International Journal of Genetic Engineering and Biotechnology.
ISSN 0974-3073 Volume 1, Number 1 (2010), pp. 35--41

http://www.irphouse.com

Insilico Analysis and Homology Modeling of Salutaridine Reductase of *Papaver somniferum*

Siddiqui M. Asif[1*], Amir Asad[1], Malik S. Anjali[1], Arya Arvind[1], Kapoor Neelesh[1], Kumar Hirdesh[1] and Verma Priyanka[1]

[1]Department of Biotechnology, Meerut Institute of Engineering and Technology, N.H. 58, Delhi-Roorkee Highway, Baghpat Road Bypass Crossing, Meerut-250005, U.P., India
**Corresponding Author E-mail: asifsiddiqui82@gmail.com*

Abstract

Salutaridine reductase (NADPH) is a key enzyme involves in morphine biosynthesis. BLASTP search waer performed to found orthologous proteins across different plant species. Domain analysis of all orthologous proteins was performed using Pfam and it was observed that Adh_short domain was conserved across all proteins. 3D structure of Salutaridine reductase was not available in PDB database, therefore Homology model of Salutaridine reductase was constructed using HOMER. A template structure (PDB ID: 2PFG) was selected from protein databank (PDB) using BLASTP with BLOSUM62 sequence alignment scoring matrix. The protein model was validated using PROCHECK. The model structure was visualized and superimposed with template using Discovery Studio Visualizer (Accelrys) and Swiss PDBviewer.

Keywords: Salutaridine reductase, *Papaver somniferum*, alkaloids, morphine, BLAST, FASTA.

Introduction

Benzylisoquinonline alkaloids are large and diverse groups of natural product containing more than 2500 defined structures found mainly in five plant families, including the *Papaveracea* (Facchini, 2001). Opium Poppy (*Papaver somniferum* L.*)* produces a large number of Benzylisoquinonline alkaloids including morphine and sanguinarine, derived from tyrosine via the branch -point intermediate (S)-reticuline (Sato *et al*., 2001).

Morphine is one of the 40 alkaloids present in opium from Papaver somniferum,

and is one of the strongest known analgesic compounds (Stefano *et al.*, 2000). Endogenous morphine has been characterized in numerous mammalian cells and tissues (Hazum *et al.*, 1981; Gintzler *et al.*, 1978; Goldstein *et al.*, 1985), and its structure is identical to that of morphine from poppy. The most interesting step in the biogenesis of morphine is undoubtedly the diphenolic coupling which forges the linkage between the two aromatic rings of (*R*)-reticuline and leads to salutaridine by the enzyme salutaridine synthase. Salutaridine is converted to salutaridinol by the enzyme salutaridine reductase (SalR), with the reduction of NADPH to $NADP^+$. Barton and Cohen originated the idea that coupling reactions of this type underlay the formation of many C–C and C–O bonds in a variety of alkaloids, including morphine (Barton and Cohen, 1957). Barton demonstrated the *in vitro* conversion of isotopically labeled reticuline to salutaridine by treatment with potassium ferricyanide (Barton *et al*., 1963).

Salutaridine reductase (NADPH), belongs to the family of oxidoreductases, specifically those acting on the CH-OH group of donor with NAD+ or NADP+ as acceptor, is a key enzyme involves in morphine biosynthesis. Thererfore the present investigation is planned to study Salutaridine reductase from *Papaver somniferum* in much detail.

A lot of information about this enzyme is available however no structural and comparative study has been done so far. Therefore an attempt was made to identify orthologous proteins of Salutaridine reductase from *Papaver somniferum* across different plants, comparative analysis of these proteins to find out conserved domains responsible for their functions and to model the 3D structure of Salutaridine reductase of *Papaver somniferum*.

