

Journal Pre-proof

Multiscale simulations reveal architecture of NOTCH protein and ligand specific features

Surabhi Rathore, Deepanshi Gahlot, Jesu Castin, Arastu Pandey, Shreyas Arvindekar, Shruthi Viswanath, Lipi Thukral

PII: S0006-3495(24)04075-X

DOI: <https://doi.org/10.1016/j.bpj.2024.12.014>

Reference: BPJ 13512

To appear in: *Biophysical Journal*

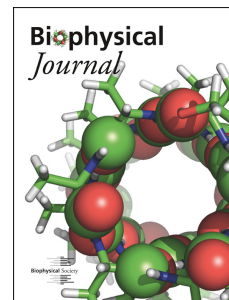
Received Date: 1 February 2023

Accepted Date: 10 December 2024

Please cite this article as: Rathore S, Gahlot D, Castin J, Pandey A, Arvindekar S, Viswanath S, Thukral L, Multiscale simulations reveal architecture of NOTCH protein and ligand specific features, *Biophysical Journal* (2025), doi: <https://doi.org/10.1016/j.bpj.2024.12.014>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2024 Biophysical Society. Published by Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.



Multiscale simulations reveal architecture of NOTCH protein and ligand specific features

Surabhi Rathore^{1,2}, Deepanshi Gahlot^{1,2}, Jesu Castin¹, Arastu Pandey³, Shreyas Arvindekar³, Shruthi Viswanath³, Lipi Thukral^{1,2,*}

1 CSIR-Institute of Genomics and Integrative Biology, Mathura Road, New Delhi- 110025, India

2 Academy of Scientific and Innovative Research (AcSIR), Ghaziabad- 201002, India

3 National Center for Biological Sciences (NCBS), Tata Institute of Fundamental Research (TIFR), GKVK, Bellary Road, Bangalore-560065, India

*Correspondence to Lipi Thukral (lipi.thukral@igib.res.in)

Abstract

NOTCH, a single-pass transmembrane protein, plays a crucial role in cell fate determination through cell-to-cell communication. It interacts with two canonical ligands, Delta-like (DLL) and Jagged (JAG), located on neighboring cells to regulate diverse cellular processes. Despite extensive studies on the functional roles of NOTCH and its ligands in cellular growth, the structural details of full-length NOTCH and its ligands remain poorly understood. In this study, we employed fragment-based modeling and multiscale simulations to study the full-length structure of the human NOTCH ectodomain, comprising 1,756 amino acids. We performed coarse-grain dynamics simulations of NOTCH in both glycosylated and non-glycosylated forms to investigate the role of glycosylation in modulating its conformational dynamics. In apo form, coarse-grained simulations revealed that glycosylated NOTCH protein can transition from an elongated structure of ~ 86 nm from the membrane surface to a semi-compact state ($\sim 23.81 \pm 9.98$ nm), which aligns with cryo-EM data. To transition from the apo form to ligand-bound forms of NOTCH, we followed an atomistic and integrative modeling approach to model the interactions between NOTCH-DLL4 and NOTCH-JAG1. Atomistic simulations of the smaller bound fragment EGF8-13 patch revealed conformational plasticity critical for NOTCH binding, while integrative modeling of full-length complexes suggested a larger binding surface than previously reported. Simulations of pathogenic mutations revealed that E360K and R448Q disrupted the NOTCH-ligand interaction surfaces, causing dissociation. In contrast, C1133Y in the Abruption domain compromised protein stability by disrupting the domain's interaction with the ligand-binding domain in the apo form of NOTCH-ECD. These findings provide a detailed molecular understanding of NOTCH and its ligands, offering insights that could enable the development of novel therapeutic approaches to selectively target pathogenic NOTCH signaling.

Significance

We report the modeled structure and dynamics of the large membrane protein, human NOTCH1 extracellular domain, using a fragment-based modelling approach. This model allowed us to explore its dynamics in both apo and ligand-bound states with the canonical NOTCH ligands DLL4 and JAG1. Key structural features

of NOTCH1, including glycosylation, ligand-binding interactions (DLL4 and JAG1), and the mapping of pathogenic mutations, were thoroughly examined. These complex structural elements, which influence protein function and receptor-ligand dynamics, provide crucial insights into the molecular basis of NOTCH signaling. Understanding these features opens new avenues for developing targeted therapeutics to modulate NOTCH receptor activity, which could be vital for treating diseases like cancer and developmental disorders.

Introduction

NOTCH is an evolutionarily conserved signaling pathway across all metazoans [1]. The central protein in this pathway, NOTCH1 (referred to as NOTCH hereafter), is a multi-domain protein consisting of approximately 2500 amino acids (Figure 1). It is initially synthesized as a single precursor, which undergoes proteolysis, yielding two fragments: the extracellular domain (ECD) and the intracellular domain (ICD) [2]. The human NOTCH-ECD is composed of 36 Epidermal Growth Factor (EGF)-like repeats followed by a Negative Regulatory Region (NRR) and a transmembrane (TM) region [3]. Apart from providing structural integrity to the NOTCH protein, the EGF-like repeats also serve as the platform for receptor-ligand binding [4]. In particular, NOTCH binds with two canonical ligands that are also transmembrane proteins, namely Delta-like protein (DLL) and Jagged (JAG) present on the neighbouring cells [5,6]. Interestingly, binding different ligands at the NOTCH-ECD results in distinct functional outcomes, but the complete structural details of NOTCH and its ligands remain largely unknown, hindering our understanding of the molecular basis of binding specificity.

Structural insights into NOTCH-ligand complexes have been reported, but these are limited to only a few EGF-like repeats. For example, in an NMR-derived structure, EGF11–13 of the NOTCH-ECD was solved, revealing that EGF11–12 adopts a distinct, elongated shape, with pair-wise domain linkage significantly affecting structural stability [7]. In 2016, another segment containing the EGF4-7 region of human NOTCH was solved using NMR and X-ray crystallography that revealed a bent conformation, thereby changing the classical view of NOTCH protein as a rigid linear structure [8]. In a recent cryo-EM study, three flexible conformations of NOTCH were observed and it was proposed that NOTCH undergoes several bending motions [9]. Furthermore, another contributing factor to NOTCH-ECD structural basis is glycosylation modifications on the EGF-like repeats. Previous studies on a single EGF-like repeat including O-glucose and O-fucose consensus sites demonstrate that the two glycans contribute to the stability of the fold [10]. However, due to the complexity of membrane association and solubility, the complete structure of NOTCH remains unresolved. In addition, ligand-bound states of NOTCH have also been recently resolved. In particular, EGF11-13 in complex with DLL4 fragment revealed that the two proteins exhibit a co-linear conformation and interact in an anti-parallel fashion [11]. Further, a three-dimensional structure of NOTCH EGF8-12 and a segment of JAG1 revealed an extended binding interface that spans all five domains [12]. These developments provide important information about ligand binding but leave a significant gap in the complete molecular understanding of NOTCH-ligand binding and its dynamics, especially with regard to mutations affecting this process. In particular, mutations in the NOTCH receptor and its ligands are associated with cancers and genetic disorders such as Alagille syndrome, CADASIL, and T-cell acute lymphoblastic leukemia (T-ALL) [13–15]. These mutations often disrupt the receptor’s ability to bind with its ligands [16], alter the intracellular signaling cascade [15], or affect receptor processing [17], resulting in dysregulated cell fate determination.

In this work, we utilised multiple contemporary computational approaches to investigate the structure and conformational dynamics of the NOTCH receptor in its apo and ligand-bound states. Employing fragment-based homology modeling, we first generated NOTCH extracellular domain with TM region and

Fig 1. A schematic overview of NOTCH receptor-mediated cell-to-cell communication with ligands. NOTCH present in one cell binds to the extracellular region of its ligands, DLL4 or JAG1 on the adjacent cell in a trans-manner to activate the signaling cascade. The NOTCH receptor primarily consists of NOTCH-ECD and NOTCH-ICD. The NOTCH-ECD consists of 36 EGF-like repeats (pink) followed by the NRR (green). The NOTCH-ECD interacts with its ligands via consecutive EGF-like repeats that result in diverse cellular outcomes.

extended fragment of C-terminal comprising ~ 1700 residues, with 36 EGF-like repeats, the catalytic domain NRR, the TM region, and a part of cytosolic C-terminal. To assess its dynamic properties, we performed coarse-grained molecular dynamics simulations of both glycosylated and non-glycosylated NOTCH-ECD in the apo state, revealing that it does not remain linear but adopts multiple folded conformations, driven by flexible N-terminal EGF-like repeats and relatively stable repeats closer to the membrane. We utilized both atomistic and integrative modeling techniques to explore its interactions with NOTCH-DLL4 and NOTCH-JAG1. Lastly, we explored whether the proposed structural framework could be used to map three key pathogenic mutations, helping us understand their impact on overall structural stability as well as local interactions. Thus, we present a comprehensive pipeline of multiscale simulations and modeling to reveal the atomistic binding of these supramolecular membrane complexes, which were previously understudied.

