ISSN: 0974-2115
www.jchps.com Journal of Chemical and Pharmaceutical Sciences

In-Silico Structural Modeling and Molecular Dynamics Simulation of Pathogen-Associated Molecular Pattern RAXX21

M H M Mubassir[1], M Abu Naser[2], Mohd Firdaus Abdul-Wahab[1], Salehhuddin HAMDAN[*1]
[1]Faculty Bioscience and Medical Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia
[2]Scientific Computing Solutions, House No. 19, Road No. 1, Block- B, Section-6, Mirpur, Dhaka-1207, Bangladesh
***Corresponding author: E-Mail: saleh65@utm.my, Cell No: + (6)0167314086, Tel: + (6)07-5557547**

ABSTRACT

The first layer of defence mechanism in plant known as pattern-triggered immunity (PTI) begins with the sense of pathogen-associated molecular patterns (PAMPs) by pattern-recognition receptors (PRRs). During the event PRRs bind with PAMPs and recruit co-receptor protein to activate the defence signal. To understand the mechanism properly, modeling of 3D structure of PAMPs and analysis of its structure is of huge importance. Although several plant PRRs have been discovered, but very few PAMPs have been discovered and characterized. This study describes the computational 3D modeling approach of a PAMP protein, RaxX21 and its important atomic characteristics. 3D structure shows that RaxX21 is mainly composed of coil structure. Also 500 ns MD simulation study reveals that sulfation at its tyrosine region which is done by the bacteria itself before secretion, remarkably improves the stability of RaxX21 which may contribute significantly in case of PTI.

KEYWORDS: PAMP, RaxX21, MD simulation, Defence, Sulhfotyrosine.

1. INTRODUCTION

Plant confers two layer of defence mechanism to protect itself from fungi, bacteria and other pathogens and microorganisms. The first layer of defense mechanism is referred as Pattern triggered immunity (PTI). Pattern recognition receptors (PRRs) which are cell surface localized receptors recognize conserved molecular signatures in this case and the immunity achieved. These molecular signatures are known as pathogen-associated molecular patterns (PAMPs) or microbial-associated molecular patterns (MAMPs) or damage-associated molecular patterns (DAMPs) secreted by microorganisms like bacteria, fungi etc.

Following the PTI mechanism, PRR Xa21 activates the defence mechanism in rice plant by recognizing its PAMP RaxX21. The early events governing RaxX21 activation can be shortly elucidated as the association of bacterial PAMP RaxX21 with rice RaxX21 and then, association of co-receptor OsSerk2 with this complex.

Xanthomonas oryzae pv. *oryzae (Xoo)* bacteria which is responsible for most destructive bacterial disease of rice secrets RaxX21 through its type 1 secretion system. But before the secretion, this RaxX21 become sulfated with the help of rax genes raxP and raxQ. These encode an ATP sulfurylase and an adenosine 5'- phosphor sulfate (APS) kinase and function in concert to produce 3'-phosphoadenosine 5'-phosphosulfate (PAPS), the universal sulfuryl group donor. This sulfation takes place at the 7th amino acid tyrosine of RaxX21 and is known as sulfotyrosine (YSU). As an important part of plant defense mechanism, study of PAMPs is of huge importance. Although most of the PRRs are identified and illustrated clearly but PAMPs for specific PRRs are still unknown. Also most of their mechanism of interaction is yet to be discovered. This is the study of modeling and analyzing PAMP RaxX21 3D structure as a first step of understanding PTI mechanism mediated by PRR Xa21 in rice plant.

2. MATERIALS AND METHODS

Sequence based analysis: As a newly discovered peptide, the Amino acid (AA) sequence of the target RaxX21 is not available yet either in NCBI or Uniprot KB. The sequence of RaxX21 was therefore retrieved from the findings of Pruitt, to have an initial idea about the physio-chemical properties and secondary structure RaxX21, primary structure was used to analyze for predicting physio-chemical properties using ProtParam tool and the secondary structure was predicted using PSIPRED (V3.3) and SOPMA. To investigate the domain architecture, InterPro and SMART were used. Also phylogeny of RaxX21 was analyzed after doing NCBI Protein Blast. Four templates were selected according to the lowest e-value and a phylogenetic tree was constructed using Clustal Omega. Also multiple sequence alignment of RaxX21 with its four selected templates was done using PRALINE tool to find out the conserved region of the protein. Again multiple sequence alignment was done with other related PAMPs flg22 and elf-18 by PRALINE tool to see the sequence similarity and conserved regions among them.

