Molecular and Cellular Probes 35 (2017) 8–19

Contents lists available at ScienceDirect

Molecular and Cellular Probes

journal homepage: www.elsevier.com/locate/ymcpr

Homology modeling and *in silico* prediction of Ulcerative colitis associated polymorphisms of NOD1

Ishani Majumdar [1], Isha Nagpal [1], Jaishree Paul*

School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India

ARTICLE INFO

Article history:
Received 16 February 2017
Received in revised form
23 May 2017
Accepted 30 May 2017
Available online 1 June 2017

Keywords:
NOD1
SNPs
Ulcerative colitis
in silico
Homology modeling
Protein function

ABSTRACT

Cytosolic pattern recognition receptors play key roles in innate immune response. Nucleotide binding and oligomerisation domain containing protein 1 (NOD1) belonging to the Nod-like receptor C (NLRC) sub-family of Nod-like receptors (NLRs) is important for detection and clearance of intra-cellular Gram negative bacteria. NOD1 is involved in activation of pro-inflammatory pathways. Limited structural data is available for NOD1. Using different templates for each domain of NOD1, we determined the full-length homology model of NOD1. ADP binding amino acids within the nucleotide binding domain (NBD) of NOD1 were also predicted. Key residues in inter-domain interaction were identified by sequence comparison with *Oryctolagus cuniculus* NOD2, a related protein. Interactions between NBD and winged helix domain (WHD) were found to be conserved in NOD1. Functional and structural effect of single nucleotide polymorphisms within the NOD1 NBD domain associated with susceptibility risk to Ulcerative colitis (UC), an inflammatory disorder of the colon was evaluated by *in silico* studies. Mutations W219R and L349P were predicted to be damaging and disease associated by prediction programs SIFT, PolyPhen2, PANTHER, SNP&GO, PhD SNP and SNAP2. We further validated the effect of W219R and L349P mutation on NOD1 function *in vitro*. Elevated mRNA expression of pro-inflammatory cytokines IL8 and IL-1β was seen as compared to the wild type NOD1 in intestinal epithelial cell line HT29 when stimulated with NOD1 ligand. Thus, these mutations may indeed have a bearing on pathogenesis of inflammation during UC.

1. Introduction

Nod-like receptors (NLRs) are cytoplasmic pattern recognition receptors that recognize intracellular pathogen derived microbial components. The NLR family consists of 22 members and is divided into four sub-families based on their *N*-terminal domains i.e. NLRA, NLRB, NLRC and NLRP. NLRA members contain the *N*-terminal transactivation domain, NLRB members have the baculovirus inhibitor of apoptosis protein repeat domain (BIR) while NLRC and NLRP members contain caspase activation and recruitment domain (CARD) and pyrin domain respectively [1]. The NLRC member, Nucleotide binding oligomerisation domain containing protein 1 (NOD1) expressed in epithelial cell, macrophages and dendritic cells plays an important role in the innate immune system [2–4]. It recognizes gamma-D-glutamyl-*meso*-DAP (iE-DAP), a conserved moiety of Gram negative bacterial cell wall [5]. Ligand recognition by NOD1 leads to signaling through NFκB, MAPK and JNK pathways initiating innate immune response and production of various pro-inflammatory cytokines [6] [7,8]. It also mediates autophagy of intracellular bacteria like *Shigella* [9].

NOD1 belongs to the Signal Transduction ATPases with Numerous Domains (STAND) family of ATPases. It is a 110 KDa protein encoded by 953 amino acids. In addition to the *N*-terminal CARD domain, NOD1 consists of the central Nucleotide Binding and oligomerisation Domain (NBD) and *C*-terminal Leucine rich Repeat (LRR) domain [10]. The CARD domain mediates downstream signal transduction through homotypic protein-protein interactions [11]. CARD domain belonging to the evolutionary conserved Death Domain (DD) super family, facilitates formation of stimuli induced oligomeric signaling complexes through both self-association and interactions with other proteins belonging to the same subfamily [12]. NBD domain is involved in nucleotide binding. The NBD domain mediates oligomerisation of the protein and plays a crucial

* Corresponding author. #424, School of Life Sciences, Jawaharlal Nehru University, New Mehrauli Road, New Delhi 110 067, India.
E-mail address: jpaul33@hotmail.com (J. Paul).
[1] Equal Contribution.

http://dx.doi.org/10.1016/j.mcp.2017.05.009
0890-8508/

role in NOD1 activation [13]. Bacterial ligand recognition by the LRR domain triggers its conformational change leading to NOD1 activation [14]. Despite both crystal and solution structures of NOD1 CARD domain along with LRR homology model being known, data for full-length NOD1 is currently lacking [15–17]. Expression and purification of full-length NLR proteins poses certain difficulties. Homology modeling of NLR family members such as NOD2 and NLRC5 has been useful in determination of their structures [13], [18].

