

SCIENTIFIC REPORTS



OPEN

Network approach of the conformational change of c-Src, a tyrosine kinase, by molecular dynamics simulation

Hyun Jung Yoon¹, Sungmin Lee^{2,3}, Sun Joo Park⁴ & Sangwook Wu¹

Non-receptor tyrosine kinase c-Src plays a critical role in numerous cellular signalling pathways. Activation of c-Src from its inactive to the active state involves large-scale conformational changes, and is controlled by the phosphorylation state of two major phosphorylation sites, Tyr416 and Tyr527. A detailed mechanism for the entire conformational transition of c-Src via phosphorylation control of Tyr416 and Tyr527 is still elusive. In this study, we investigated the inactive-to-active conformational change of c-Src by targeted molecular dynamics simulation. Based on the simulation, we proposed a dynamical scenario for the activation process of c-Src. A detailed study of the conformational transition pathway based on network analysis suggests that Lys321 plays a key role in the c-Src activation process.

The Src family of non-receptor tyrosine kinases includes c-Src, Fyn, Yes, Blk, Yrk, Fgr, Hck, Lck and Lyn¹. These tyrosine kinases act as key mediators of various cellular signal transduction processes such as migration, angiogenesis, proliferation, differentiation, survival and immune function^{2–4}. The proto-oncogene c-Src, encoded by the Src gene, was first isolated as the normal cellular homologue to the potent avian sarcoma viral transforming oncogene v-Src⁵. Crystal structures of c-Src have shown that its catalytic activity is tightly regulated by autoinhibition via multiple intramolecular interactions^{6,7}. All Src family kinases share a similar domain arrangement and consist of a unique N-terminal region, the Src homology 3 (SH3), SH2, linker and tyrosine kinase domains and the regulatory C-terminal tail. c-Src is predominantly inactive under normal cellular circumstances. The SH3 domain binds to the proline-rich region in the linker domain, and the SH2 domain binds to pTyr527, collectively resulting in a closed and inactive protein conformation^{8,9}. Many studies have reported the catalytic activation of c-Src in several human cancers^{10,11}. These studies show that inactive c-Src changes to an open, active conformation in response to physiological upstream signals.

An important mechanism of c-Src activation involves the control of its phosphorylation status. A major phosphorylation site is Tyr527, which is present in the C-terminal tail of c-Src, but not in v-Src. Dephosphorylation of pTyr527 releases the closed c-Src conformation, resulting in the activation of the SH3, SH2 and tyrosine kinase domains together with an autophosphorylation of Tyr416. This autophosphorylation at Tyr416, which is located in the activation loop of the tyrosine kinase domain, is another essential event in the process. To allow full kinase activity, c-Src must first be dephosphorylated at Tyr527 before it undergoes autophosphorylation at Tyr416^{12,13}. Thus, c-Src kinase activity is extremely dependent on the phosphorylation status of the negative-regulatory Tyr527 and the positive-regulatory Tyr416.

Network theory offers a new approach to biological problems^{14,15}. For example, Goh *et al.* successfully constructed a network of Mendelian gene–disease associations to identify unknown important genes causing disease¹⁶. The first step of this approach is to construct a biological network, in which each node is an entity such as a gene and each edge is an interaction or correlation between entities. The constructed network enables a view of the complete set of relationships. In particular, certain measurable properties of these networks facilitate the search for important biological entities. Centrality measures, developed to determine the ranking (i.e., importance) of nodes, have proven effective in uncovering important target genes relevant to disease. In this study, we

¹Department of Physics, Pukyong National University, Busan, 48513, Republic of Korea. ²Department of Energy Science, Sungkyunkwan University, Suwon, 16419, Republic of Korea. ³Department of Physics and Institute of Basic Science, Korea University, Seoul, 02841, Republic of Korea. ⁴Department of Chemistry, Pukyong National University, Busan, 48513, Republic of Korea. Correspondence and requests for materials should be addressed to S.J.P. (email: parksj@pknu.ac.kr) or S.W. (email: sangwoow@pknu.ac.kr)

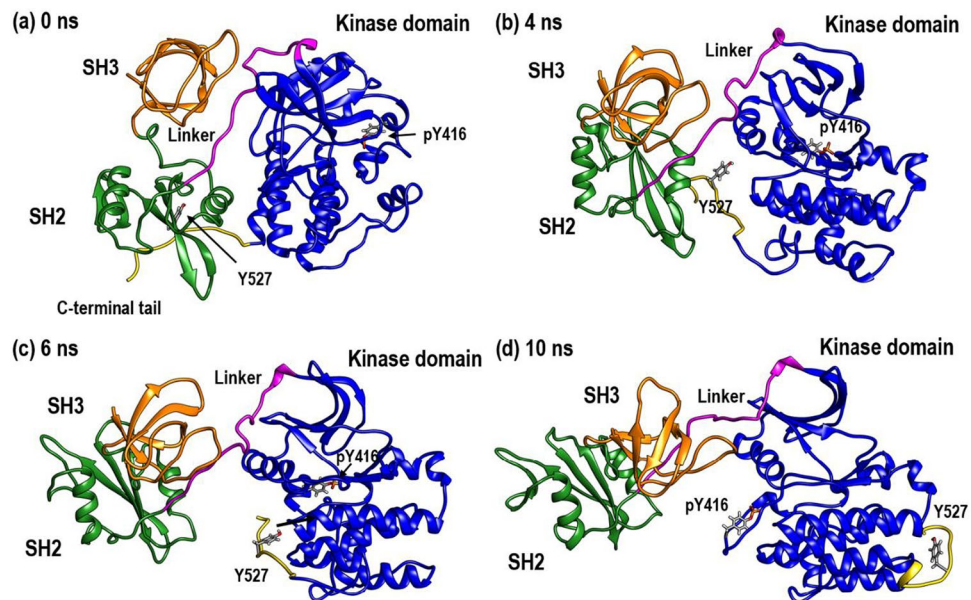


Figure 1. Four snapshots of TMD trajectory of the tyrosine kinase c-Src between 0 and 10 ns. (a) 0 ns: Starting structure (inactive conformation) (b) 4 ns: Detachment of Tyr527 in the C-terminal tail from the SH2 domain, triggering the transition from inactive to active conformation (c) 6 ns: Large-scale conformational change of c-Src (d) 10 ns: Target structure (active conformation).

utilised three centrality measures – the betweenness, closeness and degree centralities – to reveal the important residues controlling the conformational transition of the c-Src protein between the inactive and active conformations, using molecular dynamics simulation.

