Journal of Computational Intelligence in Bioinformatics
ISSN 0973-385X Volume 4, Number 2 (2011) pp. 199-207

http://www.ripublication.com/jcib.htm

Homology Modeling of Human SMAD-1 for Pulmonary Arterial Hypertension using Modeller 9v8.

Vikas Sharma[1] and Amita Mahajan[2]

[1]Research Scholar, Singhania University, Pacheri Bari, Rajasthan, India
E-mail: vikasons@yahoo.com
[2]RBIEBT, Sahaura Kharar, Dist. Mohali, Punjab, India
E-mail: amitaanil44@yahoo.com

Abstract

SMADs are group of related intracellular proteins important for transmitting signals to nucleus from transforming growth factor-β(TGF-β) superfamily at cell surface. Activation of TGF-β superfamily receptors lead to phophorylation of SMAD proteins which functions as transcription factors to regulate gene expression. Receptor regulated SMADs(R- SMADs) SMAD1/5/8 participate in Signalling downstream of BMP receptors. On the basis of crystal structure of SMAD1(PDB code 1khu) homology modeling was done using Modeller 9v8 program followed by analysis and validation using SAVES, ProSA and Procheck and the best model i.e.Model-2 is selected on basis of maximum Ramachandran.plot value and minimum bad contacts and minimum G-Value in comparison with other available models made by Modeller 9v8 for inhibitor design in Pulmonary arterial hypertension.

Keywords: BMP receptors, Modeller, R-SMADs, TGF-β

Introduction

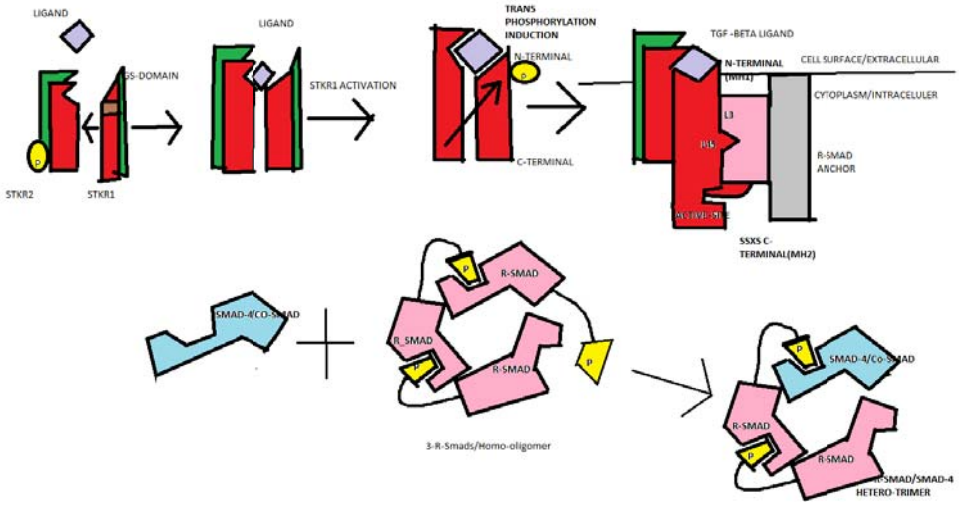
Bioinformatics has a number of applications including protein structure prediction, homology modeling, moleculer dynamics, primer designing and drug designing. Drug design is the approach of finding drugs by design, based on their biological targets. Typically a drug target is the key molecule involved in a particular metabolic or signaling pathway that is specific to a disease condition or pathology or to the infectivity or survival of a microbial pathogen. Among the different types of drug designing, Structure-based drug design is one of several methods in the rational drug

design.. SBDD uses the known 3D geometrical shape or structure of proteins to assist in the development of new drug compounds. A number of diseases have been very well treated from this approach of drug designing like diabetes, cancer, tuberculosis etc.

