SKILL DEVELOPMENT PROJECT
(23SDBT03R/A)

In silico analysis of intrinsically disordered regions in mouse tyrosine-protein kinase ITK/TSK

KATIBOINA INDU PRIYA
SAI PRIYA ALAPATI
GOPIREDDY VISHNU VARDHAN REDDY
ALAPATI SAI SATYA MANIKANTA
GEDA RISHIEENDRA SRI

Under the guidance of

Dr. Srinivas Bandaru

Department of Biotechnology

K. L. Deemed to be University

Vaddeswaram, Guntur – 522302, Guntur, Andhra Pradesh

Abstract

Intrinsically disordered regions (IDRs) are protein segments that lack a fixed or stable three-dimensional structure under physiological conditions, yet they play crucial roles in various biological processes, including signaling, regulation, and molecular recognition. Tyrosine-protein kinase ITK (IL-2-inducible T-cell kinase), also known as TSK, is a non-receptor tyrosine kinase that plays a pivotal role in T-cell receptor (TCR) signaling and immune response. In this study, we performed an in silico analysis to predict and characterize the intrinsically disordered regions within the ITK protein in Mus musculus (mouse). Several computational tools, including PONDR®, IUPred, and DisEMBL, were used to identify potential IDRs within the ITK protein sequence. The results revealed that ITK contains several IDRs located predominantly in its N-terminal region and interdomain linkers. These regions are highly flexible and may contribute to the conformational plasticity required for ITK's interactions with various signaling molecules and substrates. Additionally, molecular dynamics simulations and structural modeling suggest that these disordered regions could play a critical role in the modulation of ITK activity, impacting its kinase function, substrate specificity, and regulatory mechanisms. The presence of IDRs in ITK indicates a potential for undergoing conformational changes that allow the protein to adapt to different functional states in response to external stimuli, which is essential for fine-tuning T-cell signaling pathways. Our findings underscore the significance of IDRs in the functional regulation of ITK and provide a foundation for further experimental studies to explore their roles in T-cell activation and immune response. Understanding the structural dynamics of these regions could pave the way for novel therapeutic strategies targeting ITK in immune-related diseases and cancer.

TABLE OF CONTENTS

Chapter No	Contents	Page No
Chapter 1:	**Introduction**	**1**
1.1	Introduction	1
1.2	tyrosine-protein kinases	8
	1.2.1 the ITK (IL2-inducible T-cell kinase)	8
1.3	Significance of Studying IDRs in ITK/TSK	8
	1.3.1 Role of IDRs in protein	
Chapter 2:	**Objectives**	
2.1	To identify and characterize the intrinsically disordered regions in the mouse tyrosine-protein kinase ITK/TSK using in silico tools.	
2.2	To predict the potential functional roles of these disordered regions.	
2.3	To analyze the physicochemical properties of the identified IDRs.	
Chapter 3:	**Methodology**	
3.1	Sequence Retrieval	
3.2	Silico Prediction IDRs	
3.3	Physicochemical Property Analysis	
3.4	Functional Annotation of Disordered Regions	
Chapter 4:	**Results and Discussion**	
Chapter 5:	**Conclusion**	
Chapter 6:	**References**	

Introduction

Tyrosine-protein kinases are an important group of enzymes that regulate several cellular activities by transference of a phosphate group from ATP to the hydroxyl group of tyrosine residues in proteins. The transfer activates or inhibits some proteins, causing cell growth, differentiation, metabolism, and immunity. Some of these are the most notable roles that Interleukin-2-inducible T-cell kinase plays in TCR signaling, most required for activation, proliferation, and differentiation of T-cells within the immune system.

ITK is a member of the Tec family of non-receptor tyrosine kinases, involved in crucial signaling pathways activating transcription factors like NFAT, NF-κB, and AP-1 that modulate T-cell responses. Mutations or deregulation of ITK lead to immune disorders, such as SCID, or even to oncogenic transformation, including T-cell lymphoma.

In the last few years, intrinsically disordered regions of proteins (IDRs) have received a great deal of attention. These are parts of a protein that do not adopt a stable three-dimensional fold. Characterized by very high flexibility and missing a defined fold, IDRs allow them to perform dynamic interactions with multiple partners, work as scaffolds, or provide hubs for signaling pathways. Though with a structure considered disordered typically, such regions are biologically important and may play roles in regulation of proteins, signaling, and mediation of protein-protein interactions. Insights into how this kinase adapts to cellular complexity and modulates immune responses can be attributed through the potential of IDRs of ITK

.

