Available online at **www.joac.info**

ISSN: 2278-1862

Journal of Applicable Chemistry

2018, 7 (5): 1319-1329

(International Peer Reviewed Journal)

I-IR Interaction Exploration and Computational Investigation of Evolution of IRS1

Meenu Sharma[1] and C. S.Kapoor[2]*

1. Department of Biotechnology, (PIBS), Pacific University of Higher Education and Research, Udaipur-313003, Rajasthan, **INDIA**
2. Department of Environmental Sciences, (PIBS), Pacific University of Higher Education and Research, Udaipur-313003, Rajasthan, **INDIA**

Email:drcskapoor@yahoo.co.in

Accepted on 16th September, 2018

ABSTRACT

Insulin receptor substrate (IRS) proteins play acrucial role in insulin signalling by docking and activating insulin receptors. It is a well-known fact that IRS family emerged in vertebrates. To have a better understanding of evolution of IRS1 and its connection with the regulation of metabolism controlled by insulin, it's very important have the know-how of closely related species. The key molecular interactions playing vital role in insulin and insulin-receptor interaction are studied along with phylogenetic conservation and evolution of IRS1.Numerous methods including statistical analysis were employed to find its probability of occurrence. The studies provide a rapid way to evolution and make confident predictions.

Graphical Abstract

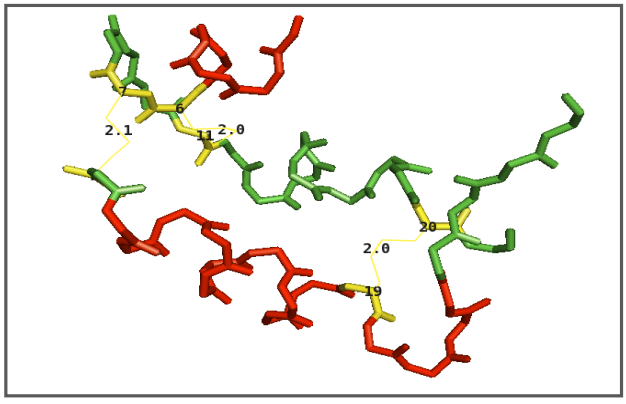

Image showing insulin three-dimensional structure represented in sticks with inter-chain and intra-chain disulphide bonds labelled with bond lengths

Keywords: Insulin receptor substrate-1 protein, Interaction visualization, Phylogeny, Evolution, Vertebrate.

INTRODUCTION

Insulin plays a vital role in body's glucose metabolism. The discovery of insulin in 1922 signifies a landmark in the field of molecular endocrinology and medicine [1]. This protein hormone has been widely used to deal with the condition called diabetes. The fact drew enormous interest towards its signalling mechanisms and involved cascading proteins for better understanding of underlying mechanism at cellular level.

In the process of insulin signalling, binding of insulin to the alpha subunit of the insulin receptor is responsible for the activation of tyrosine kinase of the beta subunit [2, 3]. The activation further initiates the auto-phosphorylation of other tyrosine residues present in the beta subunit [4]. They are renowned by insulin receptor substrate family (IRS) members or the phosphor-tyrosine-binding domains of adaptor proteins [5]. The IRS protein cascades are the common elements in the peripheral response and signalling pathway since these protein cascades are recognized by others in the signalling pathway for further downstream action. Subsequently it results in the uptake of glucose from the blood stream and stored as glycogen in tissues [6]. Hence it is very clear that the insulin receptor substrate family serves as a key role not only in signalling but also in growth and function of pancreatic beta-cell [7, 8]. It may also be responsible to cause peripheral insulin resistance and hyperinsulinemia, in case it is unable to bind to IRS cascade. [9].

From the IRS-family only six members have been isolated and they are IRS1, IRS2, IRS3, IRS4, IRS5 and IRS6, respectively. Among IRS members, the insulin receptor substrate 1 or IRS1 is well-known to be associated with the regulation of blood glucose level. For instance, an up regulation of gluconeogenic enzymes such as glucose 6 phosphatase (G6Pase) and phosphoenol pyruvate carboxy kinase (PEPCK) can be achieved by liver-specific knockdown of IRS1. Increased blood glucose level can also be achieved by decline of glucokinase (GK) expression level that causes reduction of IRS1 level [10, 11]. However, studies are needed to understand the relationship among IRS family members [7, 12]. Studies reveal that some of the family members such as IRS1 and IRS2 are widely distributed in the human body while others have restricted distribution for example IRS3 is found in adiposities and brain, IRS4 in cell lines and embryonic tissues, IRS5 in liver and kidney, while IRS6 in skeletal muscles [7, 13].

