The SCR-3/7 domains of FHR5 share significant sequence similarities with the SCR-10/14 domains of Factor H. Although templates for individual SCR3/7 domains in FHR5 were searched for in PDB-BLAST, the best choices were these domain structures from Factor H due to their direct sequence similarities (Fig. 1C). FHR5 SCR-3/4 was represented by Factor H SCR-10/11 with a high sequence identity of 57.4%. FHR5 SCR-5/6 was represented by Factor H SCR-12/13, also with a high sequence iden-

tity of 53.9%. Although FHR5 SCR-7 is similar to Factor H SCR-14, no structure existed for Factor H SCR-14. Searches showed that the best template structure for FHR5 SCR-7 was that of SCR-11 of Factor H with a sequence identity of 34.5%. The individual template-target sequence alignments (Fig. 1C) showed no significant indels in the structure, because the numbers of residues in these were well-aligned. Thus, the FHR5 SCR-7 and SCR-8/9 sequences had only one gap inserted in each. The individual modeled domains satisfied validation checks using PROCHECK, where the Ramachandran plots showed that 70% of the residues were in the most favored steric regions. The FHR5 dimer was generated from its monomer structure by aligning its SCR-1/2 domains with the crystal structure of the FHR1 SCR-1/2 dimer (see “Experimental procedures”), followed by energy minimization to relax this starting structure.



## Solution structure of FHR5



**Figure 7. Concentration dependence of the X-ray Guinier  $R_g$  and  $R_{xs}$  values.** Each value was measured in triplicate and shown as the mean  $\pm$  S.D. at low concentrations, except for the data sets above 0.2 mg/ml, which were measured once. The  $R_g$  (A) and  $R_{xs}$  (B) values are shown for PBS-50 (blue, circles), PBS-137 (purple, squares), and Tris-150 (black, triangles) buffers. The dotted line in both panels represents the mean across all values.

**Table 2**

### X-ray scattering data for FHR5 SCR-1/9

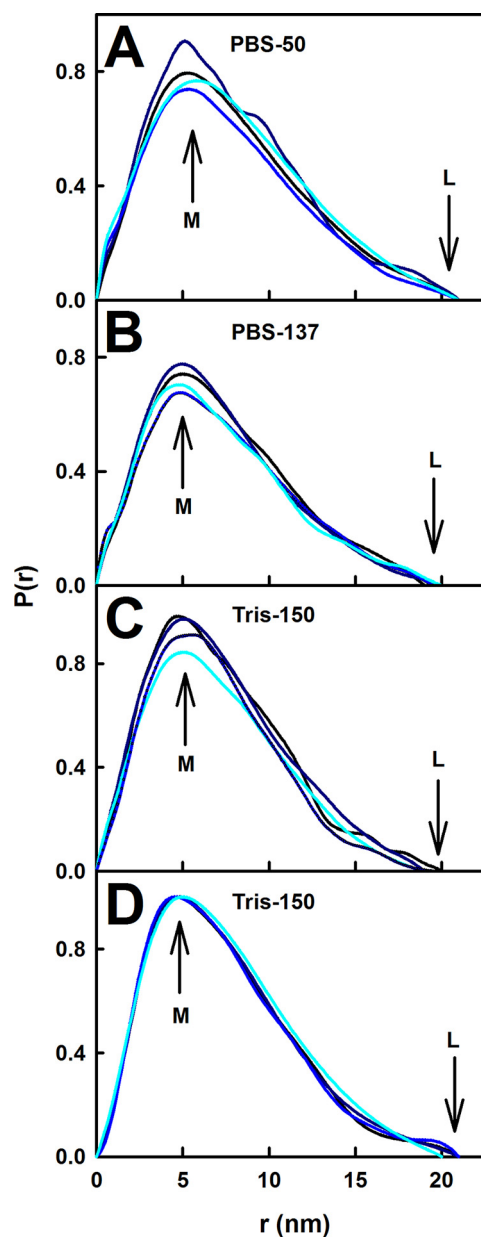
The data represent the mean values from all data sets.

Sample	$R_g^a$	$R_g^a$	$R_{xs}$	$L$	$M$
	<i>nm</i>				
<b>150 mM NaCl (Tris)</b>					
0.5 mg/ml	5.52	5.85	2.45	21.0	4.68
0.4 mg/ml	5.56	5.92	2.51	21.0	4.68
0.3 mg/ml	5.56 $\pm$ 0.07	5.90	2.49 $\pm$ 0.05	21.0	4.52
0.2 mg/ml	5.63	5.94	2.38	20.0	4.92
0.17 mg/ml	5.28 $\pm$ 0.13	5.72	2.38 $\pm$ 0.02	20.0	4.75
0.13 mg/ml	5.40 $\pm$ 0.14	5.72	2.27 $\pm$ 0.08	19.0	4.98
0.09 mg/ml	5.37 $\pm$ 0.17	5.62	2.43 $\pm$ 0.11	19.5	4.99
0.04 mg/ml	5.30 $\pm$ 0.21	5.46	2.35 $\pm$ 0.20	19.5	5.60
<b>137 mM NaCl (PBS)</b>					
0.17 mg/ml	5.55 $\pm$ 0.05	5.67	2.28 $\pm$ 0.06	21.0	4.98
0.13 mg/ml	5.35	5.80	2.43 $\pm$ 0.05	21.0	5.00
0.09 mg/ml	5.61 $\pm$ 0.11	5.85	2.21 $\pm$ 0.11	21.0	4.87
0.04 mg/ml	5.63	5.88	2.74	21.0	4.75
<b>50 mM NaCl (PBS)</b>					
0.17 mg/ml	5.96 $\pm$ 0.09	6.21	2.41 $\pm$ 0.04	20.0	5.25
0.13 mg/ml	5.82 $\pm$ 0.14	6.11	2.42 $\pm$ 0.11	20.0	5.12
0.09 mg/ml	5.84 $\pm$ 0.05	6.26	2.37 $\pm$ 0.08	19.0	5.38
0.04 mg/ml	6.01 $\pm$ 0.27	6.33	2.63 $\pm$ 0.33	19.5	5.77

<sup>a</sup> The first  $R_g$  value is from Guinier analyses, and the second one is from the  $P(r)$  analyses.

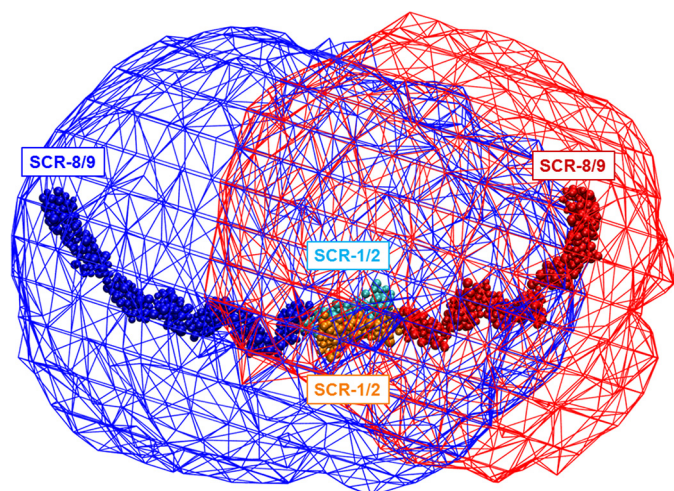
### Modeling the solution structure of the FHR5 dimer

Atomistic modeling of the FHR5 scattering data established the best-fit FHR5 dimer structures, hence providing a molecular explanation for its solution structure. The scattering curves for 0.17 and 0.5 mg/ml FHR5 in Tris-150 were used to assess good quality curves with no traces of aggregation and better signal/noise ratios at 0.5 mg/ml (Fig. 9). Data for 0.5 mg/ml were not available in PBS-137 or PBS-50, and traces of aggregates were present in PBS-50 buffer; thus, these data sets were not used.



**Figure 8. X-ray distance distribution  $P(r)$  analyses for FHR5.** The  $P(r)$  curves for FHR5 in A–D correspond to those shown in Fig. 6. In each panel, the  $P(r)$  curves were normalized for concentration and colored according to the FHR5 concentration from light blue at the lowest concentration to dark blue at the highest concentration. The maximum  $M$  depicts the most commonly occurring distances within the FHR5 structure. The length of FHR5 is signified by  $L$  at the  $r$  value where  $P(r)$  reaches zero.

The starting structure for the FHR5 dimer represented an extended conformation of the 18 SCR domains (Fig. 9). Each SCR domain was held fixed in conformation. Because as many as 14 linkers between the 18 domains were potentially variable, three different Monte Carlo conformational searches were set up. As detailed in Table 3, these varied all 14 linkers (Search 1) or eight linkers in which the crystal structure–observed linkers were kept fixed (Search 2) or four linkers after every third SCR domain (Search 3) (Fig. 1B) (see “Experimental procedures”). Initial Monte Carlo conformational simulations in Searches 1–3 gave many models that were too elongated with too large  $R_g$  values and few models with low  $R_g$  values close to the



**Figure 9.** Density plot of the conformationally varied FHR5 structures. The linear dimeric FHR5 starting structure is shown at the center in blue and red for the two monomers, with the SCR-1/2 dimer at the center of this. The grid shows the complete spatial extent covered by the 72,755 modeled conformations of Search 2 for each FHR5 monomer (Table 3), shown in blue and red.

experimental  $R_g$  value of  $\sim 5.5$  nm. Thus, in further simulations, models were selected with  $R_g$  values closer to the experimental  $R_g$  value to generate further conformers, but now using an  $R_g$  cut-off of 6.0 nm as a constraint to generate more compact FHR5 dimers. This resulted in more structures with lower  $R_g$  values; however, many of these models were rejected by the workflow because the more compact shapes gave rise to physically disallowed steric clashes between the SCR domains.