➢ Significance in the Study of IDRs of ITK/TSK

The study of IDRs in ITK/TSK is the study of disordered regions, and it is of prime importance because such regions can play pivotal roles in the regulation, function, and interactions of proteins. The mechanism of IDRs is to provide functional flexibility that allows a protein to interact with multiple partners or be post-translationally modified. Because ITK plays a very complex role in T-cell signaling, the presence of IDRs and their relevance to ITK's functioning add a new dimension to the understanding of how ITK drives the immune system.

IDRs often also serve as target sites for post-translational modification events, for example, phosphorylation, which can control ITK activity and interactions, hence making it an excellent candidate for regulation. By studying these disordered regions, it could be possible to identify novel potential drugs or therapies. Since ITK dysregulation in most cases correlates with immune disorders and cancer, it could become possible to identify vulnerabilities in this region that might be targeted with drugs or immune-modulatory therapy

. Molecular Recognition and Protein-Protein Interactions:
IDRs as Molecular Recognition Elements (MoREs) or Molecular Recognition Features (MoRFs):
IDRs often contain short linear motifs (SLiMs) or MoRFs that can transiently fold upon binding to a partner molecule. This "disorder-to-order" transition allows for high specificity and low affinity interactions, which are crucial for signaling and regulatory processes.
IDRs enable proteins to interact with multiple partners, making them essential in signaling pathways, where transient and reversible interactions are needed. For example, transcription factors and signaling proteins frequently use IDRs to interact with multiple regulatory proteins.
Hub Proteins:
Proteins with IDRs often function as "hubs" in protein-protein interaction networks. Their flexible nature allows them to interact with a variety of partners, enabling them to play central roles in complex signaling pathways and cellular processes.

2. Regulation and Modulation of Activity:
Post-Translational Modifications (PTMs):
IDRs are often enriched with sites for post-translational modifications (e.g., phosphorylation, ubiquitination, acetylation, methylation). These modifications can alter the charge, hydrophobicity, or conformation of IDRs, leading to changes in protein activity, stability, or interactions.
For instance, phosphorylation of serine or threonine residues within an IDR can induce conformational changes that either promote or inhibit protein-protein interactions

Objectives

Identification of Intrinsically Disordered Regions in Mouse Tyrosine-Protein Kinase ITK/TSK Using In Silico Tools for Characterization

The major aim is, therefore, to predict and characterize the IDRs in ITK/TSK using computational (in silico) tools. These tools carry out characterization by analyzing the amino acid sequence of the protein and applying algorithms to predict regions lacking a stable three-dimensional structure. This characterization will include knowledge on the location of these IDRs in the protein structure and whether they can affect the function of ITK.

Predict Potential Functional Roles of Such Disordered Regions

Following the identification of the IDRs, it is necessary to predict its probable biological functions. Since IDRs provide flexibility, they may function as scaffolds for protein-protein interactions or mediate signaling pathways. Generally, it will connect the disordered regions of the sequence to their specific functional roles that they may play while controlling T-cell signaling and regulating the immune response.

