Structure prediction of LDLR-HNP1 complex based on docking enhanced by LDLR binding 3D motif

Reyhaneh Esmaielbeiki[1*], Declan P. Naughton[1] and Jean-Christophe Nebel[1].

[1] Faculty of Science, Engineering and Computing, Kingston University, Kingston-upon-Thames, KT1 2EE, UK

*Correspondence to: Reyhaneh Esmaielbeiki, Faculty of Science, Engineering and Computing, Kingston University, Kingston-upon-Thames, KT1 2EE, UK. T: +44 (0)20 8417 7159. F:+44 (0)20 8417 2972. E-mail: R.Esmaielbeiki@kingston.ac.uk.

Date of initial submission: 14-07-2011

Date of revised submission: -

Date of final acceptance: 26-08-2011

ABSTRACT

Human antimicrobial peptides (AMPs), including defensins, have come under intense scrutiny owing to their key multiple roles as antimicrobial agents. Not only do they display direct action on microbes, but also recently they have shown to interact with the immune system to increase antimicrobial activity. Unfortunately, since mechanisms involved in the binding of AMPs to mammalian cells are largely unknown, their potential as novel anti-infective agents cannot be exploited yet. Following the reported interaction of Human Neutrophil Peptide 1 dimer (HNP1) with a low density lipoprotein receptor (LDLR), a computational study was conducted to discover their putative mode of interaction.

State-of-the-art docking software produced a set of LDLR-HNP1 complex 3D models. Creation of a 3D motif capturing atomic interactions of LDLR binding interface allowed selection of the most plausible configurations. Eventually, only two models were in agreement with the literature.

Binding energy estimations revealed that not only one of them is particularly stable, but also interaction with LDLR weakens significantly bonds within the HNP1 dimer. This may be significant since it suggests a mechanism for internalisation of HNP1 in mammalian cells.

In addition to a novel approach for complex structure prediction, this study proposes a 3D model of the LDLR-HNP1 complex which highlights the key residues which are involved in the interactions. The putative identification of the receptor binding mechanism should inform the future design of synthetic HNPs to afford maximum internalisation, which could lead to novel anti-infective drugs.

Keywords: Low density lipoprotein Receptor; 3D motif; protein-protein interaction; docking; human alpha defensin; human immune system.

INTRODUCTION

Human antimicrobial peptides (AMPs), including mammalian defensins, have come under intense scrutiny owing to their key multiple roles as antimicrobial agents against a range of bacteria, fungi and viruses. These roles have been reported to involve immunostimulation via chemotaxis, direct action on viral particles, and binding to, followed by internalisation, into mammalian cells where antimicrobial activity is manifested through inhibition of viral replication, via inhibition of protein kinase C signalling [1, 2, 3].

These molecules provide enormous scope for the investigation of mechanisms involved in infection, along with immune response events, and represent a reservoir of potential novel anti-infective agents. In this vein, the use of synthetic AMPs to treat HIV was reported as early as 1993 [4]. Given the appearance and growth in numbers of drug resistant infections, and the relative paucity of new clinically effective antimicrobial agents, further studies are warranted to optimise the activities of natural and synthetic AMPs.

One key step, which requires further study, is to optimise the binding of (synthetic) AMPs to mammalian cells to afford internalisation for intracellular defences to operate. Following the reported interaction of human α-defensins with a low density lipoprotein receptor (LDLR) [1, 5], a plausible approach is to study potential interactions between AMPs and the LDLR.

The LDLR family contains seven homologous members and is responsible for mediating different types of ligands especially cholesterol into the cell [6]. Their structure is composed of several domains which include a ligand binding (LB) domain [7] composed of ligand binding modules (LAs), also named complement-type repeats (CRs), a beta-propeller domain and transmembrane and cytoplasmic sections (Figure 1, top row).

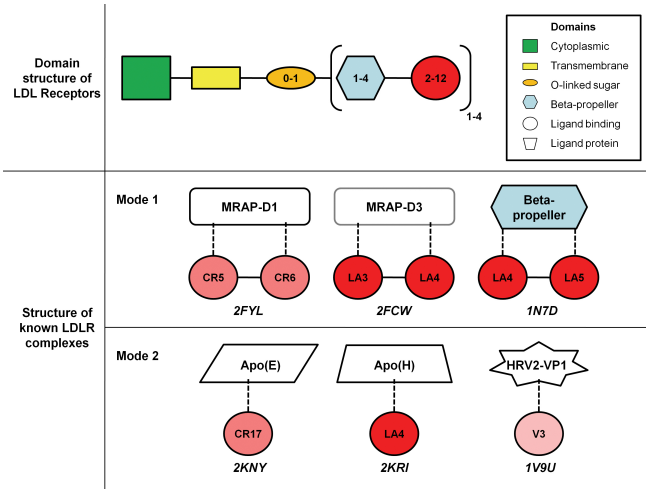

Figure **1**. Modular structure of LDLR receptor family: general domain pattern (top) and schematic representation of the LDLR-ligand binding modes of known complexes (bottom).

The Low-density lipoprotein receptor family interacts with a wide variety of human and virion proteins [8] through their homologous LA modules which are between 40-50 residues long [9,10] (Figure 2). LA's structure is stabilised by three disulfide bonds and a calcium ion [11, 12]. This ion is an essential element of the ligand binding domain conformation since it is required to establish interaction between LDLR and the ligand [8, 13].