All six analyses from the three searches at two FHR5 concentrations gave a clear single minimum in the distribution of  $R$ -factor goodness of fit values (Fig. 10). A lower  $R$ -factor indicated a better fit to experiment. Thus, all three searches successfully generated good-fit solution structures for the FHR5 dimer. Starting from 200,000–250,000 trial structures in Searches 1–3, 86,732 structures with no steric clashes were accepted for Search 1, and likewise 72,755 structures for Search 2 and 123,776 structures for Search 3 (yellow in Fig. 10). To verify the Monte Carlo-generated conformations, a grid density plot was generated for the Search 2 library of models (Fig. 9). The volumetric data showed that a full conformational range of structures had been sampled, compared with the starting FHR5 dimer model at the center of the grid. Significantly, the experimental  $R_g$  value of 5.36 nm occurred at the left of the distribution plots in Fig. 10, clearly indicating that FHR5 has a compact domain structure. In distinction, linear FHR5 models showed higher  $R_g$  values of over 8 nm.

The three sets of 72,755–123,776 models were each filtered to identify the best-fit structures (Table 3). The appearance of the  $R_g$  versus  $R$ -factor graphs was similar in all six fits (Fig. 10). This showed that the outcome of the modeling was independent of the assumption used to generate the linkers. As required, the dimer models with the lowest  $R$ -factors of 4–5% agreed well with the experimental  $R_g$  value of 5.36 nm. The most extended FHR5 structures with the largest  $R_g$  values of 8 nm and above showed the highest  $R$ -factors of  $\sim 30\%$ . No models had an  $R_g$  of 4.5 nm or less because such a dimer would be too compact to be sterically allowed. Filters were now used to reject poor-fit

structures. First, a  $\pm 5\%$  experimental  $R_g$  filter was used to reject models that had  $R_g$  values outside this range, followed by a  $\pm 5\%$   $R_{\text{xs}}$  filter. Models with an  $R$ -factor below 6% were then selected. For the two fits of Search 1 (Table 3), totals of 28 and 131 models were identified (green in Fig. 10A). For Search 2, totals of 55 and 52 models were identified (Fig. 10B). For Search 3, totals of 694 and 749 models were identified (Fig. 10C). These best-fit models formed a single cluster of fits at the  $R$ -factor minima. The best-fit models with the lowest  $R$ -factors (red in Fig. 10) had  $R$ -factors of 4.5 and 4.2% for Search 1, 4.7 and 3.9% for Search 2, and 4.3 and 3.8% for Search 3. For comparison, the parameters for the best-fit 100 models are also shown in Table 3.

Visual inspection of the fits between the theoretical and experimental SAXS  $I(Q)$  and  $P(r)$  curves showed good agreement (Fig. 11). The  $M$  and  $L$  values of the  $P(r)$  curves were well-reproduced. Kratky plots of the SAXS curves monitor whether the protein was compact and globular or was extended and disordered in its structure. The normalized Kratky analyses of  $(Q.R_g)^2 \cdot I(Q)/I(0)$  versus  $Q.R_g$  for the three best-fit models from Searches 1–3 and the experimental curve at 0.5 mg/ml showed that a clear peak was seen at a  $Q.R_g$  of 2.26 (Fig. 12). Good fits to the experimental curve were also obtained at larger  $Q.R_g$  values for all three best-fit models. The Kratky plot thus showed that FHR5 possessed a globular structure with little interdomain flexibility. In comparison, our recent Factor H models showed poorer fits at larger  $Q.R_g$  values, indicating that the 20 SCR domains in Factor H showed more flexibility (12). This comparison indicated that the structure of FHR5 was well-formed, and this was less flexible in structure than full-length Factor H.

Because all three searches gave similar good fits, Search 2 was selected for the final output because this most closely resembled the crystal structures for the SCR domain pairs used to construct it. To understand better the 55 best-fit structures from Search 2 (available in the supporting material), they were clustered into conformational families using principal component analysis (Fig. 13) (21, 22). Principal component analysis determines the correlated motions of protein residues as linearly uncorrelated variables termed principal components. These “essential motions” are extracted from a covariance matrix of the atomic coordinates of the frames in the trajectory. The eigenvectors of this matrix each have an associated eigenvalue that characterizes the clustering of the models based on structural coordinates (or variance). By this, the first three eigenvalue rankings (PC1–PC3) accounted for a variance of 68.9% in the 55 best-fit FHR5 models. The median FHR5 structure from each principal component analysis group consistently revealed folded-back N-terminal domains and extended C-terminal domains (Fig. 14).

#### Sedimentation coefficient modeling of the FHR5 dimer

As an independent test of the SAXS modeling, the theoretical  $s_{20,w}$  values were calculated using HYDROPRO for the best-fit FHR5 dimer models obtained from the three searches, Searches 1–3 (Table 3). The six best-fit models gave a mean  $s_{20,w}$  value of  $5.3 \pm 0.2$  S. This compared well with the experimental  $s_{20,w}$



## Solution structure of FHR5

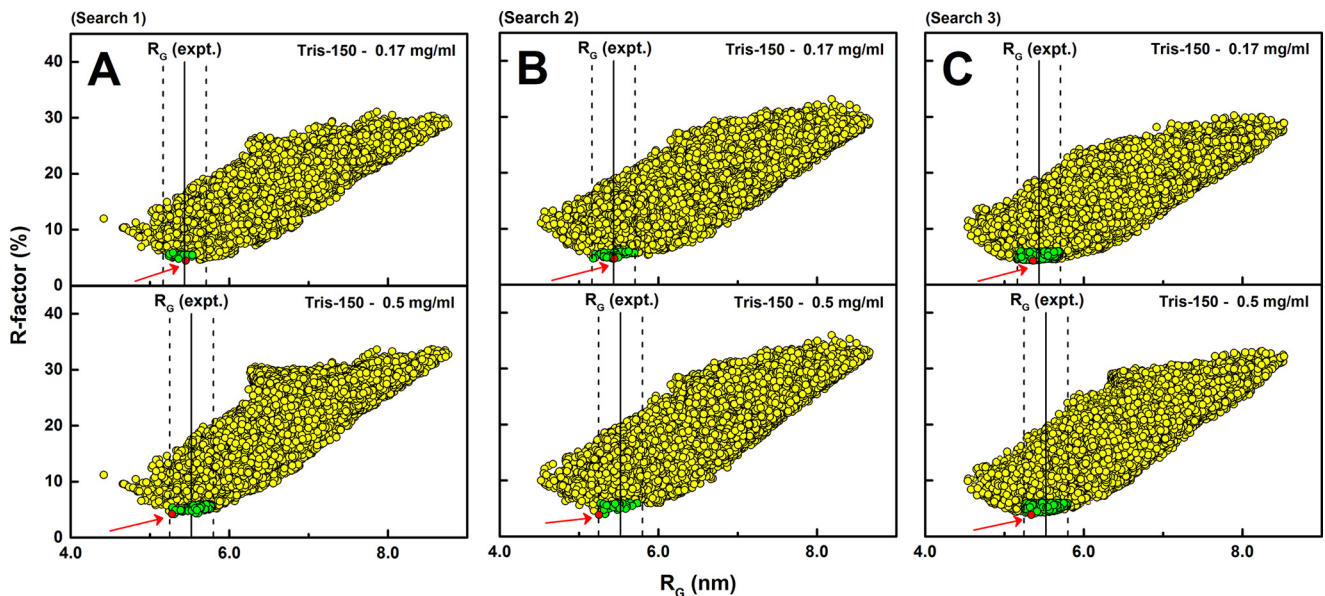
**Table 3**

Three modeling fit searches for FHR5 using X-ray scattering curve fits and sedimentation coefficients

Search	Filter	Models <sup>a</sup>	$R_g^b$	$R_g^b$	$R_{xs}$	$D_{max}$	R-factor	$s_{20,w}$
			<i>nm</i>	<i>nm</i>	<i>nm</i>	<i>nm</i>	%	<i>S</i>
Search 1 (all linkers 2–8)	None	86,732	4.42–8.29	NA	0.01–4.23	NA	4.46–27.14	NA
Tris-150 0.17 mg/ml	R-factor	100	5.24–6.02	NA	2.00–2.73	NA	4.46–5.80	NA
	R-factor, $R_g$ , $R_{xs}$	28	5.24–5.54	NA	2.25–2.48	NA	4.46–5.39	NA
	Best fit	1	5.45	5.67	2.25	20.0	4.46	5.2
	Experimental	NA	5.44	5.72	2.39	20.0	NA	NA
Tris-150 0.5 mg/ml	R-factor	100	5.24–5.84	NA	2.33–2.65	NA	4.21–5.30	NA
	R-factor, $R_g$ , $R_{xs}$	131	5.28–5.79	NA	2.33–2.57	NA	4.21–5.30	NA
	Best fit	1	5.31	5.63	2.44	20.0	4.21–5.98	5.2
	Experimental	NA	5.52	5.85	2.45	21.0	NA	NA
Search 2 (linkers 2, 4, 6, 7)	None	72,755	4.52–8.66	NA	0.01–4.66	NA	4.68–30.89	NA
Tris-150 0.17 mg/ml	R-factor	100	5.01–5.88	NA	2.19–2.71	NA	4.68–6.01	NA
	R-factor, $R_g$ , $R_{xs}$	55	5.18–5.70	NA	2.27–2.51	NA	4.68–5.84	NA
	Best fit	1	5.45	5.89	2.37	20.0	4.68	5.2
	Experimental	NA	5.44	5.72	2.39	20.0	NA	NA
Tris-150 0.5 mg/ml	R-factor	100	5.01–6.02	NA	2.19–2.82	NA	3.86–6.15	NA
	R-factor, $R_g$ , $R_{xs}$	52	5.25–5.74	NA	2.57–2.34	NA	3.86–5.99	NA
	Best fit	1	5.25	5.38	2.48	20.0	3.86	5.5
	Experimental	NA	5.52	5.85	2.45	21.0	NA	NA
Search 3 (linkers 3 and 6)	None	123,776	4.54–7.89	NA	0.004–4.57	NA	4.28–28.14	NA
Tris-150 0.17 mg/ml	R-factor	100	5.16–5.69	NA	2.14–2.53	NA	4.28–4.61	NA
	R-factor, $R_g$ , $R_{xs}$	694	5.17–5.69	NA	2.27–2.51	NA	4.28–6.00	NA
	Best fit	1	5.36	5.58	2.37	20.0	4.28	5.4
	Experimental	NA	5.44	5.72	2.39	20.0	NA	NA
Tris-150 0.5 mg/ml	R-factor	100	5.13–5.66	NA	2.22–2.69	NA	3.84–4.65	NA
	R-factor, $R_g$ , $R_{xs}$	749	5.26–5.79	NA	2.33–2.57	NA	3.84–5.99	NA
	Best fit	1	5.34	5.64	2.52	20.0	3.84	5.1
	Experimental	NA	5.52	5.85	2.45	21.0	NA	NA

<sup>a</sup> Total number of models accepted after Monte Carlo simulations and after model filtering. The best fit model corresponds to that with the lowest R-factor in the filtered models.

