JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

IDENTIFICATION AND QUANTIFICATION OF HUMAN INSULIN BY APPLYING CONSERVE DOMAIN SEARCH

Uma Kumari[1], Shanu[2], Devendra Kumar[3]
[1]Senior Bioinformatics Scientist, [2]Project Trainee, [3]Project Trainee
Bioinformatics Project and Research Institute, Noida - 201301, India

Abstract: Identification and quantification of human insulin is increasingly requested for clinical monitoring, for anti-doping purposes. Indeed, insulin analogues may be abused for suicide or homicide in different way in forensic interest. The recent technological advances the development of new extraction techniques particularly for anti-doping analyses, detection of insulins in the first validated quantitative method for analysis human insulin and its six analogues (Lispro, Aspart, Glulisine, Glargine, Detemir and Degludec) in plasma. Sequence analysis using complete amino acid sequence of human insulin-like growth factor I (IGF-I), a polypeptide isolated from serum, has been determined.Analysis PDBsum the relative disposition of the α-helices (red cylinders) Insulin-like growth factor-1 (IGF-1) act as mediator of growth hormone (GH), that promoting cell growth and differentiation in childhood have an anabolic effect in adults. IGF-1 network of growth factors, receptors and binding proteins involved in mediating cellular proliferation, differentiation and Bioavailability of IGF-1 is affected by insulin-like growth factor binding proteins (IGFBPs) which bind IGF-1 in circulation. IGFs is a (also called C-domain) single polypeptide chains which hold the insulin folding. IGF-1 and IGF-2(except that they are expressed at different stages of physiological development) have similar functional.

Keywords: Insulin, Structure Analysis, Validation, multi-domain model,COBALT, E-Value

INTRODUCTION

In these tissues the absorbed glucose gets converted into either fats (triglycerides) via lipogenesis or glycogen via glycogenesis or into both in case of liver.Beta cells are sensitive to blood sugar levels therefore they release insulin into the blood in response to high level of glucose and inhibit secretion of insulin when glucose level is slow.Insulin, it also enhances the glucose uptake and the metabolism in the cells, and thus reducing the blood sugar level.The secretion of insulin and glucagon into the blood glucose concentration is the primary mechanism of glucose homeostasis. By stimulating the gluconeogenesis and glycogenolysis in the liver glucagon increases the level of glucose in the blood.The linking (intrachain) disulphide bonds are formed at cysteine residues between the positions A7-B7 and A20-B19. There is an additional (intrachain) disulphide bond within the A-chain between cysteine residues at positions A6 and A11.The hexamer is the more stable form than the monomer, and also desirable for practical reasons; so the monomer is a much faster-reacting drug since the diffusion rate is inversely related to particle size.Insulin can aggregate and form febrile inter digitized beta-sheets. This can cause injection amyloidosis, and thus prevents the storage of insulin for long periods. An important feature the hexamer is about 36000 Dalton in size. The six molecules are linked together as three dimeric is the presence of zinc.Atoms (Zn^{2+}) on the axis of symmetry are surrounded by three histamine residues and three water molecules at position B10.In 1921, Frederick Banting and Charles Herbert Best working in the laboratory of J.R. Macleod at the University of Toronto were the first who isolate the insulin from pancreas of dog. In 1969, Dorothy Hodgkin determined the crystal structure of insulin in the solid state.Insulin, it is also the first protein which was chemically synthesized and produced by the Recombinant DNA technology.In 1978, the synthetic "human" insulin was the first genetically engineered that was produced by using E. coli.Biosynthetic recombinant "human" insulin or its analogues are the vast majority of insulin used worldwide.Recombinant insulin, it is produced either in coli or yeast (usually *Saccharomyces cerevisiae*).Fast-acting Insulin: - It is also called (rapid-acting insulin).It is absorbed quickly and after injection it starts working in about 15minutes and thus helps in lower the level of blood sugar after meals. In Random blood sugar test blood sample is taken at a random time and confirmed by the repeat testing. Blood sugar values are expressed in mill moles per litre (m/ mol) or milligrams per decilitre (mg/dl). This blood test indicates the average level of

