Supplementary Information online for:

Mechanistic implications for receptor degradation from the PCSK9/LDLR complex structure at neutral pH

Paola Lo Surdo[1,&,†], Matthew J. Bottomley[1,&,†], Alessandra Calzetta[1], Ethan C. Settembre[2,#], Agostino Cirillo[1,^], Shilpa Pandit[3], Yan G. Ni[3], Brian Hubbard[3], Ayesha Sitlani[3] & Andrea Carfi[1,#,*]

From the [1]Department of Biochemistry and Molecular Biology, IRBM *P. Angeletti*, Via Pontina Km 30.600, I-00040 Pomezia (Rome), Italy; [2]Children's Hospital, Enders Building, 320 Longwood Avenue, 02132 Boston, MA, USA; and [3]Division of Cardiovascular Diseases, Merck Research Laboratories, 126 E. Lincoln Avenue, 07065 Rahway, NJ, USA

Current author addresses are: [&]Novartis Vaccines & Diagnostics, Via Fiorentina 1, 53100 Siena, Italy; [^]Novartis Institute for Biomedical Research, Basel, Switzerland and [#]Novartis Vaccines & Diagnostics, 45 Sidney St, Cambridge, MA 02139, USA

[†] These authors contributed equally to this work.

* To whom correspondence should be addressed:
e-mail: andrea.carfi@novartis.com
Tel.: (617)-871-8363

Supplementary Methods

Supplementary References

Supplementary Figure legends

Supplementary Table

Supplementary Figures

Supplementary Methods

Purification of recombinant PCSK9 and LDLR proteins

Full-length human PCSK9 (residues M1-Q692) was cloned in a pcDNA3.1/V5 His6 vector (Invitrogen) and expressed in stably transfected HEK293 cells. The secreted PCSK9 was purified by immobilized metal affinity chromatography (IMAC; using Talon resin, Clontech) as described previously (Fisher et al, 2007), followed by ion-exchange chromatography on a RESOURCE Q column (GE Healthcare). Peak fractions were pooled and loaded on a Superdex 200 size-exclusion chromatography column (GE Healthcare) equilibrated in buffer containing 25mM HEPES pH 7.9, 150mM NaCl, 0.1mM $CaCl_2$ and 10% glycerol. The purified PCSK9 was concentrated to ~ 10 mg/ml and stored at -80°C.

A PCSK9 protein spanning residues Q31-A451 but lacking the C-terminal domain (CTD) residues G452–Q692, and therefore named PCSK9ΔC, was expressed in *E. coli* and purified as described previously (Bottomley et al, 2009). The purified PCSK9ΔC protein is fully auto-processed, thus representing the mature form of PCSK9 lacking the CTD.

The full-length LDLR ectodomain (residues A1-A699) with the LDLR signal peptide sequence at the N-terminus was cloned into a pcDNA-DEST40/V5-His6 Gateway vector (Invitrogen). Two potential N-linked glycosylation sites, N494 and N636, were mutated to Gln, to aid crystallization without affecting LDL-binding capacity (Rudenko et al, 2002). Shorter expression constructs with C-terminal His tags were prepared by sub-cloning regions of the full-length LDLR plasmid to make: L5-L6-L7-EGFPH, L7-EGFPH and EGFPH (as shown in the main text, Fig. 1A). Point mutations of LDLR and LDLR-fragment constructs were confirmed by DNA sequencing. All LDLR proteins were expressed either as stably transfected HEK293 cell lines or by transient transfection of HEK293-EBNA cells. Full-length ectodomain and truncated versions of LDLR were purified as described for PCSK9, with the final gel filtration buffer containing 25mM

Hepes pH 8.0, 150mM NaCl and 1mM $CaCl_2$. The purified LDLR proteins were concentrated to ~ 2 mg/ml and stored at -80°C.

For complex preparations, purified PCSK9 and the LDLR proteins were mixed in a 1:1 molar ratio and incubated for 30 minutes on ice. For each complex any excess of either component was removed by size-exclusion chromatography on a Superdex 200 (10/300) column equilibrated in buffer containing 25mM Hepes pH 7.0, 150mM NaCl and 0.5mM $CaCl_2$. Protein complexes were concentrated to 5mg/mL.

Crystallization of PCSK9/LDLR complexes

Protein complexes were crystallized by the vapour diffusion method at 18°C. Crystals of the full-length LDLR/PCSK9 complex were obtained against a reservoir containing 0.1M imidazole and 0.7M sodium acetate, pH 7.0, allowing structure determination and refinement to 4.2Å resolution (Supplementary Table S1 online). Mass spectrometry, SDS-PAGE analysis and N-terminal sequencing of washed crystals indicated that adventitious proteolysis of the LDLR had occurred during crystallization between repeats L4 and L5 ($G_{171}{}^{\downarrow}D_{172}$), for ~50% of receptor molecules, such that the crystals contained a mixture of full-length receptor and a shorter form lacking L1 to L4. Purification and crystallization in the presence of protease inhibitors was sufficient to prevent receptor degradation but resulting crystals did not exhibit improved diffraction.

Crystallization trials were also performed using alternative complexes of PCSK9 and N-terminally truncated LDLR proteins (Fig. 1). Crystals were obtained for PCSK9 bound to L5-L6-L7-EGFPH or L7-EGFPH in the same conditions as for the full-length PCSK9/LDLR complex, but these crystals did not diffract beyond 10Å resolution. Crystals of the PCSK9/EGFPH complex were obtained in 0.1M sodium cacodylate pH 7.0, 1.7M sodium acetate, 300mM NaCl, 5% ethanol, allowing structure determination and refinement to 3.3Å resolution (Supplementary Table S1 and Supplementary Fig. 1).

Data collection, structure determination and refinement of PCSK9/LDLR complexes

X-ray diffraction data were collected at 100K on the ID14 beam lines at the ESRF (Grenoble, France). Diffraction data were integrated with MOSFLM (Leslie, 1992) and the CCP4 suite of programs (CCP4, 1994) was used for data processing and analysis (Supplementary Table S1). Crystals of the full-length PCSK9/LDLR complex belonged to the $P6_5$ space group and contained one protein complex and ~ 80% solvent in the asymmetric unit (AU). The crystals of PCSK9 bound to the LDLR EGFPH fragment belonged to the $P2_12_12_1$ space group, the AU contained one protein complex and ~ 60% solvent (Supplementary Table S1).