<sup>b</sup> The first  $R_g$  value of the pair is from Guinier analyses, and the second is from the  $P(r)$  analyses. NA, not available.

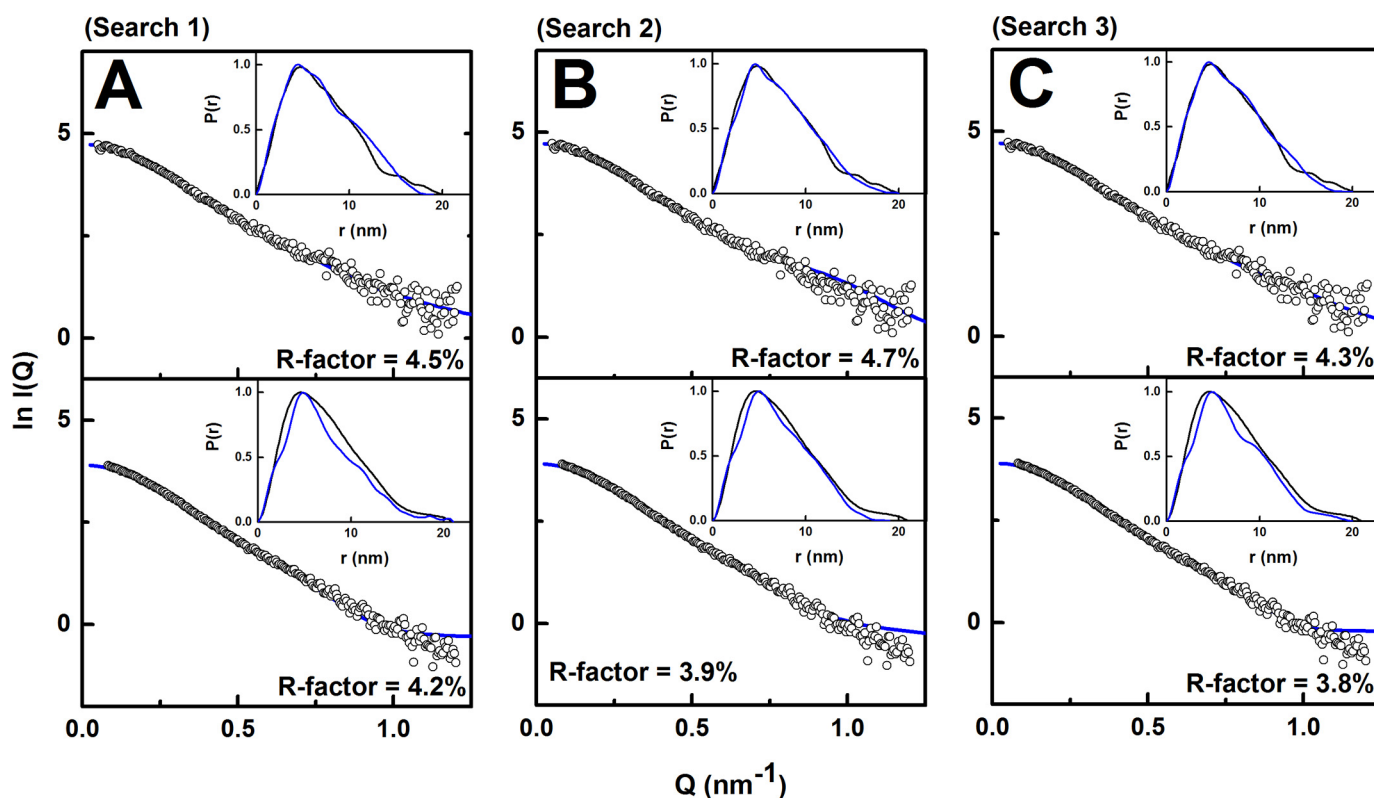


**Figure 10. R-factor analyses for the atomistic modeling of FHR5 dimers.** The goodness-of-fit R-factors are compared with the calculated  $R_g$  values for the 86,732, 72,755, and 123,776 conformationally randomized FHR5 dimer models from three different Monte Carlo searches. The yellow circles denote the models in each search. The experimental  $R_g$  values are shown by the solid vertical lines with  $R_g$  error ranges of  $\pm 5\%$  represented by the dashed lines. The 28–749 best fit models for each search after filtering for the R-factor,  $R_g$ , and  $R_{xs}$  values are shown in green. The single best-fit model with the lowest R-factor at the minima (red arrow) is in red. A, the models were compared with data for 0.17 and 0.5 mg/ml FHR5 in Tris-150, varying linkers 2–8 (Search 1). B, the models were compared with the same two data sets, but varying linkers 2, 4, 6, and 7 (Search 2). C, the models were compared with the same two data sets, but varying linkers 3 and 6 (Search 3).

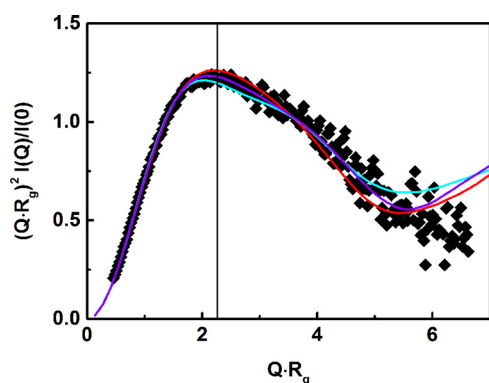
value in PBS-137 of  $5.97 \pm 0.2$  S (Table 1). The typical accuracy of the  $s_{20,w}$  calculation is  $\pm 0.21$  S (23). The difference of 0.67 S may result from potential trace aggregates remaining in the X-ray sample, which would increase the experimental and modeled  $R_g$  values of FHR5 and in turn decrease the modeled  $s_{20,w}$  value.

## Discussion

Up to now, the domain organization of FHR5 was unknown. Here we present the first protein structures for the FHR5 dimer by a combination of SAXS and AUC in conjunction with molecular simulations. Previously, it was often thought that FHR5 possessed nine SCR domains in a flexible linear conformation



**Figure 11. X-ray scattering curve fits for best-fit FHR5 models.** The best fits correspond to the models with the lowest  $R$ -factor (red in Fig. 9). The experimental data at 0.17 mg/ml (top panels) and 0.50 mg/ml (bottom panels) are indicated by circles, overlaid with the modeled scattering curve (blue line). The insets show the experimental (black) and modeled (blue)  $P(r)$  curves. A, B, and C show the best fit from Search 1, 2, and 3, respectively (Table 3).



**Figure 12. Normalized dimensionless Kratky plots of the best-fit curves of FHR5 models.** The experimental data for 0.5 mg/ml FHR5 in Tris-150 are denoted by diamonds. The modeled curves for 0.5 mg/ml in Tris-150 according to Search 1, Search 2, and Search 3 are shown in red, cyan, and purple, respectively. The vertical solid line denotes the maximum of the Kratky plot.

