Acta Biochimica Indonesiana

https://doi.org/10.32889/actabioina.143

RESEARCH ARTICLE

Open Access

Mutation analysis of SHP2, SOS1, and SOS2 related to dysregulation of Ras/MAPK pathway in Noonan syndrome

Nihayatul Karimah*

Research Center for Vaccine and Drugs, National Research and Innovation Agency (BRIN), Cibinong, Indonesia

*Corresponding author: LAPTIAB, Kawasan Puspiptek, Muncul, Kec. Setu, Kota Tangerang Selatan, Banten 15314, Indonesia. Email: nihayatul.karimah@brin.go.id

ABSTRACT

Background: Noonan syndrome, characterized by short stature, congenital heart defects, and facial dysmorphology, results from dysregulation of the Ras/MAPK pathway. Mutations in Ras/MAPK pathway proteins such as SHP2, SOS1, and SOS2 are responsible for this condition.

Objective: This study aimed to model the mutations in SHP2, SOS1, and SOS2 using FoldX and predict the structural impact.

Methods: Mutations were sourced from the OMIM Database. Protein sequence was retrieved from UniProt, and evolutionary conservation profiles were estimated by ConSurf. The structures of SHP2 and SOS1 were obtained from Protein Data Bank, while the undefined structure of SOS2 was modeled using YASARA. FoldX was used to model the mutations in two steps: structure repair and residue mutation.

Results: The evolutionary conservation profile indicated that most mutations occur in the highly conserved residues. These mutations disrupt various important interactions at domain interfaces. The total energy changes were predominantly positive, indicating instability in the mutated proteins due to the loss of the domain interactions and some unfavorable local conformational changes.

Conclusion: FoldX is a valuable tool for modeling protein mutations and predicting altered function. The models demonstrate that the mutations contribute to the aberrant autoinhibitory control and catalytic activity of the proteins.

Keywords: Noonan syndrome, protein mutations, FoldX, Ras/MAPK pathway

Introduction

Noonan syndrome (NS) is a congenital genetic disorder characterized by several key features, including short stature, congenital heart defects, and facial dysmorphology. Less common manifestations include learning disabilities, feeding and behavioral problems, increased bruising and bleeding, bone marrow complications, and infertility [1]. The genetic defects associated with Noonan syndrome predominantly involve multiple genes that encode proteins within the Ras/mitogen-activated protein kinases (Ras/MAPK) pathway. Notably, a Southeast Asian study found that 60% of Noonan syndrome cases are linked to mutations in the *PTPN11* gene, which codes for the SHP2 protein [2]. The Ras/MAPK pathway is critical for cell proliferation and differentiation. Dysregulation of this pathway can lead to developmental issues in cells. This suggests the significant physiological role of the Ras/MAPK pathway in human development from embryogenesis onward.

Important proteins in the Ras/MAPK pathway include SHP2, SOS1, and SOS2 [3]. SHP2 is a protein tyrosine phosphatase (PTP) known for removing a phosphate group from phosphorylated tyrosine residues in other phosphotyrosine-containing proteins. The structure of SHP2 comprises several domains: an amino-SH2 domain, a carboxy-SH2 domain, a PTP domain, and a carboxy-tail [4]. In its inactive state, the amino-SH2 domain's