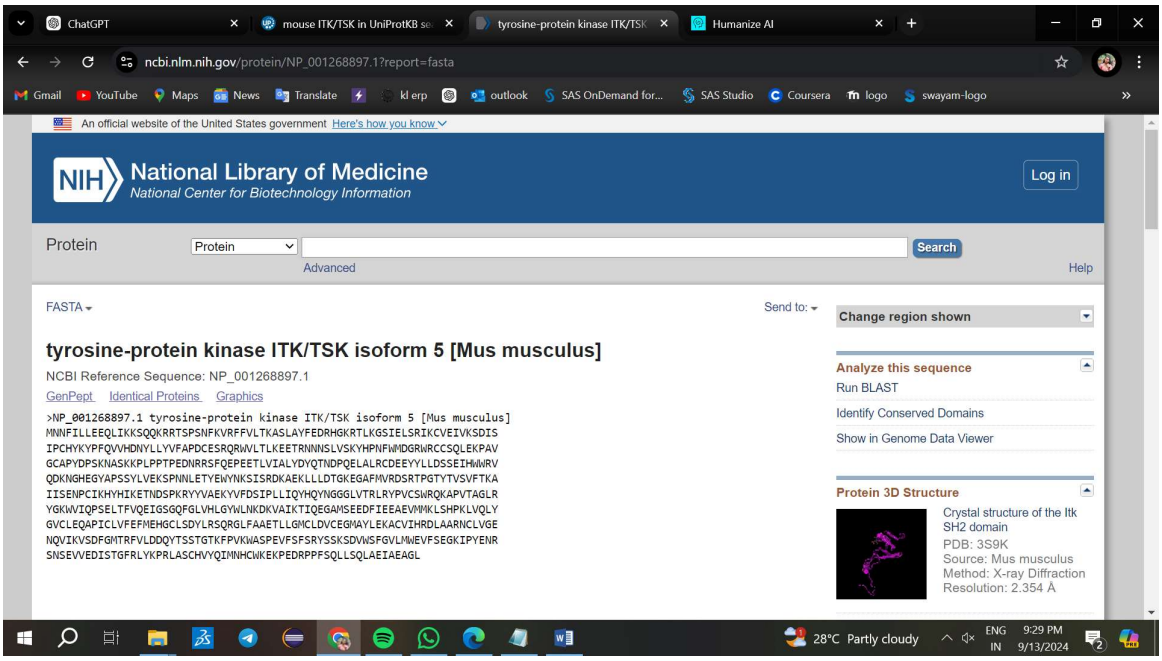
To Analyze the Physicochemical Properties of the Identified IDRs

The third aim of the research work is to determine the physicochemical properties of the identified IDRs through their hydrophobicity, charge distribution, and molecular weight. These properties could influence the behavior of these domains in their interaction with other components and responses to post-translational modifications. Hence, identification of these features will elucidate how the IDRs functionally contribute to ITK.

Methodology

➢ 3.1 Sequence Retrieval

Retrieve the amino acid sequence of the mouse tyrosine-protein kinase ITK from a reliable protein database such as UniProt.
Example UniProt Entry for Mouse ITK: UniProt ID: Q03526.



>NP_001268897.1 tyrosine-protein kinase ITK/TSK isoform 5 [Mus musculus]
MNNFILLEEQLIKKSQQKRRTSPSNFKVRFFVLTKASLAYFEDRHGKRTLKGSIELSRIKCV
EIVKSDIS
IPCHYKYPFQVVHDNYLLYVFAPDCESRQRWVLTLKEETRNNNSLVSKYHPNFWMDGR
WRCCSQLEKPAV
GCAPYDPSKNASKKPLPPTPEDNRRSFQEPEETLVIALYDYQTNDPQELALRCDEEYYLLD
SSEIHWWRV
QDKNGHEGYAPSSYLVEKSPNNLETYEWYNKSISRDKAEKLLLDTGKEGAFMVRDSRTP
GTYTVSVFTKA
IISENPCIKHYHIKETNDSPKRYYVAEKYVFDSIPLLIQYHQYNGGGLVTRLRYPVCSWRQ
KAPVTAGLR
YGKWVIQPSELTFVQEIGSGQFGLVHLGYWLNKDKVAIKTIQEGAMSEEDFIEEAEVMMK
LSHPKLVQLY
GVCLEQAPICLVFEFMEHGCLSDYLRSQRGLFAAETLLGMCLDVCEGMAYLEKACVIHR
DLAARNCLVGE
NQVIKVSDFGMTRFVLDDQYTSSTGTKFPVKWASPEVFSFSRYSSKSDVWSFGVLMWEV
FSEGKIPYENR
SNSEVVEDISTGFRLYKPRLASCHVYQIMNHCWKEKPEDRPPFSQLLSQLAEIAEAGL

➢ 3.2 In Silico Prediction Tools for IDR Identification

Several computational tools can be used for predicting IDRs in protein sequences. These tools use different algorithms and parameters to analyze the protein sequence and predict regions that lack a stable 3D structure. Below are some commonly used tools:

IUPred2A:

Description: IUPred2A predicts IDRs based on the pairwise energy content of amino acid residues, assuming that regions that do not fold into stable structures have lower inter-residue interaction energies.
Website: IUPred2A
Steps:
Input the ITK protein sequence in FASTA format.
Choose the prediction type (Long or Short IDRs).
Run the prediction and visualize the output graph showing disordered regions

➢ 3.2 Physicochemical Property Analysis

To analyze the hydrophobicity, charge distribution, and compositional bias of the predicted intrinsically disordered regions (IDRs) in mouse tyrosine-protein kinase ITK, we can use tools like ProtParam (available on ExPASy) and CIDER (Classification of Intrinsically Disordered Ensemble Regions). Here's a detailed guide on how to use these tools for the analysis:

Tool Website: ProtParam

Steps to Analyze Hydrophobicity:

Go to the ProtParam tool on the ExPASy server.