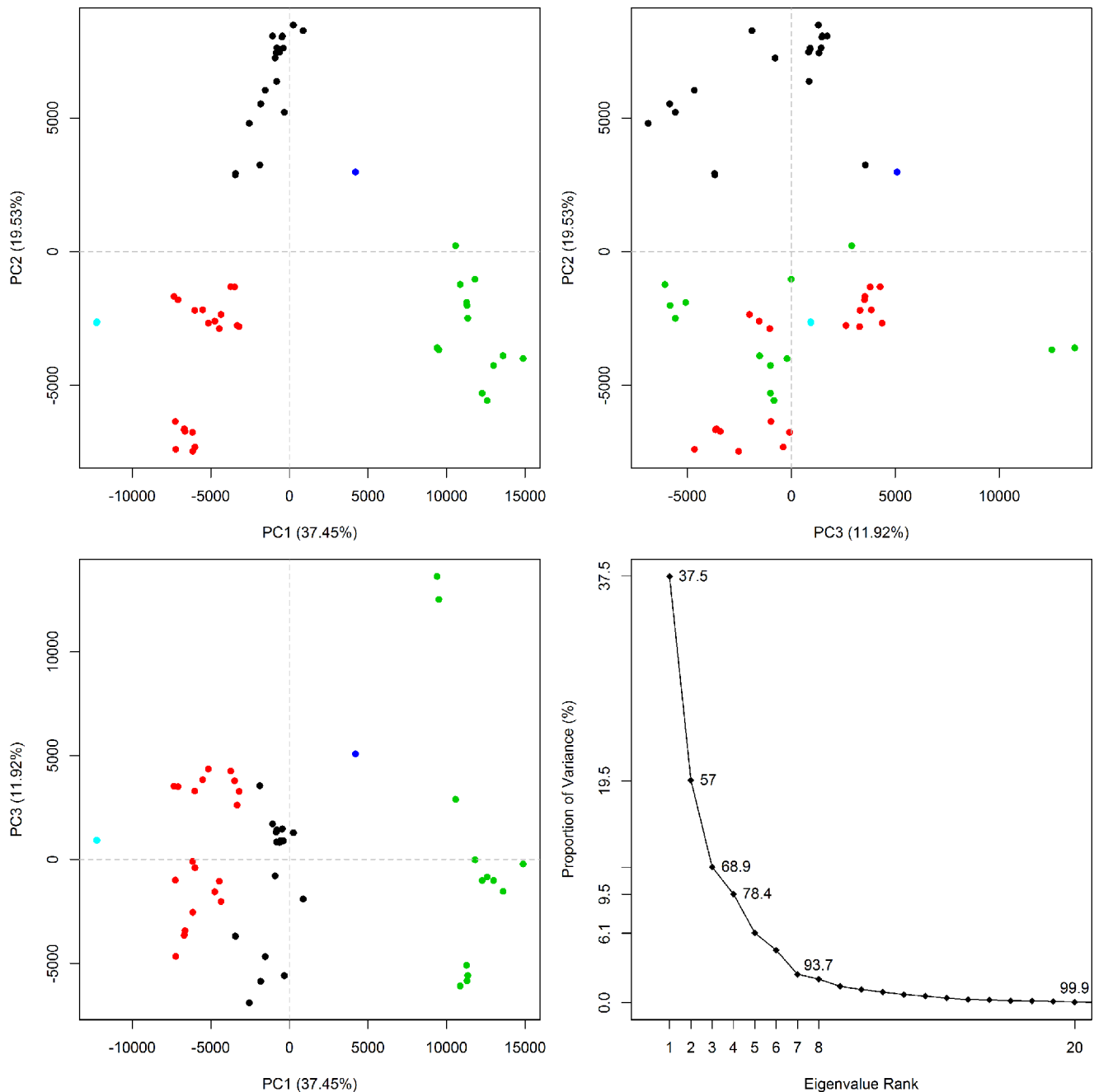
(4, 8, 24–26). Instead, our analyses now show that FHR5 is dimeric and adopts a compact domain conformation. Such a structure readily leads to FHR5 oligomer formation in the presence of mutant FHR5 protein (see below). This structure revises our understanding of how FHR5 interacts with its target ligand C3b and its C3d fragment, as well as others, such as heparin-like analogues. It also explains the molecular defect underlying CFHR5 nephropathy.

New understandings of the FHR5 solution structure were determined as follows.

(i) Our SEC-MALLS and AUC data showed that full-length FHR5 SCR-1/9 is a dimer (Figs. 3 and 4), in agreement with previous results for the FHR proteins (4, 6, 7). In addition, AUC monitors macromolecular shapes through the  $s_{20,w}$  values, which measure macromolecular elongation. Of interest here was that, not only did the  $s_{20,w}$  values correspond to a much more compact protein than expected from the 18 domains in the dimer (Table 3), but also these  $s_{20,w}$  values decreased with an increase in the NaCl concentration of the buffer. This decrease implied that the compact structure became more elongated through the weakening of charge-charge interactions between the SCR domains. The predicted pI values of the N-terminal five domains SCR-1/5 were mostly acidic at 4.6, 5.4, 8.5, 4.7, and 4.3 in that order, whereas the predicted pI values of the four C-terminal domains SCR-6/9 were mostly basic at 9.6, 6.3, 8.9, and 8.4 in that order (RRID:SCR\_018087). Differences in these individual pI values may facilitate the formation of a more compact FHR5 domain structure through charge attractions in physiological 137 mM NaCl salt.

(ii) The SAXS data provided more detailed views of the FHR5 structure. Interestingly, given that the SAXS technique is sensitive to aggregate formation, both FHR5 and Factor H turned out to be aggregation-prone. The  $R_g$  value of Factor H was originally reported to be 12.4 nm in the first SAXS studies in 1991 for reason of being aggregated; with improved Factor H purifications, this value has now diminished to 7.22–7.77 nm (12). Factor H aggregates in storage conditions. FHR5 as supplied for our study showed aggregation by SEC-MALLS and

## Solution structure of FHR5



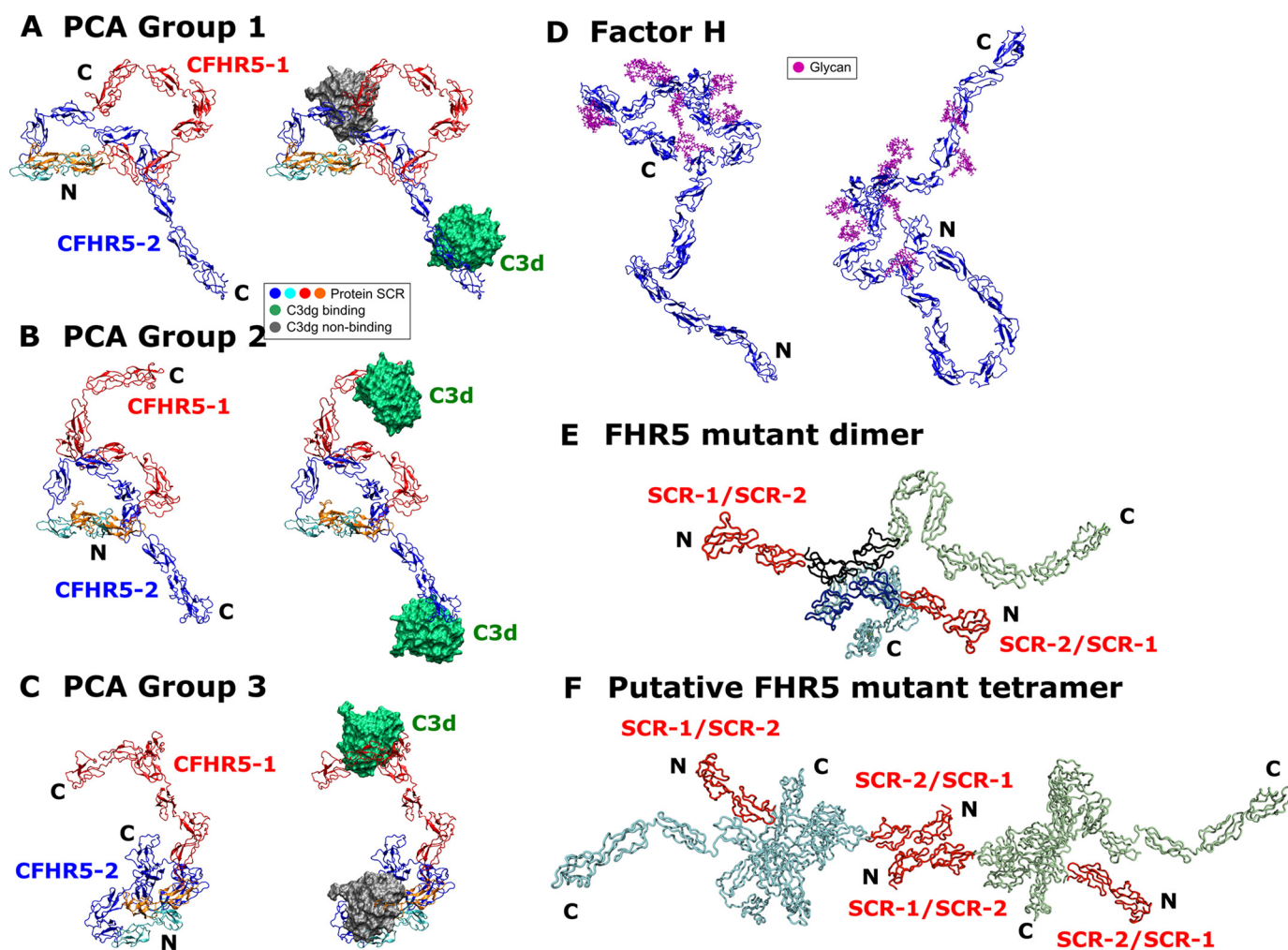
**Figure 13.** Principal component analyses of the FHR5 best-fit models from Search 2. The 55 best-fit models (Table 3) were grouped by principal component analysis into five groups, of which three groups were predominant in terms of the number of 48 models they contained. These are shown as PC1, PC2, and PC3 (black, 19 models; red, 16 models; green, 13 models) and exemplified by the first three principal components (PC2 versus PC1 and PC3 versus PC2).

AUC, and these aggregates were removed by size-exclusion chromatography. SAXS showed that the  $R_g$  and  $R_{xs}$  values for FHR5 were relatively constant in 50–150 mM NaCl and between 0.1–0.5 mg/ml, although residual trace aggregates were detectable in 50 mM NaCl buffer. The maximum length  $L$  of FHR5 was 20–21 nm in all buffers (Table 2). A single SCR domain is about 4 nm in length. A hypothetical fully extended FHR5 domain arrangement (Fig. 1A) would be predicted to be of length 64 nm, or over 3-fold longer than seen experimentally (Fig. 8). Likewise, Factor H is predicted to be 80 nm in length if fully extended, but it was observed to be only 26–29 nm in length, so again such an extended structure is also predicted to

be 3-fold longer than seen experimentally (12). Both FHR5 and Factor H thus have similar folded-back domain structures.

