Published in final edited form as:

Proteins. 2018 November ; 86(11): 1165–1175. doi:10.1002/prot.25597.

COMPARATIVE STUDY OF THE EFFECT OF DISEASE CAUSING AND BENIGN MUTATIONS IN POSITION Q92 ON CHOLESTEROL BINDING BY THE NPC1 N-TERMINAL DOMAIN:

The role of Q92 in CLR transport by NPC1

Marharyta Petukh^{a,b,1,*} and Igor B. Zhulin^{a,b,c,2}

^aDepartment of Microbiology, University of Tennessee, Knoxville, Tennessee, USA

^bComputational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

^cCenter for Bioinformatics, Pavlov First Saint Petersburg State Medical University, Saint Petersburg, Russia

Abstract

The Niemann-Pick type C1 (NPC1) protein is a large transmembrane protein located in lysosomes/endosomes. NPC1 binds cholesterol (CLR) and transports it to cellular membrane and endoplasmic reticulum. Mutations in NPC1 cause Niemann-Pick type C (NPC) disease, a rare autosomal disorder characterized by intracellular accumulations of CLR and subsequent neurodegeneration leading to premature death. Among known disease-causing mutations in NPC1, Q92R is the one that is located in the N-terminal cholesterol-binding domain (NTD). Here we study the effect of the mutation on the ability of NPC1(NTD) to bind and retain CLR in the binding pocket using structural analysis. We compare characteristics of the Q92R and Q92S mutant type (MT) protein, which is predicted to be benign. We provide detailed investigation of the CLR-NPC1(NTD) binding process; and propose the mechanism, by which Q92R mutation causes NPC disease. We show that although Q92 residue neither directly participates in catalytic activity of the NPC1(NTD), nor defines its CLR-binding specificity – it is important for the overall protein structure as well as for providing favorable electrostatic environment for CLR transfer. Our results suggest that a negative electrostatic potential of the CLR binding site (the S-opening) might promote NPC2 interaction with NPC1(NTD) and/or proper CLR orientation and its enforced transfer. We show that in contrast to the benign Q92S mutation, Q92R significantly reduces electrostatic potential around S-opening, and thus likely affects NPC1(NTD)-NPC2 interaction and/or CLR transfer from NPC2 to NPC1.

*To whom correspondence should be addressed: Marharyta Petukh, Address: Biology Department, Presbyterian College (503 S Broad St, Clinton, SC 29325), mpetukh@presby.edu.

¹Present address: Biology Department, Presbyterian College, Clinton, South Carolina, USA

²Present address: Department of Microbiology, The Ohio State University, Columbus, Ohio, USA

Conflict of Interest

Authors declare no conflict of interests.

Keywords

Niemann-Pick type C1; N-terminal domain; cholesterol; mutation; electrostatic potential

Introduction

Cholesterol is a vital component of cell membranes in eukaryotes. It plays a dual role as it rigidifies the fluid membrane, and thus reduces passive permeability and increases the mechanical durability of the lipid bilayer; and secondly it enables post-Golgi protein sorting by regulating the thickness of the lipid bilayer. Regulation of synthesis, influx and efflux keeps cellular cholesterol levels precisely controlled¹. The exogenous CLR is derived from serum low-density lipoprotein (LDL) and related plasma lipoproteins by receptor-mediated endocytosis^{1,2}. Thus, LDLs are recognized by cell-surface receptors and transported to endosomes and lysosomes. While in late endosomes/lysosomes, LDL is released from its receptor and degraded into free CLR³. Genetic and biochemical characterizations revealed that Niemann-Pick type C1 (NPC1) (transmembrane) and Niemann-Pick type C2 (NPC2) (small soluble) proteins cooperate to export LDL-derived cholesterol from late endosomes and lysosomes to the endoplasmic reticulum and other cellular compartments^{4,5}.

Previously it was demonstrated that both NPC1 and NPC2 can deliver CLR to the liposome, although while NPC2 can function on its own, NPC1 requires the NPC2 presence for CLR transport^{4,6,7}. Experimentally it was found that although both proteins can directly bind free CLR, the on- and off-rate of CLR binding for NPC2 is much higher than for NPC1^{4,8}. Recent crystallographic⁹ and cryo-EM¹⁰ structures of the human NPC1 provided a structural foundation explaining that entrance into the cholesterol-binding pocket is obstructed by helices that must move aside to permit entry to or exit out through S-opening¹¹. NPC2 facilitates CLR transfer in NPC1 – the rate of cholesterol binding to NPC1(NTD) increases more than 15-fold when the sterol is first bound to NPC2 and then transferred. Moreover, experimentally it was demonstrated that cholesterol can be transferred between NPC1 and NPC2 in a bidirectional mode⁴. Thus, at T=37 °C NPC2 efficiently transfers CLR to both NPC1 and to other NPC2 at wide range of pH, while NPC1 preferentially transfers CLR to NPC2 rather than to other NPC1 protein, with a maximum rate at acidic pH.

NPC1 is a 1278 aa transmembrane protein that has 13 transmembrane segments and three relatively large, lumenally oriented domains – NTD, middle luminal domain (MLD), and C-terminal domain (CTD)^{10–14}. MLD and CTD are flanked by transmembrane helices 2/3 and 8/9, respectively. NTD has the highest conservation rate within the entire NPC1 protein sequence¹⁵ and it contains the cholesterol-binding site^{16,17}. NTD, a 240 aa domain of NPC1, was shown to be soluble, and retain cholesterol-binding activity upon expression in vivo¹⁷. NPC1(NTD) was demonstrated to bind CLR in an orientation with the 3 β -hydroxyl site facing the interior of the domain and the iso-octyl side chain exposed, which is opposite to CLR binding by NPC2^{8,11,16}. Both MLD and CTD interact with NTD and thus modulate its activity^{14,18}.

