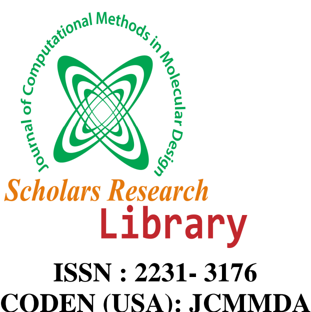
*Available online at* www.scholarsresearchlibrary.com

**Scholars Research Library**

***J. Comput. Method. Mol. Design*, 2011, 1 (4):22-27**
*(http://scholarsresearchlibrary.com/archive.html)*




**ISSN : 2231- 3176**
**CODEN (USA): JCMMDA**

# *In silico* Analysis and 3D Modelling of SORD Protein in Diabetic Retinopathy

**Vidhya VG[1], Anusha Bhaskar[2*], Parthiban Purushothaman[3]**

[1]*Department of Biotechnology, Faculty of Science and Humanities, SRM University, Kattankulathur, Chennai*
[2]*Department of Biotechnology, PRIST University, Vallam, Thanjavur*
[3]*Department of Ophthalmology, KAPV Govt. Medical College, Trichy*

---

## ABSTRACT

*Diabetic retinopathy is the leading cause of blindness among patients with diabetes mellitus. Many genes have been identified for its pathogenesis. SORD is one of the candidate gene involved in the development of diabetic retinopathy. An in silico technique was initiated to characterize the properties and structure of the protein. The SORD protein analyzed in the study showed that this is a stable protein and belong to the medium chain dehydrogenase / reductase protein family (MDR). The secondary structure prediction of the protein revealed that the presence of maximum number of random coils as its secondary structure elements. The 3D structure was modelled using Swiss model workspace and the structure was validated. The present study gave an outlook on SORD protein and further research was carried out in preventing the pathogenesis of disease.*

**Keywords:** Diabetic retinopathy, Sorbitol Dehydrogenase, Swiss model.

---

## INTRODUCTION

Diabetes and visual disability due to diabetic retinopathy is a serious health and socio-economic problem in India. Diabetic retinopathy is one of the most common microvascular complications of diabetes affecting 80 % of patients over 20 years duration of diabetes. Despite remarkable advances in the diagnosis and treatment of diabetic retinopathy and its associated complications, it remains the leading cause of blindness among working age individuals in developed countries. In developing countries like India also, it is one of the major causes of blindness in view of the prevailing diabetes epidemic [1].

Multiple biochemical pathways have been proposed to explain the pathogenesis of diabetic retinopathy all starting initially from hyperglycaemia. These may include increased polyol pathway; increased advanced glycation end products (AGE) formation; activation of protein kinase C (PKC) and increased hexosamine pathway flux.

Diabetic retinopathy progresses from mild nonproliferative abnormalities, characterized by increased vascular permeability to moderate and severe non proliferative diabetic retinopathy (NPDR) characterized by vascular closure, to proliferative diabetic retinopathy, characterized by the growth of blood vessels on the retina and posterior surface of the vitreous. Macular edema, characterized by retinal thickening from leaky blood vessels can develop at all stages of retinopathy. Pregnancy, puberty, blood glucose control, hypertension and cataract surgery can accelerate these changes [2].

The number of people worldwide at risk of developing vision loss from diabetes is predicted to double over the next 30 years, so it is imperative to develop better means to identify, prevent and treat retinopathy in its earlier stages rather than wait for the onset of vision threatening lesions.

Over the past few years, progress has been made in identifying some of the genes associated with diabetic retinopathy. Sorbitol dehydrogenase (SORD) is one of the candidate gene found to be involved in causing diabetic retinopathy.

Sorbitol dehydrogenase is a zinc containing enzyme that catalyzes the oxidation of sorbitol to fructose with $NAD^+$ as cofactor. It is a member of the multigene family that includes alcohol dehydrogenases [3], ζ crystalline [4] and threonine dehydrogenase [5]. SORD is involved in the metabolism of different polyols and is believed to cooperate with aldose reductase in osmotic regulation [6].The regulation through the polyol pathway is thought to affect the accumulation of sorbitol that is associated with diabetes mellitus and its complications, such as neuropathy [7], retinopathy [8] and cataracts [9]. SORD is also thought to affect the cytosolic ratio of $NADH/NAD^+$ impacting on several metabolic pathways [10].