(iii) Because no high-resolution FHR5 domain structures were available, the starting model for FHR5 was generated by standard homology modeling methods based on sequence similarities. The FHR5 SCR-1/2 and SCR-8/9 structures were readily modeled on other FHR proteins. These modeled domain pairs were notable for their short linker lengths of three residues each, suggesting that these linkers were relatively inflexible (Fig. 1C). The longest inter-SCR linkers occurred between SCR-3/7, which were six, six, eight, and seven residues in length, respectively. Interestingly, the same linker lengths





**Figure 14. The three best-fit FHR5 models, their interaction with C3d, and mutant FHR5.** A–C, ribbon views of single FHR5 models, these being representative of the three principal component analysis (PCA) groups determined from Search 2 (Table 3). The models correspond to the X-ray curve fits in Fig. 11B. The two monomers are shown in red and blue, with the N-terminal SCR-1/2 dimer pair denoted by N and shown in cyan and orange. The C-terminal SCR-9 domains are denoted by C. To their right are shown surface views of two C3d molecules bound to the two SCR-9 domains in the FHR5 dimer shown in the same orientation. If binding is sterically allowed, the C3d surface is shown in green; if this is sterically blocked, this is grayed out. D, the two best fit structures (to the same scale) of the X-ray scattering models of glycosylated Factor H are shown. Purple denotes the eight N-glycan chains in Factor H. E, the mutant FHR5 dimer structure was generated from the best-fit structure from Search 2 (Table 3) by the addition of two extra SCR domain pairs to represent the mutant. The native SCR-1/2 pair is shown in black and blue at the center, with the extra SCR-1/2 pair shown in red at the two N termini of the native SCR-1/2 pair. F, the putative daisy-chaining of mutant FHR5 dimers to form a tetramer. Pairs of mutant SCR-1/2 domains are formed based on the crystal structure of this dimer. Such a tetramer can be extended to form hexamers and larger oligomers, with chain extension limited by binding of a WT FHR5 molecule.

occurred in SCR-10/14 of Factor H. In fact, sequence similarities showed that these five SCR domains resembled SCR-3/7 of FHR5. These Factor H domains contributed significantly to its folded-back solution structure (11, 12, 27, 28). These long linkers in Factor H and FHR5 contained a high proportion of charged residues, particularly lysine and glutamate, and are conserved in mouse and bovine factor H (29). Indeed, SCR-10/14 of factor H not only has longer interdomain linkers, but also has shorter SCR sequences and higher glycosylation levels (30). These similarities imply that these middle domains act as conformational spacers that result in more compact domain structures that enable the multiple factor H and FHR5 binding sites to act synergistically.

(iv) The Monte Carlo simulations generated a large conformational library of possible SCR arrangements in FHR5, from which best-fit structures were identified. These best-fit structures accounted for the experimental SAXS and AUC data for

FHR5. Interestingly, the Kratky plots (Fig. 12) did not show evidence of disorder or flexibility in the FHR5 solution structure, meaning that its structure was well-defined. The molecular structures for FHR5 and Factor H show similar folded-back and compact SCR structures (Fig. 14, A–D). From the principal component analyses, the three best-fit FHR5 structures (see supporting material) showed that, whereas the SCR-1/2 dimer pair was consistently buried in the dimer core in all three structures, the two SCR-3/4 domain pairs looped back across the SCR-1/2 core in a compact arrangement with SCR-1/2. The two C-terminal ends with SCR-5/9 were solvent-exposed and either extended away from the SCR-1/4 core or looped back toward this core. The functional SCR-8/9 domains thus showed a range of folded-back or extended conformations relative to a more compact SCR-1/4 core. The three best-fit conformations were able to interact with one or two C3d ligands (Fig. 14, A–C).

## Solution structure of FHR5

In terms of new functional insight obtained from this study, FHR5 is a complement deregulator that competitively inhibits factor H, an important regulator of C3b activation at host cell surfaces (4, 5). From the nine best fit structures (Table 3), the two C3d- or C3b-binding sites found in dimeric FHR5 will have a C-terminal separation of around 10–20 nm. FHR5 would increase its avidity for C3d- or C3b-coated host cell surfaces only if bound C3d or C3b were present at a great enough spatial density on this surface, thus displacing the binding of the much more abundant Factor H. If the spatial density of C3b/C3d on surfaces is low, the much more abundant Factor H will preferentially bind to inhibit and degrade C3b there via Factor H-mediated cleavage. When the density of C3b/C3d is great enough to allow dimeric FHR5 binding to be functionally bivalent (Fig. 14B), the FHR5-C3d interaction becomes stronger than the monovalent Factor H-C3d interaction. This reasoning indicates a mechanism for FHR5 to modulate Factor H activity.

In CFHR5 nephropathy, the heterozygous duplication of SCR-1/2 results in a more elongated FHR5 molecule that is detectable in the blood of patients (Fig. 14, E and F) (8). Other heterozygous genomic rearrangements that result in the production of more elongated FHR proteins with additional N-terminal SCR-1/2 domains have been described in association with autosomal dominant C3 glomerulopathy (25, 31). Conversely, clear loss-of-function variants in *CFHR5* occur at high frequency in the population and are not known to be pathogenic. Among the ~245,000 alleles tested in GnomAD ([/gnomad.broadinstitute.org/gene/ENSG00000134389](http://gnomad.broadinstitute.org/gene/ENSG00000134389)), ~3000 variants predicted to stop FHR5 protein translation before the final exon are documented. In addition, 3% of the UK population is homozygous for a *CFHR3/CFHR1* deletion polymorphism that results in the complete deficiency of FHR1 and FHR3 (32). Together, these observations suggest that a gain-of-function mechanism underlies CFHR5 nephropathy and that tandem duplication of the two N-terminal SCR-1/2 domains is necessary and sufficient to cause this. Structural simulations using our FHR5 models show that the extra SCR-1/2 domains of the mutant FHR5 protein are readily added, and these will be accessible to other FHR5 molecules (Fig. 14, E and F). At least two distinct mechanisms can be proposed by which the mutation in CFHR5 nephropathy causes augmented function (*i.e.* increased avidity for C3-coated surfaces). In one, as proposed previously, the presence of two accessible SCR-1/2 dimerization motifs on the single mutant protein would allow trimers or higher-order oligomers to form that would be trimeric or multivalent with respect to C3d (Fig. 13, E and F). In addition, the greater length of macromolecules containing mutant FHR5 would reduce the density of C3d on a host cell surface required for multivalent binding to occur, because the longer protein would have a greater steric range.

Overall, it is expected that different tissues will function differentially with respect to FHR5- or Factor H-binding activity. It is possible that a high blood flow rate, such as that in the renal glomeruli, enables the density of C3b or C3d deposition to become high enough to allow FHR5 dimers to bind bivalently. This explains why FHR5 is enriched in C3-coated glomeruli and why *CFHR5* gain-of-function mutations result in the purely renal disease of CFHR5 nephropathy, which manifests clinically

at times of infection when the complement system is systemically activated. The striking clinical and histological similarity of IgA nephropathy to CFHR5 nephropathy, combined with the co-localization of FHR5 and C3 in the glomerulus in IgA nephropathy (1), raises the possibility that the FHR proteins, including FHR5, play an important role in both diseases. This possibility is supported by the observation that, in IgA nephropathy, a common polymorphic deletion of *CFHR1* (which encodes the smaller dimeric complement deregulator FHR1) is protective (33). Our demonstration of a compact FHR5 dimer structure at a molecular level therefore reveals new aspects of how FHR5 antagonizes Factor H function, amplifying complement activation at host cell surfaces when C3 deposition reaches a critical density and leading to renal damage.

## Experimental procedures

### Purification and composition of full-length FHR5

Mammalian-expressed (HEK293 cells) human FHR5 SCR-1/9 was purchased from Creative Biolabs (Shirley, NY, USA). This was prepared with a His tag, which was cleaved off by the manufacturer. This protein was prone to aggregation. Aggregate-free FHR5 for SAXS was successfully purified from ~1 mg of protein that was pooled and concentrated using a Vivaspin 20 spin concentrator (Sartorius) with a 10-kDa molecular mass cut-off and then purified using a Superdex 200 10/300 GL gel-filtration column (Cytiva) equilibrated in 50 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 7.4, using a Gilson HPLC system kindly made available by Dr. A. J. Beavil (Kings College London). The FHR5 concentration was checked by the absorbance reading at 280 nm. Its purity and integrity were checked by SDS-PAGE before and after each SAXS and AUC experiment under reducing and nonreducing conditions using a Novex® 8–12% Bis-Tris gel 1.0 mm (Invitrogen, Paisley, UK).

The amino acid composition of human FHR5 SCR-1/9 was determined from its sequence (SWISSPROT accession code Q9BXR6). Two potential N-linked glycan sites were present at Asn-126 and Asn-400 (Fig. 14) and may be occupied by biantennary glycans as reported for Factor H (30). However, there was no evidence that these sites were occupied, in particular at Asn-126, where glycan was not present in the crystal structure of HEK293-expressed FHR1 SCR-1/2 (Protein Data Bank (PDB) code 3ZD2) (4). Because FHR1 SCR-1/2 has the same glycosylation sequence as that in FHR5 (Fig. 1C), glycosylation was disregarded here. The mass of glycan-free WT FHR5 was predicted to be 62,377 Da from its sequence. Using the program SLUV (20), it has an unhydrated volume of 79.76 nm<sup>3</sup>, a hydrated volume of 105.23 nm<sup>3</sup>, a partial specific volume of 0.7278 nm<sup>3</sup>, and an absorption coefficient of 15.59 (1%, 280 nm, 1-cm path length).

FHR5 samples were run through SEC-MALLS. This determines protein molecular masses using a standard HPLC system equipped with a Superdex 200 Increase 5/150 GL gel-filtration column (Cytiva). The instrument was equipped with three detectors, namely a miniDawn detector (Wyatt Technology), which is a triple-angle light scattering detector; an Optilab DSP Interferometric Refractometer (Wyatt Technology), which measures refractive index changes; and an SPD-20A UV absorbance

detector (Shimadzu Scientific). In multiple runs, 60- $\mu\text{l}$  aliquots of FHR5 were loaded on the column via an injection loop. Following separation by size exclusion, the three different detectors were combined in parallel to provide a molecular mass for the eluted sample. The chromatograms were analyzed using ASTRA software (Wyatt Technology).

