Journal of Biotechnology and Bioengineering
Volume 5, Issue 1, 2021, PP 1-12
ISSN 2637-5362
DOI: https://doi.org/10.22259/2637-5362.0501001

Microsecond-Timescale Conformational Dynamics Reveals the Flexibility and Communication Pathway of the KIT Extracellular Domain in Complex with Stem Cell Factor

Wen-Shyong Tzou[1*],Chung-Hao Li[1,2], Pang-Hung Hsu[1],Tun-Wen Pai[3,4], Chin-Hwa Hu[1]

[1]Department of Bioscience and Biotechnology, National Taiwan Ocean University, Taiwan

[2]Centers for Disease Control, Taipei, Taiwan

[3]Department of Computer Science and Engineering, National Taiwan Ocean University, Taiwan

[4]Department of Computer Science and Information Engineering, National Taipei University of Technology, Taiwan

***Corresponding Author:** Wen-Shyong Tzou, Department of Bioscience and Biotechnology, National Taiwan Ocean University, Taiwan.*

ABSTRACT

Stem cell factor (SCF) (ligand)-KIT (receptor) interaction plays the pivotal roles in mediating a diverse biological process such as development, differentiation, cancer, and metabolism. For SCF-KIT dimer-dimer interaction, the binding of ligand to the receptor leads to the dimerization of the receptor and subsequent cross-phosphorylation between two KIT protomers, resulting into the relay of the signal transduction. In this study, we intend to investigate the conformational flexibility of KIT extracellular domain in complex with SCF in a molecular simulation of one microsecond. We discovered that the dimer conformation is intact during the simulation, while the plasma membrane-proximal domains of KIT extracellular domain demonstrated the conformational flexibility by conducting a large scale of pseudo-torsional motion. Furthermore, based on the dihedral angles of protein backbone extracted from the trajectories of the simulation, we found the communication pathway from the ligand through the receptor in a distance as long as 170 angstroms.

Keywords: *Molecular Dynamics Simulation; Flexibility; Communication Pathway; KIT Extracellular Domain; Stem Cell Factor*

INTRODUCTION

Stem Cell Factor (SCF) is the ligand for the tyrosine kinase activation of its cognate receptor (the receptor tyrosine kinase (RTK) KIT). KIT belongs to type III receptor tyrosine kinase (RTKIII) family which contains five immune globulin (ig)-like extracellular domains, trans membrane domain and the intracellular protein tyrosine kinase (PTK) domain with a large kinase-insert region[1]. SCF-KIT is responsible for hematopoiesis, melanogenesis and fertility [2] and its gain-of-function mutation is linked to many cancers including gastro-intestinal-stromal tumors and acute myeloid leukemia. RTKIII family (hCSF1R, Flt3, and PDGFR a/b)shares the activation mechanism by which the ligand binding to the receptor brings two protomers receptors together and the dimerization of the receptor leads to the intracellular auto-phosphorylation and initiates intracellular signaling. The crystal structure shows the dimeric SCF brings together two ecto domains of KIT. Conformational change is required for KIT to enable the lateral interaction between the membrane-proximal fourth immunoglobulin-like domains of KIT to contact each other for stabilizing the complex. Besides, the fourth and fifth interfaces from two protomers were demonstrated to be important for the signal transduction in the culture cell experiment [1].

Several studies had characterized the amino acid residues that are potentially vital to allosteric regulation on the intracellular catalytic domain of KIT. Groups of residues or protein regions that mediate short- and long-range communication were identified based on the molecular dynamics simulation of the catalytic domain of KIT [3]. Communication pathways of the catalytic domain of KIT and CSF-1R in the native and mutated states were also reported[4]. But until now there is no research aiming to investigate the conformational dynamics of KIT

extracellular domain in complex with SCF. There is also no research aiming to discovering the communication pathway in SCF-KIT complex to facilitate the long-range communication from SCF to KIT extracellular domain.

Material and Methods

Molecular Dynamics (MD) Simulation

We used the crystal structure of KIT extracellular domain in complex with SCF (SCF: 3-130; KIT: 33-507; 1206 residues in total; PDB ID: 2E9W)[1] to conduct the molecular dynamics simulation. The molecular dynamics simulation was conducted by employing CHARMM36 force field for protein and TIP3 model for water molecules. The molecular simulation system was generated by CHARM-GUI [5]. The all-hydrogen simulation system for KIT extracellular domain in complex with SCF includes 22,361 water molecules, 646 potassium and 628 chloride ions in a neutral-charge cubic box (194 angstroms x 194 angstroms x194 angstroms). We conducted energy minimization first by employing the steepest descent minimization with 8 angstroms as the non-bonded cutoff. Then an NVT (constant number (N), volume (V), and temperature (T)) for 25 ps and NPT (constant number (N), pressure (P), and temperature (T)) for 10 ns were conducted for equilibration. A productive molecular dynamics was conducted in NPT condition for one microsecond.

All molecular dynamics were conducted by employing Langevin dynamics in 303.15 K (30 degrees Celsius) with friction coefficient 1.0 ps^1. Pressure coupling employs isotropic MC barostat for NPT and 1.0 bar as the reference pressure. During simulation, SHAKE algorithm was employed by constraining bonds containing hydrogens and not calculating forces of bonds containing hydrogen. The simulation was conducted in periodic boundary condition (PBC). Particle-mesh Ewald summation was used for the long-range electrostatic interactions. During molecular dynamics, we employed 2 fs as the integration time step. Pmed. cuda in Amber [6]was employed in the molecular dynamics in a personal computer with Graphics Processing Unit (GPU) GeForce GTX 980.

Analysis of the KIT extracellular domain in complex with SCF conformation of the trajectories of Molecular Dynamics

We employed sasa, hbond, and gyrate of Gromacs [7] to compute the solvent accessible surface of whole complex, hydrogen bonding, and radius of gyration, respectively. For analysis of the geometry of KIT extracellular domain in complex with SCF, we defined the center of mass and principal axes of each extracellular ig-like domain of KIT of each protomer and SCF dimer. We wrote python scripts to compute the distance, included angle and pseudo-torsion angles between sub-domains. Computation of structurally related parameters was performed using the well-equilibrated conformations after 200 ns of the productive molecular simulation.