Results and Discussions

Targeted molecular dynamics simulation. Tyr416 is located near the activation loop (Asp404–Glu432) and Tyr527 is located in the tail region of the C-terminus. In the inactive form of c-Src (PDB id: 2SRC)¹⁷, Tyr527 is phosphorylated (pTyr527) but Tyr416 is not. Conversely, in the active c-Src conformation (PDB id: 1Y57)¹⁸, Tyr527 is not phosphorylated, whereas Tyr416 is (Supporting Information Fig. S1).

We employed targeted molecular dynamics (TMD) simulation¹⁹ to study the conformational change of c-Src tyrosine kinase. This method allows the modelling of a continuous transition pathway between two known conformations. TMD simulation can be used to model two directions for the conformational change between the inactive and active states: namely, from inactive to active, or vice versa. In this study, we focused on the activation process, i.e., from the inactive state to the active state.

The phosphorylation at Tyr527 results in multiple intramolecular interactions between pTyr527 and the SH2 domain and between the linker and the SH3 domain. In particular, the pTyr527–SH2 domain interaction plays the role of a lock switch such that the assembly of the domains remains tightly packed together. The main force keeping the assembly of domains packed into a narrow space seems to be electrostatic interactions involving the negative charge of pTyr527. Two positive residues, Arg175 and Lys203, appear to be involved in the electrostatic interaction in the SH2 domain. The positions of Tyr527, Arg175 and Lys203 are shown in Supporting Information Fig. S2.

The side chain of Tyr416 in the activation loop is buried between the two lobes of the kinase domain, effectively protecting Tyr416 from being phosphorylated. The hydrophobic interaction between the inward surface of the α -helix and the N-terminal part of the activation loop also stabilises the inactive conformation.

In the transition from the inactive to the active form, the electrostatic interaction between Tyr527 and Arg175/Lys203 becomes weaker. The C-terminal tail, including Tyr527, is completely detached from the SH2 domain during the period of 0–3 ns of the TMD simulation (see Supporting Information Fig. S3). The process of detachment stimulates the unpacking of the domain assembly. The activation processes from the inactive state (0 ns: Fig. 1a) to the active state (10 ns: Fig. 1d) in the TMD simulation are shown in Fig. 1. Between 4 and 6 ns of the simulation, the interaction between Tyr527 (in the C-terminal tail) and its neighbouring positive residues becomes weaker. The detachment of Tyr527 from the SH2 domain triggers the targeted conformational change (Fig. 1b). The detached Tyr527 moves toward the kinase domain (Fig. 1c).

At this stage, the most prominent conformational change occurs at the kinase domain. The secondary structures, including the α C-helix in the kinase domain, rotate significantly. Between 6 and 10 ns, Tyr527 moves away from the SH2 domain and toward the kinase domain. During this time, Tyr416 remains buried beneath the activation loop. By the final stage, Tyr527 has reached the far side of the kinase domain relative to the SH2 domain. The activation loop has also moved from its original position. Tyr416 is now exposed to the surface. The trajectory for conformational transition from the inactive to the active state is shown in Supporting Information Movie 1.

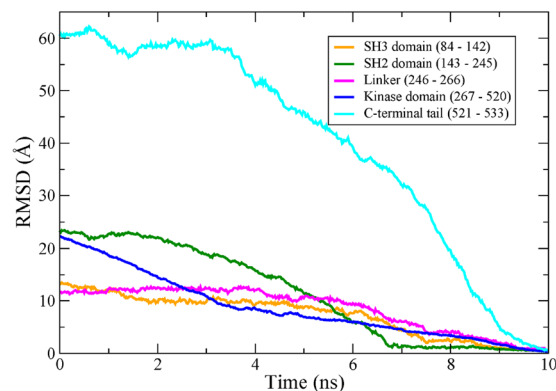


Figure 2. Backbone RMSD values for each domain of the tyrosine kinase c-Src. The starting structure is the inactive conformation and the target structure is the active conformation. The RMSD values of each domain were obtained by aligning the backbone during the TMD trajectory with that of the active conformation. SH3 (orange), SH2 (green), linker (pink), kinase (blue), and C-terminal tail (cyan).

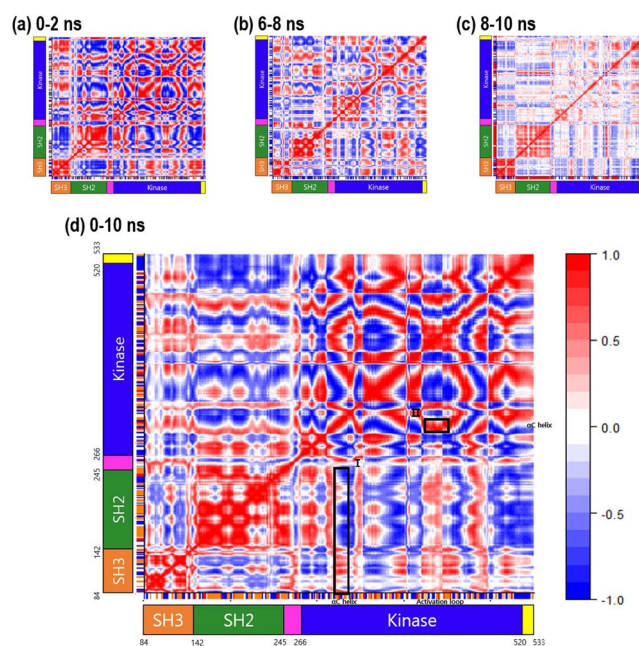


Figure 3. Dynamical cross-correlation maps (DCCMs). Correlation values are in the range of -1 to $+1$, where positive values mean that two residues are correlated and negative values mean they are anti-correlated.