For AUC and SAXS experiments, FHR5 was dialyzed into the appropriate buffer at 4 °C prior to data collection. For AUC, these buffers were PBS-20 (20 mM NaCl plus 8.2 mM  $\text{Na}_2\text{HPO}_4$ , 2.6 mM KCl, 1.5 mM  $\text{KH}_2\text{PO}_4$ , pH 7.4), PBS-50 with 50 mM NaCl, PBS-90 with 90 mM NaCl, PBS-137 with 137 mM NaCl (standard physiological salt), PBS-250 with 250 mM NaCl, and PBS-350 with 350 mM NaCl. For SAXS, three buffers were used, namely Tris-150 (50 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 7.4) and PBS-50 and PBS 137 as above. The experimental buffer densities were measured at 20 °C using an Anton Paar DMA 5000 density meter, and their theoretical values were calculated from SEDNTERP (34). The resulting densities were 1.000538 g/ml for PBS-20 (theoretical, 1.00052 g/ml), 1.001714 g/ml for PBS-50 (theoretical, 1.00176 g/ml), 1.003382 g/ml for PBS-90 (theoretical, 1.00342 g/ml), 1.005054 g/ml for PBS-137 (theoretical, 1.00524 g/ml), 1.009960 g/ml for PBS-250 (theoretical, 1.00999 g/ml), 1.013920 g/ml for PBS-350 (theoretical, 1.01406 g/ml), and 1.00650 g/ml for Tris-150 (theoretical, 1.00603 g/ml). A solvent viscosity of 0.01002 cp was used throughout in the AUC analyses.

#### Sedimentation velocity data collection and analyses for FHR5

AUC data were obtained on a Beckman XL-I instrument, equipped with an eight-hole AnTi50 rotor (Beckman-Coulter Inc., Palo Alto, CA). Ultracentrifugation caused any aggregates present to sediment rapidly, leaving the soluble FHR5 protein visible for analysis. Approximately 400  $\mu\text{l}$  of FHR5 sample was loaded into standard AUC double-sector cells for sedimentation velocity experiments at 20 °C, equipped with sapphire windows and with 12-mm column heights. Sample concentrations were 0.16 mg/ml; therefore, absorbance optics was used to collect data. Up to 500 consecutive scans were recorded until the protein had fully sedimented. The AUC runs were performed using two rotor speeds of 40,000 and 50,000 rpm to check for reproducibility.

Data analysis was performed using SEDFIT software (version 14.6) (35, 36), using direct boundary Lamm fits of up to 50 selected scans at appropriately spaced time intervals. A  $c(s)$  size-distribution analysis was carried out, which assumes that all species have the same frictional ratio  $ff_0$ . The  $c(s)$  distribution was optimized by floating the value of the meniscus and bottom of the cell positions, the baseline, and the frictional ratio  $ff_0$  (set at 1.2 to begin with). Fits were carried out until satisfactory visual fits and overall root mean square deviations were obtained. The final SEDFIT analysis used a resolution of 200, and the sedimentation coefficient  $s_{20,w}$  for FHR5 was determined from the peak maximum in the  $c(s)$  size-distribution plot. The  $c(s)$  integration function was also used to derive the percentage of oligomers in the total loading concentration if required.

#### SAXS data collection and data analyses for FHR5

SAXS experiments were carried out in one beam session on the BM29 BioSAXS beamline at the European Synchrotron Radiation Facility (Grenoble, France), operating with a ring energy of 6.0 GeV. Data were acquired using a Pilatus 1M two-dimensional detector with a pixel size of 172  $\mu\text{m}$ . The sample-to-detector distance was 3.0 m. The beamline was equipped with an automatic sample changer, and the samples were loaded using the thermo-regulated PCR tube configuration in the BsxCuBE control interface. The FHR5 samples were measured in three buffers (above) at concentrations of 0.04, 0.09, 0.13, and 0.17 mg/ml. Additional data sets were collected at 0.2, 0.3, 0.4, and 0.5 mg/ml concentrations in Tris-150 buffer. Data were collected in triplicate from a total sample volume of 50  $\mu\text{l}$  per run. An exposure time of 1 s was used, and the absence of radiation damage was monitored from continuous automatic online checks. A total of 10 frames were collected as the sample was passed continuously through a quartz capillary tube (1.8 mm in diameter) to minimize radiation damage due to exposure. The final time frames were merged, excluding any damaged data, to improve the signal/noise ratio. Between each sample measurement, the sample capillary was cleaned using Hellmanex<sup>®</sup> and water to ensure the removal of any residual protein or aggregates on the capillary walls (37).

The raw scattering data files were corrected by subtraction of the buffer data from the sample data. The resulting one-dimensional scattering curve  $I(Q)$  in a  $Q$  range between 0.05 and 2  $\text{nm}^{-1}$  (where  $Q = 4\pi \sin \theta/\lambda$ ;  $2\theta$  is the scattering angle and  $\lambda$  is the wavelength) represented the macromolecular structure. Guinier analysis of  $I(Q)$  against  $Q^2$  at low  $Q$  values gave the radius of gyration,  $R_g$ , which is a measure of structural elongation if the internal inhomogeneity of the scattering densities has no effect, and the forward scattering at zero angle  $I(0)$ .

$$\ln I(Q) = \ln I(0) - \frac{R_g^2 Q^2}{3} \quad (\text{Eq. 1})$$

$$\ln (I(Q) \cdot Q) = (I(Q) \cdot Q)_{Q \rightarrow 0} - \frac{R_{xs}^2 Q^2}{2} \quad (\text{Eq. 2})$$

The Guinier plots are usually valid in a  $Q$  range up to  $Q \cdot R_g$  values of 1.5 (38). If the macromolecular structure is elongated, the mean cross-sectional radius of gyration  $R_{xs}$  is obtained from plots of  $I(Q) \cdot Q$  against  $Q^2$  in a larger  $Q$  range than those used for the  $R_g$  values. Using the SCT software package (39), the  $Q$  ranges for the  $R_g$  and  $R_{xs}$  values were 0.1–0.27  $\text{nm}^{-1}$  and 0.32–0.55  $\text{nm}^{-1}$ , respectively.

Indirect Fourier transformation of the scattering curve  $I(Q)$  in reciprocal space (units in  $\text{nm}^{-1}$ ) into real space (units in nm) gives the distance distribution function  $P(r)$ . This transformation was carried out using the program GNOM (40).

$$P(r) = \frac{1}{2\pi^2} \int_0^\infty I(Q) Q r \sin(Qr) dQ \quad (\text{Eq. 3})$$

$P(r)$  corresponds to the distribution of interatomic distances  $r$  in the macromolecule. To obtain the distance distribution  $P(r)$  curve, the full measured scattering curve was utilized. By



## Solution structure of FHR5

specifying an assumed maximum dimension  $D_{\max}$ , the  $P(r)$  curve provides the macromolecular length  $L$  and the most common distance  $M$ . The  $P(r)$  curves also provide an alternative calculation of  $R_g$  for comparison with the Guinier analysis.

### Generating of the starting model for FHR5

Protein structural analyses of FHR5 were initiated from homology models for the nine SCR domains. First, suitable templates were selected based on high sequence and structural similarities. This was achieved by a combination of PDB-BLAST searches and sequence alignments between the five FHR-related proteins and Factor H. The final template was selected from the quality of the sequence alignment and its structural relevance. The template structures were taken from closely related structures in the PDB. In the process, the amino acid sequence of FHR5 SCR-1/9 was used to replace the sequence of the template structure. These were constructed using closely related structural templates using MODELLER (version 9.15) (41).

The closest template for each of the nine SCR domains, defined in terms of sequence identity and minimum insertions and deletions, was identified using CLUSTALO alignments (42). Four template structures for eight domains were used as follows (Fig. 1, B and C): FHR1 SCR-1/2 (PDB code 3ZD2), Factor H SCR-10/11 (PDB code 4B2R), Factor H SCR-12/13 (PDB code 2KMS), and FHR2 SCR-3/4 (PDB code 3ZD1). The ninth domain was SCR-7, for which a multiple-sequence alignment (43) was performed using the NMR structures of Factor H SCR-10/11 (PDB code 4B2R) (28) and Factor H SCR-11/12 (PDB code: 4B2S) (28), which provided an experimental structure for Factor H SCR-11. The full-length FHR5 model was evaluated using the SAVES server (RRID:SCR\_018219), which incorporated validation criteria including PROCHECK and Ramachandran plots. The secondary structure and surface accessibilities of the FHR5 model were analyzed using the Definition of Secondary Structure of Protein (DSSP) program (44). Structures were also modeled using SWISS-MODEL (45) to cross-check the models from MODELLER using another tool.

The PDB file for the dimer of FHR5 was generated by superimposing SCR-1/2 of each FHR5 monomer model onto the FHR1 SCR-1/2 dimer crystal structure (PDB code 3ZD2), using PyMOL (Schrödinger, LLC, New York). This structure was inputted directly into the atomistic modeling workflow of the SASSIE scattering curve fit package (19). First, the PDB file was manually corrected for gaps or errors in the amino acid sequence. A protein structure file (PSF), which contained molecule-specific information for the application of a force field, was generated via PSFGEN using Visual Molecular Dynamics (VMD) (version 1.9.2) (46). To create a physically realistic atomistic model, the structure was subjected to 10 ps of energy minimization using the molecular dynamics simulation package NAMD (version 2.9) (46, 47). The force field for this was CHARMM-36 (48, 49), and energy minimization was performed using the conjugate gradient method.

