doi:10.1016/j.jmb.2006.07.013 J. Mol. Biol. (2006) 362, 700–716

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Binding Site Structure of One LRP–RAP Complex: Implications for a Common Ligand–Receptor Binding Motif

Gitte A. Jensen[1], Olav M. Andersen[2], Alexandre M. J. J. Bonvin[3] Ida Bjerrum-Bohr[1], Michael Etzerodt[2], Hans C. Thøgersen[2] Charlotte O'Shea[1], Flemming M. Poulsen[1] and Birthe B. Kragelund[1]*

[1]*SBiN Lab, Institute of Molecular Biology and Physiology, University of Copenhagen, Øster Farimagsgade 2A, DK-1353 Copenhagen K, Denmark*

[2]*Laboratory of Gene Expression Department of Molecular Biology, University of Aarhus Gustav Wieds Vej 10 DK-8000 Aarhus C, Denmark*

[3]*NMR Research Group, Bijvoet Center for Biomolecular Research, Utrecht University 3584CH Utrecht The Netherlands*

**Corresponding author*

The low-density lipoprotein receptor-related protein (LRP) interacts with more than 30 ligands of different sizes and structures that can all be replaced by the receptor-associated protein (RAP). The double module of complement type repeats, CR56, of LRP binds many ligands including all three domains of RAP and α_2-macroglobulin, which promotes the catabolism of the Aβ-peptide implicated in Alzheimer's disease. To understand the receptor–ligand cross-talk, the NMR structure of CR56 has been solved and ligand binding experiments with RAP domain 1 (RAPd1) have been performed. From chemical shift perturbations of both binding partners upon complex formation, a HADDOCK model of the complex between CR56 and RAPd1 has been obtained. The binding residues are similar to a common binding motif suggested from α_2-macroglobulin binding studies and provide evidence for an understanding of their mutual cross-competition pattern. The present structural results convey a simultaneous description of both binding partners of an LRP–ligand complex and open a route to a broader understanding of the binding specificity of the LRP receptor, which may involve a general four-residue receptor–ligand recognition motif common to all LRP ligands. The present result may be beneficial in the design of antagonists of ligand binding to the LDL receptor family, and especially of drugs for treatment of Alzheimer's disease.

Keywords: NMR; HADDOCK; lipoprotein; domain; SPR

Introduction

The low-density lipoprotein receptor-related protein (LRP) is a large (600 kDa) multifunctional transmembrane cell-surface protein. The receptor recognizes several structurally different extra-cellular ligands such as lipoproteins, apolipoproteins, lipases, protease inhibitors, and protease/inhibitor complexes [1,2] and internalises the bound ligands for degradation by receptor-mediated endocytosis. LRP belongs to the low-density lipoprotein (LDL) receptor family and the ectodomain structure of the LDL receptor at low pH was solved recently by X-ray crystallography.[3] LRP is distinctively larger than the LDL receptor and contains on the extra-cellular side a long modular array of acidic cysteine-rich complement-type repeats, epidermal growth factor (EGF) repeats, and β-propeller modules. A complement-type repeat (CR) consists of approximately 40 amino

Present address: O. M. Andersen, Molecular Cardiovascular Research, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, D-13125 Berlin, Germany.

Abbreviations used: LRP, low-density lipoprotein receptor-related protein; LDL, low-density lipoprotein; RAP, receptor-associated protein; RAPd1, RAP domain 1; CR, complement-type repeat; EGF, epidermal growth factor; ER, endoplasmatic reticulum; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; HSQC, heteronuclear single quantum coherence; AIR, ambiguous interaction restraint; DQF-COSY, double-quantum filtered correlated spectroscopy; TOCSY, total correlated spectroscopy; SPR, surface plasmon resonance.

E-mail address of the corresponding author: bbk@apk.molbio.ku.dk

acid residues with three disulphide linkages and coordinates one calcium ion. In LRP, CRs are distributed in four clusters, I to IV, containing 2, 8, 10 and 11 CRs, respectively, where β-propeller modules and EGF modules flank each cluster. The CRs in the second and forth cluster of LRP are the most active in ligand binding,[4] whereas the EGF repeats and the β-propeller modules are suggested to be important for the dissociation of ligands in the endosomes[5] besides a discrete function in ligand binding. Receptor-associated protein (RAP) is an endoplasmatic reticulum (ER) resident protein that acts as a chaperone protein for LRP and prevents all types of ligands from associating with newly synthesized LRP in the rough ER.[6–8] RAP consists of 323 amino acid residues distributed into three sequence-based domains: RAPd1 (RAP:18–112), RAPd2 (RAP:113–218), and RAPd3 (RAP:219–323), as shown by sequence alignments and ^{1}H-NMR analysis.[9] The solution structure of a C-terminal truncated form of RAPd1, RAPd1T (RAP:17–97) has been solved by NMR spectroscopy and contains three α-helices composed of residues 23–34 (H1), 39–65 (H2) and 73–88 (H3).[10] The 15 C-terminal residues (amino acid residues 98–112) are not organized in any regular secondary structures or part of tertiary interactions in the non-bound conformation of domain 1, but it has been proposed that they form a more regular structure in the complex with a receptor molecule or in the context of the entire RAP molecule.[10]

Several structures of the isolated CRs of the LDL and LRP receptors have been solved both individually by NMR[11–15] or by X-ray crystallography[16,17] and as double modules.[18,19] All CR modules have the same basic three-dimensional fold with a short antiparallel β-sheet, and for some CRs one single turn of an α-helix, as the only secondary structural elements. Furthermore, the common CR fold consists of two loops connected by a disulphide bridge between the second and fifth cysteine at one end. An antiparallel β-sheet arranged as a β-hairpin stabilizes, together with a disulphide bond between the first and third conserved cysteine residues, the N-terminal loop, whereas the C-terminal loop is stabilized through its organization around a bound calcium ion and a conserved disulphide bond between the forth and sixth conserved cysteine residues. The structures solved so far have an overall backbone RMSD to the average structure of 3.0(±0.9) Å and structural studies of double modules point to distinct modular structural independency.[19,20]

LRP acts as a scavenger in the clearance of numerous different ligands from the blood stream. These ligands deviate both in structure and function and, characteristically, do not cross-compete for binding. All ligands can, however, be replaced by RAP.[2,20–22] Multiple mutagenesis and binding studies have aimed at determining the residues in LRP responsible for binding to RAP,[20,23] α_2-macroglobulin,[24,25] uPA/PAI1,[26,27] apolipoprotein E,[28] and lipoprotein lipase.[29] Generally, the binding of all these ligands map to the region of residues 776–1399 spanning the CRs in the second cluster of LRP.[30] LRP–ligand interactions are typically highly electrostatic in nature and involve contacts between basic residues of the ligand and acidic residues of LRP. Several independent binding sites for RAP have been suggested, and at least three sites have been determined.[31–33] Surface plasmon resonance experiments show that complement-type repeats 5 and 6 (CR56, LRP residue Ser932 to His1013) of cluster II is a minimal RAP-binding receptor unit of LRP, and that no single repeat of cluster II shows high-affinity binding to RAP.[31,34] The affinities of the three RAP domains for CR56 range from mM (d2) to nM (d3).[34] Apart from binding to all three RAP domains, the same minimal binding unit, CR56, of LRP binds preferentially over other consecutive repeats to the uPA–PAI1 complex,[27] showing that the CR56 double module contains binding determinants for many of the LRP ligands.

How does LRP, or more specifically, how do the CR repeats that are so similar in sequence recognize so many distinctly different ligands with high affinity? Structures of the individual repeats have revealed that given their small sizes and the open loop structure, the sequence variability results in completely different contour surfaces and charge distributions that might explain the diverse properties in ligand binding activity. However, structures of the LDL receptor at low pH (pH 5.3) and of the VLDLR bound to a minor group human rhinovirus,[35] collectively point to an additional fingerprint motif in recognition and binding that is not linked to secondary structure disposition. Until now no structures of complexes between the LRP receptor and its natural ligands have been solved. This study is aimed at unravelling the binding specificity of LRP using RAP as an all-ligand binding analogue.

We describe here the NMR structure of the CR56 double module from LRP together with a mapping of the interactions between CR56 and RAPd1 as determined by NMR. Based on ^{1}H and ^{15}N chemical shift perturbations on both binding partners upon complex formation, a structural model of the complex between CR56 and RAPd1 has been calculated using the program HADDOCK.[36] From this model the interactions between RAP and LRP are described and linked to suggested binding motifs and recently published structures. We suggest that many of the natural LRP ligands may present a common minimal binding motif within a recognition helix and that the specificity of binding should not only be sought in the charge distribution of this α-helix, but certainly also in a four-residue fingerprint-binding motif.

Results

The CR56 structure

The structure determination by NMR spectroscopy of the double module CR56 was severely

hampered by the 57% sequence identity between the two modules, CR5 (LRP residue Ser932–Cys971) and CR6 (LRP residue Ala972–His1013) (Figure 1(a)). This gave rise to many spectral overlaps and imposed severe complications for the assignments. A total of 675 non-redundant distance restraints