The nucleotide sequence of the mRNA covers 2471 bp including an open reading frame that yields a protein of 356 amino acid residues. The gene structure of SORD spans approximately 30 kb divided into 9 exons and 8 introns. The gene was localized to chromosome 15q21.1 by *in situ* hybridization [11].

## MATERIALS AND METHODS

### Primary structure prediction

For physio-chemical characterization, theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient [12], instability index [13], aliphatic index [14] and grand average hydropathy (GRAVY) [15] were computed using the Expasy ProtParam server (http://expasy.org/cgi-bin/protpraram).

### Secondary structure prediction

Secondary structure of this protein was predicted using the FASTA sequences of 56 k protease, and predicted using GOR IV and SOPMA.

### Protein functional sites

Interproscan and Fingerprint scan were the tools used to predict the signatures and the motif regions in the sequence.

### Homology modeling

The protein sequence was subjected for comparative homology modeling via Swiss model (**[16]** Arnold K *et al. Bioinformatics* 2006 **22**: 195 [PMID: 16301204 1250 [PMID: 12217917]) to generate putative 3D model. The Swiss model performs the sequence alignments and searches for the putative template protein for

generating the 3D model and the ESyPred incorporated with Modeller (version 6v2) generated the putative 3D model.

**Validation**

The validation of the modelled structure was carried out using Protein Structure Validation Suite (PSVS) tool. Structural analysis was performed and the 3-d coordinate file was visualized and analyzed in Rasmol.

## RESULTS AND DISCUSSION

The similarity search for the sequence was carried out with the help of BLAST tool. The results indicated 100 % similarity to L-iditol – 2 dehydrogenase [*Homo sapiens*]
[Table 1].

**Primary structure prediction**

In this study primary structure of sorbitol dehydrogenase were predicted using Expasy's ProtParam server (http://expasy.org/cgi-bin/protparam) using the gene sequence and the results are shown in Table 1. Results showed that sorbitol dehydrogenase had 357 amino acid residues and the estimated molecular weight 38324.5. The maximum number of amino acid present in the sequence was found to be glycine (9.8 %) and the least was that of glutamine (1.7 %). The total number of positively charged residues (Asp + Glu) was 36 and the total number of negatively charged residues (Arg + Lys) was 39.The calculated isoelectric point (pI) is useful for at pI the solubility is least and the mobility in an electric field is zero. Isoelectric point (pI) is the pH at which the surface of protein is covered with charge but net charge of protein is zero. The calculated isoelectric point (pI) was computed to be 8.23. The computed value is more than 7 indicate that the protein is basic. The very high aliphatic index (93.39) indicates that this protein is stable for a wide range of temperature range. While the instability index (30.53) provides the estimate of the stability of protein in a test tube. The Grand Average Hydropathicity (GRAVY) value is low 0.012, indicates better interaction of the protein with water.

**Secondary structure prediction**

The secondary structure is composed of alpha helix and beta sheets and the secondary structure is predicted using GOR IV and SOPMA. Table 2 presents the comparative analysis of GOR IV and SOPMA from which it is clear that random coil is predominantly present when the structure was predicted both by SOPMA and GOR, followed by extended strand and alpha helix. The secondary structure prediction was done and random coil was found to be frequent (53.78%) followed by Extended strand (25.77%) and alpha helix was found to be least frequent (20.45%). This is graphically represented in Figure 2.

The domain search was done by Conserved domain search on the Blast site and it showed two domains – the medium chain reductase/dehydrogenases (MDR) / zinc dependent alcohol dehydrogenase like family and L- idonate 5 dehydrogenase family [Figure 1].

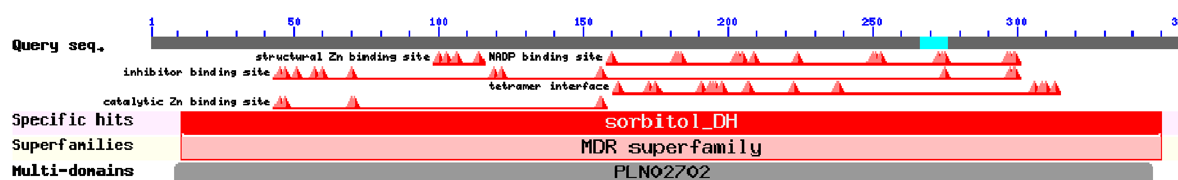
The fingerprint scan of the sequence showed ten fingerprints having two motifs for each fingerprint in the sequence [Table 3].