### Molecular simulations and SAXS fitting of FHR5

By excluding the dimerization interface at SCR-1/2 and linker L1, which do not vary in conformation (Fig. 1A), FHR5 contains seven potentially flexible inter-SCR linkers, L2–L8 (Fig. 1B). The linkers were subjected to peptide dihedral angle variations in the Monte Carlo simulations through the Markov sampling of backbone torsion angles (19). This allowed the rapid generation of a large conformational library of physically realistic atomistic models of the FHR5 SCR-1/9 dimer through the Complex Monte Carlo module of SASSIE. The same linkers on either monomer of the dimer were varied independently of each other; thus, the resulting dimer structures were asymmetric in shape. In Search 1, all seven linkers (L2–L8) were varied. These were defined as follows: L2, <sup>141</sup>SFTKGE<sup>146</sup>; L3, <sup>202</sup>KGQVRS<sup>207</sup>; L4, <sup>263</sup>VEQVKT<sup>268</sup>; L5, <sup>323</sup>VATHQLKR<sup>330</sup>; L6, <sup>382</sup>TEKREQF<sup>388</sup>; L7, <sup>443</sup>VESTAY<sup>448</sup>; and L8, <sup>504</sup>LDP<sup>506</sup>. In Search 2, only L2, L4, L6, and L7 were varied, because these were not part of the crystal structures that were used (Fig. 1B). In Search 3, only linkers L3 and L6 were varied as a control of Searches 1 and 2. This strategy of independent simulations (Table 3) checked whether extra or fewer constraints in the linkers affected the resulting best-fit structures. During the Monte Carlo simulations, models with steric overlaps that were generated by SASSIE were excluded by specifying an atomic overlap distance cut-off of 0.3 nm. Simulations were continued to produce models with  $R_g$  values close to that of 6.0 nm obtained experimentally by filtering for a fixed range of  $R_g$  values in the FHR5 dimer models. The outputted structures were generated as binary format DCD files and visualized on VMD. In the three searches, a total of up to 250,000 models were generated to sample a sufficient number of conformations for the two monomers in the dimer.

Using the SCT module (39) in SASSIE, a theoretical scattering curve was calculated for each of the FHR5 dimer models. The atomic coordinates were converted into small spheres to generate a coarse-grained sphere model. A cube side length of 0.53 nm in conjunction with a cut-off of four atoms was used to generate unhydrated sphere models. Because the hydration shell was visible by X-rays, a hydration shell containing 0.3 g of H<sub>2</sub>O/g of protein was added to each of the models by HYPRO (50). The theoretical scattering curve  $I(Q)$  for each model was calculated using the Debye equation adapted to spheres (39, 51).

The theoretical scattering curves for the dimer models were compared with the experimental SAXS curves. In the SCT Analyze module of SASSIE, the  $R_g$  and  $R_{xs}$  values were calculated from the modeled curves using the same  $Q$  ranges that were used for the experimental Guinier analyses. The curve fits were compared and filtered based on their  $R_g$  and  $R_{xs}$  values as well as their goodness-of-fit  $R$ -factor values defined as follows,

$$R \text{ factor} = \frac{\sum \| \| I_{\text{Expt}}(Q) - \eta \| I_{\text{Theor}}(Q) \| \|}{\sum \| I_{\text{Expt}}(Q) \|} \times 100 \quad (\text{Eq. 4})$$

where  $I_{\text{Expt}}(Q)$  and  $I_{\text{Theor}}(Q)$  were the experimental and theoretically calculated scattered intensities, and  $\eta$  was a scaling factor used to match the theoretical and experimental  $I(0)$  values. Typical best-fit  $R$ -factors for SAXS modeling are between 2 and 8% (14). To visualize the initial and best-fit models for

the FHR5 dimer, density plots were generated using the Density Plot module in SASSIE. The envelope was generated for the sterically accepted trial models, sampled to produce the volumetric data, using the Gaussian cube file format. This was superimposed onto the initial FHR5 dimer model. The output files were rendered, analyzed, and annotated in VMD. Once the best-fit dimer models were chosen, their sedimentation coefficients were calculated for comparison with the AUC data, based on the atomic coordinates using the HYDROPRO shell modeling program (52).

### Data availability

All data are contained within the article. The 55 best-fit models from Search 2 and the six best-fit structures from Fig. 11 are available in the [supporting material](#).

**Acknowledgments**—We thank Dr. P. Pernot (ESRF, Grenoble) for excellent X-ray instrumental support, Dr. David W. Wright for expert computational support, and Dr. Andrew J. Beavil (King's College London) for expert assistance with SEC-MALLS. Initial work in this project won an award for Best Poster Presentation at the XXVIth International Complement Workshop in Kanazawa, Japan, September 4–8, 2016.

**Author contributions**—N. K.-K. and J. S. B. formal analysis; N. K.-K., J. S. B., and L. C. M. validation; N. K.-K. and J. G. investigation; N. K.-K., J. S. B., J. G., and L. C. M. methodology; N. K.-K. and S. J. P. writing—original draft; N. K.-K., J. S. B., D. P. G., and S. J. P. writing—review and editing; D. P. G. and S. J. P. supervision; D. P. G. and S. J. P. funding acquisition; D. P. G. and S. J. P. project administration.

**Funding and additional information**—N. K. K. was supported by the Rosetrees Trust for Medical Research (reference M187) and a UCL Impact Studentship. D. P. G. was supported by a Medical Research Council Clinician Scientist Fellowship. J. S. B. and S. J. P. were supported by a joint Engineering and Physical Sciences Research Council (EPSRC) (EP/K039121/1) and National Science Foundation (NSF) (CHE-1265821) grant for the CCP-SAS project.

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: FHR5, Factor H-related 5; SCR, short complement regulator; AUC, analytical ultracentrifugation; MD, molecular dynamics; SAXS, small-angle X-ray scattering; SEC, size-exclusion chromatography; MALLS, multiple-angle laser light scattering; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; PDB, Protein Data Bank.

### References

- Murphy, B., Georgiou, T., Machet, D., Hill, P., and McRae, J. (2002) Factor H-related protein-5: a novel component of human glomerular immune deposits. *Am. J. Kidney Dis.* **39**, 24–27 [CrossRef Medline](#)
- McRae, J. L., Duthy, T. G., Griggs, K. M., Ormsby, R. J., Cowan, P. J., Cromer, B. A., McKinstry, W. J., Parker, M. W., Murphy, B. F., and Gordon, D. L. (2005) Human factor H-related protein 5 has cofactor activity,

inhibits C3 convertase activity, binds heparin and C-reactive protein, and associates with lipoprotein. *J. Immunol.* **174**, 6250–6256 [CrossRef Medline](#)

- Soares, D., and Barlow, P. N. (2005) Complement control protein modules in the regulators of complement activators. In *Structural Biology of the Complement System* (Morikis, D., and Lambris, J. D., eds) pp. 19–62, Taylor & Francis, Boca Raton, FL
- Goicoechea de Jorge, E., Caesar, J. J., Malik, T. H., Patel, M., Colledge, M., Johnson, S., Hakobyan, S., Morgan, B. P., Harris, C. L., Pickering, M. C., and Lea, S. M. (2013) Dimerization of complement factor H-related proteins modulates complement activation *in vivo*. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 4685–4690 [CrossRef Medline](#)
- Jokiranta, T. S., Jaakola, V. P., Lehtinen, M. J., Pärepaalo, M., Meri, S., and Goldman, A. (2006) Structure of complement factor H carboxyl-terminus reveals molecular basis of atypical haemolytic uremic syndrome. *EMBO J.* **25**, 1784–1794 [CrossRef Medline](#)
- van Beek, A. E., Pouw, R. B., Brouwer, M. C., van Mierlo, G., Rispens, T., Kuijpers, T. W., and Wouters, D. (2016) Dimerization of complement factor H-related (FHR) proteins: FHR-5 forms homodimers whereas FHR-1 and FHR-2 both homodimerize and heterodimerize with each other. *Immunobiology* **221**, 1182–1183 [CrossRef](#)
- Ruseva, M. M., Malik, T. H., and Pickering, M. C. (2016) Insights into the role of FHR5 in C3 glomerulopathy. *Immunobiology* **221**, 1167 [CrossRef](#)
- Gale, D. P., Goicoechea de Jorge, E., Cook, H. T., Martinez-Barricarte, R., Hadjisavvas, A., McLean, A. G., Pusey, C. D., Pierides, A., Kyriacou, K., Athanasiou, Y., Voskarides, K., Deltas, C., Palmer, A., Frémeaux-Bacchi, V., de Cordoba, S. R., *et al.* (2010) Identification of a mutation in complement factor H-related protein 5 in patients of Cypriot origin with glomerulonephritis. *Lancet* **376**, 794–801 [CrossRef Medline](#)
- Athanasiou, Y., Voskarides, K., Gale, D. P., Damianou, L., Patsias, C., Zavaros, M., Maxwell, P. H., Cook, H. T., Demosthenous, P., Hadjisavvas, A., Kyriacou, K., Zouvani, I., Pierides, A., and Deltas, C. (2011) Familial C3 glomerulopathy associated with CFHR5 mutations: clinical characteristics of 91 patients in 16 pedigrees. *Clin. J. Am. Soc. Nephrol.* **6**, 1436–1446 [CrossRef Medline](#)
- DiScipio, R. G. (1992) Ultrastructures and interactions of complement factors H and I. *J. Immunol.* **149**, 2592–2599 [Medline](#)
- Okemefuna, A. I., Nan, R., Gor, J., and Perkins, S. J. (2009) Electrostatic interactions contribute to the folded-back conformation of wild type human factor H. *J. Mol. Biol.* **391**, 98–118 [CrossRef Medline](#)
- Osborne, A. J., Nan, R., Miller, A., Bhatt, J. S., Gor, J., and Perkins, S. J. (2018) Two distinct conformations of factor H regulate discrete complement-binding functions in the fluid phase and at cell surfaces. *J. Biol. Chem.* **293**, 17166–17187 [CrossRef Medline](#)
- Perkins, S. J., Nan, R., Li, K., Khan, S., and Abe, Y. (2011) Analytical ultracentrifugation combined with X-ray and neutron scattering: experiment and modelling. *Methods* **54**, 181–199 [CrossRef Medline](#)
- Perkins, S. J., Wright, D. W., Zhang, H., Brookes, E. H., Chen, J., Irving, T. C., Krueger, S., Barlow, D. J., Edler, K. J., Scott, D. J., Terrill, N. J., King, S. M., Butler, P. D., and Curtis, J. E. (2016) Atomistic modelling of scattering data in the Collaborative Computational Project for Small Angle Scattering (CCP-SAS). *J. Appl. Crystallogr.* **49**, 1861–1875 [CrossRef Medline](#)
- Perkins, S. J., Nan, R. D., Li, K. Y., Khan, S., and Miller, A. (2012) Complement factor H-ligand interactions: self-association, multivalency and dissociation constants. *Immunobiology* **217**, 281–297 [CrossRef Medline](#)
- Perkins, S. J., and Goodship, T. H. (2002) Molecular modelling of the C-terminal domains of factor H of human complement: a correlation between haemolytic uraemic syndrome and a predicted heparin binding site. *J. Mol. Biol.* **316**, 217–224 [CrossRef Medline](#)
- 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., Skerka, C., Zipfel, P. F., and 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 [CrossRef Medline](#)
- Rodríguez, E., Rallapalli, P. M., Osborne, A. J., and Perkins, S. J. (2014) New functional and structural insights from updated mutational databases for