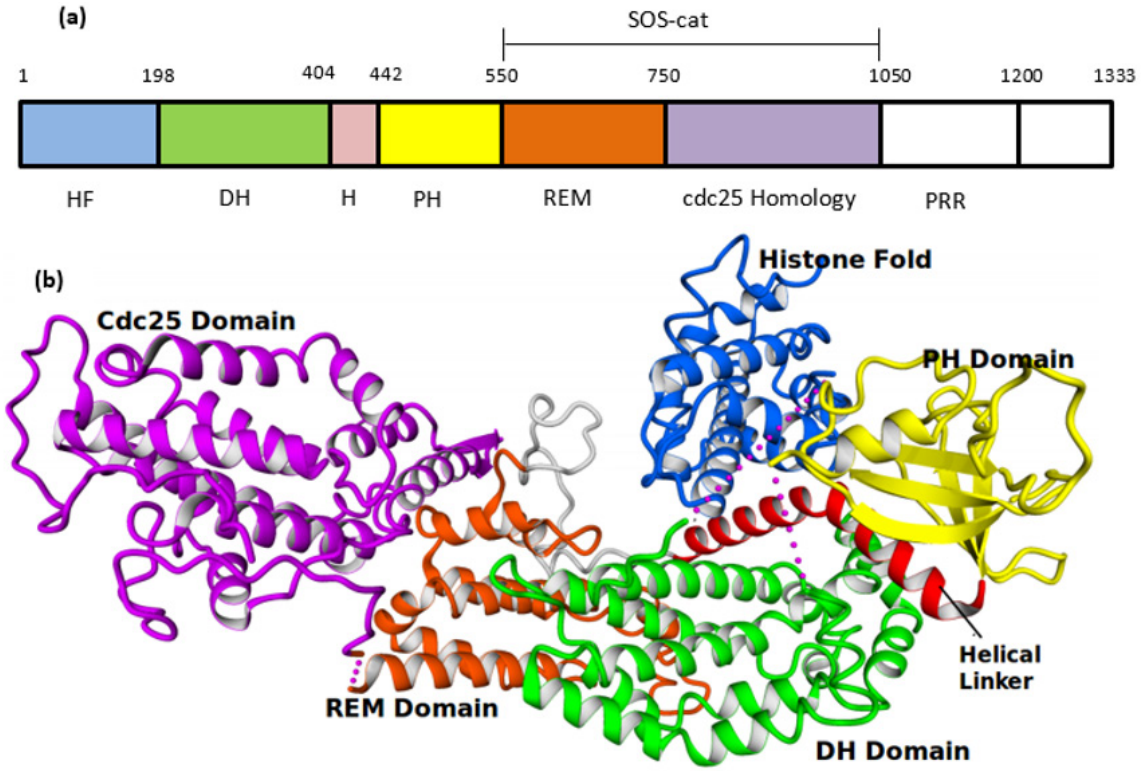

Figure 1. Structural overview and inactive conformation of SOS1. (a) The SOS structure is divided into a histone folds (HF) region, a Dbl Homology (DH) domain, a helical linker (H) region, a Pleckstrin Homology (PH) domain, a Ras Exchanger Motif (REM), a cdc25 homology domain, and a proline-rich region (PRR). (b) DH domain locks the REM domain, while the histone fold region restricts the PH domain from interacting with the effector. The helical linker connects the PH and REM domains and stabilizes the inactive conformation of SOS1. The colors used in Figure 1a correspond to the colors in Figure 1b. PDB ID: 3KSY.

phosphotyrosine binding pocket interacts with the catalytic residues of the PTP domain, resulting in autoinhibition. Meanwhile the carboxy-SH2 domain and the carboxy-tail remain undisturbed.

Two activation mechanisms of SHP2 have been proposed. The first mechanism is evidenced by a crystallographically data, where a small binding protein interacts with the carboxy-SH2 domain via its phosphotyrosine residue. This interaction facilitating the amino-SH2 domain's binding to another phosphotyrosine in the binding protein, leading to SHP2's open conformation. The second mechanisms involves the binding of two phosphotyrosines in the carboxy-tail, each to the amino- and carboxy-SH2 domains, freeing the PTP domain for activity. Following phosphorylation by receptor tyrosine kinase (RTK), SHP2 can function as an adaptor for the Grb2/SOS complex, triggering the activation of the Ras/MAPK pathway [5].

SOS, encoded by the *Son of Sevenless* genes, serves as a guanine nucleotide exchange factor (GEF) and comprises several functional domains that facilitate its interaction with Ras proteins during catalysis. This large protein is divided into distinct regions: a histone folds region (HF), a Dbl Homology (DH) domain, a Pleckstrin Homology (PH) domain, a Ras-exchanger motif (REM) domain, a cdc25 homology domain, and a proline-rich region (Figure 1a) [6]. The histone regulates the inactive conformation of SOS (Figure 1b) folds region and DH-PH domains, in which the DH domain locks the REM domain. In contrast, the histone fold conformation prevents SOS from binding with PIP2 in the membrane [7].

SOS activation is initiated by binding the SH3 domain of Grb2 with the proline-rich region of SOS. This interaction causes a conformational change in the histone folds and the DH domain, allowing negatively charged residues of PIP2 to interact with the arginine or lysine residues in the PH domain. This engagement drives SOS to interact with Ras through its cdc25 homology domain. Within this

domain, a helical hairpin, stabilized by the REM domain, facilitates the loosening of the switch-1 region of Ras, enabling guanine nucleotide exchange during catalysis [6]. SOS1 and SOS2 are highly homologous, but they appear to be functionally different proteins in the cytosol.

To study the mutational impact of specific amino acid residues in a protein structure, FoldX is beneficial to use, especially if a good quality crystal structure is available [8,9]. It will inform whether the mutations cause destabilization and influence its structure and function, as certain studies have reported that the application is more accurate in predicting destabilizing mutations [9]. One of the protein roles in the signal transduction pathway is becoming a molecular switch. The proteins will appear between active and inactive states with distinctive conformation. Analyzing mutation using an inactive state of structure could give insight into the influence of the mutation on inactivation control.

To study the mutational impact of specific amino acid residues in a protein structure, FoldX is particularly beneficial, especially if a high-quality crystal structure is available [8,9]. This tool helps determine whether mutations cause destabilization and influence the protein's structure and function, with studies indicating its accuracy in predicting destabilizing mutation [9]. Proteins in signal transduction pathways often act as molecular switches, alternating between active and inactive states with distinct conformations. Analyzing mutations in the inactive state can provide insights into how these mutations affect protein inactivation control. Conversely, examining mutations in the active state can shed light on potential changes in protein activity caused by the mutations. This study utilized FoldX to model mutations in SHP2, SOS1, and SOS2, which are known to dysregulate the Ras/MAPK pathway and contribute to the development of Noonan syndrome. The impact of these mutations on the protein structure was predicted based on the results from the modelling.

Method

Data retrieval

Data on the genetic heterogeneity of Noonan syndrome was retrieved from the Online Mendelian Inheritance in Man (OMIM) database [10]. The mutations in SHP2 (OMIM 176876), SOS1 (OMIM 182530), and SOS2 (OMIM 601247) associated with Noonan syndrome was taken from the allelic variant table of OMIM for each causative gene/locus. Information regarding protein function, family, domain, and sequence was retrieved from UniProt [11]. Similarities of presumed homolog sequences were examined using the BLAST feature on the MRS server [12]. Three-dimension crystal structures with resolutions < 4 Å were obtained from the Protein Data Bank (PDB) [13]. Evolutionary conservation profiles, estimated by ConSurf [14], provided conservation scores for each amino acid in the protein sequence, categorizing residues into non-conserved (scores 1, 2, or 3), conserved (scores 4, 5, or 6), and highly conserved (scores 7, 8, or 9).