Modeling of RaxX21: NCBI blastp and HHpred analysis of RaxX21 protein AA sequence was carried out against Protein Data Bank (PDB) using default parameter values to search for the suitable template for RaxX21 single template modeling. All these tools suggested no suitable template for modeling of RaxX21. Then different automated online modeling approach was carried out using modeler (in linked with HHpred server), CPH model 3.2, Geno3D, Sparks-X, Swiss-Model, Phyre2 intensive modeling, Raptor-X, Quark and AIDA. Then the sulfation in tyrosine region was done using Molden tool.

Structural validation: To evaluate the structural and geometrical consistency and reliability of the modeled proteins, several approaches were adopted. ERRAT was used to study the non-bonded interactions between different atoms types while, Verify-3D was subjected to assess the compatibility of the atomic models with its own AA sequence. To study the geometrical consistency of the modeled proteins, Ramachandran plot generated from RAMPAGE were assessed.

Molecular dynamic simulation of RaxX21: To refine and obtain the stable structure of RaxX21 protein, the protein modeled by Quark was subjected to molecular dynamic (MD) simulation with GROMACS software suite. Here both the normal and sulfated peptide was taken for MD simulation. The Gromacs 96 54a7 united force field was used to run the simulations. As YSU is a new residue, the YSU was introduced to the force field. Before running the simulation, the systems were solvated, neutralized, energy minimized and equilibrated. In case of solvation, the proteins were taken into a cubic box with a minimum distance 1Å between the protein surfaces and edges. Then the boxes with the proteins inside were solvated with simple point charge (SPC) water model. The systems were neutralized with one counter sodium (Na^+) ion by genion tool before energy minimization. Energy minimization was done through 1000 steps of steepest descent minimization to remove bad van der Waals contacts and generate a good starting structure for MD simulation. Then the systems were equilibrated for 2 ns NPT ensemble followed by 1 ns NVT ensemble maintaining a constant 1 atm. pressure and 300 K temperature, respectively. Finally a 500 ns MD simulation was carried out for both the systems maintaining a constant 1 atm. pressure and 300 K temperature. To treat the long range electrostatic interactions, particle mesh Ewald (PME) method was applied. Root mean square deviation (RMSD), radius of gyration (Rg), energy and root mean square fluctuations (RMSFs**),** solvent accessible surface area (SASA), and secondary structure of RaxX21 was calculated using GROMACS tools to monitor conformational changes over the simulation time. Also PyMol tool and VMD was used to monitor the changes in structure over the simulation period of time.

3. RESULTS AND DISCUSSION

Sequence based analysis: ProtParam analysis revealed that RaxX21 is of 21 AA with molecular weight 2162.35 Da. The isoelectric point (pI) of RaxX21 was computed 6.92 consistent with the slightly acidic properties and the aliphatic index computed 18.57 and the instability of RaxX21 is computed to be 50.79 which classifies the protein as unstable. Finally the grand average of hydrophobicity (GRAVY) index was computed -1.59 which indicates RaxX21 as hydrophilic protein. The secondary structure predicted by both PSIPRED and SOPMA showed that RaxX21 protein is mainly composed of random coil. SOPMA resulted RaxX21 protein composed of random coil (85.71%) while having extended strand (9.52%) and beta turn (4.76%). SOPMA secondary structure prediction also predicted alpha helix, random coil, 3_{10} helix, pi helix, beta-bridge, bend region and ambiguous states 0.00%. No domain were predicted by SMART and Interpro tool. Binding site prediction by Predict Protein shows that there are four binding site regions which are AA 1-3, AA 6-9, AA 11, AA 19-21. NCBI Protein blast resulted the PDB ID 2DRM, Chain E, Acanthamoeba Myosin I Sh3 Domain Bound to Acan125 as the best template for RaxX21 with 67% sequence identity. Also the phylogenic tree constructed by Clustal Omega resulted RaxX21 and 2DRM in the same node of the tree (Fig.1c).