Recently it was demonstrated that CTD interacts with NTD (residues of the loop 227–238) maintaining 410 Å² interface, that is almost twice larger than the interface formed between NTD and MLD¹⁴. MLD is known to directly interact with NPC2-CLR complex, and thus assist in the transfer of CLR from NPC2 onto NPC1 protein^{4,11,18}. The binding occurs at the distance ~90 Å from the lysosome's limiting membrane¹³. Interestingly, NPC1(MLD) binds NPC2-CLR complex only at acidic pH¹⁸. Crystal structure of the NPC1(NTD)-NPC2-CLR complex shows that L175, K179, D180, and D182, G199 and Q200 residues of NPC1 are involved in direct interactions with NPC2¹³. NPC2-NPC1 interaction promotes the CLR transfer from NPC2 to NPC1 by facilitating reorientation of the helical subdomain, enlarging opening (“S-opening”) of the CLR-binding pocket in NPC1 and thus releasing CLR^{4,11,13,19,20}. The type of interaction between NPC1 and NPC2 for CLR transport was termed a “hydrophobic hand-off” or “sliding” transfer model, where close NPC2-NPC1 interaction allows CLR transfer from NPC2 to NPC1(NTD) without being exposed to the water phase^{11,18–20}. It is suggested that the CLR release from the NPC1(NTD) through the endosome/lysosome membrane requires transmembrane (TM) subdomains of NPC1^{9,11}. According to the current model, after CLR binding to NPC1(NTD), the polyproline segment located between the NPC1(NTD) and TM1 reorients in such a way as to make it possible for CLR delivery to the sterol-sensing domain formed by conserved TM helices of NPC1^{9,21,22}.

Mutations in any of the two genes – NPC1 or NPC2 – cause the Niemann-Pick disease type C (NPC) – a lysosomal storage disease characterized by intracellular accumulation of LDL-derived cholesterol, sphingomyelin, and other lipids in endosomes and lysosomes and delayed induction of cholesterol homeostatic reactions. NPC patients may develop progressive neurodegeneration, hepatosplenomegaly, and premature death^{23,24}. The most common form of the disease, NPC1, results from mutations in the NPC1 gene. As an example, naturally occurred variant Q92R (VAR_043174), located in NTD of NPC1 protein, was reported as a disease-causing mutation in several independent studies^{25–27}.

Here we report the effect of Q92R mutation on the functionality of NPC1(NTD). We applied structural, and physico-chemical analysis to decipher the mutation consequences on the structural integrity of the NTD and its CLR binding ability. To reveal a potential mechanism of the disease-causing Q92R mutation, we also analyzed structural and physico-chemical consequences of Q92S mutation that was predicted to be benign^{28–31}.

Methods

Predicting benign substitutions in NPC1

The SAVER algorithm²⁸ was used to predict damaging and benign mutations and the SAVER NPC1 website (<http://saver.zhulinlab.org/>) was used to predict benign substitution in NPC1 in position 92. Substitution Q92S was selected for further analysis, because it was predicted to be the only benign amino acid substitution not only by SAVER, but also by other algorithms, specifically PolyPhen2²⁹, SIFT³⁰, and PROVEAN³¹.

Structure preparation and MD simulation

The NPC1(NTD) crystal structure (NPC1 residues 23–247), with cholesterol bound to the ligand-binding pocket, was retrieved from the Protein Data Bank (PDB ID: 3JD8). The Q92R and Q92S mutations were introduced by applying *mutate* subroutine in VMD³² using the wild type (WT) protein as the template. To generate the protein in apo state, the CLR was removed from the ligand-binding pocket.

As it was previously discussed, NPC1(NTD) is located in luminal part of endosomes/lysosomes^{4,5}, where pH is in a range of 4.0–6.5^{33–35} and CLR transport from NPC2 to NPC1 is pH depend – it is twice slower at pH=7.0 compared to pH=5.5¹⁰. We accounted for the pH conditions of the protein environment by assigning appropriate protonation state of charged residues according to DelPhiPKa webserver³⁶. Calculations were made at pH=4.5. Out of all residues that might hold charge (K, H, R, D, E), all 5 Arg and 10 Lys residues in NPC1(NTD) are charged; the only His residue H215 is neutral; 3 out of 11 Glu residues (E61, E127, E229) and 1 out of 16 Asp residues (D162) was assigned as neutral. VMD was applied to add appropriate number of hydrogen atoms to each residue as needed using Charmm36 force field. Topology and parameters of CLR were generated by SwissParam webserver based on the ligand structure from 3JD8 structural file³⁷.

As it was previously mentioned, all 18 cysteine residues in NPC1(NTD) form 9 cysteine bridges¹¹ – this feature was reinforced while generating the PSF file in VMD by applying specific patch for each pair of closest cysteine residues. This allowed to conserve all cysteine bridges while applying NAMD.

Once all structures (WT, MTs in apo state and with bound CLR) were built, they were relaxed and simulated using NAMD³⁸. The initial structural clashes were avoided by structures relaxation for 5000 steps via conjugate gradient algorithm. The minimized structures were further relaxed by gradually increasing the temperature of the system from 0 K to 298 K by 1 K every 100 steps (0.2 fs) of MD simulations for the total number of 29800 steps (59.6 fs). The next production phase of MD simulations lasted 50 ns at constant temperature (T=298 K). Every 1000 steps (2 fs), a snapshot of the system was exported to a DCD trajectory file that was later unzipped into multiple PDB snapshots and used for statistical evaluation. In energy minimizations, relaxation, and production phases of MD the water phase was modeled implicitly with Generalized Born model. The solvent and solution dielectric constants were set to $\epsilon=1$ and 80 respectively. The control over constant temperature (T=298 K) was maintained with Langevin Dynamics algorithm in NAMD.