Materials and Methods

Protein sequence of Salutaridine reductase from *Papaver somniferum* was searched through NCBI (http://www.ncbi.nlm.nih.gov/) using Entrez search tool and one record was found (Acc. No. ABC47654). The record was downloaded and the protein sequence was stored in FASTA format, in text file. BLASTP search was performed to find out the orthologous proteins of Salutaridine reductase. All proteins sequences showing more than 50% similarity to the query protein are retrieved and their protein sequence was stored in FASTA format. Pfam (http://pfam.sanger.ac.uk/) was used for domain analysis. Template selection was done using BLASTP for the query sequence against PDB (Protein Data Bank) available at NCBI. Sequence alignment for template 2PFG|A and query was done by using BLAST (bl2seq). The alignment along with template structure was submitted to HOMER (http://protein.bio.unipd.it/homer/) for homology modeling. The validation for predicted structure model was performed by using PROCHECK (Laskowski *et al*., 1996) and energy minimization performed by Verify3D (Bowie *et al*., 1991). The overall stereochemical quality of the protein was assessed by Ramchandran plot analysis (Ramachandran *et al*., 1963). The structures were visualized and superimposed using Discovery Studio Visualizer (Accelrys) and Swiss PDBviewer v 4.0.1.

Results and Discussion

Salutaridine reductase was found conserved among different plants, on analyzing the similar proteins retrieved in BLAST search and the number of protein sequence similar to Salutaridine reductase differ among different plants. The number of proteins similar to Salutaridine reductase in *Arabidopsis thaliana, Capsicum annuum, Medicago truncatula, Papaver bracteatum, Populus trichocarpa, Ricinus communis* and *Vitis vinifera* are 5, 2, 3, 1, 9, 5 and 9 respectively (Table 1). Pfam resuls shows that Adh_short domain was conserved across all plants species, NAD_binding_4 was absent in *Capsicum annuum,* KR and Epimerase were absent in *Papaver bracteatum*. Beside these some other domains were also present in different plants with varying frequency (Table 2).

For homology modeling a template, Crystal Structure of Human Cbr1 in Complex with Bigf2 (PDB ID: 2PFG) was selected. This sequence showed a highest sequence homology of 34%, with atomic resolution of its X ray crystal structure being 1.54 A and R value being 0.115. The alignment obtained between query and 2PFG is shown in Figure 1. The 2PFG structure was used as a template for homology modeling HOMER. The predicted model (Figure 2) was also checked for psi and phi torsion angles using the Ramchandran plots. The molecular visualization program Discovery Studio Visualizer (Accelrys) was used to manipulate the models based on residue interactions, energy minimization and steric hinderance. The model predicted by HOMER was used for further analysis by PROCHECK (Laskowski *et al*., 1996). Ramchandran plot analysis (Figure 3) shows 88.4% of the residues in the most favored region, 8.3 % in the additional allowed, 2.3% in the generously allowed regions and 0.9% in the disallowed region. Modeled structure of Salutridine reductase was visualized (Figure 2) using Discovery Studio Visualizer (Accelrys) and superimposed (Figure 4) with 2PFG using SWISS PDB viewer.

Table 1 : Orthologous proteins of *Papaver somniferum* Salutaridine reductase found in BLASTp search

Organism	Protein	CDS	Gene	Expect	Identities
Arabidopsis thaliana	NP_191681	NM_115986	816953	4e-85	167/303
	NP_179996	NM_127980	825294	2e-83	163/303
	NP_001077951	NM_001084482	816953	3e-83	163/303
	AAF78417	AC009273	--	2e-80	156/302
	NP_563635	NM_100063	839259	2e-80	156/302
Capsicum annuum	sp|B2X050	--	--	2e-84	163/314
	ABM54181	EF025511	--	2e-82	161/314
Medicago truncatula	ACJ84741	BT052079	--	2e-80	157/301
	ACJ84918	BT052256	--	7e-80	158/303
	ABD28440	AC148817	--	1e-78	161/302
Papaver	ABO93462	EF184229		7e-175	298/311
Populus	XP_002301348	XM_002301312	7494064	8e-92	179/298