The RMSD of each domain during the 10 ns TMD trajectory is shown in Fig. 2. The domain RMSD values express the deviation of the backbone alignment during the TMD trajectory with respect to the backbone of the active conformation (PDB id: 1Y57). The domain with the largest initial RMSD is the C-terminal tail region (cyan). The RMSD of the C-terminal tail begins at 60 Å and gradually decreases throughout the trajectory.

Dynamical cross-correlation of c-Src tyrosine kinase. We calculated a dynamical cross-correlation map (DCCM)^{20–22} to analyse the correlations among the positions of residues during the transition from the inactive to the active conformation. In this map, C_{ij} ranges between -1 and $+1$. Positive values mean that two C α atoms are correlated: they tend to move with same period and phase. Conversely, negative values mean that two C α atoms are anti-correlated: the two C α atoms move in the same period but opposite phase. The DCCMs corresponding to four intervals of the TMD simulation are shown in Fig. 3: 0–2 ns (Fig. 3a), 6–8 ns (Fig. 3b), 8–10 ns (Fig. 3c) and 0–10 ns (Fig. 3d). In the DCCMs, the five domains are shown in different colours: SH3 domain (orange), SH2 domain (green), linker (magenta), kinase domain (blue) and C-terminal tail (yellow). During 0–2 ns (Fig. 3a), the intra-domain SH3–SH3 and SH2–SH2 correlations are positive. Conversely, the inter-domain SH3–SH2 correlation is negative. Therefore, during the early stage of the transition, the SH3 and SH2 domains move in an anti-correlated manner. During this stage, the linker region (pink) shows correlated movement with the SH3 domain and anti-correlated movement with the SH2 domain. The C-terminal tail, however, shows

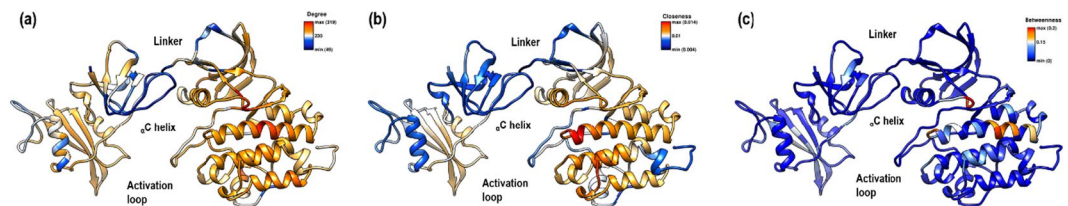


Figure 4. Centrality measures during the 10 ns TMD simulation. (a) Degree centrality (b) Closeness centrality (c) Betweenness centrality. We mapped the centrality measures onto the active conformation.

	Betweenness	Closeness	Degree
1	Glu320	Val377	Val323
2	Lys321	Arg379	Val402
3	Ala368	Glu378	Lys321
4	Leu322	Met380	Ile370
5	Met380	Thr508	Thr301

Table 1. Residues with top 5 values of each centrality measure.

the opposite pattern: anti-correlated movement with SH3 and correlated movement with SH2. During 6–8 ns (Fig. 3b), we observe two different correlation patterns for the linker region. The Pro246–Asp258 region of the linker shows correlation with the SH2 domain, whereas the Asp258–Ser266 region shows anti-correlation with the SH2 domain. The activation loop in the kinase domain is anti-correlated with the SH3 domain and correlated with the SH2 domain. Overall, the C-terminal tail shows anti-correlation with the SH2 domain. However, the Lys203–Tyr230 region shows correlation with the SH2 domain. During 8–10 ns, many of the correlation values are close to zero (Fig. 3c). This indicates that the strongest correlation among the residues occurs in the early stage of the transition from the inactive to the active conformation. The DCCM for the entire TMD simulation (0–10 ns) is shown in Fig. 3d. Overall, the residues of the linker region have only a weak correlation with the kinase domain. The α C helix has correlation and anti-correlation with the SH3 and SH2 domains, respectively (denoted as box I in Fig. 3d). In addition, the α C helix has a strong correlation with the activation loop (denoted as box II in Fig. 3d).

Network analysis. The degree, closeness and betweenness centralities were measured for the conformational change that occurred during the 10 ns TMD simulation (Fig. 4). The residues with high values of these centralities are not those in the activation loop, linker, or α C-helix. Instead, the residues with high centrality measures are mainly confined to the helix region adjacent to the α C-helix. The residues with the top 5 values of each centrality measure are listed in Table 1.

Histograms of the degree, closeness and betweenness centralities are shown in Fig. 5. In particular, the distribution of the betweenness centrality shows a power-law decrease. By fitting this distribution to the power law^{23,24} $p(x) = Cx^{-\alpha}$, the exponent α was calculated as 1.85 ± 0.17 . The exponent α is thus less than 2, indicating that the mean of the betweenness centrality distribution is infinite in the limit of large system size. Values of $\alpha < 2$ are typically found in complex systems, such as the intensity of solar flares ($\alpha = 1.83$)²⁵, intensity of wars ($\alpha = 1.80$)²⁵ and frequency of family names ($\alpha = 1.94$)²⁵. We obtained the same value for the exponent α when we calculated the DCCM with a different threshold, $C_{thr} = 0.5$ (see Supporting Information Fig. S4).

