Journal of Clinical and Nursing Research, 2024, Volume 8, Issue 8
http://ojs.bbwpublisher.com/index.php/JCNR
Online ISSN: 2208-3693
Print ISSN: 2208-3685

Effect of *PKD1L1* Mutation on its Interaction with *PKD2* to Cause Situs Inversus Totalis

Fatima Haroon[1]*, Salma Saeed Khan[2], Assad Ullah[3]

[1]Institute of Molecular Biology and Biotechnology (IMBB), the University of Lahore, Pakistan
[2]Biomedical and Life Sciences, Kohsar University Murree, Murree 47150, Punjab, Pakistan
[3]Department of Biosciences, COMSATS University Islamabad, Park Rd, Islamabad Capital Territory 45550, Pakistan

Corresponding author: Fatima Haroon, fatimaharoon322@gmail.com

Copyright: © 2024 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Situs inversus totalis (SIT) is a rare homozygous recessive disease caused by the mutation in *PKD1L1*, which is required for normal interaction with *PKD2* and leads to different complications such as respiratory disorders, brain disorders and even obesity. The present study was designed to find out the mutational effect on the binding of *PKD2* with mutated *PKD1L1*, which leads to SIT. The three-dimensional (3D) structure of wild type and mutated *PKD1L1* was predicted with > 90% confidence using different online tools. The different online tools that were employed were SWISS-MODEL, Phyre2 (normal & intensive) and i-TASSER. To compute the physiochemical properties of *PKD1L1* (wild & mutated) and *PKD2 in silico* approaches were employed using the ExPASy ProtParam tool. Physicochemical properties such as molecular weight, isoelectric point, the total number of negatively and positively charged residues, extinction coefficient, half-life, instability and aliphatic index, grand average of hydropathicity, and amino acid percentage were calculated. A lot of variability was observed in these parameters among *PKD1L1* and *PKD2*, which accounted for diversification in their functional properties. The theoretical pI points showed that *PKD1L1* (whole) is more basic with 6.64 pI compared to its first chain TOPO_DOM (amino acids from 1–1748) has a pI of 5.62 which means it is basic while *PKD2* have the lowest pI point of 5.34. Docking was performed using the PatchDock and ClusPro online tools.

Keywords: SIT; *PKD1L1*; *PKD2*; SWISS-MODEL; ClusPro

Online publication: September 3, 2024

1. Introduction

Situs inversus totalis is a rare X-linked recessive disorder in relevance to congenital heart disease (CHD) that exhibits complete reversal of thoracic and abdominal viscera. People suffering from SIT could lead a normal, healthy life, but in some medical emergencies, the complete mirror imaging of the viscera generates a myriad of functional malformations in the cardiac, spine, and airway passages, which, if accompanied together, could increase the difficulty levels for anesthesiologists [1]. The incidence of occurrence through various prevalence reports varies comprehensively and has been estimated at 1/16,000–1/33,000 births [2]. Individuals with situs inversus have a roughly 50% probability of developing congenital heart disease. However, the frequency of

congenital heart defects in situs inversus totalis is modest, ranging from 2% to 5% [3]. An abnormality of the situs may manifest as a distinct symptom of a condition encompassing a broader spectrum of abnormalities. There are significant signs that situs inversus in association with Kartagener's syndrome has a co-relation with mucociliary failure and is, therefore, classified as an immotile ciliary syndrome- a rare, ciliopathic genetic disorder [4]. It is present in 25% of patients with SIT and is associated with an increased risk of developing respiratory complications. In humans, the genetic cause of situs inversus totalis is extremely heterogeneous [5]. Genetic causes account for less than 20% of cases and their spectrum is highly heterogeneous. Several genes involved in the formation of the left-right axis have been identified through animal studies [6]. These genes demonstrate a complicated genetic cascade of left-right differentiation that occurs before the manifestation of physical asymmetry [7]. The majority of situs abnormalities are caused by random mutations, and the disorder can be caused by a variety of distinct genetic variables or genes in various persons or families. Environmental and probabilistic factors may also be involved, given that in a significant number of instances, no clear monogenetic basis for their disease can be identified [8].

The implication of polycystic kidney disease protein 1-like 1 (*PKD1L1*) in the establishment of laterality in humans has been observed. Upon whole-exome sequencing (WES), homozygous mutations in *PKD1L1* have been recently identified [9]. Major genes contributed at the outset of SIT have shown their association with NODAL/TGFβ signaling, hedgehog signaling, and mono-cilia functions [10]. It is vital for the movement of fluids and particles across epithelial surfaces, and cilia play critical roles in a variety of signal transduction pathways. For cilia to develop, a complex genetic program must be maintained [11]. Mutations in genes that contribute to ciliogenesis have been attributed to ciliopathies, including primary ciliary dyskinesia (*DNAH5*, *DNAH11*, *DNAI1*). Ciliopathies are characterized by a wide range of disease and developmental mutant features that can be organ-specific or have broad pleiotropic effects [12]. Initiation of left-right asymmetry in vertebrates begins during early embryonic development, entailing nodal cilia and a network of signaling transduction cascades [13].

In humans, *PKD1L1* is a paralog of *PKD1* (polycystin-1 [MIM: 601313]), which belongs to the polycystin cation channel family 1 (PC-1) and consists of 58 exons that cover 187 kb of genomic DNA [14]. *PKD2* contains two immunoglobulins (Ig)-like polycystic kidney disease (*PKD*) domains in the N-terminal extracellular region, a small receptor for egg jelly (REJ) domain, a GPS motif, 11 putative transmembrane segments, a polycystin-1, lipoxygenase, alpha-toxin (PLAT) domain, and a C-terminal intracellular coil (CC). The expression of *PKD1L1* in the adult and fetal heart, as well as in the testis, has been established by RNA dot-blot analysis of numerous human tissues [15]. It has been hypothesized that *PKD1L1* and *PKD2* form complexes within cilia that sense nodal flow and are involved in the left-sided activation of the nodal signaling cascade through their C-terminal coiled-coil (CC) domains in the node. Within the extracellular domains at the N-terminal, specifically in the domains like REJ (receptor egg jelly), GPS (G-protein coupled receptor proteolytic site), extracellular domains between transmembrane domains, and the intracellular CC domain, bi-allelic missense, splice site, and nonsense variations have been identified [13]. In addition, it has been discovered that mutations in the CC domain disrupt the interaction between *PKD1L1* and the *PKD2* protein. *PKD1L1* is required for nodal activation to be restricted to the left side downstream of nodal flow [16]. It has also been demonstrated in many types of research that:

(1) *PKD1L1* is required for nodal activation to be restricted to the left side downstream of nodal flow;

(2) *PKD1L1* can mediate fluid flow-induced Ca-signal response in vitro, suggesting that it may play an analogous role in vivo in the elicitation of nodal flow response [17].

PKD1L1 is a protein that is a constituent in the ciliary calcium channel and effectively controls the concentration of calcium without having any influence on the cytoplasmic concentration of calcium. It has been shown that *PDKD1L1* and *PKD2* interact with each other and are involved in L-R patterning in mouse, medaka

fish, and zebrafish models [18]. In addition, *PKD1L1* aids in the detection of left/right patterning through cilia by forming a complex with another protein named *PKD2*. However, in humans, *PKD2* heterozygous mutations induce polycystic kidney disease 2 (MIM: 613095), a type of kidney disease caused by anomalies of the renal primary cilium that affects the function of the kidney [19]. Individuals with laterality problems have also been documented to have heterozygous mutations in the *PKD2* gene. However, this is a rare occurrence. Asymmetric gene expression in embryonic nodal cilia causes left-sided laminar flow, which facilitates left-right asymmetry establishment, which is aided by *PKD1L1* as it is involved in the specification of the left/right axis downstream of nodal flow [20].

By conducting studies on orthologous genes in medaka fish and mice, it has been evident that loss of function mutation in *PKD1L1* has its role in laterality defects. Moreover, the interaction between *PKD1L1* and *PKD2* gets disturbed due to mutation in the CC domain of *PKD1L1* [21]. Studies showed that SIT is a recessive genetic disorder, although all SIT patients do not have PCD, in some patients with SIT, Primary Ciliary Dyskinesia (PCD) was also observed. In those SIT patients in which PCD is also present, the body also developed into a literal position due to brain development mechanisms and, therefore, is mostly left-handed [22]. The results from one of the precious papers showed that both SIT and PCD are recessive mutations linked to genes. The result showed that two SIT patients, but non-PCD had a mutation in PCD genes. This showed that due to mutation, they were non-PCD and, therefore, cases of absence of PCD in a few SIT-positive patients [23]. The result of the same study showed that one patient with SIT but was non-PCD had a recessive mutation in the *PKD1L1* domain which confirms that the cause of SIT is because of the presence of a recessive mutation in the *PKD1L1* gene [24]. Studies showed that Situs inversus totalis (SIT) is due to mutations in the axonemal heavy chain DNAH11 (axonemal heavy chain dynein type 11), which is linked with primary ciliary dyskinesia disorder. The studies showed that in PCD-positive patients, the various chains of dynein arms of the axoneme are absent or defective [25]. This present study cohort will try to explore the cause of situs inversus totalis which might be due to mutated *PKD1L1* and mutated *PKD1L1* do not interact properly with the *PKD2* and results in SIT.

2. Methodology

The amino acid sequences of human *PKD1L1* and *PKD2* were retrieved from the UniProt database. Using standard parameters, the homologous sequence for *PKD1L1* was browsed and selected from the Orthologous Matrix (OMA) browser and visualized using the open-source software Aliview (Larsson, 2014). Later on, multiple sequence alignment (MSA) of *PKD1L1* and selected orthologues were produced using Clustal Omega from EMBL-EBI Web Services. The PDB protein files were visualized using the PyMol and Chimera software.

The three-dimensional (3D) structures of wild and mutated human *PKD1L1* and *PKD2* models were predicted by using different available online tools, the SWISS MODELING approach, which is an automated homology modeling online server for the prediction of protein structure. In SWISS MODEL, the selection of the top model is based on the highest scores of the resolved structure. The highest score represents the accuracy and the stereo-chemical quality of the predicted model. The FASTA format from UniProt was submitted and by using the BLAST and HHblits methods, a template library was created for the query protein. The SWISS-MODEL predicted the structure based on the evolutionary-related data. The ranking of the template models is based on the quality of the predicted models, which is based on Global Model Quality Estimate (GMQE) and Quaternary Structure Quality Estimate (QSQE) values.

Phyre2 predicts the 3D model of query protein by advanced remote homology detection methods. Besides 3D model prediction, Phyre2 also predicts the ligand binding sites in the query protein and the effect of

mutation at specific sites. The analysis of mutation is shown in the form of a graph, which represents the effect of mutations at a specific position of the query protein sequence. In Phyre2, SuSPect method is applied for the determination of conservation in the sequence. The third method used for structure prediction was by i-TESSAR. It is an automated 3D structure prediction and structure-based function annotation server. It generates the full-length structural models from multiple threading alignments after the submission of the query protein sequence. After the structural assembly simulation, refinement of the generated model is performed. From the known protein function database (PDB) the biological function of protein is predicted, which includes the ligand binding site in the query protein. Both the selected mutated structures (Cys-Ser) were also tried to design by altering the residual positions within the structure. Mutated structures were analyzed through multiple online tools such as Maestro pred, proven respectively and the effect of mutated amino acids was checked on the overall protein stability and lethality.