MATERIALS AND METHODS

Data Collection: We have collected the sequence data on IRS1from the National Center for Biotechnology Information database (NCBI) [14]. The present study also includes the interaction study between insulin and insulin receptor and for the purpose amino acid sequence of insulin peptide was obtained from PubMed against ID-P01308.1 for organism type-*Homo sapiens*. The functional protein sequences (in FASTA format) were gathered from the NCBI database and cross referred with expose for annotation and are further analyzed using various online and offline bioinformatics tools.

Interaction study between insulin and insulin receptor: Pymol software is used to study three-dimensional visualization of insulin and to explore important interactions between insulin and insulin receptor [15]. Amino acid level images are captured for detailed analysis.

Insilco alignments of IRS1 sequence using align bl2seq: NCBI protein blast is used to carry out insilico sequence alignment of IRS1 using blast bl2seq. Blast parameters thus obtained are recorded for Score, Query cover, E-value, Identities, Positives and Gaps.

Multiple sequences alignment and generation of scores: The sequences are subjected to NCBI-Cobalt and MEGA-crustal for multiple sequence alignment [16] with gap open penalty of 10 and gap extension penalty of 0.1 for pair wise alignment and with gap open penalty of 10 and gap extension penalty of 0.2 for multiple alignment with PAM matrix and 4 gap separation distance.

Phylogenetic Tree Construction: On the basis of the results obtained in the last step, we constructed the phylogenetic tree. The evolutionary history was inferred using the UPGMA (unweighted pair-group method using arithmetic averages) method. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed as number of amino acid substitutions per site using the Poisson correction method and all positions containing gaps and missing data were eliminated

Analysis using MEGA7: Composition distance, pairwise distance and gamma parameters were obtained for the resulting 20 blast hits for each receptor **[17]**. The software calculates the composition distance for a given pair of sequences as a measure of the difference in amino acid composition. It is one half the sum of squared difference in counts of bases (or residues). MEGA 4 computes and presents the Composition Distance per site,

The estimated value of the shape parameter for the discrete Gamma Distribution was calculated by Maximum Likelihood method and substitution pattern and rates were estimated under the Jones-Taylor-Thornton model (+G). Then to model evolutionary rate differences among five different categories we used discrete Gamma distribution [+G]. For estimating maximum Log values, a tree topology was computed using MEGA7.

Tajima's Neutrality Test was carried out. All positions containing gaps and missing data were eliminated and number of sequences, total number of sites, Number of segregating sites and nucleotide diversity was calculated using divergence ratio.

RESULTS AND DISCUSSION

Insulin sequence under study is P01308 and consists of 110 amino acids stretch. This sequence is for *Homo sapiens* and can be studied in four parts:

- Amino acid “M”, it is the very first amino acid and is the initiation amino acid in the sequence.
- Amino acids “MALWMRLLPLLALLALWGPDPAAA” these 24 amino acids are a part of signal peptide.
- The signal peptide sequence is then followed by a stretch of 30 amino acids “FVNQHLCGSHLVEALYLVCGERGFFYTPKT” that is known as Chain B. This is a part of mature insulin sequence.
- Chain C consists of “EAEDLQVGQVELGGGPGAGLQPLALEGSLQ” and gets cleaved during process of insulin maturation.
- Amino acid sequence “GIVEQCCTSICSLYQLENYCN” is Chain A of mature amino acid.

Insulin has very unique disulphide bonds and is believed that vital structural information is bestowed in the chains that are present in mature insulin, i.e. chain-A and chain-B [**18-20**]. The chain-C peptides are not present in mature insulin but play vital role in proper positioning and interaction of cystine residues of chain-A and chain-B for the formation of disulphide bonds [**21**].

Insulin Structure Visualization: Insulin in itself is very important and research studies have established a well-defined insulin structure. The protein structure of insulin is visualized in pymol-3D viewer for a better in-depth knowledge of characteristics of its three -dimensional structure.

In the figure 1, in this visualization the insulin chains are shown linked by two inter-chain disulphide bonds (chainA7-chainB7 and chainA20-chainB19) and have an intra-chain disulphide bond present within the chain A (chainA6-chainA11). All these disulphide bonds are found in the range of 2 Armstrong. These two segments have well established two inter-chain disulphide bonds and one intra-chain disulphide bond in chain-A. Its unique binding to site-1 is followed by binding of site-2 to insulin binding domain of insulin receptor. The major residues that play vital role in the binding events are present in the exposed area of the binding cavity.