The structures were determined by the molecular replacement method with PHASER software (McCoy et al, 2005) using as search models the structures of PCSK9 (PDB code 2QTW), the LDLR β-propeller-EGF(C) fragment (1IJQ), the L7 repeat (1XFE) and the EGF(A) domain (3BPS, 2W2N). After rigid body refinement, 2|Fo|-|Fc| electron density maps were calculated and revealed the EGF(B) domain. Due to the higher resolution of the diffraction data collected for the PCSK9/EGFPH complex, model building and refinement were performed first for this structure, using Coot (Emsley & Cowtan, 2004) and REFMAC (Murshudov et al, 1997). The resulting model was then used to further refine the 4.2Å resolution LDLR/PCSK9 structure. Towards the end of the refinement the L7 domain from the low pH LDLR structure (1N7D) was fitted as a rigid body into 2|Fo|-|Fc| electron density maps. No electron density was observed for the L1-L6 region of the receptor despite the presence of both L5 and L6 in the crystals (as confirmed by mass spectrometry). Thus, the L6 and L5, and presumably also L1 to L4, are likely flexible in the crystals. Atomic coordinates have been deposited in the PDB with accession codes 3P5B (3.3Å resolution structure) and 3P5C (4.2Å resolution structure). Figures were prepared using Pymol (www.pymol.org).

Furin digestion experiments

Furin digestion experiments were performed for 18 hrs at 23°C in buffer containing 100mM Hepes pH 7.5, 150mM NaCl, 3mM $CaCl_2$, and 0.5% Triton X-100. Furin (New England Biolabs)

was added in a ratio of 7 units enzyme per 12.5μg of protein sample. The minimal furin cleavage site, Arg-X-X-Arg$^{\downarrow}$, is present in PCSK9 at residues $R_{215}FHR^{\downarrow}Q_{219}$. The samples were analysed using 4-12% SDS-PAGE gels with Coomassie staining.

Surface plasmon resonance (SPR) experiments

All SPR experiments were performed using Biacore instruments equilibrated at 25°C, essentially as described previously (Fisher et al, 2007). In all cases, the LDLR proteins were immobilized on the sensor surface and PCSK9 was injected in optimized running buffer. For each titration, the LDLR fragment protein was first covalently immobilized by amine-coupling on a carboxymethylated dextran sensor chip (CM-5, GE Healthcare). The reciprocal set-up was not viable since PCSK9 retained very little binding activity when covalently immobilized on the sensor chip. Amine-coupling reactions for immobilization of the LDLR fragments were performed using purified protein at ~ 2μg/mL in 10mM sodium acetate buffer pH 4.5 until ~ 100 response units (RU) were captured.

Titration experiments were performed by injecting at 50μL/min the purified full-length wild type PCSK9 in optimized running buffer (filtered and degassed) containing 150mM NaCl, 2mM $CaCl_2$, 0.0025% P-20 detergent and 0.05% BSA (added to minimize non-specific interactions), with buffering by either 25mM Hepes pH 7.5 or 25mM sodium acetate pH 5.5. Following each injection, sensor chip surfaces were regenerated twice, with a 5 second injection of 10mM NaOH at pH 11, followed by a 5 second injection of 10mM HCl at pH 2. Each titration series contained 8-12 analyte injections, including duplicates, with maximum used PCSK9 concentrations of 7.5μM at neutral pH and 0.5μM at acidic pH.

The Biacore data were analysed using the *BIAevaluation* software version 4.1, provided by the manufacturer specifically to accompany the Biacore instruments. In order to yield curves representing specific binding of the injected analyte to the immobilized ligand, an injection of buffer only was subtracted from each curve, and reference sensorgrams (obtained by simultaneously

injecting PCSK9 over a blank sensor chip surface, i.e. a surface lacking any LDL receptor) were subtracted from experimental sensorgrams. Each injection of the titration series was performed for 120 seconds, allowing the binding response (RU, resonance units) to reach equilibrium for the large majority of cases. At the injection end point, the 'steady-state' value was reached, where the RU measured represented equal rates of association and dissociation (R_{eq}). Next, following the manufacturer's instructions, the steady-state analysis approach was used to plot equilibrium binding response (R_{eq}) against analyte (PCSK9) concentration in order to obtain the dissociation constants (K_D) for the interaction of PCSK9 (analyte) with the LDLR (ligand). The *BIAevaluation* software includes a steady-state affinity model that allows calculation of binding affinity from such a plot of the steady-state data R_{eq} vs Conc. The software uses a non-linear least-squares minimization routine in order to fit the experimental data to a curve described by the equation:

$$R_{eq} = K_A \times Conc \times R_{max} / (K_A \times Conc \times n+1)$$

where K_A is the equilibrium association constant (and $K_A = 1/K_D$), Conc is the molar analyte (PCSK9) concentration, R_{max} is the maximum analyte binding capacity of the ligand-loaded sensor chip, and *n* is the stoichiometry (n=1 for the PCSK9/LDLR interaction). The standard deviation reported for the K_D value determined is the root mean square deviation for all data points on the curve.

The BIAevaluation software provided by the manufacturer also allows for extraction of the kinetic parameters of association rate constant (k_{on}) and dissociation rate constant (k_{off}) from a titration series of injections, thus permitting an alternative route for determination of the affinity constant K_D (since $K_D = k_{off} / k_{on}$). This approach is an alternative to the steady-state approach, and is routinely implemented in SPR analyses. The kinetic approach was implemented for analysis of the PCSK9/LDLR-L626E mutant interaction at acidic pH (as noted in Table 1 by the annotation #), because the experimental data were unsuitable for steady-state analysis due to steady-state not being reached for most injections of that particular titration series (see Figure S3, panel 10). First, we validated this kinetic approach by analyzing the PCSK9/LDLR (wild type) interaction, where there

was essentially no difference between the K_D values determined via the two different approaches (see Table 1 and Figure S3, panels 9a and 9b), confirming that both approaches are compatible, give similar results and are valid under the conditions employed herein.