Mutation. We conjectured that the residues with high centrality measures would play a critical role in the conformational change of c-Src. To verify this, we selected a residue with high centrality measures, then mutated it, and monitored the conformational transition during a TMD simulation of the mutant c-Src. As the mutation target we selected Lys321, since this residue ranked highly in both the betweenness and degree centralities. Lys321 was mutated into Ala. Subsequently, a 10 ns TMD simulation was performed for the mutant Lys321Ala under the same conditions as previously (i.e., the same spring constant). The RMSD values of the wild type (WT) and mutant for the SH3, SH2, linker, kinase domain (267–520) and C-terminal tail regions are shown, respectively, in Fig. 6a–e. The overall RMSD values for the mutant are shown in Fig. 6f. The conformational change from the inactive to the active conformation failed to reach completion. In particular, the mutation of Lys321 to Ala321 caused notable changes in the linker, kinase domain and C-terminal tail regions during the conformational transition of c-Src. The RMSD values in the linker region are shown in blue for the WT and red for the mutant Lys321Ala (Fig. 6c). For the first 8 ns of the 10 ns TMD simulation, the RMSD values for the WT and mutant show similar trends. However, from 8 ns, the two values diverge. For the mutant, the RMSD value of the linker region sharply decreases at this point. However, for the WT, the RMSD value of the linker region shows a slight recovery at 8 ns before decreasing more gradually. The RMSD values in the kinase domain show a more pronounced difference between the WT and mutant (Fig. 6d). The two RMSD values in the kinase domain diverge at 2 ns, after which the RMSD for the WT decreases more steeply than that for the mutant. The RMSD value of the mutant recovers slightly after 8.5 ns and approaches a final value of approximately 4 Å. This deviation from zero indicates that the kinase domain did not converge to the target structure (the active conformation) under identical TMD

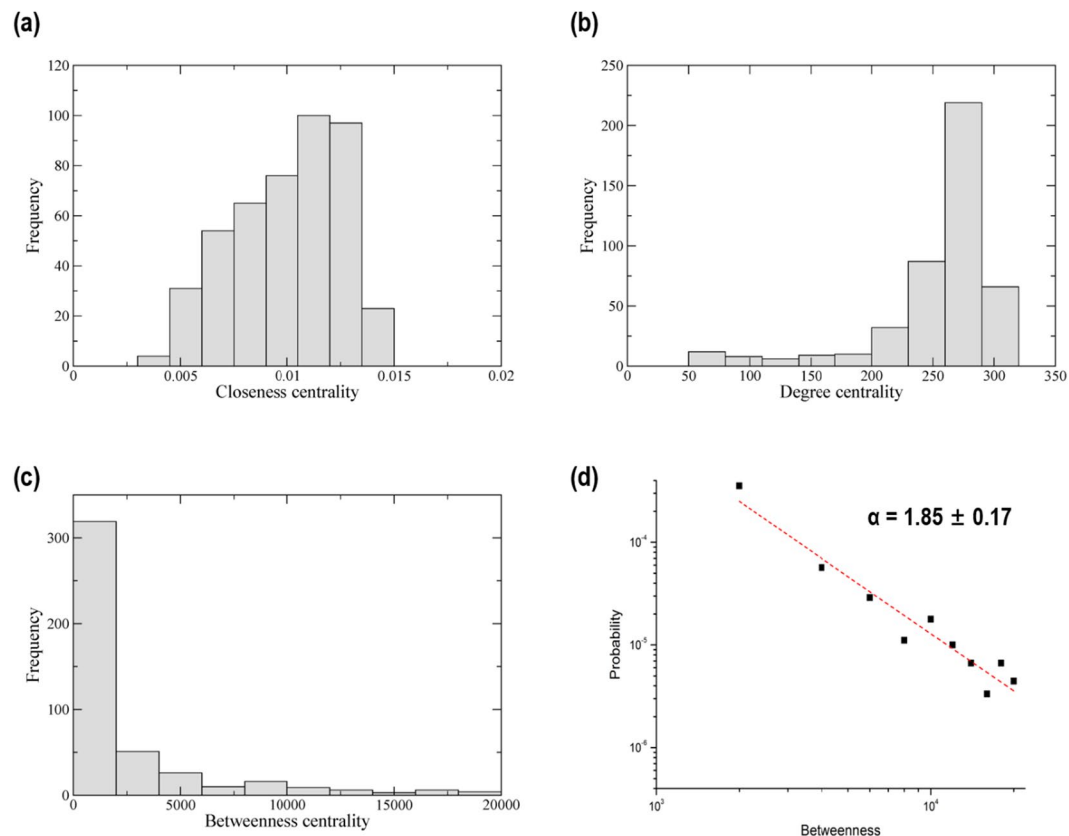


Figure 5. Distributions of the centrality measures during the 10 ns TMD simulation. (a) Histogram of closeness centrality (b) Histogram of degree centrality (c) Histogram of betweenness centrality. (d) Power law for the distribution of betweenness centrality.

simulation conditions as for the WT. An incomplete transition is also observed in the C-terminal tail of the mutant (Fig. 6e). Overall, therefore, the mutation of Lys321 gives rise to an incomplete conformational transition from the inactive to the active conformation.

The betweenness centrality values for the WT and the mutant during the conformational change, calculated with $C_{thr} > 0.3$, are shown in Fig. 7. The larger the size of each circle, and the more deeply red in colour, the higher the value of the betweenness centrality for the corresponding residue. In Fig. 7, only the edges with weights (w_{ij}) in the range of 0 to 0.2 are shown, for clarity of visualisation. The inactive structure of c-Src (PDB id: 2SRC) was employed as the network topology. Figure 7 shows that the residues with the highest betweenness centralities are heavily concentrated in the kinase domain, with a large number of edges for both the WT and mutant. However, despite sharing the common feature that the residues with the highest betweenness centralities are localised around the kinase domain, there is a distinct difference between the network patterns for the WT and the mutant in detail. For the WT, Arg318–Leu322, Tyr376, Glu378, Met380 and Asn381, which are all located close to Glu320, the residue with the highest value of the betweenness centrality, also show relatively high betweenness centralities. For the mutant, however, the residues with relatively high betweenness centralities are more or less dispersed throughout the kinase domain. The replacement of the charged residue Lys321 by the hydrophobic residue Ala causes the redistribution of the network pattern in the mutant, leading to a different pattern in the conformational change. The residues with the highest values of the betweenness centrality for the WT and the mutant are listed in Supporting Information Table S1. In order to demonstrate the robustness of the betweenness centrality ranking, we analyzed the additional paths generated by WISP (Weighted Implementation of Suboptimal Paths)²⁶. The analysis on the additional paths by WISP are shown in Supporting Information (Figs S6, S7, S8, and Table S3).