trichocarpa	XP_002336437	XM_002336398	7457079	4e-87	160/303
	XP_002301343	XM_002301307	7457079	6e-87	164/304
	XP_002301346	XM_002301310	7461754	1e-86,	162/308
	XP_002301344	XM_002301308	7494060	3e-86	162/307
	XP_002336130	XM_002336091	7456996	2e-84	158/301
	XP_002320111	XM_002320075	7464917	2e-81	157/293
	XP_002336414	XM_002336375	7457053	1e-80	159/301
	XP_002320110	XM_002320074	7496945	3e-79	152/293
Ricinus communis	EEF50493	EQ973775	--	1e-87	170/300
	EEF50491	EQ973775	--	5e-82	160/308
	EEF50490	EQ973775	--	2e-81	160/313
	EEF50492	EQ973775	--	1e-80	159/308
	EEF51148	EQ973774	--	5e-48	96/177
Vitis vinifera	CAO39528	CU459257	--	6e-92	177/300
	CAO39529	CU459257	--	5e-91	176/300
	CAO39527	CU459257	--	8e-91	178/306
	CAN77657	AM484013	--	4e-86	170/305
	CAO39530	CU459257	--	1e-81	160/288
	CAO16754	CU459242	--	9e-81	162/306
	CAO16753	CU459242	--	1e-79	157/299
	CAO16751	CU459242	--	3e-67	137/274
	CAO39532	CU459257	--	1e-71	153/264

Table 2 : Conserved domain present across seven different plant species.

Name of domain	**Ricinus communis**	**Vitis venifera**	**Papaver bracteatum**	**Arabidopsis thaliana**	**Populus trichocarpa**	**Capsicum annuum**	**Medicago truncatula**
Adh_short	Y	Y	Y	Y	Y	Y	Y
NAD_binding_4	Y	Y	Y	Y	Y	-	Y
KR	Y	Y	-	Y	Y	Y	Y
Epimerase	Y	Y	-	Y	Y	Y	Y
ADH_zinc_N	-	Y	-	Y	Y	-	-
3Beta_HSD	Y	Y	-	-	Y	-	-
Shikimate_DH	-	-	-	Y	-	-	-
LeuA_dimer	-	-	Y	-	-	-	-
DUF1667	-	Y	-	-	-	-	-
Ponericin	Y	-	-	-	-	-	Y
Polysacc_synt_2	-	-	-	-	Y	-	-
Polysacc_synt_2	-	-	-	Y	-	-	-
PFK	-	-	-	Y	-	-	-
3Beta_HSD	-	-	-	-	Y	-	-
SASP	-	-	-	Y	-	-	-

```
Query  15   AVVTGGNKGIGFEI----CKQLSSNGIMVVLTCRDVTKGHEAVEKLKNSNHENVVFHQLD  70
            A+VTGGNKGIG  I    C+  S +   VVLT RDVT+G  AV++L+ +   +  FHQLD
Sbjct  7    ALVTGGNKGIGLAIVRDLCRLFSGD---VVLTARDVTRGQAAVQQLQ-AEGLSPRFHQLD  62

Query  71   VTDPIATMSSLADFIKTHFGKLDILVNNAGVAGFSVDADRFKAMISDIGEDSEELVKIYE  130
            + D + ++ +L DF++  +G LD+LVNNAG+A    D   F
Sbjct  63   IDD-LQSIRALRDFLRKEYGGLDVLVNNAGIAFKVADPTPFHIQ----------------  105

Query  131  KPEAQELMSETYELAEECLKINYNGVKSVTEVLIPLLQLSDSPRIVNVSS--STGSLKYV  188
                        AE  +K N+ G + V   L+PL++     R+VNVSS  S  +LK
Sbjct  106  --------------AEVTMKTNFFGTRDVXTELLPLIKPQG--RVVNVSSIMSVRALKSC  149

Query  189  SNETALEILGDGDALTEERIDMVVNMLLKDFKENLIETNGWPSFGAAYTTSKACLNAYTR  248
            S E  L+     + +TEE +  ++N  ++D K+ + +  GWPS  +AY  +K  +   +R
Sbjct  150  SPE--LQQKFRSETITEEELVGLMNKFVEDTKKGVHQKEGWPS--SAYGVTKIGVTVLSR  205

Query  249  VLANKIP------KFQVNCVCPGLVKTEMNYGIGNYTAEEGAEHVVRIALFPDD--GPSG  300
            + A K+       K  +N  CPG V+T+M       + EEGAE  V +AL P D  GP G
Sbjct  206  IHARKLSEQRKGDKILLNACCPGWVRTDMAGPKATKSPEEGAETPVYLALLPPDAEGPHG  265

Query  301  FF  302
              F
Sbjct  266  QF  267
```