Protein-protein docking of mutated *PKD1L1* with *PKD2* was assisted by using different online protein-protein docking servers such as ClusPro and PatchDock. To generate rigid-body docking conformations, PatchDock was used with default search parameters. The ID of *PKD1L1*, Q8TDX9 was retrieved from Uniport and was downloaded in the FASTA format. Submitted in the ExPASy ProtParam tool for the analysis of Physio-chemical properties. Same for *PKD2* the sequence was submitted for the calculations of Physio-chemical properties.

3. Results

The sequence of *PKD1L1* and *PKD2* was retrieved by using the UniPort. The *PKD1L1* consists of 2849 amino acids. Sequencing information showed that there are five domains in the *PKD1L1* protein. These are *PKD1* (508 to 590), *PKD2* (592 to 673), REJ (674 to 1571), GPS (1688 to 1734), and PLAT (1796 to 1913), respectively (**Figure 1**).

Table 1. Different variants of *PKD2*

Sr. No	Variant ID	Position	Mutation	Description
1	VAR_058822	306	R > Q	In *PKD2*; dbSNP:rs9909329471
2	VAR_058823	322	R > Q	In *PKD2*; dbSNP:rs145877597
3	VAR_058824	322	R > W	In *PKD2*; dbSNP:rs1553925453
4	VAR_009195	414	W > G	In *PKD2*, affects channel activity as it can abolish channel currents induced by the gain-of-function artificial mutant P-604
5	VAR_058825	420	R > G	In *PKD2*, affects channel activity as it can abolish channel currents induced by the gain-of-function artificial mutant P-604
6	VAR_011074	479	> missing	In *PKD2*; somatic mutation
7	VAR_011075	504–512	> missing	In *PKD2*; somatic mutation
8	VAR_058827	511	D > V	In *PKD2*; loss of function in the release of Ca^{2+} stores from the endoplasmic reticulum; no effect on location at the endoplasmic reticulum; affects channel activity as it can abolish channel currents induced by the gain-of-function artificial mutant P-604; dbSNP:rs121918043
9	VAR_011076	684	> missing	In *PKD2*; somatic mutation; abolishes channel currents induced by the gain-of-function artificial mutant P-604; dbSNP:rs1578144872
10	VAR_058830	807	R > Q	In *PKD2*; dbSNP:rs147654263

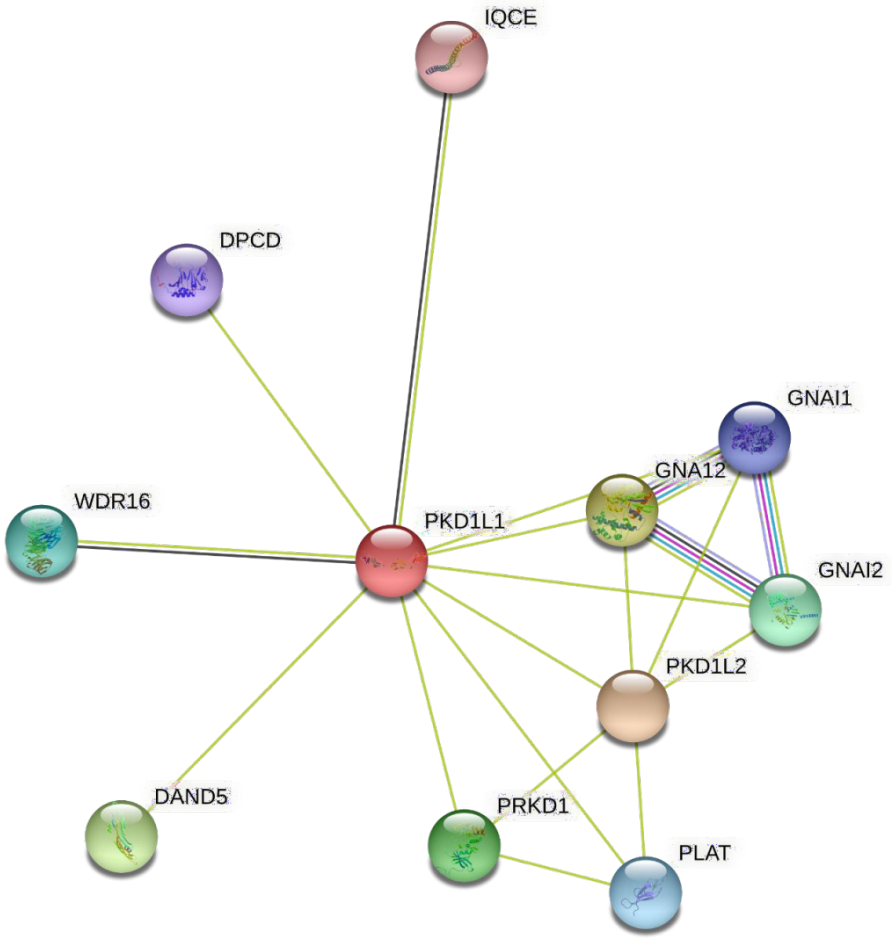

Figure 1. The different proteins which are produced by a single *PKD1L1* protein-coding gene locus and also represent the different functional partners.

PKD2 polycystine-2 consists of 968 amino acids and one domain only. *In silico* analysis showed that mutations at different positions in its variants lead to disease conditions. Ten different variants of *PKD2* are predicted, in which different types of mutations at specific amino acids are identified (shown in **Table 1**).

After submitting the sequence of *PKD1L1* in OMA for retrieval of the homologous orthologous group, the result showed that no orthologous group (HOG) was retrieved for *PKD1L1*. However, two genomes were retrieved by OMA for *PKD1L1* (as shown in **Table 2**).

For further determination of orthologues, other tools were tried to be used. An online available tool that was DIOPT (Drosophila RNAi Screening Center Integrative Ortholog Prediction Tool) ortholog finder was used to retrieve the orthologue of *PKD1L1*. The results obtained by this tool identify the orthologues (as shown in **Table 3**) for *PKD1L1* by using different approaches. The sensitivity of this tool is more compared to other tools.

Table 2. The list of genomes homologous to *PKD1L1*

Sr. No	Code/Name	Type	NCBI Taxon ID	Number of sequences/ Genes	Scientific name	Code	Domains
1	PARL1	Extant	402281	3579	Parvibaculum lavamentivorans (strain DS-1/ DSM 13023 / NCIMB 13966)	PARL1	B
2	BIFL1	Extant	565040	1972	Bifidobacterium longum subsp. infantis (strain 157F)	BIFL1	B

Table 3. Orthologs of *PKD1L1* predicted by DRSC integrative ortholog prediction tool

Human Gene	Human Gene ID	HG NCI D	Species 2	HGN CID	Huma n symbol	DI OP T Sc ore	Weighted Score	Ran k	Best Score e	Best Score Reverse	Prediction derived from
PKD1L1	168507	18053	Huma n	9015	PKDR EJ	4	3.94	High	Yes	Yes	Compara, eggNOG, OrthoDB, Panther
PKD1L1	168507	18053	Huma n	21716	*PKD1 L3*	4	3.94	High	Yes	Yes	Compara, eggNOG, OrthoDB, Panther
PKD1L1	168507	18053	Huma n	21715	*PKD1 L2*	3	3.04	Low	No	No	Compara, OrthoDB, Panther
PKD1L1	168507	18053	Huma n	9012	*PKD2 L2*	3	2.93	Low	No	No	Compara, eggNOG, Panther
PKD1L1	168507	18053	Huma n	9009	PKD2	3	2.93	Low	No	No	Compara, eggNOG, Panther
PKD1L1	168507	18053	Huma n	9011	*PKD2 L1*	3	2.93	Low	No	No	Compara, eggNOG, Panther
PKD1L1	168507	18053	Huma n	26521	*LOXH D1*	3	2.86	Mode rate	No	Yes	Compara, eggNOG, RoundUp
PKD1L1	168507	18053	Huma n	9008	PKD1	2	1.91	Mode rate	No	Yes	eggNOG, OrthoDB

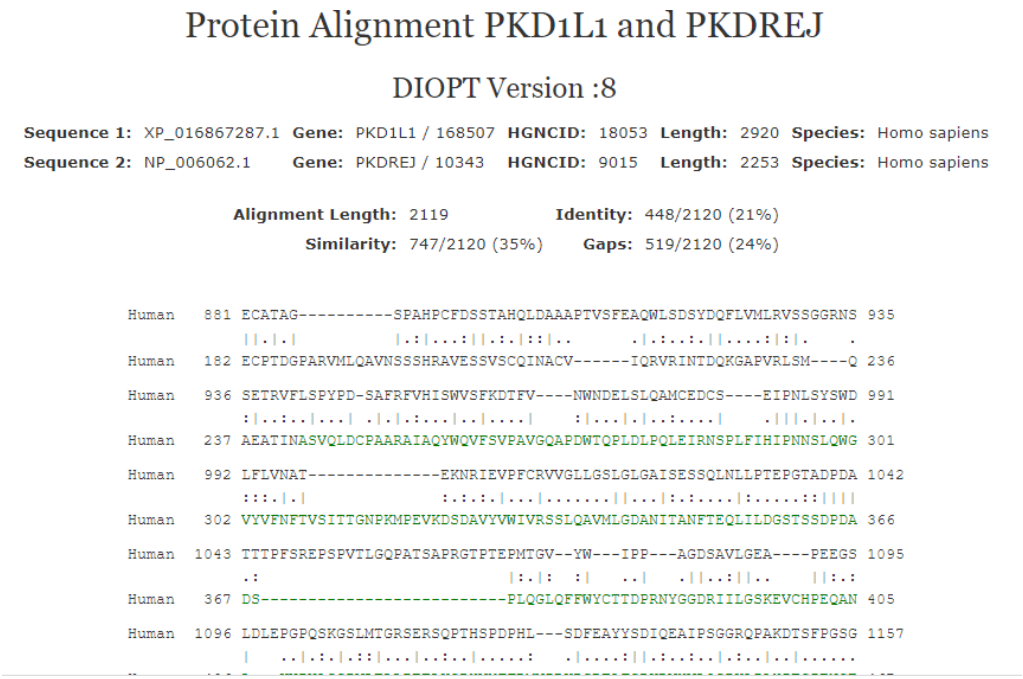

Figure 2. Protein alignment of *PKD1L1* with its orthologue predicted by DIOPT Tool version 8.