Briefly, in the kinetic analysis, the sensorgram curves for each analyte concentration injected are analyzed by the *BIAevaluation* software through a non-linear least-squares minimization routine in order to fit simultaneously the experimental data of association and dissociation to a 1:1 interaction model. For example, where injected analyte A binds to immobilized ligand B to form a complex AB (the Langmuir model), association is described by the term:

$$d[AB]/dt = k_{on} \times [A] \times [B];$$

and dissociation by the term:

$$-d[AB]/dt = k_{off} \times [AB];$$

such that the net rate equation is given by:

$$d[AB]/dt = k_{on} \times [A] \times [B] - k_{off} \times [AB];$$

where A is the molar analyte concentration in solution (maintained constant by the injection flow system); AB is the concentration of the complex which is measured directly as R, the SPR response in response units (RU); and B is the ligand on the surface for which the total concentration can be expressed in RU as the maximum binding capacity R_{max}, such that the concentration of free B is $R_{max} - R$. Thus, the net rate equation can be rewritten as follows:

$$dR/dt = k_{on} \times [A] \times [R_{max} - R] - k_{off} \times R.$$

Analysis of the experimental SPR data using this Langmuir model allows determination of the association and dissociation rate constants k_{on} ($M^{-1}.s^{-1}$) and k_{off} (s^{-1}), which can then be used to derive the affinity constant K_D (with a standard deviation reflecting the root mean square deviation upon fitting of the model curves to the experimental curves), as reported here in Table 1 in the main text. Further details of analysis via the steady-state and kinetic models can be found in the literature (Karlsson, 1999) and in the Biacore users handbook (Biacore T100 Software Handbook, GE Healthcare).

Supplementary References

Bottomley MJ, Cirillo A, Orsatti L, Ruggeri L, Fisher TS, Santoro JC, Cummings RT, Cubbon RM, Lo Surdo P, Calzetta A, Noto A, Baysarowich J, Mattu M, Talamo F, De Francesco R, Sparrow CP, Sitlani A, Carfi A (2009) Structural and biochemical characterization of the wild type PCSK9-EGF(AB) complex and natural familial hypercholesterolemia mutants. *J Biol Chem* **284:** 1313-1323

CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50:** 760-763

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126-2132

Fisher TS, Lo Surdo P, Pandit S, Mattu M, Santoro JC, Wisniewski D, Cummings RT, Calzetta A, Cubbon RM, Fischer PA, Tarachandani A, De Francesco R, Wright SD, Sparrow CP, Carfi A, Sitlani A (2007) PCSK9-dependent LDL receptor regulation: Effects of pH and LDL. *J Biol Chem* **282:** 20502-20512

Karlsson R (1999) Affinity analysis of non-steady-state data obtained under mass transport limited conditions using BIAcore technology. *J Mol Recognit* **12:** 285-292

Leslie AGW (1992) Recent changes to teh MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **26**

McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61:** 458-464

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53:** 240-255

Rudenko G, Henry L, Henderson K, Ichtchenko K, Brown MS, Goldstein JL, Deisenhofer J (2002) Structure of the LDL receptor extracellular domain at endosomal pH. *Science* **298:** 2353-2358

Supplementary Figure legends

Supplementary Figure S1: Electron density map.

A sigma weighted 2|Fo|-|Fc| electron density map drawn at 1.3 σ from the PCSK9/EGFPH structure refined to 3.3Å resolution (PDB accession code 3P5B) in proximity of the LDLR linker-β-propeller interface. Backbone and side chain sticks are shown for the linker (green) and the β-propeller (cyan) regions of the LDLR.

Supplementary Figure S2: LDLR protects PCSK9 from furin protease activity

An SDS-PAGE analysis of the susceptibility to furin cleavage of (i) PCSK9 alone and (ii) a purified equimolar complex of PCSK9/LDLR proteins reveal that the presence of LDLR is sufficient to protect PCSK9 from furin cleavage. Presence or absence of furin is indicated by + or – above each lane of the gel. Key: lanes 2-3 show PCSK9 alone without/with furin treatment; lanes 4-5 show the PCSK9/LDLR complex without/with furin treatment. The bands are labeled as follows, CCTD: PCSK9 153-692; CCTD1: PCSK9 218-692 the larger product after furin-cleavage; Pro: PCSK9 prodomain. The gel is representative of several repeated experiments, typically performed at concentrations of PCSK9 and LDLR 10-fold greater than the K_D of the complex.

Supplementary Figure S3: Surface plasmon resonance data

Experimental SPR data is shown for PCSK9 or PCSK9ΔC (the injected analytes) binding to various LDLR receptor proteins (the immobilized ligands) under neutral or acidic pH buffer conditions. The left panels show the experimental sensorgram data, while the right panels show the resulting plots of 'R_{eq} vs PCSK9 concentration' for each titration, where R_{eq} is the SPR response in response units (RU) at equilibrium. For the data in panels 9-10 the curves were fit using a 1:1 binding model in order to derive K_D from kinetic analysis; the steady-state analysis could not be used in this case because steady-state was not reached for several

of the of injections. Comparison of panels 9a and 9b shows that the steady-state and kinetic analyses yield comparable results, as discussed in detail in the Supplementary Methods.

Supplementary Figure S4: SDS-PAGE analyses of proteins studied

SDS-PAGE (4-12% gradient Bis-Tris buffered gel) analysis of the LDLR and L7 proteins expressed and purified for analysis by SPR.

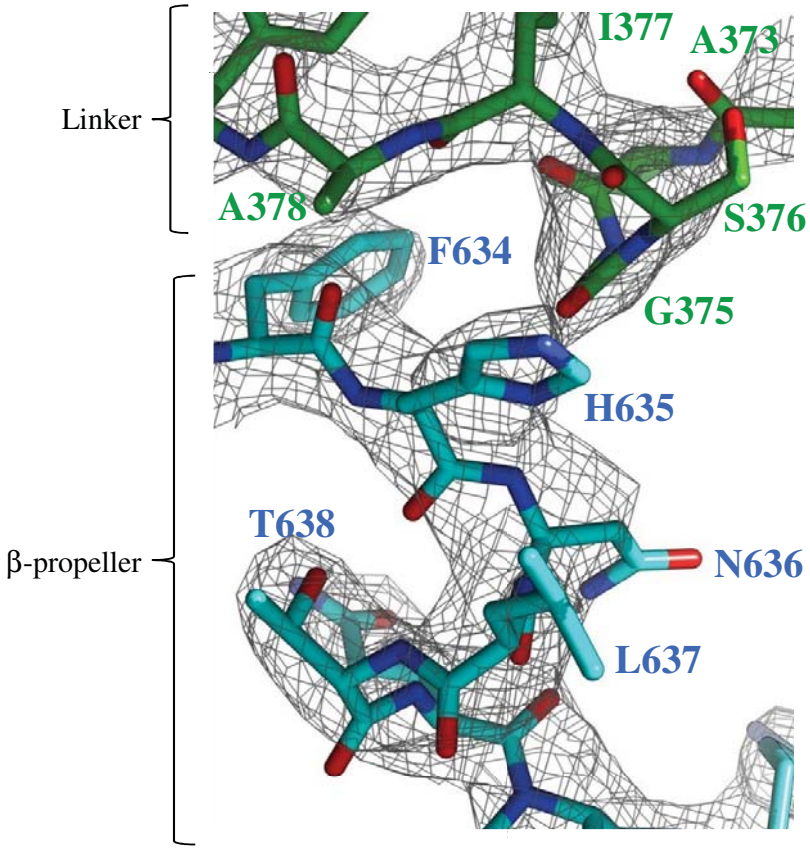
Supplementary Table S1 (Lo Surdo et al.)