The tertiary structure was modelled by Swiss model workspace by using the templates from PDBSum [Figure 3] the modelled structure showed 256 H bonds, 15 helices, 27 strands and 39 turns. The modelled structure was validated by PSVS and Ramachandran plot was plotted [Figure 4].

**Table 1. BLAST result of the SORD protein**

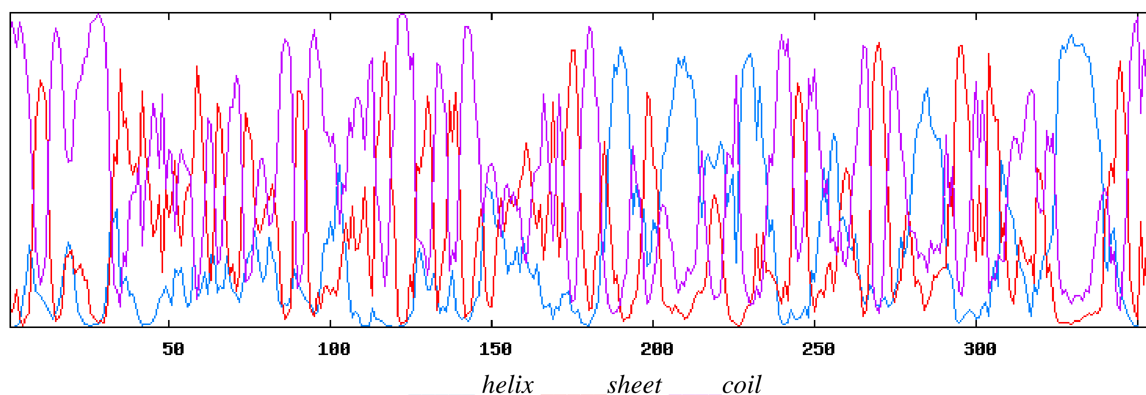
| Accession | Description | Score | E value |
|---|---|---|---|
| NP_003095.2 | sorbitol dehydrogenase [Homo sapiens] | 733 | 0.0 |
| AAA66064.1 | sorbitol dehydrogenase [Homo sapiens] | 730 | 0.0 |
| AAA80565.1 | L-iditol-2 dehydrogenase [Homo sapiens] | 727 | 0.0 |
| 3QE3_A | Chain A, Sheep Liver Sorbitol Dehydrogenase | 651 | 0.0 |

**Figure 1 : Graphical representation of conserved domain in SORD protein**




The analysis of SORD protein showed sequence similarity mostly to L-iditol – 2 dehydrogenase [*Homo sapiens*]. The two domains were identified by conserved domain search. Sorbitol dehydrogenase is a zinc containing enzyme, involved in the metabolism of different polyols and is believed to be cooperate with aldose reductase in osmotic regulation (Burg, 1988). The regulation through the polyol pathway is thought to affect the accumulation of sorbitol that is associated with diabetes mellitus and its complications such as retinopathy (Robison et al., 1983). The modelled structure revealed that 86.1 % are allowed regions. Further research involving development of appropriate strategies for studying this protein could be of significance in preventing diabetic retinopathy.

**Figure 2: Graphical representation of Secondary elements in SORD protein**




*helix* *sheet* *coil*

**Table 2: Secondary structure of cysteine protease by SOPMA and GOR**

| Secondary structure | SOPMA | GOR |
|---|---|---|
| Alpha helix | 30.25% | 20.45% |
| 310 helix | 0.00% | 0.00% |
| Pi helix | 0.00% | 0.00% |
| Beta bridge | 0.00% | 0.00% |
| Extended strand | 21.57% | 25.77% |
| Beta turn | 8.96% | 0.00% |
| Bend region | 0.00% | 0.00% |
| Random coil | 39.22% | 53.78% |
| Ambiguous states | 0.00% | 0.00% |
| Other states | 0.00% | 0.00% |
| Sequence length | 110 | 110 |