blood sugar for the past two to three months and also measures the percentage of blood sugar attached to the oxygen-carrying protein in red blood cells (haemoglobin). Higher the blood sugar level, more the haemoglobin will be with sugar attached. On two separate tests an A1C level of 6.5 percent or higher suggests the diabetes. The super family of insulin includes many types of homologous proteins such as Insulin, IGF-1 (insulin-like growth factor-1), IGF-2. Insulin-like peptides (INSL3, INSL4, INSL5, INSL6, ILP-1 and ILP-2) (Antonova Y *et al.*, 2012)Bombyxin(Tamura*et al.*, 1986)Locust insulin related peptides, Molluscan insulin-related peptides (Smit*et al.*,1988) and Caenorhabditiselegans insulin-like peptides (Bairoch*et al.*,1998) which are found in mostly of the animal phylum. Insulin-like growth factors (IGF-1 and IGF-2) are responsible for regulating the growth and metabolism of all cells and which primarily produce by as an endocrine hormone. Unlike insulin, IGFs is a (also called C-domain) single polypeptide chains which hold the insulin folding. IGF-1 and IGF-2(except that they are expressed at different stages of physiological development) have similar functional. IGF-1 which expressed in after birth whereas IGF-2 which is expressed during early embryonic and fetal development in most of the somatic cells (Engstrom W, 2013).In insect largely found Bombyxin in the lepidopteron class. Prothoracicotropic hormone is previously name of bombyxin. (PTTH)(Tamura *et al.*,1986) stimulates the ecdysone are release by the prothoracic gland. These hormones are regulating the insect moulting and metamorphosis (Suzuki A, 1994). It also has a role in trehalose sugar metabolism (Okamoto*et al.*,2013) is known that the action of IGFs is regulated by IGF-binding proteins (IGFBPs). IGFBPs bind to IGFs which are responsible for increase the half-life of IGFs in the regulate receptor binding, thus regulate their signals and circulation(Kim *et al.*, 1997). There are six IGFBPs (IGFBP-1 to IGFBP-6) bind to IGFs with equal or greater affinity than the IGF-1R (Allard JB *et al.*,2018).As mentioned above, and insulin is the well-known member of this family. Maintain optimal blood glucose levels by facilitating glucose uptake by cell, and then it is mediate by interaction of insulin with its receptor is the main functions of insulin.They play distinct physiological roles in mammals and other animals of insulin and insulin like growth factor which are homologous. Whilst the former is the primary regulator of carbohydrate homeostasis and has effects on lipid and protein metabolism (Saltiel AR *et al.*,2001, Kitamura T *et al.*, 2003) the latter stimulate cell growth, replication and differentiation (Liu JP *et al.*,1993, Adams TE *et al.*,2000). The main mechanism of action of these hormones is mediated by particular Insulin Receptor (IR) or the Insulin-like Growth Factor Receptor (IGF1R) type 1. The IR and IGF1R, along with the IR-Related Receptor (IRR) (Ebina Y *et al.*,1985, Ullrich A *et al.*, 1986, Shier P *et al.*,1989) form subclass II of the Receptor Tyrosine Kinase (RTK) super family (Hubbard SR*et al.*,2000), and unlike the other members which dimerize or oligomerise upon ligand binding, (Heldin CH *et al.*,1996)the IR family members are pre-formed covalently-linked homo dimers ($\alpha 2\beta 2$) consisting of quite a few structural domains.Computational analysis work shows that It is potential that these receptors also function as hetero dimers, since IR/IGF1R hybrids have been found in all tissues expressing both receptors (Benyoucef S *et al.*,2007, Bailyes EM*et al.*,1997).