Data collection and refinement statistics

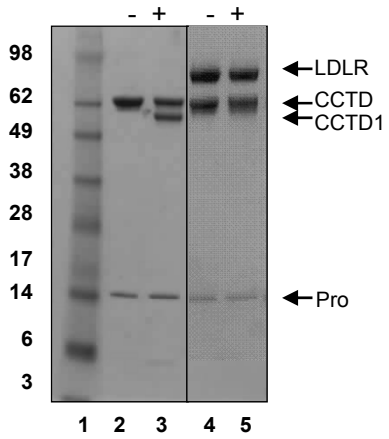
	PCSK9 + FL LDLR ectodomain	**PCSK9 + EGFPH**
Space Group (angles)	$P6_5$ (90, 90, 120)	$P2_12_12_1$ (90, 90, 90)
a, *b*, *c* (Å)	320.9 x 320.9 x 77.1	77.0 x 109.7 x 178.5
Solvent content (%)	78	60
Resolution range (Å)	60.0-4.20 (4.43-4.20)†	50.0-3.30 (3.48-3.30)
Reflections: total	108345 (15165)	78131 (11518)
Reflections: unique	33296 (4783)	23117 (3346)
Completeness (%)	99.1 (98.8)	98.8 (99.4)
Redundancy	3.3 (3.2)	3.4 (3.4)
<Mean (I)/s.d.>	6.4 (2.0)	8.1 (2.1)
R_{sym} (%)	16.5 (63.0)	17.1 (64.9)
R_{cryst} (%)	32.7	27.0
R_{free} (%)	35.1	29.9
RMSD Bond length (Å)	0.015	0.007
RMSD Bond angle (°)	1.63	0.95
PDB entry code	3P5C	3P5B

† Values in parentheses correspond to the statistics for the highest resolution shell.

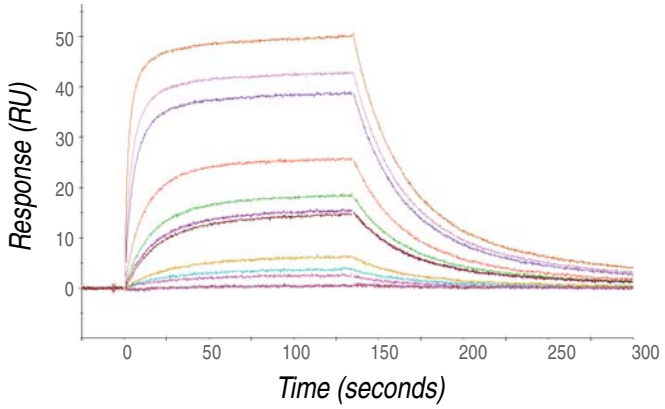
Supplementary Figure S1 (Lo Surdo et al.)

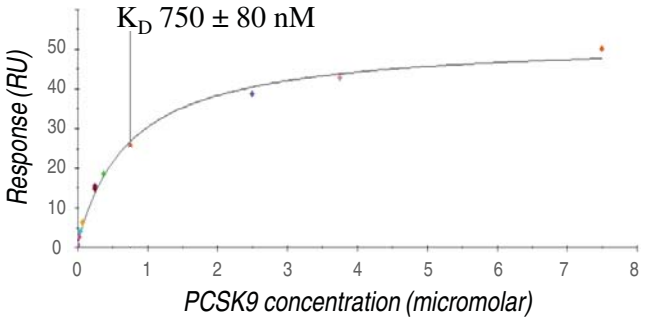


Supplementary Figure S2 (Lo Surdo et al.)