Materials and methods

Generation of the full-length human NOTCH-ECD receptor

Figure S1 shows a schematic of the workflow to model the NOTCH-ECD structure and a detailed description of the modeling procedure is listed below for replicability. The initial step involved compiling input from available experimental data. The human NOTCH-ECD consists of ~ 1700 amino acids, organised into 36 EGF-like repeats. The experimental structures for 10 out of 36 EGF-like repeats; EGF4-7 (PDB ID: 5FM9), EGF8-12 (PDB ID: 5UK5), EGF11-13 (PDB ID: 2VJ3, 4XLW) were available [8, 11, 12, 18]. The EGF-like repeats display conserved structural features, including a characteristic anti-parallel β -sheet fold, and six cysteines that form three disulfide bonds [7, 19]. Additionally, structural information was available for the NRR domain, determined via X-ray crystallography, and for the transmembrane (TM) region, characterized by NMR conformers [20, 21]. Collectively, these known structures covered approximately $\sim 42\%$ of the complete NOTCH-ECD. Thus, the unresolved regions in the complete human NOTCH-ECD correspond to the missing structures of EGF1-3 and EGF14-36.

The second step involved fragment-based homology modeling to reconstruct the missing EGF-like repeats (EGF1-3 and EGF14-36). Using the PROBCONS standalone version 1.12 [22], pairwise sequence identity percentages between the missing EGF sequences and the known structures of EGF4-13 were computed. Based on the sequence identity matrix shown in Figure S2, templates for each missing EGF-like repeat were selected, with identity scores ranging from approximately $\sim 40\text{-}62\%$. Homology modeling for the individual repeats was then performed using the I-TASSER webserver, and the models with the highest confidence scores (C-scores) were selected for further analysis (Table S1) [23]. The structural models were evaluated for three key conserved characteristics: (i) the anti-parallel β -sheet structure, (ii) the presence of six cysteines, and (iii) the formation of three disulfide bridges. Additional structural validation was performed using the SAVES v5.0 webserver, which included assessments with the Ramachandran plot, ERRAT, and Verify3D scores. These tools provide a statistical measure of the local environment preference for each residue, considering factors such as buried area, side-chain orientation, and local secondary structure (Table S2) [24].

The modeled EGF-like repeats were further subjected to refinement using the ModRefiner webserver in order to obtain the refined high-resolution secondary structures [25].

The third step involved generating eight larger EGF fragments and performing MD simulations. EGF-like repeats were concatenated using standard peptide bond criteria (C-N length=1.33 Å, dihedral angles $\omega = -90^\circ$, $\phi = 120^\circ$) to build structural patches. Two fragments (EGF4-9 and EGF10-13) were available from NMR data [8]. To account for the flexible EGF9-10 interface, multiple orientations were generated by adjusting torsional angles. In total, eight fragments (EGF1-3, EGF4-13, EGF14-17, EGF18-20, EGF21-24, EGF25-28, EGF29-32, EGF33-36) were built and refined through 500 ns of MD simulations in a solvated triclinic box with Na⁺/Cl⁻ ions for neutralization. Each system underwent two-step energy minimization using the steepest descent followed by the conjugate gradient. EGF-like repeat patches were position-restrained at the first and last residues, equilibrated at 310 K and 1 bar with a v-rescale thermostat and Parrinello-Rahman barostat. PME summation was used for long-range electrostatic interactions, with cutoffs set at 1 nm for real-space and van der Waals interactions. A control test was done to select force-field and two simulations for each EGF patch were performed using OPLS-AA and Charmm36-AA force fields. The average structural deviation in EGF patches simulated using OPLS and Charmm36-AA was observed to be 1.07 +/- 0.20 nm and 1.41 +/- 0.23 nm, respectively. Since OPLS-AA provided more stable secondary structure and dynamics, it was used for further refinement (Figure S3-S4). The final production run with all-bond constraints was performed in GROMACS using the OPLS-AA force field.

The final step involved assembling the complete NOTCH-ECD structure and validating the model extensively. Eight refined EGF patches were combined into a single structure using standard peptide bond criteria. The EGF-like repeat domain was linked to the NRR crystal structure (PDB ID: 3ETO) and anchored to the TM region (PDB ID: 5KZO) for full-length assembly [20,21]. Geometric checks, including main-chain bond lengths and angles, were performed using Ramachandran plot and Procheck [24]. The model underwent refinement through the steepest descent and conjugate gradient energy minimization. Domain-wise validation was conducted against available X-ray and NMR structures for NOTCH EGF4-13, NRR, and TM regions.

AlphaFold, a cutting-edge tool for three-dimensional structure prediction [26,27] was reported after our model was generated. While AlphaFold has been effectively used in our previous studies to investigate autophagy pathways [28], its application to large transmembrane receptors, such as NOTCH-ECD, remains challenging. The full-length AlphaFold-predicted structure exhibited implausible domain-domain orientations, as illustrated in Figure S5. However, a comparison of EGF-like repeat models of EGF1-3 and EGF14-36 predicted by I-TASSER and AlphaFold (version 2.3.1) provided similar results (Figure S6). AlphaFold models were predicted with high confidence, achieving average pLDDT scores of 80.24 to 95.47. Structural comparison showed RMSD deviations of 0.43–0.96 Å between the I-TASSER and AlphaFold models (Figure S7).

Coarse-grain molecular dynamics simulations of the full-length non-glycosylated and glycosylated apo NOTCH-ECD

The full-length all-atom human NOTCH-ECD structure was converted into the coarse-grained model using the Charmm-GUI Martini Solution Maker and employing the Martini elnedyn22 force field [29,30]. The NOTCH-ECD transmembrane region was position restraint applied to the TM region and the C-terminal (1736-1771 a.a.) with a force constant of 1000 kJ mol⁻¹ nm⁻². The structure was placed in a box with dimensions of 60 x 60 x 110 nm. 103 Na⁺ ions were added to neutralize and Martini water representation was used to solvate both the NOTCH-ECD systems. The simulations were carried out

using GROMACS version 2018.3 by employing the Martini elnodyn22 force field [31]. The systems were subjected to energy minimization using the steepest descent method. In order to understand the impact of glycans on NOTCH folding dynamics, we also generated the fully-glycosylated model of coarse-grained NOTCH-ECD. NOTCH-ECD harbours 36 Ser/Thr consensus sites, C2-X-X-X-X-(S/T)-C3 (where X is any amino acid) which undergo glycosylation (Table S3) [32–34]. As shown in Figure S8, the core part of the coarse-grained structure for glucose, fucose, and GlcNAc consists of three beads, while the fourth bead in GlcNAc represents the N-acetyl group [35]. The fully glycosylated coarse-grained NOTCH-ECD model was placed in a box with dimensions of 120 x 120 x 120 nm with position restraint applied to the TM region and the C-terminal (1736-1771 a.a.) at a force constant of 1000 kJ mol⁻¹ nm⁻². Three replicas for the non-glycosylated form of NOTCH-ECD were subjected to production MD of 10 μ s each. The fully glycosylated NOTCH-ECD system was subjected to 1 μ s of simulation time. The end-to-end distance between the first and last EGF-like repeats was computed in VMD across three trajectories, where the backbone atoms of the first residue of EGF1 (ARG20) and the last residue of EGF36 (HIS1426) were selected as the two reference points. Contact matrices (cut-off = 1 nm) to track regions in contact during NOTCH-ECD folding were generated using MDAnalysis [36].

Generation of atomistic NOTCH complexes with DLL4 and JAG1

We analyzed the sequence similarity between human and rat NOTCH1 (91%), DLL4 (87%), and JAG1 (96%) proteins. Using the truncated crystal structures of NOTCH-DLL4 (PDB ID: 4XLW) and NOTCH-JAG1 (PDB ID: 5UK5) from rat, we generated equivalent complexes with interfaces containing NOTCH EGF8-13 [11,37]. The NOTCH in the NOTCH-DLL4 complex structure (PDB ID: 4XLW) contains EGF11-13 fragments for NOTCH and MNNL, DSL, and EGF1-2 units for DLL4. Whereas, the NOTCH-JAG1 complex (PDB ID: 5UK5) comprises NOTCH EGF8-12 and MNNL, DSL, and EGF1-3 units of JAG1. Since the interface for these complexes is not comparable, we generated the NOTCH-DLL4 and NOTCH-JAG1 complexes with equivalent interfaces. The NOTCH-DLL4 complex model had an RMSD of 1.02 Å compared to the crystal structure, while the NOTCH-JAG1 complex showed no deviation. Ligand mutations from Luca et al. to improve affinity were retained in the NOTCH-JAG1 model (S32L, R68G, D72N, T87R, and Q182R), specifically in the MNNL region of JAG1.