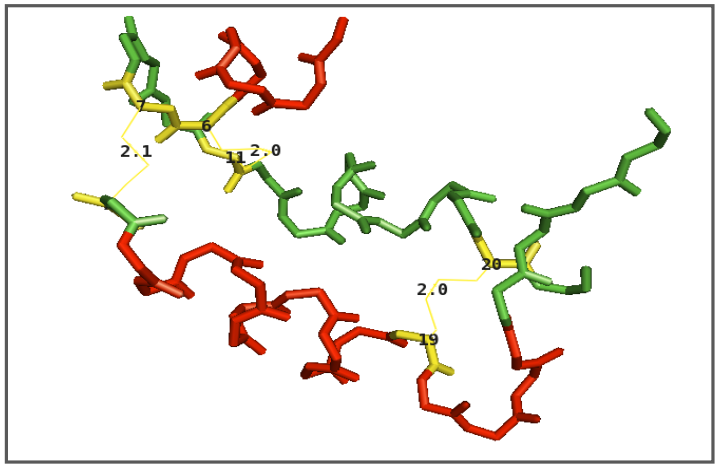

Figure 1. Image showing insulin three-dimensional structure represented in sticks with inter-chain and intra-chain disulphide bonds labelled with bond lengths.

Visualization of interactions between Insulin and insulin-receptor: In this section we consider the important interactions that play an indispensable role in the formation of insulin and insulin-receptor complex (INS-IR). Insulin binds to the ligand binding domain of α-chains of the insulin-receptor echo-domain through various interactions figure 2 shows the position of insulin and insulin-receptor as a singlet for better understanding of binding site location and domains.

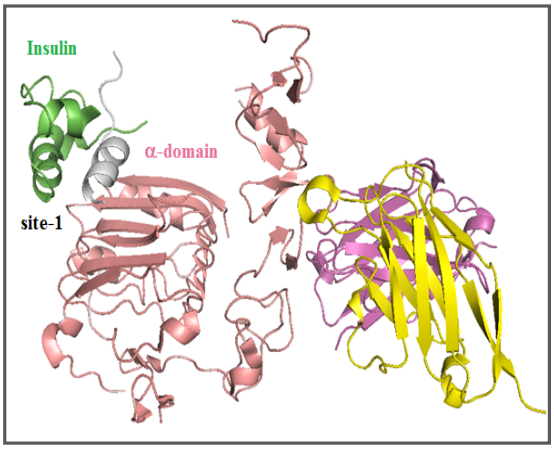

Figure 2. Insulin Receptor visualization: Image showing various segments of Insulin Receptor along with Insulin (green) bound to site-1 domain of insulin receptor.

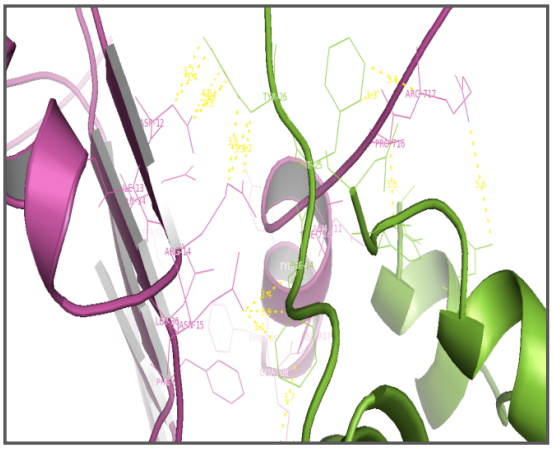

Figure 3. Interaction study between Insulin (INS) and insulin receptor (IR): Pymol image showing interactions among asn-12,arg-14 (IR) and tyr26 (INS); asn-15 (IR) and phe-24 (INS) and Pro-716, arg-717 (IR) and phe-25 (INS) with IR presented in purple and INS in green color.

9+8 when moved to greater depth of molecular interactions and visualizations it is observed that residues of Insulin-receptor namely, asn-12, arg-14, asn-15, pro-716, and arg-717 exhibit significant interaction with insulin residues namely, tyr-26, phe-24, phe-25 (Figure 3).

In addition to these interactions some other residues (Figure 4) phe-714 of insulin-receptor and asn-18 are also involved in σ and π interactions that play a vital role in the binding process at site-1 followed by binding of insulin to site-2.

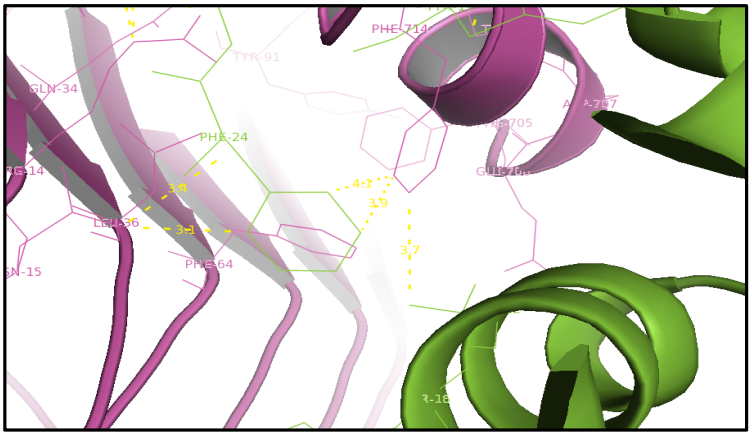

Figure 4. Interaction study between Insulin (INS) and insulin receptor (IR): Pymol image showing interactions among phe714 (IR) and phe24 and asn18(INS) with IR presented in purple and INS in green color.