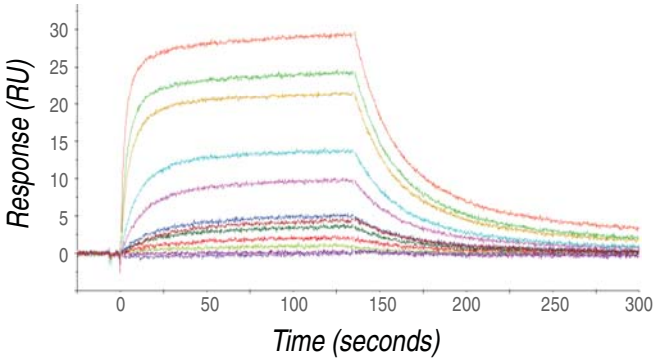


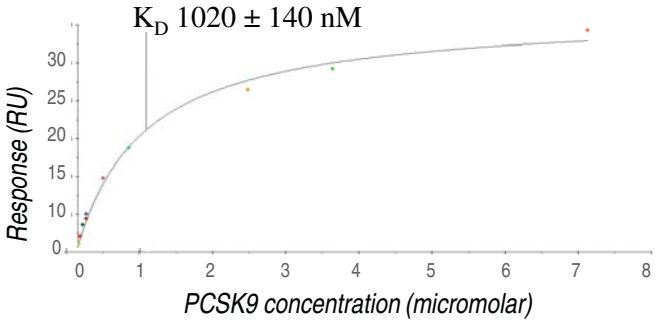
1. LDLR wild type, pH 7.5



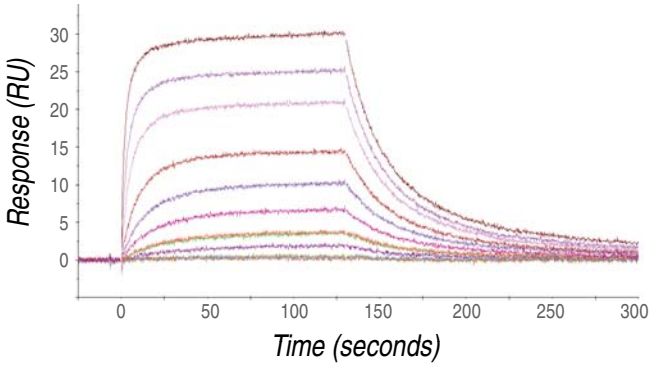


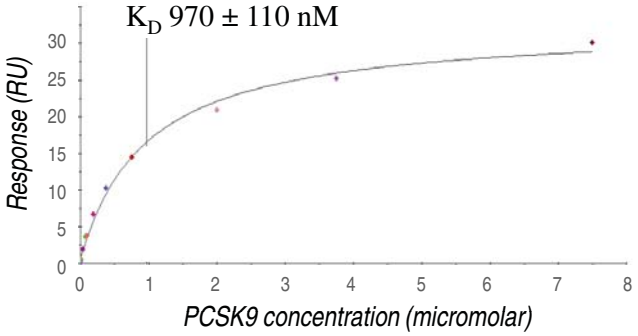
2. LDLR L626A , pH 7.5



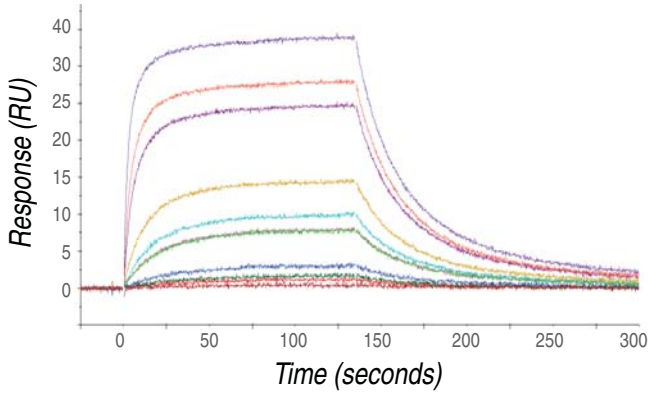


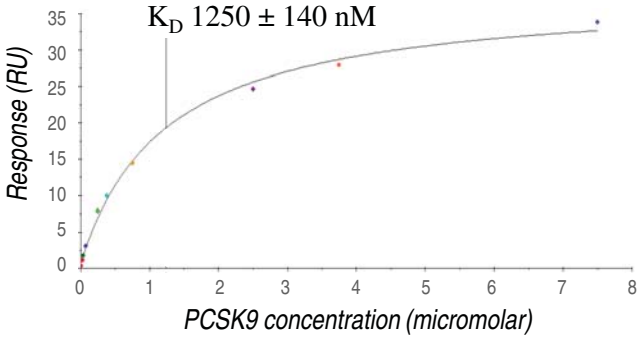
3. LDLR L626R , pH 7.5





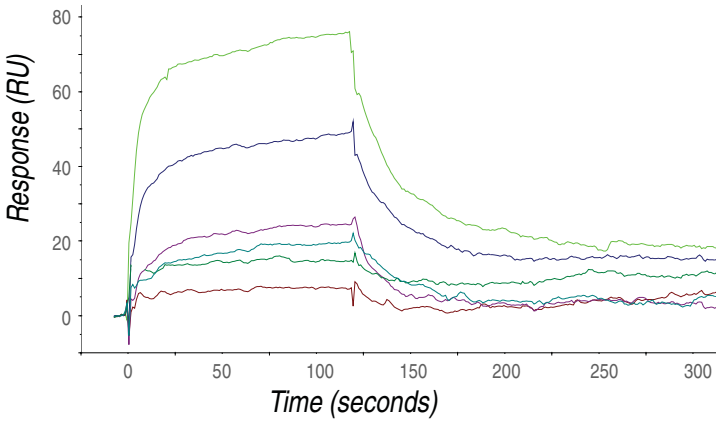
4. LDLR L626E , pH 7.5

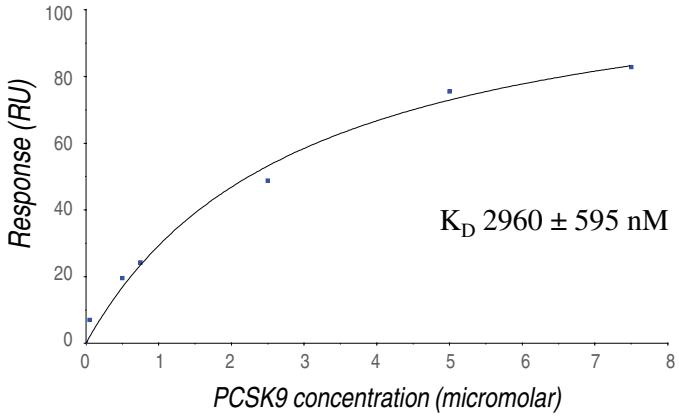




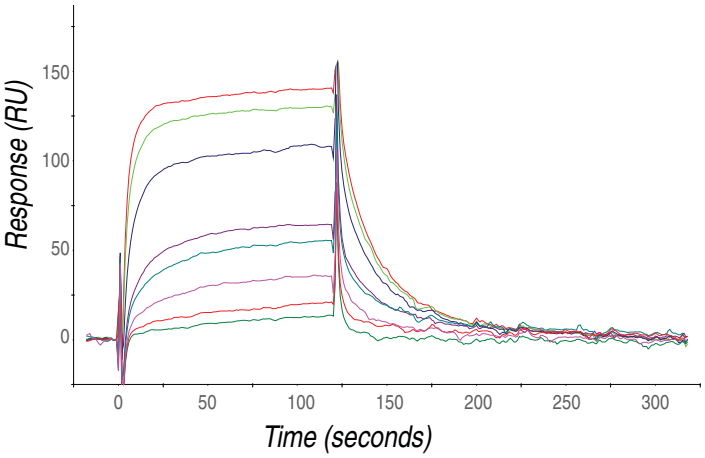
Supplementary Figure S3 (Lo Surdo et al.)