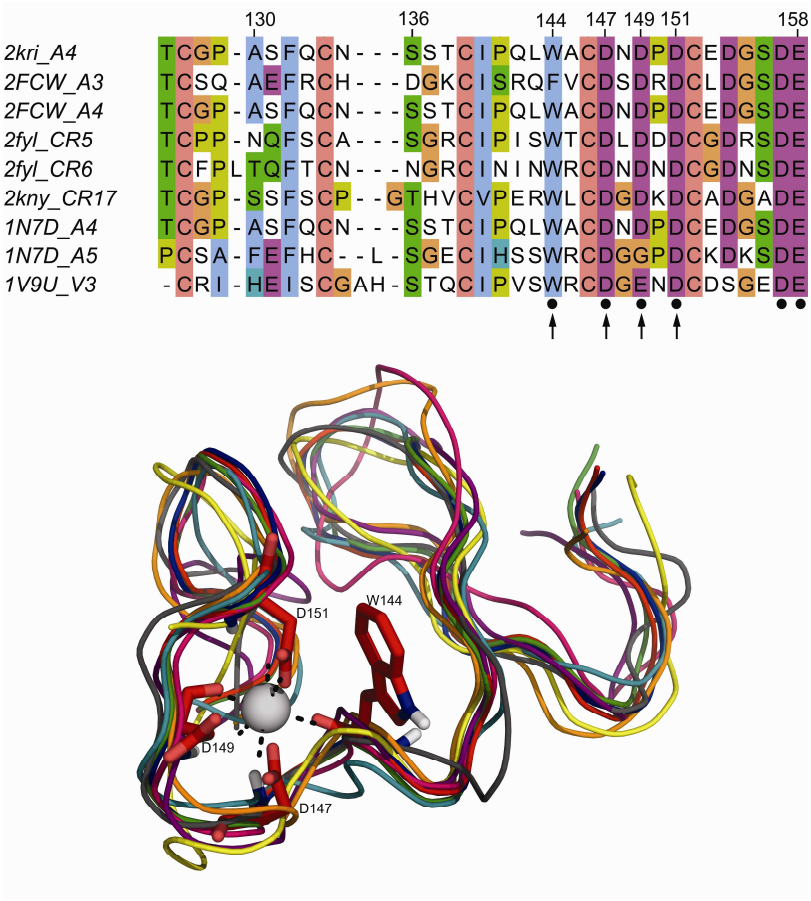

Figure **2**. Multiple sequence and structure alignment of ligand binding domains of LDLR family complexes. In the sequence alignment, residues involved in calcium interaction are denoted by dots. The three conserved acidic residues and conserved tryptophan/phenylalanine are highlighted with black arrows. Sequence numbering is based on 2KRI:B. In the structure alignment the three conserved acidic residues and conserved tryptophan are shown on 2KRI structure. Residues numbering is based on 2KRI:B. The ligand binding domains associated with each colour are: 2KRI-A4: red, 2FCW-A3:green, 2FCW-A4:dark blue, 2FYL-CR5:purple, 2FYL-CR6:yellow, 2KNY-CR17:orange, 1N7D-A4:deep teal, 1N7D-A5: gray, 1V9U-V3: pink.

High-resolution crystal structures of the available LDLR complexes have revealed that electrostatic forces play an important role in interactions [8]. This key function is captured by the minimal interaction motif described by Jensen et al. [14] (Figure 3), which also highlights a hydrophobic element in the interaction. The receptor's conserved acidic residues (ASP/GLU) interact with a ligand's lysine through a salt bridge creating a hydrophobic environment for the side chain of a receptor's tryptophan (TRP). In addition, a hydrophobic side chain, ψ, (usually Leucine or isolecine) from the ligand sits next to TRP.

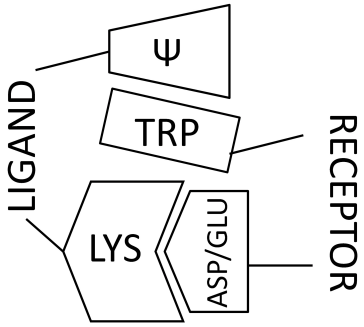

Figure **3**. Minimal binding motif defined by Jensen et al. [14].

This paper investigates the mode of interaction between a class of α-defensins and LDLR by the mean of predicted structural models [1, 5]. First, we produce a novel 3D motif which describes the binding characteristics of LDLR-ligand interactions. Then, the motif is used as constraint to evaluate LDLR-α-defensin complex models generated by state of the art docking software.

MATERIALS AND METHODS

Protein Data sets

Our study relies on the investigation of all 3D complexes involving a ligand binding domain of the LDLR family. Query of the RCSB Protein Data Bank [15] using Blast [16] on March 2011 revealed that the structures of six complexes have been resolved (Table 1). They all involve human proteins belonging to three members of the LDLR family, i.e. Low-density Lipoprotein receptor (LDLR), lipoprotein receptor-related protein 1 (LRP1) and Very low-density lipoprotein receptor (VLDLR). The sequences of their ligand binding domain were extracted from Uniprot [17], where their accession numbers are P01130: LDLR, P98155: VLDLR and Q07954: LRP1, respectively. Although their ligand binding modules are named LA, CR and V for LDLR, LRP and VLDR, respectively, in this paper we use LA when referring to any of them.

Table **1**. Known 3D structures of complexes involving members of the LDLR family.

PDB Code	**Receptor and Domains**	**Ligand**	**Ligand Complete Name**
2FCW [8]	LDLR LA3,LA4	MRAP D3	Alpha-2-macroglobulin receptor-associated protein Domain 3
2FYL [14]	LRP CR5,CR6	MRAP D1	Alpha-2-macroglobulin receptor-associated protein Domain 1
2KRI [9]	LDLR LA4	Apo(H)	Beta-2-glycoprotein 1
2KNY [11]	LRP CR17	Apo(E)	Apolipoprotein E
1N7D [33]	LDLR LA4,LA5	LDLR beta propeller	-
1V9U [34]	VLDLR V3	HRV2 VP1	Human rhinovirus 2 Viral Protein 1

In agreement with the existing 2D motif [14], sequence alignment of the LA modules (Figure 2) using ClustalW [18] shows highly conserved acidic residues and a tryptophan/phenylalanine (TRP/PHE) - pairwise E-values were calculated using Blast [16] (Table 2). Structural conservation of the ligand binding domains of the receptors, i.e. LA3, LA4, LA5, CR5, CR6, CR17 and V3, are illustrated (Figure 2) and quantified (Table 2) using the 3D alignment tool Pymol [19].

Since LA4 is the domain which is the most common in this set - in three cases out of six – it is used as representative for the purpose of α-defensin docking. Similarly, among these AMPs, defensin Human Neutrophil Peptide-1 (HNP1), which has been specifically shown to interact with LDLR [1, 5], is selected as representative. Sequences and structures of HNP1 and LA4 were extracted from the PDB [15]: 3GNY [20] and 2KRI [9] codes respectively.

Table **2.** E-value between sequences of the Ligand binding domains and RMSD between their 3D structures.