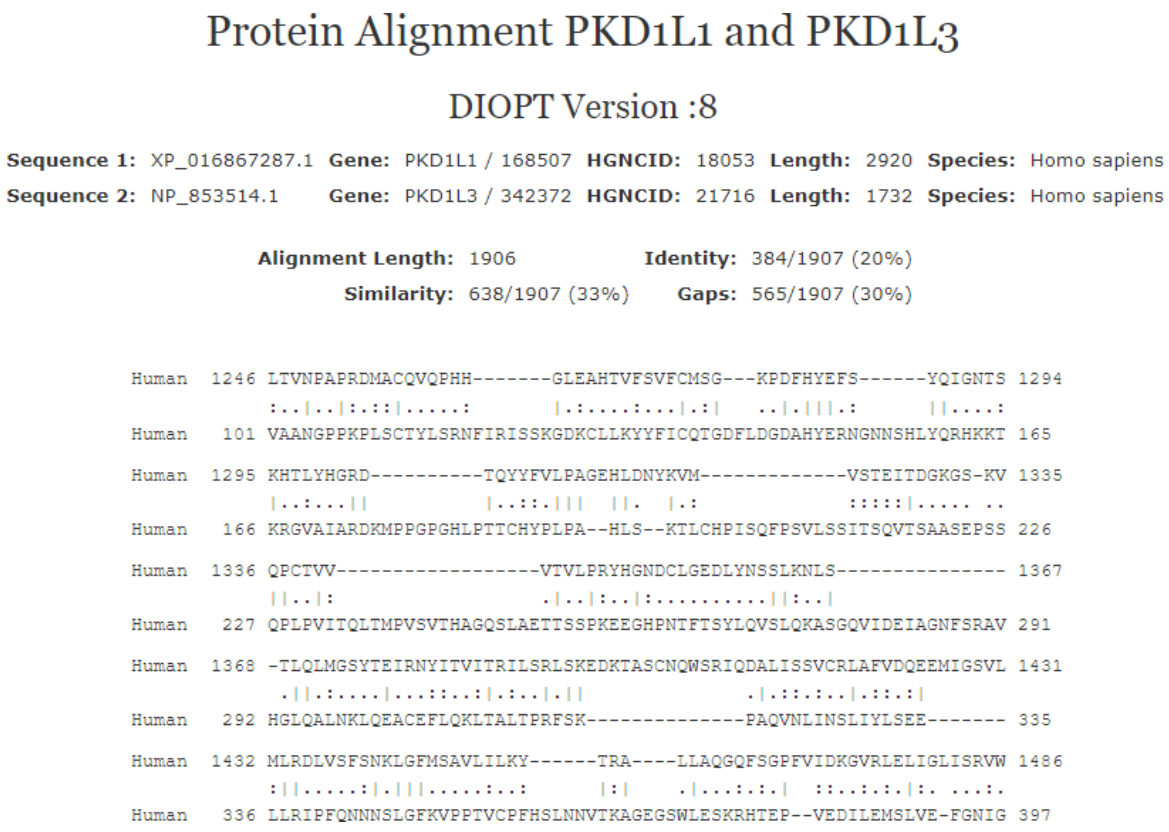

Figure 3. Protein alignment of *PKD1L1* with its moderate orthologue predicted by DIOPT Tool version 8.

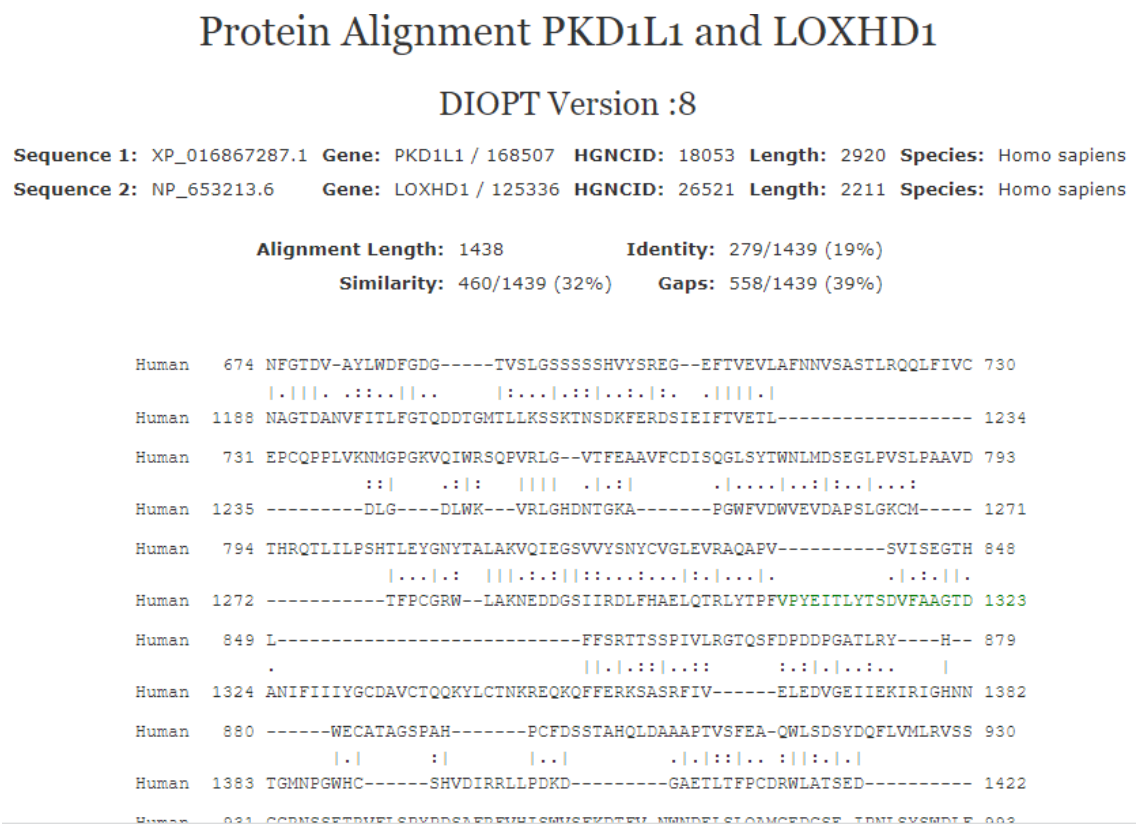

Figure 4. Protein alignment of *PKD1L1* with its low-ranked orthologue predicted by DIOPT Tool version 8.

The orthologues retrieved by DIOPT version 8 for *PKD1L1* were also aligned with its different orthologues, and the division orthologues were divided into high, moderate, and low-ranked functional similarity. The 3D structure of *PKD1L1* was predicted by using different online available tools. The predicted model for *PKD1L1* protein is determined by SWISS-MODEL, Phyre2 and i-TASSER, respectively. The predicted models were evaluated by Ramachandran plot analysis. The first model of *PKD1L1* was predicted by Phyre2 and SWISS-MODEL (**Figure 5**) and models were further validated by Ramachandran plot analysis.

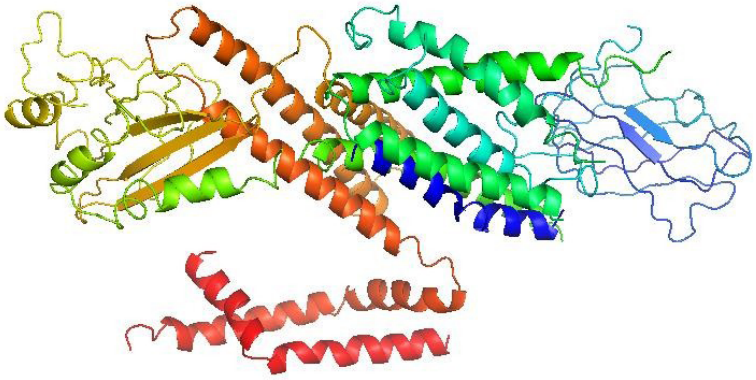

Figure 5. The 3D model of *PKD1L1* predicted by Phyre2 (whole protein sequence).

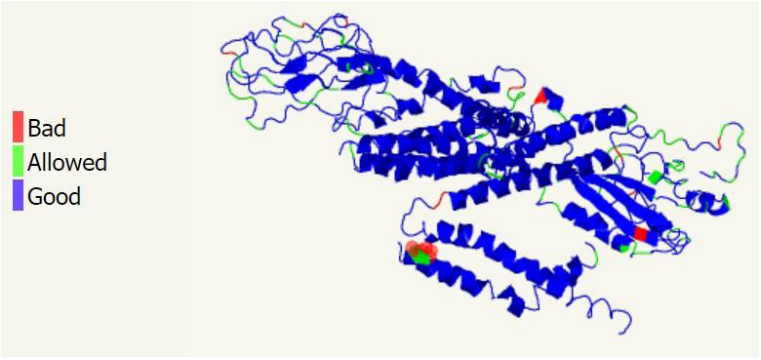

Figure 6. Ramachandran Analysis for the validation of predicted structure by Phyre2.

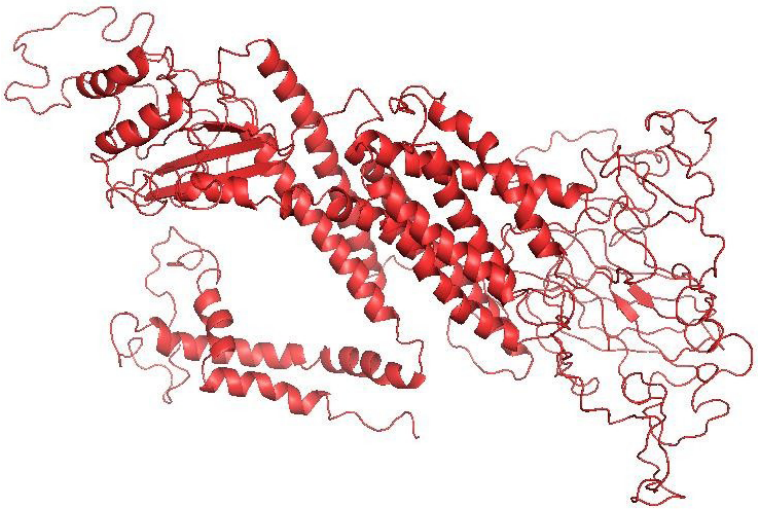

Figure 7. The 3D model of *PKD1L1* predicted by SWISS-MODEL.

SWISS-MODEL also predicted the 3D structure of *PKD1L1*, as shown in **Figure 6**. The predicted model was further validated by the Ramachandran plot analysis generated by the SWISS-MODEL tool (**Figure 7**).

The *PKD1L1* sequence was further submitted to SWISS-MODEL and Phyre2 for the prediction of the model, which contained the mutated residual region. In the case of Phyre2, only selected amino acid residues (1010–1724) were submitted and processed for the generation of the model. Thus, 3D structure was resolved. The zoom-in structure of *PKD1L1* shows the position of amino acids cysteine 1691. The validation of the predicted result was done by Ramachandran plot analysis which was generated by Phyre2. The residues that were present in the allowed and disallowed regions according to the Ramachandran analysis were represented in blue (good), green (allow) and red (disallowed) colors, respectively (**Table 4**).