The side chain interaction of binding site residues plays very important role in stabilizing the insulin and insulin-receptor complex. Various crystallographic and NMR studies have been done to elucidate the structure at different stages of the event. The complex formation which takes place by dimerization of insulin receptor is followed by the activation of intracellular cascade that in-turn is responsible for the translocation of GLUT transporters and causes glucose uptake from the blood. The event is so vital that any imbalance results in diabetic condition that comprises a major section of human population.

Phylogenetic analysis: We have developed phylogram, cladogram and binary tree, which is equivalent to cladogram and our findings show significant relationships among the various IRS1 family members (Figure 5). The phylogenetic analysis of IRS1 was depicted using amino acid sequences of individual proteins. In the phylogenetic tree, the distance of branches was developed from the likelihood ratio mapping the evolutionary relationships.

The fact that insulin plays the vital role in glucose regulation and metabolism in the body provides a strong motivation to find similar proteins sequences in other organisms that exhibit such a similar property. We further moved forward and conducted sequence alignment of insulin protein sequence using protein-protein blast against organism types. The matrix used for the purpose is BLOSUM62 with gap penalty of 11.1 against non-redundant protein sequences. Table 1 summarizes the important parameters considered for sending the blast query for searching hits. The blast results are further filtered to obtain more authentic results result with is obtained by further filtering of the results. In the process ignored the predicted and hypothetical hits and finally ended up with 20 organism hits with a very high conservation rate. These filtered results are shown in table 2 that lists the Score, Query cover, E-value, Identities, Positives and Gaps of the matched hits.

Table 1. Search parameters that are used in the blast query

Algorithm	blastp (protein-protein BLAST)
Matrix used	BLOSUM62
Hit list size	100
Gap penalty	11.1
Expect threshold	10
Word size	6
Window size	40
Genetic code	1
Databases	Non-redundant protein sequences (nr)

Table 2. In silico alignment of IRS1 sequence using align bl2seq

S.No	Organism / Accession No	Score	Query cover	E-value	Identities	Positives	Gaps
1	gi|5031805|ref|NP_005535.1|;gi|547738|sp|P35568.1|IRS1_HUMAN;	2522	100	0	100	100	0
2	gi|246466|gb|AAB21608.1|	2511	100	0	99.759	99.76	1
3	gi|817353823|ref|XP_012301742.1|	2048	100	0	93.344	95.35	5
4	gi|829972321|ref|XP_012604932.1|	2006	100	0	93.087	94.61	4
5	gi|1191840537|ref|XP_020930260.1|	2005	100	0	92.271	94.44	1
6	gi|347300209|ref|NP_001231418.1|	2001	100	0	92.11	94.36	1
7	gi|640797999|ref|XP_008055159.1|	1995	100	0	93.018	94.54	2
8	gi|444730884|gb|ELW71257.1|	1985	100	0	91.186	93.19	4
9	gi|513013338|ref|XP_004868032.1|;gi|513013340|ref|XP_004868033.1|	1978	100	0	91.465	93.88	4
10	gi|1187583352|ref|XP_020744382.1	1976	100	0	90.024	92.44	3
11	gi|1040211151|gb|OBS69920.1|	1975	100	0	89.237	91.49	8
12	gi|1685085|gb|AAC60050.1|	1972	100	0	90.032	91.72	6
13	gi|524965330|ref|XP_005082573.1|	1965	100	0	88.987	91.72	7
14	gi|532081555|ref|XP_005326528.1|	1955	100	0	91.559	93.73	5
15	gi|537223161|gb|ERE82762.1|	1938	100	0	88.532	91.26	8
16	gi|6981106|ref|NP_037101.1|;gi|547740|sp|P35570.1|IRS1_RAT;gi|56504|emb|CAA41264.1|;gi|227862|prf||1712323A	1932	100	0	88.542	91.43	9
17	gi|1211401704|ref|XP_021490768.1|	1922	100	0	87.811	90.46	10
18	gi|1195545336|ref|XP_021053120.1|	1919	100	0	88.424	91.24	10
19	gi|29825829|ref|NP_034700.2|;gi|297914|emb|CAA49378.1|	1910	100	0	88.193	91	10
20	gi|1212175265|ref|XP_021538627.1	1910	100	0	90.909	93.08	3

Our results also goes parallel with the fact that IRS1 was the first insulin receptor docking protein identified in mammalian cells [22]. This conservation is also studied for evolutionary prospects using phylogeny reconstruction in MEGA7. A more comprehensive result is obtained from the phylogenetic analysis that provides a better insight into the evolutionary history of such proteins. During the process mainly 100 bootstrap replicated were used for the analysis using poisson substitution model with nearest-neighbour heuristic approach. The analysis results of maximum likely hood statistical approach are shown in figure 5. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model. The tree with the highest log likelihood

(-2166.5245) is shown. For the heuristic search initial tree(s) were obtained automatically by applying BioNJ and Neighbour-join algorithms to a matrix of pairwise distances that was estimated using a JTT model, and later on selecting the topology that possess superior log likelihood value. The analysis involved 20 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 86 positions in the final dataset when evolutionary analyses were conducted in MEGA.