Pulmonary arterial hypertension (PAH) is rare condition characterized by increase in pulmonary vascular resistance(PVR) leads to right heart failure and without treatment, death occurs within three years of diagnosis. Etiology of PAH is multifactoral. In some patients there is a major genetic predisposition in the form of heterozygous germ line mutations in a transforming growth factor- β-superfamily receptor, the bone morphogenetic typeII receptor(BMPR-II)in addition, it is likely that additional factors, such as inflammation, are important to manifest disease. (TuderRM, MareckiJC, RichterA, et., al., 2007).Potential mechanism involved in the development of PAH includes TGF-β superfamily which has two branches TGF-β 1 ligand for which superfamily is named(SMAD2, 3)and BMP(Bone Morphogenetic Proteins)which includes SMAD1, 5, 8 or R-SMADs or Receptor Regulated SMADs along with GDFs Growth and Differntiation Factors, AMH i.e Anti Mullerian Hormone(Kawabata and Miyazono, 2000).Both branches of TGF-β superfamily make compex with Co-SMAD or Common SMAD(SMAD4) by sharing by R-SMADs in the ratio 2:1.Phosphorilation of BMPR1 initiates Phosphorilation of R-SMADs which excites conformational change in R-SMADs making an interaction between MH1 and MH2 domain that inhibit DNA binding.This opening or conformational change of R-SMADs permits their interaction with Co-SMAD and their translocation in the nucleus for regulation of cell proliferation of SMCs and apoptosis.

SMAD proteins are signal transducers and transcriptional, co-modulators of tgf-β superfamily of ligands(Attisano and Wrana, 2000;Derynck et al., 1998;Heldin et al., 1997;Massague and Wotton., 2000;Roberts., 1999).Phosphorylation of the C-terminal serine residues in R-SMADs by type1 receptor kinases is a crucial step in TGF-β family signaling(Abdollah., 1997;Macias Silva et al, 1996;Souchelnytskyi et al., 1997).Substrate specificity is determined by the L45 loop in the type 1 receptors and primarily, by the L3 loop in the R-SMAD MH2 domain, thus TGf-β and activin receptors phosphorylates SMAD2 and SMAD3, BMP receptors phosphorylate SMAD1, SMAD5 and SMAD8 (Chen et al., 1998).ligand binding to specific transmembrane receptor kinases induces receptor oligomerisation and phosphorylation of the receptor specific SMAD protein. R-SMADs in the cytoplasm(Lagna etal., 1996;Warna eet al., 1994;Zhang et al., 1996).The R-SMAD protein regulate distinct signaling pathways.SMAD1, SMAD5 and SMAD8 mediate the signals of bone morphogenetic proteins(BMPs), while SMAD2 and SMAD3 mediate the signals of TGF-βs and activins. The conserved signaling mechanism is the formation of a heteromeric complex between the phosphorylated R-SMADs and the common mediator SMAD4.The heteromeric complex enters the nucleus to regulate transcription of target genes. The R-SMADs and SMAD4 share a common domain configuration consisting of a conserved N-terminal DNA binding domain (MH1 domain)and a C-terminal MH2 domain separated by a variable linker region. The MH2 domain of an R-SMAD, but not SMAD4, has been shown to homo-oligomerise(Wu etal., 1997).Phosphorylation triggered heteromeric assembly between

SMAD 4 and R-SMAD is mediated by the C-terminal MH2 domain.The sites of phosphorylation have been mapped to the last two serine residues with in the conserved C-terminal SSXS sequence of the R-SMADs (Abdollah et al., 1997;Souchelnytskyi et al., 1997).however, the role of phosphorylation in subunit assembly as well as the stoichiometry of the heteromeric complex remains controversial. Phosphorylation has been proposed to contribute directly to subunit assembly by bridging the MH2 domain interaction(Chacko et al., 2001).Another model suggests that phosphorylation uncouples the intramoleculer inhibitory activity of the MH1 domain on the MH2 domain;allowing the MH2domains to associate constitutively(Hata et al., 1997).The heteromeric complex between SMAD4 and R-SMAD has been proposed to be hexameric, trimeric, and dimeric (Chacko et al., 2001;Kawabata et al., 1998;Shi et al., 1997;Wu et al., 2001).Here they report the crystal structure of SMAD1 MH2 domain in a confirmation, which together with the biochemical evidence reveals the structural basis of phosphorylation...The phosphorylated C-terminal tail of SMAD1 functions as a subunit assembly switch by forming specific contacts with the phosphor serine binding pockets of the neighbouring molecule further more the MH2 domain undergoes concerted conformational changes upon trimerisation, which may serve as a signaling switch.