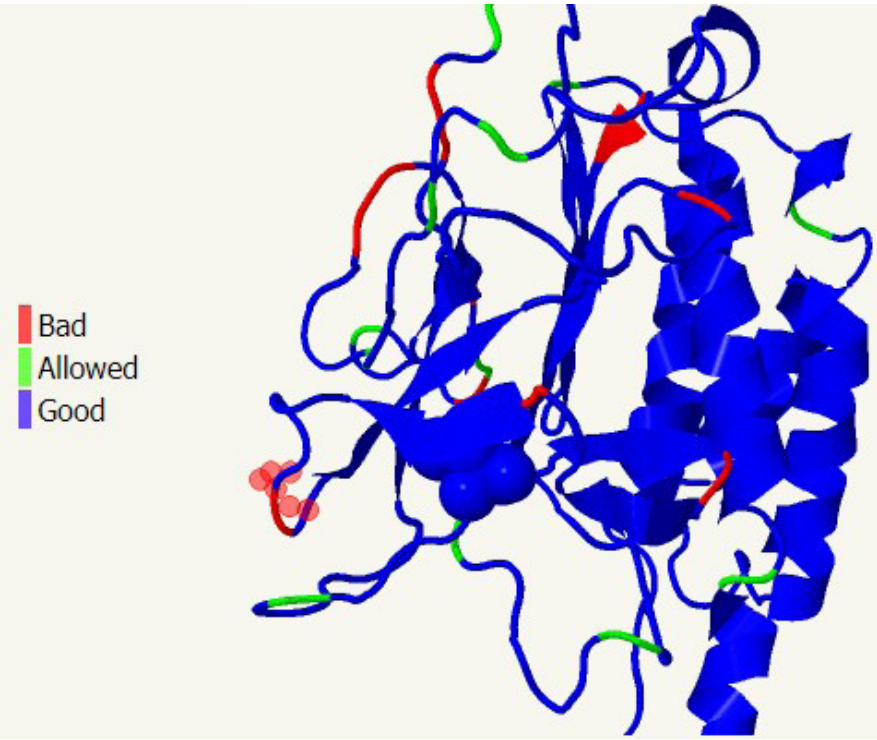

Figure 8. 3D view of *PKD1L1* (residue from 1010 to 1724) illustrating the Ramachandran regions.

Red regions show the disallowed region, green for allowed and blue for good, respectively. Blue clustered balls depicting the amino acid 682 (1691 Cyst).

Table 4. Ramachandran Plot analysis of the model of *PKD1L1* (residue 1010–1724)

Residue number	Allowed/Disallowed	Color
394	Allowed	Blue
434	Allowed	Blue
455	Allowed	Blue
476	Allowed	Blue
498	Disallowed	Red
499	Disallowed	Red
507	Allowed	Blue
511	Disallowed	Red

Table 4 (Continued)

Residue number	Allowed/Disallowed	Color
512	Allowed	Blue
514	Disallowed	Red
518	Disallowed	Red
519	Disallowed	Red
524	Allowed	Blue
529	Allowed	Blue
532	Allowed	Blue
546	Disallowed	Red
547	Disallowed	Red
563	Allowed	Blue
570	Allowed	Blue
579	Disallowed	Red
602	Allowed	Blue
687	Disallowed	Red
694	Disallowed	Red
702	Disallowed	Red

Using the SuSPect tool by Phyre2, the mutational sensitivity of the model was predicted. Positions are colored according to the average effect of the 20 possible mutations at that position. The result of the analysis showed that the target amino acids, which are at 682 positions (1691 Cys), have shown high sensitivity with a high score of 9 for mutational activity. The amino acid in which mutation leads to SIT is shown in the form of orange balls clusters together (**Figure 9**). The accuracy of the results is > 80% on a large benchmark test set.

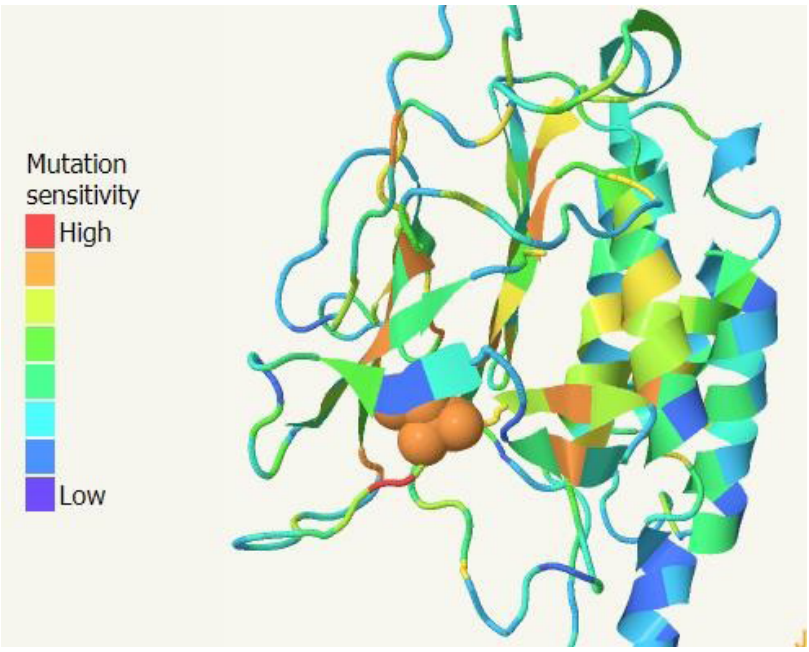

Figure 9. Mutational sensitivity analysis of *PKD1L1* (1010–1724).

For the determination of the active sites in the protein large pocket was tried to predict by using the pocket2 program by Phyre2. The active site for pockets predicted in the protein is shown in wireframe mode and shown in red color (**Figure 10**).

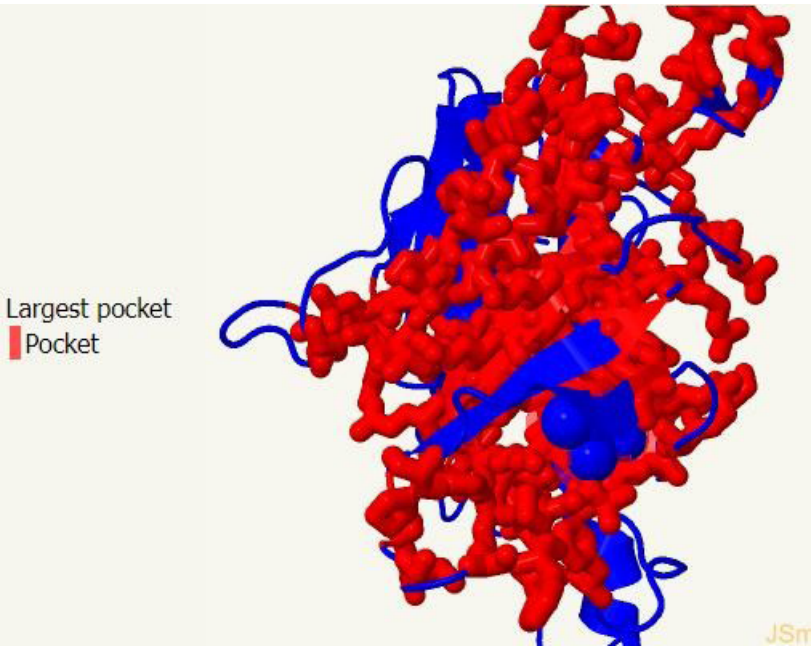

Figure 10. Pocket detection by pocket for *PKD1L1* (1010–1724).

The result analysis for conservation concerning function and structure showed that all residues in a protein are not equally important. Some are essential for the proper structure and function of the protein, whereas others can be readily replaced. Phyre2 analyses conservation in the *PKD1L1* sequence using an information-theoretic approach based on Jensen-Shannon divergence (**Figure 11**).

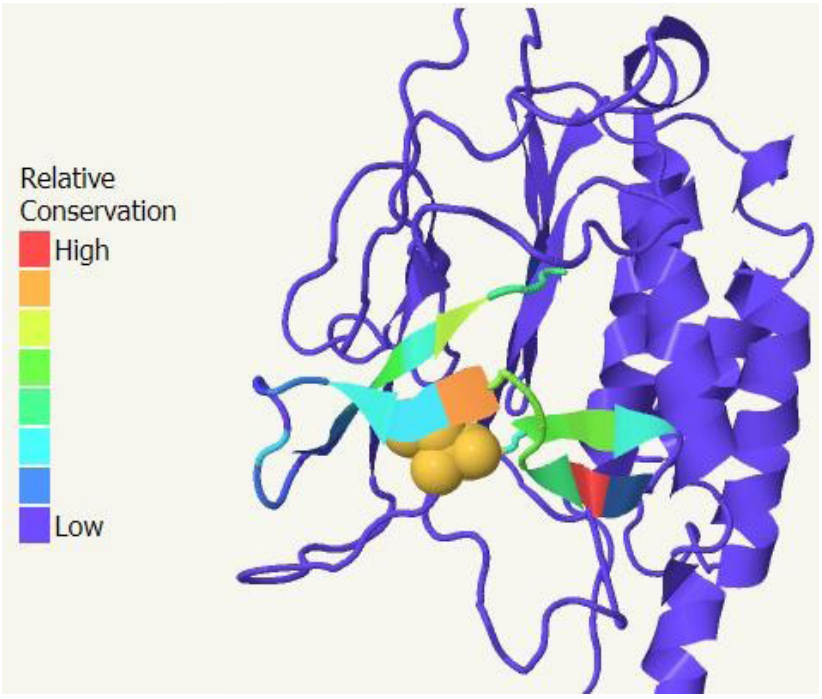

Figure 11. Analysis of conservation of function and structure in *PKD1L1*.

The quality of the predicted model was determined using the ProQ2 model quality assessment algorithm. ProQ2 uses support vector machines to predict the local and global quality of protein models. The result showed that the quality of the model is good (**Figure 12**).

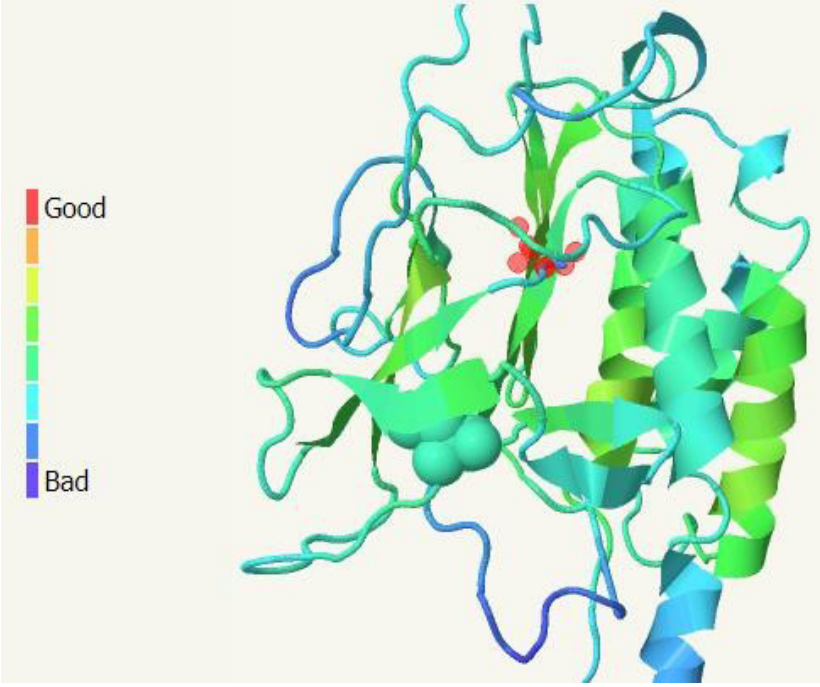

Figure 12. Quality assessment of *PKD1L1* (1010–1691).