Sequence Similarity									
	2FCWA3	**2FCWA4**	**2FYLCR5**	**2FYLCR6**	**2KRIA4**	**2KNYCR17**	**1N7DA4**	**1N7DA5**	**1V9UVD3**
2FCWA3	1e-19	2e-09	3e-08	2e-08	2e-09	4e-08	1e-07	4e-07	2e-04
2FCWA4		1e-23	2e-08	6e-08	2e-21	1e-10	1e-18	9e-08	2e-06
2FYLCR5			3e-23	2e-11	2e-08	3e-09	3e-07	4e-08	1e-06
2FYLCR6				1e-22	6e-08	2e-06	1e-07	1e-08	2e-06
2KRIA4					2e-21	1e-10	1e-18	8e-08	4e-06
2KNYCR17						9e-30	6e-08	5e-05	4e-06
1N7DA4							2e-23	2e-07	3e-06
1N7DA5								1e-23	2e-07
1V9UVD3									8e-22
RMSD Between Structures (Å)									
	2FCWA3	**2FCWA4**	**2FYLCR5**	**2FYLCR6**	**2KRIA4**	**2KNYCR17**	**1N7DA4**	**1N7DA5**	**1V9UVD3**
2FCWA3	0	0.29	0.84	1.16	0.43	0.77	0.96	0.68	0.38
2FCWA4		0	0.6	3.18	0.43	0.77	1.07	0.67	0.64
2FYLCR5			0	3.19	0.73	1.04	1.24	0.85	1.00
2FYLCR6				0	3.89	3.01	4.62	4.27	3.47
2KRIA4					0	0.93	1.67	0.92	0.53
2KNYCR17						0	1.43	1.06	1.28
1N7DA4							0	1.41	1.20
1N7DA5								0	0.84
1V9UVD3									0

Creation of 3D motif

Extending the existing 2D motif [14] using the approach suggested by Nebel et al. [21, 22], we produce a 3D motif which describes the conserved 3D positions of the key atoms involved in LDLR-ligand interaction (Figure 4.A). From the receptor, the conserved acidic residues and TRP/PHE are represented by their alpha carbon atoms. In addition, in order to add a constraint regarding interaction with the calcium ion, we include the oxygen

atom of the TRP/PHE carboxyl group with whom it interacts. On the ligand side, the basic residue interaction is expressed by the side-chain nitrogen atom(s) which form(s) hydrogen bond(s) with acidic residues of the receptor.

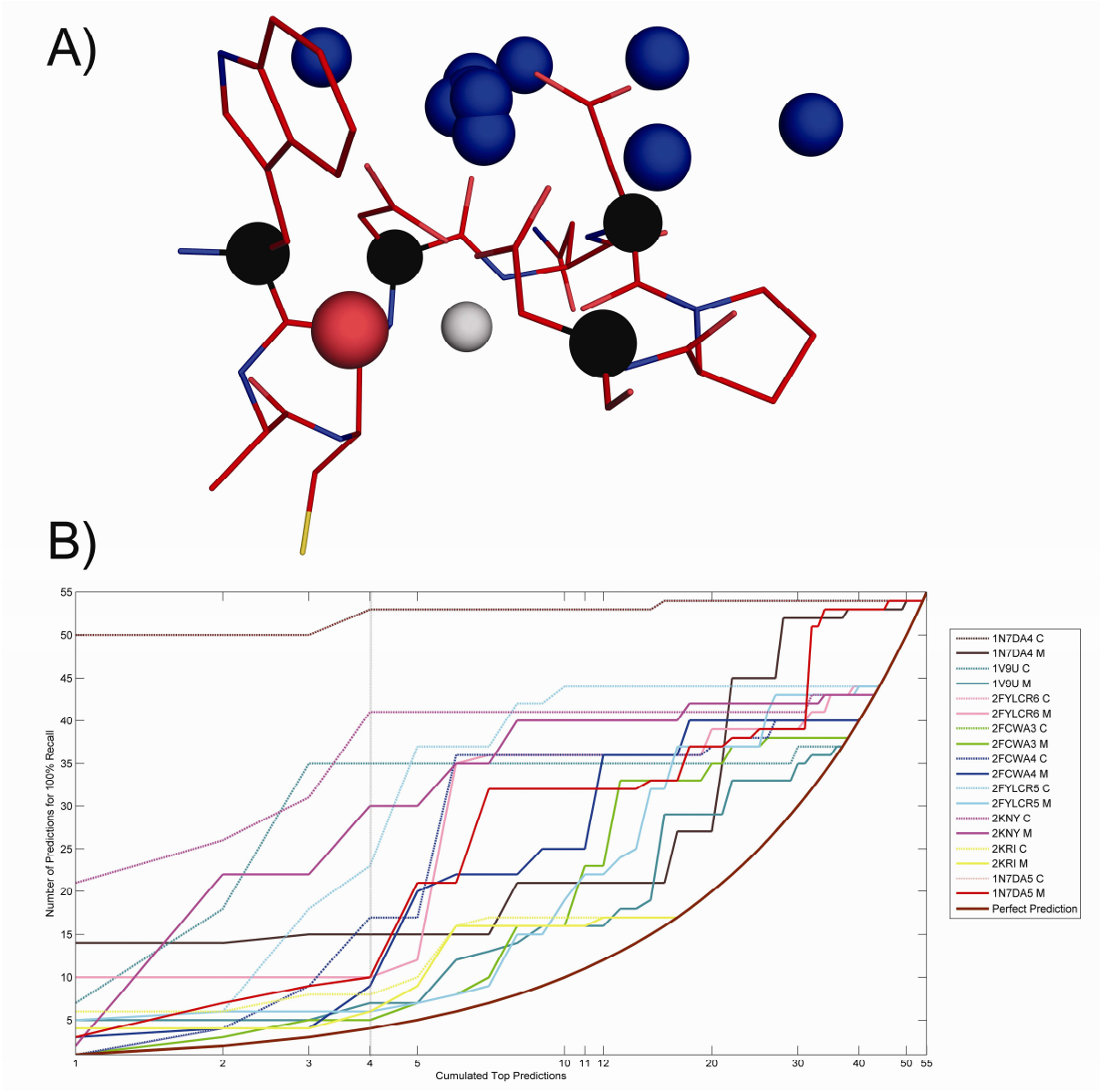

Figure **4**. (A) The 3D motif is represented by spheres. The blue ones show positions of N atoms from the ligand. The black ones are the C-alpha atoms of the ASP and TRP and the red sphere is the O atom of TRP. Location of the calcium is marked by a grey sphere. Image produced using Pymol. (B) Number of ranks to achieve 100% recall of the top predictions (or recall of top predictions with top positions). In the legend the complexes names are followed by C and M for curves based on cluster size and 3D motif method, respectively.

The actual coordinates of the consensus atoms forming the 3D motif are calculated by the multiple structure alignments of these atoms using all available receptor-protein complexes. Here, only atoms from the receptor side are used as superimposition constraints. Since their 3D structures are very well conserved - their average RMSD is 0.28 Å - positions of all receptor atoms in the 3D motif are approximated by the average coordinates

of the aligned atoms. On the other hand, given that every ligand displays a very different receptor binding sites, there is no consensus 3D position regarding the location of the nitrogen atom(s) of the basic residue(s). However, there must be specific constraints in terms of their distance and orientation from the receptor. In our 3D motif, we express implicitly these constraints by storing all the actual nitrogen positions available in our training set.