MATERIAL AND METHODS

NCBI is the branch of the National Institutes of Health (NIH). The headquarters of NCBI is located in Bethesda, Maryland and U.S. The NCBI is a series of databases. It is relevant to biotechnology and biomedicine. Backbone of bioinformatics as of database NCBI gives information about advances science and health biomedical and genomic information. Major databases include Gen Bank for DNA sequences and PubMed, a bibliographic database for biomedical literature.The FAST format is a text-based format. It is used for representing either nucleotide sequences or amino acid (protein) sequences in which base pair or amino acid are represented using single letter codes. A sequence in the FASTA format starts with a single line description,followed by the line of sequence data. It is recommended that all line of text be shorter than 80 characters in length. FASTA format also allows the sequence naming and comments to introduce the sequences. The FASTA format originates from the FASTA software package, but as now become a near universal standard in the field of bioinformatics.In bioinformatics, BLAST (basic local alignment search tool) is an algorithm and program used for finding region of local similarity between sequence .The program compare nucleotide or protein sequence to sequence database and calculate the statically significance. In data analysis of sample for the search of of an unknown gene in the mouse, the scientific perform a blast search of the human genome. If human carry a similar gene then blast will identify sequence in the human genome. It resembles on the mouse gene based on similarity sequence. This program comparesthe six frame conceptual translation products of a nucleotide query sequence (both strands) against a protein sequence database.The sensitivity of the commonly used progressive multiple sequence alignment method has been greatly improved for the alignment of divergent protein sequence. It calculate the best match for the selected sequence and lines them up so that identifies, similarity, difference can be seen.Clustal Omega is a multiple sequence alignment program. It uses seeded guide tree and HMM profile - profile technique to generate between three or more sequence. In order to produce a multiple alignment cluster omega require a guide tree which defines the order in which sequence profile are aligned.

RESULTS AND DISCUSSION

Human Insulin is the name which describes synthetic insulin which is laboratory grown in mimic the insulin in human. Human insulin control high blood sugar but doesn't cure diabetes. Three main groups of insulin, fast acting insulin, and Intermediate acting insulin long acting insulin. Rapid acting insulin– Rapid-acting insulin, also called fast-acting insulin, is a type of synthetic (man-made). It is used to control the blood sugar during meals and snacks and to correct high blood sugar. Intermediate acting insulin - It is used to control the blood sugar overnight, while fasting and between meals. Long acting insulin- long-acting insulin is also known as Basal insulin, and is a type of insulin which gives a slow steady release of insulin which helps to control the blood sugar between meals, and overnight.Premixed human insulin: human M2, M3, M5, Insuman comb 15,50NPH-Neutral Protamine HagedornBLAST-BLAST is a family of user friendly sequence similarity search tools on the web. The BLAST server is support through NCBI, USA. This tool identifies potential homologues for a given sequence. It analyze both DNA and protein sequences. BLAST is an example of pair wise. Both natural and recombinant forms of insulin are used therapeutically and treat type 1 diabetes while insulin itself is not hepatotoxic and also, it has not been linked to serum enzyme elevation or instances of clinically apparent liver injury, high doses insulin and glucose can result in hepatic glucogenesis and serum amino transferase elevation.BLAST P- It compares the submitted protein sequence against a protein database.E value analysis of BLASTp:In the Analysis of homo sapiens genome of sequence producing insulin growth factor like family member, precursor the e value

can be interpreted as meaning that in a database of the size used in the search one might expect to see no match with a similar score simply by random.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
insulin growth factor-like family member 1 precursor [Homo sapiens]	Homo sapiens	228	228	100%	3e-75	100.00%	110	NP_940943.1
insulin growth factor-like family member 1 [Pan troglodytes]	Pan troglodytes	224	224	100%	7e-74	97.27%	110	XP_024207236.1
insulin growth factor-like family member 1 [Nomascus leucogenys]	Nomascus leucogenys	219	219	100%	1e-71	96.40%	111	XP_003277651.1
insulin growth factor-like family member 1 [Pongo pygmaeus]	Pongo pygmaeus	218	218	98%	3e-71	97.22%	108	XP_054322254.1

Figure 1: Analysis by BLASTp (Description part)

The interpretation of the E value in the analysis of the Homo sapiens genome sequence that generates the precursor for insulin growth factor-like family member 1 suggests that in a database of the same size used in the search, a similar score might occur by chance alone without any intended match.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Homo sapiens isolate IRS1T2D14 insu...	Homo sa...	722	722	100%	0.0	100.00%	400	KF725074.1
PREDICTED: Pan paniscus insulin rece...	Pan pani...	718	718	100%	0.0	99.75%	4122	XM_008975275.3

Figure 2: Alignment view

In the biological sequence of insulin growth factor-like family member 1 precursor [Homo sapiens], mismatches are highlighted in red by default, and gray lines indicate gaps.