The differences in the network patterns of the betweenness centrality between the WT and the mutant are visualised in Fig. 8 using the community structure of the network. Only meaningfully large communities are shown (three communities for the WT and two for the mutant). The community structure was generated with the aid of the topology algorithm in the program Cytoscape²⁷. The edge width in the community structure is represented as the inverse of the weight w_{ij} as defined in Equation 3 in the Methods section. Figure 8 shows the relative “distance” between the nodes (i.e., residues) in the network for the WT and the mutant. By definition, residues within a community have strong correlation while residues in different communities have weak correlation. Importantly, the number of communities is increased from 24 to 130 by the mutation. This indicates that the residues are scattered into smaller communities and the strong correlations between residues are reduced (Table S2). Clearly, the networks for the WT and the mutant differ markedly as a result of this point mutation, implying that

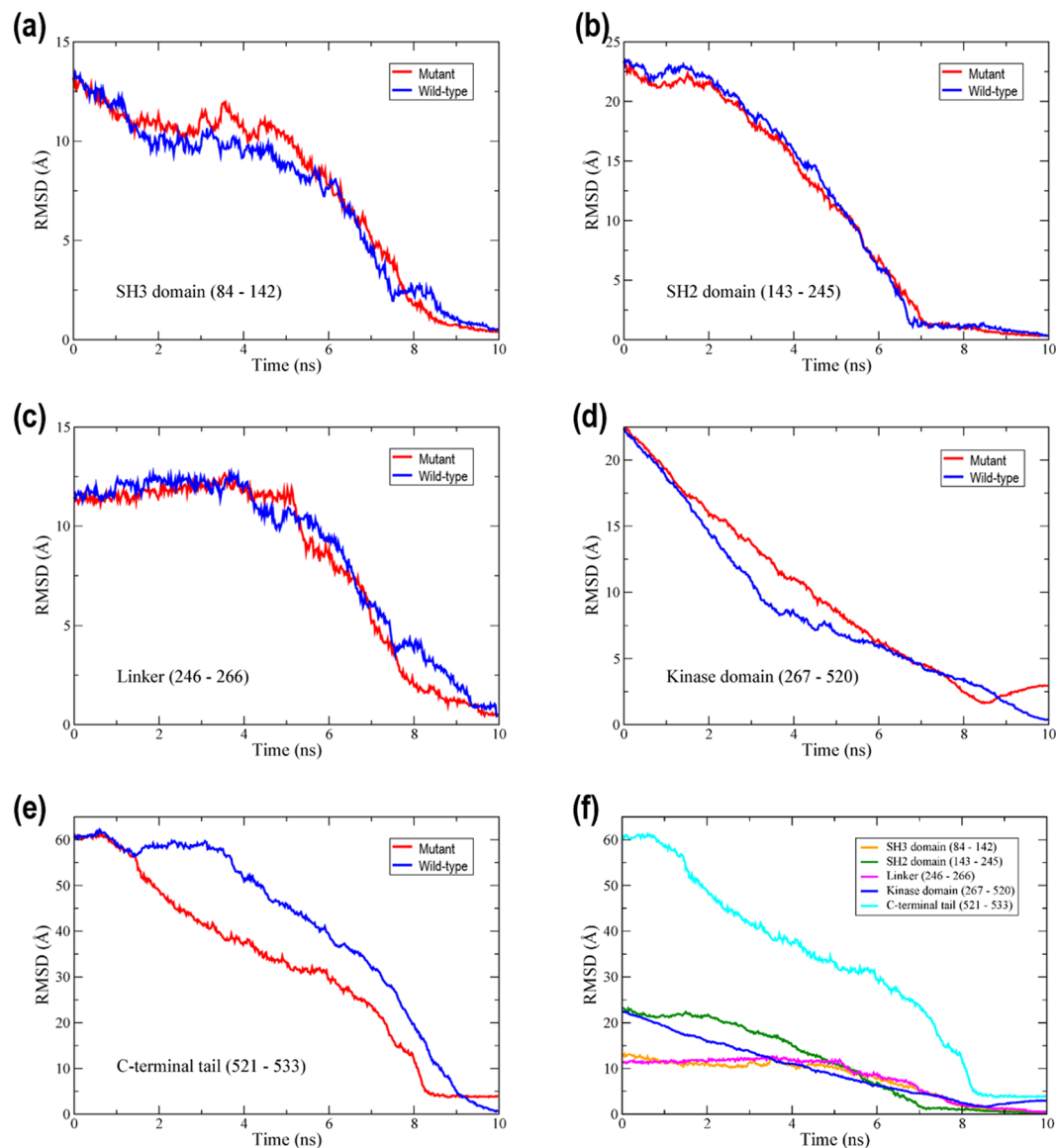


Figure 6. RMSD values for the wild type and the mutant (K321A) during 10 ns TMD simulation for (a) SH3 domain, (b) SH2 domain, (c) Linker, (d) Kinase domain, (e) C-terminal tail, (f) Overall RMSD values for the mutant.

Lys321 is a “hub” in the context of the network. The distribution of the residues among each community is listed in Supporting Information Table S2.

Conclusion

We studied the conformational change of the tyrosine kinase c-Src from the inactive to the active conformation by targeted MD simulation. From the simulation trajectory, a sequential dynamic pathway of the conformational change could be observed. The detachment of Tyr527 in the C-terminal tail from the SH2 domain triggers the conformational change of c-Src. The subsequent domain–domain interactions lead to the exposure of Tyr416, which is initially buried beneath the activation loop in the kinase domain but moves to the surface.

Once it is detached from the SH2 domain, Tyr527 moves from the C-terminal tail toward the kinase domain. Finally, the salt-bridge between Lys295 and Glu310 is formed to complete the conformational change (see Supporting Information Fig. S5).