**Figure 3: Three dimensional structure of SORD protein**

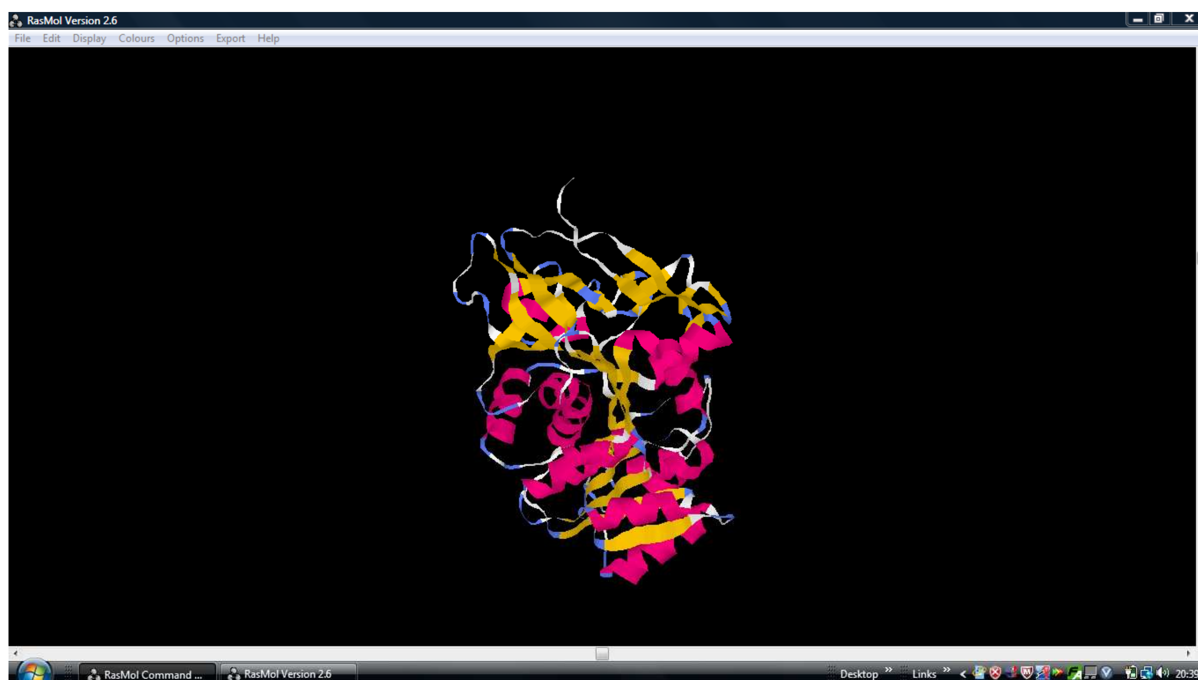



**Figure 4: Graphical representation of Ramachandran plot by PSVS**

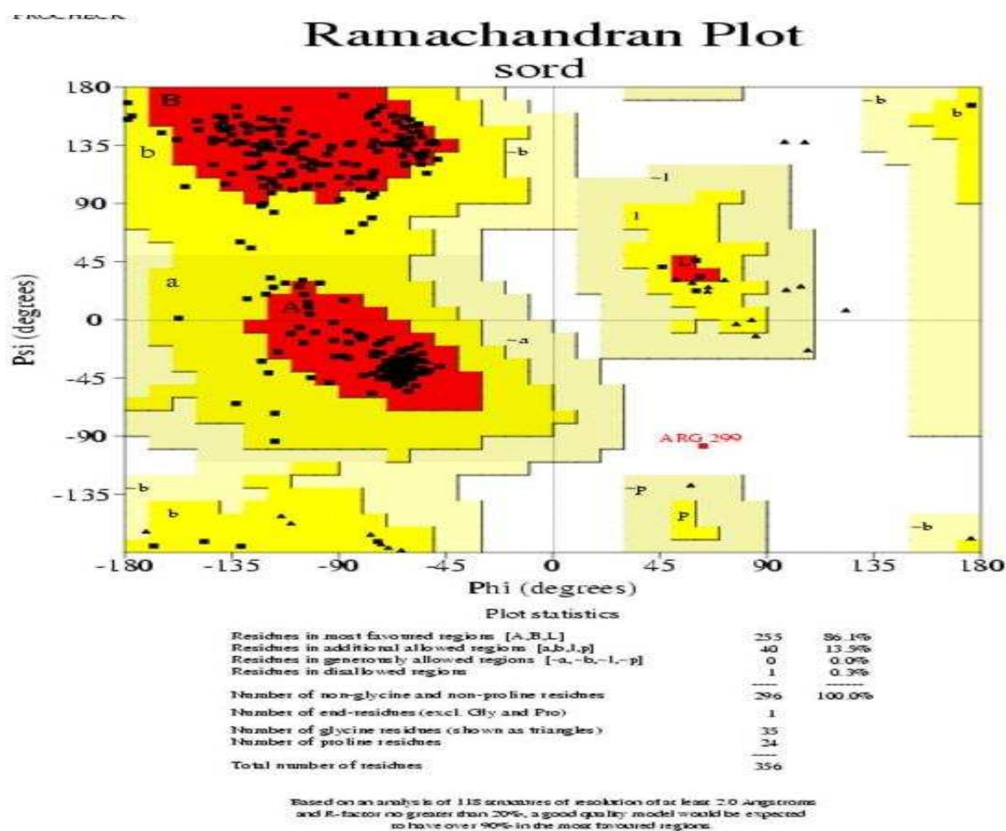


**Table3: FingerPRINT scan result of SORD protein**

| FingerPRINT | No.of Motifs |
|---|---|
| FADPNR | 2 |
| DNAREPAIRADA | 2 |
| PNDRDTASEII | 2 |
| DISEASERSIST | 2 |
| HGRDTASE | 2 |
| RAD23PROTEIN | 2 |
| KIR22CHANNEL | 2 |
| PAXILLIN | 2 |
| KIR12CHANNEL | 2 |
| PHEHYDRXLASE | 2 |

## CONCLUSION

In the present study the sequence and structure analysis of SORD protein was done by various tools and softwares. Based on the findings it could be concluded that further characterization of human SORD gene is novel and will be important for evaluating how the regulation of this gene is related in the complications connected to diabetes mellitus.

## REFERENCES

[1]. M.Balasubramanyam, M.Rema and C.Premanand, *Current Science*, 83, (**2002**), 1506 – 1511.
[2]. DS. Fong, L Aiello, TW Gardner, GL King, G Blankenship, JD Cavallerano, FL Ferris, R Klein, *Diabetes Care*, 26 (**2003**), 226 - 229.