Note that among the 9 binding sites of the available structures, we excluded that of 2FYLCR6 in the construction of our motif since its LA module is structurally different from the others as measured by an average RMSD of 2.97 Å (Table 2).

Docking

Docking predictions are performed using the ClusPro 2.0 docking program [23], which, in addition to be freely available for academic research, has demonstrated to be the state of the art at CAPRI 2009 (Critical Assessment of Predicted Interactions) [24, 25, 26]. Cluspro works by initially calculating 70,000 docking models. Then, the 1000 models with the best energy conformation are selected and clustered using PIPER [27]. Models with the most neighbours within a 9 Å C-alpha RMSD cut-off are chosen as cluster representatives and are qualified by the size of their associated cluster.

The ClusPro docking results are generated according to different constraints. For each category, software produces a set of predicted models ranked according to their cluster size. Since previous studies have highlighted the important role of electrostatic and hydrophobic interactions in LDLR complexes [8, 9, 14], we only consider predictions generated under 'electrostatic favoured' and 'Van der Waals + Electrostatic forces (VDW/elec)' modes. In this work default software parameters are used.

3D motif evaluation

Our 3D motif was evaluated in docking prediction task using a leave-one-complex-out cross validation. First, a resolved 3D complex involving LDLR is selected. Secondly, a 3D motif is produced using all the other available LDLR complexes. Thirdly, the two chains involved in the complex are submitted to ClusPro which generates a set of putative complex models. Then, the fitting of the 3D motif to each model is used to score predicted complexes. Finally, the produced ranked list is compared with the list of models ranked according to their quality as expressed by their RMSD with respect to the actual resolved structure.

3D motif fitting is performed by superimposition on the binding site of the predicted LDLR-ligand complex using receptor atoms as constraints. We define the quality of a prediction as the shortest distance between the nitrogen of the basic residue of the ligand and those present in the 3D motif.

LDLR-HNP1 model prediction

Using the procedure previously described, LDLR-HNP1 complex estimates are generated by ClusPro and ranked using our 3D motif fitting measure. Then, the best models according to that score are further analysed in order to establish which ones are in agreement with the literature.

Finally, the stability of the remaining modelled complexes is quantified by both calculating the number of intermolecular contacts and estimating pair wise interaction energies between the different chains involved in those complexes. Detailed information on residue-residue and atom-atom contacts is provided by the Contact Map Analysis server which is part of the software suite SPACE [28]. In addition, since previous studies [29, 30] have shown good correlation between experimental measurements and energy calculations produced by the FoldX software [31, 32], its latest version, v3.0 beta5.1 (http://foldx.embl.de/) has been selected to evaluate binding energy between the two HNP1 monomers and between LDLR and each of the HNP1 chains.

RESULTS

Modes of interactions of LDLR-ligand complexes

Within the LDLR-ligand complexes, two modes of interaction between the LB module and the ligand have been identified (Figure 1, bottom rows). In the first mode, two ligand binding modules of LDLR are required to establish an interaction with the ligand. In 2FCW [8] the third and fourth modules of the ligand binding domain (LA3,4) bind to MRAP domain 3 (MRAPD3). In 2FYL [14], two modules of complement-type ligand binding repeats (CR5,6) interact with two different sections of MRAP domain 1 (MRAPD1). Similarly, LA4,5 of 1N7D [33] bind to two different sites of LDLR beta propeller.

In the second mode, only one ligand binding module of LDLR binds to the ligand. Apo(H) and Apo(E) bind to A4 in 2KRI and CR17 in 2KNY, respectively. In 1V9U [34], the third LB module of VLDLR (V3) interacts with Human rhinovirus 2 (HRV2) viral proteins VP1.

As a whole, the available six structures describe 9 different binding sites, since three complexes operate in the first mode of interaction.

3D motif validation

Our 3D motif, displayed in Figure 4.A, is evaluated against predictions of 9 binding sites. Results are reported in Figure 4.B, where the number of ranks required to achieve 100% recall, $r_{100\%recall}$, is expressed as a function of the number of top quality predictions, *t*. A perfect prediction evaluation scheme would place the *t* best predictions on the *t* top-most positions of the ranked list, whereas the worst evaluation scheme would require the whole list to recall the *t* best predictions.

Although Cluspro developers do not recommend judging the produced models according to their associated cluster size, software output shows models ranked according to that score which obviously influence user's usage of these models. Therefore, we also show on Figure 4.B how cluster size would perform if used to rank models.

In every case, ranking based on 3D motif fitting produces curves closer to the perfect prediction than those generated from cluster size ranking. As a consequence fewer models are needed to recall the top quality predictions when the 3D motif is used to access LDLR interaction predictions. If the LDLR-Apo(E) (2KNY) complex is excluded, our 3D motif allows the discovery of the 4 best quality models within a shortlist of 15. Usage of the cluster score would require listing 53 models to achieve the same outcome. The different behaviour displayed by 2KNY could be explained by the fact that this model is not a true complex since the fragment of Apo(E) has been fused with a linker to CR17 to ensure interactions between both domains [11].

This experiment validates the usage of the LDLR 3D motif as a good indicator of model quality.

Literature study of LDLR-HNP1 complex

Since HNP1 has a hydrophobic and cationic face that resembles the binding patch of ligands which interact with LDLR [5, 35, 36] (Figure 5.A), its mode of interaction may be similar to those previously studied. In addition, this area belongs to a pocket detected by both Fpocket [37] and CastP software [38] (Figure 5.B).

Regarding the hydrophobic aspect, Ala-scanning mutational study of HNP1 revealed tryptophan26 (W26) is a key residue in direct interaction with target proteins and enables the peptide to form dimmers [39]. In addition, either W26 or phenylalanine28 (F28) mutation decreases HNP1 antibacterial activity. The importance of W26 is further highlighted by the fact it is either conserved or replaced by an amino acid displaying an aromatic side chain in other human α-defensins.