The mutated sequence (at position 682(1796)) of *PKD1L1* was submitted in Phyre2, and a 3D model was predicted (**Figure 13**). Using the intense method of Phyre2, 16 templates were selected to model the *PKD1L1* protein based on heuristics to maximize confidence, percentage identity and alignment coverage. While 402 residues of *PKD1L1* were modeled ab initio, 19% of residues of *PKD1L1* (1010–1724) were modeled with > 90% confidence (**Figure 13**).

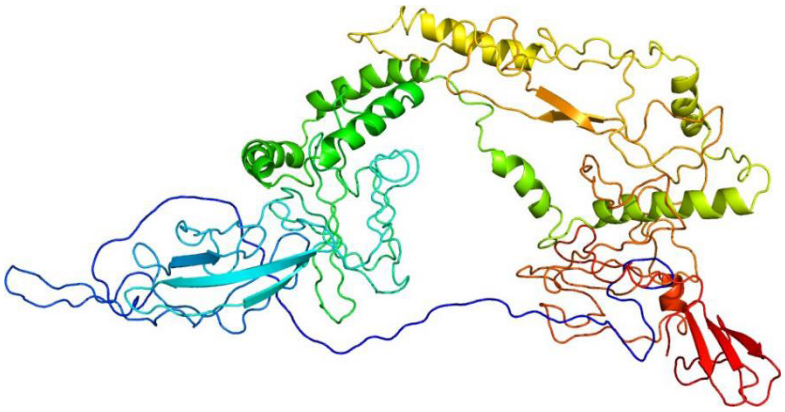

Figure 13. Mutated *PKD1L1* 3D view of the predicted structure generated by Phyre2 Intensive.

Using the i-TASSER MTD, the 3D model of *PKD1L1* was generated. The top five models were generated based on eTM-score, eRMSD and energy. The accuracy of the model is evaluated based on the estimated eTM-

score and eRMSD, which are calculated based on the significance of the structural analogous templates for domain model assembly, convergence parameters of the domain assembly simulations, satisfaction degrees of the inter-domain distances/interfaces, and the estimated accuracy of the individual domain model of *PKD1L1*. eTM-score is typically in the range of (0,1), where an eTM-score of higher value signifies a model with high confidence and vice-versa. The top-ranked model predicted by i-TASSER was taken as a finalized model of *PKD1L1*. Its eTM and eRMSD are 0.61 ± 0.15 and 21.0 ± 3.9, respectively. The most important thing to consider is the top-ranked model as a final model because it contains the amino acids of our interest, which are 1691 respectively.

According to the analysis results five domains were identified in *PKD1L1*. The first domain contained amino acids from 1 to 766 and is shown in red color, the second domain is from 767 to 1561 and depicted in blue color, and the third domain is from 1562 to 1681 and shown in green color. Fourth domain is from 1682 to 1738 and represented by cyan color and the fifth domain is from 1739 to 1890 and represented in violet-blue color as shown in **Figure 14**. The predicted model depicts that there are lots of coils and helices in the structure, while strands are very less. The amino acid 1691 is also forming the coil structure.

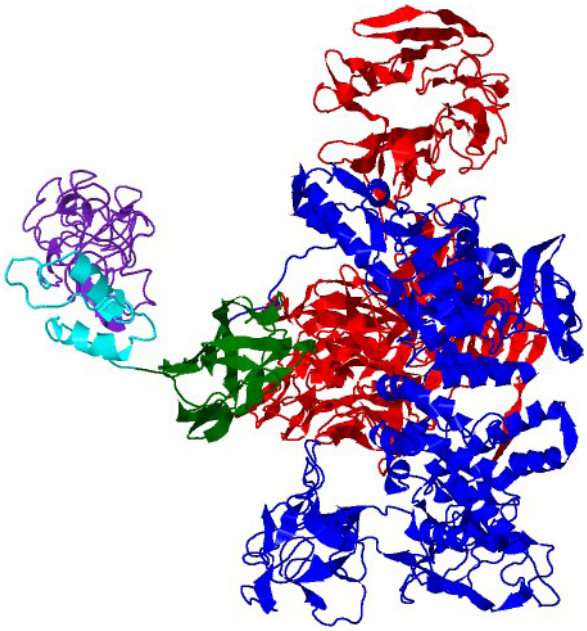

Figure 14. 3D Structure of *PKD1L1* modeled by i-TASSER -MTD tool.

The physicochemical properties of the *PKD1L1* were calculated using ExPASy tool, which include number of amino acids, molecular weight, theoretical isoelectric point (pI), total number of negatively and positively charged residues, atomic composition, formula, extinction coefficient, estimated half-life, instability index (II), aliphatic index (AI), total number of atoms in a chain, and grand average of hydropathicity (GRAVY) (**Table 6**). After the submission of Uniport ID in ExPASy, the whole protein sequence was divided into different chain lengths (**Table 5**).

Table 5. Division of whole *PKD1L1* protein sequence into different chain lengths

Sr. No.		Chain	Sequence	
1	FT	CHAIN	1–2849	Polycystic kidney disease protein 1-like 1
2	FT	TOPO_DOM	1–1748	Extracellular
3	FT	TRANSMEM	1749–1769	Helical
4	FT	TOPO_DOM	1770–1956	Cytoplasmic
5	FT	TRANSMEM	1957–1977	Helical
6	FT	TOPO_DOM	1978–1992	Extracellular
7	FT	TRANSMEM	1993–2013	Helical
8	FT	TOPO_DOM	2014–2135	Cytoplasmic
9	FT	TRANSMEM	2136–2156	Helical
10	FT	TOPO_DOM	2157–2174	Extracellular
11	FT	TRANSMEM	2175–2195	Helical
12	FT	TOPO_DOM	2196–2281	Cytoplasmic
13	FT	TRANSMEM	2282–2302	Helical
14	FT	TOPO_DOM	2303–2522	Extracellular
15	FT	TRANSMEM	2523–2543	Helical
16	FT	TOPO_DOM	2544–2562	Cytoplasmic
17	FT	TRANSMEM	2563–2583	Helical
18	FT	TOPO_DOM	2584–2616	Extracellular
19	FT	TRANSMEM	2617–2637	Helical
20	FT	TOPO_DOM	2638–2646	Cytoplasmic
21	FT	TRANSMEM	2647–2667	Helical
22	FT	TOPO_DOM	2668–2711	Extracellular
23	FT	TRANSMEM	2712–2732	Helical
24	FT	TOPO_DOM	2733–2849	Cytoplasmic
25	FT	DOMAIN	508–590	PKD 1
26	FT	DOMAIN	592–673	PKD 2
27	FT	DOMAIN	674–1571	REJ
28	FT	DOMAIN	1688–1734	GPS
29	FT	DOMAIN	1796–1913	PLAT
30	FT	REGION	970–1068	Disordered
31	FT	REGION	1081–1118	Disordered
32	FT	REGION	2023–2089	Disordered
33	FT	COMPBIAS	970–1009	Polar residues
34	FT	COMPBIAS	1047–1064	Polar residues

In the case of *PKD1L1*, the whole chain (1–2849 amino acids) and chain TOPO_DOM (amino acids from 1–1748) were selected and the physio-chemical properties were calculated. Same for *PKD2* the whole protein sequence was submitted, and the physiochemical properties were calculated (**Table 6**).

Table 6. Physiochemical properties calculation of *PKD1L1* (whole), Chain 1 & *PKD2* (whole)

Sr. No.	Physio- chemical properties	CHAIN 1-2849Polycystic kidney disease protein 1-like 1	Chain TOPO_DOM 1–1748 Extracellular	PKD2
1	Number of amino acids	2849	1748	805
2	Molecular weight	315434.93	192675.72	91981.97
3	Theoretical pI	6.64	5.62	5.34
4	Total number of negatively charged residues (Asp + Glu)	269	181	95
5	Total number of positively charged residues (Arg + Lys)	254	139	75
6	Atomic composition/ Formula	C14067H21897N3837O4146S135	C8557H13250N2306O2615S75	C4190H6403N1083O1202S25
7	Total number of atoms	44082	26803	12903
8	Extinction coefficients	415785	241430	146400
9	Estimated half-life	30 hours (mammalian reticulocytes, in vitro). > 20 hours (yeast, in vivo). > 10 hours (*Escherichia coli*, in vivo)	30 hours (mammalian reticulocytes, in vitro). > 20 hours (yeast, in vivo). > 10 hours (*Escherichia coli*, in vivo)	30 hours (mammalian reticulocytes, in vitro). > 20 hours (yeast, in vivo). > 10 hours (*Escherichia coli*, in vivo).
10	Instability index	45.80/unstable	46.85/unstable	46.96/unstable
11	Aliphatic index	84.64	80.29	90.56
12	GRAVY	-0.137	0.205	-0.136

Figure 15. Solution 1 by PatchDock analysis between *PKD1L1* (structure predictedby Phyre2 1010-1724, purple color) and *PKD2* (red-orange color).

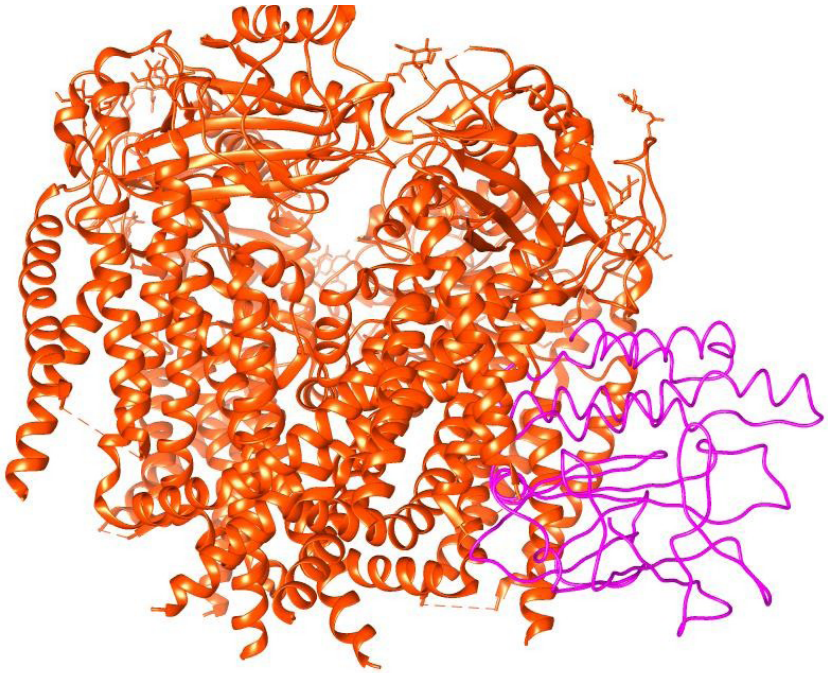

Figure 16. Solution 2 by PatchDock analysis between *PKD1L1* (structure predictedby Phyre2 1010-1724, purple color) and *PKD2* (red-orange color).