RMSD/RMSF analysis

Once the production phase of MD simulations for each of the systems was finished, the snapshots were aligned in a way to minimize RMSF of all cysteine residues within the snapshot (as they were shown to be important in supporting protein structure by forming cysteine bridges¹¹). Aligned snapshots allowed generating the average protein structure. The backbone atoms of the aligned proteins snapshots were used to calculate RMSD. This parameter allows to monitor the system deviation from the initial conformation over the time

of MD simulations. The averaged protein structure (CA atoms) was used to calculate the RMSF for each residue.

Native contacts

A pair of residues were considered to form a native contact if any two heavy atoms of their sidechain are within 4 Å from each other. To eliminate trivial cases, we removed pairs of residues that are within three residues from each other in sequence.

Solvent accessible surface area

The exposure of protein residues to the solvent was calculated with VMD software using *measure* subroutine³². The relative solvent accessible surface area (SASA) of CLR in the protein-CLR system was calculated by the following equation

$$SASA_r = \frac{SASA(ligand, protein)}{SASA(ligand)} \cdot 100 \% \quad (1),$$

where $SASA(ligand, protein)$ - SASA of the ligand bound to the protein; $SASA(ligand)$ - SASA of the free ligand.

Specific bonding

Hydrogen bonds—The VMD software³² was used to identify residue pairs forming hydrogen bonds (HB). Standard parameters for HB geometry were used with a distance cutoff between donor and acceptor 3.0 Å and the cutoff angle between donor, proton and acceptor 30 degrees.

Salt bridges—The VMD software³² was used to identify residue pairs forming salt bridges (SB). Standard parameters for SB geometry were used with a distance cutoff between nitrogen atoms of positively charged residues and oxygen atoms of COO⁻ moiety of negatively charged residues set to 4.0 Å.

Hydrophobic contacts—Hydrophobic contacts (HydroPhob) between Ala, Val, Leu, Ile, Phe, Trp, and Met residues were identified in the MD trajectory by calculating the distance between sidechain atoms. The cutoff distance between any pairs of sidechain atoms was set to 4 Å.

Electrostatic potential map

DelPhi software was used to calculate electrostatic potential around the NPC1(NTD) with the following parameters: scale was 1 Å per grid, perfil 70 %, dielectric constants for protein and water were 1 and 80, respectively³⁹. The electrostatic potential was visualized with VMD software, where the electrostatic potential distribution was shown as the forcefield lines.

Geometric parameters

All geometric characteristics of NPC1(NTD)-CLR system were analyzed with in-house developed algorithms written in *Python* programming language.

Secondary structure elements

STRIDE software⁴⁰ was used to assign secondary structure for each residue in NPC1(NTD). Parts of the protein forming helix or strand were considered to be “structured”, while the rest – “unstructured” regions.

Results

Among other disease-causing mutations in NPC1, Q92R is the one located at the luminal NTD. Previously it was demonstrated that recombinant NTD of NPC1 is soluble and it can effectively bind CLR¹⁷ that suggests that the analysis made on the domain can be attributed to the full protein. Here we study the physico-chemical properties of the NPC1(NTD) in apo state and with bound CLR. The effect of the mutation on the protein activity is deciphered by comparing structural and physico-chemical parameters with the corresponding one of the WT protein. The difference in the change in the protein functionality caused by disease-causing (Q92R) and benign (Q92S) mutations compared to the WT used to decipher the role of Q92 residue in NPC1(NTD).

After 5000 steps of structural adjustment due to energy minimization and subsequent equilibration, we carried out 50 ns of MD simulations for WT and Q92R and Q92S mutants in both apo state and with bound CLR. The snapshots of protein structures were taken every 0.1 ns resulting in 500 snap shots. RMSD analysis shows that during the production time the protein structures do not significantly change compared to their averaged one (data not shown). This allows to compare the physico-chemical characteristics of WT and MTs with already equilibrated structures.

Characterizing NPC1(NTD) residue types composition

The composition of NPC1(NTD) in terms of amino acid types is shown in Figure 1. As it is expected, hydrophobic residues (~30 % of all residues) are mostly aligning CLR-binding pocket. This ensures energetically favorable binding of hydrophobic CLR to the protein. Residues Q92 is located in the vicinity of S-opening – where CLR is transferred from NPC2 to the ligand-binding pocket of NPC1(NTD)¹¹.

Flexibility

The typical RMSF signatures of WT NPC1(NTD) in apo state and with bound CLR are shown in Figure 2. Data suggests that CLR reduces the flexibility of the protein upon binding by ~10 %. The effect is more pronounced at the C- and N-termini of the domain. As expected, the most flexible regions in WT (apo and with bound CLR) are in unstructured regions. The effect of mutations (Q92R and Q92S) on the protein flexibility is shown in Figure 3. Thus, the disease-causing mutation Q92R slightly (~3 %) increases the flexibility of apo NPC1(NTD), mostly through unstructured regions, but it significantly (~15 %) decreases the flexibility of CLR-bound NPC1(NTD), affecting both structured and

unstructured regions of the domain. In contrast, the potentially benign mutation at the same position, Q92S, increases flexibility of the protein in both apo and CLR-bound states. In the apo state, the MT has increased flexibility of structured area by 16 %, and unstructured one by 7 % compared to WT. In the presence of CLR the mutation increases flexibility of mostly unstructured regions (38 %) without a noticeable effect of the structured ones.

Structural compactness

To characterize the compactness of the NPC1(NTD) structure we used two parameters – number of native contacts and SASA of the protein. According to Table 1, the presence of CLR in the binding pocket makes the domain structure more compact in WT (by increasing the number of native contacts by ~2 %) and in both MTs (by less than 1 %). In the apo state Q92R mutation slightly increases (0.4 %), whereas Q92S mutation decreases (0.7%) the compactness of the structure compared to the WT. In the presence of CLR both mutations make the protein structure less compact (0.8 % for Q92R, and 2.1% for Q92S).