Structural modeling

Structural or homology modeling aims to predict the 3D structure of a protein using the known structure of a homologous protein as a template. In this study, structural modelling was essential for defining the previously undefined structure of SOS2. According to Sander and Schneider [15], a higher percentage of sequence identity between the template and the model sequence typically results in a better model, with >90% sequence identity indicating an excellent model. For longer sequence, the acceptable sequence identify can be as low as 30%. The protein sequence for SOS2, provided by UniProt (accession code: Q07890), was used as input to model the unknown protein structure. The structure was built using the homology modeling macro in YASARA [16], which includes steps: template recognition and alignment, alignment correction, backbone generation, loop modeling, side-chain modeling, and energy minimization.

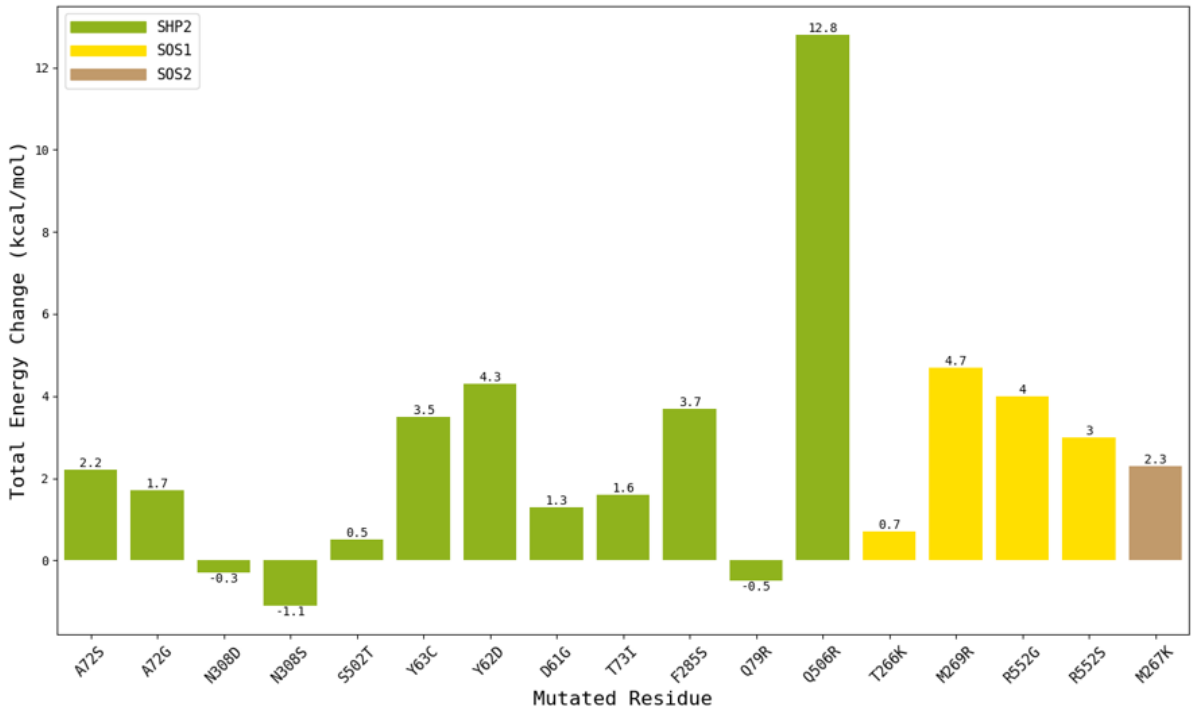

Figure 2. Total energy change caused by mutations in Noonan Syndrome-related proteins. Most of the mutations cause protein destabilization, as indicated by the positive value of energy change calculated using FoldX.

Energy change calculation

Mutation analysis was conducted using FoldX [17], which includes repairing the 3D structure and mutating the residue of interest. Initially, the 3D structures obtained from the PDB were repaired to achieve lower free energy by optimizing the side chains of amino acids. Repairing the 3D structure was conducted only for the structures from the PDB to improve the side-chain contacts to obtain better free energy. The side chain optimization was done by eliminating the bad contacts or the Van der Waals clashes. The residue mutation in the structure or the model structure was performed using the "build model" function in FoldX. This function also optimizes neighboring residues around the mutation site, shows the hydrogen-bond network change or Van der Waals clashes, and calculates the resulting free energy change caused by the mutation.

Results

The complete 3D structures of SHP2 (PDB ID: 2SHP) and SOS1 (PDB ID: 3KSY) are available in the PDB, although both are in the inactive states. The structure of SOS2 can be modeled using the SOS1 structure as a template. Mutations in SHP2 were detected in both the PTP domain and the N-SH2 domain, while mutations in SOS proteins were predominantly found in DBL homology and helical linker regions (Table 1).

The evolutionary conservation profile indicates that most mutations occur in highly conserved residues, which are crucial for the protein function. Some mutations were also found in conserved and poorly conserved residues, which nonetheless play roles in protein regulation. For example, Y62 in SHP2 is critical for forming aromatic interaction that contributes the maintain the protein in its inactive state. The application of mutation using FoldX generally resulted in in a positive value of total energy change for most proteins (Figure 2).