Figure 1 : Alignment of target sequence (ABC47654) with template (PDB ID: 2PFG)

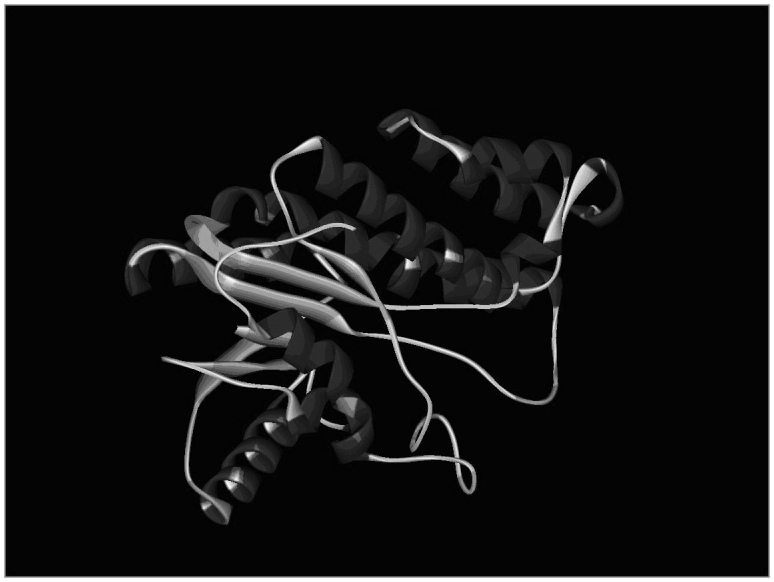

Figure 2 : Modeled structure of Salutridine reductase (ABC47654) visualized by Discovery Studio Visualizer (Accelrys) (Color by Secondary structure).