All-atom molecular dynamics simulations of NOTCH-DLL4 and NOTCH-JAG1

The NOTCH-DLL4 and NOTCH-JAG1 complexes containing equivalent NOTCH (EGF8-13) were pre-processed and placed in a triclinic box of size approximately 10 x 20 x 10 nm. TIP3P water representation was used to solvate the system and for neutralization 15 and 12 Na⁺ ions were added to N1-DLL4 and N1-JAG1 systems, respectively. The simulations were carried out using GROMACS version 5.1.4 by employing OPLS all-atom force field [31]. Particle Mesh Ewald (PME) summation using the grid spacing of 0.16 nm was used combined with a fourth-order cubic interpolation to deduce the forces and potential in-between grid points. The van der Waals cut-off was set to 1.0 nm. The initial systems were subjected to energy minimization using the steepest descent method followed by the conjugate gradient. Temperature and pressure were maintained at 310 K and 1 bar using the V-rescale thermostat and Parrinello-Rahman barostat, respectively [38, 39]. Finally, these prepared systems for NOTCH-DLL4 and NOTCH-JAG1 complexes were subjected to the production MD run in four replicates for 5 μ s each corresponding to a total runtime of 40 μ s. Trajectory analysis was carried out using various gmx modules and hydrogen bonds calculation was performed using gmx hbond module with the default cut-off as 0.35 nm. All the molecular

images were rendered using Chimera, ChimeraX, and VMD [40–42]. The graphs were generated using MATLAB version R2022a, GRACE-5.1.22 program, and Python libraries.

Integrative complexes of full-length NOTCH-ligand extracellular domains

Integrative structure determination of NOTCH-ligand complexes was carried out as shown in Figure S9. The atomistic model of NOTCH-ECD (1-1756 a.a.) was utilized to generate the NOTCH coarse-grained model, while the full-length extra-cellular regions of DLL4 and JAG1 were modeled using experimental inputs, such as truncated crystal structures of constituent protein domains, identification of flexible and rigid domains in NOTCH and its ligands from X-ray crystallography, Ca^{+2} binding experiments, NMR (heteronuclear Overhauser effects, residual dipolar couplings), and stereochemistry, i.e., excluded volume and connectivity information in Modeller (Table S4) [8, 11, 12, 43–46]. Followed by which the NOTCH-DLL4 and NOTCH-JAG1 complexes were generated in 1:1 stoichiometry. Each protein was represented by a series of connected spherical beads coarse-grained at multiple scales. Regions of the unknown structure were represented by 30-residue beads modeled as flexible strings of beads. In contrast, regions with known crystal co-ordinates were represented simultaneously by 1- and 10- residue beads and modeled as rigid bodies, where the relative distances between beads is fixed during sampling.

To balance computational efficiency and modeling accuracy, multiple domains with known crystallized fragments were grouped in a single rigid body. The grouping of domains in rigid bodies was based on the identification of flexible and rigid domains in NOTCH and its ligands as shown in previous structural studies (Figure S9) [8, 43, 44]. In addition, we included the experimental data suggested by NMR and Ca^{+2} binding measurements that show consecutive calcium-binding EGF-like repeats in NOTCH form a rigid structure, while adjacent non-calcium binding EGF-like repeats are flexible [8]. Similarly, based on truncated crystal complexes of NOTCH-DLL4 and NOTCH-JAG1, the ligands are expected to be largely linear and rigid from the MNL to EGF3 and EGF5-6 region, with a bend between EGF4 and EGF5 [8, 44]. A restraint on the end-to-end NOTCH-ECD distance, excluded volume restraint, and connectivity restraints. The distance between NOTCH residues 20 and 1735 was restrained to 85.5 nm +/- 1 nm to model the extended conformation of NOTCH in the full-length atomistic model of NOTCH-ECD. For each NOTCH-ligand complex, starting from initial random configurations for the flexible beads and rigid bodies, a total of 40 million models were sampled using Gibbs's sampling Replica Exchange Monte Carlo Markov Chain (MCMC) algorithm (Figure S9) [47, 48]. Each Gibbs sampling step at a certain temperature consisted of a cycle for Monte Carlo steps that moved every flexible bead by a maximum random translation of 2 Å (NOTCH-DLL4) or 2.5 Å (NOTCH-JAG1). While each rigid body moved by a maximum random translation of 0.5 Å (NOTCH-DLL4) or 2.5 Å (NOTCH-JAG1) and a maximum random rotation of 0.1 rad. Models were generated by the Gibbs sampling in parallel at four temperatures in the range 1-2.5 (NOTCH-DLL4) or 1-4 (NOTCH-JAG1) with periodic exchanges of models at adjacent temperatures following the Replica Exchange algorithm. The scoring function in the MCMC sampling was composed of the following restraints: a restraint on the end-to-end NOTCH-ECD distance, excluded volume restraint, and connectivity restraints. The transmembrane regions of all proteins were fixed during sampling. Filtering of sampled models, analysis of sampling exhaustiveness, estimating sampling precision, and model clustering followed previously established protocols [48–50]. Integrative modeling for NOTCH-DLL4 resulted in effectively a single cluster of models (98% of 20000 analyzed models) with a precision of 30 Å corresponding to the average RMSD between a cluster model and the cluster centroid (Figure S10). Similarly, a single cluster of models (99% of 20000 analyzed models) with a precision of 28 Å was obtained for NOTCH-JAG1 (Figure S11). Modeling and analysis were performed with the IMP version 2.15.0 and PMI version 2.15.0 [51].

Molecular dynamics simulations of three NOTCH pathogenic mutations

Three clinically relevant missense mutations for molecular dynamics simulations were selected. In particular, two mutations in the NOTCH ligand binding domain (R448Q and E360K) and one in the Abruption (Ax) domain (C1133Y) were selected based on pathogenicity. These mutations were introduced into the NOTCH-DLL4 and NOTCH-JAG1 atomistic models using Chimera, followed by protonation. For the ligand binding regions, four mutant systems were constructed: i) R448Q NOTCH-DLL4, ii) R448Q NOTCH-JAG1, iii) E360K NOTCH-DLL4, and iv) E360K NOTCH-JAG1. These mutant systems were based on atomistic 8-13 EGF wild-type ligand systems described above and were subjected to similar steps undertaken for wild-type NOTCH-ligand simulations with the OPLS all-atom force field using Gromacs version 2021.4. Each mutated complex was subjected to a 1 μ s long production run. For the third mutation, C1133Y located at the Abruption domain, amino acid substitution was introduced to the full-length NOTCH-ECD model, and the resulting mutated model was converted into the coarse-grained representation with Charmm-GUI. The same simulation protocols for the full-length WT NOTCH-ECD were applied to the mutant systems. The C1133Y mutant system was subjected to a 10 μ s production run.

Results and Discussion

Full-length atomistic model of NOTCH-ECD reveals variability in surface features

Fig 2. Structural characterization of human NOTCH-ECD model. (A) The schematic representation and the proposed extended conformation of the full-length atomistic model for the human NOTCH-ECD is shown in cartoon and surface representation, where 36 EGF-like repeats are colored in green and blue, alternatively; followed by NRR subdomains and TM region. (B) The proline (purple) and glycine (cyan) residue clusters at the hinge EGF-like repeats (EGF18, 22, and 32) of the bent regions. The zoomed-in panels show the relative orientation of proline & glycine with their side-chains shown in ball & stick representation. (C) Asymmetric charge distribution mapped on the surface of EGF-like repeats (EGF1-36) has been shown. The zoomed-in surface view of the slightly electropositive region; EGF1-5, electronegative EGF11-19, and EGF20-26 have been shown. In the color scale red corresponds to electronegative, while the blue color refers to the electropositive potential.

A full-length model of the human NOTCH-ECD, including the transmembrane region, was generated to understand its role in ligand binding and cell communication. The model integrates experimentally resolved structures (42%) for the EGF-like repeats (EGF4-7; 8-13), NRR domain, and transmembrane region, with the remaining EGF repeats (EGF1-3 and EGF14-36) built using fragment-based homology modeling and refined through 4 μ s of MD simulations.

The extended conformation of human NOTCH-ECD measures \sim 86.5 nm, with some inherent kinks (Figure 2A). We assessed the NOTCH-ECD extended model for these kinks that were primarily observed in three regions: EGF17-19, EGF21-23 & EGF31-33 (Figure 2B). Interestingly, the connecting EGF-like repeat in these three regions (EGF18, 22, 32) showed a significant percentage of proline and glycine residues, indicating their probable contribution towards NOTCH structural architecture as previous studies have suggested proline-glycine content to be an important structural determinant [52]. For example, EGF18 harbors proline (8.1%) and glycine (16.2%) in a 1:2 ratio, for EGF22 these percentages were found to be 10.3% each and for EGF32 the proline and glycine content were observed to be 13.3% and 11.1%, respectively. The NOTCH structure also showed asymmetric charge distribution on the protein surface (Figure 2C).

While the 36 EGF-like repeats exhibit a negative electrostatic potential ($-8.50e+01$ e), the N-terminus EGF1-5 displays a slightly positive potential ($1.00e+00$ e) at pH 7. This asymmetric electrostatic charge distribution on the elongated EGF-like repeats might influence the folding as well as the ligand binding of the receptor. Analysis of inter-domain contacts within a 5\AA cut-off revealed that EGF-like repeats near the N-terminus (EGF1-21) had fewer inter-domain contacts compared to those near the TM region (EGF21-36), with a few exceptions (Figure S12). These structural and electrostatic features provide insights into the functional modularity of the NOTCH-ECD.