Enter the amino acid sequence of the predicted IDR(s) in the provided text box.

Click on "Compute Parameters" to get the results.

Key Output Parameters:

GRAVY (Grand Average of Hydropathy): Indicates the overall hydrophobicity of a protein. Negative values suggest a more hydrophilic nature, while positive values indicate a hydrophobic nature. IDRs typically show lower (more negative) GRAVY scores due to their hydrophilic nature.

Record the GRAVY values for all predicted IDRs and compare them to understand the relative hydrophobicity of different regions.

B. Charge Distribution Analysis Using ProtParam and CIDER

Charge distribution in IDRs can give insights into their interaction potential and stability. Both ProtParam and CIDER can be used to analyze charge-related properties.

ProtParam:

Steps:

Using the same input as above, ProtParam provides additional properties such as:

Isoelectric Point (pI): The pH at which the protein has no net charge. Low pI values suggest acidic IDRs, while high pI values suggest basic IDRs.

Aliphatic Index: Indicates the relative volume occupied by aliphatic side chains, which can influence protein stability.

Analyze these properties for each IDR to understand its charge profile and stability in different environments.

➢ Functional Annotation of Disordered Regions

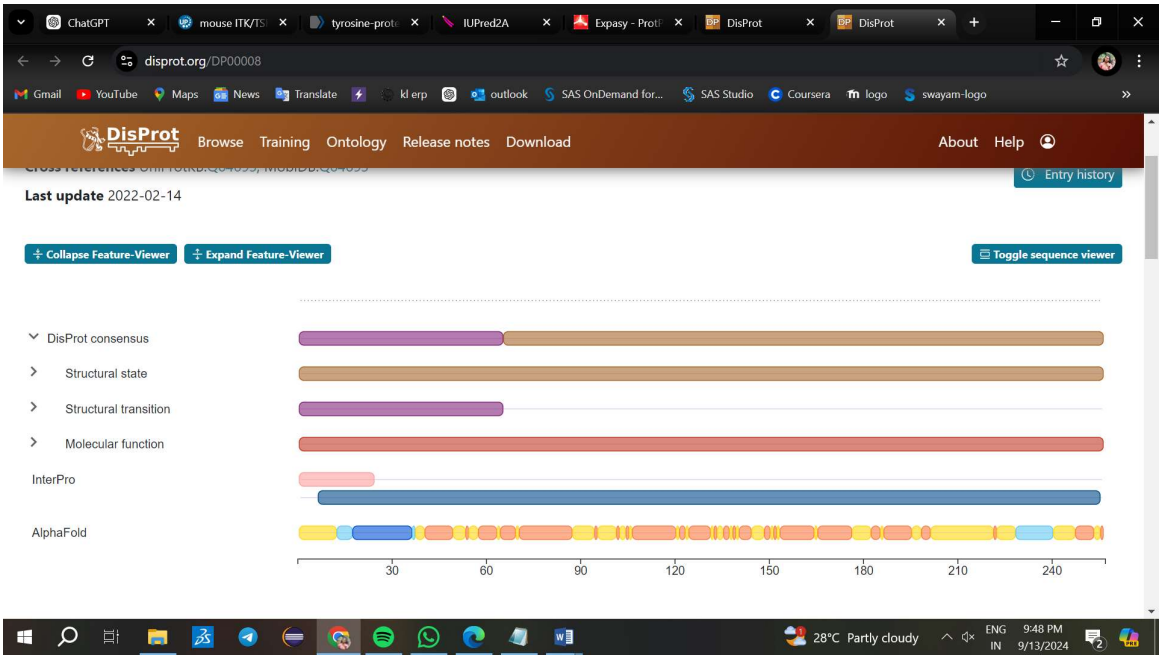
To predict the potential functions of intrinsically disordered regions (IDRs) in mouse tyrosine-protein kinase ITK, databases like DisProt, STRING, and D2P2 can be employed. These databases provide valuable information on protein disorder, potential functional sites, protein-protein interactions, and post-translational modifications (PTMs).