We quantified the inter/intra-domain correlations of the positions of residues during the conformational change using DCCM. We also constructed a dynamical protein network based on the dynamical interaction among the residues. The degree, closeness and betweenness centralities were measured for the network during the TMD simulation. Notably, we observed that the betweenness distribution follows a power law with the exponent $\alpha < 2$. By analogy with the “80–20 rule” (wherein 80% of the wealth belongs to the richest 20% of people)²⁸, the emergence of an exponent α with a value < 2 in the betweenness distribution implies that a very few residues control the conformational transition of c-Src during the conformational change. Indeed, the conformational

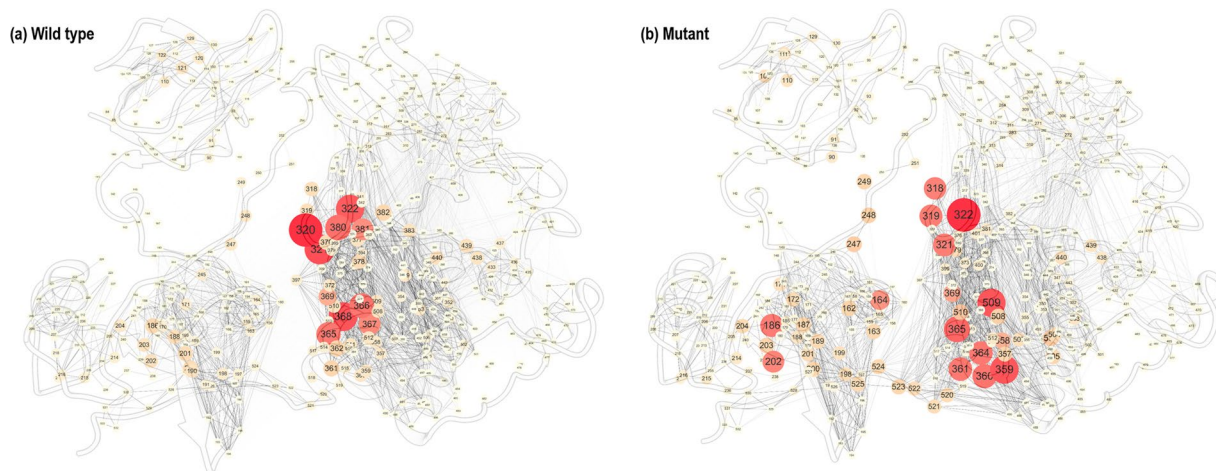
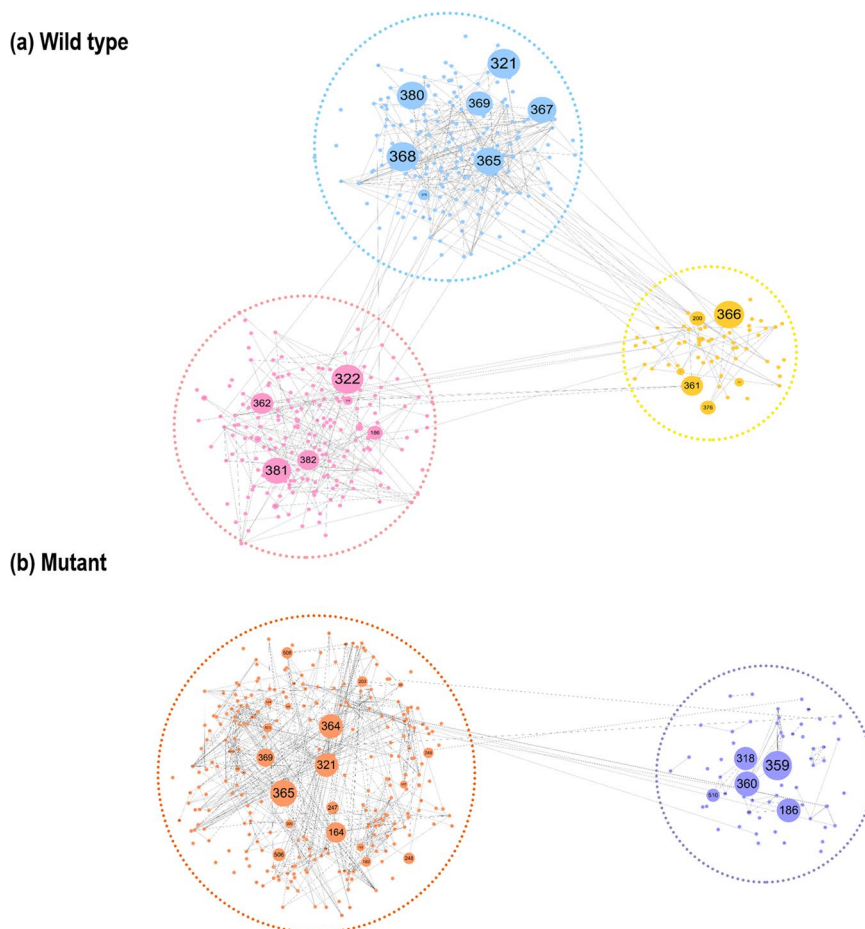


Figure 7. The distribution of the residues with the highest values of the betweenness centrality for (a) the WT, (b) the mutant during the conformational change. The network topology is the inactive conformation of c-Src (PDB id: 2SRC).