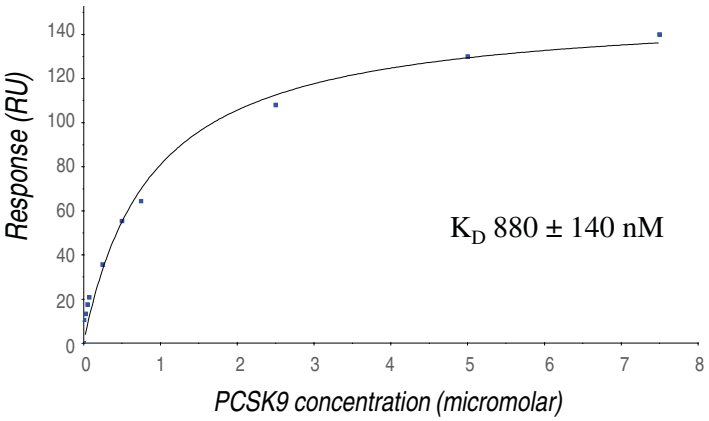
5. LDLR H635N , pH 7.5



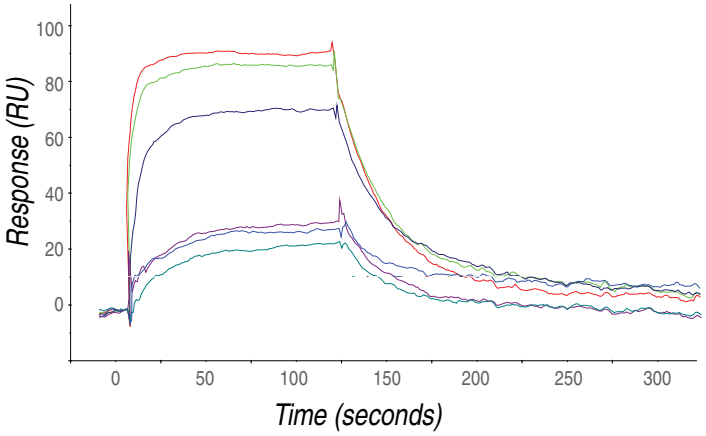


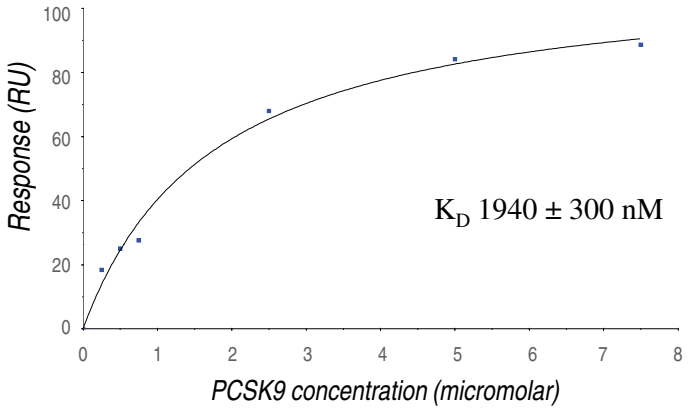
6. L7-EGFPH, pH 7.5



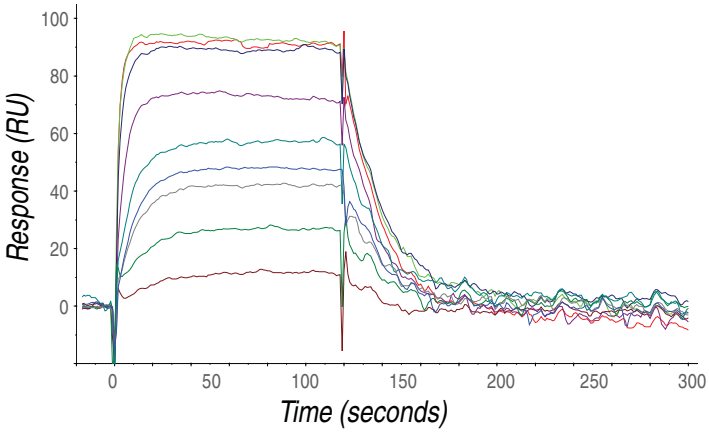


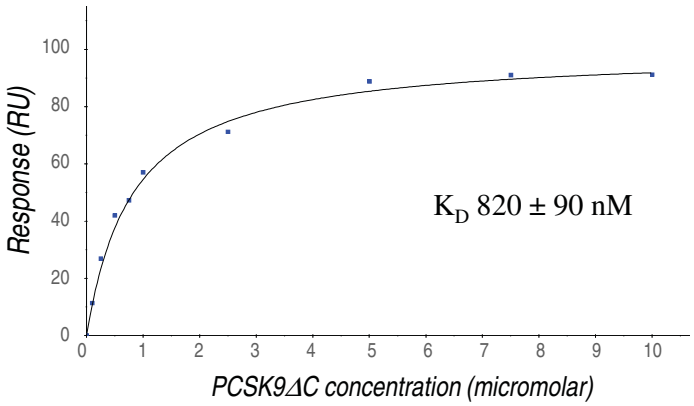
7. L7-EGFPH GS→PP, pH 7.5



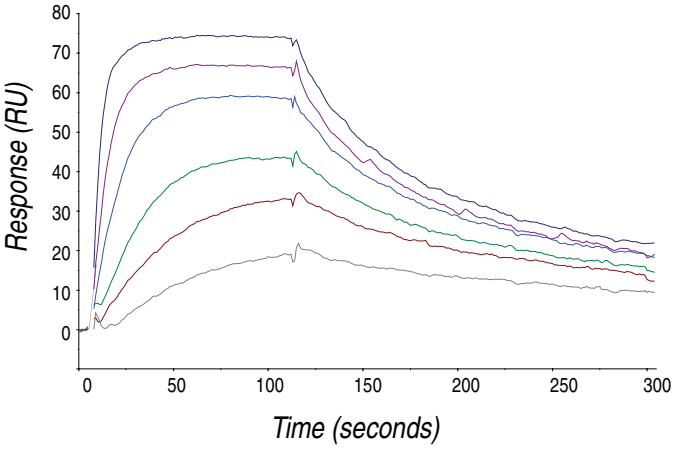