The SASA analysis of the protein shows that in WT binding of CLR does not significantly change the protein SASA – the average value is reduced by only 0.7 % (Figure 4). In the apo state, Q92R mutation reduces SASA by 2.5 % compared for the WT, whereas Q92S mutation does not change the protein SASA. In presence of CLR both mutations increase the SASA of the protein – by 0.7 % for Q92R and 1.3 % for Q92S. These observations correlate well with the effect of mutations on proteins flexibility.

Structure stabilizing factors

Among structure stabilizing factors we consider the ability of protein residues to form specific bonds such as HB, SB, and be involved in hydrophobic contacts.

Residues involved in stable (more than 50 % of the time during MD simulation), non-neighboring (further than 2 residues apart in the sequence) HB/SB in WT and both MTs are in apo state and with bound CLR are shown in Figure 5. The comparative analysis shows that there are three stable HB/SB that are conserved in all studied systems and are distant from the ligand-binding pocket: R96-D242; R161-D182; and D131-T134 (marked in bold in Figure 5). In addition to that the WT has two SBs E127-K142 and R39-D131 in both apo and CLR-bound states and HB D128-T139 in apo state only which put further mutual constraints on the N-terminus loop and strands. The Q92R mutation slightly rearranges the HB/SB in apo state – abolishes HB D128-T139 and SB R39-D131 to form two SBs D131-K138 and R78-D131 – but preserving the total number of stable specific interactions as in WT-apo. With bound CLR Q92R MT loses additional constrain within strands (E127-K142) making it more flexible (Figure 3, upper left). The Q92S mutation does not affect the number of stable HB/SB in apo state, but in comparison with WT it substitutes HB D128-T139 within strands with SB K38-D85 between N-terminus loop and one of the domain helices. In presence of CLR Q92S increases the number of stable HB/SB in NPC1(NTD) by transferring SB R39-D131 between strands and N-terminus loop to SB R78-D131 between strands and one of the domain helices and adding two new bonds SB D131-K138 and HB S239-D242 that stabilizes strands and C-terminal loop accordingly. Interestingly, neither in WT no MTs there are no stable HB/SB formed between helices that allows to assume that

this feature might be crucial for the domain functioning as adding specific interactions imposes constraints that potentially can restrict helices motion and conformational changes during CLR binding. The average number of HB/SB during the production phase of MD simulations for WT and MTs in apo state and with bound CLR is shown in Table 2. Data shows that binding of CLR reduces the number of HB/SB in all cases. Both mutations reduce the number of this type of specific interactions as well.

Structural analysis of NPC1(NTD) during the production phase of MD simulations shows that the domain has a broad network of hydrophobic interactions – the average number of hydrophobic interactions is listed in Table 3. Data shows that both mutations reduce the number of hydrophobic interactions compared to the WT protein. However, in case of Q92S mutation the number of significant contacts in apo protein reduces only by 4, while in case of Q92R – by 8, that is ~10 % of the total number of significant hydrophobic interactions in WT. With CLR present in the ligand-binding pocket, both mutations equally reduce the number of hydrophobic pairs by ~5 %.

Interactions between NPC1(NTD) and CLT.

The next step is to understand the mechanism of CLR binding to the ligand-binding pocket of NPC1(NTD), and the effect of mutations on this process. In this section we describe the geometrical characteristics of the pocket, orientation of CLR upon binding, are protein residues interacting with CLR.

S-opening. Geometry.—As it was previously mentioned, CLR enters NPC1(NTD) through S-opening formed by three helices – $\alpha 3$, $\alpha 7$, and $\alpha 8$ ¹¹ (Figure 6). It was suggested that this part of the protein should be flexible to allow forming the CLR transfer tunnel between NPC1 and NPC2¹³. We analyzed the geometry of S-opening by characterizing distances between CA atoms of its forming residues – L89, N198, and L176. The average values and standard deviations are shown in Figure 7. The latest parameter is used for the evaluation of S-opening flexibility. The calculations were performed for WT and both MTs in apo state and with bound CLR. In WT placing CLR in the binding pocket does not change the average distance between residues L89 and N198, increases the distance between residues L89 and L176 by 1.5 Å, and decreases the distance between residues L176 and N198 by 1.5 Å. CLR slightly reduces distance fluctuations between residues L89 and N198 and residues L176 and N198, making S-opening more rigid. In apo state, the disease-causing mutation Q92R affects only the average distance between residues L176 and N198 by reducing it by 1 Å but compensates it by increased fluctuations of the distance between residues L89 and L176 (more than twice). Q92R mutation has almost no effect on geometry of S-opening in the presence of CLR. The predicted benign mutation Q92S affects mostly the distance between residues L89 and N198. Thus, in apo state, the average distance reduction by 1 Å is accompanied by significant rigidity compared to the WT. In the presence of CLR, the Q92S mutation also decreases this average distance by 1 Å although without significant effect on its fluctuations.

The change in the diameter of the inscribed sphere in the triangle made of CA atoms of residues L89, N198, and L176 was used as another characteristic of the S-opening protein. It

defines the probability of CLR to enter/exit NPC1(NTD) through S-opening from geometrical prospect. This parameter is not significantly different between apo structure and CLR bound structure (Figure 7). Moreover, both mutations (Q92R and Q92S) have no distinguishable influence on this characteristic in either state of the protein compared to the WT.