This suggests that the mutations tend to destabilize the proteins, potentially altering their activation states by affecting the binding affinity with effectors that activate the proteins.

Mutations in SHP2: A72S and A72G

In the inactive state, the wild-type alanine at position 72 (A72) in N-SH2 domain of SHP2 forms

Table 1. The mutated residues, location in the structure, and their evolutionary conservation

Protein	Genomic variation	Amino acid mutation	Location	Conservation score	Conservation level
SHP2	NM_002834.5 (PTPN11) : c.214 G>T	A72S	N-SH2 Domain	7	Highly conserved
	NM_002834.5 (PTPN11) : c.215 C>G	A72G	N-SH2 Domain	7	Highly conserved
	NM_002834.5 (PTPN11) : c.922 A>G	N308D	PTP Domain	8	Highly conserved
	NM_002834.5 (PTPN11) : c.923 A>G	N308S	PTP Domain	8	Highly conserved
	NM_002834.5 (PTPN11) : c.1504 T>A	S502T	PTP Domain	9	Highly conserved
	NM_002834.5 (PTPN11) : c.188 A>G	Y63C	N-SH2 Domain	5	Conserved
	NM_002834.5 (PTPN11) : c.184 T>G	Y62D	N-SH2 Domain	2	Poorly conserved
	NM_002834.5 (PTPN11) : c.182 A>G	D61G	N-SH2 Domain	6	Conserved
	NM_002834.5 (PTPN11) : c.218 C>T	T73I	N-SH2 Domain	7	Highly conserved
	NM_002834.5 (PTPN11) : c.854 T>C	F285S	PTP Domain	6	Conserved
	NM_002834.5 (PTPN11) : c.236 A>G	Q79R	N-SH2 Domain	5	Conserved
	NM_002834.5 (PTPN11) : c.1529 A>G	Q506R	PTP Domain	9	Highly conserved
	NM_002834.5 (PTPN11) : c.179_181del	DEL G60	N-SH2 Domain	7	Highly conserved
SOS1	NM_005633.4 (SOS1) : c.797 C>A	T266K	DBL Homology domain	5	Conserved
	NM_005633.4 (SOS1) : c.806 T>G	M269R	DBL Homology domain	8	Highly conserved
	NM_005633.4 (SOS1) : c.1654 A>G	R552G	Helical linker	9	Highly conserved
	NM_005633.4 (SOS1) : c.1656 G>T NM_005633.4 (SOS1) : c.1656 G>C	R552S	Helical linker	9	Highly conserved
SOS2	NM_006939.4 (SOS2) : c.800 T>A	M267K	DBL Homology domain	8	Highly conserved

a hydrogen bond with Q506 and hydrophobic contact with I282 and I463 in the PTP domain. The A72S mutant has a larger side chain with a polar hydroxyl group, while the A72G mutant lack a side chain, providing a provides flexible backbone. These changes in side chain size, flexibility, and polarity abolish the hydrophobic interaction of A72 with I282 and I463 in PTP domain. Both mutations, A72S and A72G, reduce the stability of the inactive state of SHP2. This destabilization impairs the ability of the N-SH2 domain to inhibit the catalytic cleft of the PTP domain, resulting in increased SHP2 activity and enhanced dephosphorylation of the substrate.

Mutations in SHP2: N308D and N308S

These mutations are located in the PTP domain, near the interface between PTP and N-SH2 domains. In the wild type, N308 forms hydrogen bonds with T288 and R501. In the N308D mutant, D308 forms hydrogen bond with N306, altering the contact and conformation of R501 near the Q-loop. Normally, R501 maintains a hydrogen bond network with the backbone of A461 in the P-loop and with N308. Consequently, the D308 mutant disrupts these interactions, potentially altering the local conformation of the P-loop and Q-loop. The N308S mutation, with a smaller side chain than the wild-type N308, diminshes the polar interactions at residue 208 with R501, causing changes in the hydrogen bond network of R501 with the Q-loop and P-loop. Despite this, R501 may form hydrogen bonds with other surrounding residues to anchor the catalytic cleft of the PTP domain. Both mutations, N308D and N308S, affect the the conformation of residues in the catalytic cleft, potentially altering the activity level of the PTP domain. These changes can lead to enhanced SHP2 activity.

Mutations in SHP2: S502T

The wild-type residue S502 is located near the P-loop and Q-loop in the PTP domain. It plays a crucial role in maintaining the inactivate state of SHP2 by forming a hydrogen bond with E76 in the N-SH2 domain, which stabilizes the interaction between the PTP and N-SH2 domains. The S502T mutation introduces a beta-branched side chain, which diminishes the polar interaction with E76. This change may decrease the stability of the inactive conformation of SHP2, potentially leading to enhanced SHP2 activity.

Mutation in SHP2: Y63C

In the inactive state of SHP2, the tyrosine residue at position 63 (Y63) in the N-SH2 domain is located at the interface between the N-SH2 domain and PTP domains. Y63 forms Van der Waals contacts with Q255 in the PTP domain. The Y63C mutation replaces this tyrosine with a cysteine, a smaller residue, disrupting these contacts. Additionally, the substitution with C63 creates a significant void and abolishes hydrophobic interactions with surrounding isoleucine residues, leadin to the destabilization of the inactive conformation of SHP2. This destabilization can result in enhanced SHP2 activity.