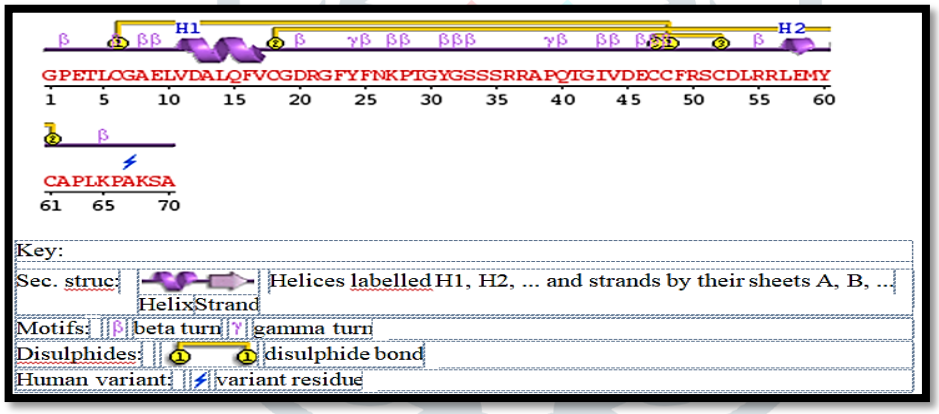

Figure 3.E- value analysis of BLASTN

Analysis of Homo sapiens isolate IRS1T2D14 insulin receptor substrate the E value analysis can be interpreted as meaning that in a database of the search, one might expect to see match with a similar score simply by random. Then analysis for plugin for an efficient 3D protein structure was done. Protein databank 3D structure (1BQT) structure of Human insulin-like growth factor-I (IGF-I) determined by IH-NMR the measurement geometry, 6 structures. The domain structure of human insulin-like growth factor-I has been determined through a technique NMR measurements and distance geometry calculations. A total of 320 inter atomic distance constraints, including 12 related to the disulphide bridges, that used in bond and element calculation. The resulting structure is characterized by the presence of three helical rods corresponding to the sequence regions, the residue Ala8-Cys18, Gly42-Cys48 and Leu54-Cys61, a turn structure and an extended structure exist in the amino (as a residue) Gly19-Gly22 and Phe23-Asn26 regions, respectively.Identification of the N- and C-terminal in the 3D structure or sequence , with their the r.m.s.d. value is 1.9 A for backbone atoms. And as result analysis the three alpha-helical regions are 1.0,0.9 and 0.8A, respectively,1.8 shows the their spatial arrangements. In sequence alignment or structure analysis obtained here shows that the human IGF-I molecule folds into a spatial structure very similar to that of insulin in an aqueous solution.

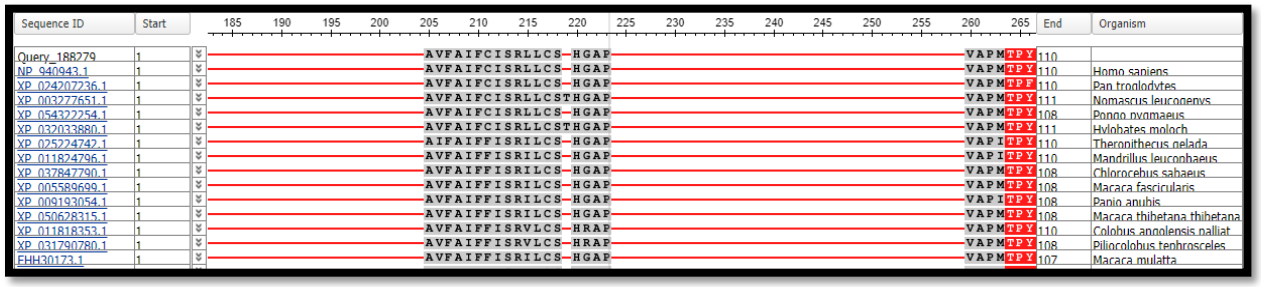

Figure 4: Wiring diagram of 3D Structure of human insulin-like growth factor-I (IGF-I)