- complement factor H, factor I, membrane cofactor protein and C3. *Biosci. Rep.* **34**, 635–649
19. Curtis, J. E., Raghunandan, S., Nanda, H., and Krueger, S. (2012) SASSIE: A program to study intrinsically disordered biological molecules and macromolecular ensembles using experimental scattering restraints. *Comput. Phys. Commun.* **183**, 382–389 [CrossRef](#)
  20. Perkins, S. J. (1986) Protein volumes and hydration effects: the calculation of partial specific volumes, neutron scattering matchpoints and 280-nm coefficients for proteins and glycoproteins from amino acid sequences. *Eur. J. Biochem.* **157**, 169–180 [CrossRef](#) [Medline](#)
  21. Hui, G. K., Wright, D. W., Vennard, O. L., Rayner, L. E., Pang, M., Yeo, S. C., Gor, J., Molyneux, K., Barratt, J., and Perkins, S. J. (2015) The solution structures of native and patient monomeric human IgA1 reveal asymmetric extended structures: implications for function and IgAN disease. *Biochem. J.* **471**, 167–185 [CrossRef](#) [Medline](#)
  22. David, C. C., and Jacobs, D. J. (2014) Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol. Biol.* **1084**, 193–226 [CrossRef](#) [Medline](#)
  23. Perkins, S. J., Okemefuna, A. I., Nan, R., Li, K., and Bonner, A. (2009) Constrained solution scattering modelling of human antibodies and complement proteins reveals novel biological insights. *J. R. Soc. Interface* **6**, S679–S696 [CrossRef](#) [Medline](#)
  24. Holers, V. M. (2013) Human C3 glomerulopathy provides unique insights into complement factor H-related protein function. *J. Clin. Invest.* **123**, 2357–2360 [CrossRef](#) [Medline](#)
  25. Tortajada, A., Yébenes, H., Abarregui-Garrido, C., Anter, J., García-Fernández, J. M., Martínez-Barricarte, R., Alba-Domínguez, M., Malik, T. H., Bedoya, R., Cabrera Pérez, R., López Trascasa, M., Pickering, M. C., Harris, C. L., Sánchez-Corral, P., Llorca, O., *et al.* (2013) C3 glomerulopathy-associated CFHR1 mutation alters FHR oligomerization and complement regulation. *J. Clin. Invest.* **123**, 2434–2446 [CrossRef](#) [Medline](#)
  26. Józsi, M., and Meri, S. (2014) Factor H-related proteins. *Methods Mol. Biol.* **1100**, 225–236 [CrossRef](#) [Medline](#)
  27. Schmidt, C. Q., Herbert, A. P., Mertens, H. D., Guariento, M., Soares, D. C., Uhrin, D., Rowe, A. J., Svergun, D. I., and Barlow, P. N. (2010) The central portion of factor H (modules 10–15) is compact and contains a structurally deviant CCP molecule. *J. Mol. Biol.* **395**, 105–122 [CrossRef](#) [Medline](#)
  28. Makou, E., Mertens, H. D., Maciejewski, M., Soares, D. C., Matis, I., Schmidt, C. Q., Herbert, A. P., Svergun, D. I., and Barlow, P. N. (2012) Solution structure of CCP modules 10–12 illuminates functional architecture of the complement regulator, factor H. *J. Mol. Biol.* **424**, 295–312 [CrossRef](#) [Medline](#)
  29. Aslam, M., and 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 [CrossRef](#) [Medline](#)
  30. Fenaille, F., Le Mignon, M., Groseil, C., Ramon, C., Riandé, S., Siret, L., and Bihoreau, N. (2007) Site-specific N-glycan characterization of human complement factor H. *Glycobiology* **17**, 932–944 [CrossRef](#) [Medline](#)
  31. Chen, Q., Wiesener, M., Eberhardt, H. U., Hartmann, A., Uzonyi, B., Kirschfink, M., Amann, K., Buettner, M., Goodship, T., Hugo, C., Skerka, C., and Zipfel, P. F. (2014) Complement factor H-related hybrid protein deregulates complement in dense deposit disease. *J. Clin. Invest.* **124**, 145–155 [CrossRef](#) [Medline](#)
  32. Holmes, L. V., Strain, L., Staniforth, S. J., Moore, I., Marchbank, K., Kavanagh, D., Goodship, J. A., Cordell, H. J., and Goodship, T. H. J. (2013) Determining the population frequency of the CFHR3/CFHR1 deletion at 1q32. *PLoS ONE* **8**, e60352 [CrossRef](#) [Medline](#)
  33. Kiryluk, K., Li, Y., Scolari, F., Sanna-Cherchi, S., Choi, M., Verbitsky, M., Fasel, D., Lata, S., Prakash, S., Shapiro, S., Fischman, C., Snyder, H. J., Appel, G., Izzì, C., Viola, B. F., *et al.* (2014) Discovery of new risk loci for IgA nephropathy implicates genes involved in immunity against intestinal pathogens. *Nat. Genet.* **46**, 1187–1196 [CrossRef](#) [Medline](#)
  34. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., eds) pp. 90–125, Royal Society of Chemistry, Cambridge, UK
  35. Schuck, P. (1998) Sedimentation analysis of non-interacting and self-associating solutes using numerical solutions to the Lamm equation. *Biophys. J.* **75**, 1503–1512 [CrossRef](#) [Medline](#)
  36. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619 [CrossRef](#) [Medline](#)
  37. Pernot, P., Round, A., Barrett, R., Antolinos, A. D., Gobbo, A., Gordon, E., Huet, J., Kieffer, J., Lentini, M., Mattenet, M., Morawe, C., Mueller-Dieckmann, C., Ohlsson, S., Schmid, W., Surr, J., *et al.* (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664 [CrossRef](#) [Medline](#)
  38. Glatter, O., and Kratky, O. (1982) *Small-angle X-ray scattering*, Academic Press, Inc., New York
  39. Wright, D. W., and Perkins, S. J. (2015) SCT: a suite of programs for comparing atomistic models with small-angle scattering data. *J. Appl. Crystallogr.* **48**, 953–961 [CrossRef](#) [Medline](#)
  40. Semenyuk, A. V., and Svergun, D. I. (1991) GNOM—a program package for small angle scattering data-processing. *J. Appl. Crystallogr.* **24**, 537–540 [CrossRef](#)
  41. Šali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 [CrossRef](#) [Medline](#)
  42. McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P., and Lopez, R. (2013) Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res.* **41**, W597–W600 [CrossRef](#) [Medline](#)
  43. Larsson, P., Wallner, B., Lindahl, E., and Elofsson, A. (2008) Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Sci.* **17**, 990–1002 [CrossRef](#) [Medline](#)
  44. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 [CrossRef](#) [Medline](#)
  45. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 [CrossRef](#) [Medline](#)
  46. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 [CrossRef](#) [Medline](#)
  47. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 [CrossRef](#) [Medline](#)
  48. MacKerell, A. D., Jr., Brooks, B., Brooks, C. L., Nilsson, L., Roux, B., Won, Y., and Karplus, M. (1998) CHARMM: the energy function and its parameterization. *Encyclopedia of Computational Chemistry*, John Wiley & Sons, Inc., New York
  49. Best, R. B., Zhu, X., Shim, J., Lopes, P. E., Mittal, J., Feig, M., and MacKerell, A. D., Jr. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone  $\phi$ ,  $\psi$  and side-chain  $\chi_1$  and  $\chi_2$  dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 [CrossRef](#) [Medline](#)
  50. Ashton, A. W., Boehm, M. K., Gallimore, J. R., Pepys, M. B., and Perkins, S. J. (1997) Pentameric and decameric structures in solution of the serum amyloid P component by X-ray and neutron scattering and molecular modelling analyses. *J. Mol. Biol.* **272**, 408–422 [CrossRef](#) [Medline](#)
  51. Perkins, S. J., and Weiss, H. (1983) Low resolution structural studies of mitochondrial ubiquinol-cytochrome c reductase in detergent solutions by neutron scattering. *J. Mol. Biol.* **168**, 847–866 [CrossRef](#) [Medline](#)
  52. Garcia de la Torre, J., Huertas, M. L., and Carrasco, B. (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* **78**, 719–730 [CrossRef](#) [Medline](#)