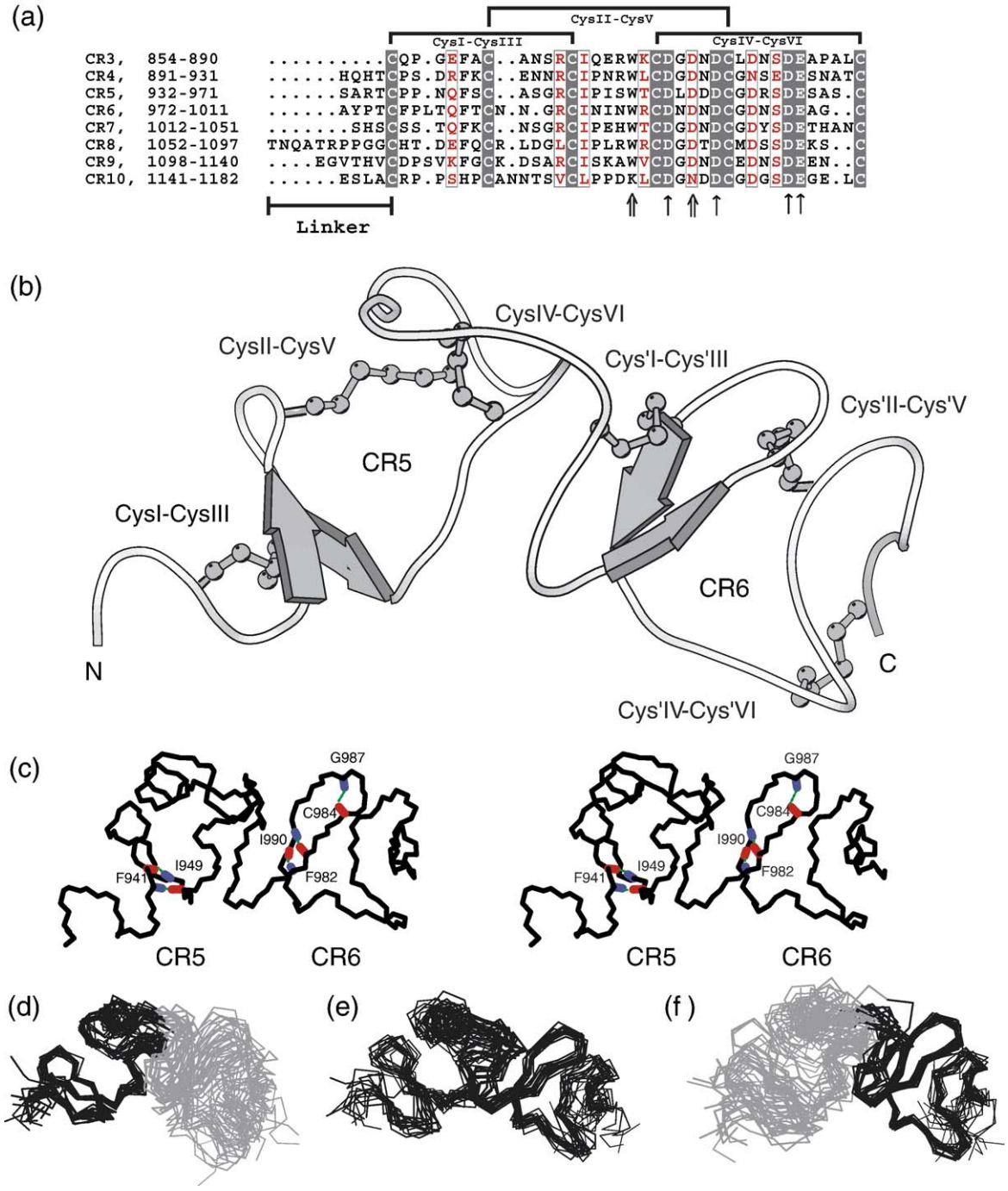

Figure 1. NMR solution structure of CR56. (a) Sequence alignments of complement-type repeats CR3-CR10 from cluster II of LRP with residues coordinating calcium *via* their side-chain indicated with a ↑ and *via* their backbone carbonyl with a ⇑. The linker regions are indicated with a horizontal bar below the alignment, and the three conserved disulphide bridges are indicated above the alignment. (b) Lowest energy structure presented as a smooth coil through the C^{α} atoms, and showing the disulphide bridges and secondary structure elements of β-hairpins (Phe941 - Ile949 (CR5) and Phe982 – Ile990 (CR6)). The six disulphide bridges are indicated as Cys–Cys (CR5) and Cys′–Cys′ (CR6) pairs with the cysteine residues in each repeat numbered from I to VI corresponding to the six conserved cysteine residues in complement-type repeats. (c) Stereo view of the backbone trace of the lowest energy structure, showing the five backbone hydrogen bonds involving the three residue pairs Phe941/Ile949, Phe982/Ile990, and Gly987/Cys984. (d)–(f) Cα traces for the representative ensemble of 15 structures fitted with respect to the backbone atoms for residues (d) Cys936-Cys971 (CR5), (e) Cys936-Cys1011 (CR56), and (f) Cys976-Cys1011 (CR6).

were obtained from ^{1}H-^{1}H nuclear Overhauser effects (NOEs), and 21 of these were inter-modular effects between CR5 and CR6. From NOEs and from a set of 29 phi-angle restraints and restraints for calcium coordination, an ensemble of 15 NMR structures has been calculated to represent the structure of the CR56 double module in solution. A ribbon representation of the lowest energy structure and the ensemble of 15 structures with no violations greater than 0.5 Å (see Materials and Methods), are shown in Figure 1(b)–(f). A summary of structural restraints and statistics is presented in Table 1. The 15 structures superimpose well on the CR5 and CR6 backbone separately, with a backbone RMSD from the average structure of 1.17(±0.30) Å from Cys936 to Cys971 (CR5) and 1.00(±0.30) Å from Cys976 to Cys1011 (CR6) (Figure 1(d)–(f) and Table 1). The backbone RMSD to the average structure for the entire CR56 backbone is, however, 1.58(±0.28) Å from Cys936 to Cys1011, indicating that the linker region between CR5 and CR6 is flexible to some given extent. However, observed inter-modular NOE effects between Ser952–Thr954 (CR5) and Leu979 (CR6) and between Leu957 (CR5) and Arg988-Cys989 (CR6) do indicate a preferred orientation of the two modules, respectively. Results from residual dipolar couplings and relaxation data (see Supplementary Data Figures), suggest that CR5 behaves differently from CR6 and from the linker region dynamically. Measured backbone ^{15}N-R1 and ^{15}N-R2 values for CR56 are, on average, the largest for the CR5 module, with maximum values for ^{15}N-R1 of 2.82(±0.17) s^{-1} (CR5, Thr954) and 2.87(±0.34) s^{-1} (CR6, Ala972), and with maximum values for ^{15}N-R2 of 24.5(±2.34) s^{-1} (CR5, Leu957) and 16.57(±0.96) s^{-1} (CR6, Gly1003). The H^N-N residual dipolar couplings data set for CR56 likewise show the existence of different ranges of observed values for the two modules, resulting either from a less flexible CR5 compared to the linker region and CR6 or from stronger alignment with the anisotropic media of CR5 due to interactions with the bicelles. These data support the existence of a non-random orientation of the two modules with respect to each other, although the low number of inter-modular NOEs is not enough for establishing a precise modular arrangement.

The structures of the individual modules of CR5 and CR6 show the typical fold of a CR module with an incoming short β-hairpin of ten residues (Phe941–Ile949 of CR5 and Phe982–Ile990 of CR6) followed by the calcium-binding loop. Only few residues adopt conformations in regular secondary structures and this fact is manifested in the distribution of backbone torsion angles for the lowest energy structure, with 51.5% of the residues in most favoured regions, 33.3% in allowed regions, 9.1% in additionally allowed regions, and 6.1% in the disallowed regions (Ser945, Ser952, Asp958, and Cys976). This is also seen for other CR modules like LB2[12] and LB6[13] and reflects most likely the low content of regular secondary structure elements of these modules and possibly also dynamics. Calcium coordination is accomplished by backbone atoms of Trp953, Asp958 (CR5) and Trp994, Asp999 (CR6) and by side-chain atoms of Asp956, Asp960, Asp966, Glu967 (CR5) and Asp997, Asp1001, Asp1007, Glu1008 (CR6). Five of the six conserved acidic residues in both CR5 and CR6 (Figure 1(a)), are taking part in calcium coordination, and, as these are conserved in all modules, other or additional residues must be responsible for the specific binding and discrimination of ligands. Interestingly, CR5 has one extra acidic residue compared to other CR modules, Asp959, which does not coordinate calcium, whereas CR6 does not have this extra acidic residue. This residue has been shown by mutagenesis to be important for

Table 1. Summary of experimental restraints and of structural statistics for the NMR solution structure of CR56

A. *Experimental NMR restraints*	
Total NOE distance restraints	675
CR5 intra module[a]	307
CR6 intra module[b]	347
CR56 inter module[c]	21
CR56 intra residue ($\|i-j\|=0$)	75
CR56 inter residue short range ($\|i-j\|<=4$)	383
CR56 inter residue long range ($\|i-j\|>4$)	217
Total dihedral φ angle restraints	29
B. *Deviations from experimentally derived restraints*	
NOE violations > 0.5 Å	0.0±0.0
NOE violations > 0.3 Å	2.4±1.7
Dihedral angle violations > 5°	1.4±1.1
C. *Deviation from ideal geometry (RMSD)*	
Impropers (°)	2.4±0.6
Bonds (Å)	0.0±0.0
Angles (°)	0.7±0.2
D. *Energies (kcal·mol^{-1})*	
Total	−2263±590
Bond	33±9
Angle	169±47
Impropers	130±38
van der Waals[d]	436±113
Electrostatics[d]	−3027±785
NOE restraint	89±26
Dihedral angle restraint	101±34
E. *Independent RDC R-factors*[e]	
R-factor (CR5/CR6/CR56)	0.48/0.62/0.62
RMS	11.5/6.7/12.6
F. *RMSD of atomic positions (Å)*[f]	
Backbone atoms (Cα, N, C)	1.17/0.99/1.58
All heavy atoms	1.66/1.45/1.93
G. *PROCHECK report*	
Most favourable (%)	40.6/51.2/44.9
Allowed (%)	44.9/39.2/42.8
Additionally allowed (%)	10.1/8.5/9.7
Disallowed (%)	4.3/1.0/2.5

[a] LRP residues Ser932-Tyr973.
[b] LRP residues Pro974-His1013.
[c] LRP residues Ser932-Tyr973, Pro974-His1013.
[d] The non-bonded energies were calculated with the OPLS parameters using a 8.5 Å cut-off.[63]
[e] Calculated using PALES.[64]
[f] The three number are for residue Cys936–Cys971/Cys976–Cys1011/Cys936–Cys1011, respectively.

RAP binding[23] and could be an active player in selective ligand binding.

The 15 structures have been analysed thoroughly for hydrogen bonds and salt bridges. Backbone hydrogen bonds have been found between the following five HN-O pairs: Phe941/Ile949, Ile949/Phe941, Phe982/Ile990, Gly987/Cys984, Ile990/Phe982, and they define the two β-hairpins as shown in Figure 1(c). All acidic and basic residues of CR56, except three calcium-coordinating aspartic acids (Asp956, Asp966 and Asp1007) have their charged side-chains relatively highly exposed to the solvent, with relative side-chain residue surface accessible areas higher than 20% as calculated by NACCESS†.

Interaction with RAP

The NMR spectra of the individual RAP domains are of different qualities, suggesting different populations of structures with structural averaging in the population of RAPd3, and to a smaller extent also in RAPd2 but not observed for RAPd1.[24] For studies of the receptor-ligand interplay, RAPd1 is an appropriate ligand for several reasons. It binds to CR56 with a dissociation constant of 2.8 μM,[34] it inhibits LRP-binding of α_2-macroglobulin that promotes the catabolism of the Aβ-peptide implicated in Alzheimer's disease, it has a well-defined three-dimensional structure, and the NMR assignments are available.[37] The complex between CR56 and RAPd1 does not preserve long-term stability and together with an extremely low refolding outcome of the CR56 double module in labelled media, a full structure determination of the complex was impracticable. Instead, a HADDOCK model based on chemical shift perturbations for the backbone ^{1}H and ^{15}N nuclei for both CR56 and RAPd1 upon complex formation could be achieved.

Chemical shift perturbations of both proteins are shown in Figure 2(a) and (b). The ^{15}N-^{1}H heteronuclear single quantum coherence (HSQC) spectra of the titration clearly show that free RAPd1 and free CR56 are in slow exchange on the NMR time scale with a one-to-one stoichiometric CR56-RAPd1 complex (Figure 2(c)). Quantification of the peak corresponding to Glu40 of CR56 enables the determination of a binding constant, $K_D \sim 2$ μM, that corresponds nicely to a previously determined affinity.[34] In the exchange spectrum recorded on a half-saturated sample (see Materials and Methods), each amide will give rise to four peaks, one from the free form, one from the bound form and two peaks from magnetization exchange between the two forms. Due to spectral overlap in these spectra the assignment of all residues in their bound form was difficult and thus not complete, especially for CR56, as these spectra were recorded on dilute samples. Therefore, these data were combined with assignments obtained from a 0.08 mM ^{13}C,^{15}N-labelled sample of the complex, and in that way we obtained 56 out of 78 signals for [^{15}N]RAPd1 and 68 out of 77 signals for [^{15}N]CR56 were assigned. As can be seen in Figure 2(a) and (b) the largest perturbations within CR56 are seen for residues Ala944, Cys948, Trp953, Thr954, Asp956, Cys961, Gly962, Asp966, Ser968, Ser970 of the CR5 module, for residues Ala972 and Thr975 of the linker region between CR5 and CR6, and for residues Cys976, Thr980, Gly987, Asp997, Gly1003 of CR6. All the perturbed residues in CR56 are exposed to the surface, except Cys961, Gly962, and Cys976 (Figure 2(d)). In RAPd1, large perturbations are seen for residues Phe20–Leu24, Leu28, Gln33 in helix I (RAP(23–34)), residues Leu35, His36 and Leu37 between helix I and helix II (RAP(39–65)), the C-terminal residues Ile51, Lys61-Asp65 of helix II, residues Glu74 and Leu81 of helix III (RAP(73–88)) and the C-terminal residues Asp91, Asp92, flanking helix III. All the perturbed RAPd1 residues are exposed to the surface, except Leu37 (Figure 2(e)). Importantly, no significant shift changes were observed for the residues in the C terminus of RAPd1(17–112), which agrees well with the use of the structure of the truncated RAPd1 (17–97) in the modelling procedures. The average values for the combined chemical shift perturbations for the backbone ^{1}H and ^{15}N chemical shifts ((|ΔδH|+¼|ΔδN)) were found to be 0.24(±0.25) ppm for RAPd1 and 0.15(±0.25) ppm for CR56. The highest contribution to the combined total CR56 shifts is found for CR5 (0.21(±0.34) ppm for CR5 (Ser932–Cys971)) compared to CR6 (0.11(±0.074) ppm for CR6 (Cys972–His1013)).