Mutation in SHP2: Y62E

Residue Y62 in the N-SH2 domain forms an aromatic interaction with Y279 in the PTP domain. The Y62E mutation causes the aromatic side chain of Y279 to flip and induces a slight conformational change in D64. This mutation creates a significant void at the interface with the P-loop, exposing the P-loop to the surface and potentially increasing PTP activity.

Mutation in SHP2: D61G

The D61G mutation in the N-SH2 domain also exposes the P-loop to the surface because glycine, with its lack of side chain steric hindrance, leaves a significant void at the interface. This mutation disrupts the hydrogen bond network involving D61 with K366 and S460 in PTP domain, causing local conformational change in PTP domain residues. Consequently, this mutation destabilizes the inactive state, leading to enhanced SHP2 activity.

Mutation in SHP2: T73I

The T73I mutation occurs at the interface between N-SH2 and PTP domains in the inactive state, near the Q-loop. The larger and more hydrophobic I73 residue causes local conformational change and the loss of hydrogen bonds among S502, G258, and G76. These hydrogen bonds are crucial for anchoring the interaction between the N-SH2 and PTP domains. Therefore, this mutation destabilize the inactive state and enhances SHP2 activity.

Mutation in SHP2: F285S

F285 in the PTP domain, near the Q-loop, engages in aromatic interaction with Y263 and H287. It also contributes to the interaction between the PTP and N-SH2 domains in the inactive state through Van der Waals interaction with E76. The S285 mutation creates a significant void in the closed conformation of SHP2, disrupting many local interactions provided by the wild-type F285. This destabilization of the inactive state leads to enhanced SHP2 activity.

Mutation in SHP2: Q79R

Q79R is located at the interface of N-SH2 and PTP domains, but far from the catalytic loops. This mutation induces a local conformational change in E83 and E76, leading to the loss of a hydrogen bond between E83 and Q79. The introduction of the positively charged Q79R disrupts the hydrophobic environment created by neighboring residues L262 and Y263 in the PTP domain. This disruption may destabilize the inactive state of SHP2, thereby enchanicng SHP2 activity.

Mutation in SHP2: T411M

T411M differs from the other Noonan syndrome-related mutations in SHP2 as it located on the

surface of the PTP domain, far from both the catalytic cleft and the domain interface. With a low conservation score (~1) among PTPase, T411 is a variable residue in the SHP2 sequence. Nonetheless, the T411M mutation introduces a hydrophobic residue on the protein surface, potentially impairing protein folding. This mutation also abolishes hydrogen bonds between T411, K405, and Q408.

Mutation in SHP2: Q506R

Q506 is located exactly in the Q-loop and on the interface of the N-SH2 and PTP domains. The Q506R mutation significantly impacts the local environment of the Q-loop by introducing a larger, positively charged side chain. The guanidinium group placement changes the local conformation of N58, R498, and Q255, resulting in the loss of hydrogen bonds between Q506 and N58/A72, as well as between R498 and Q255. The hydrogen bond network among R498, M504, and V505 is also affected. Given the high conservation score (~9) and a substantial reduction in protein stability (~12.8 kcal/mol), this mutation severely destabilizes the inactive state of SHP2 and alters the PTP domain's activity.

Mutation in SHP2: Deletion of G60

Residue G60 is crucial for maintaining the tight interaction between N-SH2 and PTP domains in the inactive state of SHP2, anchoring Q510 in the Q-loop through a hydrogen bond. Deleting G60 disrupts this bond and causes local conformational changes in Q506, T507, and Q510 in the catalytic Q-loop. Accordingly, this disruption could eliminate the autoinhibitory role of N-SH2 domain, thereby altering the phosphatase activity.

Mutations in SOS1: T266K and M269R

Both T266K and M269R mutations are located in the DH domain, right at the interface between DH domain and REM domain. The T266K mutation introduces a positively charged lysin with a longer side chain than the wild type threonine. This change is detrimental to domain interaction, as the positively charged lysine interferes with the conformation of W729 conformation in the REM domain. The M269R mutation replaces a hydrophobic methionine with a hydrophilic arginine, abolishing the hydrophobic interaction between M269 in the DH domain and W729 and L687 in the REM domain. Additionally, the introduction of a positive charge causes local conformational change in R694 and N691 in the REM domain. Therefore, both mutations disrupt the local environment at the DH and REM domain interface, potentially enhancing SOS1 activity since the DH domain naturally has auto-inhibitory function for SOS.

Mutations in SOS1: R552G and R552S

The wild-type R552 resides in the helical linker (or helical hairpin) connecting the DH and REM domains, as well as the PH and cdc25 homology domains. In the inactive state, R552 is situated at the interface between the helical hairpin and the histone folds. The R552G and R552S mutations replace the large arginine with much smaller residues, creating a significant void at the domain interface. These mutations disrupt the crucial electrostatic interaction between R552 in the helical hairpin and D140 in the histone fold, while this interaction is very important for the inactive state of SOS1. The mutations also cause local conformational change in D140, S548, and L144. Given the high conservation score of R552 (9) and the significant instability caused by the mutations (total energy change = 3 kcal/mol), these alterations likely make the PH domain more accessible to PIP2 in the membrane, thus enhancing SOS1 activation.