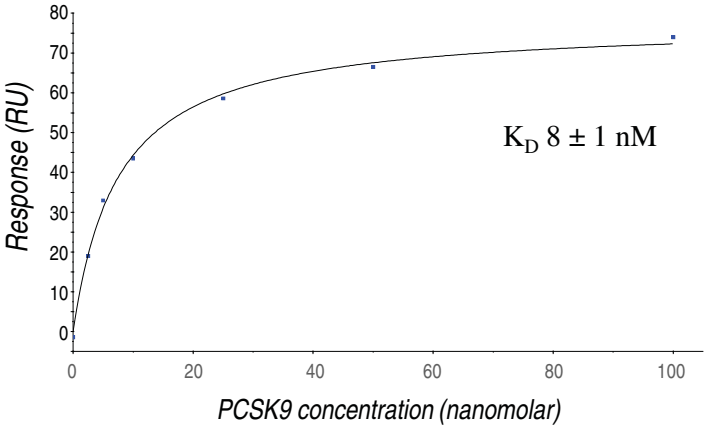
8. L7-EGFPH, pH 7.5, injecting PCSK9ΔC



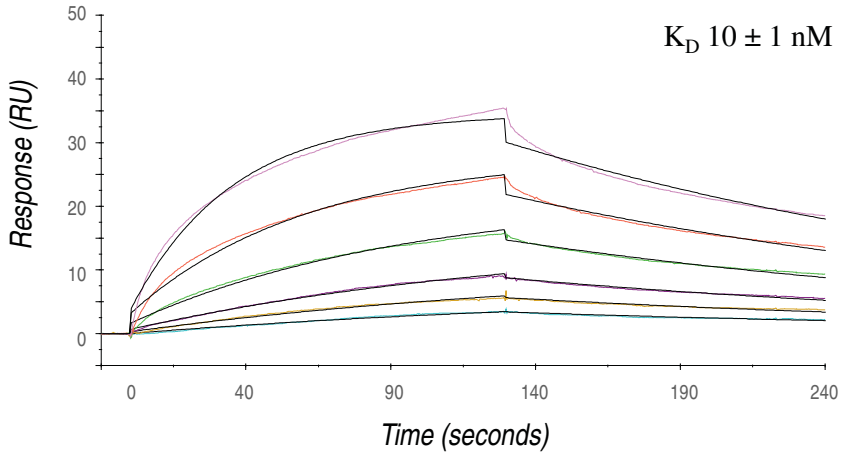


9a. LDLR wild type, pH 5.5, steady-state analysis

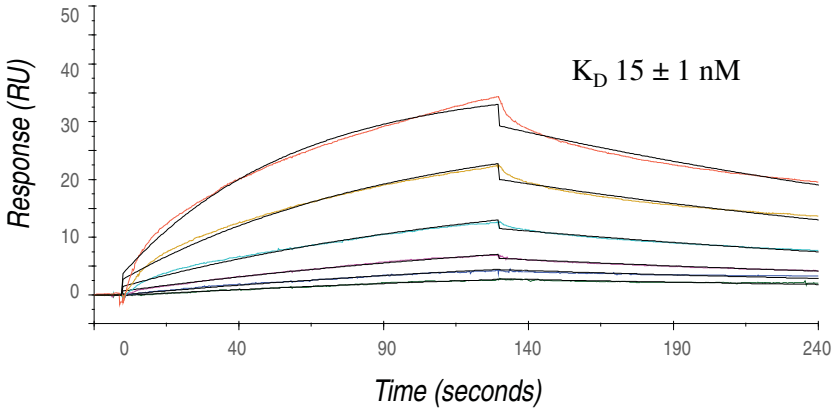




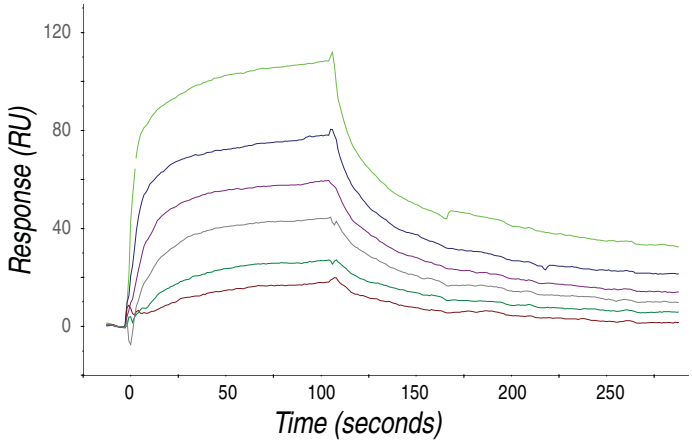
9b. LDLR wild type, pH 5.5, kinetic analysis