Multiple sequence alignment (MSA) of RaxX21 with its four selected templates using PRALINE tool showed conserved region of the proteins. With the alignment score 6642 a maximum conservation of (score 7) was observed for 12^{th} AA alanine of the RaxX21 (Fig.1a). Also 4^{th}, 10^{th}, 14^{th}, 17^{th}, 20^{th}, and 21^{th} AA also found moderately conserved. Also the MSA of RaxX21 with its closest PAMPs shows good consistency with alignment score 378. The 3^{rd} AA glycine of RaxX21 showed maximum conserved score with PAMPs flg22 and elf-18 (Fig.1b). Again the 5^{th}, 6^{th}, 10^{th}, 13^{th}, 17^{th} and 20^{th} AA was observed moderately conserved among the PAMPs.

Modeling of RaxX21: Different modeling approach was carried out for modeling of RaxX21. Among those modeler (in linked with HHpred server), Sparks-X, AIDA and Quark modeled RaxX21. The other tools like CPHmodel 3.2, Geno3D, Swiss-Model, Phyre2 intensive modeling, Raptor-X could not generate any model due to its small AA sequence (21 AA) and their failure to generate any template/psiblast. The model which are generated shows that the RaxX21 protein is composed of only coil (Fig. 1 d) which is consistent with the result of secondary structure predicted by PSIPRED and SOPMA. Then the sulfation in tyrosine region which is known as sulfotyrosine (YSU) was done using Molden tool.

Structural validation: To validate the protein models modeled different tools, ERRAT, Varify3D and Ramachandran plot generated from RAMPAGE analysis was carried out. Among the tools only protein modeled by Quark tool gave value (30.769) for ERRAT model validation. Although the ERRAT value falls below the ideal protein quality value (50.00) it was improved after energy minimization, equilibration and MD simulation. Other modeled protein from tools like modeller (in linked with HHpred server), Sparks-X and AIDA resulted ERRAT value 0 (Table.1). In case of Verify-3D, all models 100% AA amino acids have scored >= 0.2 in the 3D/1D profile.

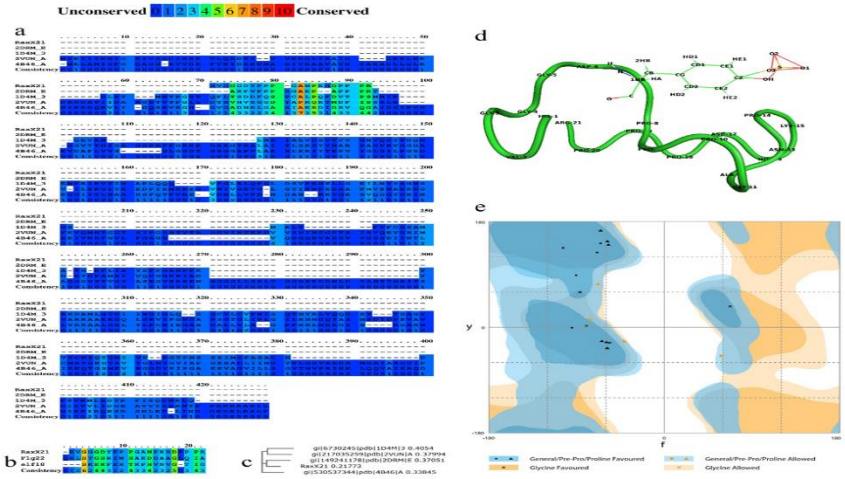


Figure.1.a) MSA of RaxX21 with its a) four selected templates, b) MSA of RaxX21 with its PAMPs, c) Phylogenic tree of RaxX21 showing the closest family members, d) Cartoon with line 3D structure of RaxX21-sY indicating sulfation at tyrosine region, e) Ramachandran plot summery generated by RAMPAGE server

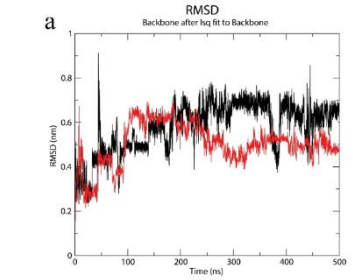
Table.1. Validation of model RaxX21 by different tools