It is seen from the chemical shift perturbations that the interaction of CR56 with RAPd1 predominantly affects residues in CR5, mainly in the calcium-binding loop, but also residues in the linker region and in the beginning of the N-terminal β-hairpin loop of CR6 are influenced. Earlier work by mutagenesis has shown Trp953 and Asp958 of CR5 and Trp994 and Asp999 of CR6 to be almost of equal importance for RAP binding.[27,34] Chemical shift perturbations are indeed observed for the two residues of CR5, and for Asp999 of CR6; however, neither the backbone amides nor the side-chain Nε, Hε^{1} of Trp994 of CR6 were markedly influenced by RAPd1 binding. The shift changes observed for these fingerprint residues upon binding are certainly directly pointing to the participation of these in binding.

Model of the CR56–RAPd1 complex

Modelling of protein–protein complexes from NMR chemical shift data, mutagenesis data, or both has recently been highly successful in generating structures within 2 Å of already known X-ray structures of protein complexes using the robust modelling program HADDOCK.[36,38,39] Based on the observed chemical shift perturbations of CR56 and of RAPd1 upon complex formation, 21

† Hubbard, S. J. and Thornton, J. M. (1993). NACCESS. Department of Biochemistry and Molecular Biology, University College London.

ambiguous interaction restraints (AIRs) were defined for residues at the interface (see Materials and Methods). The active and passive residues of CR56 and RAPd1 are listed in Table 2 together with the defined semi and fully flexible segments used in docking.

Analysis of the final 200 water-refined HADDOCK program models for the CR56–RAPd1 complex resulted in 14 different clusters. Statistics of the top 5 clusters are presented in Table 3. Cluster 2 has the best average HADDOCK score (see Material and Methods). Based on this we present here cluster 2 as the HADDOCK model that represents best the structure of the CR56–RAPd1 complex (Figure 4(a)). The five best structures in cluster 2 have an average backbone RMSD value from the average of 2.40(±0.80) Å for the interface, 2.55(±1.40) Å for the entire backbone (Figure 3(a)); 2.75(±0.80) Å for the CR56 backbone (Figure 3(b)), and 1.85(±0.50) Å for the RAPd1 backbone (Figure 3(c)).

The structures of RAPd1 and CR56 in their bound form are different to the structures in their free form, with backbone RMSD values to the mean structure of 0.79(±0.51) Å for residues Glu19–Asp91 of RAPd1 and 2.32(±0.95) Å for residues Cys936–Cys1011 for CR56. These RMSD values are based on the 15 lowest total energy structures from each set of free structures and the five lowest inter-molecular energy structures from cluster 2 of the HADDOCK model of the CR56-RAPd1 structure. The smallest changes occur in helix 1 of RAPd1 (0.19(±0.08) Å *versus* 0.27(±0.12) Å for helix 2 and 0.26(±0.15) Å for helix 3) whereas the changes in CR56 are largest for CR6 (1.59(±0.78) Å for CR6 *versus* 1.16(±0.31) Å for CR5). Interestingly, both of the individual modules of CR56 are well defined in the complex, whereas the entire double module is not (Figure 3(a) and (d)). This is due to the fact that the linker region moves, whereby CR6 reorients with respect to CR5 (Figure 5(a)). This process is distinctly allowed for in the HADDOCK runs.

Discussion

The structure of a minimal ligand binding module of LRP consisting of the double module CR56 has been solved by NMR spectroscopy and its interaction with the first domain of RAP analysed by chemical shift mapping. From these data a model representing a structure of a complex between LRP and RAP has been obtained using the program HADDOCK. This first structural model between RAP and LRP is described and further analysed in the context of the recently published LDL and VLDL receptor structures.

The LRP–RAP interface

The large interface between LRP and RAP (~2400 Å^2) as shown in Figure 4(a) contains a set of inter-molecular salt bridges and hydrogen bonds that provides the specificity in the binding recognition pattern of RAPd1 together with a large number of hydrophobic contacts. Typically, both modules of CR56 each interact with RAPd1 through a hydrophobic patch and through a network of buried salt bridges, although the interacting network of CR5 and RAPd1 is more distinct amongst the set of structures of cluster 2. van der Waals' contacts are seen between RAPd1 residues Leu28 and Leu64, together with the hydrophobic part of three lysine side-chains, Lys31, Lys60, and Lys63 and one glutamine, Gln27 (stacking above Lys63), with residues Ile949, and Trp953 of CR5 and Ile990, Trp994, Ala1009 and the hydrophobic part of the side-chain of Arg1004, of CR6 (Figure 4(b)). All basic/acidic residues with less than 5% relative solvent accessibility, as calculated by the program NACCESS‡, are forming interactions with oppositely charged acidic/basic partner residue nearby. Three networks of buried salt bridges are observed to be present in more than 40% of the five structures of cluster 2: one major network of four cross-interacting residues is located within the hydrophobic interface and two minor ones of a total of five residues, are located at the rim of the hydrophobic interface, tethering the C terminus of RAPd1 to the complex, (Figure 4(c)-i,ii). Contacts are seen between the following CR56/RAPd1 residue pairs: network I: Asp959/Lys24, Asp958/Lys24, Asp959/Lys60 and Asp958/Lys60 (Figure 4(c)-i**)**; network II: Asp997/Asp999/Lys93 (Figure 4(c)-ii); network III: Asp1004/Arg34 (Figure 4(c)-iii). The electrostatic surfaces of the interface are distinctively charged as can be seen in Figure 5(b). The interactions between LRP and RAP in the form of the complex between CR56 and RAPd1, involve as suggested from earlier studies distinct acidic residues together with the conserved tryptophan side-chain of each CR repeats. Clearly, this fits nicely with the lysine side-chains from RAPd1 that serve dual roles, one in specific electrostatic interactions, the other undoubtedly in providing hydrophobic platforms for interacting non-polar side-chains of the receptor.

High affinity binding of RAP to LRP has earlier been attributed to clusters of arginine residues in the C termini of RAPd2 Arg-Leu-Arg-Arg(203–206) and RAPd3 Arg-Val-Ser-Arg-Ser-Arg-Glu-Lys(282–289),[40] but also two lysine residues at position 256 and 270 (Lys256 and Lys270) have been shown to be critical for LRP binding.[41] From studies on α_2-macroglobulin, and from sequential alignments of RAP and apolipoproteins, a common binding motif for the LRP–ligands was proposed involving two lysine (basic) residues separated by two to five residues and both basic residues N-terminally flanked by hydrophobic residues, preferentially leucine or valine.[24] RAPd1 contains several basic motifs analogous to the proposed α_2-macroglobulin binding motif, positioned in RAP residues 45–50 (LHADLK,

‡ Hubbard, S. J. and Thornton, J. M. (1993). NACCESS. Department of Biochemistry and Molecular Biology, University College London.

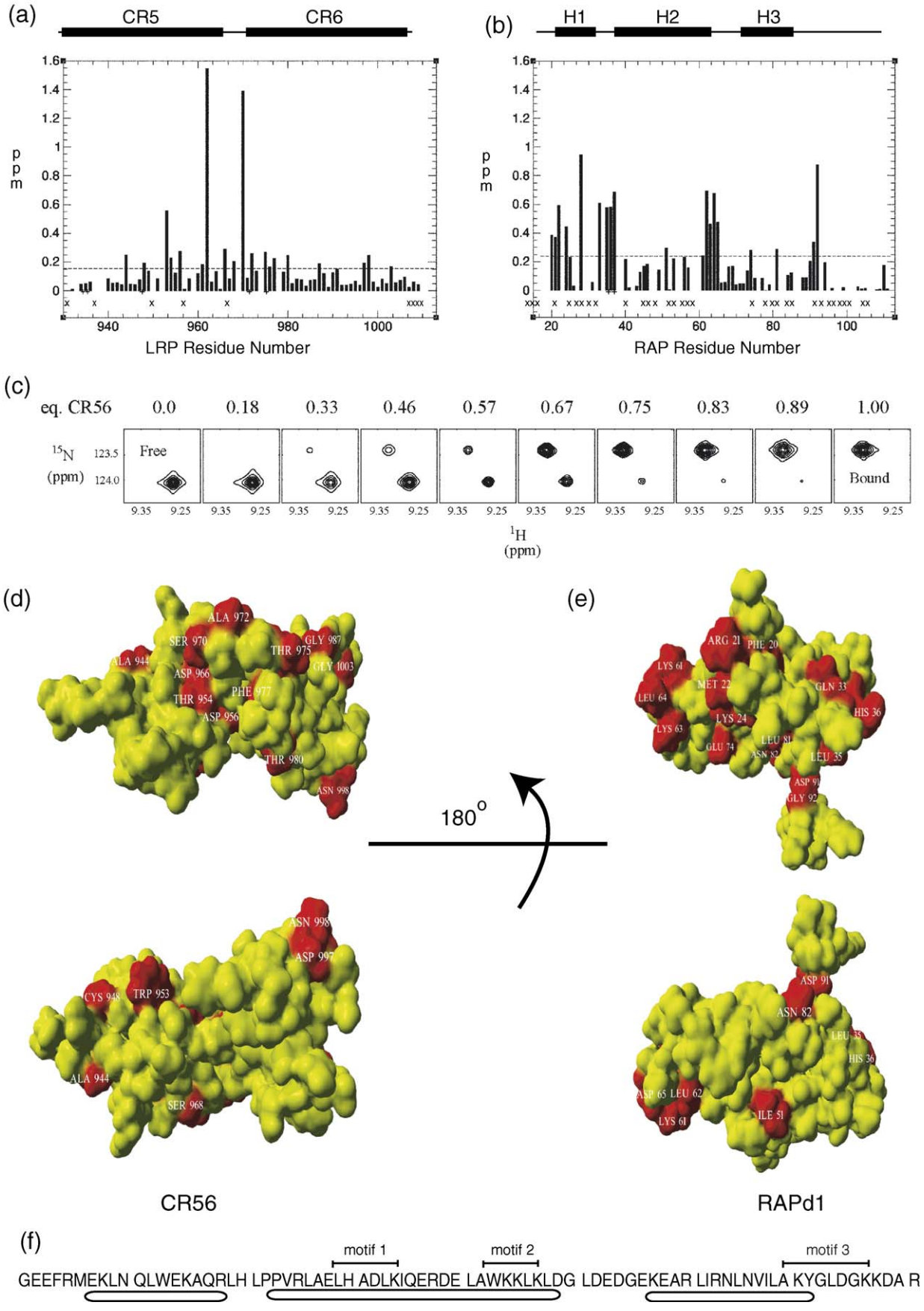

Figure 2 (*legend on opposite page*)

Table 2. Active and passive residues used in the modeling of the CR56–RAPd1 complex with HADDOCK

Active residues	CR56	Ala944, Cys948, Trp953, Cys961, Ser968, Thr980, Asp997, Asn998 and Gly1003
	RAPd1	Arg21, Lys24, Ala33, Arg35, Leu36, Lys63, Leu64, Asp65, Glu74, Asp91 and Gly92
Passive residues	CR56	All other solvent accessible residues (>55%)
	RAPd1	All other solvent accessible residues (>55%)
Semi flexible segments	CR56	Cys936–Ser968, Phe977–Ala1009
	RAPd1	Glu23–Ala43, Ala58–Gly92
Fully flexible segments	CR56	Ser932–Thr935, Ala969–Cys976, Gly1010–His1013
	RAPd1	Gly17–Met22, Gly92–Gly97

motif 1), 59–63 (WKKLK, motif 2) and a less similar motif in position 86–93 (AKYGLDGK, motif 3) (Figure 2(f)).

One lysine of RAPd1, Lys60, participates bi-functionally in the binding to the LRP receptor in that it adds both to affinity and specificity. This lysine is positioned in the C terminus of helix 2**.** The sequence around this lysine has the features of the common binding motif suggested earlier from studies of α_2-macroglobulin,[24] and includes the motif 2 residues ΨKXΨK (WKKLK in RAPd1; Ψ is hydrophobic, X is any residue). Both of the basic residues in the motif are flanked on their N-terminal side by a large hydrophobic residue, and are separated by less than five residues. Earlier suggestions to the position of a common LRP binding motif in RAPd1 was 15 residues upstream starting at position Leu45[27]; however, none of these suggested residues show any particular chemical shift changes upon complex formation, and they are not involved in binding. The role of the basic residues suggested from work on α_2-macroglobulin was to provide specificity. Both lysine residues in the structure of the receptor-binding domain (RBD) of α_2-macroglobulin are pointing to the solvent with no electrostatic interaction partners. In the free form of RAPd1, both Lys60 and Lys63 are also solvent exposed, with no interaction partners found. In the complex of RAPd1 and CR56, however, the two lysine residues as well as Trp59, and Leu64 are taking part in the hydrophobic interface. Additionally, Lys60 is involved in inter-molecular electrostatic interactions of network I described above. Lys63 is, in contrast to earlier suggestions, heavily involved in intra-molecular electrostatic interactions with side-chains of both Asp68 and Glu74 of RAPd1 and thus does not in this complex contribute to the interactions with CR56.