Conformational dynamics of the apo-NOTCH receptor with and without glycans show EGF-domain compactness

Intracellular communication in the NOTCH protein is speculated to be regulated by structural changes in the NOTCH-ECD [44, 53]. Due to the system length complexity, we constructed a coarse-grained model of the NOTCH-ECD to investigate its dynamic behavior. The minimized all-atom linear model of NOTCH-ECD was converted into a coarse-grained model and position restraints were applied to the transmembrane region to mimic membrane anchoring. In total, three replicas of $10\ \mu\text{s}$ simulations were generated. The apo NOTCH-ECD is highly dynamic and each EGF-like repeat shows unique flexibility that drives EGFs to come together in compact states. Figure 3A, S13 shows that the protein simulation started with an initial extended conformation of the NOTCH-ECD (≈ 80.06 nm), and within 500 ns the distance between the ARG20 (first residue) to HIS1426 (last residue) reduced significantly attaining an average of $\approx 23.81 \pm 9.98$ nm across three $10\ \mu\text{s}$ trajectories leading to a compact structure. The above results are in good agreement with the previous cry-EM report that NOTCH exhibits bending motions and adopts multiple conformations [9].

Fig 3. Apo NOTCH-ECD attains multiple collapsed conformations. (A) Surface representation of the human NOTCH-ECD showing its transition from linear to compact conformations, highlighting the reduction in end-to-end distance across the EGF-like repeats. EGF-like repeats, NRR, and TM are color-coded in light pink, green, and grey, respectively. (B) A schematic representation of the 36 glycosylation sites distributed across the EGF-like repeats in NOTCH-ECD, with the structural model of glycosylated NOTCH-ECD shown in surface view. Glycans (Glucose, Fucose, and GlcNAc) are represented as blue circles, red triangles, and blue squares, respectively, following the Consortium for Functional Glycomics guidelines. (C) Time evolution of the distance measured between the backbone atoms of ARG20 (EGF1) and HIS1426 (EGF36), indicates the folding behavior of the NOTCH-ECD structure in its apo form. (D) A box plot showing the average root mean square deviation (RMSD) values for individual EGF-like repeats in glycosylated vs non-glycosylated NOTCH-ECD. (E) NOTCH residues previously reported to bind ligands are shown in light green, while glycans (Glucose, GlcNAc, and Fucose) on the EGF8-21 region surrounding the binding site are represented in bead form (blue and red, respectively).

Next, we set out to understand glycan modifications on the apo-NOTCH structure which is known to extensively regulate signaling in distinct contexts, such as protein folding, localization, and ligand binding [54, 55]. There are 36 sites with Ser/Thr containing consensus sequences which get modified with three main forms of O-linked glycans identified as O-glucose, O-fucose, and O-GlcNAc [56–58]. Figure 3B shows the glycans at the respective sites on NOTCH-ECD and the glycosylated structure model was parameterised for coarse-grained simulations (see Methodology for details). The end-to-end distance and compact conformations of glycosylated NOTCH-ECD displayed a similar trend as compared to the

non-glycosylated form, where the average distance was observed to be 48.92 ± 2.24 nm within 1 μ s time (Figure 3C).

Interestingly, the presence of specific glycans induced alterations in the local geometry of the structure (Figure 3D). While the complete non-glycosylated structure showed structural deviations from the starting structure (20.71 ± 1.87 nm), NOTCH EGF7-8, 24-25 displayed relatively higher dynamics with a median value ranging from 0.35 to 0.52 nm. In a previous cryo-EM study, substantial rearrangements of the individual EGF-like repeats have been suggested to regulate compact conformations of the NOTCH protein [9]. Distance-based metrics also showed that overall knot formation within specific EGFs remained the same. Both sets of simulations showed similar regions to be in contact, in particular, NOTCH EGF10-11 was in contact with the EGF15-16 region (Figure S14). Further, extended time-scale simulations for non-glycosylated structure revealed a second smaller knot between EGF-like repeats 1 and 2 with either EGF18 or EGF27-28 (Figure S15).

However, EGF-like repeats participating in folding with minimal contributions were observed to be EGF5-6, 19-21, 29-31, and 34-36. Hence, these results confirm that EGF1-16 folds into knot or cluster, while EGF-like repeats proximal to the TM region (EGF25-36) participate minimally in clustering and may rather impart rigidity to the NOTCH-ECD structure. Additionally, we observed a cluster of glycans around the known binding region; primarily EGF11-12, upon NOTCH-ECD folding. As shown in Figure 3E, the glycans were observed to orient outwards on the surface surrounding the binding region which may have a significant role in enhancing the accessibility for ligand-protein interaction. In a recent study by Zeronian et al., 2021, cross-linking mass spectrometry was employed to analyze intra-NOTCH and NOTCH-ligand interactions [59]. We compared the distances of residue pairs from the experimental data and simulations to investigate NOTCH intramolecular interactions. Figure S16 displays the NOTCH-ECD models, where 60% of crosslinks are satisfied, with each residue pair highlighted in red and the crosslink shown as a yellow line (Table S5). Although the datasets are not directly comparable, as the experimental study was conducted with ligands, it is still valuable to observe how contacts within NOTCH can form and how these insights align with our simulations. Together these results suggest that apo form of NOTCH-ECD on membrane prefers a compact conformation driven by specific N-terminal EGF-like repeats. The glycosylation significantly alters the local structural changes and is also abundant at functional regions that include binding sites.

Differential dynamics and preferential binding of NOTCH emerges from distinct EGF-like repeats upon binding with DLL4 and JAG2

NOTCH protein shows diverse cellular outcomes upon binding with its canonical ligands, Delta-like protein and Jagged via EGF-like repeats [60,61]. At the atomistic level, the truncated crystal complexes of NOTCH with DLL4 and JAG1 have been experimentally resolved, where NOTCH EGF11-13 and NOTCH EGF8-12 are bound to DLL4 and JAG1, respectively [11,12]. Using modeling, we constructed equivalent protein complexes of similar NOTCH EGF8-13 bound to its ligands (*see methods*). To study the dynamics of these complexes, we performed all-atom molecular dynamics simulations for four replicas of 5 μ s each, totaling 40 μ s. Compared to the NOTCH-JAG1 complex, the NOTCH-DLL4 complex exhibited higher dynamics, with an average RMSD of 2.38 ± 0.17 nm for DLL4, versus 1.55 ± 0.18 nm for JAG1 (Figure 4A). Snapshots at different time points revealed distinct NOTCH conformations, showing a compact orientation when bound to DLL4 and a more linear form with JAG1. While the dynamics of NOTCH with DLL4 (2.01 ± 0.11 nm) and JAG1 (1.71 ± 0.15 nm) were similar, DLL4 exhibited higher dynamics (1.88 ± 0.07 nm) compared to JAG1 (1.16 ± 0.10 nm). The trajectories revealed key flexibility determinants primarily from specific EGF-like repeats (Figure 4B). NOTCH EGF10 showed higher RMSD with DLL4 (0.75 ± 0.04 nm)

compared to JAG1 (0.61 ± 0.05 nm), while NOTCH EGF12 exhibited lower RMSD with DLL4 (0.53 ± 0.09 nm) than with JAG1 (0.75 ± 0.09 nm). NOTCH EGF9 and EGF11 displayed similar deviations in both complexes.

Fig 4. Differential dynamics of NOTCH upon binding to DLL4 and JAG1. (A) Representative snapshots of the NOTCH-DLL4 and NOTCH-JAG1 complexes from all four simulations are shown in cartoon representation. NOTCH bound to DLL4 and JAG1 are colored dark blue and light blue, respectively, while the ligands are shown in grey. These snapshots illustrate the dynamic behavior of NOTCH in its varying conformations in both complexes. The frequency distribution plot displays the overall backbone deviation of the NOTCH-DLL4 (dark blue) and NOTCH-JAG1 (light blue) complexes across the 5 μ s trajectory of all four simulations. (B) Distribution plots of the RMSD for individual NOTCH EGF-like repeats (EGF8-13) upon binding with DLL4 and JAG1 across the simulation trajectories, with NOTCH-EGFs bound to DLL4 and JAG1 highlighted in green and purple, respectively. (C) A line graph showing the average number of hydrogen bonds formed by NOTCH EGF9 and EGF10 with DLL4 and JAG1 across all four simulations (5 μ s each). (D) Zoomed-in structural snapshots of NOTCH EGF-like repeats 9 and 10 in the NOTCH-DLL4 and NOTCH-JAG1 complexes, with residues contributing to greater than 50% occupancy shown in ball and stick representation. Common residues interacting with both DLL4 and JAG1 are highlighted in grey, while those exclusively interacting with NOTCH-DLL4 and NOTCH-JAG1 are shown in dark orange and dark blue, respectively.

Ligand mobility was also studied, both DLL4 and JAG1 have well-defined domains, including MNNL (module at the N-terminus of the NOTCH ligands) and DSL (Delta/Serrate/Lag-2), followed by EGF-like repeats (Figure S17). We found that DLL4 EGF2 (0.53 ± 0.04 nm) displayed higher dynamics compared to JAG1 EGF2 (0.32 ± 0.02 nm), indicating greater stability when bound to JAG1. RMSF analysis showed that EGF2 of DLL4 had higher fluctuations, while EGF2 of JAG1 was more stable (Figure S18).