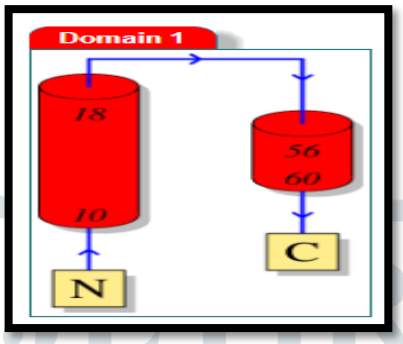

Figure 5: Diagram illustrating the protein's topology

The schematic diagram in figure 5 illustrates the protein's topology arranged into the relative disposition of α- strands helices (red cylinders).

CONCLUSION

Identification of sequence in insulin by applying computational technique the Blast P algorithm compare the nucleotide and protein database sequence to find out the parameter and describe the expectation value (E-value), graphic summary, description and global and local alignment. Expectation value E-value shows the similar quality (score) in biological database. The smaller the E-value(small E-value: low number of hits, but of high quality) , better the match. E = m x n/ 2bit-scorem - query sequence length n - total database length (sum of all sequences).BLAST N: - [E-value -0.0] E-VALUE of Insulin growth factor-like family member 1 precursor [Homo sapiens] through BLAST P: -[E-VALUE-2e-75]E-VALUE of Insulin receptor [Homo sapiens] through BLAST X:-[E-value-2e-26]The Graphic Summary shows alignments (as colored boxes) of database matches to our Query sequence (solid red bar under the color key). In graphics color bar summarizes the BLAST result at the top in linear map represents a protein that matches the query sequence. In fast alignment method the sequence of Human Insulin pair wise sequence alignment tool showed input as nucleotide or protein sequences and compared it with existing databases. It usually leads to fundamental biological insight into structure-function-sequence relationships of protein or nucleotide sequence families. IGF, insulin like growth factors; specific to vertebrates. Members include number of peptides that contain insulin-like growth factors I and II, that play a variety of roles in controlling processes such as growth, differentiation, and reproduction. On a cellular level they affect cell cycle, cell migration, apoptosis, proliferation, and differentiation. In analysis these peptide hormones are single chains that are cross-linked by three disulphide bonds.

REFERENCES:

1. Adams, T.E., Ela, V.C., Garrett, T.P., and Ward, C.W. (2000). Structure and function of the type1insulin-like growth factor receptor. Cell Mol Life Sci.57:1050–1093.
2. Allard, J. B., and Duane, C. (2018). IGF-Binding Proteins: Why Do They Exist and Why Are There So Many? Front. Endocrinol. 9:113–129.
3. Antonova, Y., Arik, A. J., Moore, W., Riehle, M. A., and Brown, M. R. (2012). Insulin –like Peptides: Structure, Signaling, and Function. Elsevier, 63-92.
4. Bailyes, E.M., Nave, B.T., Shoos, M.A., Orr, S.R., Hayward, A.C., et al. (1997). Insulin receptor/IGF-I receptor hybrids are widely distributed in mammalian tissues: quantification of individual receptor species by selective immune precipitation and immune blotting. Biochem J.327(Pt1):209–215.
5. Bairoch, A., Lyon, B.,Cedex, F.V., and Biomedical, G. (1998). New Insulin-Like Proteins with A typical Disulphide Bond Pattern Characterized in Caenorhabditiselegans by Comparative Sequence Analysis and Homology Modeling. Genome Res. 8:348–353.
6. Benyoucef, S., Surinya, K.H., Hadaschik, D., and Sidle, K.(2007). Characterization of insulin/IGF hybrid receptors: contributions of the insulin receptor L2 and Fn1 domains and the alternatively spliced exon 11 sequence to ligand binding and receptor activation. Biochem J.403:603–613.