Tool	ERRAT	Verify 3D (%)	Ramachandran Plot Summary from RAMPAGE (%)		
			*FR	*AR	*OR
Quark	30.769	100	78.9	21.1	0
Modeller	0	100	89.5	10.5	0
Sparks-X	0	100	100	0	0
AIDA	0	100	89.5	10.5	0
CPHmodel 3.2	Could not generate any model				
Geno3D	Could not generate any model				
Swiss-Model	Could not generate any model				
Phyre2	Could not generate any model				
Raptor-X	Could not generate any model				

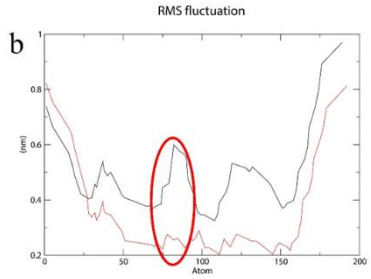
*FR, Favoured Region; AR, Allowed Region; OT, Outlier Region

RAMPAGE server was used to generate the result for Ramachandran plot. These servers resulted that none of the proteins are in outlier region (Table.1). In the case of Quark, 78.9% are in favoured region and 21.1% are in allowed region (Fig. 1e). No AA was found in outlier region which again proves the good quality structure of the protein. But as only Quark resulted ERRAT value, only this model was taken for further analysis and molecular dynamic simulation.

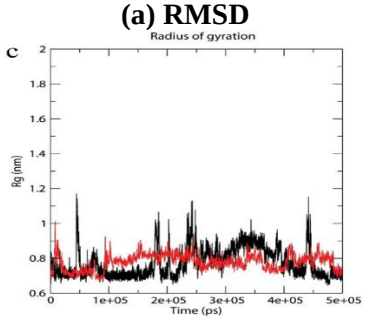
Molecular dynamic simulation of RaxX21-sY: To analyze the nature of RaxX21-sY in water different analysis was done. Analysis of the RMSD value showed upward trend up to 120 ns (Fig.2a). Then the protein becomes stable and the RMSD goes downward maintain a stable condition at 0.5 Aº until the end of the simulation. On the contrary, the RaxX21 shows flexible nature and higher RMSD compared to the sulfated one. Although it shows stable nature during 250 ns to 350 ns period (Fig.2a) but the overall RMSD graph shows unstable nature of the protein compared to RaxX21-sY. This proves that the sulfation in RaxX21 tyrosine region makes the protein more stable. RMS fluctuation was dominant at the terminal regions of RaxX21 for both proteins. For N-terminal region the RaxX21-sY RMS fluctuation is high whereas for the C-terminal region the RMS fluctuation in RaxX21-sY is low compared to RaxX21. Also, there is significant RMS fluctuation variation at the middle region of the protein. For the sulfated protein, the RMS fluctuation is very less whereas the normal protein the RMS fluctuation is very high. This also depicts that the sulfation is may be the reason for making the protein structure more stable at that region. The overall radius of gyration value showed almost stable condition for C-alpha atoms for RaxX21-sY (Fig.2c) and didn't show any significant drift while there are mild fluctuations observed up to 1.2 nm through the simulation period of time. Gyration radius stable value again supports the result of RMSD and RMS fluctuation for RaxX21-sY stable behavior compared to RaxX21. The solvent accessible surface area for both RaxX21 and RaxX21-sY showed almost stable behavior at a certain point (Fig.2d). Although the stable behavior is more prominent in case of RaxX21-sY. The DSSP analysis of the secondary structure analysis of the both RaxX21 protein and RaxX21-sY showed that the coil and bend were prominent over the 500 ns period of time (Fig.2e). Also there were turn and glimpse of beta-bridge and 3-helix throughout the simulation period of time (Fig.2f).