The interaction between *PKD1L1* and *PKD2* was further determined by docking them by using another online tool The *PKD1L1* (Phyre2 1010–1724 amino acids) was taken as a receptor while *PKD2* was taken as a ligand. Both wild type and mutated *PKD1L1* were docked with *PKD2* provided in **Figure 17–19**.

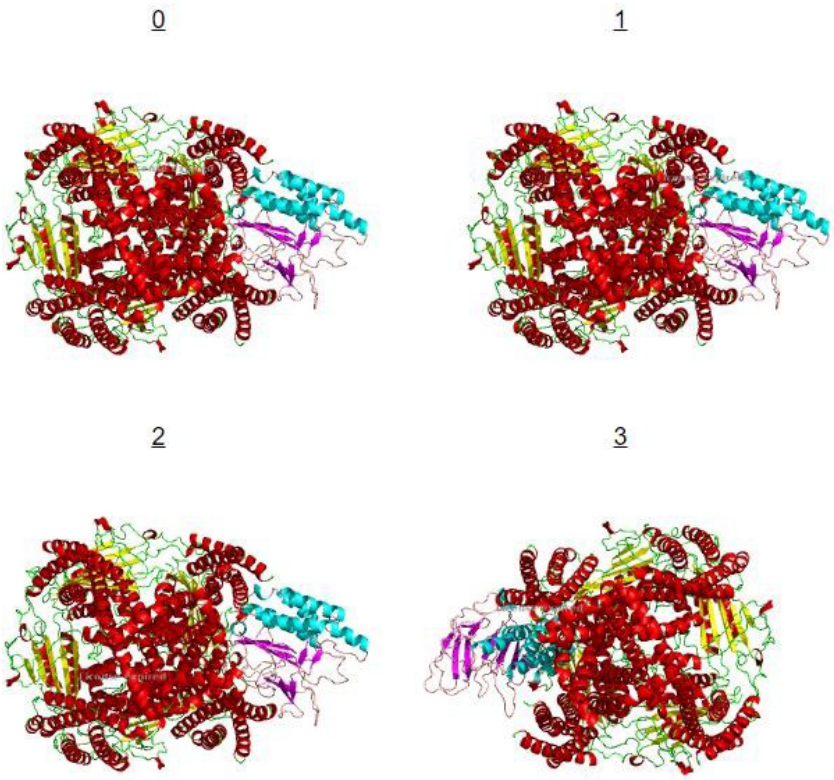

Figure 17. ClusPro Docking analysis solutions 0–3.

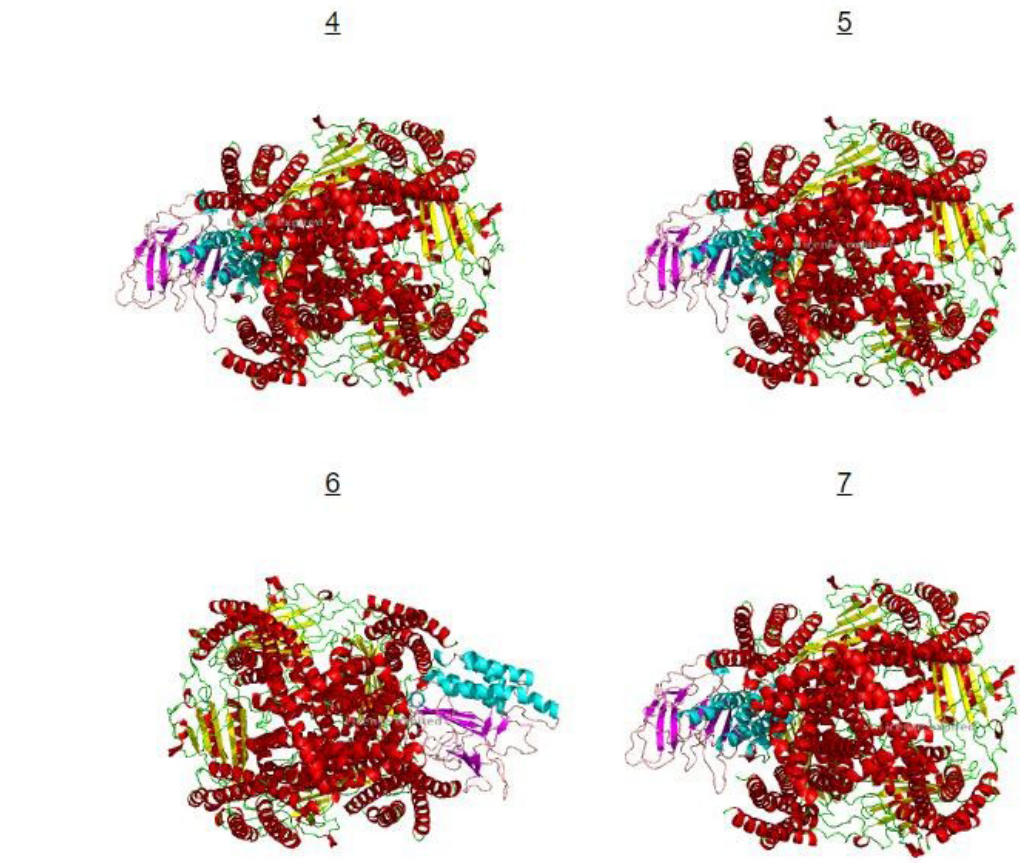

Figure 18. ClusPro Docking analysis solutions from 4–7.

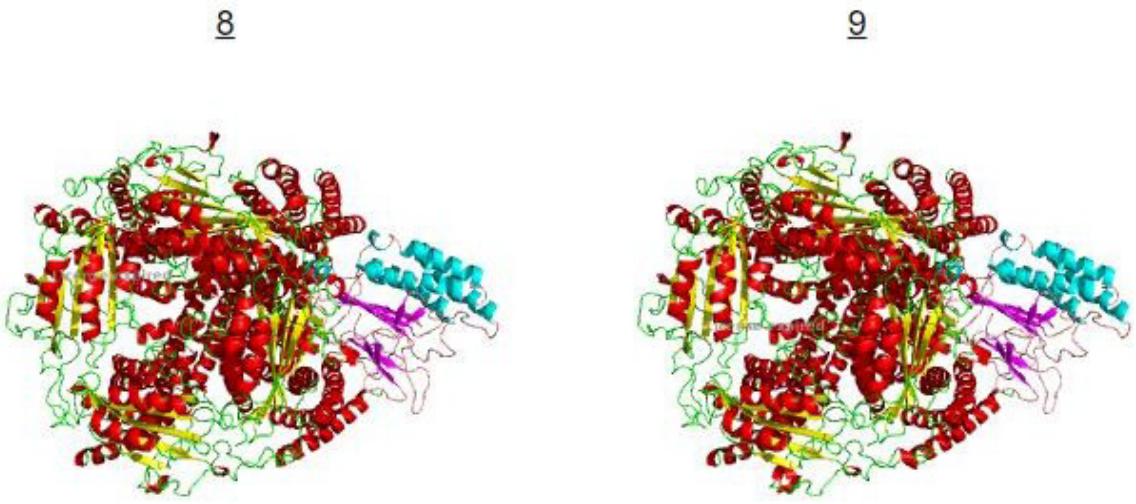

Figure 19. ClusPro Docking analysis solutions from 8–9.

The mutated *PKD1L1* (1010–1724 amino acids) 3D model generated by the intensive Phyre2 prediction tool was docked with *PKD2*. After docking, the following models and results were generated (**Figures 20–22**).

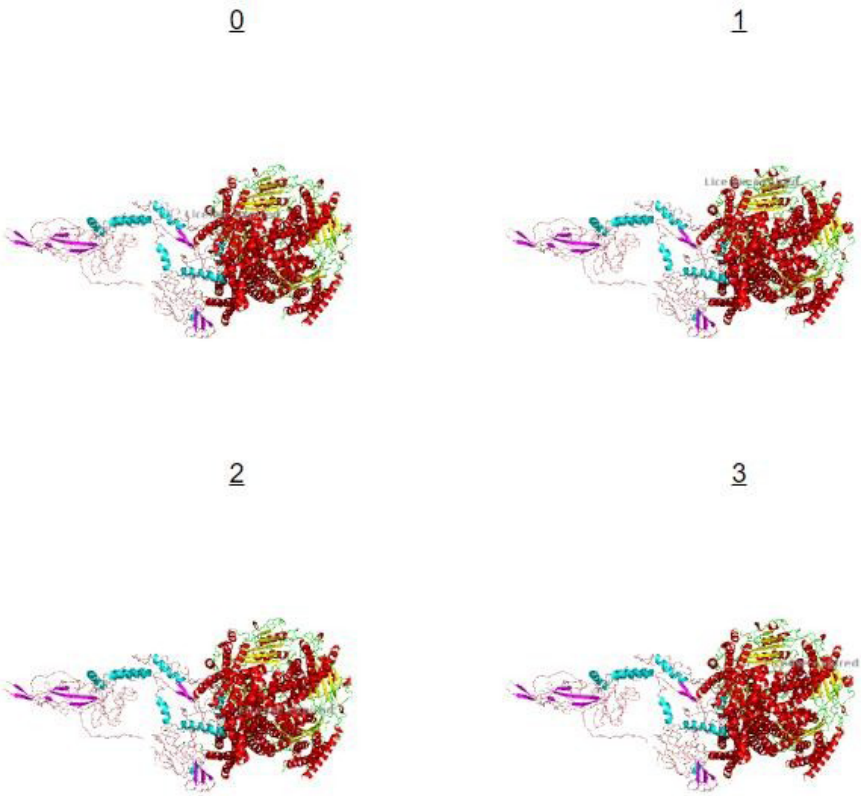

Figure 20. ClusPro Docking analysis of mutated (1691 Cyst/Ser) solutions from 0-–3.