Given that NOD1 plays an important physiological role in defense against microbial infection, its genetic variation can contribute to dysregulated signaling causing damage to the host. Indeed, polymorphisms in *NOD1* gene have been linked with various auto-immune diseases. Single Nucleotide Polymorphisms (SNPs) are single base changes leading to DNA variation with more than or equal to 1% allelic frequency in a population. SNPs may exist in the non-coding regions (introns, 5′UTR, 3′UTR) or in coding sequences of the gene. Non-synonymous SNPs (nsSNPs), a type of SNP within the coding regions, alter the amino acid sequence of the protein which can affect protein stability [19] or disrupt its functions [20]- [21]. nsSNPs make up about 2% of the genetic variations associated with genetic diseases [22]. SNPs located in an intron of *NOD1* are associated with autoimmune inflammatory diseases like asthma [23]. The intestinal mucosa is constantly exposed to commensal bacteria in the gut lumen. NOD1 is important in the intestinal epithelial cells for bacterial clearance via production of NF-κB dependent anti-microbial peptides. NOD1 signaling also coordinates the immune response by T cells and B cells with Toll-like receptors in the colon. Studies suggesting role of *NOD1* SNPs in causing genetic susceptibility to Inflammatory bowel disease (IBD), a chronic, relapsing and inflammatory autoimmune disorder of the gastrointestinal tract are population dependent. It results from a dysregulated immune response towards individuals own commensal microbiota [24]. A complex indel polymorphism in *NOD1* is strongly associated with IBD in Caucasian population [25]. While exonic SNPs found in New Zealand Caucasian population are not associated with the disease [26]. Indel polymorphisms found in German and UK populations also did not cause IBD susceptibility risk [27,28]. In patients of North India suffering from Ulcerative colitis (UC), a clinical sub-type of IBD causing mucosal inflammation of the colon, both exonic and intronic SNPs have shown association with disease activity [29,30]. The impact of SNPs lying within exon six of *NOD1* that encodes the NBD domain is unknown [29]. The NBD domain is crucial for NOD1 function as it carries out the switch between inactive monomeric and active oligomeric state. A model of NOD1 activation has been proposed based on site directed mutation studies [31]. According to this model, elicitor recognition leads to conformational change in NOD1 receptor that is followed by exchange of ADP bound to the NBD domain with ATP. Hydrolysis of ATP causes oligomerisation and activation of NOD1.

Keeping the above facts in consideration, we predicted the full length homology model of NOD1. We also identified residues involved in NOD1 inter-domain interactions by sequence comparison with NOD2, a NLRC member closely resembling NOD1 in structure and function. It was also pertinent to determine the impact of the exonic SNPs on NOD1 structure and function bioinformatically. Further, effect of these SNPs on NOD1 function was also validated experimentally.

2. Materials and methods

2.1. Homology modeling and ADP binding pocket prediction

Homology model of NOD1 was prepared by submitting its protein sequence (UniProt ID:Q9Y239) to RaptorX server (http://raptorx.uchicago.edu/) [32]. Multiple protein structures were obtained for each domain of NOD1 as potential templates. The protein structure having the highest alignment score with the target sequence was used as the template for modeling of NOD1. The quality of the predicted model was indicated by p-value that measures relative quality of the model. ADP binding pocket in NOD1 model was predicted using the Raptor X binding server.

2.2. Energy minimization

The homology model was subjected to protein preparation wizard of GLIDE v.8 from Schrodinger suite to rectify the atomic asymmetries, which includes addition of missing hydrogen atoms, refinement of hydrogen bonds, allocation of bonds and bond orders, generation of disulphide bonds, capping the terminal ends of protein with ACE and NME groups, resolving protonation states of histidine, fixing and assigning missing residues [33]. After which the protein's hydrogen bonds and charges were optimized followed by minimization step using OPLS_2005 force field which aids in the refinement of the protein structure. The homology model was validated using RAMPAGE (http://www-cryst.bioc.cam.ac.uk/rampage/) [34].

2.3. Databases

rsIDs of *NOD1* SNPs reported by Verma et al. [29] were retrieved from db SNP database (http://www.ncbi.nlm.nih.gov/SNP/) [35]. Amino acid sequence of NOD1 NBD domain was identified from the NOD1 protein sequence entry (UniProt ID:Q9Y239) in UNIPROT database (www.uniprot.org/) [36]. *Homo sapiens* NOD1 NBD homologs were recovered using NOD1 NBD sequence as query in BLASTP (https://blast.ncbi.nlm.nih.gov/) [37].

2.4. Multiple sequence alignment

Multiple sequence alignment of NOD1 NBD mammalian orthologs was performed using MUSCLE v 3.8 (www.ebi.ac.uk/Tools/msa/muscle/) [38]. Consensus sequence images of multiple sequence alignment were generated using Web Logo v 3.4 (weblogo.berkeley.edu/logo.cgi) [39].

2.5. Calculation of evolutionary conservation

The degree of evolutionary conservation of the amino acids undergoing mutation due to ns SNPs was examined using ConSurf web server (consurf.tau.ac.il/) [40]. The multiple sequence alignment of 21 NOD1 NBD mammalian homologs was submitted as input.

2.6. Prediction of functional impact of SNPs

Impact of the SNPs on NOD1 function was predicted by Sorting Intolerant from Tolerant (SIFT), Phenotyping Polymorphism (PolyPhen) 2 v 2.2, Protein Analysis Through Evolutionary Relationships (PANTHER), SNPs and GO, PhD SNP, Screening for Non-Acceptable Polymorphisms (SNAP) and SNP effect 4.0.

2.6.1. SIFT

SIFT algorithm predicts potential effects of nsSNPs as tolerant or damaging on protein function (http://sift.jcvi.org) [41]. NOD1 protein sequence was used as query to which the SIFT server searched and aligned homologous protein sequences of suitable diversity. Sequence diversity was indicated by median information content score of 2.31. The classifier model HumDiv and genome assembly GRCh37/hg19 were used for the prediction.