Mutual Information and Communication Pathway

One microsecond trajectory for KIT extracellular domain in complex with SCF was used to infer the correlation (mutual information) of dihedral angles of protein backbone by essentially following the procedures employed in G-protein coupled receptors [8]. The dihedral angles of each amino acid were extracted from each frame of the trajectory. For each pair of residues, the amount of mutual information was defined by calculating the Shannon entropy [9] according to the following formula,

$$S(x) = R(ln(Nh(n)) - \frac{1}{N}\sum_{n=1}^{nrBins}\{k(n)\ln(k(n))\} - \frac{nrBins-1}{2N}$$

where R is the gas constant, $h(n)$ the bin size (10 degrees for the dihedral angle), N the total number of datapoints, $k(n)$ the number of datapoints in bin n, and nr Bins is 35 for the number of the bins. Pair-wise mutual information (MI) is computed as $I(x,y) = S(x) + S(y) - S(x,y)$, where $I(x,y)$ denotes the mutual information (MI) between two amino acids, $S(x)$ and $S(y)$ denote the Shannon entropy of amino acids x and y, respectively, and $S(x,y)$ is the Shannon entropy between amino acid x and y. Subsequently, for each pair of residues, a pathway was established by maximizing the summed mutual information along the connecting path and minimizing the number of the intermediate residues along the path. A "pipeline" was therefore constructed by clustering the pathways, which maximized the overlap of pathways in the same pipeline and minimized the overlap of pathways belonging to

the different pipelines. The method was implemented using a commercial software package (MATLAB 2016, The Math Works Inc., Natick, MA).

RESULTS

Conformational Analysis of KIT Extracellular Domain in Complex with Scf during Molecular Dynamics Simulation

We employed the crystal structure of KIT extracellular domain in complex with SCF (Additional file 1: Fig. S1) (PDB ID: 2E9W)[1], immersed the protein in a water box, and conducted the molecular dynamics simulation for one microsecond. Close investigation of the trajectory of molecular dynamics simulation reveals that SCF-KIT interaction and dimer structures remain intact during the simulation; however, KIT extracellular domain exhibits conformation flexibility. The average rmsd (root mean square of deviation) of backbone of SCF-KIT is 5.7 angstroms. We did not find significant changes of the secondary structures, radius of gyration, solvent accessible surface of whole complex, and hydrogen bonding (in protein and with water molecules) (Additional file 1: Fig. S2~7).

Analysis of the Geometry of KIT Extracellular Domain in Complex with SCF during Molecular Dynamics Simulation

We further examined the geometric features of KIT extracellular domain during molecular dynamics simulation. We demarcated five extracellularig-like domains of receptor KIT extracellular domain into D1, D2, D3, D4 and D5 of two protomers (by following the nomenclature in [1]).We defined the center of mass and principal axes of each ig-like domain of KIT of each protomer. By analyzing the distance, included angle and pseudo-torsion angles between neighbored domains (Additional file 1: Fig. S8), we found most of the relative motion of domains proceeded in a quite stable manner. The standard deviations of the distance between domains during molecular dynamics simulation are mostly less than one angstrom. The standard deviations of the included angle between domains during molecular dynamics simulation are mostly less than six degrees. The standard deviations of the pseudo-torsion angle between domains during molecular dynamics simulation are mostly less than 12 degrees (Additional file 1: Fig. S9~S13). Exceptionally, the pseudo-torsional motion of D4 vs D5 (39.5 and 34.1 degrees for each protomer, respectively) (Fig. 1) varies a lot. We concluded that the conformational dynamics of SCF-KIT dimer proceeds in a way that still preserves the integrity of dimer. But some neighbored ig-like domains conducted a significant pseudo-torsional motion.

Molecular Dynamics Reveals Long-Distance Communication Pathway of KIT Extracellular Domain in Complex with SCF

In order to discover the communication pathway facilitated by the concerted movements of KIT extracellular domain in complex with SCF, we extracted the dihedral angles of the protein backbone from the trajectory of molecular dynamics simulation. We utilized the dihedral angles to calculate the time-averaged mutual information based on Shanon entropy. The higher the time-averaged mutual information between two residues, the more likely the two residues have correlated motion during the molecular dynamic simulation. When the mutual information is binned by the distance between amino acid residues, we found an interesting pattern of time-averaged mutual information (Fig.2). There is high time-averaged mutual information in the short distance (4 angstroms of inter-residue C-alpha distance), showing the direct communication in the short distance. Time-averaged mutual information drops abruptlyin the 8 angstroms. Importantly, the time-averaged mutual information regains to the similar value in 168 angstroms to that in the short inter-residue C-alpha distance. This pattern suggested a long-distance communication pathway existing from the ligand SCF to the receptor KIT.

We employed the “pipeline analysis” to reveal the communication pathway. The longest communication pathway starts from ligand SCF and travels throughig-like domains D3, D4 and D5 of one receptor KIT protomer, and then to domain D5 of another KIT protomer (Fig. 3). The residues linking the communication pathway are Val15(SCF)-ASP14(SCF)-LYS78(SCF)-HIS263(D3)-ARG271(D3) TYR259(D3)-LEU275(D3)-THR276(D3)-PHE229(D3)-ARG224(D3)-LYS310(D3)-VAL399(D4)-VAL394(D4)-GLN347(D4)-PHE355(D4)-MET351(D4)-GLY388(D4)-GLY470(D5)-VAL473(D5)-THR461(D5)-ASP458(D5)-LEU455(D5)-SER453(D5)-CYS450(D5)-PHE504(D5)-ASP419(D5). Furthermore, CYS450 (D5) of one receptor KIT protomer is also linked to ASN505(D5) and PHE483(D5) of another receptor KIT protomer.