The dynamic properties of NOTCH-ligand interactions were found to differ, with binding mechanisms influenced by core structure-dynamics relationships, spatial distribution, and association pathways. NOTCH-ligand complexes exhibit "catch bond" behavior, where weak bonds become stronger under tensile forces [12, 62]. To analyze these weak interactions, we calculated the hydrogen bonding (H-bond) network in both complexes (Figure 4C). The NOTCH EGF9 was observed to form a higher number of H-bonds when bound to DLL4 as compared to JAG1. Whereas, NOTCH EGF10 showed a higher number of H-bonds in the NOTCH-JAG1 complex. Interestingly, NOTCH EGF11-12 displayed a similar number of hydrogen bonds upon binding to DLL4 and JAG1 (Figure S19), consistent with previous studies [63–65].

We analyzed the occupancy of NOTCH residues, and none of the residues in NOTCH-EGF10 reached more than 50% occupancy in the NOTCH-DLL4 simulations. NOTCH-EGF9 showed more interacting residues with DLL4 than JAG1, while NOTCH-EGF12 had slightly more residues interacting with JAG1 (Figure S20). NOTCH-EGF9 had 9 residues exclusively interacting with DLL4 and 7 with JAG1 (Figure 4D). NOTCH-EGF10 showed more residues interacting with JAG1, while NOTCH-EGF11 had shared interacting residues with both ligands. These findings suggest that dynamic regions of NOTCH play a key role in regulating receptor-ligand selectivity.

Fig 5. Integrative models of the NOTCH-DLL4 and NOTCH-JAG1 complexes reveal variable interacting regions. (A) The representative cluster center models of the major cluster for integrative models of the full-length ectodomain complexes NOTCH-DLL4 and NOTCH-JAG1 are shown, where NOTCH, DLL4, and JAG1 are highlighted in green, pink and blue colors, respectively. (B) The localization probability density maps display the positioning of different protein domains in the NOTCH-DLL4 and NOTCH-JAG1 clusters. These maps specify the probability of any voxel being occupied by the corresponding domain in the major cluster. Interacting domains of NOTCH, DLL4, and JAG1 are colored in darker shades of green, pink and blue, respectively. All the maps are contoured at 10% of their respective maximum voxel values. (C) The enlarged view of the marked regions in panel B is shown, where the differentially interacting domains of NOTCH with DLL4 and JAG1 are marked. NOTCH EGF8-13 (dark green) interacting with the DLL4 via its N-terminal to EGF2 (dark pink), while NOTCH N-terminal to EGF13 (dark green) interacting with JAG1 (N-terminal to EGF5; dark blue) are shown. (D) The schematic highlights the interacting NOTCH-ligand domains in the integrative complex models. The grey blocks represent the previously known interacting domains of NOTCH with DLL4 and JAG1. The above-mentioned interacting regions of NOTCH, DLL4, and JAG1 are highlighted with colored dashed boxes.

Integrative models of NOTCH-ECD with ECD of DLL4 and JAG1 show variability in the interacting regions

To explore the NOTCH dynamics upon ligand binding in the context of intracellular trans-activation, we further explored integrative modeling of full-length extracellular domains of NOTCH and its binding partners; DLL4 and JAG1. We utilized our full-length NOTCH-ECD atomistic model to generate the coarse-grained NOTCH receptor (1-1756 a.a.) spanning from N-terminal (1-19 a.a.) to the TM region. The ligand proteins; DLL4 and JAG1 are also single-pass transmembrane with an extracellular region comprising 550 and 1093 residues, respectively. To build the coarse-grained models for DLL4 and JAG1, integrative modeling was performed utilizing experimental structural details of the previously resolved fragments and data from biochemical experiments (see methods). Next, we generated the full-length extracellular domain complex models of NOTCH-DLL4 and NOTCH-JAG1, where each protein was represented by a series of connected spherical beads coarse-grained at multiple scales (Figure 5A). Regions of unknown structure were represented by 30-residues per bead modeled as flexible string of beads. In contrast, regions with known crystal co-ordinates were represented simultaneously by 1- and 10- residues per bead and modeled as rigid bodies, where the relative distances between beads are fixed during sampling.

In both integrative models, the NOTCH N-terminus was observed to bend away from the linear structure of the rest of the complex. In addition, the NOTCH N-terminus exhibited a low precision in both complexes, indicating its highly flexible nature, consistent with the observations from the coarse-grained simulations of NOTCH-ECD. On the other hand, the ligands DLL4 and JAG1 adopted a linear conformation extending from the transmembrane region of one cell (Figure 5B). Since there is no available information on binding EGF-like repeats beyond EGF8-13, we utilized these large complexes to gain insights about specific interacting regions. Interestingly, we observed that although NOTCH EGF8-13 mediates contact with both ligands, an additional region from EGF5-7 of NOTCH was observed to be in proximity specifically with JAG1 (Figure 5C-D, S21). Due to the anti-parallel orientation of the NOTCH protein with its ligands, the corresponding regions of DLL4 and JAG1 observed to be interacting were; N-terminal to EGF2 and N-terminal to EGF5, respectively. Together, these findings point towards the flexibility of the NOTCH N-terminus and highlight the EGF-specific differences in the interaction of NOTCH with its ligands.

Pathogenic NOTCH mutations affect interaction surfaces and intrinsic dynamics

We studied pathogenic mutations to understand their effects on ligand interactions and structural stability. Targeting the ligand-binding domain, we introduced two pathogenic mutations- E360K and R448Q, located in EGF9 and EGF11, respectively, to assess their impact on interaction stability with DLL4 and JAG1. These mutations are known to significantly disrupt the structural integrity of NOTCH-DLL4 complexes. The pathogenicity of R448Q is implicated in cardiac abnormalities [66]. E360 is particularly implicated in oncogenesis, highlighting its relevance in cancer studies [67–69]. Secondly, we introduced mutation C1133Y in the Abruptex domain that can lead to changes in receptor conformation that disrupt NOTCH signaling, potentially contributing to diseases like cancer. For example, the C1133Y mutation has been linked to increased proliferation and invasion in oral squamous cell carcinoma, indicating its role in the inactivation of NOTCH activity [70].

In complex with DLL4 and JAG1, we subjected E360K and R448Q mutant systems to 1 μ s all-atom MD simulations (Figure 6A-B). The residue R448 is reported to be critical for the binding with DLL4 by forming key interactions with D218, F195, and Y216 of DLL4 [12]. While the mutation is on EGF11, we observed significant reorientation to the local EGF site along with adjoining EGF9 and 12. In particular, R448Q mutation resulted in the movement of NOTCH_{EGF11} away from the DLL4_{EGF4} and DLL4_{DSL} to 7.12 ± 0.46 nm and 1.73 ± 0.06 nm (Figure S22). While these spatial reorientations result in local instability to NOTCH_{EGF9}, there was also an overall decrease in the number of hydrogen bonds, especially between NOTCH_{EGF10-11} and DLL4, from an average number of approximately 10 to 4 (Figure 6C). The second ligand binding mutation was a charge-reverting mutation at NOTCH_{EGF9} – E360K. As shown in Figure 6D, we observed a relatively higher disassociation of receptor-ligand interactions. The NOTCH_{EGF8-10} showed complete detachment from the binding interface formed with DLL4_{EGF1-3}, leading to the complete abrogation of interaction (Figure S22). Additionally, higher dynamics for mutant E360K NOTCH protein were observed, with RMSD of the mutant shifting from 2.01 ± 0.47 nm to 3.74 ± 0.21 nm.

In comparison with DLL4, mutant simulations with JAG1 exhibited less effect on complex stability (Figure 6E-F). However, it is noteworthy that these mutations have local structural perturbations. The R448Q mutant led to reduced interactions of NOTCH_{EGF9-10} with the JAG1. The mutation primarily perturbs the average hydrogen bonding of NOTCH_{EGF10} with JAG1 (Figure 6G). The E360K mutation also affected the dynamics and local interaction network of NOTCH. NOTCH_{EGF11} showed a higher RMSD of 1.13 ± 0.23 nm [WT NOTCH_{EGF11} RMSD: 0.47 ± 0.09], while the interactions between NOTCH_{EGF9} and JAG1 were disrupted, leading to the loss of two H-bonding contacts (Figure 6H).

Next, to investigate the impact of mutations in the folding dynamics of apo NOTCH, we studied the C1133Y mutation on the Ax domain that ranges from EGF24 to EGF29. The Abruptex domain is proposed to mediate both inter- and intra- molecular interactions of the NOTCH receptor, particularly through contacts involving the ligand-binding region, such as NOTCH_{EGF11-EGF12}, as described by Pei & Baker (2008) (Figure 6I) [71]. Mutations in the Ax domain, such as C1133Y in EGF29, have been linked to increased proliferation in oral squamous cell carcinoma, with experimental evidence suggesting that this mutation disrupts NOTCH signaling, leading to its inactivation.