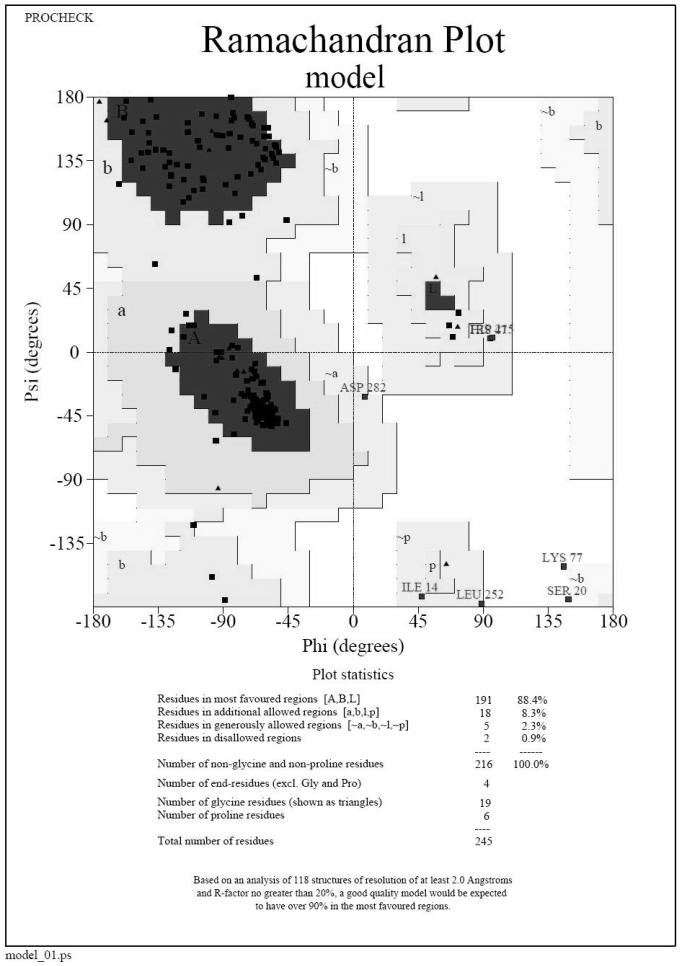

Figure 3 : Ramachandran Plot analysis of predicted model for Salutridine reductase

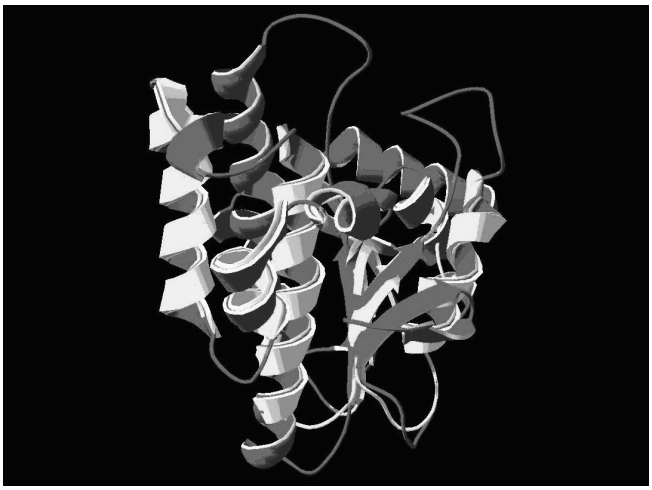

Figure 4 : Superimposed structures of target sequence (ABC47654 yellow) with template sequence (2PFG pink)

References

[1] Barton, D.H.R. and Cohen, T., 1957, "Festschrift Arthur Stoll", *Birkhauser*, Basel, p. 117.

[2] Barton, D.H.R., *et al*., 1963, "Proceedings of the Chemical Society". July, *Proc. Chem. Soc.*, 189-228, 203.

[3] Facchini, P.J., 2001, "Alkaloid biosynthesis in plants: Biochemistry, cell biology, molecular regulation, and metabolic engineering applications". *Annual Review of Plant Physiology and Plant Molecular Biology*, 52, 29-66.

[4] Gintzler, A.R., *et al*., 1978, "A nonpeptide morphine-like compound: immunocytochemical localization in the mouse brain". *Science*; 199:447–448.

[5] Goldstein, A., *et al*., 1985, "Morphine and other opiates from beef brain and adrenal". *Proc Natl Acad Sci* U S A.; 82:5203–5207.

[6] Hazum, E., *et al*., 1981, "Morphine in cow and human milk: could dietary morphine constitute a ligand for specific morphine (mu) receptors?". *Science*; 213:1010–1012.

[7] Sato, F., *et al*., 2001, "Metabolic engineering of plant alkaloid biosynthesis". *Proceedings of the National Academy of Sciences, USA* 98, 367–372.

[8] Stefano, G.B., *et al*., 2000, "Endogenous morphine". *Trends Neurosci*.; 23:436–442.

[9] Laskowski RA, *et al*., 1996, "AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR". *J Biomol. NMR.*, 8: 477.

[10] Bowie, J.U., *et al*., 1991, "A method to identify protein sequences that fold into a known three-dimensional structure". *Science*, 253: 164

[11] Ramachandran, G.N., *et al*., 1963, "Stereochemistry of polypeptide chain configurations". *J. Mol. Biol.*, 7: 95.