To evaluate the importance of Lys60 of RAPd1 in the binding interface and to provide independent experimental evidence for the obtained HADDOCK model, Lys60 was mutated to alanine, K60A. Additionally, one lysine, Lys50, which is positioned three helical turns upstream from Lys60, and which in the HADDOCK model of the complex is peripheral to the binding site was likewise mutated to alanine, K50A. For both mutant proteins and for wild-type RAPd1 interactions with CR56 was analysed by surface plasmon resonance (Figure 6). As expected both wild-type RAPd1 and RAPd1-K50A bound to CR56 with a K_D value of approximately 1 μM whereas K60A did not show any detectable binding, even at high molar excess to U-CR56 (5 μM). Structural integrity of all three proteins was confirmed from far-UV CD spectra, which were indistinguishable (data not shown). Thus, Lys60 is indeed essential for RAPd1 binding to CR56, whereas Lys50 is not, strongly supporting the validity of the model presented.

Interestingly, the position of the two important lysine residues in the structure of the receptor-binding domain of α_2-macroglobulin is in the single α-helix,[42,43] which structurally corresponds nicely to the α-helical position of the binding motif of RAPd1. Accordingly, the two motifs align very well with an all atoms RMSD value of 1.56 Å of the two lysine pairs. This similar mode of binding of both RAPd1 and α_2-macroglobulin suggests that the natural ligands of LRP may posses a potential receptor recognition helix containing this motif which, we believe, is the most prominent LRP binding site.

A general recognition motif

To obtain further insight into the determinants of binding we compared the available complexes

Figure 2. Chemical shift perturbations of CR56 and RAPd1 upon complex formation. (a) Combined chemical shift changes ($|\Delta\delta H|+1/4\,|\Delta\delta N|$) (bars) *versus* residue number for [^{15}N]CR56 upon binding of RAPd1 and (b) for [^{15}N]RAPd1 upon binding of CR56. Proline residues (+) and non-assigned residues in the bound forms (x) are indicated. Also indicated (dotted lines) are the chemical shift cut-off values used for the definition of AIRs as input for the HADDOCK calculations (0.15 ppm in (a) and 0.24 ppm in (b)). The CR5 and CR6 domain boundaries and the three helices H1–H3 of RAPd1 are indicated with bars above the respective plots. (c) Sections of ^{15}N-^{1}H HSQC spectra showing the peaks from RAPd1 residue Glu40 during the titration of ^{15}N-labelled RAPd1 with increasing amounts of unlabelled CR56. (d)–(e) Surface plots showing residues with significantly perturbed chemical shifts upon complex formation in (d) CR56 (Ala944, Cys948, Trp953, Thr954, Asp956, Asp966, Ser968, Ser970, Ala977, Thr975, Phe977, Thr980, Gly987, Asp997, Asn998 and Gly1003) and (e) in RAPd1 (Phe20, Arg21, Met22, Lys24, Leu28, Gln33, Leu35, His36, Ile51, Lys61, Leu62, Lys63, Leu64, Asp65, Glu74, Leu81, Asn82, Asp91 and Gly92). (f) Sequence of human RAPd1T (RAP residues Gly17–Arg97). The three lysine-rich motifs of RAPd1 are indicated with bars above the sequence and the α-helices with rounded rectangles below.

Table 3. Statistics of the top 5 CR56 - RAPd1 clusters obtained with HADDOCK

Cluster	Haddock score[a]	RMSD-E_{min}[b]	N[c]	E_{vdw}[d]	E^{d}_{elec}	E_{AIR}[e]	BSA (Å^2)	E_{desolv}[f]
Clust-2	−132 (19)	3.4 (1.8)	13	−52 (19)	−744 (108)	315 (43)	2413 (319)	37 (4)
Clust-1	−126 (13)	6.4 (0.9)	24	−55 (10)	−688 (79)	289 (76)	2263 (175)	38 (7)
Clust-5	−103 (17)	7.0 (0.5)	8	−44 (10)	−642 (70)	315 (49)	2029 (166)	37 (6)
Clust13	−93 (12)	10.4 (0.7)	4	−48 (9)	−624 (28)	388 (96)	2280 (182)	42 (4)
Clust-8	−87 (24)	11.1 (0.2)	6	−60 (9)	−495 (107)	386 (73)	2311 (233)	33 (6)

[a] Final HADDOCK score calculated as the weighted sum of: $0.2*E_{elec}+1.0*E_{vdW}+1.0*E_{desolv}+0.1*E_{AIR}$.
[b] Overall backbone RMSD from the lowest energy structure.
[c] Number of structures in a given cluster.
[d] The intermolecular energies (kcal mol^{-1}) were calculated with the OPLS parameters using a 8.5 Å cut-off.
[e] HADDOCK ambiguous interaction restraint energy (kcal mol^{-1}).
[f] The desolvation energy is calculated using the atomic desolvation parameters by Fernandez-Recio *et al.*[65]

between a ligand and CR domains from lipoprotein receptors. The following complexes are available. The crystal structures of the extra cellular domain of the LDL receptor,[3] the complex of the human rhinovirus and VLDLR-ligand binding domains,[35] and the present HADDOCK structure of the complex of RAP domain 1 and LRP modules CR56. We aimed at identifying a general recognition pattern that could be proposed to form a minimal set of interactions between a lipoprotein receptor binding module and a target protein. A comparison of the three interfaces so far available at atomic resolution all show the same pattern of interacting residues (Figure 7) forming (at least) a four-residue fingerprint motif (two from each protein). The receptor modules have a conserved acidic amino acid residue (aspartic or glutamic acid) and a conserved tryptophan residue, whereas the interacting protein has a conserved lysine residue that specifically interacts with the acidic side-chain *via* a salt bridge and forms a hydrophobic platform for the tryptophan side-chain. An additional hydrophobic side-chain from the interacting protein, typically a leucine or an isoleucine, seal the hydrophobic stack after the tryptophan side-chain. The fingerprint residues thereby become Asp/Glu(R)-Lys(L)-Trp(R)-Leu/Ile/Gln(L) (R = receptor, L = ligand). It is likely that this fingerprint is specific and is supplemented with additional van der Waals' interaction energy. Mutation of the Asp and Trp residues of the CR56 motif (Trp953 and Asp958 of CR5 and Trp994 and Asp999 of CR6) has clearly shown that binding to RAP is almost abolished, which support the importance of this motif.[34] Likewise, the mutation of the lysine of the RBD of α_2-macroglobulin (Lys1374) has a profound effect of binding to LRP.[41] A significance of the localisation of the minimal set of interactions is that the residues of this fingerprint are not coupled to any secondary structure preference of the interacting protein. In both the RBD of α_2-macroglobulin and in RAPd1 this motif is found within an α-helix, whereas in the two other structures, this motif is conferred from an assembly of loop structures. So, whether a recognition helix is a trait of a set of LRP ligands or from natural ligands only, remains to be shown, but secondary structure prediction programs do indicate that the residues of the arginine-rich motifs of RAPd2 and RAPd3 have high preference for being in α-helical environments. The proposition that LRP ligand binding could be ascribed to motifs of secondary or tertiary structure of the ligands and not to primary sequences alone has been hypothesized earlier.[44] Here we present direct results on LRP–RAP binding; these, together with comparisons to the structure of the RBD of α_2-macroglobulin and to the structure of intramolecular LDLR interactions

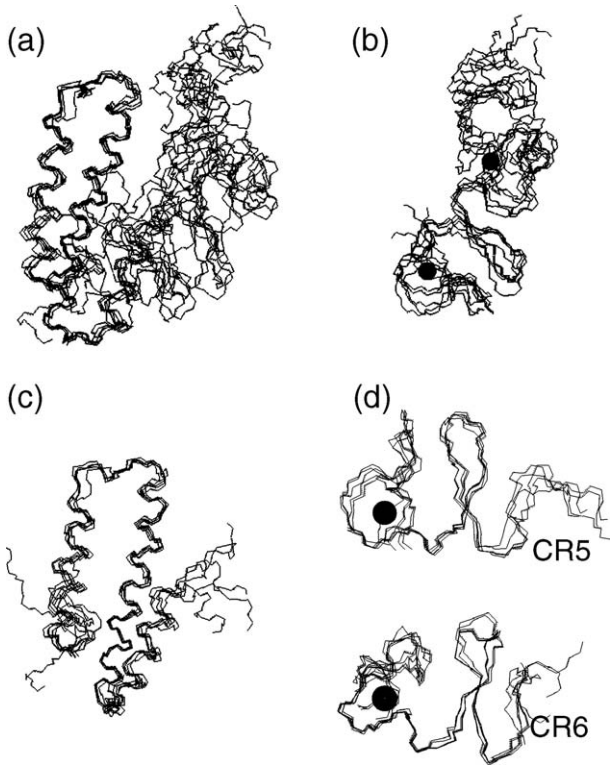

Figure 3. HADDOCK program models. Superimposition of the five lowest energy structures of the best HADDOCK cluster (2) with (a) the five lowest energy structures fitted with respect to the RAPd1 Cα, C, N backbone atoms of the three α-helices highlighting the overall definition of the two proteins with respect to the interface, (b) five lowest energy structure fitted to the backbone atoms C, N, Cα of the residues 936–971, 976–1011 of CR56, (c) of the residues in the three helices of RAPd1, (d) to each of the modules 936–971 (CR5) and 976–1011 (CR6), respectively.