S-opening. Electrostatics.—Previously it was demonstrated that CLR enters NPC1(NTD) through S-opening with its 3 β -hydroxyl group facing the inner part of the protein¹⁶. Because this group is polar, it is reasonable to assume, that electrostatic properties of the protein (S-opening specifically) may facilitate and guide CLR entering the binding pocket. Thus, we calculated the electrostatic potential distribution around the WT protein in apo state and compared it with both mutant proteins. In Figure 8 we show the electrostatic potential as forcefield lines around S-opening (green circle) formed by three helices - α 3, α 7, and α 8 (shown in yellow). The density of the lines corresponds to the strength of electrostatic potential, while the color – to its polarity. Thus, the presence of mostly red forcefield lines in WT and both MTs indicates that electrostatic potential in S-opening is rather negative. However, Q92R mutation significantly reduces the electrostatic potential strength, while Q92S – slightly increases it. This suggests that in case of the disease-causing mutation substitution of polar Q with positively charged R in position 92 reduces electronegativity and thus might affect the ability of S-opening to guide CLR binding and/or NPC1-NPC2 interaction. In contrast, the Q92S substitution preserves the overall polarity and thus retains the negative electrostatic potential of the S-opening.

CLR-binding pocket.—One of the major factors allowing NPC1(NTD) to accept CLR from NPC2 and successfully transfer it through the lysosome/endosome membrane is an optimal storage of the ligand in the ligand-binding pocket. It is expected that any deviation from this optimum location as well as abnormality in the CLR dynamics would negatively affect CLR transportation. Here we evaluate the behavior of CLR in terms of its motion and interaction with NPC1(NTD) residues while being placed in the ligand-binding pocket of WT over the production phase of MD simulations. The results are used to evaluate the effect of disease causing (Q92R) and benign (Q92S) mutations.

We started with the geometric location of CLR in the ligand-binding pocket. By monitoring the angle between the ligand and the protein we can define the change of CLR orientation during MD simulations. We chose atoms of CLR molecule (C3 and C17) that designate the direction of four hydrocarbon rings, and CA atoms of the protein that define direction of helix α 3 (residues R78, L89) and helix α 4 (residues T112, L105) (see insert in Figure 9). We chose those helices because they are located within the ligand-binding pocket and are less flexible than other parts of the domain. Analysis shows that in WT CLR almost aligned with α 3 helix by composing an angle of ~5 degrees (left set of graphs in Figure 9), standard deviation ~4 degrees. Also, data implies that both mutations (Q92R and Q92S) do not change this angle compared to the WT. The angle between CLR and α 4 is ~30 degrees in WT with higher standard deviation of 6 degrees (right graphs in Figure 9). Q92R reduces, while Q92S increases this angle by ~1degree. Both mutations also decrease standard deviation by 1.5 degrees.

Taking into the consideration the previous finding that CLR is aligned with $\alpha 3$ helix, we studied the possible displacement of the ligand related to this helix by monitoring the change in two distances: 1) between CLR atom C3 and CA atom of protein residue 78; and 2) between CLR atom C17 and CA atom of protein residue 89 (upper and bottom graph on Figure 10, correspondingly). For 3β -hydroxyl site of CLR the distance between C3 atom of CLR and CA atom of the protein residue 78 does not significantly change over the time of MD simulation. The average value for WT and both MTs are $\sim 9 \text{ \AA}$ with standard deviation of $\sim 0.5 \text{ \AA}$. This suggests that 3β -hydroxyl site of CLR participates in specific interactions that are not affected by the mutations in position 92. In contrast, the isooctyl site of CLR was found to be more mobile – the average distance between C17 atom of the ligand and CA atom of residue 89 in WT protein is $8.7 \text{ \AA}$, SD of $\sim 0.9 \text{ \AA}$. The Q92R mutation does not change the average distance between the two atoms, but it slightly reduces the fluctuation (SD= $0.8 \text{ \AA}$). Contrarily, the Q92S mutation affects both the distance between C17 or the ligand and CA of residue 89 of the protein (reduced to $8.4 \text{ \AA}$), and fluctuation (SD= $0.7 \text{ \AA}$). This data indicates that for normal functioning the motion of 3β -hydroxyl site of CLR is more constrained by residues in the ligand-binding pocket than the isooctyl site; out of two mutations only benign Q92S has influence on free motion of the CLR, its isooctyl site.

The chemical structure of CLR suggests that most of its molecule is hydrophobic, i.e. it does not prefer to be exposed to the water solution. Thus, the ligand-binding pocket, where the ligand is located after transfer from NPC2 and before transport across the membrane, should minimize interactions of CLR with water molecules. To monitor the degree of ligand exposure to the solution, we monitored its SASA during the production phase of MD simulations in WT and both MTs with bound CLR. We found that in all cases CLR exposes only 6 % of its surface to the water phase (data not shown).

For characterization of residues surround CLR while being placed in the ligand-binding pocket, we accounted for only those which sidechain is within $5 \text{ \AA}$ of the ligand. Because we found that the location of CLR does not change much during MD simulations, we considered only those residues of NPC1(NTD) that are close to CLR more than 50 % of the time during the production phase of MD simulation – shown in Figure 11. Thus, in WT, twenty-four residues are surrounding the ligand with more than 50 % of frequency. These include eight polar (E30, N41, Q79, T82, N86, T112, Y192, Q200), thirteen hydrophobic (W27, L87, L89, F93, L104, F108, L111, L176, M193, F194, A201, F203, I205) residue and two prolines (P90 and P202). The Q92R substitution removes L87 from the vicinity of CLR but substitutes it with equal L144 interaction. The Q92S mutation results in adding K38 residue in the CLR surrounding that further restricts the motion of the polar site of the ligand, and at the same time abolishes interaction with I205 close to its isooctyl site.

We also analyzed the ability of CLT to establish HB with residues of the protein in the ligand-binding pocket. We found that only two residues of NPC1(NTD) are capable to form direct HB with the ligand – N41 and Q79 (see Table 4). The Q92R mutation does not significantly affect HB formation between CLR and N41/Q79 residues. The Q92S mutation slightly increases the probability of hydrogen bonding for both residues.