Mutations in SOS2: M267K

The M267K mutations in SOS2 is located at the interface between DH domain and REM domain. The introduction of a positively charged lysine at this position causes significant conformational changes in W727 and R692 in the REM domain. This mutation also abolishes the hydrophobic interaction between the wild-type M267 and

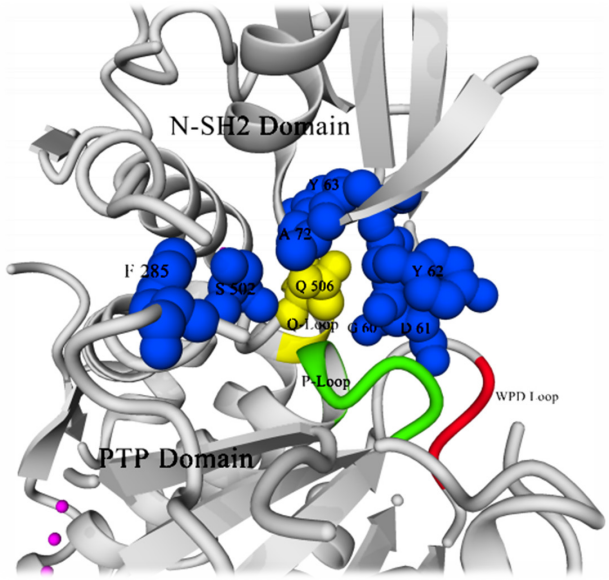

Figure 3. Autoinhibitory control by N-SH2 domain in SHP2. The N-SH2 domain of SHP2 performs auto-inhibitory control over the catalytic cleft of the PTP domain. The ball residues represent the mutated residues involved in direct domain contact. The Q-loop is colored yellow, the P-loop is colored green, while WPD loop is colored red. PDB ID: 2SHP.

W727, potentially resulting in a partially active SOS2. Consequently, the REM domain becomes more accessible to Ras-GTP, enhancing the catalytic activity of SOS2 toward Ras protein.

Discussion

Mutations that reduce auto-inhibitory control tend to occur in residues located at domain interfaces, where one domain inhibits another to keep the protein or enzyme inactive. The diminished control of auto-inhibition is predicted to ultimately lead to the partially active proteins and enhanced cellular MAPK pathway signaling.

In SHP2, most mutations are found at the interface between the N-SH2 and PTP domain (Figure 3), likely down-regulating the auto-inhibitory control exerted by the N-SH2 domain over the PTP domain. Similarly, in SOS proteins, mutations that introduce positively charged residues diminish the auto-inhibitory control of the DH domain over the REM domain. Additionally, substitutions with smaller residues reduce the auto-inhibitory control of the histone fold over the helical linker.

The reduction in auto-inhibitory control in SHP2, SOS1, and SOS2 is represented by the positive total energy changes calculated by FoldX, indicating instability of the inactive state due to the loss of the domain interactions and unfavorable local conformational changes. These findings align with previous studies that have shown mutantions in SHP2 activate the protein and contribute to the pathogenesis of Noonan syndrome [18]. Therefore, modelling mutations using FoldX is useful for predicting changes in protein function caused by conformational destabilization.

The 3D crystal structure of SHP2 in its inactive state reveals key residues involved in the interaction between N-SH2 and PTP domain. These residues, which are highly conserved, often undergo mutation and deletion in Noonan syndrome, underscoring their importance for the overall SHP2 structure. Notably, residues such as A72 and G60 in the N-SH2 domain substantially contribute to locking the Q-loop in the PTP domain by direct anchoring of Q506 and Q510, respectively. Other residues conceal the catalytic loops of the PTP domain through steric hindrance and mutual interactions with the neighboring loop residues.

Mutation models generated by FoldX demonstrate the disruption of various interactions, including hydrogen bonds, pi-pi stacking, hydrophobic contact, and van der Waals interaction. Specifically, three mutations in SHP2—N308D, N308S, and Q506R—are predicted to alter SHP2 catalytic activity by directly affecting the catalytic P-loop and Q-loop in the PTP domain. These mutations disrupt the hydrogen bond network and alter the local conformation of loop residues, potentially increasing SHP2 activity and enhancing the cellular MAPK pathway.

Mutants D308 and S308 seem to stabilize the P-loop and Q-loop by causing conformational and hydrogen bond network changes around these loops. The Q506R mutation in Q-loop may have the most severe impact on SHP2 activity because Q506 is crucial for both catalytic activity and auto-inhibitory control (Figure 4). Furthermore, the wild-type Q506 is highly conserved among SHPs, and the R506 mutation strongly destabilizes

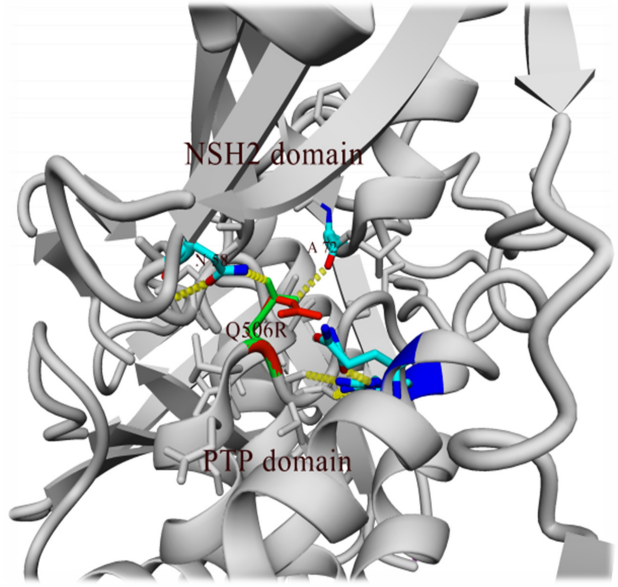

Figure 4. Impact of Q506R mutation in the Q-loop of SHP2. The Q506R mutation in the Q-loop of SHP2 abolishes hydrogen bond interaction with N58 and A72 in the N-SH2 domain, causing local conformational change around catalytic cleft. This mutation may diminish auto-inhibitory control and alter the catalytic activity level.