Materials and Methods

An attempt will be made to model a target protein by steps described below:

Step-1: Target Identification: To start drug designing one needs to have sound knowledge about the disease and the potential targets those can be used for the generation of the drug molecule the most promising drug target is selected for drug development. On the basis of Literature study Human SMAD 5 is the target protein selected whose structure is un-known.

Step-2: Target Sequence Retrieval: To work with the protein (Human SMAD1), the sequence of this protein encoding for the disease retrieved. This is done by using the

protein sequence database NCBI. The sequence in FASTA format is obtained through these online databases.

Step-3: Target Sequence Alignment: Next step in queue is alignment of the sequence retrieved from NCBI is aligned with sequences from PDB database using BLASTp (Basic Local Alignment Search Tool) from sequence analysis and "Run Blast" option on right side of Genpept entry of Human SMAD1.

Step-4: Modelling the Target: Modeller (http://www.salilab.org/modeller/) 9v8 program is used for performing the comparative 3D modelling of the protein structure. The following three input files are required: 1.Atom file:Saved with extension.atm as file name.atm is obtained from PDB.

Alignment file: the homology program, MODELLER uses a special form of PIR format where the information about sequene numbering and chain codes are written into the "Description" line between the PIR protein tag and protein alignment entry.

Script file: this is an ordinary python 2.3 script to specify the alignment and the atom file for the modeller is given to Modeller 9v8.The modellar generates the models using the information provided as input files. The best model among these is selected for further process.

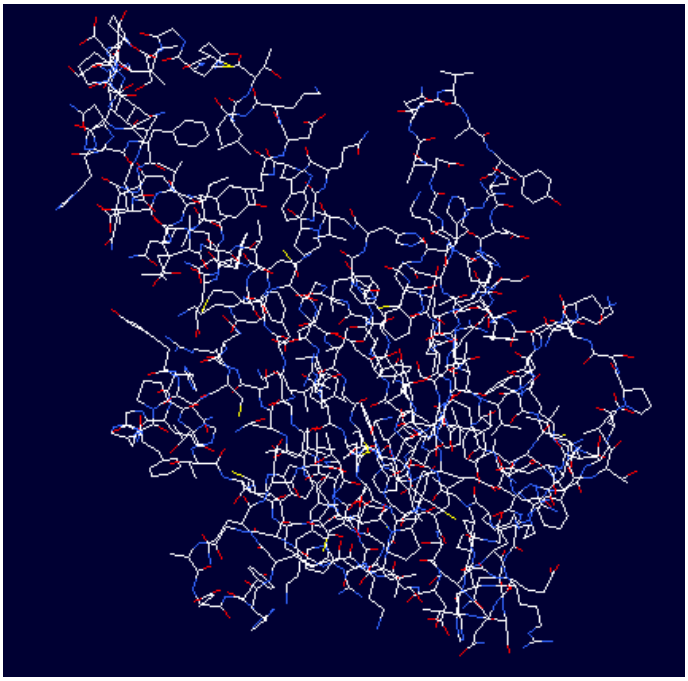
Step-5: Visualization of the Models: Deepview or SPDBV is used to visualize the models generated by the modeller. This gives the idea of the 3D configuration of the protein models generated by the modeller having minor differences in their and orientation. This is calculated by using Structure Analysis and Validation Server(SAVES)

Step-6: Analysis and Validation of Models: Structural analysis and validation server (SAVES) is used for a single validation program i.e Procheck. The structure submitted at SAVES are validated after validation model with maximum core value and zero bad contacts are used for further process ProSA (Protein structure analysis)online server to get z score for each model. After the final analysis and validation process the best protein model of the target protein used for active site prediction.