Methods

Targeted molecular dynamics (TMD) simulation. Performing a TMD simulation requires an external potential, U_{TMD} , to be defined. U_{TMD} is given as¹⁹

$$U_{TMD} = \frac{k}{2N} [RMSD(t) - RMSD^*(t)]^2 \quad (1)$$

RMSD(t) is the root-mean-square deviation (RMSD) of the simulated structure from the target structure at time t . RMSD*(t) is the RMSD value at time t assuming a linear decrease from the initial to the target structure. The inactive form (PDB ID: 2SRC) of c-Src is defined as the initial conformation and the active form (PDB id: 1Y57) as the target conformation. The two conformations were manipulated to have the same number of atoms for the TMD simulation. The spring constant k was set as 2500 kcal/mol · Å² for 3619 atoms ($=N$), excluding hydrogen atoms. We performed the simulations using the NAMD 2.9 package²⁹ with the CHARMM 27 force field³⁰ and with protein parameters incorporating the CMAP corrections³¹. In TMD, the external force guides the conformational change from the inactive to the active conformation within a reasonable time scale. In our calculation, time was set to 10 ns. We employed as small as possible value of the spring constant (2500 kcal/mol · Å²) to undergo the conformational change within the time scale, minimizing the biasing effect in the system. The TIP3P water model³² was employed. The particle mesh Ewald (PME) method was used with a direct space cut-off of 12 Å³³. The damping coefficient for the Langevin dynamics simulation was 5 ps⁻¹. The Nosé–Hoover method was used to maintain constant pressure (1 atm)³⁴. The TMD trajectory was performed for 10 ns of simulation time in the NPT ensemble at 310 K. A -2 charge was assigned to the benzene ring of the phosphorylated Tyr416.

Dynamical cross-correlation analysis. Dynamical cross-correlation^{20–22} is a useful method to analyse the correlation between residues in trajectories of MD simulations.

$$C_{ij} = \frac{\langle (r_i(t) - \langle r_i(t) \rangle) (r_j(t) - \langle r_j(t) \rangle) \rangle}{\sqrt{(\langle r_i^2(t) \rangle - \langle r_i(t) \rangle^2) (\langle r_j^2(t) \rangle - \langle r_j(t) \rangle^2)}} \quad (2)$$

where $r_i(t)$ and $r_j(t)$ are the atomic positions of the i^{th} and j^{th} C α atoms at time t . The quantity $r_i(t) - \langle r_i(t) \rangle$ corresponds to the fluctuation of the i^{th} atom. In a similar way, $r_j(t) - \langle r_j(t) \rangle$ corresponds to the fluctuation of the j^{th} atom. We obtained a 450 × 450 correlation map for the conformational changes of all the C α atoms during the 10 ns TMD simulation.

Network analysis: centrality measure. Network analysis has become extremely useful in a wide variety of complex systems^{35–38}. The quantity C_{ij} (Eq. 1) in the DCCM can be interpreted as an adjacency matrix. We constructed a network based on it. For this purpose, we defined the weight w_{ij} of the edge between the nodes i and j as^{39,40}

$$w_{ij} = -\log|C_{ij}| \quad (3)$$

The weight is the probability of information transfer across the edge, as measured by the DCCM obtained from the TMD simulation. Since C_{ij} has few 0 values the constructed network is almost fully connected. For simplicity of analysis, we introduced the threshold C_{thr} : if $C_{ij} < C_{thr}$ the edge between i and j has no weight (disconnected, $w_{ij} = 0$). We performed the following analysis using several values of C_{thr} . Since the results were qualitatively the same in each case, we present the results only for $C_{thr} = 0.3$. In the constructed network, each node corresponds to a C α atom and each edge is an information transfer probability (i.e., cross-correlation).

To identify and quantify the nodes that occupy critical positions in a network, several centrality measures have been proposed, including the degree, betweenness, eigenvector and closeness centralities^{40–42}. In this work, we used three centrality measures based on the DCCM: degree centrality, closeness centrality and betweenness centrality. The degree centrality measures the number of edges incident on a node in a network, thus expressing the “popularity” of the node^{41,42}.

$$C_D(v_i) = d_i = \sum_j A_{ij} \quad (4)$$

where A_{ij} is the adjacency matrix: if $w_{ij} > 0$ then $A_{ij} = 1$, otherwise $A_{ij} = 0$. The closeness centrality is defined as the average length of the shortest paths between a node and all the other nodes in a network. This is a measure of how long it will take information to spread from a given node to the other nodes. The closeness centrality is defined as^{41,42}

$$C_C(v_i) = \frac{n-1}{\sum_{j \neq i} g(v_i, v_j)} \quad (5)$$

where $g(v_i, v_j)$ is the shortest path with a weight between two nodes i and j . The betweenness centrality is a measure of how many information pathways flow through a node in a network. The betweenness of a node i is the fraction of the shortest paths between pairs of nodes that pass through node i . It is defined as^{41–43}

$$b_i = \frac{\sum_{s < t} g_i^{st} / n_{st}}{\frac{1}{2}n(n-1)} \quad (6)$$

where g_i^{st} is the number of shortest paths from s to t with a weight that pass through node i and n_{st} is the total number of shortest paths from s to t . The three centrality measures were obtained using Bio3d^{44–46}.

The images of the nodes of the network (Fig. 8) were generated with the “refuse force directed layout” algorithm in Cytoscape²⁷.