Fig 6. Pathogenic mutations affect the ligand binding and intrinsic dynamics of NOTCH (A-B) Overlaid conformations of wild-type (WT) NOTCH_{LBD}-DLL4_{LBD} from 100 ns to 1 μ s, at 100 ns intervals, shown in cartoon representation. The positions of R448 and E340 are marked as red spheres. The spatial orientation of NOTCH EGFs relative to the DLL4 domains is shown in a schematic next to the snapshot, with R448 and E360 indicated as stars. (C-D) Ensemble of conformations for R448Q and E360K NOTCH_{LBD}-DLL4_{LBD}, shown in cartoon representation with respective mutations highlighted as red spheres. The schematic representation of the mutated NOTCH-DLL4 orientation is shown next to the snapshot. The probability distribution compares the RMSD of mutated NOTCH_{LBD}, i.e., EGF9 and EGF13 with respect to WT NOTCH. Bar plots display the average number of hydrogen bonds (H-bonds) between NOTCH_{EGF10-11} and DLL4 for R448Q mutant, and between NOTCH_{EGF9} and DLL4 for E360K mutant. (E-F) Overlaid conformations of WT NOTCH_{LBD}-JAG1_{LBD} from 100 ns to 1 μ s, at 100 ns intervals, in cartoon representation. R448 and E340 positions are marked as red spheres. (G-H) Ensemble of conformations for R448Q and E360K NOTCH_{LBD}-JAG1_{LBD}, shown in cartoon representation with respective mutations as red spheres. The schematic of the mutated NOTCH-JAG1 orientation is shown next to the snapshot. The probability distribution compares the RMSD of mutated NOTCH and EGF11 with respect to WT NOTCH. Bar plots highlight the average number of H-bonds observed between NOTCH_{EGF9-10} and JAG1 for R448Q, and NOTCH_{EGF9-11} and JAG1 for E360K. (I) Full-length coarse-grained model of NOTCH-ECD, where the ligand-binding domain (LBD) and Abruption (Ax) domain are highlighted in dark brown and salmon, respectively. The C1133 position is highlighted in red. (J) Normalized average contacts between LBD and Ax domain, calculated based on inter-EGF distances for WT and C1133Y mutant, shown in purple and green, respectively (K) Distribution plots of distances between LBD (EGF9-11) and Ax domain EGFs 24 and 28, with WT and mutant data highlighted in purple and green, respectively. Representative snapshot of WT NOTCH-ECD in folded conformation, with the LBD in closer proximity to the Ax domain, compared to the C1133Y NOTCH-ECD, showing a distal orientation of the LBD with respect to the Abruption domain. The position of the C1133Y mutation is marked in red.

We introduced the mutation at the Ax domain in the coarse-grained structure, as simulated in wild-type NOTCH full-length ECD domain. In two out of three wild-type simulations, three ligand-binding EGFs were bound to the Ax domain (Figure 6J). However, upon the introduction of the C1133Y mutation, these contacts are completely abrogated, and the distance between EGF centers in WT varies from \sim 8 nm to \sim 12 nm, while upon mutation, the distance increases to \sim 11 to \sim 17 nm (Figure 6K). Consistent with Pei & Baker’s findings, our simulations of the wild-type NOTCH receptor demonstrated interactions between the Ax domain and the ligand-binding domain, supporting their functional interplay. Together, these mutations not only impair stable receptor-ligand interactions but also alter intrinsic dynamics and inter-domain communication within NOTCH, highlighting their role in disease pathogenesis through the disruption of structural and functional integrity.

Conclusion

NOTCH is a crucial stem cell receptor involved in cell-to-cell communication through its extracellular domain, and its mutations are implicated in various diseases, including cancer and rare genetic disorders. Over 50% of NOTCH mutations occur in the EGF-like repeat domains [72], but a complete structural understanding of the NOTCH ectodomain remains elusive, with therapeutic strategies focusing mainly on limited regions like the NRR domain [73–75]. Understanding how the NOTCH receptor achieves specificity in binding to its ligands has been explored in various contexts [37, 60]. Force-clamp spectroscopy of

truncated NOTCH-ligand complexes has demonstrated differential binding affinities, mediated by the "catch bond" nature of interactions between the proteins. Additionally, atomic force microscopy studies suggest strong adhesion between cells expressing NOTCH and DLL proteins, highlighting the robust interaction between these molecules [12, 76]. In this study, we used a modeling and multiscale simulation approach to investigate two key aspects of NOTCH. First, we modeled the full-length human NOTCH-ECD, revealing its dynamic behavior in the apo form with glycan moieties, showing that it adopts both extended and compact conformations. Second, we examined NOTCH binding to its canonical ligands, DLL4 and JAG1 using atomistic simulations of shorter truncated bound complexes and integrative modeling for full-length NOTCH-receptor complexes. Our simulations revealed distinct conformations for each ligand, and clear dynamics from the N-terminal of NOTCH, particularly in the more dynamic DLL4-bound form. Three clinically relevant mutations—R448Q and E360K in the NOTCH ligand-binding domain and C1133Y in the Abruption region—were introduced into NOTCH-ligand complexes and NOTCH-ECD models, respectively, for molecular dynamics simulations. Comparative analysis with wild-type NOTCH-DLL4 and NOTCH-JAG1 simulations showed that E360K and R448Q disrupted interaction surfaces, while C1133Y destabilized the Abruption domain in NOTCH-ECD apo form. These results provide a comprehensive molecular understanding of NOTCH and its interaction with ligands, shedding light on the structural and dynamic aspects that govern receptor-ligand binding and signaling. By identifying how specific mutations disrupt binding surfaces and stability, we gain crucial insights into the mechanisms that underlie NOTCH-related diseases, such as cancer and genetic disorders. In a study by Wang et al., specific mutations such as D469G in the EGF-like domain impair calcium coordination critical for ligand binding, while R1594Q in the juxtamembrane region disrupts ligand-mediated activation of NOTCH1 [77]. Mutations in the RAM domain, like P1770S, further impair transcription complex formation, collectively disrupting NOTCH signaling and driving tumorigenesis. These findings emphasize the need for more structural and genomic studies to better understand the tumor-suppressive role of NOTCH in cancers. For future studies, detailed knowledge can help guide the design of more precise therapeutic interventions that target NOTCH signaling, either by restoring proper receptor-ligand interactions or by selectively modulating dysfunctional pathways associated with pathogenic mutations.

Supporting information

This article contains supporting information (Figures S1-S22, Table S1-S5).

Data availability

The data files for the full-length NOTCH-ECD atomistic as well as coarse-grained models along with NOTCH-DLL4 and NOTCH-JAG1 dynamics can be accessed from <https://github.com/Surabhi-CSB/NOTCH>. The input data, scripts, and results for multiscale integrative modeling of NOTCH-ligand complexes are available at http://github.com/isblab/notch_ligand.

Author contributions

SR: Conceptualization, Investigation, Data curation, Formal analysis, Visualization, Writing-original draft; DG: Data curation, Visualization; JC: Data curation, Formal analysis AP: Data curation, Formal

analysis; SA: Data curation, Visualization; SV: Conceptualization, Formal analysis; LT: Conceptualization, Supervision, Writing-original draft, Project administration.

Declaration of Interests

The authors declare that they have no conflicts of interest with the contents of this article.

Acknowledgments

This work was supported by funding from the Department of Science and Technology (DST) and Council of Scientific and Industrial Research (CSIR). SR was supported by DST-Inspire fellowship and DG by CSIR-NET fellowship. SV acknowledges Science Engineering and Research Board (SERB, grant number SPG/2020/000475) and Department of Atomic Energy (DAE, grant number RTI 4006) for the support. LT is thankful to the support from the CSIR-Young scientist grant OLP1151 and DBT NNP project (BT/PR40265/BTIS/137/62/2023). We acknowledge Prof. Christina Redfield, University of Oxford for sharing the PDB structures of EGF4-9 and 10-13 with us. We also acknowledge the support from CSIR-IGIB for infrastructure and CSIR-4PI for supercomputing facilities.

References

1. Gazave E, Lapébie P, Richards GS, Brunet F, Ereskovsky AV, Degnan BM, et al. Origin and evolution of the Notch signalling pathway: an overview from eukaryotic genomes. *BMC evolutionary biology*. 2009;9(1):1-27.
2. Logeat F, Bessia C, Brou C, LeBail O, Jarriault S, Seidah NG, et al. The Notch1 receptor is cleaved constitutively by a furin-like convertase. *Proceedings of the National Academy of Sciences*. 1998 jul;95(14):8108-12.
3. Pancewicz J, Nicot C. Current views on the role of Notch signaling and the pathogenesis of human leukemia. *BMC Cancer*. 2011 nov;11(1).
4. Sakamoto K, Chao WS, Ichi Katsube K, Yamaguchi A. Distinct roles of EGF repeats for the Notch signaling system. *Experimental Cell Research*. 2005 jan;302(2):281-91.
5. Kakuda S, LoPilato RK, Ito A, Haltiwanger RS. Canonical Notch ligands and Fringes have distinct effects on NOTCH1 and NOTCH2. *Journal of Biological Chemistry*. 2020;295(43):14710-22.
6. Kopan R, Ilagan MXG. The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell*. 2009;137(2):216-33.
7. Hambleton S, Valeyev NV, Muranyi A, Knott V, Werner JM, McMichael AJ, et al. Structural and Functional Properties of the Human Notch-1 Ligand Binding Region. *Structure*. 2004 dec;12(12):2173-83.
8. Weissshuhn PC, Sheppard D, Taylor P, Whiteman P, Lea SM, Handford PA, et al. Non-Linear and Flexible Regions of the Human Notch1 Extracellular Domain Revealed by High-Resolution Structural Studies. *Structure*. 2016 apr;24(4):555-66.