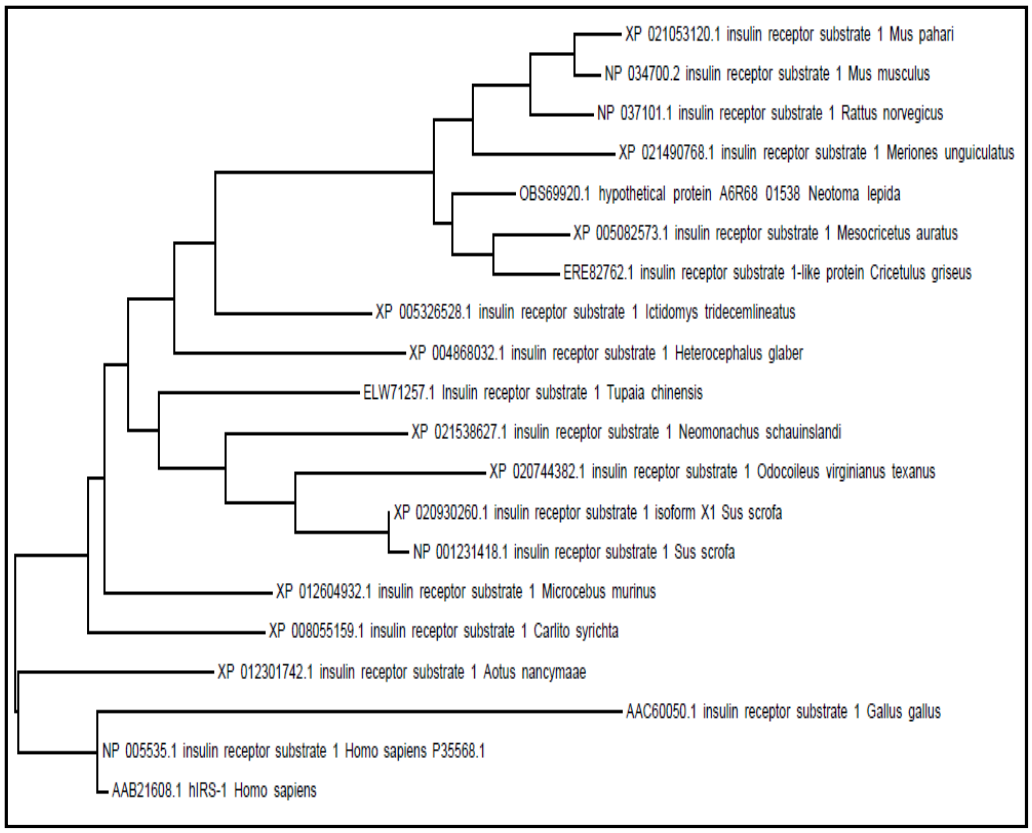

Figure 5. Phylogenetic tree obtained during molecular phylogenetic analysis by Maximum Likelihood method for IRS1

In the present work first of all comparative sequence similarity study is conducted among vertebrates to find the conservation of insulin hormone among them. Irs1 is found in many classical targets of insulin action and is important for insulin sensitivity and embryonic and postnatal body growth [23]. Here we found that most of the signal peptide, chain-A and chain-B amino acids are conserved among *Homo sapiens, Mus pahari, Rattus norvegicus, Merionesunguiculatus, Neotoma lepida, Mesocricetus auratus, Cricetulusgnseus, Ictidomystridecem lineatus, Heterocephalus glaber, Tupaiachinensis, Neomonachusschauinslandi, Odocoileus virginianus texanus, Sus scrofa, Microcebus murinus, Carlito syrichta, AotusNancymaae, Gallus gallus*, etc. The work provides the reason behind using animal insulin as a substitute for insulin for diabetic human population.

Pie chart in figure 6 shows the composition distance of homo sapiens obtained against various organism hits of IRS1 sequence. Maximum similarity of 10% is obtained in case of Aotusnancymaae and Heterocephalusglaber followed by muspahari, Gallus gallus (8% each) and Tupaiachinensis, Cricetulus griseus (7% each) against Homo sapiens sequence of IRS1.