7. Chena, R., Sugawara, H., Koike Lopez Gibson, T.J., *et al.* (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic acid Res. 31:3497-3500.
8. Ebina. Y., Edery, M., Ellis, L., Standring, D., Beaudoin, J., *et al*. (1985). Expression of a functional human insulin receptor r from a cloned cDNA in Chinese hamster vary cells. ProcNatlAcadSci USA.82:8014–8018.
9. Engstrom, W. (2013) Insulin-Like Growth Factor 2 in Development and Disease: A Mini-Review. Gerontology. 59: 240–249
10. Heldin, C.H., Ostman, A. (1996). Ligand-induced dimerization of growth factor receptors: variations on the theme. Cytokine Growth Factor Rev. 7:3–10.
11. Hubbard, S.R., and Till, J.H. (2000). Protein tyrosine kinase structure and function. Annul Rev Biochem. 69:373–398.
12. Kim, H., Rosenfeld, R. G., and Oh, Y. (1997). Biological roles of insulin-like growth factor binding proteins (IGFBPs). Exp. Mol. Med. 29:85–96.
13. Kitamura, T., Kahn, C.R., andAcini, D. (2003). Insulin receptor knockout mice. Annul Rev Physiol. 65:313–332.
14. Kumari,Uma and Choudhary, Ashok Kumar. (2016). Genome Sequence Analysis of SolanumLycopersicum by Applying Sequence Alignment Method to Determine the Statistical Significance of an Alignment. IJBTR. 6(3): 9-12.
15. Liu, J.P., Baker, J., Perkins, A.S., Robertson, E.J., and Eastertides, A. (1993). Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r). Cell. 75:59–72.
16. Okamoto, N.,&Mizoguchi, A. (2013). Insulin-like and IGF-like peptides in the silkmothBombyxmori: discovery, structure, secretion, and function. Frontiers in physiology. 4, 217.
17. Saltiel, A.R., Kahn, C.R. (2001). Insulin signaling and the regulation of glucose and lipid metabolism. Nature. 414:799–806.
18. Shier, P., and Watt, V.M. (1989). Primary structure of a putative receptor for a ligand of the insulin family. J Biol Chem.264:14605–14608.
19. Smit, A.B., Vreugdenhil,E.,Ebberink, R.H.M.,Geraerts, W.P.M.,et al. (1988) Growth-controlling molluscan neurons produce the precursor of an insulin related peptide. Lett.toNat. 331:535–538.
20. Suzuki, A.(1994). The brain secretary peptides that control moulting and metamorphosis of the silk moth, Bombyx.Int.J.Dev.Biol.38:301–310.
21. Tamura, S., Suzuki, A., Mizoguchi, A., Fujiwara, Y., Suzuki, A., Takahashi, S. Y., and Ishizaki, H. (1986). Amino acid sequence of a prothoracicotropic hormone of the silkworm Bombyx. Proc. Natl. Acad. Sci. U.S.A. 83:5840–5843.
22. Ullrich, A., Gray, A., Tam, A.W., Yang-Feng, T., Subokawa, M., *et al.*(1986). Insulin-like growth factor I receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity. Embo J. 5:2503–2512.
23. Uma Kumari and Ashok Kumar Choudhary. (2016). Computational Analysis of sequences to determine expectation value commonly used in Bioinformatics Database. IOSR-JCE. 18(6):pp20-22.
24. Virk N and Kumari U. (2022). Genome Sequence Analysis of Lungs Cancer Protein WDR74 (WD Repeat-Containing Protein). IJRASET. 10(V):4533-4537.
25. VirkNavjot&KumariUma. (2022). Identification of new potential drugs for lung adenocarcinoma causing protein RMB10 using computer-aided drug design approach. IJBTR. 12(2): 1-8.