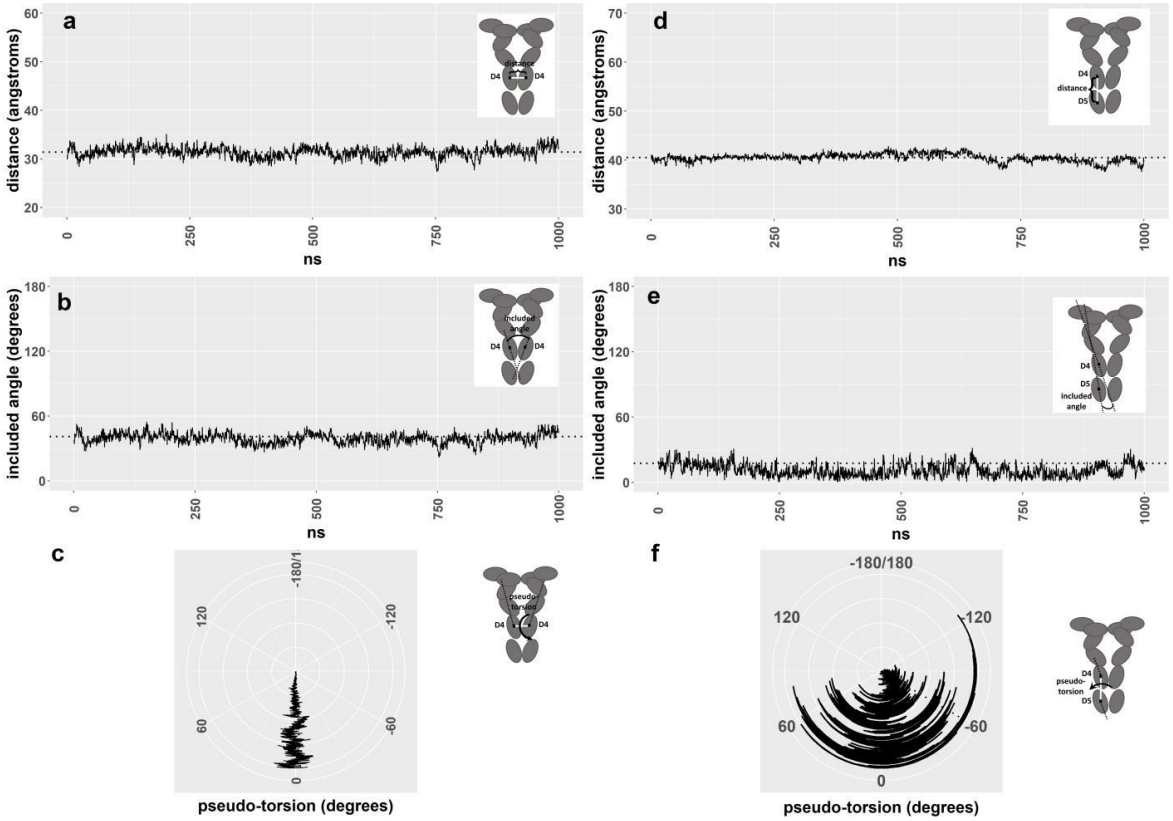


Figure 1. *The geometry of KIT extracellular domain in complex with SCF during molecular dynamics simulation. (a, b, c) The distance, included angle, and pseudo-torsion angles between two extracellular ig-like domains D4 from two protomers of receptor KIT.(d, e, f) The distance, included angle, and pseudo-torsion angles of extracellular ig-like domains D4 vs. D5 of the same protomer of receptor KIT. Schematic representation of the corresponding geometric parameters is also shown.*

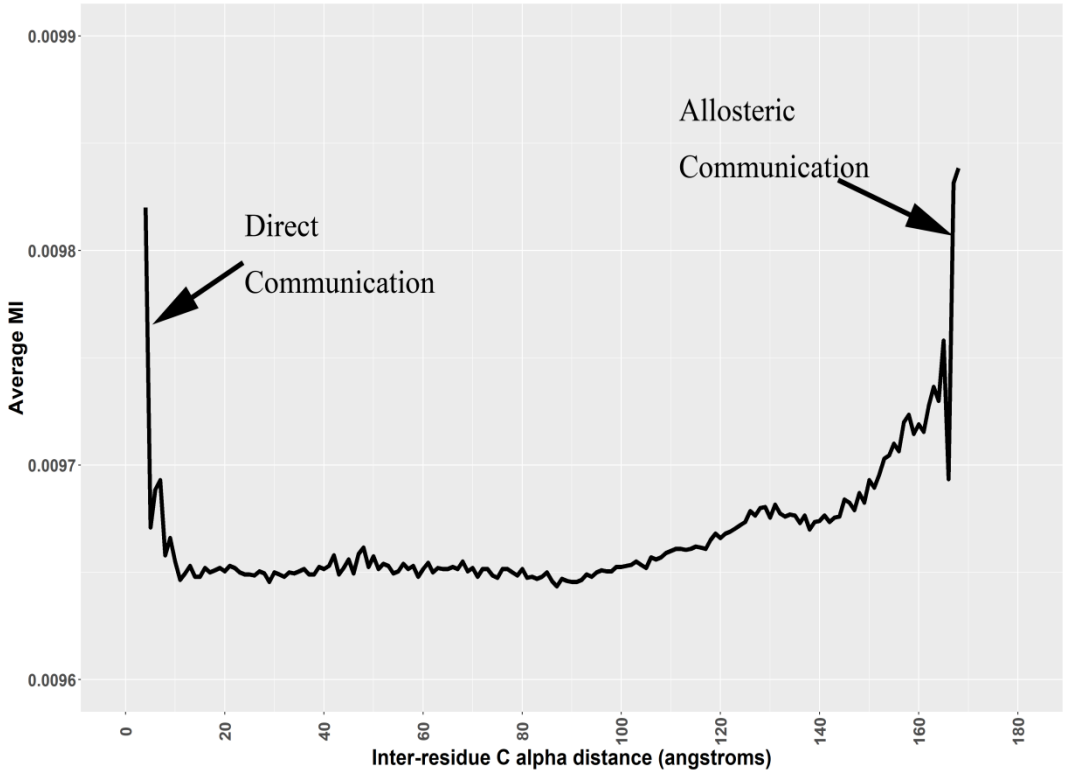


Figure 2. *Mutual information vs. C-alpha distance of all amino acid pair of protein. Higher mutual information between two amino acids as far as 170 angstroms away is apparent.*