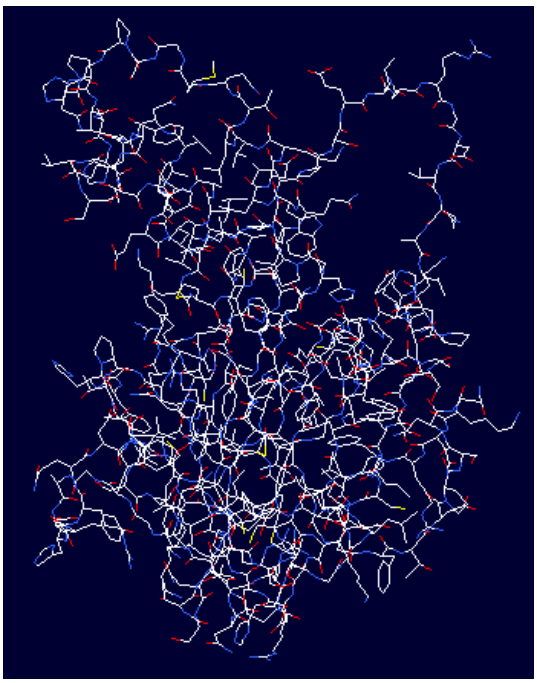
Results & Discussion

Homology Modelling of SMAD1 was done using modeller9v8 program. As output of the program we got five models in dot pdb format which were validated using SAVES and ProSA online tools getting their Ramachandran plot value and minimum bad contacts and minimum negative overall G-Value. Out of Five models, Model 2 is selected the best model because of it satisfied the set criteria for the best Model. Studies on SMAD proteins revealed the signal transduction pathway of the members in the TGF-β superfamily

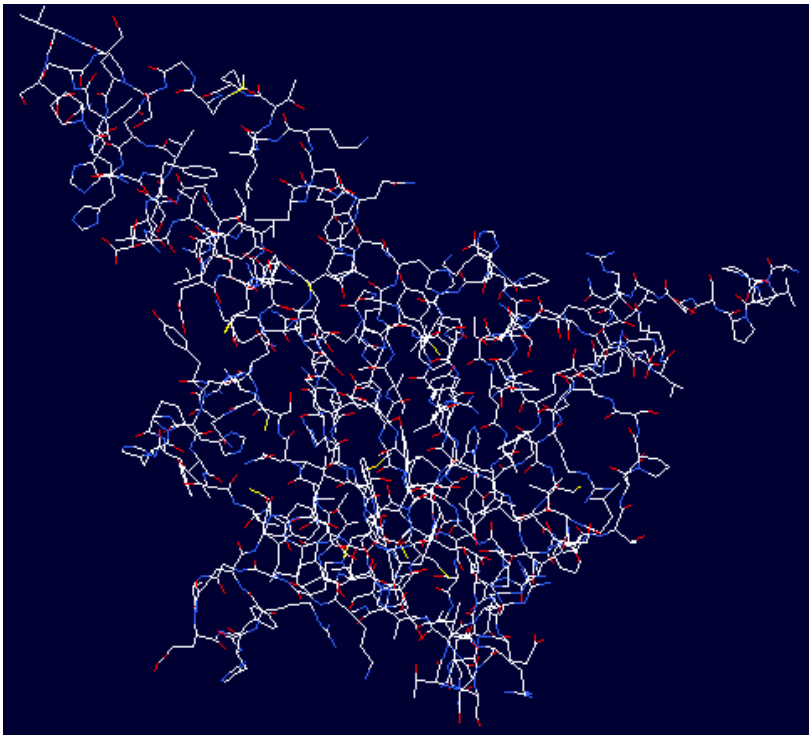
Models generated by Modeller 9v8 are the following