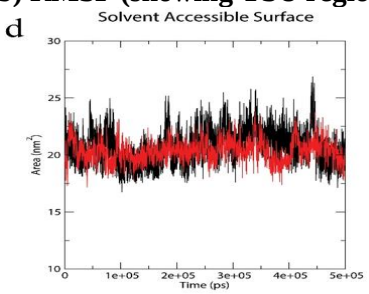
(a) RMSD



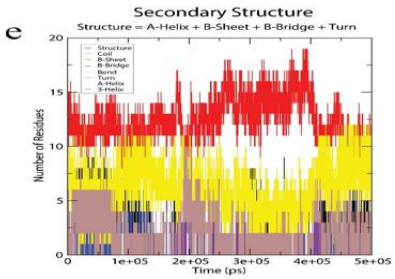
(b) RMSF (showing YSU region)



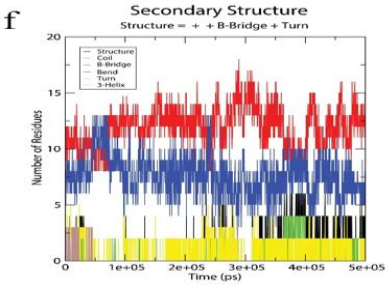
(c) Rg



(d) SASA



(e) secondary structure of RaxX21



(f) secondary structure of RaxX21-sY protein over 500ns period of time.

Figure.2. (a), (b), (c) & (d) express the values for RaxX21-sY whereas the black graph is the values for RaxX21

4. CONCLUSION

The *in-silico* 3D structure prediction and molecular dynamic (MD) simulation analysis of PAMP RaxX21-sY can contribute significantly as an initial step towards understanding the defence mechanism of rice plant mediated by PRR Xa21. The results proved that the structure of RaxX21-sY is mainly composed of coil. Moreover, 500 ns MD simulation study showed that the sulfation in RaxX21 shows significant stable behavior compared to the nonsulfated one. This stability of the protein is dominant in the sulfation area of the protein which may be a significant characteristics of RaxX21 in case of attacking the plant. Further research needed to be done to understand the key role of sulfation at tyrosine region and the proper action mechanism of RaxX21-sY.

5. ACKNOWLEDGEMENT

The authors are grateful to Asita Elengoe for her valuable suggestions on result analysis.

REFERENCES

Buchan DW, Ward S, Lobley AE, Nugent T, Bryson K, Jones DT, Protein annotation and modelling servers at University College London, Nucleic acids research, gkq427, 2010.

Combet C, Jambon M, Deléage G, Geourjon C, Geno3D, automatic comparative molecular modelling of protein, Bioinformatics, 18 (1), 2002, 213-214.

DeLano WL, PyMOL: An Open-Source Molecular Graphics Tool, CCP4 Newsletter on Protein Crystallography, 40, 2002, 44-53.

Eswar N, Webb B, Marti-Renom MA, Madhusudhan M, Eramian D, Shen My, Comparative protein structure modeling using Modeller, Current protocols in bioinformatics, 5.6 (1-5), 2006, 6.30.

Fiser A, Protein structure modeling in the proteomics era, Expert review of Proteomics, 1 (1), 2004, 97-110.

Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A, Protein identification and analysis tools on the ExPASy server, Springer, 2005.

Geourjon C, Deleage G, SOPMA, significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments, Computer applications in the biosciences, CABIOS, 11 (6), 1995, 681-684.

Gomez-Gomez L, Boller T, FLS2, An LRR receptor like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis, Molecular cell, 5 (6), 2005, 1003-1011.

Huang YJ, Mao B, Aramini JM, Montelione GT, Assessment of template-based protein structure predictions in CASP10, Proteins, Structure, Function, and Bioinformatics, 82 (S2), 2014, 43-56.

Humphrey W, Dalke A, Schulten K, VMD, visual molecular dynamics, Journal of molecular graphics, 14 (1), 1996, 33-38.

Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Inter Pro, the integrative protein signature database, Nucleic acids research, 37 (1), 2009, D211-D215.

Kallberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, eTemplate-based protein structure modeling using the RaptorX web server, Nature protocols, 7 (8), 2012, 1511-1522.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ, The Phyre2 web portal for protein modeling, prediction and analysis, Nature protocols, 10 (6), 2015, 845-858.

Liithy R, Bowie J, Eisenberg D, Assessment of protein models with three-dimensional profiles, Nature, 356 (6364), 1992, 83-85.

Lovell SC, Davis IW, Arendall WB, de Bakker PI, Word JM, Prisant MG, Structure validation by Cα geometry, ϕ, ψ and Cβ deviation, Proteins, Structure, Function, and Bioinformatics, 50 (3), 2003, 437-450.