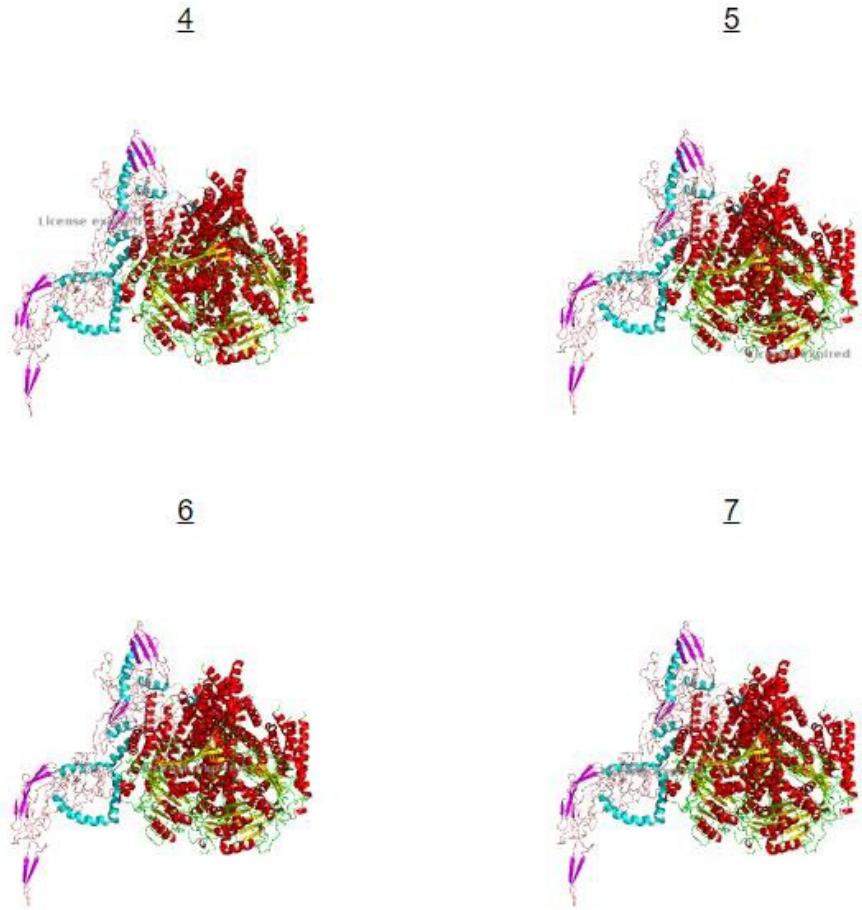

Figure 21. ClusPro Docking analysis of mutated (1691 Cyst/Ser) solutions from 4–7.

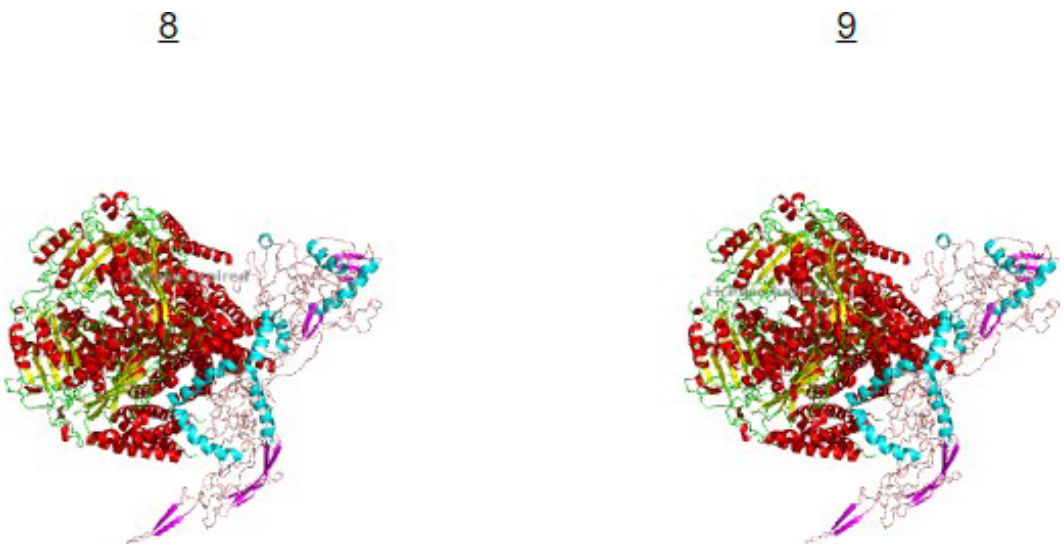

Figure 22. ClusPro Docking analysis of mutated (1691 Cyst/Ser) solutions from 8–9.

The result analysis showed that among *PKD1L1* and *PKD2*, *PKD1L1* was the longest protein with 2849 amino acids compared to *PKD2*, which was the shortest with 807 amino acids. The computed pI for the *PKD1L1* (whole) was more than compared to its first chain, which contained 1691 important amino acids and *PKD2*, respectively. Overall, the number of negatively charged amino acids was more compared to positively charged amino acids, but in the case of *PKD1L1*, the number is greater than the first chain and *PKD2*. Extinction coefficients (ECs) for *PKD1L1* were determined at 280 nm with the assumption that all pairs of Cys residues form cysteines, which is 415785 M^{-1} cm^{-1} (**Table 6**), while for its first chain, TOPO_DOM was 241430, and for *PKD2*, the value was 146400 M^{-1} cm^{-1}. The instability index (II) for *PKD1L1* (whole) was 45.80/ unstable, for TOPO_DOM 46.85/unstable, and for *PKD2* it was 46.96/ unstable. In addition to II, the aliphatic index (AI) for *PKD1L1* (whole) was also computed, which was 84.64, for TOPO_DOM 80.29 and *PKD2* it was 90.56. The GRAVY score observed for *PKD1L1* was -0.137, while for TOPO_DOM it was 0.205 and for *PKD2* -0.136 (**Table 6**).

4. Discussion

The mutation in the *PKD2* domain of *PKD1L1* results in the disorder i.e., SIT. HTX is a recessive disorder that occurs at the developmental stages [26]. The protein sequence for *PKD1L1* (ID: Q8TDX9) was retrieved from UniProt and depicts that there are five domains in *PKD1L1* protein. These are *PKD1* (508 to 590), *PKD2* (592 to 673), REJ (674 to 1571), GPS (1688 to 1734), and PLAT (1796 to 1913), respectively. Analysis of the STRING tool showed that the *PKD1L1* protein is functionally and physically associated with different proteins, different proteins that are produced by a single *PKD1L1* protein-coding gene locus are *PKD1*L2, PRKD1, PLAT, GNA12, GNAI1, GNAI2, DAND5, WDR16, DPCD, and IQEC which can also be considered as different functional partners of *PKD1L1*. The colored lines represent the protein-protein association with respect to shared function but do not interact or show physical binding. The color code for each line is as follows, the purple line represents the experimentally determined interaction of connected proteins, the green line represents the neighborhood gene, the red line represents a fused gene, the blue line represents the co-occurrence of gene, the yellow line represents the text mining, the black line represents co-expression. The cyan line represents the protein homology, respectively. The yellow color line between *PKD1L1* and GNA12,

PRKD1, PLAT, *PKD1*L2, GNAI1, and GNAI2 showed that these are all co-mentioned in the PubMed abstracts (text mining). The multiple-colored lines between GNA12, GNAI1 and GNAI2 represent that in the curated database, they are associated with each other, co-expression is not significant, but putative homologs are co-expressed in other organisms with a score of 0.062 for GANI1 and both GAN12 vs GANI1 proteins have some sequence similarities with 282.3 bits over 347 amino acids. While the multiple-colored lines between GAN12 vs. GANI2, in addition to putative homologs that are co-expressed in other organisms, both also represent co-expression with a functional score of 0.069. Alignment data showed that both GNA12 and GNAI2 have some sequence similarity with 273.1 bits over 354 amino acids.

For *PKD1L1*, the mutation in the natural variant, VAR_077879, is the leading cause of situs inversus totalis (SIT). That type of missense mutation could be the leading cause of SIT [27]. The other variants of *PKD1L1* are not associated with any diseases. The sequence of *PKD2* (polycystine-2) was retrieved from the NCBI, and that protein consists of 968 amino acids. Previous studies showed that it consists of one domain only [28]. *In silico* analysis showed that mutations at different positions in its different variants lead to different disease conditions including the SIT. There are ten different variants of *PKD2* predicted, in which different types of mutations at specific amino acids are identified (shown in **Table 2**). All these mutated variants are associated with different abnormalities [29].

By using the STRING tool, different functionally associated protein networks were retrieved for the *PKD2* protein. The STRING analysis result showed that different neighboring proteins such as PACS1, HAX1, PKHD1, TRPV4, ITPR1, RYR2, *PKD1*, TRPC1, and EXOC5 are associated structurally and functionally with *PKD2* protein, respectively. Each node in STRING analysis represents all the proteins produced by a single, protein-coding gene locus. The query protein is situated in the center (red color node or circle). The connecting lines represent the proteins that jointly contribute to the shared function, although proteins do not bind with each other if they have functions in common. Analysis of the result showed that the interaction of *PKD2* with *PKD1* is experimentally determined as the purple connecting line between *PKD2* and *PKD1*. *PKD1* and EXOC5 show a curated database with *PKD2*, while similarity in homology is only demonstrated by *PKD1* with *PKD2*, while all others did not show homology with *PKD2*.

The OMA tools helped predict the different orthologous of the specific gene among different genome groups. The OMA analysis result showed that for *PKD1L1*, no orthologue group exists. For the fast detection of orthologue for genes, the suitable tool is DIOPT. This tool showed more sensitivity compared to other orthologue retrieving tools. Compared to OMA analysis, the DIOPT tool identifies the putative orthologues for *PKD1L1*. Eight orthologues were identified, shown in **Table 2**, from which two showed the highest score and two showed moderate scores, while the other four showed a very low score, which means low orthologue relation with *PKD1L1*. The highest DIOPT score for orthologue identified by the tool was with three different genes, which are PKDREJ (Polycystin family receptor for egg jelly which might have a role in fertilization of egg), DRscDB, and Human Protein Atlas (HPA) having Ensemble ID ENSG00000130943. The alignment of *PKD1L1* with its orthologue. There was a high alignment score of *PKD1L1* with PKDREJ identified, and *PKD1*L3 had a moderate alignment score while LOXHD1 (**Figure 19**) showed the lowest alignment score.

The first thing needed was protein structure to investigate the interaction of *PKD1L1* with *PKD2*. For *PKD2*, the NMR PDB structure was given, but for *PKD1L1*, no NMR crystallized structure is available in PDB. Therefore, the next step was the structure modeling of *PKD1L1*. For this different free available online tools were used. Before going towards the finalization of the modeled structure of *PKD1L1* by these tools, we kept this thing on priority, while for the selection of the predicted model, the predicted model must include the mentioned mutational amino acids. The structure was modeled by using Phyre2, and SWISS-MODEL. The

validation of the generated model was confirmed by Ramachandran plot analysis. However, the main problem with both generated models was that both didn't cover the 1691 amino acids. SWISS-MODEL generated the 3D structure, which has amino acids from 1751 to 2748. Therefore, different amino acid sequences were submitted and tried to model the structure, which contained 1691 Cyst amino acids.

After many attempts, only through Phyre2, the model was obtained which also modeled the 1691 amino acid. The sequence from 1010 to 1724 was selected randomly and submitted for the prediction of the model. 245 residues (34% of sequence) have been modeled by the phyre2 with 98.9% confidence with the single highest scoring template. The model was based on template c6v55a. c6v55a is a protein for adhesion g protein-coupled receptor gpr126 functioning. The predicted model was validated by Ramachandran plot analysis, which showed that the model is of good quality (blue color).