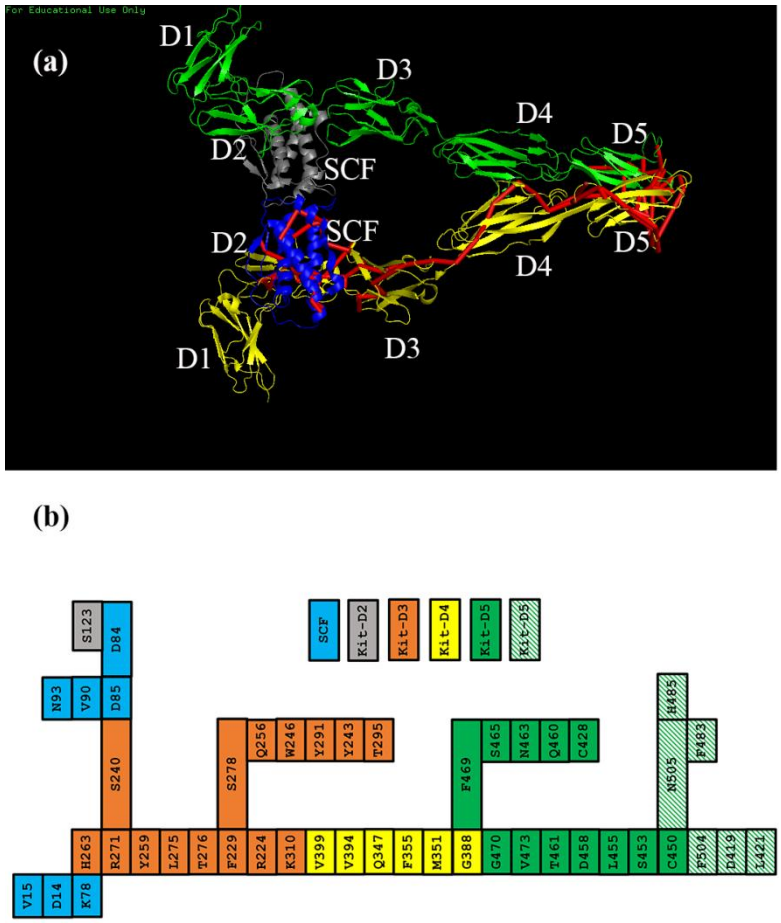


Figure 3. *Communication pathway of KIT extracellular domain in complex with SCF derived from the molecular dynamics simulation. (a) Dimeric SCF (ligand) is shown in ribbon and colored in blue and grey for two protomers. Dimeric extracellular domain of KIT (receptor) is shown in ribbon and colored in yellow and green for two protomers. Communication pathways are shown in pipe shape rendered in red. It is apparent that the communication pathway starts from ligand SCF and travels throughig-like domains D3, D4 and D5 of one KIT protomer, and then to domain D5 of another protomer of KIT. (b) Schematic representation of the residues linking the communication pathway. Residues belonging to various ig-like domains are colored accordingly*

ADDITIONAL FILE 1:

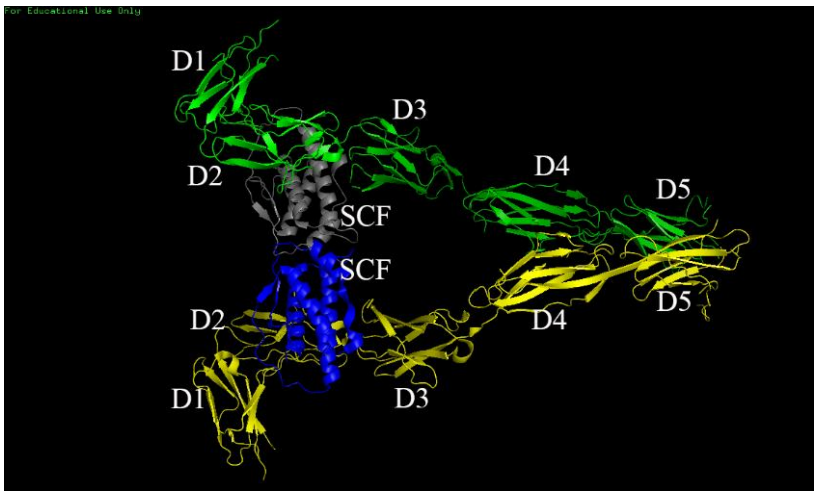


Figure S1: *Crystal structure of KIT extracellular domain in complex with SCF. Two protomers of extracellular domain of receptor KIT are shown in ribbons and colored in yellow and green; two protomers of ligand SCF are shown in ribbons and colored in blue and grey. Five ig-like domains of receptor KIT (D1, D2, D3, D4 and D5) of two protomers are also labeled.*

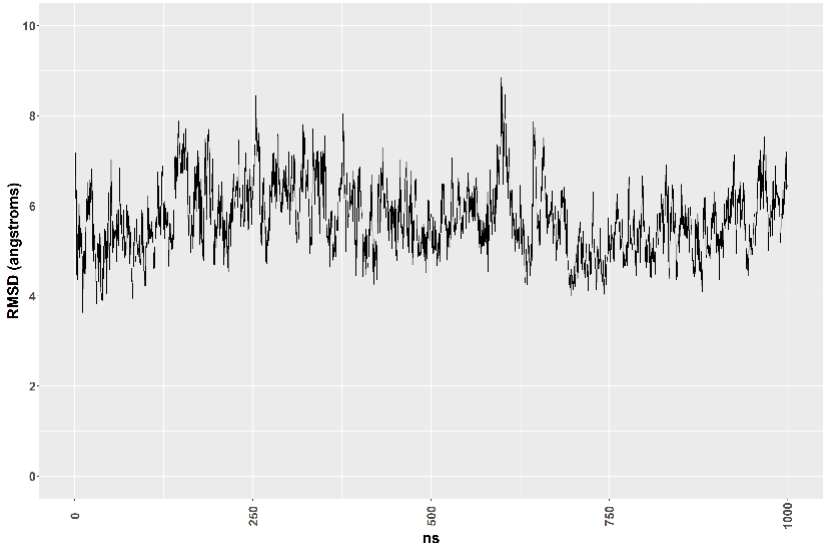


Figure S2: *Root-mean-square deviation of C alpha atoms of KIT extracellular domain in complex with SCF in the molecular dynamics simulation vs. the crystal structure.*