[3]. H. Jornvall, H. Von bahr-Lindstorm. J Jefferey, *Eur.J.Biochem*. 140 (**1989**), 17-23.
[4]. T Borras, B. Persson, H. Jornvall, *Biochemistry* 28 (**1989**), 6133-6139.
[5]. BD. Aronson, RL.Somerville, BR.Epperly, EE. Dekker, *J.Biol.Chem* 264 (**1989**), 5226 – 5232.
[6]. MB. Burg, *Kidney Int*. 33 (**1988**), 635 – 641.
[7]. KH. Gabbay, *N.Engl. J.Med* 288 (**1973**), 831 – 836.
[8]. WG.Robison, PF. Kador Jr, JH. Kinoshita, *Science* 221 (**1983**), 1177 – 1179.
[9]. JH. Kinoshita, *Invest. Ophthalmol* 13 (**1974**), 713 – 724.
[10]. JR.Williamson, K. Chang, M. Frangos, KS. Hasan, Y. Ido, T. Kawamura, JR. Nyengaard, M.Van den Enden, C. Kilo, RG. Tilton, *Diabetes* 42 (**1993**), 801 – 813.
[11]. I. Takeshi, CP. Nicholaas, DB. Zimonjic, C. Karlsson, H. Jan-Olov, G. Vaca, IR. Rodrigeuz, D. Carper, *Genomics* 26 (**1995**), 55-62.
[12]. SC. Gill, PH.Von Hippel, *Anal. Biochem*, 182 (**1989**), 319.
[13]. K.Guruprasad, BVP. Reddy, MW. Pandit, *Prot. Eng* 4 (**1990**), 155 – 161.
[14]. AJ. Ikai, *J. Biochem*, 88 (**1980**), 1895-1898.
[15]. J. Kyte, RF. Doolottle, *J. Mo Biol* 157 (**1982**), 105- 132.