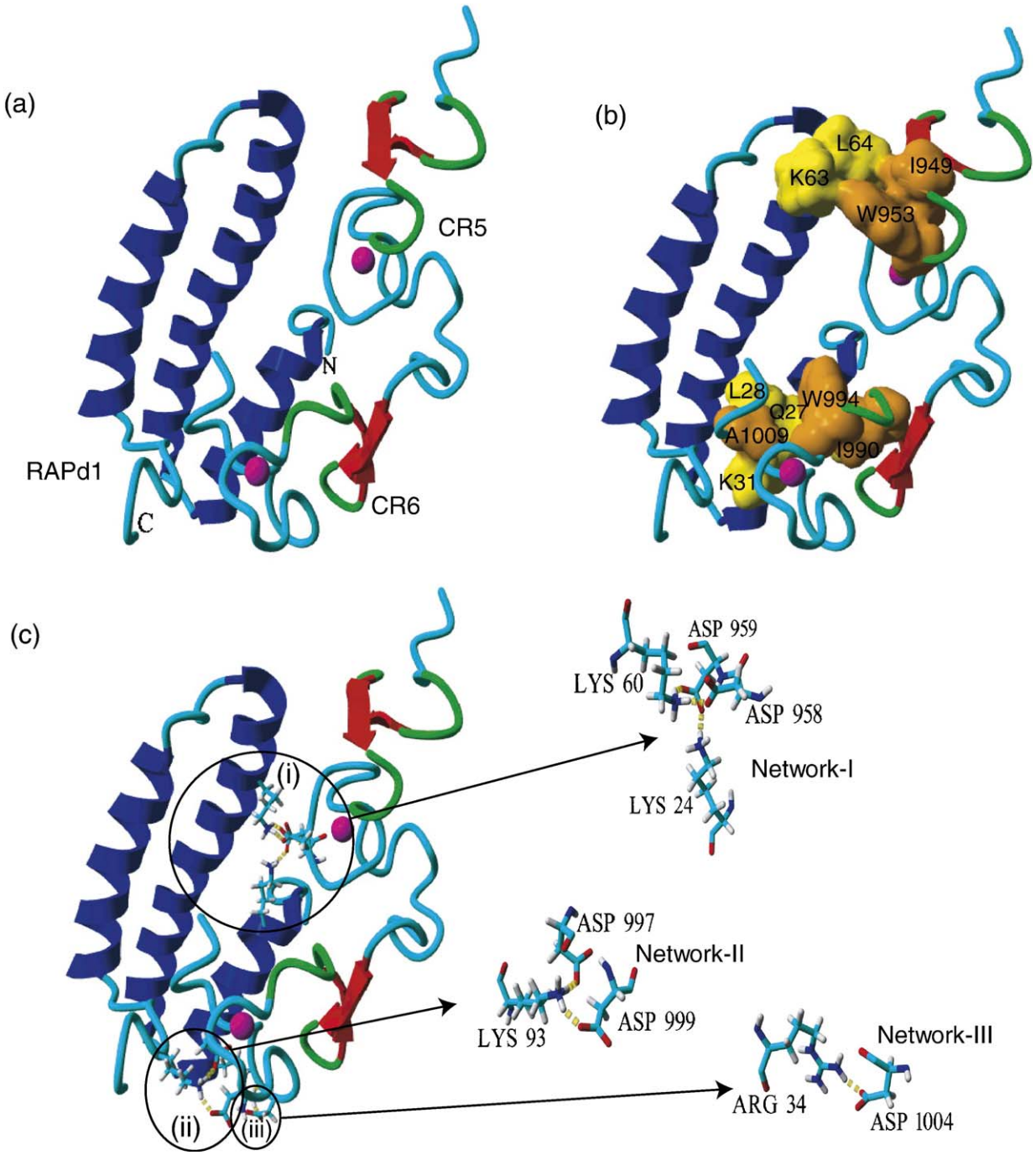

Figure 4. HADDOCK program model of the complex between RAPd1 and CR56 of LRP. (a) Backbone trace showing the three α-helices of RAPd1 and the two CR modules of CR56 with β-hairpins. (b) Hydrophobic interface with residues of Rapd1 and CR56 involved shown in van der Waals' surface representation in yellow and orange, respectively, and residues named accordingly. (c) The three networks of salt bridges between residues of RAPd1 and CR56. Network I is located below the hydrophobic cluster shown in (b) and network II at the rim of the interface.

and of a VLDLR–rhinovirus complex, strongly point to the existence of a common motif in LRP recognition.

Further relevance of the identified motif is supported by the recently described species-specific interaction of human rhinovirus serotype 1A (HRV1A) with murine but not human LDLR.[45] By inspection of the fingerprint residues (coordinating calcium *via* their backbone carbonyls) between these highly homologous receptor variants, notably they differ at a single position in the fifth LB repeat of the LDLR cluster. These residues are normally strictly conserved among species (unpublished observation), but whereas the mouse LA5 fingerprint is (Trp; Glu) the human variant contains a (Trp; Gly)-pair at these positions. Elegant studies recently conducted by Herdy and colleagues,[46] showed that substitution of the glycine within the

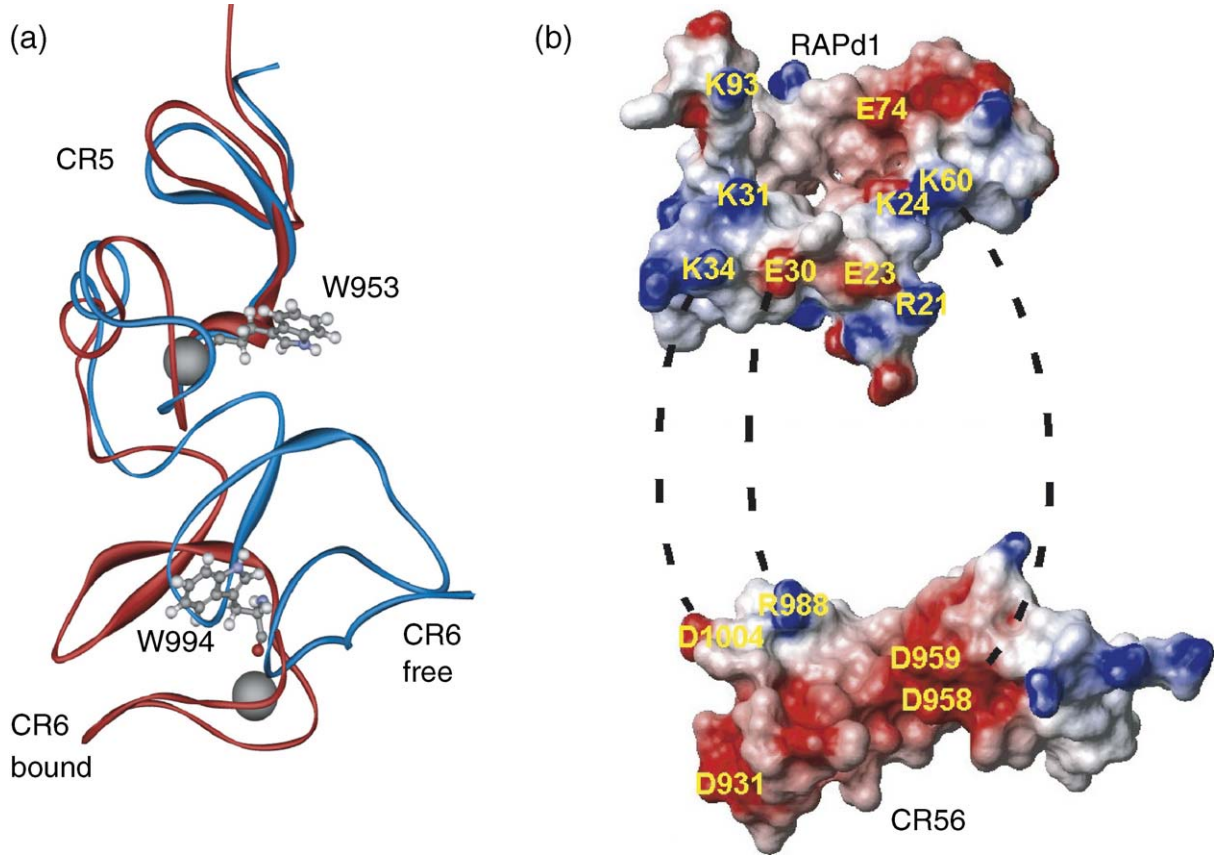

Figure 5. (a) Comparison of CR56 in its free and bound forms. The structures are aligned to residues 936–971 (CR5), and the change in modular orientation is apparent. The tryptophan of each module in the bound form is highlighted in ball-and stick. Free CR56 (blue ribbon), bound CR56 (red ribbon). (b) Electrostatic surface plots of the two interaction partners. CR56 (bottom) and RAPd1 (top). The complex is "opened", and the two partners separated and rotated each by 90° to display their interface. Residues involved in salt bridges are indicated and form the following CR56/RAPd1 pairs: Asp958/Lys24, Asp958/Lys60, Asp963/Lys93, and Arg964/Asp91.

human receptor by the acidic glutamic acid residue from the mouse receptor led to substantial higher affinity for HRV1A, again indicating the important function of the fingerprint motif in recognition of several ligands.

Models for ligand binding

The mode of action of the RAP molecules has been discussed and two predominant models have been proposed, one in which RAP acts to competitively and sterically hinder the binding of other ligand molecules,[47] and one in which RAP induces an allosteric conformational change in the LRP molecules and thereby mask the binding sites for other molecules.[4,48] In the case of the CR56 double module, given the electrostatic nature of the interaction and the limited number of free charges on the LRP molecule, it is very likely that RAP or domains of RAP occupy and mask all possible binding surfaces thereby sterically hindering binding of all other ligands. However, in light of the LDLR structure at low pH, which forms a closed end-over-end structure, it may also well be that RAP acts to preserve a closed structure of the receptor preventing the opening for constructive interaction.

Receptor-associated protein–domain 3

While this paper was under revision, two papers reported structures of unbound RAP domain 3[49] and of RAPd3 in complex with the double module LA34 of the LDL receptor.[50] The co-crystallized complex showed two almost identical, but discontinuous, docking sites for RAP, one for each module, and both having the suggested fingerprint motif, with an extended circular acidic necklace surrounding the essential lysine. The two essential lysine residues, Lys256 (binding to LA3) and Lys270 (binding to LA4), belong to the same α-helix, and are in the complex separated by four helical turns. For RAPd1 the essential lysine of the common binding motif, Lys60, is analogous to Lys270 of RAPd3, whereas four turn downstream no lysine is seen. One could argue that Lys50, which is three turn upstream from Lys60 could be the essential lysine of the second motif of RAPd1; however, mutation of this lysine to alanine has no effect on CR56–RAPd1 binding (Figure 6(b)). The mode of binding CR56 to RAPd1 is thus different to the binding of LA34 to RAPd3. Instead, CR56 binds to the opposite face of the N-terminal helix of the RAP domain 1, and anchors to Lys93. This mode of binding is not as strong as

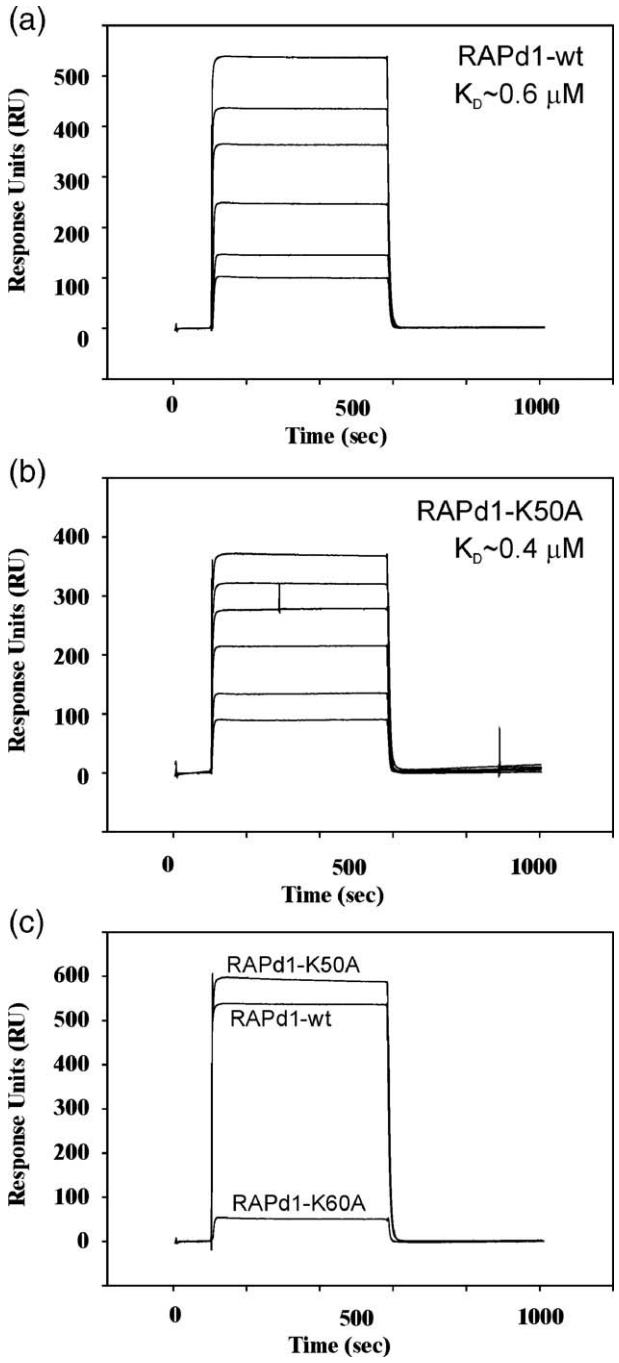

Figure 6. SPR analysis of RAPd1 binding to immobilized CR56. U-CR56 was immobilized on a biosensor chip with a density of 300 fmol/mm^2 and representative sensorgrams from the binding analysis are shown for (a) binding of wild-type (wt)-RAPd1 to U-CR56 (20, 50, 100, 200, 500, 1000 and 5000 nM) and (b) binding of RAPd1-K50A to U-CR56 (20, 50, 100, 200, 500, 1000 and 5000 nM). (c) Comparison of wt-RAPd1, RAPd1–K50A and RAPd1–K60A binding to U-CR56, all at 5 μM RAPd1. Plateau response levels in each sensorgram series correspond directly to RAP concentrations.

binding using two identical motifs as for RAPd3, which is reflected in the lower binding affinity observed for RAPd1 and also in the extent of chemical shift changes observed for CR6.