Mahram A, Herbordt MC, Fast and accurate NCBI BLASTP, acceleration with multiphase FPGA-based prefiltering, Proceedings of the 24th ACM International Conference on Supercomputing, ACM, 2010.

Medzhitov R, Janeway CA, Innate immunity: the virtues of a nonclonal system of recognition, Cell, 91 (3), 1997, 295-298.

Nekrasov V, Li J, Batoux M, Roux M, Chu ZH, Lacombe S, Control of the pattern-recognition receptor EFR by an ER protein complex in plant immunity, The EMBO journal, 28 (21), 2009, 3428-3438.

Petrov D, Margreitter C, Grandits M, Oostenbrink C, Zagrovic B, A systematic framework for molecular dynamics simulations of protein post-translational modifications, PLoS Comput Biol, 9 (7), 2013, e1003154.

Pruitt RN, Schwessinger B, Joe A, Thomas N, Liu F, Albert M, The rice immune receptor XA21 recognizes a tyrosine-sulfated protein from a Gram-negative bacterium, Science advances, 1 (6), 2015, e1500245.

Ronald PC, Beutler B, Plant and animal sensors of conserved microbial signatures, Science, 330 (6007), 2010, 1061-1064.

Ronald PC, The Role of RaxST, a Prokaryotic Sulfotransferase, and RaxABC, a Putative Type I Secretion System, in Activation of the Rice XA21-Mediated Immune Response, Scientifica, 2014.

Schaftenaar G, Noordik JH, Molden, a pre-and post-processing program for molecular and electronic structures, Journal of computer-aided molecular design, 14 (2), 2000, 123-134.

Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, Definition and testing of the GROMOS force-field versions 54A7 and 54B7, European biophysics journal, 40 (7), 2011, 843-856.

Schultz J, Milpetz F, Bork P, Ponting CP, SMART, a simple modular architecture research tool, identification of signaling domains, Proceedings of the National Academy of Sciences, 95 (11), 1998, 5857-5864.

Schwede T, Kopp J, Guex N, Peitsch MC, SWISS-MODEL, an automated protein homology-modeling server, Nucleic acids research, 31 (13), 2003, 3381-3385.

Shen Y, Sharma P, Silva FGd, Ronald P, The Xanthomonas oryzae pv.◊ oryzae raxP and raxQ genes encode an ATP sulphurylase and adenosine-5′-phosphosulphate kinase that are required for AvrXa21 avirulence activity, Molecular microbiology, 44 (1), 2002, 37-48.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Molecular systems biology, 7 (1), 2011, 539.

Simossis VA, Heringa J, Praline, a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information, Nucleic acids research, 33 (suppl 2), 2005, W289-W294.

Soding J, Biegert A, Lupas AN, The HHpred interactive server for protein homology detection and structure prediction, Nucleic acids research, 33 (2), 2005, W244-W248.

Song W.Y, Wang G.L, Chen L.L, Kim H.S, Pi L.Y, Holsten T, A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21Science, 270 (5243), 1995, 1804-1806.

Sun Y, Li L, Macho AP, Han Z, Hu Z, Zipfel C, et al. Structural basis for flg22-induced activation of the Arabidopsis FLS2-BAK1 immune complex, Science, 342 (6158), 2013, 624-628.

Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ, GROMACS, fast, flexible, and free, Journal of computational chemistry, 26 (16), 2005, 1701-1718.

Van der Spoel D, Van Maaren PJ, Berendsen HJ, A systematic study of water models for molecular simulation, derivation of water models optimized for use with a reaction field, The Journal of chemical physics, 108 (24), 1998, 10220-10230.

Wallner B, Elofsson A, Can correct protein models be identified? Protein Science, 12 (5), 2003, 1073-1086.

Xu D, Jaroszewski L, Li Z, Godzik A, AIDA, ab initio domain assembly server, Nucleic acids research, 42 (W1), 2014, W308-W313.

Xu D, Zhang Y, Ab initio protein structure assembly using continuous structure fragments and optimized knowledge based force field, Proteins, Structure, Function, and Bioinformatics, 80 (7), 2012, 1715-1735.

Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JD, Boller T, Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts, Agrobacterium-mediated transformation Cell, 125 (4), 2006, 749-760.