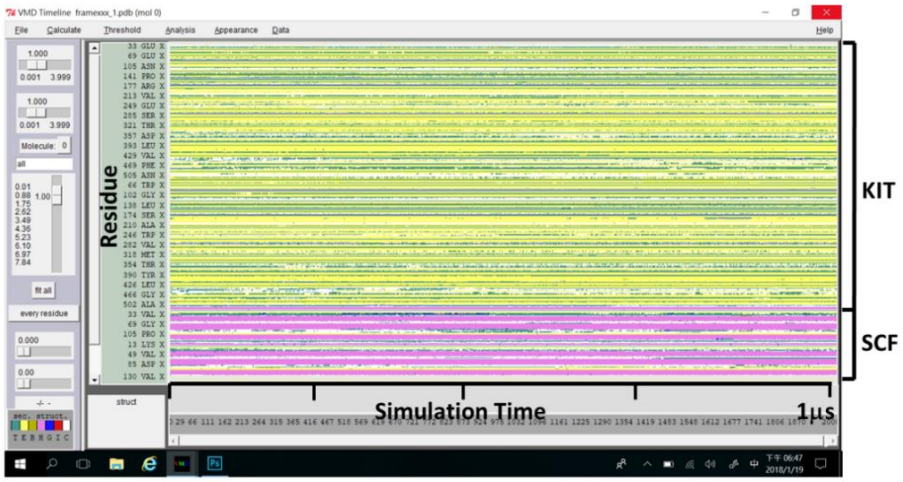


Figure S3: *Secondary structure of KIT extracellular domain in complex with SCF during one microsecond simulation of molecular dynamics.*

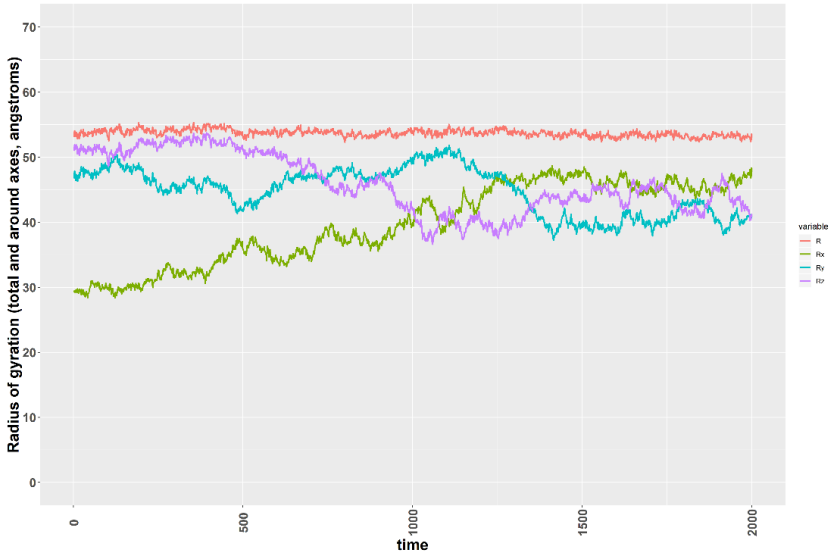


Figure S4: *Radius of gyration of KIT extracellular domain in complex with SCF in the molecular dynamics simulation. R: total. Rx, Ry, Rz: around the axes*

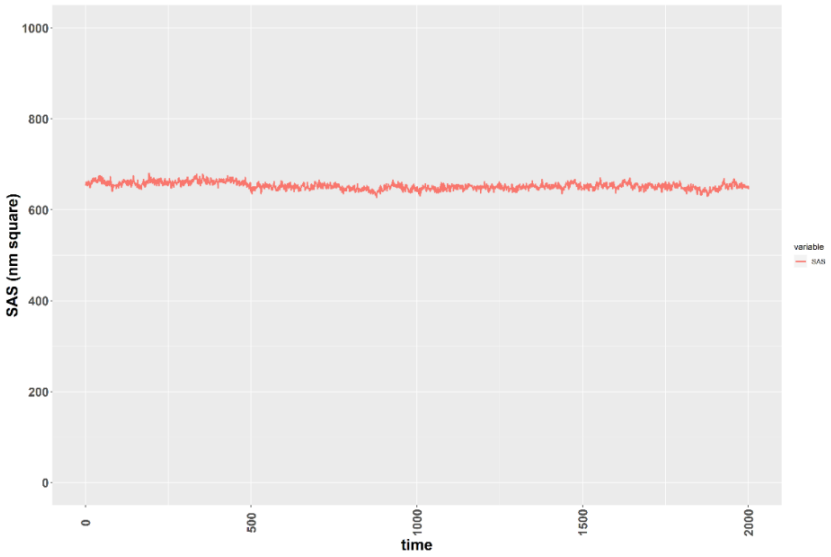


Figure S5: *Solvent accessible surface area of KIT extracellular domain in complex with SCF in the molecular dynamics simulation.*