Conclusion

The NMR structure determination of the LRP double module CR56 and the binding studies with RAPd1 have resulted in a qualitative HADDOCK model of the complex between RAPd1 and CR56 from which the inter-atomic interactions important for receptor recognition could be described, and the model supported by site-directed mutant binding studies. Receptor interactions are accomplished *via* a combination of networks of charge–charge or polar contacts between RAP and the complement-type repeat itself together with significant hydrophobic interactions. The diversity in binding affinity among the RAP domains is a result of different binding motifs. Domain 1 contains a lysine-rich binding motif highly similar to that described for α_2-macroglobulin and for LDLR in its low pH form. We propose that the minimal common interactions motif between lipid receptor binding modules and ligands consists of the four-residue Asp/Glu-Lys-Trp-Ile/Leu fingerprint. Many of the natural LRP ligands may present this fingerprint within a recognition helix and the specificity of binding should be sought not only in the charge distribution of this α-helix, but certainly also in the four-residue fingerprint-binding motif.

Materials and Methods

Construction of expression plasmids and site-directed mutagenesis

The ubiquitin (U)-fused CR56 construct was produced as described by Andersen *et al*.[31] Mutations in the RAPd1 domain[9] were performed using Quick Change (Stratagene). For the RAPd1–K50A mutant, the primer pairs used were 5′- CTC CAC GCT GAT CTG GCG ATA CAG GAG AGG GAC-3′ and 5′- GTC CCT CTC CTG TAT CGC CAG ATC AGC GTG GAG-3′. For the RAPd1–K60A mutant, the primer pairs were 5′- GAC GAA CTC GCC TGG GCG AAA CTA AAG CTT GAC GGC-3′ and 5′- GCC GTC AAG CTT TAG TTT CGC CCA GGC GAG TTC GTC -3′. Primers were produced by TAG, Copenhagen.

Protein expression and purification

Recombinant CII-fusion proteins of RAPd1 and the RAPd1-K50A and RAPd1-K60A mutants were expressed in *Escherichiacoli* BL21 (DE3) cells, harvested by centrifugation and resuspended in 50 mM Tris–HCl (pH 8), 0.5 mM EDTA and 0.5 M NaCl before being sonicated. Following centrifugation at 30,000 g for 50 min, cell extract was loaded onto a Ni^{2+}- activated nitrilotriacetic acid Sepharose column (Qiagen) and washed with 50 mM Tris–HCl (pH 8), 2 mM $CaCl_2$ and 0.5 M NaCl. The protein was eluted from the Ni^{2+}- column with washing buffer, supplemented with EDTA to 10 mM and then dialyzed against 20 mM Tris–HCl (pH 8), 1 mM $CaCl_2$ and 50 mM NaCl before being cleaved with factor X_a for 2.5 h at room temperature. Factor X_a was removed by affinity chromatography using X_a Removal Resin

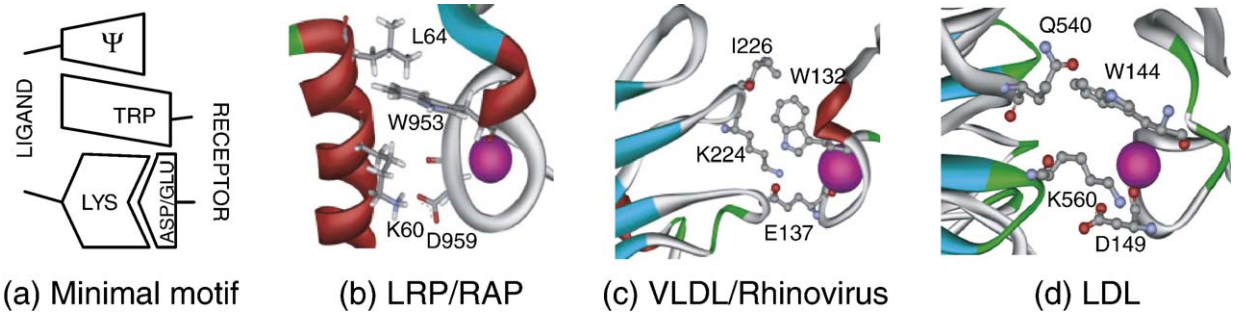

Figure 7. Suggested LRP recognition motif. (a) Schematic representation of the four residues D/E-K-W-Ψ interaction motif. Ψ is a hydrophobic residue or a residue with a significant non-polar part. (b)–(d) Views of the identified motifs in (b) LRP–RAP represented by the present HADDOCK model for CR56–RAPd1, (c) the structure of VLDLR in complex with the human rhinovirus (PDB accession code 1N7D) and (d) the low pH structure of LDL with interactions between LB5 and the β-propeller (PDB accession code 19VU).

(Qiagen). Before being loaded onto an anion exchange HiTrap Q column (GE Healthcare), the cleaved protein was once again applied to the Ni^{2+}- column, in order to remove non-cleaved protein, and then desalted on a PD10 column (GE Healthcare). Finally, the protein was concentrated and buffer exchanged into 10 mM Hepes (pH 7.4), 150 mM NaCl, 1.5 mM $CaCl_2$, 1 mM EGTA using Amicon Ultra Centrifugal filter units, molecular mass cut-off (MWCO) 5000 (Millipore). For NMR samples, RAPd1 (RAP residues Gly17–Leu112) was produced as described.[9] An N-terminal His6-tagged CR56 (LRP residues Ser932–His1013) was expressed in *E. coli*, and partly purified by conventional Ni^{2+}-affinity chromatography as reported.[31] Disulphide linkages were introduced by dialysis against a glutathione containing redox buffer system, and the native folded CR-domain pair was affinity purified using RAPd1-coupled Sepharose (binding in 10 mM Hepes (pH 7.4), 140 mM NaCl, 1 mM $MgCl_2$, and 12 mM $CaCl_2$, elution with 100 mM glycine (pH 3.0), 150 mM NaCl, and 20 mM EDTA). The His_6-tag was removed by factor Xa-processing of the fusion protein (in 50 mM Tris–HCl (pH 8.0), 500 mM NaCl, 5 mM $CaCl_2$) and passage over Ni^{2+}-NTA and factor Xa inhibitor Sepharoses, before the mature CR56 was concentrated using spin columns with a 3 kDa cut-off filter (Amicon). The production of ^{15}N-labelled proteins followed the protocols for unlabelled proteins, where bacteria were grown in defined minimal medium containing either $^{15}NH_4Cl$ or $^{15}NH_4SO_4$ as a single source for nitrogen. The production of ^{15}N,^{13}C-labelled proteins followed the same protocols as for the unlabelled proteins with $^{15}NH_4SO_4$ and [^{13}C]glucose as single sources for nitrogen and carbon, respectively.

Surface plasmon resonance (SPR) analysis

A CM5 BIAcore sensor chip was activated and purified U-CR56 protein was immobilized as described by Andersen *et al.*[31] Protein binding analysis of purified RAPd1, RAPd1–K50A and RAPd1–K60A was preformed in 10 mM Hepes (pH 7.4), 150 mM NaCl, 1.5 mM $CaCl_2$, 1 mM EGTA buffer supplemented with 0.005% (v/v) Tween-20 at a flow rate of 5 μl/min. This buffer was also used to equilibrate the chip before loading the proteins, and as running buffer. Injections of samples were preformed using the KINJECT option, and regeneration of the sensor chip was done using alternating pulses of 20 mM EDTA, 500 mM NaCl, glycine–HCl (pH 4.0) with 0.005% Tween-20 and of 0.01% (w/v) SDS. The binding data were analysed using the BIA Evaluation Program Version 3.0 (Biacore AB).

NMR spectroscopy

All NMR samples contained 100 mM NaCl, 10 mM $CaCl_2$, 1 mM $MgCl_2$ at pH 7.0 in 9:1 $H_2O/^2H_2O$, with protein concentrations from 0.02 mM to 0.5 mM. NMR spectra were recorded at 298 K on Varian Unity 750 MHz or 800 MHz spectrometers. Spectra were processed with MNMR§ or nmrPipe[51] and analysed using PRONTO3D.[52] Backbone and side-chain ^1H and ^{15}N resonances for CR56 were assigned using homonuclear 2D DQF-COSY, TOCSY (τ_m=46 ms), and NOESY spectra (τ_m=200 ms) supplemented with 3D ^1H-^{15}N TOCSY-HSQC (τ_m=46 ms) and ^1H-^{15}N NOESY-HSQC (τ_m=150 ms) spectra. Homonuclear 2D spectra were recorded with spectral widths of 10,000 Hz in both dimensions, and 4096 and 1024 data points were recorded in t_2 and t_1. The 3D ^1H-^{15}N TOCSY-HSQC spectrum was recorded with 1024, 64 and 332 data points in t_3, t_2 and t_1 and with corresponding spectral widths of 6000, 3000 and 12001 Hz. The 3D ^1H-^{15}N NOESY-HSQC spectrum was recorded with 1024, 48 and 188 data points in t_3, t_2, and t_1, and with corresponding spectral widths of 6504, 3200 and 13008 Hz. NOEs were assigned from the NOESY type spectra and classified into three classes: weak (1.8 Å–5.0 Å), medium (1.8 Å–3.3 Å), and strong (1.8 Å–2.7 Å), based on their intensities. The program DIANA[53] was used to filter out redundant ^1H-^1H NOEs. An additional qualitative NOE class was introduced (1.8 Å–6.0 Å) to account for very weak NOEs. Dihedral angle restraints for φ were obtained from $^3J_{HN}$-H^α coupling constants measured with water-flip-back constant-time (CT) HMQC-J experiments.[54] The CT-HMQC-J experiments were recorded with 3072 and 256 data points in t_2 and t_1 and with corresponding spectral widths of 12,000 and 3000 Hz. A set of ten CT-HMQC-J spectra were measured with the following evolution periods for the H^N-H^α *J* couplings: 48.2, 52.2, 68.2, 74.2, 80.2, 90.1, 100.2, 116.2, and 134.2 ms. The peak intensities as a function of this de-phasing period were measured using PRONTO3D[52] and fitted as described[54] using gnuplot. The H^N-H^α J couplings were translated into qualitative bounds for the corresponding φ angles as –60(±30)° for

§ Kjær, M., Andersen, K. V. and Rischel, C. (1992). MNMR Package for Multi-Dimensional NMR. Carlsberg Laboratory, Dept. of Chemistry, Copenhagen, Denmark.

H^N-H^α J couplings less than 5 Hz and as −139(±30)° for H^N-H^α J couplings greater than 8 Hz.

Perturbations of backbone ^1H and ^{15}N chemical shifts for CR56 and RAPd1(17-112) upon complex formation were obtained from ^1H-^{15}N HSQC spectra of the following samples: ^{15}N-labelled CR56, ^{15}N-labelled RAPd1(17-112), ^{15}N-labelled CR56 with two equivalent RAPd1, and ^{15}N-labelled RAPd1(17–112) with two equivalent CR56. Samples with ^{15}N-labelled RAPd1 were 0.1 mM, whereas samples with ^{15}N-labelled CR56 were 20 μM. A detailed NMR titration analysis was performed for the sample with ^{15}N-labelled CR56, where two equivalent of RAPd1 was added to the ^{15}N-labelled CR56 sample in ten steps and an ^1H-^{15}N-HSQC spectrum was recorded at each step. This was repeated twice for independent samples to check the reliability of the results. A corresponding single backward analysis was also performed for the two samples of ^{15}N-labelled RAPd1 with zero and two equivalent of CR56, respectively. For assignment of the bound forms of the proteins, exchange ^1H-^{15}N HSQC spectra[55] were measured of a sample with one equivalent of ^{15}N-labelled CR56 and half an equivalent of RAPd1, and a sample with one equivalent of ^{15}N-labelled RAPd1 and half an equivalent of CR56, respectively. To confirm the assignments of bound CR56 2D-^{15}N,^1H-HSQC, 3D-HNCA and 3D-HN(CO)CA spectra from the Varian Protein Pack version 1.6c were recorded on an 0.08 mM sample of a [^{13}C,^{15}N]CR56/[^{13}C,^{15}N]CRAPd1 1:1 complex and assigned.