The different functional and structural properties of the predicted model of *PKD1L1* (1010–1724) were further analyzed. The mutational sensitivity analysis showed that the amino acid 1691 (682 position concerning the submitted sequence which was from 1010 to 1725) is highly sensitive to mutation. Using the fpocket2 the active binding sites were tried to retrieve. For the amino acid 1691 (position 682) no pocket binding site was predicted. Functional and structural conservation showed that all residues are not important. The result analysis showed that no functional and structural conservation was identified. The ProQ2 model quality assessment algorithm was used to assess the quality of the predicted model. And result analysis showed that the model is of good quality. The Cysteine amino acid at the 1691 position was mutated with serine and a 3D model of the protein was predicted with a confidence of the predicted structure was 98% for 19% residues.

The current study sheds light on the variations in the vital properties of *PKD1L1* and *PKD2* such as molecular weight, isoelectric point, total number of negatively and positively charged residues, extinction coefficient, instability index, aliphatic index and grand average of hydropathicity. In addition to this, these calculation results showed that the pI point of *PKD1L1* is greater than compared to *PKD2*, the *PKD1L1* contained more negatively charged amino acids compared to *PKD2* and *PKD1L1* first chain, TOPO_DOM (**Table 6**).

After predicting the 3D structure of *PKD1L1*, it was docked with *PKD2*. The 3D structure obtained from SWISS, Phyre2 (normal) and i-TASSER (normal and MTD) showed different interactions during docking. In the case of SWISS, Phyre2 (normal), i-TASSER (normal and MTD), the amino acid which was the target of the study was either not resolved in structure (in the case of SWISS-MODEL and Phyre2 normal) or if resolved, the amino acid was located away from the rest of the protein structure or 3D predicted structure by i-TASSER MTD. Therefore, the 3D structure resolved through Phyre2 Intensive was used for docking analysis. For docking different tools were tried, but we were unsuccessful in docking using Autodock Vina, Hex 8, ZDOCK, HADDOCK and FireDock because the size of our ligand (*PKD2*) was too large to fit in the frame. Therefore, PatchDock and ClusPro were used for checking the interaction between *PKD1L1* and *PKD2*.

The results from PatchDock showed that, the target amino acid still not interacting well with *PKD2*. In the case of ClusPro, the interaction between *PKD1L1* (Phyre2 3D, 1010-1724 amino acids, wild & mutated) and *PKD2* were determined and to some extent, the interactions were present, which showed that mutation in *PKD1L1* has an important impact in binding with *PKD2*. The 3D structure of *PKD1L1* was predicted by using online tools. The model predicted by the SWISS-MODEL didn't have all the residues of the protein, especially the amino acid at position 1691. The 3D structure predicted had residues from 1751 to 2748, which means the residues of the end region were modeled only. Although the quality of the predicted model was good, with above 90% confidence, but lacked that chain which had amino acid of interest. Later, different sequences of amino acids were picked, keeping 1691 amino acids in all, and submitted in SWISS-MODEL and Phyre2.

SWISS-MODEL modeled the structure but could not model the amino acid of interest. The model predicted by Phyre2 had good coverage, and the quality of the model was also good. The mutated structure of *PKD1L1* was also predicted and validated by Ramachandran plot analysis. The confidence of the predicted model was greater than 90%.

5. Conclusion

All this discussion concludes that mutation in *PKD1L1* can be a leading cause of SIT, besides other abnormalities. Mutation in different variants of *PKD2* made SIT symptomatic from asymptomatic. For *in silico* analysis, the presence of the protein structure is essential. Although online tools predicted the model of a protein, which is accurate with 90% confidence, there is still a need for NMR or X-ray crystallography to predict the 3D structure of the protein. Concerning the Pakistani population the wet experimental analysis can make that research work very excellent. Still, it was tried to collect the sampling for SIT patients but could not collect the positive sample. After a long struggle, only one patient with symptoms was found SIT positive. However, due to insufficiency of resources were unable to analyze that sample.

Author contribution

Literature review: Fatima Haroon
Experiment design: Fatima Haroon, Salma Saeed Khan, Assad Ullah
Experimental work: Fatima Haroon
Result analysis: Fatima Haroon

Disclosure statement

The author declares that they have no conflict of interest.

References

[1] Eitler K, Bibok A, Telkes G, 2022, Situs Inversus Totalis: A Clinical Review. International Journal of General Medicine, 2437–2449.

[2] Manti S, Hwang DY, 2023, Case Reports in Pulmonary Medicine, thesis, University School of Medicine, New Haven.

[3] Tofigh AM, Nematihonar B, Azimi B, et al., 2023, Three Surgical Cases of Situs Inversus Totalis with Individual Challenges; Case Report and Literature Review. International Journal of Surgery Open, 59: 100689.

[4] Weis F, Degrandi C, 2021, Situs Inversus Totalis. Anesthesiology & Intensivmedizin, 62.

[5] Chen W, Guo Z, Qian L, Wang L, 2020, Comorbidities in Situs Inversus Totalis: A Hospital-Based Study. Birth Defects Research, 112(5): 418–426.

[6] Hirano N, Iseki M, Nakagawa K, et al., 2024, A Case Report of Perihilar Cholangiocarcinoma in a Patient with Situs Inversus Totalis. Clinical Journal of Gastroenterology, 1–8.

[7] Whitchurch JB, Schneider S, Hilger AC, et al., 2024, PKD1L1 Is Involved in Congenital Chylothorax. Cells, 13(2): 149.

[8] Dardas Z, Fatih JM, Jolly A, et al., 2024, NODAL Variants Are Associated with a Continuum of Laterality Defects from Simple D-Transposition of the Great Arteries to Heterotaxy. Genome Medicine, 16(1): 53.

[9] Hafez AS, Asar MM, Farid S, et al., 2024, Laparoscopic Retroperitoneal Lymphadenectomy for Ovarian Mixed Germ Cell Tumor in a Patient with Situs Inversus Totalis; Reporting the First Case Worldwide with Literature Review and In Silico Analysis. Pathology-Research and Practice, 155228.

[10] Correa ARE, Endrakanti M, Naini K, et al., 2021, Hydrops Fetalis in PKD1L1-Related Heterotaxy: Report of Two Foetuses and Expanding the Phenotypic and Molecular Spectrum. Annals of Human Genetics, 85(3–4): 138–145.

[11] Burwick RM, Govindappagari S, Sanchez-Lara PA, 2021, Situs Inversus Totalis and Prenatal Diagnosis of a Primary Ciliary Dyskinesia. Journal of Clinical Ultrasound, 49(1): 71–73.

[12] de Oliveira Garcia AC, Caputo LRG, de Andrade WV, Latorraca EF, 2024, Nodal Flow and Situs Inversus: A Review of the Literature. Seven Editora, Chapter 15.

[13] Valet M, Siggia ED, Brivanlou AH, 2022, Mechanical Regulation of Early Vertebrate Embryogenesis. Nature Reviews Molecular Cell Biology, 23(3): 169–184.

[14] Gu H, Yuan ZZ, Xie XH, et al., 2022, A Novel Nonsense PKD1L1 Variant Cause Heterotaxy Syndrome with Congenital Asplenia in a Han Chinese Patient. Journal of Human Genetics, 67(10): 573–577.

[15] Hellen DJ, Bennett A, Malla S, et al., 2023, Liver-Restricted Deletion of the Biliary Atresia Candidate Gene PKD1L1 Causes Bile Duct Dysmorphogenesis and Ciliopathy. Hepatology, 77(4): 1274–1286.

[16] Antony D, Gulec Yilmaz E, Gezdirici A, et al., 2022, Spectrum of Genetic Variants in a Cohort of 37 Laterality Defect Cases. Frontiers in Genetics, 13: 861236.

[17] Leslie JS, Rawlins LE, Chioza BA, et al., 2020, MNS1 Variant Associated with Situs Inversus and Male Infertility. European Journal of Human Genetics, 28(1): 50–55.

[18] Esarte Palomero O, Larmore M, DeCaen PG, 2023, Polycystin Channel Complexes. Annual Review of Physiology, 85: 425–448.

[19] Graziani L, Zampatti S, Carriero ML, et al., 2023, Co-Inheritance of Pathogenic Variants in PKD1 and PKD2 Genes Determined by Parental Segregation and De Novo Origin: A Case Report. Genes, 14(8): 1589.

[20] Sezgin İ, Kayataş M, Kurtulgan HK, et al., 2020, Analysis of PKD1 and PKD2 Gene Mutations for Autosomal Dominant Polycystic Kidney Disease Cases in Turkish Population. Turkish Journal of Nephrology (Online), 29(4): 304–309.

[21] Miyanishi H, Uchida K, 2021, Establishment of a Simplified System to Evaluate Salinity Preference and Validation of Behavioral Salinity Selection in the Japanese Medaka, Oryzias latipes. Fishes, 6(2): 18.

[22] Ciki KI, Turer Ö, Hizal M, et al., 2020, A Rare Cause of Acute Abdominal Pain in a Patient with Primary Ciliary Dyskinesia with Situs Inversus Totalis. Turkish Journal of Pediatrics, 62(1).

[23] Abdali HA, Duddu JR, Mubarak MJ, Mohamed AS, 2021, Rare Association of Klippel-Feil Syndrome with Situs Inversus Totalis and Review of the Genetic Background. BMJ Case Reports CP, 14(5): e241906.

[24] Postema MC, Carrion-Castillo A, Fisher SE, et al., 2020, The Genetics of Situs Inversus Without Primary Ciliary Dyskinesia. Scientific Reports, 10(1): 3677.

[25] Sodeifian F, Samieefar N, Shahkarami S, et al., 2023, DNAH11 and a Novel Genetic Variant Associated with Situs Inversus: A Case Report and Review of the Literature. Case Reports in Medicine, 2023(1): 8436715.

[26] Alzahrani OR, Alatwi HE, Alharbi AA, et al., 2022, Identification and Characterization of Novel Mutations in Chronic Kidney Disease (CKD) and Autosomal Dominant Polycystic Kidney Disease (ADPKD) in Saudi Subjects by Whole-exome Sequencing. Medicina, 58(11): 1657.

[27] Cheng C, Li X, Zhao S, et al., 2022, Compound Heterozygous Variants in DYNC2H1 in a Foetus with Type III Short Rib-Polydactyly Syndrome and Situs Inversus Totalis. BMC Medical Genomics, 15(1): 55.

[28] Dyer LM, 2022, Unravelling the Genetic and Molecular Mechanisms of the Left-Right Patterning Genes PKD1L1 and PKD2 (Doctoral Dissertation, University of Oxford).

[29] Zhang Z, Bai H, Blumenfeld J, et al., 2021, Detection of PKD1 and PKD2 Somatic Variants in Autosomal Dominant Polycystic Kidney Cyst Epithelial Cells by Whole-genome Sequencing. Journal of the American Society of Nephrology, 32(12): 3114–3129.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.