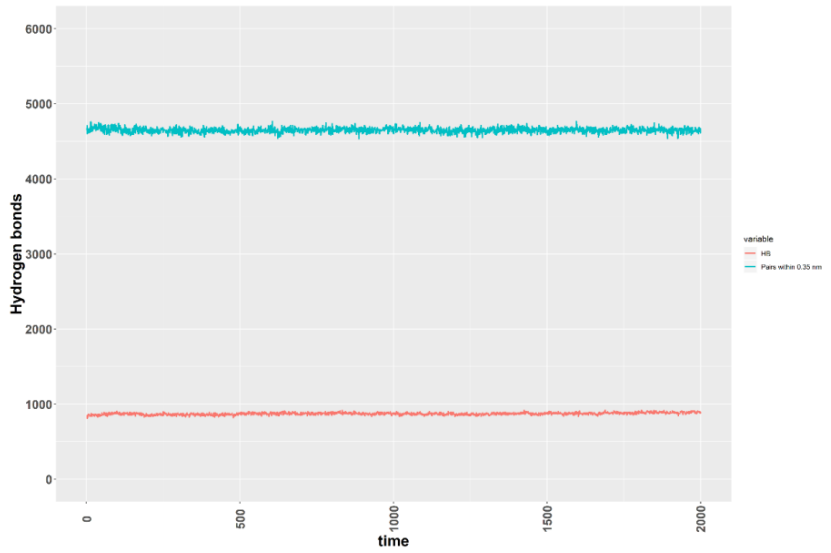
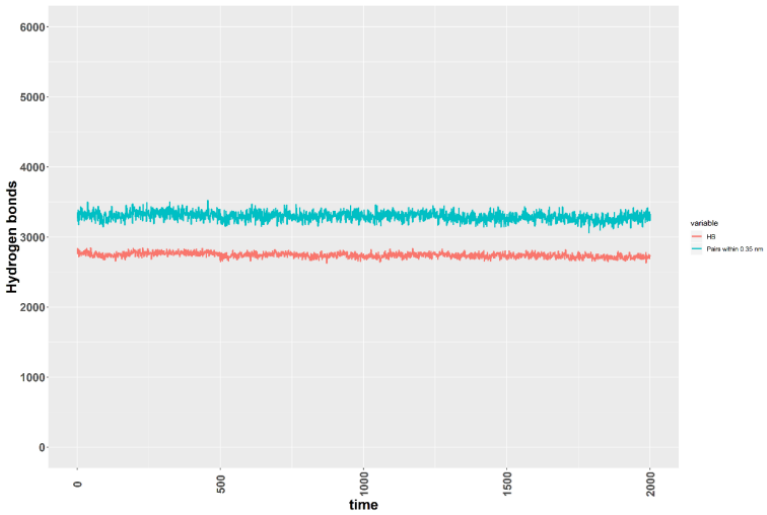


Figure S6: *The number of the intramolecular hydrogen bonds of KIT extracellular domain in complex with SCF in the molecular dynamics simulation.*



FigureS7: *The number of the inter-molecular hydrogen bonds between KIT extracellular domain in complex with SCF and water molecules in the molecular dynamics simulation.*

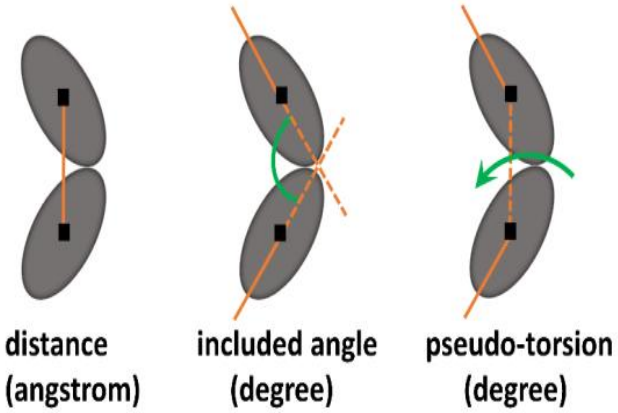


Figure S8: *Schematic representation of the geometry (the distance, included angle and pseudo-torsion angle) of KIT extracellular domain in complex with SCF. The center of mass and principal axes of the extracellular ig-like domain of KIT protomer under consideration were shown.*

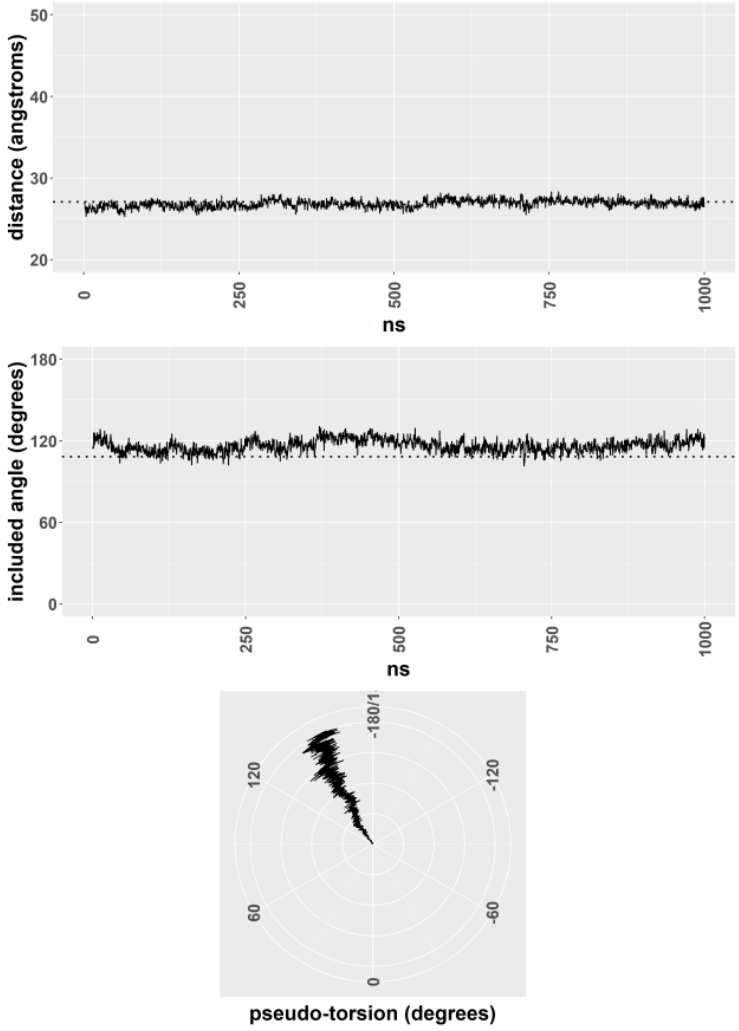


Figure S9: *The distance, included angle, and pseudo-torsion angles of the extracellular ig-like domains D1 vs. D2 of the same protomer of receptor KIT.*