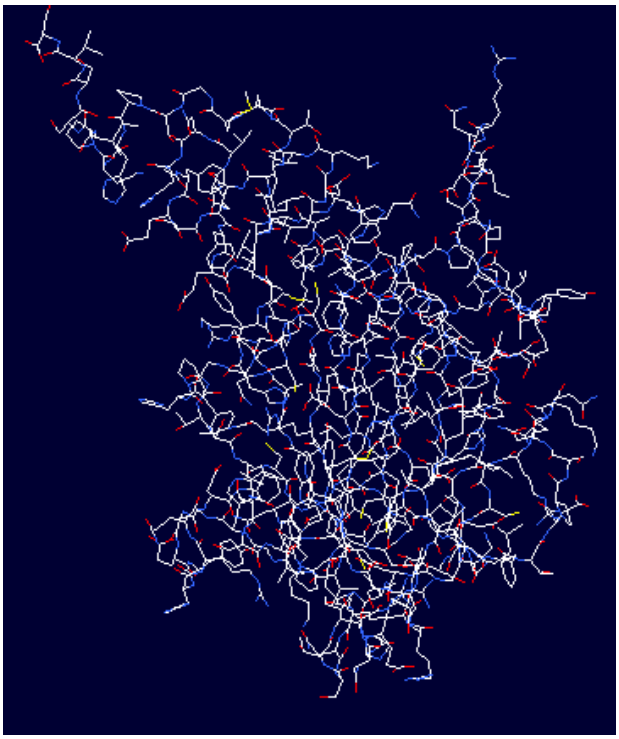
Model-1



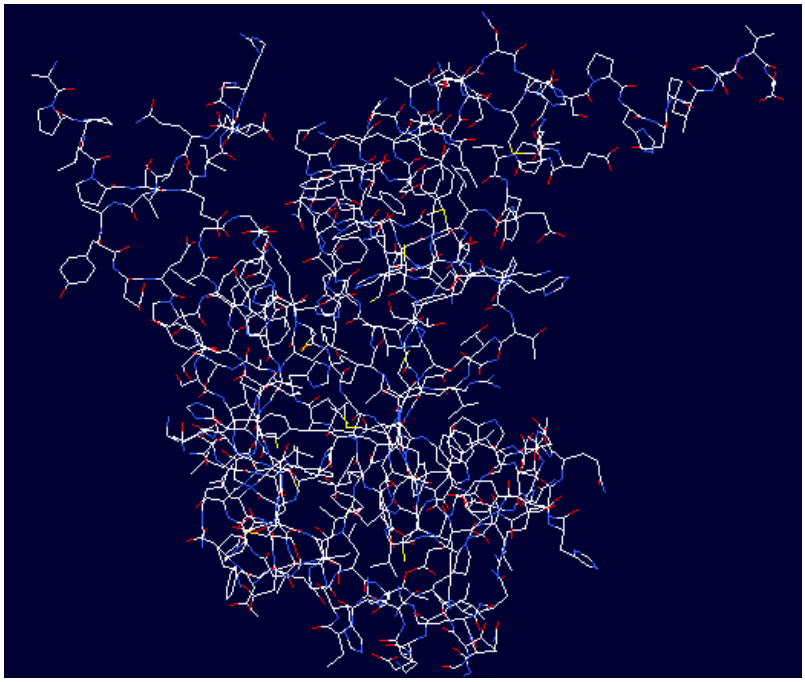
Model-2



Model-3



Model-4



Model-5

Comparing Five models obtained from Modeller 9v8 from SAVES and ProSA

Model	R.Plot	Core	Allow	Generous	Bad contacts	G-Value
Model1	91.6%	6.8%	1.1%	0.5%	4	-0.06
Model2	93.2%	5.3%	1.1%	0.5%	4	-0.06
Model3	90.5%	7.4%	1.6%	0.5%	5	-0.08
Model4	90.5%	7.4%	1.6%	0.5%	3	-0.08
Model5	90.5%	8.4%	0.5%	0.5%	6	-0.05

Model 2 is selected as the best model because of maximum R. plot value and optimum bad contacts and optimum negative overall G-Value. If we compare Model 2 with Model 4 in terms of Bad contacts, Model 4 has minimum bad contacts as compared to all available models but at the same time it has a less R.Plot value. Also if we compare Model 2 with Model 5, model 5 has minimum G-Value among all the models but has a low R.Plot value and Maximum bad contacts.

Homology Modelling of SMAD1 was done using modeller9v8 program. As output of the program we got five models in dot pdb format which were validated using SAVES and ProSA online tools getting their Ramachandran plot value and minimum bad contacts and minimum negative overall G-Value.Out of Five models, Model 2 is selected the best model because of it satisfied the set criteria for the best Model. Studies on SMAD proteins revealed the signal transduction pathway of the members in the TGF-β superfamily. SMAD1, 5 and MADH6 belong to same