Discussion and Conclusion

This study examined the potential impact of the disease-causing mutation Q92R on the ability of NPC1(NTD) to bind CLR, to store it in the ligand-binding pocket, and to release it for further transfer across the membrane. As a control, we analyzed another mutation at the same position, Q92S, which was predicted to be benign by comprehensive evolutionary analysis^{28–31}. Residue Q92 is located at the N-site of $\alpha 3$ helix, within S-opening that is known to be the CLR entrance/exit point in NPC1(NTD)^{11,13}. The comprehensive study of the effect of Q92R mutation on the functionality of NPC1(NTD) and comparing it with the effect of benign Q92S mutation served two purposes – provided insight into the role of Q92 residue in CLR-NPC1(NTD) binding process; and proposed the mechanism, by which Q92R mutation causes NPC1 disease.

We found that upon CLR binding the NPC1(NTD) does not significantly change its SASA, but it causes the rearrangement of contact residues in a way that promotes general structural rigidity. We revealed that the orientation of CLR within the ligand-binding pocket does not change over time – it is well aligned with $\alpha 3$ (at 5-degree angle) and constructs an angle of 30 degrees with $\alpha 4$. This strictly defined position is dictated by geometry of the ligand-binding pocket and by the physico-chemical characteristics of the neighboring residues. Furthermore, we noticed that 3 β -hydroxyl site of CLR is involved in two hydrogen bonds (with N41 and Q79) that significantly restricts its motion. In contrast, the CLR isooctyl site is not engaged in any specific interactions and thus is more mobile. We also examined the properties of S-opening – the known entrance/exit point for CLR during NPC1(NTD)-CLR transfer. We showed that although the shape of S-opening is slightly different in NPC1(NTD) for apo state of the protein and with bound CLR, its capacity for the ligand transfer (in terms of inscribed radius) remains unchanged. It suggests that if S-opening needs to be enlarged for CLR entering NPC1(NTD)^{4,11,13,19,20} with the aid from NPC2, the CLR exiting from NPC1(NTD) should also be facilitated, e.g. by interaction between NPC1(NTD) and the NPC1 sterol-sensing domain^{9,21,22}. One of the most important properties of the S-opening is that it possesses a highly negative electrostatic potential that might be critical for NPC2-CLR binding and /or transfer. The latter statement can be strengthened by the fact that CLR transfer from NPC2 to NPC1(NTD) occurs at very specific ligand orientation – its polar (3 β -hydroxyl) group facing the interior of the domain^{8,11,16} that in fact can be reinforced by negative electrostatic potential of S-opening.

The study of the effects of Q92R and Q92S mutations on the protein structure revealed that the disease-causing mutation rigidify the protein in presence of CLR in the ligand-binding pocket, while the benign one makes it more flexible for both apo state and with bound CLR. This finding suggests that if S-opening is required to be flexible for structural rearrangements and subsequent CLR exit, the disease-causing mutation will interfere the process. Interestingly, both mutations affect HB/SB network in strands located at the opposite site of NPC1(NTD). This suggests that Q92 residue is important for the structure arrangement. Moreover, our data points out that the disease-causing mutation Q92R causes significant changes in the electrostatic potential distribution around S-opening. We found that substitution of polar Q92 with positively charged R in the vicinity of S-opening results in electrostatic neutralization of S-opening, which in turn could noticeably reduce

electrostatic attraction between NPC1(NTD) and NPC2 and/or the drugging force acting on CLR. In contrast, a substitution to other polar residue, serine (benign mutation), conserves or slightly increases the electrostatic polarity of S-opening that might indeed increase both NPC1(NTD)-NPC2 interaction and/or the speed of CLR transfer from NPC2 to NPC1.

Acknowledgments

Funding

This work was supported in part by the National Institutes of Health grant GM072285 to I.B.Z.

References

1. Simons K, Ikonen E. How cells handle cholesterol. *Science*. 2000;290(5497):1721–1726. [PubMed: 11099405]
2. Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. *Science*. 1986;232(4746):34–47. [PubMed: 3513311]
3. Goldstein J, Dana S, Faust J, Beaudet A, Brown M. Role of lysosomal acid lipase in the metabolism of plasma low density lipoprotein. Observations in cultured fibroblasts from a patient with cholesteryl ester storage disease. *Journal of Biological Chemistry*. 1975;250(21):8487–8495. [PubMed: 172501]
4. Infante RE, Wang ML, Radhakrishnan A, Kwon HJ, Brown MS, Goldstein JL. NPC2 facilitates bidirectional transfer of cholesterol between NPC1 and lipid bilayers, a step in cholesterol egress from lysosomes. *Proceedings of the National Academy of Sciences*. 2008;105(40):15287–15292.
5. Sleat DE, Wiseman JA, El-Banna M, et al. Genetic evidence for nonredundant functional cooperativity between NPC1 and NPC2 in lipid transport. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101(16):5886–5891. [PubMed: 15071184]
6. Cheruku SR, Xu Z, Dutia R, Lobel P, Storch J. Mechanism of cholesterol transfer from the Niemann-Pick type C2 protein to model membranes supports a role in lysosomal cholesterol transport. *Journal of Biological Chemistry*. 2006;281(42):31594–31604. [PubMed: 16606609]
7. Babalola JO, Wendeler M, Breiden B, et al. Development of an assay for the intermembrane transfer of cholesterol by Niemann-Pick C2 protein. *Biological chemistry*. 2007;388(6):617–626. [PubMed: 17552909]
8. Xu S, Benoff B, Liou H-L, Lobel P, Stock AM. Structural basis of sterol binding by NPC2, a lysosomal protein deficient in Niemann-Pick type C2 disease. *Journal of Biological Chemistry*. 2007;282(32):23525–23531. [PubMed: 17573352]
9. Li X, Wang J, Coutavas E, Shi H, Hao Q, Blobel G. Structure of human Niemann–Pick C1 protein. *Proceedings of the National Academy of Sciences*. 2016;113(29):8212–8217.
10. Gong X, Qian H, Zhou X, et al. Structural insights into the Niemann-Pick C1 (NPC1)-mediated cholesterol transfer and Ebola infection. *Cell*. 2016;165(6):1467–1478. [PubMed: 27238017]
11. Kwon HJ, Abi-Mosleh L, Wang ML, et al. Structure of N-terminal domain of NPC1 reveals distinct subdomains for binding and transfer of cholesterol. *Cell*. 2009;137(7):1213–1224. [PubMed: 19563754]
12. Zhao Y, Ren J, Harlos K, Stuart DI. Structure of glycosylated NPC1 luminal domain C reveals insights into NPC2 and Ebola virus interactions. *FEBS letters*. 2016;590(5):605–612. [PubMed: 26846330]
13. Li X, Saha P, Li J, Blobel G, Pfeffer SR. Clues to the mechanism of cholesterol transfer from the structure of NPC1 middle luminal domain bound to NPC2. *Proceedings of the National Academy of Sciences*. 2016;113(36):10079–10084.
14. Li X, Lu F, Trinh MN, et al. 3.3 Å structure of Niemann–Pick C1 protein reveals insights into the function of the C-terminal luminal domain in cholesterol transport. *Proceedings of the National Academy of Sciences*. 2017;201711716.