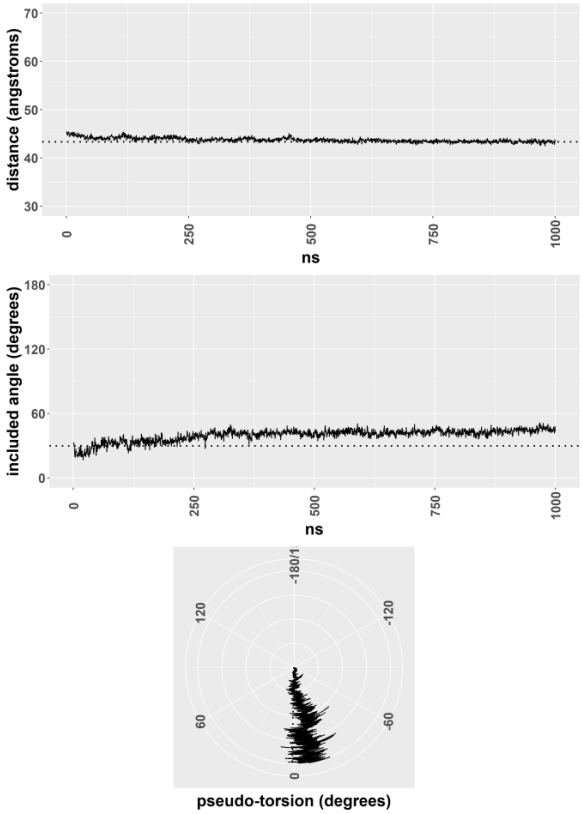


Figure S10: *The distance, included angle, and pseudo-torsion angles of the extracellular ig-like domains D2 vs. D3 of the same protomer of receptor KIT.*

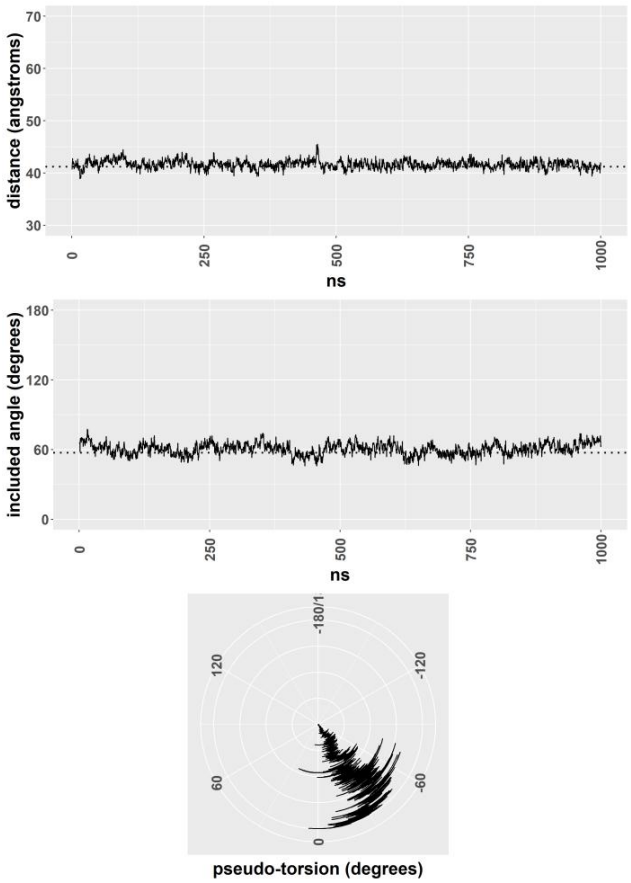


Figure S11: *The distance, included angle, and pseudo-torsion angles of the extracellular ig-like domains D3 vs. D4 of the same protomer of receptor KIT.*

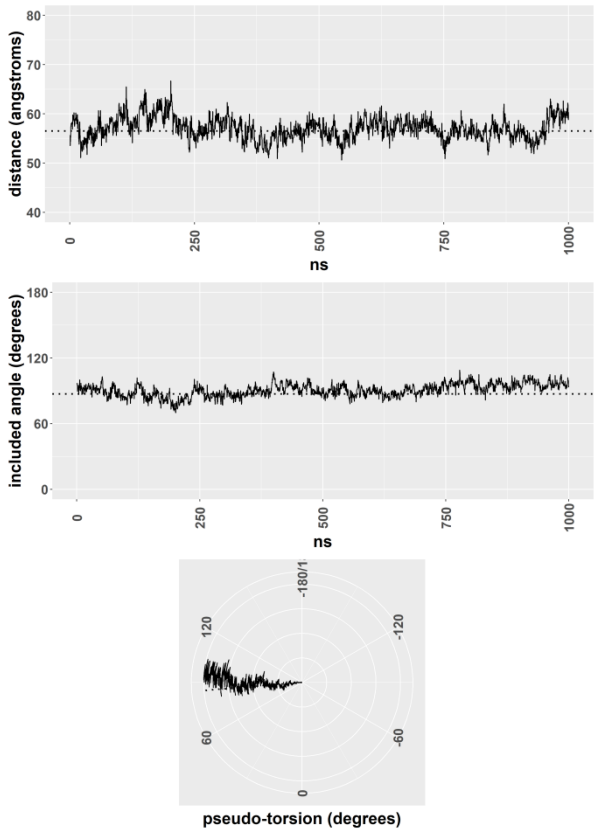


Figure S12: *The distance, included angle, and pseudo-torsion angles between two extracellular ig-like domains D3 from two protomers of receptor KIT.*

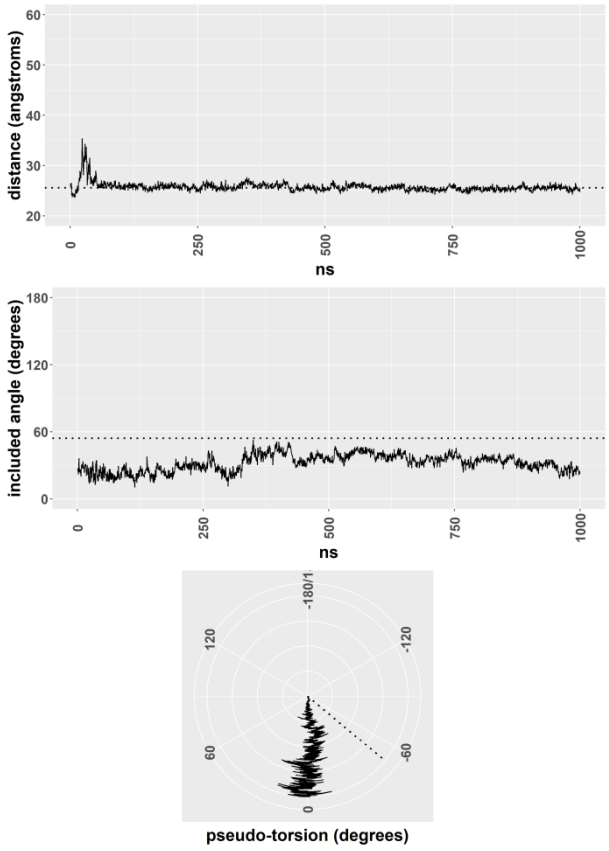


Figure S13: *The distance, included angle, and pseudo-torsion angles between two extracellular ig-like domains D5 from two protomers of receptor KIT.*