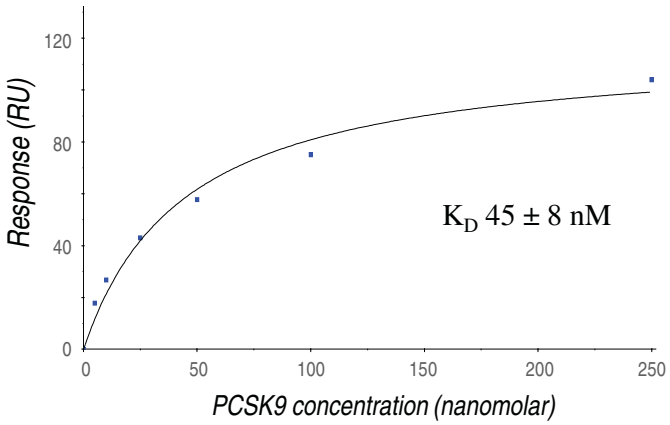


10. LDLR L626E, pH 5.5, kinetic analysis

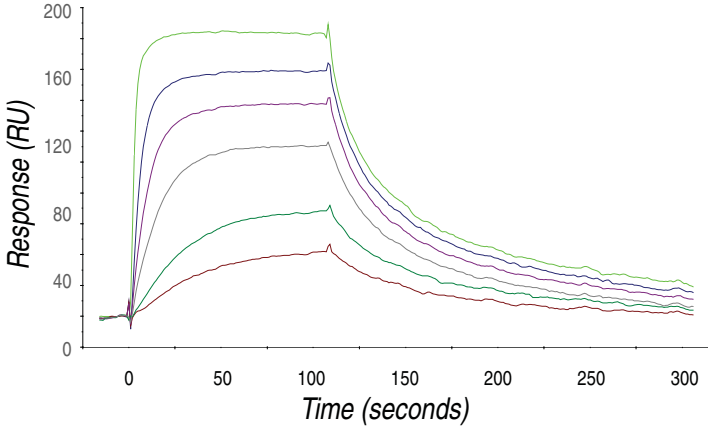


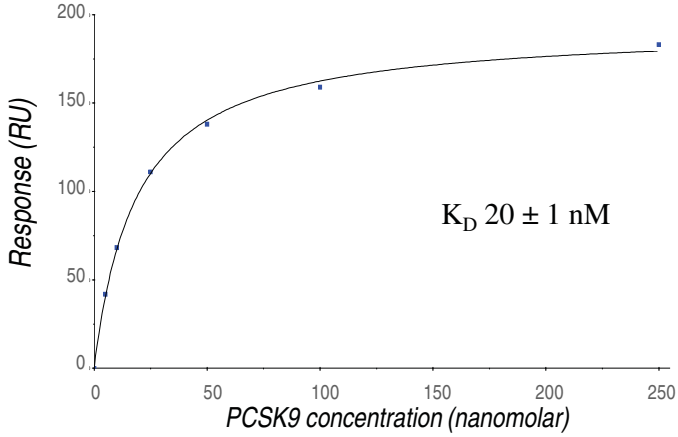
11. LDLR H635N , pH 5.5



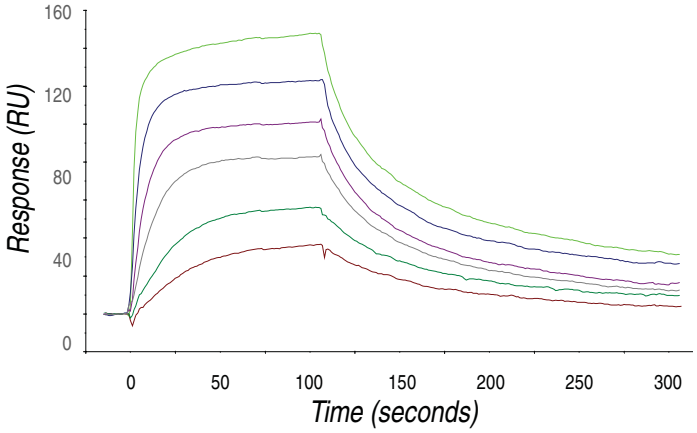


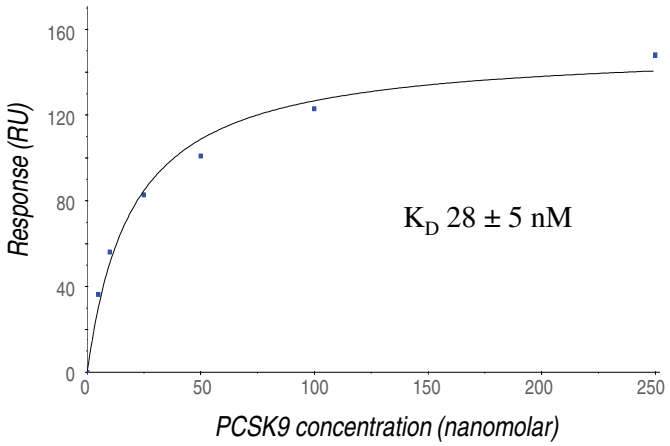
12. L7-EGFPH, pH 5.5



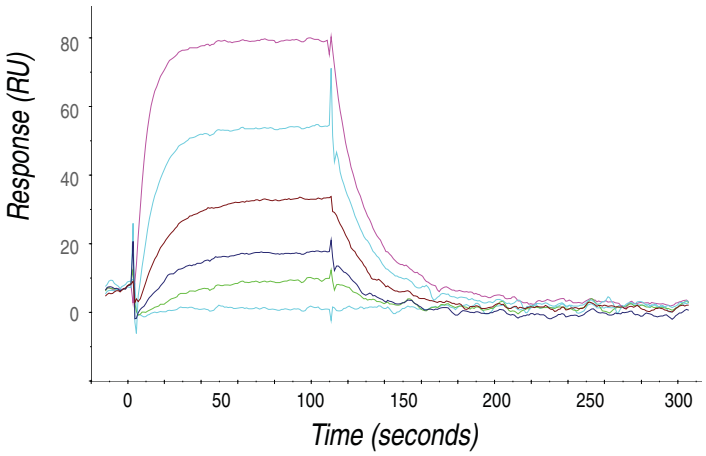


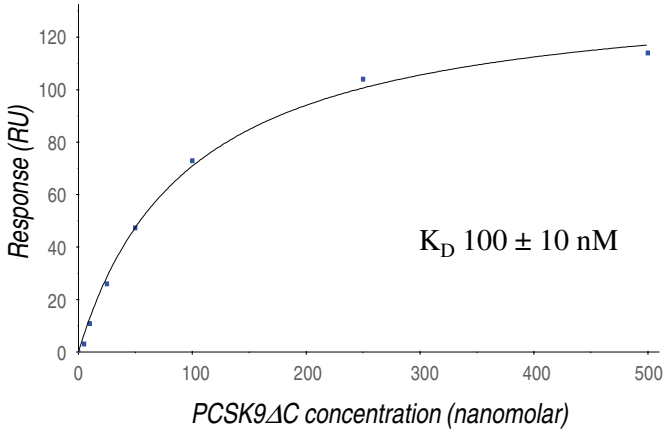
13. L7-EGFPH GS→PP, pH 5.5



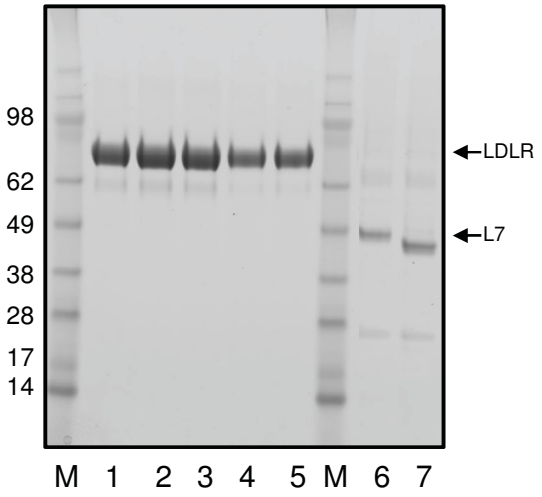


14. L7-EGFPH, pH 5.5, injecting PCSK9ΔC





Supplementary Figure S4 (Lo Surdo et al.)



Key to Lanes:

1. LDLR
2. LDLR *H635N*
3. LDLR *L626A*
4. LDLR *L626R*
5. LDLR *L626E*
6. L7-EGFPH
7. L7-EGFPH *GS-PP*

M: Molecular weight markers