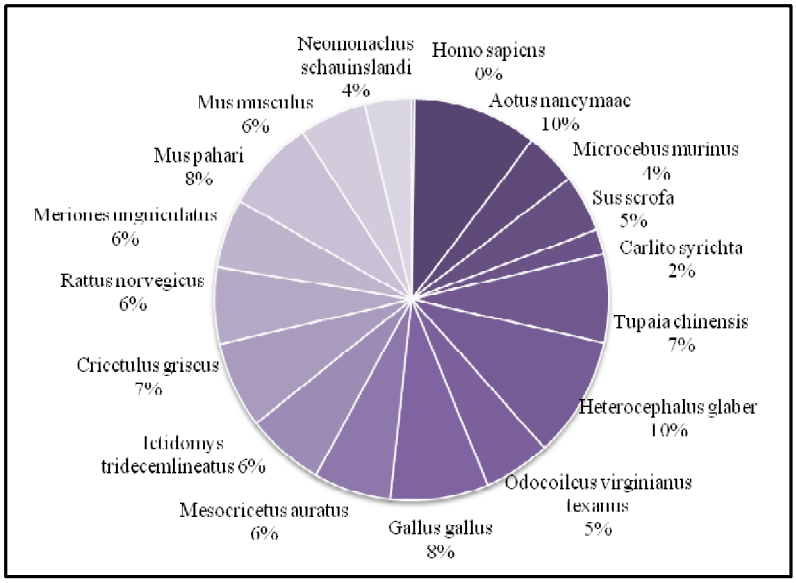

Figure 6. Pie chart showing composition distance of homo sapiens against various organism hits of IRS1 sequence.

In figure 7 again Pie chart are used to display pairwise distance obtained for homo sapiens against various organism hits of IRS1 sequence. In this maximum pairwise distance is obtained against muspharai and Merionesunguiculatus (8% each) and a minimum pairwise distance is obtained in case of aotusnancymaae, microcebusmurinus, Carlito syrichta (4% each) followed by Tupaiachinensis (5%).

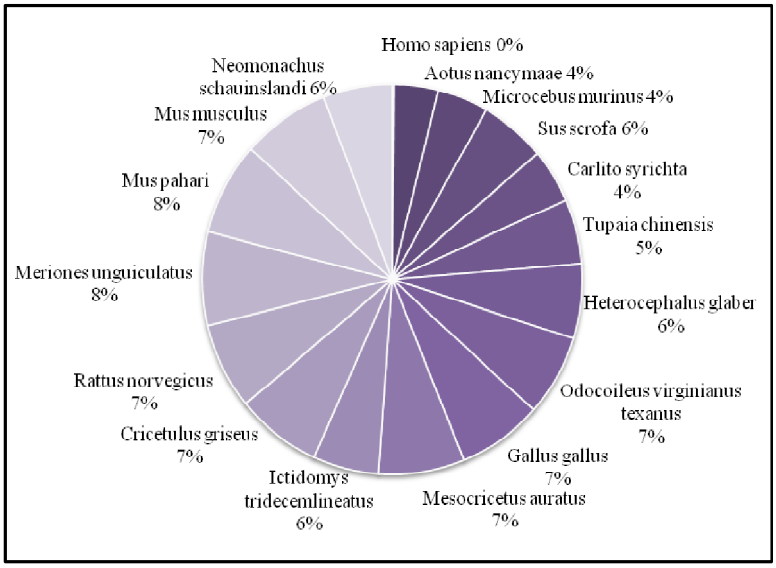

Figure 7. Pie chart showing pairwise distance of homo sapiens against various organism hits of IRS1 sequence.

Several categories of rates were used with equal probability to approximate the gamma distribution for each category. The mean of each category is then used to represent all the rates falling in the category. Maximum Likelihood Estimate of Gamma Parameter for Site Rates was also calculated using MEGA7. Refer table 3 for IRS1 Maximum Likelihood Estimate of Gamma Parameter for Site Rates. In this table maxima and minima were displayed for various Gamma Categories for #5

positions. The line graph in figure 8 shows the relative plot of minima and maxima of five Gamma Categories using Maximum Likelihood Estimate of Gamma Parameters for Site rates for IRS1

Table 3.IRS1 Maximum Likelihood Estimate of Gamma Parameter for Site Rates

Value Type	Gamma Categories				
	#1	#2	#3	#4	#5
Max	0.292	0.28	0.315	0.969	3.64
Min	0	0	0	0.001	0.03