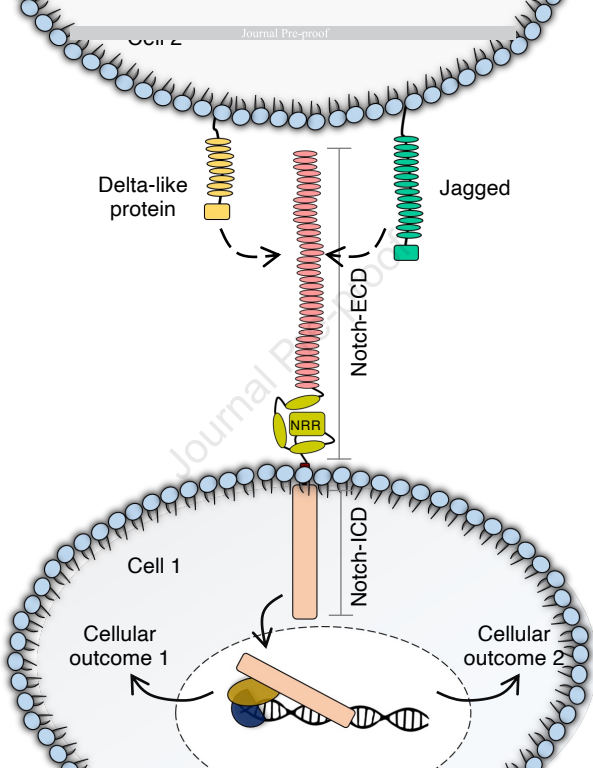
9. Kelly DF, Lake RJ, Middelkoop TC, Fan HY, Artavanis-Tsakonas S, Walz T. Molecular structure and dimeric organization of the Notch extracellular domain as revealed by electron microscopy. *PLoS one*. 2010;5(5):e10532.
10. Takeuchi H, Yu H, Hao H, Takeuchi M, Ito A, Li H, et al. O-Glycosylation modulates the stability of epidermal growth factor-like repeats and thereby regulates Notch trafficking. *Journal of Biological Chemistry*. 2017;292(38):15964-73.
11. Luca VC, Jude KM, Pierce NW, Nachury MV, Fischer S, Garcia KC. Structural basis for Notch1 engagement of Delta-like 4. *Science*. 2015 feb;347(6224):847-53.
12. Luca VC, Kim BC, Ge C, Kakuda S, Wu D, Roein-Peikar M, et al. Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. *Science*. 2017 mar;355(6331):1320-4.
13. Weng AP, Ferrando AA, Lee W, Morris IV JP, Silverman LB, Sanchez-Irizarry C, et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science*. 2004;306(5694):269-71.
14. Louvi A, Artavanis-Tsakonas S. Notch and disease: A growing field. *Seminars in Cell & Developmental Biology*. 2012 jun;23(4):473-80.
15. Zhou B, Lin W, Long Y, Yang Y, Zhang H, Wu K, et al. Notch signaling pathway: architecture, disease, and therapeutics. *Signal Transduction and Targeted Therapy*. 2022 mar;7(1).
16. Chillakuri CR, Sheppard D, Lea SM, Handford PA. Notch receptor–ligand binding and activation: Insights from molecular studies. *Seminars in Cell & Developmental Biology*. 2012;23(4):421-8.
17. Sulis ML, Williams O, Palomero T, Tosello V, Pallikuppam S, Real PJ, et al. NOTCH1 extracellular juxtamembrane expansion mutations in T-ALL. *Blood, The Journal of the American Society of Hematology*. 2008;112(3):733-40.
18. Cordle J, Johnson S, Tay JZY, Roversi P, Wilkin MB, de Madrid BH, et al. A conserved face of the Jagged/Serrate DSL domain is involved in Notch trans-activation and cis-inhibition. *Nature Structural & Molecular Biology*. 2008 jul;15(8):849-57.
19. Rand MD, Lindblom A, Carlson J, Villoutreix BO, Stenflo J. Calcium binding to tandem repeats of EGF-like modules. Expression and characterization of the EGF-like modules of human Notch-1 implicated in receptor-ligand interactions. *Protein Science*. 2008 dec;6(10):2059-71.
20. Deatherage CL, Lu Z, Kroncke BM, Ma S, Smith JA, Voehler MW, et al. Structural and biochemical differences between the Notch and the amyloid precursor protein transmembrane domains. *Science Advances*. 2017 apr;3(4).
21. Gordon WR, Roy M, Vardar-Ulu D, Garfinkel M, Mansour MR, Aster JC, et al. Structure of the Notch1-negative regulatory region: implications for normal activation and pathogenic signaling in T-ALL. *Blood*. 2009 apr;113(18):4381-90.
22. Do CB, Mahabhashyam MSP, Brudno M, Batzoglou S. ProbCons: Probabilistic consistency-based multiple sequence alignment. *Genome Research*. 2005 feb;15(2):330-40.
23. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Research*. 2015 apr;43(W1):W174-81.

24. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography*. 1993;26(2):283-91.
25. Xu D, Zhang Y. Improving the Physical Realism and Structural Accuracy of Protein Models by a Two-Step Atomic-Level Energy Minimization. *Biophysical Journal*. 2011 nov;101(10):2525-34.
26. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-9.
27. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic acids research*. 2022;50(D1):D439-44.
28. Malhotra N, Khatri S, Kumar A, Arun A, Daripa P, Fatih S, et al. AI-based AlphaFold2 significantly expands the structural space of the autophagy pathway. *Autophagy*. 2023;19(12):3201-20.
29. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *The Journal of Physical Chemistry B*. 2007 jun;111(27):7812-24.
30. Qi Y, Ingólfsson HI, Cheng X, Lee J, Marrink SJ, Im W. CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *Journal of Chemical Theory and Computation*. 2015 aug;11(9):4486-94.
31. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015 sep;1-2:19-25.
32. Takeuchi H, Haltiwanger RS. Significance of glycosylation in Notch signaling. *Biochemical and biophysical research communications*. 2014;453(2):235-42.
33. Pandey A, Niknejad N, Jafar-Nejad H. Multifaceted regulation of Notch signaling by glycosylation. *Glycobiology*. 2021;31(1):8-28.
34. Varshney S, Stanley P. Multiple roles for O-glycans in Notch signalling. *FEBS letters*. 2018;592(23):3819-34.
35. Shivgan AT, Marzinek JK, Huber RG, Krah A, Henchman RH, Matsudaira P, et al. Extending the Martini coarse-grained force field to N-glycans. *Journal of Chemical Information and Modeling*. 2020;60(8):3864-83.
36. Gowers RJ, Linke M, Barnoud J, Reddy TJ, Melo MN, Seyler SL, et al. MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. In: *Proceedings of the 15th python in science conference*. vol. 98. SciPy Austin, TX; 2016. p. 105.
37. Suckling RJ, Korona B, Whiteman P, Chillakuri C, Holt L, Handford PA, et al. Structural and functional dissection of the interplay between lipid and Notch binding by human Notch ligands. *The EMBO journal*. 2017;36(15):2204-15.
38. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *The Journal of chemical physics*. 2007;126(1):014101.

39. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*. 1981 dec;52(12):7182-90.
40. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science*. 2017 sep;27(1):14-25.
41. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera?A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*. 2004;25(13):1605-12.
42. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*. 1996 feb;14(1):33-8.
43. Chillakuri CR, Sheppard D, Ilagan MXG, Holt LR, Abbott F, Liang S, et al. Structural Analysis Uncovers Lipid-Binding Properties of Notch Ligands. *Cell Reports*. 2013 nov;5(4):861-7.
44. Kovall RA, Gebelein B, Sprinzak D, Kopan R. The Canonical Notch Signaling Pathway: Structural and Biochemical Insights into Shape, Sugar, and Force. *Developmental Cell*. 2017 may;41(3):228-41.
45. Rudenko G, Henry L, Henderson K, Ichtchenko K, Brown MS, Goldstein JL, et al. Structure of the LDL Receptor Extracellular Domain at Endosomal pH. *Science*. 2002 dec;298(5602):2353-8.
46. Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology*. 1993;234(3):779-815.
47. Saltzberg D, Greenberg CH, Viswanath S, Chemmama I, Webb B, Pellarin R, et al. Modeling Biological Complexes Using Integrative Modeling Platform. In: *Methods in Molecular Biology*. Springer New York; 2019. p. 353-77.
48. Saltzberg DJ, Viswanath S, Echeverria I, Chemmama IE, Webb B, Sali A. Using Integrative Modeling Platform to compute, validate, and archive a model of a protein complex structure. *Protein Science*. 2020 dec;30(1):250-61.
49. Arvindekar S, Jackman MJ, Low JK, Landsberg MJ, Mackay JP, Viswanath S. Molecular architecture of nucleosome remodeling and deacetylase sub-complexes by integrative structure determination. *Protein Science*. 2022;31(9):e4387.
50. Viswanath S, Chemmama IE, Cimermanic P, Sali A. Assessing Exhaustiveness of Stochastic Sampling for Integrative Modeling of Macromolecular Structures. *Biophysical Journal*. 2017 dec;113(11):2344-53.
51. Russel D, Lasker K, Webb B, Velázquez-Muriel J, Tjioe E, Schneidman-Duhovny D, et al. Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS biology*. 2012;10(1):e1001244.
52. Rauscher S, Baud S, Miao M, Keeley FW, Pomes R. Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. *Structure*. 2006;14(11):1667-76.
53. Greenwald I, Kovall R. Notch signaling: genetics and structure. *WormBook: The Online Review of C elegans Biology [Internet]*. 2018.