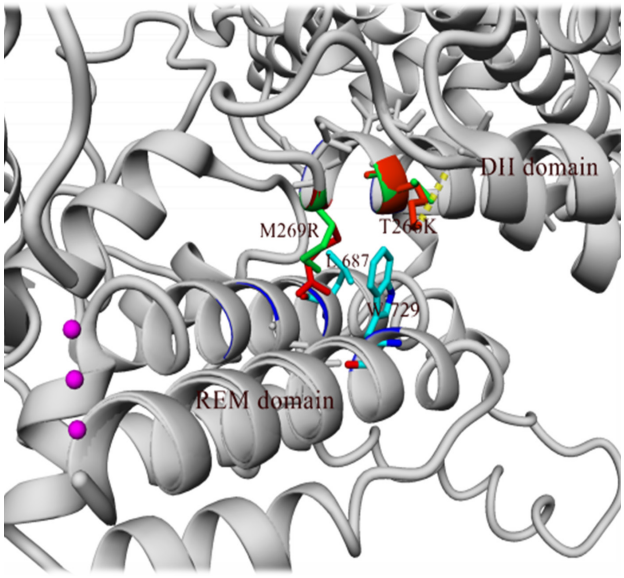

Figure 5. Disruption of auto-inhibitory control in SOS1 by K266 and R269 mutations. The positively charged mutations K266 and R269 in SOS1 may disrupt the auto-inhibitory control of the DH domain over the REM domain by introducing a charge near the hydrophobic residues W729 and L687.

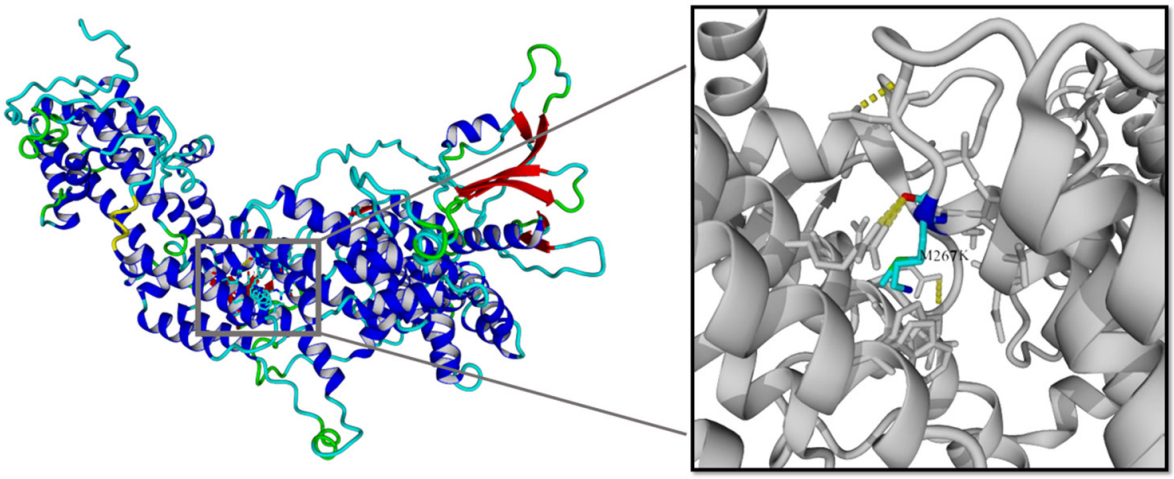

Figure 6. Structural model of SOS2 highlighting M267K mutation. 3D structural model of SOS2 visualizing the M267K mutation, which changes the interaction network of W727 and R692 in the REM domain.

the protein structure in its inactive state. The arginine at position 506 enhances catalytic activity by providing more stable coordination to water during SHP2-PO_4 hydrolysis in the end of dephosphorylation.

SOS1 and SOS2 also experience reduced auto-inhibitory control due to mutations leading to Noonan syndrome. All mutated residues are highly conserved among SOS family. In SOS1, the DH domain and histone fold carry out the auto-inhibitory function to maintain the protein's inactive state. Positively charged mutations like K266 and R269 in the DH domain disrupt the hydrophobic environment around W729 and L687 in the REM domain (Figure 5). This disruption is also observed in SOS2, where the M267K mutation in the DH domain converts a hydrophobic methionine to a positively charged lysine, destabilizing the

hydrophobic interaction and releasing DH domain control over the REM domain (Figure 6). Similarly, smaller mutations like G552 and S552 in SOS1 disrupt the crucial electrostatic interaction between the histone fold and helical linker, crucial for auto-inhibition. The positive total energy change caused by these mutations indicates a decrease in SOS1 structural stability.