DISCUSSION

Recently there are many researches aiming to discover the long-range communication pathway in the interior of the protein. For example, Rivalta *et al* [10] employed molecular dynamics (MD) to investigate the mechanism underlying the activation of the glutaminase catalysis by the ligand PRFAR (N'-[(5'-phosphoribulosyl) formimino]-5-amino imidazole-4-carboxamide ribonucleotide) as far away as 25 angstroms from the catalysis site. They utilized the correlation of protein motions and networks analysis to derive the allosteric pathways (communication pathway) in the PRFAR bound enzyme. The pathways involve the conserved residues that correlate motion of the PRFAR binding loop to motion at the protein-protein interface, and eventually at the glutaminase active site [10]. Bhattacharya and Vaidehi investigated the communication pathway of ligand binding in the extracellular domain to the intracellular domain of G protein-coupled receptor (GPCR) [8]. They analyzed the trajectory of the molecular dynamics and infer the communication pathway from the extracellular domain to the intracellular domain of GPCR. In our research of SCF-KIT complex, we found a communication pathway as long as 170 angstroms, spanning from the ligand SCF to the membrane-proximal ig-like domain of KIT. Although the communication pathway does not prescribe the direction of the correlated motion of protein backbone, it vividly demonstrates the communication of motion facilitated by the conformational flexibility of KIT extracellular domain in complex with SCF. The communication pathway can be regarded as providing a necessary condition for tipping the equilibrium toward the formation of SCF-KIT dimer, as suggested in "Conformational Equilibria Model", in which it proposed that different conformation of protein exists in equilibrium.[11-13]. We even speculate that the communication pathway might be transmitted further to the transmembrane domain and the catalytic domain of receptor KIT. We are now investigating this hypothesis.

To our knowledge, this is the first time that the concerted motion of extra cellular ig-like domains of KIT receptor is discovered through molecular dynamics simulation. We found two neighbored domains (D4 and D5) undergoing the pseudo-torsional movement. Since KIT belongs to type III receptor tyrosine kinase (RTKIII) family, it is interesting to investigate if other family members (hCSF1R, Flt3, and PDGFR a/b) share similar features to those of SCF-KIT complex.

ACKNOWLEDGEMENTS

This work was financially supported by the Center of Excellence for the Oceans, National Taiwan Ocean University and Minister of Science and Technology, Taiwan.

REFERENCE

[1] Yuzawa S, Opatowsky Y, Zhang Z, Mandiyan V, Lax I, Schlessinger J: Structural basis for activation of the receptor tyrosine kinase KIT by stem cell factor. Cell 2007, 130(2):323-334.

[2] Ashman LK: The biology of stem cell factor and its receptor C-kit. Int J Biochem Cell Biol 1999, 31(10):1037-1051.

[3] Karami Y, Laine E, Carbone A: Dissecting protein architecture with communication blocks and communicating segment pairs. BMC Bioinformatics 2016, 17 Suppl 2:13.

[4] Allain A, Chauvot de Beauchene I, Langenfeld F, Guarracino Y, Laine E, Tchertanov L: Allosteric pathway identification through network analysis: from molecular dynamics simulations to interactive 2D and 3D graphs. Faraday Discuss 2014, 169:303-321.

[5] Jo S, Kim T, Iyer VG, Im W: CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem 2008, 29(11):1859-1865.

[6] Case DA, Cheatham TE, 3rd, Darden T, Gohlke H, Luo R, Merz KM, Jr., Onufriev A, Simmerling C, Wang B, Woods RJ: The Amber biomolecular simulation programs. J Comput Chem 2005, 26(16):1668-1688.

[7] Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D et al: GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 2013, 29(7):845-854.

[8] Bhattacharya S, Vaidehi N: Differences in allosteric communication pipelines in the inactive and active states of a GPCR. Biophys J 2014, 107(2):422-434.

[9] Killian BJ, Kravitz JY, Somani S, Dasgupta P, Pang YP, Gilson MK: Configurational entropy in protein-peptide binding: computational study of Tsg101 ubiquitin E2 variant domain with an HIV-derived PTAP nonapeptide. J Mol Biol 2009, 389(2):315-335.

[10] Rivalta I, Sultan MM, Lee NS, Manley GA, Loria JP, Batista VS: Allosteric pathways in imidazole glycerol phosphate synthase.

Proceedings of the National Academy of Sciences of the United States of America 2012, 109(22):E1428-1436.

[11] Monod J, Wyman J, Changeux JP: On the Nature of Allosteric Transitions: A Plausible Model. J Mol Biol 1965, 12:88-118.

[12] Motlagh HN, Wrabl JO, Li J, Hilser VJ: The ensemble nature of allostery. Nature 2014, 508(7496):331-339.

[13] Tsai CJ, Nussinov R: A unified view of "how allostery works". PLoS Comput Biol 2014, 10(2):e1003394.

Citation: *Wen-Shyong Tzou, Chung-Hao Li, Pang-Hung Hsu, Tun-Wen Pai, Chin-Hwa Hu, "Microsecond-Timescale Conformational Dynamics Reveals the Flexibility and Communication Pathway of the KIT Extracellular Domain in Complex with Stem Cell Factor", Journal of Biotechnology and Bioengineering, 5(1), 2021, pp 1-12. DOI: https://doi.org/10.22259/2637-5362.0501001*

Copyright: *© 2021 Wen-Shyong Tzou., et al, This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*