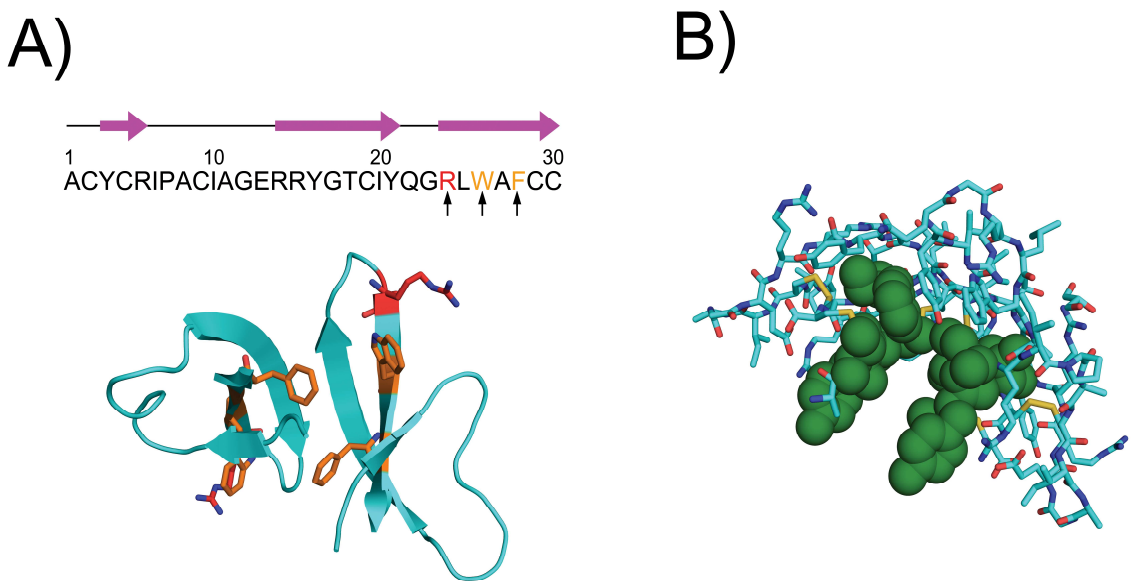

Figure **5**. (A) HNP1 sequence and 3D structure of the HNP1 dimer. The secondary structure of HNP1 is shown above the sequence. W26 and F28 are highlighted using arrows in the sequence and orange sticks in the 3D structure. R24 is also marked in red. (B) The pocket detected for HNP1 dimer using CastP software [32]. Images produced using Pymol.

As for the cationic face, HNP1 sequence comprises four basic residues, i.e. arginines, which could play a role similar to the lysines present in the studied LDLR-ligand complexes. Among these basic residues, arginine24 (R24) has been reported as an important residue for interacting with bacterial lipids [40].

Although beta sheets are dominant in HNP1 and LDLR structures, the study of known LDLR-ligand complexes does not support the involvement of beta sheets in their interactions. Actually, this study suggests formation of a salt bridge between HNP1's R24 and LDLR acidic residues and that either W26 or F28 plays the role of ψ in the minimal motif (Figure 3).

Docking prediction of LDLR-HNP1 complex

Cluspro produced a total of 43 predicted models using both the electrostatic and VWD/elec categories. Those models were ranked using our 3D motif and, as suggested by our previous experiment, only the top 15 are considered for further analysis (Table 3). Since R24 and either W26 or F28 are expected to be involved in the interaction, only Model.002.01, Model.006.02 and Model.006.18 are in agreement with literature findings.

Pairwise structural alignment reveals high similarity between Model.006.18 and Model.002.01 (1.61 Å RMSD). This shows that Cluspro converged towards a specific docking configuration from two different sets of constraints. Model.002.01 is chosen as representative of this configuration. In addition, as required by the minimal motif (Figure 3), Model.002.01 and Model.006.02 have candidates for the role of ψ since the TRP 144 of LA4 interacts with both W26:B and F28:A of HNP1 (Figure 6). Both models position their R24 N atoms at

similar locations (RMSD < 0.2Å). However, there is approximately a 90-degree angle between the positions of the ligands which leads to a 13.13 Å RMSD between those two putative complex configurations.

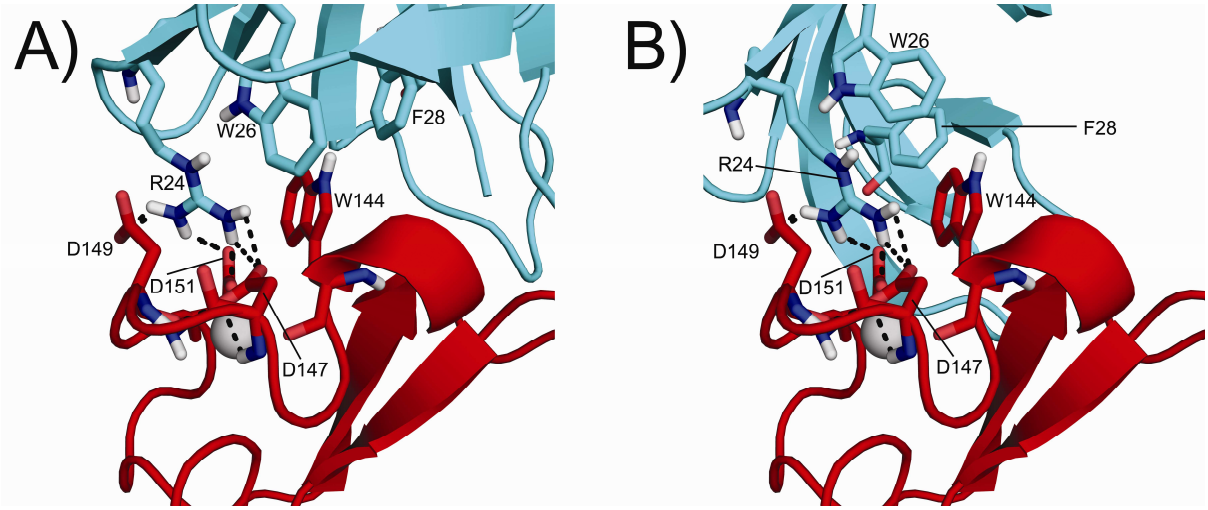

Figure **6**. Proposed LDLR-HNP1 interaction models for (A) model.002.01 and (B) model.006.02. Structures of HNP1 and A4 are shown in cyan and red, respectively. Calcium ion is represented as a grey sphere. R24 creates salt bridge with the aspartic residues which are shown as black dashed lines. W144 of A4 and F28 and W26 of HNP1 provide the hydrophobic interactions. Images produced using Pymol.

Table **3.** Residues involved in interaction between LDLR and HNP1 according to docking results. Model IDs starting by 002 and 006 are produced according to electrostatic and VDW+elec constraints, respectively. Contacts between residues are identified by SPACE [28].