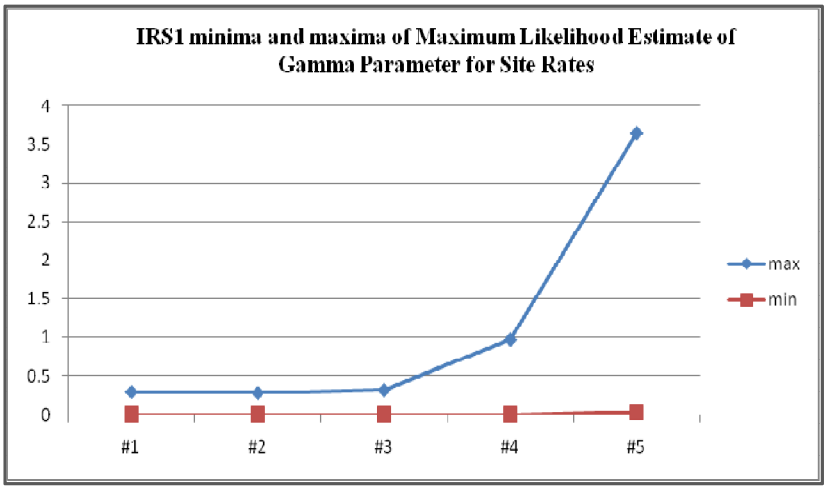

Figure 8. Line graph showing relative plot of minima and maxima of five Gamma Categories using Maximum Likelihood Estimate of Gamma Parameters for Site rates for IRS1

Tajim's Test: Tajima's test of neutrality [24, 25] that compares the number of segregating sites per site. It's an important statistic that is widely used in population genetics. The site is considered independently segregating if one can find two or more nucleotides at that site in comparison to m number of sequences under study. In our case, during simulation study Tajima's D never converges to 0. We found the D value as a small negative number which is not 0 but very close to zero. This is because of finite number of samples (Table 4).

Table 4. Tajima's Neutrality test statistics for IRS1 sequences

Peptide Name	*m*	*S*	p_s	Θ	π	*D*
IRS 1	19	372	0.313395	0.08966	0.084448	-0.245374

m = number of sequences, *n* = total number of sites, *S* = Number of segregating sites, $p_s = S/n$, $\Theta = p_s/a_1$, π = nucleotide diversity, and *D* is the Tajima test statistic

Here we expect a positive selection or we say selective sweeps within the population which doesn't undergo any sort of demographic changes like population contraction, expansion, immigration, migration, etc. Due to this population haplotypes will remain same and if any mutation occurs will be a rare chance. Then if we have lots of rare occurring mutations, our Tajima's D will be negative. It is now clear that IRS1 is a very important molecule. It is an intracellular protein that is responsible for communicating various extracellular signals within the cells. It is also confirm that IRS1 is the major substrate of insulin-like growth factor 1 receptor (IGF1R) [26]. IRS1 has multiple sites that act as docking sites for multiple SH2-containing proteins like Crk, SHP2, PI3K, Grb2 etc [27].

APPLICATION

It is a rapid way to calculate amino acid sequences in terms of evolutionary conservation rates and provides vital information about regions of structural and functional importance.

CONCLUSION

In this work, we have applied an innovative and rapid approach to study the structural, functional and phylogenetic relationship among the insulin receptor substrate proteins. Our study shows a rapid way to calculate amino acid sequences in terms of evolutionary conservation rates and provides vital information about regions of structural and functional importance. The information also provides greater in site into metabolism regulation controlled by insulin.

Conflict of Interest: We declare that we have no conflict of interest.