15. Carstea ED, Morris JA, Coleman KG, et al. Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. *Science*. 1997;277(5323):228–231. [PubMed: 9211849]
16. Infante RE, Radhakrishnan A, Abi-Mosleh L, et al. Purified NPC1 Protein II. Localization of sterol binding to a 240-amino acid soluble luminal loop. *Journal of Biological Chemistry*. 2008;283(2):1064–1075. [PubMed: 17989072]
17. Infante RE, Abi-Mosleh L, Radhakrishnan A, Dale JD, Brown MS, Goldstein JL. Purified NPC1 protein I. Binding of cholesterol and oxysterols to a 1278-amino acid membrane protein. *Journal of Biological Chemistry*. 2008;283(2):1052–1063. [PubMed: 17989073]
18. Deffieu MS, Pfeffer SR. Niemann–Pick type C 1 function requires luminal domain residues that mediate cholesterol-dependent NPC2 binding. *Proceedings of the National Academy of Sciences*. 2011;108(47):18932–18936.
19. Wang ML, Motamed M, Infante RE, et al. Identification of surface residues on Niemann-Pick C2 essential for hydrophobic handoff of cholesterol to NPC1 in lysosomes. *Cell metabolism*. 2010;12(2):166–173. [PubMed: 20674861]
20. Estiu G, Khatri N, Wiest O. Computational studies of the cholesterol transport between NPC2 and the N-terminal domain of NPC1 (NPC1 (NTD)). *Biochemistry*. 2013;52(39):6879–6891. [PubMed: 24001314]
21. Kuwabara PE, Labouesse M. The sterol-sensing domain: multiple families, a unique role? *TRENDS in Genetics*. 2002;18(4):193–201. [PubMed: 11932020]
22. Goldstein JL, DeBose-Boyd RA, Brown MS. Protein sensors for membrane sterols. *Cell*. 2006;124(1):35–46. [PubMed: 16413480]
23. Vanier MT. Niemann-Pick disease type C. *Orphanet journal of rare diseases*. 2010;5(1):16. [PubMed: 20525256]
24. Garver WS, Francis GA, Jelinek D, et al. The National Niemann–Pick C1 disease database: report of clinical features and health problems. *American Journal of Medical Genetics Part A*. 2007;143(11):1204–1211.
25. Ribeiro I, Marcão A, Amaral O, Miranda MS, Vanier MT, Millat G. Niemann-Pick type C disease: NPC1 mutations associated with severe and mild cellular cholesterol trafficking alterations. *Human genetics*. 2001;109(1):24–32. [PubMed: 11479732]
26. Sun X, Marks DL, Park WD, et al. Niemann-Pick C variant detection by altered sphingolipid trafficking and correlation with mutations within a specific domain of NPC1. *The American Journal of Human Genetics*. 2001;68(6):1361–1372. [PubMed: 11349231]
27. Mavridou I, Dimitriou E, Vanier MT, et al. The Spectrum of Niemann-Pick Type C Disease in Greece In: *JIMD Reports*, Volume 36 Springer; 2017:41–48. [PubMed: 28105569]
28. Adebali O, Reznik AO, Ory DS, Zhulin IB. Establishing the precise evolutionary history of a gene improves prediction of disease-causing missense mutations. *Genetics in Medicine*. 2016;18(10).
29. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*. 2013;7 20.21–27.20.41. [PubMed: 23315928]
30. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nature protocols*. 2016;11(1):1. [PubMed: 26633127]
31. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*. 2015;31(16):2745–2747. [PubMed: 25851949]
32. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33–38, 27–38. [PubMed: 8744570]
33. Diering GH, Numata M. Endosomal pH in neuronal signaling and synaptic transmission: role of Na⁺/H⁺ exchanger NHE5. *Frontiers in physiology*. 2014;4:412. [PubMed: 24454292]
34. Hu Y-B, Dammer EB, Ren R-J, Wang G. The endosomal-lysosomal system: from acidification and cargo sorting to neurodegeneration. *Translational neurodegeneration*. 2015;4(1):18. [PubMed: 26448863]
35. Casey JR, Grinstein S, Orlowski J. Sensors and regulators of intracellular pH. *Nature reviews Molecular cell biology*. 2010;11(1):50. [PubMed: 19997129]
36. Wang L, Zhang M, Alexov E. DelPhiPKa web server: predicting p K a of proteins, RNAs and DNAs. *Bioinformatics*. 2015;32(4):614–615. [PubMed: 26515825]