Model ID	**ASP Residue(s) (LDLR)**	**ARG residue: Chain (HNP1)**	**Hydrophobic residue: Chain (HNP1)**
model.002.01	D147,149,151	R24:B	W26:B, F28:A
model.006.19	D149,151	R15:A	W26:A,I6:A,L25:A
model.006.02	D147,149,151	R24:B	W26:B,F28:A
model.006.05	D147,151	R14:B	I10:B
model.006.28	D147,151	R14:B	-
model.006.03	D147,149,151	R24:B	I6:B
model.006.12	D147,151	R15:A	-
model.006.23	D149	R15:B	I20:A
model.002.15	D147,149	R14:A	W26 :B,F28:A
model.006.17	D149,151	R14:B	-
model.006.06	D149,151	R24:B	I6:B
model.006.18	D149,151	R24:B	W26:B,F28:A
model.006.00	D147,149	R14:B	-
model.006.13	D149,151	R14:A	-
model.006.22	D147,149,151	R15:A	-

Complex stability analysis based on FoldX binding energy calculations (Figure 7) reveals that Model.002.01 is a much more stable LDLR-HNP1 complex than Model.006.02. Although Cluspro energy values (-712.5 and -143.3 Kcal/mol for Model.002.01 and Model.006.02 respectively) are not particularly accurate [41], they are in agreement with FoldX conclusions. In addition, the fact that Cluspro simulations based on two different sets of constraints led to the configuration exemplified by Model.002.01 supports the presumption of its higher stability. It is interesting to notice that, for this model, the strength of the LDLR-HNP1 bonds weakens the bond between the two HNP1 monomers (Figure 7).

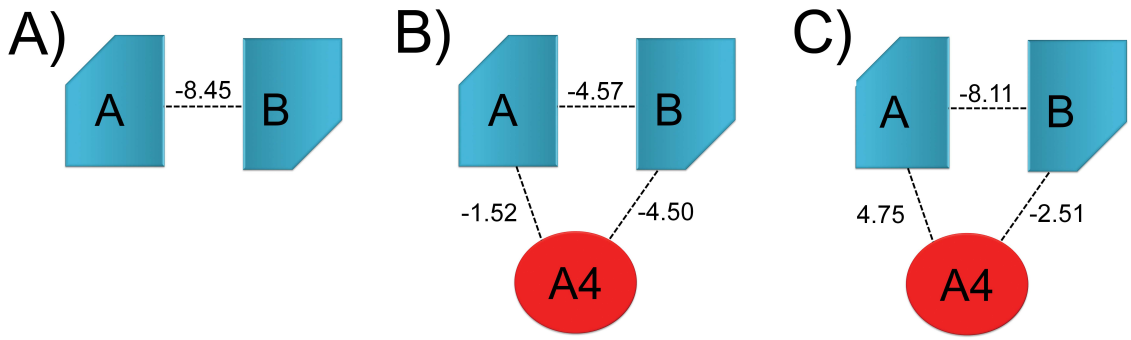

Figure **7**. Complex stability expressed by interaction energy estimated by FoldX for the structures. HNP1 dimer and A4 are shown by rectangle and circle, respectively. (A) HNP1 dimer (PDB Code: 3GNY), (B) Model.002.01, (C) Model.006.02. Energies are in Kcal/mol.

DISCUSSION

The major objective of this investigation was to establish whether a LDLR-HNP1 interaction can occur based upon computational models. Previous reports of i) the versatility of ligand recognition exhibited by the LDLR family [6], ii) an interaction between human α-defensins with LDLR [1,5], and iii) its role in internalising ligands (such as cholesterol and amyloid-beta [1,3,43] led to examination of the putative interaction.

The study relating to HNP1 dimer formation, conducted using SPACE [28], revealed that the structure contains 33 intermolecular contacts including 3 hydrogen bonds (see Table S1 and Table S2). The dimer binding energy calculation of -8.45 Kcal/mol (Figure 7.A) is commensurate with several models where stable interactions occur [30, 41].

The major observation from the modelling is that interactions between the different chains of Model002.01 are very strong, -10.59 Kcal/mol as a whole (Figure 7.B). For model.006.02, a very different scenario is depicted where the energy saving for interaction with the dimer is greatly diminished where the interaction with one monomer requires 4.75 Kcal/mol (Figure 7.A). This thermodynamically unfavourable scenario points to Model.002.01 as preferential.

The strength of binding seen in Model.002.01 is reflected in the levels of intra-molecular interactions. In addition to the contacts present in the dimer, binding to the receptor generates a further 48 contacts including 8 as hydrogen bonds and 3 as electrostatic interactions, (see Table S3 and Table S4).

Within the complex, an intriguing feature is the modulation of dimer interaction energies depending on which model is studied. A considerably weaker level of dimer binding strength is observed for Model.002.01 which may have ramifications for internalisation should this step proceed through the monomer form. In contrast, for Model.006.02, the binding interaction for the dimer remains strong.

One aim of this study involved identification of the receptor binding mechanism for the purposes of informing the future design of synthetic HNPs to afford maximum internalisation. This report highlights the key putative contacts between HNP1 and the LDLR, and moreover, emphasises the potential importance of maintaining the HNP1 dimer form for binding and potentially for internalisation. Further computational studies are required to clarify the mechanism of internalisation and interaction with membrane [42, 43].

These insights, from computation study based drug design, provide a number of avenues towards novel synthetic antimicrobial peptides which can be synthesised and tested through conventional assays. Strengthening or weakening LDLR-HNP interactions may have synergistic or dysergistic effects on the two key aspects, namely docking and internalisation. In this vein, strengthening the links that make the HNP1 dimer, even to the extent of forming fixed permanent bond to anchor the dimer link, may be an avenue to greater efficacy in some forms of antimicrobial activity.

LIST OF ABBREVIATIONS

LDLR: Low Density Lipoprotein Receptor

HNP1: Human Neutrophil Peptide 1

HIV: Human Immunodeficiency Virus

AVAILABILITY

3D motif and LDLR-HNP1 models are available upon request from the authors.

SUPPLEMENTARY MATERIAL

Table S1: Contacts between HNP1's monomers.

Table S2: Statistics of contacts in HNP's monomers.

Table S3: Contacts between A4 and HNP1's dimer in Model.006.02 and Model 002.01.

Table S4: Statistics of contacts in Model.006.02 and Model 002.01.

ACKNOWLEDGEMENTS

This work was in part supported by grant 6435/B/T02/2011/40 of the Polish National Centre for Science.