REFERENCES

[1]. F. G. Banting, C. H Best, The internal secretion of the pancreas. 1922, *Indian J Med Res*, **2007**, 125(3), 251-266.
[2]. C. Chakrabort, Biochemical and molecular basis of insulin resistance, *Curr Protein Pep Sci*, **2006**, 7: 113–21.
[3]. X. J. Sun, P Rothenberg, C. R. Kahn, J. M. Backer, E. Araki, P. A. Wilden, D. A. Cahill, B. J. Goldstein, M. F. White. Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein, *Nature*, **1991**, 352, 73–77.
[4]. Obberghen E Van, V. Baron, L. Delahaye, B. Emanuelli, N. Filippa, S. Giorgetti-Peraldi, P. Lebrun, I. Mothe-Satney, P. Peraldi, S. Rocchi, D. Sawka-Verhelle, Surfing the insulin signaling web, *Eur J Clin Invest*, **2001**, 31, 966–977.
[5]. J. M. Lizcano, D. R. Alessi, The insulin signaling pathway, *CurrBiol*, **2002**, 12, 236–238.
[6]. S. Anand, V. S. Muthusamy, S. Sujatha, K. N. Sangeetha, R. Bharathi, S. Anand, V. S. Muthusamy, S. Sujatha, K. N. Sangeetha, R. B. Raja, S Sudhagar, N. P. Devi, BS Lakshmi, Aloe emodin glycosides stimulates glucose transport and glycogen storage through PI3K dependent mechanism in L6 myotubes and inhibits adipocyte differentiation in 3T3L1 adipocytes, *FEBS Lett*, **2010**, 584: 3170-8.
[7]. G. Sesti, M. Federici, M. L. Hribal, D. Lauro, P. Sbraccia, R. Lauro, Defects of the insulin receptor substrate (IRS) system in human metabolic disorders, *FASEB J*, **2001**, 15, 2099-2111.
[8]. R. N. Kulkarni, J. N. Winnay, M. Daniels, J. C. Bruning, S. N. Flier, D. Hanahan, C. R.. Kahn, Altered function of insulin receptor substrate-1-deficient mouse islets and cultured betacell lines. *J Clin Invest*, **1999**, 104, R69-75.
[9]. M. F. White, IRS proteins and the common path to diabetes, *Am J Physiol Endocrinol Metab*, **2002**, 283: E413–E422.
[10]. M. V. Angela, L. Margarita, P. Sebastian, F. W. Morris, Insulin receptor substrate (IRS) proteins IRS-1 and IRS-2: Differential signaling in the insulin/insulin-like growth factor-I pathways in fetal brown adipocytes. *Mol Endocrinol*, **1998**, 12, 688–697.
[11]. S. Balakrishnan, S. Kumpatla, V. Viswanathan, Comparative analysis of transcription factors of insulin signalling, *Indian J Biotech*, **2010**, 9, 24–30.
[12]. D. Cai, S. Dhe-Paganon, P. A. Melendez, J. Lee, S. E. Shoelson, Two new substrates in insulin signaling, IRS-5/DOK4 and IRS-6/DOK5, *J BiolChem*, **2003**, 278: 25323–25330.
[13]. C. M. Taniguchi, B. Emanuelli, C. R. Kahn, Critical nodes in signalling pathways: insights into insulin action, *Nat Rev Mol Cell Biol*, **2006**, 7: 85–96.
[14]. D. L. Wheeler, T. Barrett, D. A. Benson, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. DiCuccio, R. Edgar, S Federhen, L. Y. Geer, Database resources of the National Center for Biotechnology Information, *Nucleic Acids Res*., **2007**, 35, D5–12.
[15]. PyMOL v2.0 Release Notes. pymol.org/dokuwiki/?id=media:new2. Rretrived march, **2018**.

[16]. R, Chenna, H, Sugawara, T, Koike, R, Lopez, T. J. Gibson, D. G. Higgins, J. D. Thompson, Multiple sequence alignment with the Clustal series of programs,. *Nucleic Acids Res*, **2003**, 31, 3497-500.
[17]. S. Kumar, G Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol BiolEvol*, **2016**, 33(7), 1870-1874. doi:10.1093/molbev/ msw054.
[18]. D. N. Brems, P. L. Brown, L. A. Heckenlaible,B. H. Frank, Equilibrium denaturation of insulin and proinsulin, *Biochemistry*, **1990,** 29, 9289-9293
[19]. P. Sieber,, K. Eisler,, B. Kamber, B. Riniker, W. Rittel, F. Marki, de Gasparo, M., Synthesis and biological activity of two disulphide bond isomers of human insulin: [A7- A11,A6-B7-cystine]- and [A6-A7,A11-B7-cystine]insulin (human) Hoppe-Seyler's Z. *Physiol. Chem*., **1978**, 359, 113−123.
[20]. C. C. Wang, C. L Tsou, The insulin A and B chains contain sufficient structural information to form the native molecule. *Trends Biochem. Sci*., **1991**, 16, 279−281.
[21]. D. F. Steiner, On the role of the proinsulin C-peptide. *Diabetes*, **1978**, 27, Suppl. 1, 145−148.
[22]. X. J. Sun, P. Rothenberg, C. R. Kahn, J. M. Backer, E. Araki, P. A. Wilden, D. A. Cahill, B. J. Goldstein, MFl White. Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein, *Nature*, **1991**, 352, 73-7.
[23]. E. Araki, M. A. Lipes, M. E. Patti, J. C. Brüning, B. Haag, R. S. Johnson, C. R. Kahn, Alternative pathway of insulin signalling in mice with targeted disruption of the IRS-1 gene. *Nature*, **1994**, 372, 186–90.
[24]. F Tajima, Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics*, **1989**, 123, 585-595.
[25]. S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, **2016**, 33(7), 1870-1874.
[26]. R. K. Dearth, X. Cui, H. J. Kim, D. L. Hadsell, A. V. Lee, Oncogenic transformation by the signaling adaptor proteins insulin receptor substrate (IRS)-1 and IRS-2, *Cell Cycle*, **2007**, Mar 15, 6(6), 705-13. Epub 2007 Mar 20. (REVIEW) PMID 17374994.
[27]. K. Mardilovich, S. L. Pankratz, L. M. Shaw. Expression and function of the insulin receptor substrate proteins in cancer, *Cell Commun Signal*., **2009**, Jun 17, 7, 14. doi: 10.1186/1478-811X-7-14. PMID 19534786.