[^{15}N]R1 and [^{15}N]R2 spectra[56] of samples with ^{15}N-labelled CR56 (1 mM) were recorded with 256 and 2048 data points in t_1 and t_2. Spectral widths were 3200 Hz in t_1 and 12,000 Hz in t_2. An array of 12 [^{15}N]R1 spectra was recorded with delay times of 10, 30, 50, 70, 90, 100, 150, 170, 210 and 250 ms, and an array of 12 [^{15}N]R2 spectra was recorded with delay times of 10, 30, 50, 100, 200, 300, 450, 600, 800, 1000, 1500 and 2000 ms. Each array of [^{15}N]R1 or [^{15}N]R2 spectra was recorded interleaved.

S3CT spectra[57] were recorded for RDC measurements at 37 °C with 175 and 4096 data points in t_1 and t_2, and spectral widths of 3200 Hz and 12,000 Hz in t_1 and t_2. Isotropic and anisotropic S3CT spectra were recorded in an aqueous solution of 100 mM NaCl, 10 mM $CaCl_2$, 1 mM $MgCl_2$ (pH 7.0) (isotropic) and in an aqueous solution with 7.5% (w/v) DMPC/DHPC/CTAB (30:10:1) (bicelles) in 100 mM NaCl, 10 mM $CaCl_2$, 1 mM $MgCl_2$ (pH 7.0) (anisotropic). The protein concentrations in the NMR samples were 0.43 mM (isotropic) and 0.19 mM (anisotropic).

Structure calculations

The experimentally determined distance and dihedral angle restraints were applied in an *ab initio* simulated annealing protocol with the X-PLOR[58] program. Distance restraints were introduced for the four ^1H-^1H NOE classes and dihedral angle restraints were introduced for the two classes of φ angles. Covalent patches for the six disulphide-bridges were included for the following cysteine/cysteine pairs, as inferred from previous data: LRP residues Cys936/Cys948, Cys943/Cys961, Cys955/Cys971, Cys976/Cys989, Cys984/Cys1002, and Cys996/Cys1011. Furthermore, six distance restraints were defined per calcium site to ensure calcium coordination throughout the structure calculations. These were inferred from comparative evaluation of the X-ray structure of LA5 from the canonical LDL receptor.[16] The following simulated annealing protocol was used: 10 ps at 3000 K in 5000 time steps followed by 6 ps with the temperature lowered linearly from 3000K to 100 K in 3000 time steps, and a final Powell minimization with 1000 steps. Force constants for the NMR restraints during X-PLOR calculations were set to 30 kcal/Å^2 for the H–H distance bounds and to 30 kcal/rad^2 for the φ angles. A total of 200 structures were calculated and the 20 lowest energy structures were selected and further refined with CNS in an explicit solvent layer of water, as described.[59] During the water refinement the calcium ions were removed and pseudo-calcium restraints were introduced to keep the calcium binding residues in an octahedral coordination. Force constants for the NMR restraints during CNS calculations were set to 50 kcal/Å^2 for the H–H distance bounds and to 200 kcal/rad^2 for the φ angles. Fifteen structures were selected out of the 20 calculated, having no ^1H-^1H NOE distance violations greater than 0.5 Å. The structural quality was evaluated using PROCHECK-NMR.[60] Figures showing three-dimensional structures and molecular surfaces were prepared using MOLSCRIPT,[61] YASARA)‖ or MOLMOL.[62]

Modelling of the CR56-RAPd1 complex

A model of the CR56-RAPd1 complex was calculated using the protein–protein docking program HADDOCK[36] (high ambiguity driven protein protein docking). The docking was performed from ensembles of structures. Since inter-module rearrangements could be involved in the binding to RAPd1, we first pre-sampled the inter-module orientation by running the HADDOCK protocol, but for CR56 only (PDB accession code 2FYJ, this work). CR56 was considered fully flexible and the structural integrity was maintained by using all experimental NOEs except for the inter-module ones involving residues Cys971 to Pro974. The resulting ensemble of 20 conformations was used for cross-docking against ten RAPd1 structures taken from the NMR structure of a C-terminal truncated form of RAPd1 (PDB accession code 1LRE; RAP residues 17–91).

The chemical shift perturbations observed for CR56 and RAPd1(17–112) upon complex formation were used to define ambiguous interaction Restraints (AIRs) for residues at the interface. Active residues were defined as those having chemical shift perturbations larger than the average + 0.5 standard deviation (0.243 ppm for RAPd1 and 0.168 ppm for CR56) and a relative residue accessible surface area larger than 55% for either side-chain or backbone atoms, as calculated with NACCESS¶). Note that the reported average chemical shift perturbation values were obtained in two steps: first by calculating the average over all residues and then by excluding residues with shifts larger than the first average value plus four times the standard deviation. This was done to avoid that very large shifts might result in too restrictive cut-offs for the definition of active residues. Passive residues were defined as all other surface accessible residues (relative residue accessible surface area larger than 55% for either side-chain or backbone atoms). AIRs were defined between every active residue of the first protein and all active and passive residues of the other protein, and *vice-versa*. Semi-flexible segments were defined as all active and passive residues ± 2 sequential residues. Residues

‖ http://www.yasara.org

¶ Hubbard, S. J. and Thornton, J. M. (1993). NACCESS. Department of Biochemistry and Molecular Biology, University College London.

within the segment from Ser970 to Phe977 constituting the linker, the last cysteine in CR5 and the first in CR6 together with the residue just before and just after these, whose chemical shift perturbation might reflect inter-domain rearrangements rather than direct interaction with RAP, were excluded from the AIR definition. Likewise, all NOEs involving residues Cys971-Pro974 were removed from the NOE list. In addition to the AIR restraints, pseudo-calcium restraints and 642 ^{1}H-^{1}H NOE distance restraints were included for CR56.

The docking was performed with the development version of HADDOCK[36] (version 2.0_devel). Starting from 20 CR56 pre-sampled structures (see above) and ten RAPd1 structures 2000 rigid-body solutions were generated (each of the possible 200 different combination of starting structure was thus used ten times). The best 400 solutions according to the HADDOCK rigid-body score were selected for semi-flexible refinement in torsion angle space; the top 200 structures were finally refined in explicit water. The score in the new version of HADDOCK is defined as a weighted sum of inter-molecular electrostatic (Elec) and van der Waals (vdW) energies, buried surface area (BSA), desolvation energy (Desolv) and AIR energy: rigid-body score = 1.0*Elec + 0.01*vdW − 0.01*BSA + 2.5*Desolv + 0.1*AIR, final score = 0.2*Elec + 1.0*vdW + 1.0*Desolv + 0.1*AIR). The desolvation energy is calculated using the atomic desolvation parameters of Fernandez-Recio *et al.*[65] A dielectric constant of 10 was chosen for the vacuum stages of the docking protocol (rigid-body and semi-flexible refinement). The final ensemble of 200 solutions was analysed and clustered based on the pair-wise RMSD matrix calculated over the backbone atoms of the interface residues of CR56 after fitting on the interface residues of RAPd1. This new way of calculating RMSD values in HADDOCK results in high values that emphasize the differences between docking solutions. For this reason the clustering was performed using a 7.5 Å cut-off. The clusters were ranked based on the averaged HADDOCK score of their top five members.

Coordinates

The structural coordinates have been deposited in the RCSB Protein Data Bank under the accession codes 2FYJ (CR56) and 2FYL (CR56:RAPd1 cluster2, model 5, lowest energy structure) and all NMR assignments and restraints have been deposited in the BioMagResBank under the accession code 5961.

Acknowledgements

This work is a contribution from the SBiN-Lab, and was supported by grants from the Danish Research Councils (to B.B. K.), the John and Birthe Meyer foundation (to F.M. P.), the Marie Curie foundation, and the Hede Nielsen foundation. A.M. J.J. B. acknowledges financial support from a "Jonge Chemici" grant from the Netherlands Organization for Scientific Research (NWO). M. E. and H.C. T. acknowledge financial support from the Danish Research Councils. We thank Peter Holme Jensen for recording of the first set of NMR spectra, Cyril Dominguez for his assistance with the initial HADDOCK calculations and Kresten Lindorff-Larsen with the PALES calculations. Signe A. Sjørup and Ove Lillelund are thanked for excellent technical assistance.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.07.013

References

1. Hussain, M. M. (2001). Structural, biochemical and signaling properties of the low-density lipoprotein receptor gene family. *Front Biosci.* **6**, D417–D428.
2. Nykjær, A., Petersen, C. M., Moller, B., Jensen, P. H., Moestrup, S. K., Holtet, T. L. *et al.* (1992). Purified alpha 2-macroglobulin receptor/LDL receptor-related protein binds urokinase plasminogen activator inhibitor type-1 complex. Evidence that the alpha 2-macroglobulin receptor mediates cellular degradation of urokinase receptor-bound complexes. *J. Biol. Chem.* **267**, 14543–14546.
3. Rudenko, G., Henry, L., Henderson, K., Ichtchenko, K., Brown, M. S., Goldstein, J. L. & Deisenhofer, J. (2002). Structure of the LDL receptor extracellular domain at endosomal pH. *Science*, **298**, 2353–2358.
4. Neels, J. G., van Den Berg, B. M., Lookene, A., Olivecrona, G., Pannekoek, H. & van Zonneveld, A. J. (1999). The second and fourth cluster of class A cysteine-rich repeats of the low density lipoprotein receptor-related protein share ligand-binding properties. *J. Biol. Chem.* **274**, 31305–31311.
5. Davis, C. G., Goldstein, J. L., Sudhof, T. C., Anderson, R. G., Russell, D. W. & Brown, M. S. (1987). Acid-dependent ligand dissociation and recycling of LDL receptor mediated by growth factor homology region. *Nature*, **326**, 760–765.
6. Li, Y., Lu, W., Schwartz, A. L. & Bu, G. (2002). Receptor-associated protein facilitates proper folding and maturation of the low-density lipoprotein receptor and its class 2 mutants. *Biochemistry*, **41**, 4921–4928.
7. Willnow, T. E., Rohlmann, A., Horton, J., Otani, H., Braun, J. R., Hammer, R. E. & Herz, J. (1996). RAP, a specialized chaperone, prevents ligand-induced ER retention and degradation of LDL receptor-related endocytic receptors. *EMBO J.* **15**, 2632–2639.
8. Bu, G. (2001). The roles of receptor-associated protein (RAP) as a molecular chaperone for members of the LDL receptor family. *Int. Rev. Cytol.* **209**, 79–116.
9. Ellgaard, L., Holtet, T. L., Nielsen, P. R., Etzerodt, M., Gliemann, J. & Thøgersen, H. C. (1997). Dissection of the domain architecture of the alpha2macroglobulin-receptor-associated protein. *Eur. J. Biochem.* **244**, 544–551.
10. Nielsen, P. R., Ellgaard, L., Etzerodt, M., Thøgersen, H. C. & Poulsen, F. M. (1997). The solution structure of the N-terminal domain of alpha2-macroglobulin receptor-associated protein. *Proc. Natl Acad. Sci. USA*, **94**, 7521–7525.
11. Daly, N. L., Scanlon, M. J., Djordjevic, J. T., Kroon, P. A. & Smith, R. (1995). Three-dimensional structure of a cysteine-rich repeat from the low-density lipoprotein receptor. *Proc. Natl Acad. Sci. USA*, **92**, 6334–6338.
12. Daly, N. L., Djordjevic, J. T., Kroon, P. A. & Smith, R. (1995). Three-dimensional structure of the second