REFERENCES

[1] Nassar, T.; Akkawi, S.; Bar-Shavit, R.; Haj-Yehia, A.; Bdeir, K.; Al-Mehdi, A.B.; Tarshis, M.; Higazi, A.A. Human alpha-defensin regulates smooth muscle cell contraction: a role for low-density lipoprotein receptor-related protein/alpha 2-macroglobulin receptor. *Blood*, **2002**, *100*(12), 4026–4032.

[2] Chang, T.L.; Vargas, J. Jr.; DelPortillo, A.; Klotman, ME. Dual role of alpha-defensin-1 in anti-HIV-1 innate immunity. *J. Clin. Invest.*, **2005**, *115*(3), 765-773.

[3] Fuentealba, R.A.; Liu, Q.; Zhang, J.; Kanekiyo, T.; Hu, X.; Lee, J.M.; LaDu, M.J.; Bu, G. Low-density lipoprotein receptor-related protein 1 (LRP1) mediates neuronal Abeta42 uptake and lysosomal trafficking. *PLoS. One.*, **2010**, *5*(7), e11884.

[4] Nakashima, H.; Yamamoto, N.; Masuda, M.; Fujii, N. Defensins inhibit HIV replication in vitro. *AIDS*, **1993**, *7*(8), 1129.

[5] Higazi, A.A.; Nassar, T.; Ganz, T.; Rader, D.J.; Udassin, R.; Bdeir, K.; Hiss, E.; Sachais, B.S.; Williams, K.J.; Leitersdorf E.; Cines D.B. The alpha-defensins stimulate proteoglycan-dependent catabolism of low-density lipoprotein by vascular cells: a new class of inflammatory apolipoprotein and a possible contributor to atherogenesis. *Blood*, **2000**, *96*(4), 1393–1398.

[6] Blacklow, S.C. Versatility in ligand recognition by LDL receptor family proteins: advances and frontiers. *Curr. Opin. Struct. Biol.*, **2007**, *17*(4), 419–426.

[7] Guttman, M.; Prieto, J.H.; Croy, J.E.; Komives, E.A. Decoding of Lipoprotein−Receptor Interactions: Properties of Ligand Binding Modules Governing Interactions with Apolipoprotein E. *Biochemistry*, **2010**, *49*(6), 1207–1216.

[8] Fisher, C.; Beglova, N.; Blacklow, S.C. Structure of an LDLR-RAP complex reveals a general mode for ligand recognition by lipoprotein receptors. *Mol. Cell*, **2006**, *22*(2), 277–283.

[9] Beglov, D.; Lee, C.J.; De Biasio, A.; Kozakov, D.; Brenke, R.; Vajda, S.; Beglova, N. Structural insights into recognition of beta2-glycoprotein I by the lipoprotein receptors. *Proteins*: *Struct., Funct., Bioinf.*, **2009**, *77*(4), 940-949.

[10] Herz, J.; Bock, H.H. Lipoprotein receptors in the nervous system. *Annu. Rev. Biochem.*, **2002**, *71*, 405–434.

[11] Guttman, M.; Prieto, J.H.; Handel, T.M.; Domaille, P.J.; Komives, E.A. Structure of the Minimal Interface Between ApoE and LRP. *J. Mol. Biol.*, **2010**, *398*(2), 306–319.

[12] Rudenko, G.; Deisenhofer, J. The low-density lipoprotein receptor: ligands, debates and lore. *Curr. Opin. Struct. Biol.*, **2003**, *13*(6), 683–689.

[13] Dirlam-Schatz, K.A.; Attie, A.D. Calcium induces a conformational change in the ligand binding domain of the low density lipoprotein receptor. *J. Lipid Res.*, **1998**, *39*(2), 402-411.

[14] Jensen, G.A.; Andersen, O.M.; Bonvin, A.M.; Bjerrum-Bohr, I.; Etzerodt, M.; Thogersen, H.C.; O'Shea, C.; Poulsen, F.M.; Kragelund, B.B. Binding site structure of one LRP-RAP complex: implications for a common ligand-receptor binding motif. *J. Mol. Biol.*, **2006** , *362*(4), 700–716.

[15] Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.*, **2000**, *28*, 235-242.

[16] Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **1997**, *25*(17), 3389-3402.

[17] Suzek, B.E.; Huang, H.; McGarvey, P.; Mazumder, R.; Wu, C.H. UniRef: Comprehensive and Non-Redundant UniProt Reference Clusters. *Bioinformatics*, **2007**, *23*(10), 1282-1288.

[18] Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T.J.; Higgins, D.G.; Thompson, J.D. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.*, **2003**, *31*(13), 3497-3500.

[19] The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC. http://pymol.org/.

[20] Wei, G.; de Leeuw, E.; Pazgier, M.; Yuan, W.; Zou, G.; Wang, J.; Ericksen, B.; Lu, W.Y.; Lehrer, R.I.; Lu, W. Through the looking glass: mechanistic insights from enantiomeric human defensins. *J. Biol. Chem.*, **2009**, *284*(42), 29180-29192.

[21] Nebel, J.C. Generation of 3D templates of active sites of proteins with rigid prosthetic groups. *Bioinformatics*, **2006**, *22*(10), 1183-1189.

[22] Nebel, J.C.; Herzyk, P.; Gilbert, D.R. Automatic generation of 3D motifs for classification of protein binding sites. *BMC Bioinf.*, **2007**, *8*, 321.

[23] Comeau, S.R.; Gatchell, D.W.; Vajda, S.; Camacho, C.J. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, **2004**, *20*(1), 45-50.

[24] Kozakov, D.; Hall, D.R.; Beglov, D.; Brenke, R.; Comeau, S.R.; Shen, Y.; Li, K.; Zheng, J.; Vakili, P.; Paschalidis, I.Ch.; Vajda, S. Achieving reliability and high accuracy in automated protein docking: Cluspro, PIPER, SDU, and stability analysis in CAPRI rounds 13–19. *Proteins*: *Struct., Funct., Bioinf.*, **2010**, *78*(15), 3124–3130.

[25] Comeau, S.R.; Kozakov, D.; Brenke, R.; Shen, Y.; Beglov, D.; Vajda, S. ClusPro: performance in CAPRI rounds 6–11 and the new server. *Proteins*: *Struct., Funct., Bioinf.*, **2007**, *69*(4), 781–785.

[26] Shen, Y.; Brenke, R.; Kozakov, D.; Comeau, S.R.; Beglov, D.; Vajda, S. Docking with PIPER and refinement with SDU in rounds 6–11 of CAPRI. *Proteins*: *Struct., Funct., Bioinf.*, **2007**, 69(4), 734–742.