54. Okajima T, Xu A, Lei L, Irvine KD. Chaperone activity of protein O-fucosyltransferase 1 promotes notch receptor folding. *Science*. 2005;307(5715):1599-603.
55. Sasamura T, Ishikawa HO, Sasaki N, Higashi S, Kanai M, Nakao S, et al. The O-fucosyltransferase O-fut1 is an extracellular component that is essential for the constitutive endocytic trafficking of Notch in *Drosophila*. *Development*. 2007 Apr;134(7):1347-56. Epub 2007 Feb 28.
56. Harvey BM, Rana NA, Moss H, Leonardi J, Jafar-Nejad H, Haltiwanger RS. Mapping sites of O-glycosylation and fringe elongation on *Drosophila* Notch. *Journal of Biological Chemistry*. 2016;291(31):16348-60.
57. Rana NA, Nita-Lazar A, Takeuchi H, Kakuda S, Luther KB, Haltiwanger RS. O-glucose trisaccharide is present at high but variable stoichiometry at multiple sites on mouse Notch1. *Journal of Biological Chemistry*. 2011;286(36):31623-37.
58. Urata Y, Saiki W, Tsukamoto Y, Sago H, Hibi H, Okajima T, et al. Xylosyl extension of O-glucose glycans on the extracellular domain of NOTCH1 and NOTCH2 regulates Notch cell surface trafficking. *Cells*. 2020;9(5):1220.
59. Zeronian MR, Klykov O, Portell i de Montserrat J, Konijnenberg MJ, Gaur A, Scheltema RA, et al. Notch–Jagged signaling complex defined by an interaction mosaic. *Proceedings of the National Academy of Sciences*. 2021;118(30):e2102502118.
60. Benedito R, Roca C, Sörensen I, Adams S, Gossler A, Fruttiger M, et al. The Notch Ligands Dll4 and Jagged1 Have Opposing Effects on Angiogenesis. *Cell*. 2009 jun;137(6):1124-35.
61. Nandagopal N, Santat LA, LeBon L, Sprinzak D, Bronner ME, Elowitz MB. Dynamic ligand discrimination in the notch signaling pathway. *Cell*. 2018;172(4):869-80.
62. Sokurenko EV, Vogel V, Thomas WE. Catch-bond mechanism of force-enhanced adhesion: counter-intuitive, elusive, but... widespread? *Cell host & microbe*. 2008;4(4):314-23.
63. Fehon RG, Kooh PJ, Rebay I, Regan CL, Xu T, Muskavitch MA, et al. Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in *Drosophila*. *Cell*. 1990;61(3):523-34.
64. Klose R, Berger C, Moll I, Adam MG, Schwarz F, Mohr K, et al. Soluble Notch ligand and receptor peptides act antagonistically during angiogenesis. *Cardiovascular research*. 2015;107(1):153-63.
65. Rebay I, Fleming RJ, Fehon RG, Cherbas L, Cherbas P, Artavanis-Tsakonas S. Specific EGF repeats of Notch mediate interactions with Delta and Serrate: implications for Notch as a multifunctional receptor. *Cell*. 1991;67(4):687-99.
66. Southgate L, Sukalo M, Karountzos AS, Taylor EJ, Collinson CS, Ruddy D, et al. Haploinsufficiency of the NOTCH1 receptor as a cause of Adams–Oliver syndrome with variable cardiac anomalies. *Circulation: Cardiovascular Genetics*. 2015;8(4):572-81.
67. Zhang L, Zhou Y, Cheng C, Cui H, Cheng L, Kong P, et al. Genomic analyses reveal mutational signatures and frequently altered genes in esophageal squamous cell carcinoma. *The American Journal of Human Genetics*. 2015;96(4):597-611.

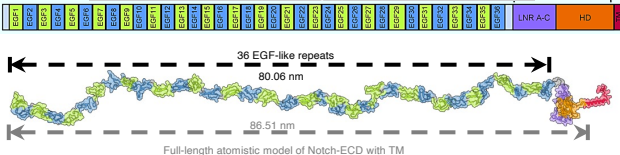
68. Chang J, Tan W, Ling Z, Xi R, Shao M, Chen M, et al. Genomic analysis of oesophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and genomic alterations. *Nature communications*. 2017;8(1):15290.
69. Cheng C, Cui H, Zhang L, Jia Z, Song B, Wang F, et al. Genomic analyses reveal FAM84B and the NOTCH pathway are associated with the progression of esophageal squamous cell carcinoma. *Gigascience*. 2016;5(1):s13742-015.
70. Zheng Y, Wang Z, Ding X, Zhang W, Li G, Liu L, et al. A novel Notch1 missense mutation (C1133Y) in the Abruptex domain exhibits enhanced proliferation and invasion in oral squamous cell carcinoma. *Cancer Cell International*. 2018;18:1-15.
71. Pei Z, Baker NE. Competition between Delta and the Abruptex domain of Notch. *BMC developmental biology*. 2008;8:1-14.
72. Mutvei AP, Fredlund E, Lendahl U. Frequency and distribution of Notch mutations in tumor cell lines. *BMC cancer*. 2015;15(1):1-11.
73. Aste-Amézaga M, Zhang N, Lineberger JE, Arnold BA, Toner TJ, Gu M, et al. Characterization of Notch1 antibodies that inhibit signaling of both normal and mutated Notch1 receptors. *PloS one*. 2010;5(2):e9094.
74. Falk R, Falk A, Dyson MR, Melidoni AN, Parthiban K, Young JL, et al. Generation of anti-Notch antibodies and their application in blocking Notch signalling in neural stem cells. *Methods*. 2012;58(1):69-78.
75. Sharma A, Gadkari RA, Ramakanth SV, Padmanabhan K, Madhumathi DS, Devi L, et al. A novel monoclonal antibody against Notch1 targets leukemia-associated mutant Notch1 and depletes therapy resistant cancer stem cells in solid tumors. *Scientific reports*. 2015;5(1):1-18.
76. Ahimou F, Mok LP, Bardot B, Wesley C. The adhesion force of Notch with Delta and the rate of Notch signaling. *The Journal of cell biology*. 2004;167(6):1217-29.
77. Wang NJ, Sanborn Z, Arnett KL, Bayston LJ, Liao W, Proby CM, et al. Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. *Proceedings of the National Academy of Sciences*. 2011;108(43):17761-6.



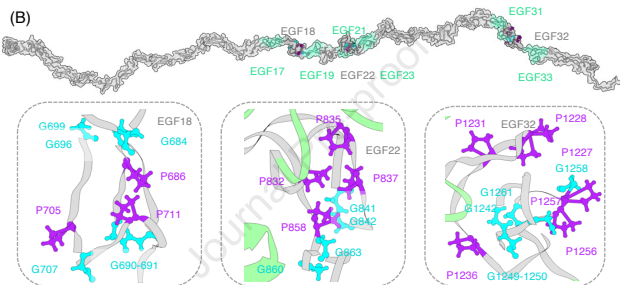
(A)

Journal Pre-proof

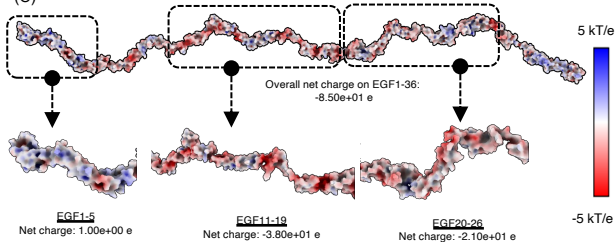
NRR



(B)

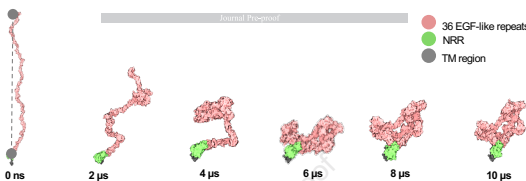


(C)

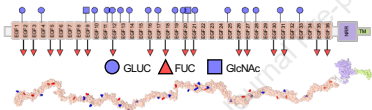


- 36 EGF-like repeats
- NRR
- TM region

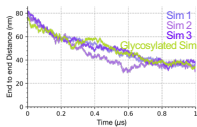
(A)



(B)



(C)



(D)



(E)