37. Zoete V, Cuendet MA, Grosdidier A, Michielin O. SwissParam: a fast force field generation tool for small organic molecules. *Journal of computational chemistry*. 2011;32(11):2359–2368. [PubMed: 21541964]
38. Phillips JC, Braun R, Wang W, et al. Scalable molecular dynamics with NAMD. *Journal of computational chemistry*. 2005;26(16):1781–1802. [PubMed: 16222654]
39. Li L, Li C, Sarkar S, et al. DelPhi: a comprehensive suite for DelPhi software and associated resources. *BMC biophysics*. 2012;5(1):9. [PubMed: 22583952]
40. Heinig M, Frishman D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic acids research*. 2004;32(suppl_2):W500–W502. [PubMed: 15215436]

Figure 1.
Cartoon presentation of WT NPC1(NTD) with bond CLR (yellow). Hydrophobic residues (ALIMWVF) are shown in orange; positively charged (RKH) in blue; negatively charged (DE) in red; polar (STQN and neutral EDHRK) in purple; C in pink; PG in white. Residues forming hydrogen bond with 3 β -hydroxyl group of CLR are marked with black arrows. Residues participating in stable HB/SB – with red arrows.

Figure 2. Residues RMSF mapped on the average WT structure in apo state (left) and with bound CLR (right). Red color corresponds to more flexible parts of the structure, while blue – to rigid ones.

Figure 3.

The change in RMSF for each residue in NPC1(NTD) in apo state (left) and with bound CLR (right) for both mutants Q92R (top) and Q92S (bottom) in comparison with the correspondent data for WT protein is mapped on the average protein structure. Parts of the structure with increased flexibility is shown in red, while with increased rigidity – in blue. The change in the sum of all residues RMSF is shown below for the whole protein, structured, and unstructured part of protein.

Figure 4. Distribution of SASA of NPC1(NTD) for WT(upper) and MTs Q92R (middle) and Q92S (bottom) in apo state (open circles) and with bound CLR (solid circles) during production phase of MD simulations.

Figure 5.

Frequency of HB/SB formation for WT (left) and MTs – Q92R (middle), Q92S (right) in apo state (top, open circles) and with CLR (bottom, close circles) for stable, non-neighboring pair residues. OX- and OY-axis represent pair of residue IDs, the size of circles reflects the value of frequency.

Figure 6.
Cartoon presentation of NPC1(NTD) with bound CLR (green). Helices $\alpha 3$, $\alpha 7$, and $\alpha 8$ are in orange. Residues L89, N198, and L176 are marked in red. Distances between those residues are labeled.

Figure 7.

Geometric characteristics of S-opening in NPC1(NTD) for WT and both MTs in apo state (light grey) and with bound CLR (dark grey). The average distance was monitored between CA atoms of residues pairs 1: L89-N198 (upper left); 2: L89-L176 (upper right); 3: N176-L198 (bottom left). At the bottom right is the average diameter of the sphere that can be inserted in triangle made of CA atoms of residues (L89, L176, N198). Standard error is shown as error bars. Numbers above show the SD of the average.

Figure 8. Electrostatic potential of the average structure NPC1(NTD) mapped on the protein structure, displayed as forcefield lines. S-opening is shown in green circle. Positive electrostatic field lines are shown in blue, whereas negative ones – in red. Data is shown for WT (left), Q92R (middle), and Q92S (right) MTs in apo state. Helices α_3 , α_7 , and α_8 are shown in yellow.

Figure 9.

Orientation of CLR in the ligand-binding pocket of WT (circle) and MTs Q92R (up triangle) and Q92S (down triangle) calculated as the angle between vectors: 1) atoms C17; C3 of CLR and CA atoms of residues (86;78) of helix α_3 (left); and 2) atoms C17; C3 of CLR and CA atoms of residues (105;89) of helix α_4 (right). The inserted picture is cartoon presentation of the NPC1(NTD) (light grey) with bound CLR (dark grey). Atoms used for CLR orientation calculations and correspondent distances are marked in black.

Figure 10.

Location of CLR in relation to $\alpha.3$ helix of NPC1(NTD) defined by two distances: between C3 atom of CLR and CA atom residue 78; and between C17 atom of CLR and CA atom residue 89. Inserted figure shows cartoon presentation of the protein (light grey) with bound CLR (dark grey) with marked distances and involved atoms in black.

Figure 11. Residues forming the ligand-binding pocket. OY-axis corresponds to the frequency with which the side chains of residues (on the OX axis) are within 5 Angstroms from CLR.

Table 1.

The average number of native contacts in NPC1(NTD) for WT and MT in apo state and with bound CLR. The standard deviation is provided.

	Mean	SD	Mean	SD
	apo		CLR	
WT	453.6	13.2	463.7	13.7
Q92R	455.6	12.1	459.8	11.4
Q92S	450.0	13.2	454.0	13.1

Table 2.

The average number of HB/SB in NPC1(NTD) for WT and MT in apo state and with bound CLR. The standard deviation is provided.

	Mean	SD	Mean	SD
	apo		CLR	
WT	13.1	2.4	12.13	2.3
Q92R	12.3	2.4	11.37	2.4
Q92S	11.9	2.5	11.81	2.6

Table 3.

The average number of hydrophobic interactions in NPC1(NTD) for WT and MT in apo state and with bound CLR. The standard deviation is provided.

	Mean	SD	Mean	SD
	apo		CLR	
WT	120.2	5.8	123.9	5.3
Q92R	116.0	5.3	123.7	4.6
Q92S	116.1	5.7	119.1	5.0

Table 4.

Residues of NPC1(NTD) involved in HB formation with CLR in WT protein and in Q92R and Q92S mutants.

Residue	Frequency, %		
	WT	Q92R	Q92S
N41	25.1	26.3	28.9
Q79	18.4	16.2	20.8