[27] Kozakov, D.; Brenke, R.; Comeau, S.R.; Vajda, S. PIPER: an FFT-based protein docking program with pairwise potentials. *Proteins*: *Struct., Funct., Bioinf.*, **2006**, *65*(2), 392–406.

[28] Sobolev, V.; Eyal, E.; Gerzon, S.; Potapov, V.; Babor, M.; Prilusky, J.; Edelman, M. SPACE: a suite of tools for protein structure prediction and analysis based on complementarity and environment. Nucleic Acids Res., **2005**, *33*, W39-43.

[29] Sánchez, I.E.; Beltrao, P.; Stricher, F.; Schymkowitz, J.; Ferkinghoff-Borg, J.; Rousseau, F.; Serrano, L. Genome-wide prediction of SH2 domain targets using structural information and the FoldX algorithm. *PLoS Comput. Biol.*, **2008**, *4*(4), e1000052.

[30] Kiel, C.; Wohlgemuth, S.; Rousseau, F.; Schymkowitz, J.; Ferkinghoff-Borg, J.; Wittinghofer, F.; Serrano, L. Recognizing and defining true Ras binding domains II: in silico prediction based on homology modelling and energy calculations. *J. Mol. Biol.*, **2005**, *348*(3), 759-775.

[31] Guerois, R.; Nielsen, J.E.; Serrano, L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.*, **2002**, *320*(2), 369-387.

[32] Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L. The FoldX web server: an online force field. *Nucleic Acids Res.*, **2005**, *33*, W382-388.

[33] Rudenko, G.; Henry, L.; Henderson, K.; Ichtchenko, K.; Brown, M.S.; Goldstein, J.L.; Deisenhofer, J. Structure of the LDL receptor extracellular domain at endosomal pH. *Science*, **2002**, *298*(5602), 2353-2358.

[34] Verdaguer, N.; Fita, I.; Reithmayer, M.; Moser, R.; Blaas, D. X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. *Nat. Struct. Mol. Biol*., **2004**, *11*(5), 429-434.

[35] Quinn, K.; Henriques, M.; Parker, T.; Slutsky, A.S.; Zhang, H. Human neutrophil peptides: a novel potential mediator of inflammatory cardiovascular disease. *Am. J. Physiol. Heart Circ. Physiol*., **2008**, *295*(5), 1817-1824.

[36] Soman, S.S.; Sivakumar, K.C.; Sreekumar, E. Molecular dynamics simulation studies and in vitro site directed mutagenesis of avian beta-defensin Apl_AvBD2. *BMC Bioinf*., **2010**, *11*(Suppl 1), S7.

[37] Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinf.*, **2009**, *10*, 168.

[38] Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res*., **2006**, *34*, W116-W118.

[39] Wei, G.; Pazgier, M.; de Leeuw, E.; Rajabi, M.; Li J.; Zou, G.; Jung, G.; Yuan, W.; Lu, W.Y.; Lehrer, R.I.; Lu, W.Trp-26 imparts functional versatility to human alpha-defensin HNP1. *J. Biol. Chem*., **2010**, *285*(21), 16275-16285.

[40] Zhang, Y.; Lu, W.; Hong, M. The membrane-bound structure and topology of a human α-defensin indicate a dimer pore mechanism for membrane disruption. *Biochemistry*, **2010**, *49*(45), 9770-9782.

[41] Ponomarev, S.Y.; Audie, J. Computational prediction and analysis of the DR6-NAPP interaction. *Proteins*: *Struct*., *Funct*., *Bioinf*., **2011**, *79*(5), 1376-1395.

[42] Fleming, E.; Maharaj, N.P.; Chen, J.L.; Nelson, R.B.; Elmore, D.E. Effect of lipid composition on buforin II structure and membrane entry. *Proteins*: *Struct*., *Funct*., *Bioinf*., **2008**, *73*(2), 480-491.

[43] Lazaridis, T. Implicit solvent simulations of peptide interactions with anionic lipid membranes. *Proteins*: *Struct*., *Funct*., *Bioinf*., **2005**, *58*(3), 518-527.

FIGURE LEGENDS

Figure 1. Modular structure of LDLR receptor family: general domain pattern (top) and schematic representation of the LDLR-ligand binding modes of known complexes (bottom).

Figure 2. Multiple sequence and structure alignment of ligand binding domains of LDLR family complexes. In the sequence alignment, residues involved in calcium interaction are denoted by dots. The three conserved acidic residues and conserved tryptophan/phenylalanine are highlighted with black arrows. Sequence numbering is based on 2KRI:B. In the structure alignment the three conserved acidic residues and conserved tryptophan are shown on 2KRI structure. Residues numbering is based on 2KRI:B. The ligand binding domains associated with each colour are: 2KRI-A4: red, 2FCW-A3:green, 2FCW-A4:dark blue, 2FYL-CR5:purple, 2FYL-CR6:yellow, 2KNY-CR17:orange, 1N7D-A4:deep teal, 1N7D-A5: gray, 1V9U-V3: pink.

Figure 3. Minimal binding motif defined by Jensen et al.[14].

Figure 4. (A) The 3D motif is represented by spheres. The blue ones show positions of N atoms from the ligand. The black ones are the C-alpha atoms of the ASP and TRP and the red sphere is the O atom of TRP. Location of the calcium is marked by a grey sphere. Image produced using Pymol. (B) Number of ranks to achieve 100% recall of the top predictions (or recall of top predictions with top positions). In the legend the complexes names are followed by C and M for curves based on cluster size and 3D motif method, respectively.

Figure 5. (A) HNP1 sequence and 3D structure of the HNP1 dimer. The secondary structure of HNP1 is shown above the sequence. W26 and F28 are highlighted using arrows in the sequence and orange sticks in the 3D structure. R24 is also marked in red. (B) The pocket detected for HNP1 dimer using CastP software [32]. Images produced using Pymol.

Figure 6. Proposed LDLR-HNP1 interaction models for (A) model.002.01 and (B) model.006.02. Structures of HNP1 and A4 are shown in cyan and red, respectively. Calcium ion is represented as a grey sphere. R24 creates salt bridge with the aspartic residues which are shown as black dashed lines. W144 of A4 and F28 and W26 of HNP1 provide the hydrophobic interactions. Images produced using Pymol.

Figure 7. Complex stability expressed by interaction energy estimated by FoldX for the structures. HNP1 dimer and A4 are shown by rectangle and circle, respectively. (A) HNP1 dimer (PDB Code: 3GNY), (B) Model.002.01, (C) Model.006.02. Energies are in Kcal/mol.