Conclusion

Modelling studies using FoldX, combined with structural biology knowledge, clearly demonstrate that mutations disrupt the autoinhibitory control and catalytic activity of SHP2, SOS1, and SOS2 proteins, contributing to Noonan syndrome. Most of these mutations are gain-of-function mutations, which may enhance signal transduction in the Ras/MAPK pathway. FoldX is a useful tool for modelling protein mutations that lead protein destabilization and distinctive protein function. Further research, particularly site-directed mutagenesis, is recommended to elucidate the detailed mechanisms and impacts of each mutation. Understanding the structural details of mutation in Noonan syndrome-related proteins can be instrumental in developing effective prevention and treatment strategies.

Acknowledgments

The author thanks Prof. Dr. Gert Vriend and Dr. Hanka Venselaar from Centre of Molecular and Biomolecular Informatics (CMBI), Radboud University Nijmegen for giving the insight about Noonan syndrome and inspiring the author to apply the knowledge of structural biology in this study.

Funding

This work was supported by Centre of Molecular and Biomolecular Informatics (CMBI), Radboud University Nijmegen, and Indonesia Endowment Fund for Education (LPDP) Scholarship.

Declaration of interest

None.

Received: September 13, 2023
Revised: January 4, 2024
Accepted: February 7, 2024
Published online: August 4, 2024

References

1. Allen MJ, Sharma S. Noonan Syndrome. StatPearls. Treasure Island (FL): StatPearls Publishing; 2023. Available: http://www.ncbi.nlm.nih.gov/books/NBK532269/
2. Koh A, Tan E, Brett MS, Lai AHM, Jamuar SS, Ng I, et al. The spectrum of genetic variants and phenotypic features of Southeast Asian patients with Noonan syndrome. Molec Gen & Gen Med. 2019;7: e00581. https://doi.org/10.1002/mgg3.581
3. Tartaglia M, Zampino G, Gelb BD. Noonan Syndrome: Clinical Aspects and Molecular Pathogenesis. Mol Syndromol. 2010;1: 2-26. https://doi.org/10.1159/000276766
4. Diop A, Santorelli D, Malagrinò F, Nardella C, Pennacchietti V, Pagano L, et al. SH2 Domains: Folding, Binding and Therapeutical Approaches. IJMS. 2022;23: 15944. https://doi.org/10.3390/ijms232415944
5. Buday L, Vas V. Novel regulation of Ras proteins by direct tyrosine phosphorylation and dephosphorylation. Cancer Metastasis Rev. 2020;39: 1067-1073. https://doi.org/10.1007/s10555-020-09918-2
6. Findlay GM, Pawson T. How is SOS activated? Let us count the ways. Nat Struct Mol Biol. 2008;15: 538-540. https://doi.org/10.1038/nsmb0608-538
7. Bandaru P, Kondo Y, Kuriyan J. The Interdependent Activation of Son-of-Sevenless and Ras. Cold Spring Harb Perspect Med. 2019;9: a031534. https://doi.org/10.1101/cshperspect.a031534
8. Van Durme J, Delgado J, Stricher F, Serrano L, Schymkowitz J, Rousseau F. A graphical interface for the FoldX forcefield. Bioinformatics. 2011;27: 1711-1712. https://doi.org/10.1093/bioinformatics/btr254
9. Buß O, Rudat J, Ochsenreither K. FoldX as Protein Engineering Tool: Better Than Random Based Approaches? Computational and Structural Biotechnology Journal. 2018;16: 25-33. https://doi.org/10.1016/j.csbj.2018.01.002
10. Amberger JS, Bocchini CA, Scott AF, Hamosh A. OMIM.org: leveraging knowledge across phenotype-gene relationships. Nucleic Acids Research. 2019;47: D1038-D1043. https://doi.org/10.1093/nar/gky1151
11. The UniProt Consortium, Bateman A, Martin M-J, Orchard S, Magrane M, Ahmad S, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Research. 2023;51: D523-D531. https://doi.org/10.1093/nar/gkac1052

12. Hekkelman ML, Vriend G. MRS: a fast and compact retrieval system for biological data. Nucleic Acids Research. 2005;33: W766-W769. https://doi.org/10.1093/nar/gki422

13. Burley SK, Bhikadiya C, Bi C, Bittrich S, Chao H, Chen L, et al. RCSB Protein Data Bank (RCSB.org): delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. Nucleic Acids Research. 2023;51: D488-D508. https://doi.org/10.1093/nar/gkac1077

14. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 2016;44: W344-W350. https://doi.org/10.1093/nar/gkw408

15. Sander C, Schneider R. Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins. 1991;9: 56-68. https://doi.org/10.1002/prot.340090107

16. Land H, Humble MS. YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations. In: Bornscheuer UT, Höhne M, editors. Protein Engineering. New York, NY: Springer New York; 2018. pp. 43-67. https://doi.org/10.1007/978-1-4939-7366-8_4

17. Delgado J, Radusky LG, Cianferoni D, Serrano L. FoldX 5.0: working with RNA, small molecules and a new graphical interface. Valencia A, editor. Bioinformatics. 2019;35: 4168-4169. https://doi.org/10.1093/bioinformatics/btz184

18. Zhu G, Xie J, Kong W, Xie J, Li Y, Du L, et al. Phase Separation of Disease-Associated SHP2 Mutants Underlies MAPK Hyperactivation. Cell. 2020;183: 490-502.e18. https://doi.org/10.1016/j.cell.2020.09.002