structural subgroup and mediate BMP signaling(Hoodless PA, Abdollah S, Stapleton M, Attisano L.et al., 1996).Homology Modelling predicted the models of SMAD1 based on available crystal structure of SMAD Homolog have a very high accuracy and can provide base for interaction of SMAD1, 5 with SMAD4 for studying structure function relation in the biochemical pathway(Hariharan R, Pillai MR., 2006).These further can be used for inhibitor design for multifactorial, Pulmonary arterial hypertension, Also as SMADs have a central and interactive role in number of pathways related to a number of diseases like Cancer, Sporodic PPH(JR Thomson, RD Machado et al, 2003), Osteogenesis(Bing Liu, Ning Mao, 2003), Cardiac Hematopoiesis(Lieve Umans, Marc Tjwa et al, 2007), Apoptosis, Mutation Analysis of Human Tumors(Yang Ke, William P Bennett., et al, 1998) etc.It can be used in structure based ligand design (Joseph McCarty, D, 1999) or structure based drug design(Antel, J., 1999) (Powers, R, Shoichet, B., 2002) (McMartin, C, Bohacek, R, 1997) for an anti cancer compound that if fits in terms of toxicity and bioavailability can be sent for clinical trails.Synthesis of such drug compounds do have a pharmceutical value in the market if it promises to be a potent inhibitor drug compound. SMAD proteins are highly conserved and their evolutionary study can help in elucidating their basic mechanism of action at various levels of gene expression.(Arnold SJ, Maretto S, Robertson EJ et al.). Developmental biology study can be done using early mouse embryos for getting an insight of genetic approach of SMADs.

References

[1] Aubin, J., Davy, A. & Soriano, P. 2004. In vivo convergence of BMP and MAPK signaling pathways: impact of differential SMAD1 phosphorylation on development and homeostasis. Genes Dev. 18, 1482–1494.

[2] A.L.Morris, M.W.MacArthur, E.G.Hutchinson, J.M.Thornton, Stereochemical quality of protein structure coordinates, proteins 12(1992)345.

[3] Bin Y Qin, Benoy M.Checko, Suvana S.Lam, Mark P.de Caestecker, John J.Correia & Kai Lin, Structural basis of SMAD1 Activation by Receptor kinase Phosphorylation, J, Mol.Cell, Vol.8, 1303-1312, Dec, 2001.

[4] Chen, Y. G. & Massague, J. 1999. SMAD1 recognition and activation by the ALK1 group of transforming growth factor-beta family receptors. J. Biol. Chem. 274, 3672–3677.

[5] Hata, A., Lagna, G., Massague, J., & Hammati Brivanlou, A(1998a)SMAD6 inhibits BMP/SMAD1 signalling by specifically competing with SMAD4 tumor suppressor.genes Dev.12, 186-197.

[6] JMGraff, A.Bansal, D.A.Melton, Xenopus Mad proteins transduce distinct subsets of signals for the TGF-β superfamily, Cell 85(1996)479.

[7] M.Li.BWang.Homology modeling and examination of the effect of the D92E mutation on the H5n1 known structural protein ns1 effector domain, J.Mol.Model.13(2007)1237.

[8] Miyazono, K., Maeda, S. & Imamura, T. 2005. BMP receptor signaling: transcriptional targets, regulation of signals, and signaling cross-talk. Cytokine Growth Factor Rev. 16, 251–263.
[9] Massague, J., Seoane, J. & Wotton, D. 2005. SMAD transcription factors. Genes Dev. 19, 2783–2810.
[10] Pera, E. M., Ikeda, A., Eivers, E. & De Robertis, E. M. 2003. Integration of IGF, FGF, and anti-BMP signals via SMAD1 phosphorylation in neural induction. Genes Dev. 17, 3023–3028.
[11] P.ten Dijke, K.Miyazono, C.H.Heldin, Signalling inputs converge on nuclear effectors in TGF-β signaling, Trends Biochem.Sci.25(2000)64.
[12] P.Nagaranjan, Xubao lin et al, Characterisation of DNA-Binding properties of SMAD5.
[13] R.Hariharan, M.R.Pillai, Homology Modelling of the DNA-binding domain of human SMAD5:A Moleculer model for inhibitor design, J.Mol, Graph Model.24(2006)271.
[14] Rui Hao, Lei chan, Jai Wei, Zhi-Xin wang, Structure of drosophila Mad MH2 domain.
[15] Shi, Y., Wang, Y. F., Jayaraman, L., Yang, H., Massague, J. & Pavletich, N. P. 1998. Crystal structure of a SMAD MH1 domain bound to DNA: insights on DNA binding in TGF-beta signaling. Cell 94, 585–594.
[16] T.Imamura, M.Takase, A.Nishihara, E.Oeda, J.Hanal, M.Kawavata, K.Miyazono, SMAD6 inhabits signaling by the TGF-beta superfamily, Nature 389(1997)622.