References

1. Thomas, S. M. & Brugge, J. S. Cellular functions regulated by Src family kinases. *Annu. Rev. Cell Dev. Biol.* **13**, 513–609 (1997).
2. Frame, M. C. Newest findings on the oldest oncogene; how activated src does it. *J. Cell Sci.* **117**, 989–998 (2004).
3. Summy, J. M. & Gallick, G. E. Treatment for advanced tumors: SRC reclaims center stage. *Clin. Cancer Res.* **12**, 1398–1401 (2006).
4. Yeatman, T. J. A renaissance for SRC. *Nat. Rev. Cancer* **4**, 470–480 (2004).
5. Brown, M. T. & Cooper, J. A. Regulation, substrates and functions of src. *Biochim. Biophys. Acta* **1287**, 121–149 (1996).
6. Boggon, T. J. & Eck, M. J. Structure and regulation of Src family kinases. *Oncogene* **23**, 7918–7927 (2004).
7. Roskoski, R. Jr. Src protein-tyrosine kinase structure and regulation. *Biochem. Biophys. Res. Commun.* **324**, 1155–1164 (2004).
8. Okada, M. & Nakagawa, H. A protein tyrosine kinase involved in regulation of pp60c-src function. *J. Biol. Chem.* **264**, 20886–20893 (1989).
9. Zrihan-Licht, S. *et al.* Association of csk-homologous kinase (CHK) (formerly MATK) with HER-2/ErbB-2 in breast cancer cells. *J. Biol. Chem.* **272**, 1856–1863 (1997).
10. Biscardi, J. S., Ishizawa, R. C., Silva, C. M. & Parsons, S. J. Tyrosine kinase signalling in breast cancer: epidermal growth factor receptor and c-Src interactions in breast cancer. *Breast Cancer Res.* **2**, 203–210 (2000).
11. Roskoski, R. Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors. *Pharmacol. Res.* **94**, 9–25 (2015).
12. Sun, G., Sharma, A. K. & Budde, R. J. Autophosphorylation of Src and Yes blocks their inactivation by Csk phosphorylation. *Oncogene* **17**, 1587–1595 (1998).
13. Boerner, R. J. *et al.* Correlation of the phosphorylation states of pp60c-src with tyrosine kinase activity: the intramolecular pY530-SH2 complex retains significant activity if Y419 is phosphorylated. *Biochemistry* **35**, 9519–9525 (1996).
14. Maayan, A. Introduction to Network Analysis in Systems Biology. *Sci Signal.* **4**, 1–11 (2011).
15. Almaas, E. Biological impacts and context of network theory. *J. Exp. Biol.* **210**, 1548–1558 (2007).
16. Goh, K. I. *et al.* The human disease network. *Proc Natl Acad Sci USA* **104**, 8685–8690 (2007).
17. Xu, W., Doshi, A., Lei, M., Eck, M. J. & Harrison, S. C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell* **3**, 619–638 (1999).
18. Cowan-Jacob, S. W. *et al.* The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **13**, 861–871 (2005).
19. Schlitter, J. S., Engels, M., Kruger, P., Jacoby, E. & Wollmer, A. Targeted molecular dynamics simulation of conformational change – Application to the T ↔ R transition in insulin. *Mol. Simul.* **10**, 291–308 (1993).
20. Hünenberger, P. H., Mark, A. E. & Van Gunsteren, W. F. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *J. Mol. Biol.* **252**, 492–503 (1995).
21. Karplus, M. & Ichiye, T. Comment on a “fluctuation and cross correlation analysis of protein motions observed in nanosecond molecular dynamics simulations”. *J. Mol. Biol.* **263**, 120–122 (1996).
22. Kasahara, K., Fukuda, I. & Nakamura, H. A. novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer–DNA complex. *PLoS one* **9**, e112419 (2014).
23. Barabási, A. L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
24. Albert, R. & Barabási, A. L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
25. Newman, M. E. J. Power laws, Pareto distributions and Zipf’s law. *Contemp. Phys.* **46**, 323–351 (2005).
26. VanWart, A. T., Durrant, J., Votapka, L. & Amaro, R. E. Weighted Implementation of Suboptimal Paths (WISP): An optimized algorithm and tools for dynamical network analysis. *J. Chem. Theory Comput.* **10**, 511–517 (2014).
27. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Gen. Res.* **13**, 2498–504 (2003).
28. Pareto, V. *Cours d’Economie Politique*. (Droz, Geneva, 1896).
29. Phillips, J. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
30. MacKerell, A. D., Banavali, N. & Floppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **56**, 257–265 (2000).
31. MacKerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25**, 1400–1415 (2004).
32. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
33. Darden, T., York, D. & Pedersen, L. G. Particle mesh Ewald: An N · log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
34. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* **103**, 4613–4621 (1995).
35. Barabási, A. L. & Oltvai, Z. N. Network biology: understanding the cell’s functional organization. *Nat. Rev. Genet.* **5**, 101–113 (2004).
36. Boccaletti, S., Latora, V., Moreno, Y., Charez, M. & Hwang, D. U. Complex networks: Structure and dynamics. *Phys. Rep.* **424**, 175–308 (2006).
37. Wasserman, S. & Faust, K. *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences Cambridge University Press (1994).
38. Newman, M. E. J. The structure and function of complex networks. *SIAM Rev.* **45**, 167–256 (2003).
39. Sethi, A., Eargle, J., Black, A. A. & Luthey-Schulten, Z. Dynamical networks in tRNA: protein complexes. *Proc. Natl. Acad. Sci. USA* **106**, 6620–6625 (2009).
40. VanWart, A. T., Eargle, J., Luthey-Schulten, Z. & Amaro, R. E. Exploring residue component contributions to dynamical network models of allostery. *J. Chem. Theory Comput.* **8**, 2949–2961 (2012).
41. Freeman, L. C. Centrality in social networks conceptual clarification. *Soc. Networks* **1**, 215–239 (1978).
42. Mason, O. & Verwoerd, M. Graph theory and networks in biology. *IET Syst. Biol.* **1**, 89–119 (2007).
43. Newman, M. E. J. A measure of betweenness centrality based on random walks. *Soc. Networks* **27**, 39–54 (2005).
44. Grant, B. J., Rodrigues, A. P. C., Elsayy, K. M., McCammon, J. A. & Caves, L. S. D. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695–2696 (2006).
45. Skjæræven, L., Yao, X. Q., Scarabelli, G. & Grant, B. J. Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC Bioinformatics* **15**, 399 (2014).
46. Skjæræven, L., Jariwala, S., Yao, X. Q. & Grant, B. J. Online interactive analysis of protein structure ensembles with Bio3D-web. *Struct. Bioinformatics* **32**, 3510–3512 (2016).

Acknowledgements

This research was supported by the National Research Foundation (NRF) of Korea under the Grant Number (2015R1D1A1A01061125) (SW) and (2015R1D1A3A01018506) (SJP). This work was also supported by the Pukyong National University Research Funding in 2014 (CD20140872) (SW).

Author Contributions

H.J.Y. performed MD simulation and analysed data; S.L. analysed network data; S.J.\P. designed research; S.W. performed MD simulation and analysed data; S.L., S.J.P., and S.W. wrote manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-23964-5>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018