cysteine-rich repeat from the human low-density lipoprotein receptor. *Biochemistry*, **34**, 14474–14481.
13. North, C. L. & Blacklow, S. C. (2000). Solution structure of the sixth LDL-A module of the LDL receptor. *Biochemistry*, **39**, 2564–2571.
14. Dolmer, K., Huang, W. & Gettins, P. G. (2000). NMR solution structure of complement-like repeat CR3 from the low density lipoprotein receptor-related protein. Evidence for specific binding to the receptor binding domain of human alpha(2)-macroglobulin. *J. Biol. Chem.* **275**, 3264–3269.
15. Huang, W., Dolmer, K. & Gettins, P. G. (1999). NMR solution structure of complement-like repeat CR8 from the low density lipoprotein receptor-related protein. *J. Biol. Chem.* **274**, 14130–14136.
16. Fass, D., Blacklow, S., Kim, P. S. & Berger, J. M. (1997). Molecular basis of familial hypercholesterolaemia from structure of LDL receptor module. *Nature*, **388**, 691–693.
17. Simonovic, M., Dolmer, K., Huang, W., Strickland, D. K., Volz, K. & Gettins, P. G. (2001). Calcium coordination and pH dependence of the calcium affinity of ligand-binding repeat CR7 from the LRP. Comparison with related domains from the LRP and the LDL receptor. *Biochemistry*, **40**, 15127–15134.
18. Kurniawan, N. D., Atkins, A. R., Bieri, S., Brown, C. J., Brereton, I. M., Kroon, P. A. & Smith, R. (2000). NMR structure of a concatemer of the first and second ligand-binding modules of the human low-density lipoprotein receptor. *Protein Sci.* **9**, 1282–1293.
19. North, C. L. & Blacklow, S. C. (1999). Structural independence of ligand-binding modules five and six of the LDL receptor. *Biochemistry*, **38**, 3926–3935.
20. Mikhailenko, I., Battey, F. D., Migliorini, M., Ruiz, J. F., Argraves, K., Moayeri, M. & Strickland, D. K. (2001). Recognition of alpha 2-macroglobulin by the low density lipoprotein receptor-related protein requires the cooperation of two ligand binding cluster regions. *J. Biol. Chem.* **276**, 39484–39491.
21. Moestrup, S. K. & Gliemann, J. (1991). Analysis of ligand recognition by the purified alpha 2-macroglobulin receptor (low density lipoprotein receptor-related protein). Evidence that high affinity of alpha 2-macroglobulin-proteinase complex is achieved by binding to adjacent receptors. *J. Biol. Chem.* **266**, 14011–14017.
22. Orth, K., Madison, E. L., Gething, M. J., Sambrook, J. F. & Herz, J. (1992). Complexes of tissue-type plasminogen activator and its serpin inhibitor plasminogen-activator inhibitor type 1 are internalized by means of the low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor. *Proc. Natl Acad. Sci. USA*, **89**, 7422–7426.
23. Warshawsky, I., Bu, G. & Schwartz, A. L. (1995). Sites within the 39-kDa protein important for regulating ligand binding to the low-density lipoprotein receptor-related protein. *Biochemistry*, **34**, 3404–3415.
24. Nielsen, K. L., Holtet, T. L., Etzerodt, M., Moestrup, S. K., Gliemann, J., Sottrup-Jensen, L. & Thøgersen, H. C. (1996). Identification of residues in alpha-macroglobulins important for binding to the alpha2-macroglobulin receptor/low density lipoprotein receptor-related protein. *J. Biol. Chem.* **271**, 12909–12912.
25. Andersen, O. M., Christensen, P. A., Christensen, L. L., Jacobsen, C., Moestrup, S. K., Etzerodt, M. & Thøgersen, H. C. (2000). Specific binding of alpha-macroglobulin to complement-type repeat CR4 of the low-density lipoprotein receptor-related protein. *Biochemistry*, **39**, 10627–10633.
26. Nykjær, A., Kjoller, L., Cohen, R. L., Lawrence, D. A., Garni-Wagner, B. A., Todd, R. F., III *et al.* (1994). Regions involved in binding of urokinase-type-1 inhibitor complex and pro-urokinase to the endocytic alpha 2-macroglobulin receptor/low density lipoprotein receptor-related protein. Evidence that the urokinase receptor protects pro-urokinase against binding to the endocytic receptor. *J. Biol. Chem.* **269**, 25668–25676.
27. Andersen, O. M., Petersen, H. H., Jacobsen, C., Moestrup, S. K., Etzerodt, M., Andreasen, P. A. & Thogersen, H. C. (2001). Analysis of a two-domain binding site for the urokinase-type plasminogen activator-plasminogen activator inhibitor-1 complex in low-density-lipoprotein-receptor-related protein. *Biochem. J.* **357**, 289–296.
28. Kuchenhoff, A., Harrach-Ruprecht, B. & Robenek, H. (1997). Interaction of apo E-containing lipoproteins with the LDL receptor-related protein LRP. *Am. J. Physiol.* **272**, C369–C382.
29. Williams, S. E., Inoue, I., Tran, H., Fry, G. L., Pladet, M. W., Iverius, P. H. *et al.* (1994). The carboxyl-terminal domain of lipoprotein lipase binds to the low density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor (LRP) and mediates binding of normal very low density lipoproteins to LRP. *J. Biol. Chem.* **269**, 8653–8658.
30. Moestrup, S. K., Holtet, T. L., Etzerodt, M., Thøgersen, H. C., Nykjaer, A., Andreasen, P. A. *et al.* (1993). Alpha 2-macroglobulin-proteinase complexes, plasminogen activator inhibitor type-1-plasminogen activator complexes, and receptor-associated protein bind to a region of the alpha 2-macroglobulin receptor containing a cluster of eight complement-type repeats. *J. Biol. Chem.* **268**, 13691–13696.
31. Andersen, O. M., Christensen, L. L., Christensen, P. A., Sorensen, E. S., Jacobsen, C., Moestrup, S. K. *et al.* (2000). Identification of the minimal functional unit in the low density lipoprotein receptor-related protein for binding the receptor-associated protein (RAP). A conserved acidic residue in the complement-type repeats is important for recognition of RAP. *J. Biol. Chem.* **275**, 21017–21024.
32. Iadonato, S. P., Bu, G., Maksymovitch, E. A. & Schwartz, A. L. (1993). Interaction of a 39 kDa protein with the low-density-lipoprotein-receptor-related protein (LRP) on rat hepatoma cells. *Biochem. J.* **296**, 867–875.
33. Obermoeller, L. M., Warshawsky, I., Wardell, M. R. & Bu, G. (1997). Differential functions of triplicated repeats suggest two independent roles for the receptor-associated protein as a molecular chaperone. *J. Biol. Chem.* **272**, 10761–10768.
34. Andersen, O. M., Schwarz, F. P., Eisenstein, E., Jacobsen, C., Moestrup, S. K., Etzerodt, M. & Thøgersen, H. C. (2001). Dominant thermodynamic role of the third independent receptor binding site in the receptor-associated protein RAP. *Biochemistry*, **40**, 15408–15417.
35. Verdaguer, N., Fita, I., Reithmayer, M., Moser, R. & Blass, D. (2004). X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. *Nature Struct. Biol.* **10**, 1–6.
36. Dominguez, C., Boelens, R. & Bonvin, A. M. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737.
37. Nielsen, P. R. (1996). *The Structural Basis for Receptor Binding*, Ph. D. thesis, Department of Chemistry, Carlsberg Laboratory, Denmark.

38. van Dijk, A. D., de Vries, S. J., Dominguez, C., Chen, H., Zhou, H. X. & Bonvin, A. M. (2005). Data-driven docking: HADDOCK's adventures in CAPRI. *Proteins: Struct. Funct. Genet.* **60**, 232–238.
39. van Dijk, A. D., Fushman, D. & Bonvin, A. M. (2005). Various strategies of using residual dipolar couplings in NMR-driven protein docking: application to Lys48-linked di-ubiquitin and validation against ^{15}N-relaxation data. *Proteins: Struct. Funct. Genet.* **60**, 367–381.
40. Melman, L., Cao, Z. F., Rennke, S., Marzolo, M. P., Wardell, M. R. & Bu, G. (2001). High affinity binding of receptor-associated protein to heparin and low density lipoprotein receptor-related protein requires similar basic amino acid sequence motifs. *J. Biol. Chem.* **276**, 29338–29346.
41. Migliorini, M. M., Behre, E. H., Brew, S., Ingham, K. C. & Strickland, D. K. (2003). Allosteric modulation of ligand binding to low density lipoprotein receptor-related protein by the receptor-associated protein requires critical lysine residues within its carboxyl-terminal domain. *J. Biol. Chem.* **278**, 17986–17992.
42. Jenner, L., Husted, L., Thirup, S., Sottrup-Jensen, L. & Nyborg, J. (1998). Crystal structure of the receptor-binding domain of alpha 2-macroglobulin. *Structure*, **6**, 595–604.
43. Huang, W., Dolmer, K., Liao, X. & Gettins, P. G. (2000). NMR solution structure of the receptor binding domain of human alpha(2)-macroglobulin. *J. Biol. Chem.* **275**, 1089–1094.
44. Hussain, M. M., Strickland, D. K. & Bakillah, A. (1999). The mammalian low-density lipoprotein receptor family. *Annu. Rev. Nutrit.* **19**, 141–172.
45. Reithmayer, M., Reischl, A., Snyers, L. & Blass, D. (2002). Species-specific receptor recognition by a minor-group human rhinovirus (HRV): HRV serotype 1A distinguishes between the murine and the human low-density lipoprotein receptor. *J. Virol.* **76**, 6957–6965.
46. Herdy, B., Snyers, L., Reithmayer, M., Hinterdorfer, P. & Blass, D. (2004). Identification of the human rhinovirus serotype 1A binding site on the murine low-density lipoprotein receptor by using human-mouse receptor chimeras. *J. Virol.* **78**, 6766–6774.
47. Bu, G. & Marzolo, M. P. (2000). Role of rap in the biogenesis of lipoprotein receptors. *Trends Cardiovasc. Med.* **10**, 148–155.
48. Horn, I. R., van Den Berg, B. M., van der Meijden, P. Z., Pannekoek, H. & van Zonneveld, A. J. (1997). Molecular analysis of ligand binding to the second cluster of complement-type repeats of the low density lipoprotein receptor-related protein. Evidence for an allosteric component in receptor-associated protein-mediated inhibition of ligand binding. *J. Biol. Chem.* **272**, 13608–13613.
49. Lee, D., Walsh, J. D., Mikhailenko, I., Yo, P., Migliorini, M., Harris, B. *et al.* (2006). RAP uses a histidine switch to regulate its interaction with LRP in the ER and golgi. *Mol. Cell*, **22**, 423–430.
50. Fisher, C., Berglova, N. & Blacklow, S. C. (2006). Structure of an LDLR-RAP complex reveals a general mode of ligand recognition by lipoprotein receptors. *Mol. Cell*, **22**, 277–283.
51. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
52. Kjær, M., Andersen, K. V. & Poulsen, F. M. (1994). Automated and semiautomated analysis of homo- and heteronuclear multidimensional nuclear magnetic resonance spectra of proteins: the program Pronto. *Methods Enzymol.* **239**, 288–307.
53. Guntert, P., Braun, W. & Wuthrich, K. (1991). Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517–530.
54. Kuboniwa, H., Grzesiek, S., Delaglio, F. & Bax, A. (1994). Measurement of HN-H alpha J couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods. *J. Biomol. NMR*, **4**, 871–878.
55. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. (1994). A heteronuclear correlation experiment for simultaneous determination of 15N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR*, **4**, 727–734.
56. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G. *et al.* (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry*, **33**, 5984–6003.
57. Lerche, M. H., Meissner, A., Poulsen, F. M. & Sorensen, O. W. (1999). Pulse sequences for measurement of one-bond (15)N-(1)H coupling constants in the protein backbone. *J. Magn. Reson.* **140**, 259–263.
58. Brünger, A. T. (1992). *X-PLOR 3.1:* a system for X-ray crystallography and NMR. Yale University, New Haven.
59. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M. & Nilges, M. (2003). Refinement of protein structures in explicit solvent. *Proteins: Struct. Funct. Genet.* **50**, 496–506.
60. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.
61. Kraulis, P. J. (1991). MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
62. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.
63. Jorgensen, W. L. & Tirado-Rives, J. (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **110**, 1657–1666.
64. Zweckstetter, M. & Bax, A. (2000). Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. *J. Am. Chem. Soc.* **122**, 3791–3792.
65. Fernandez-Recio, J., Totrov, A. & Abagyan, R. (2004). Identification of protein–protein interaction sites from docking energy landscapes. *J. Mol. Biol.* **335**, 843–865.

Edited by M. F. Summers

(Received 29 March 2006; received in revised form 5 July 2006; accepted 7 July 2006)
Available online 15 July 2006