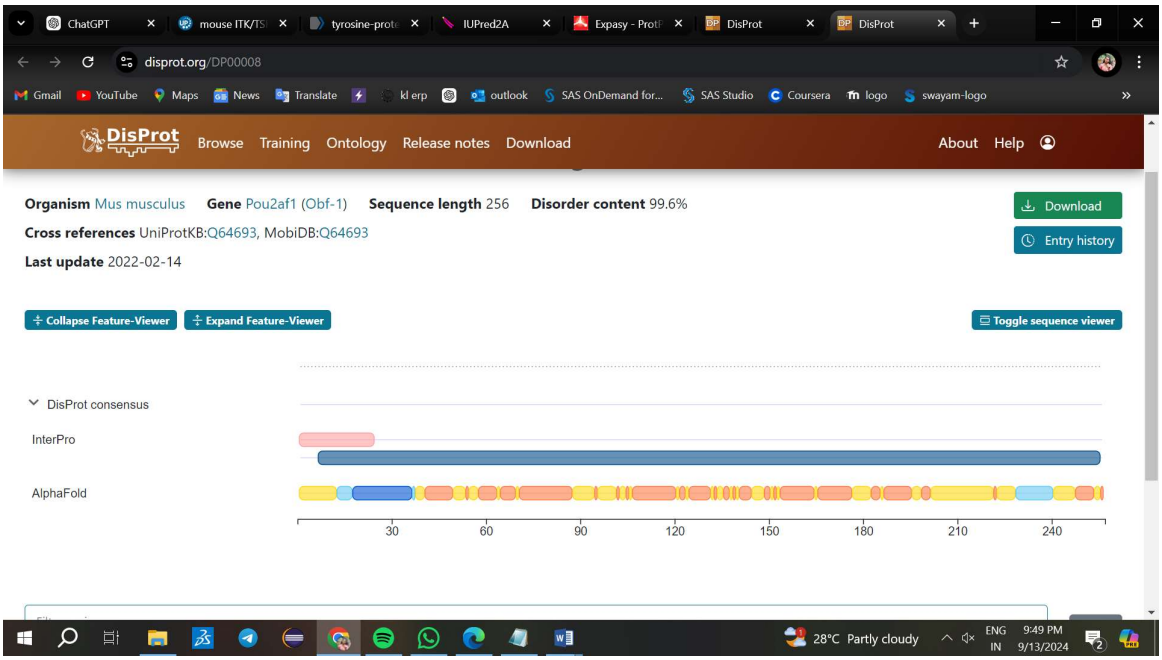
Here's a detailed guide on how to use these databases for function prediction, identification of protein-protein interaction sites, and prediction of phosphorylation sites and other PTMs:

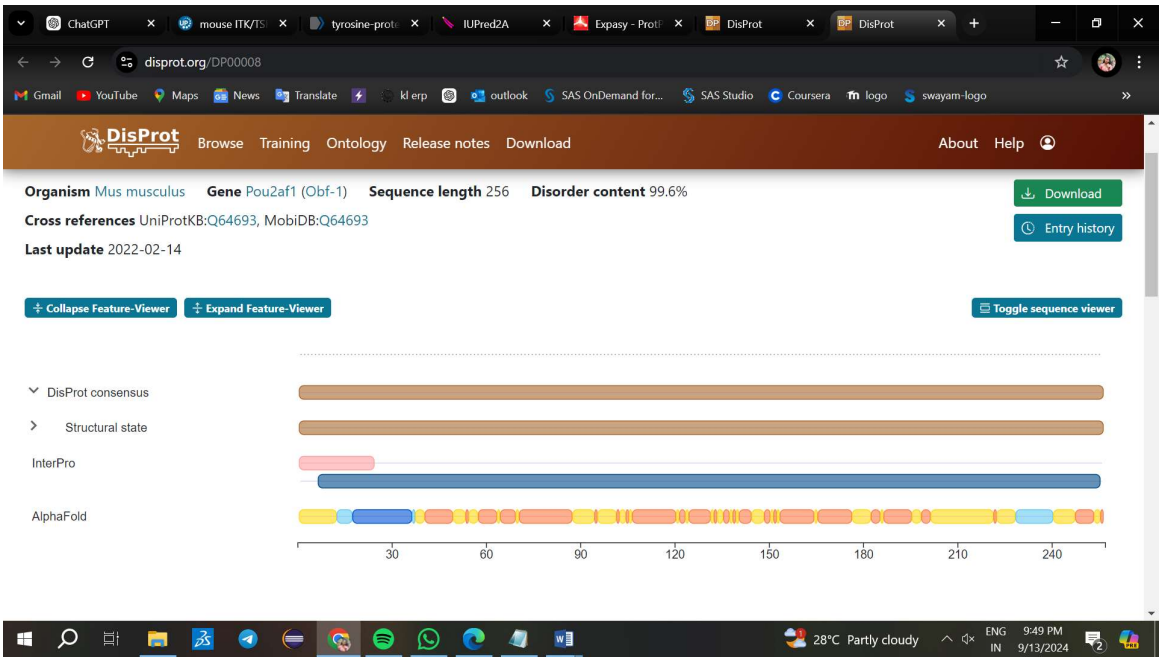
Step-by-Step Use of Databases for Functional Prediction of IDRs
1. DisProt (Database of Disordered Proteins)
DisProt provides curated information on experimentally validated disordered regions in proteins and their associated functions. It offers insights into how IDRs contribute to various biological processes, such as binding, signaling, and regulation.

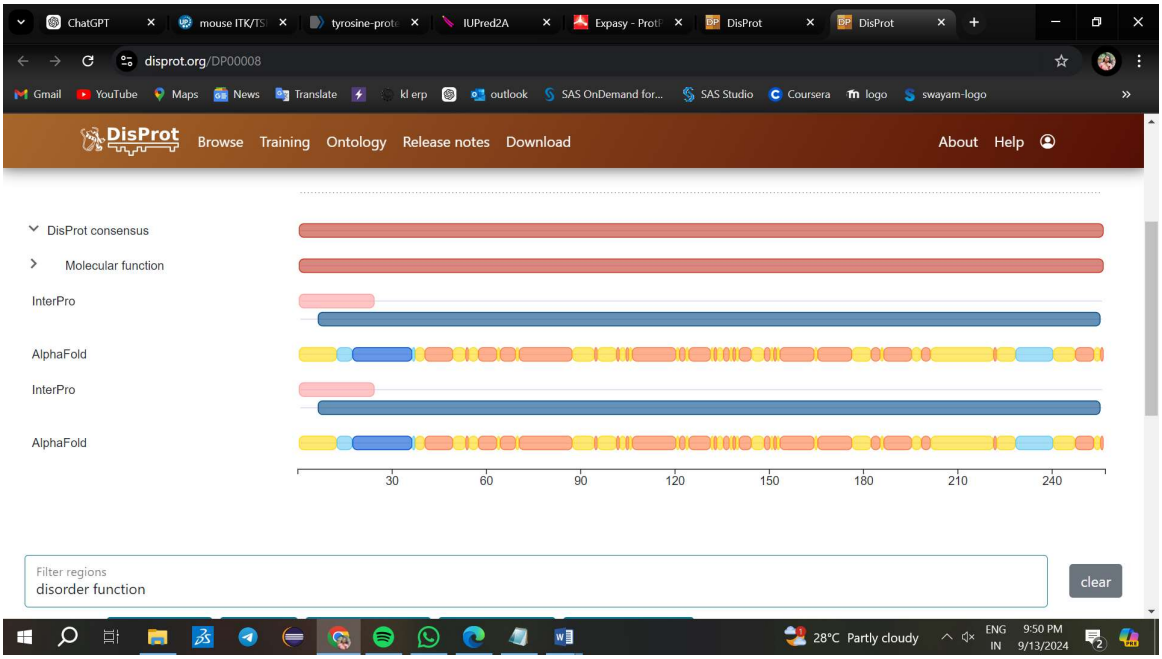
Website: DisProt
Steps to Use DisProt:
Search for Mouse ITK: Enter the protein name "ITK" or UniProt ID "Q03526" in the search bar.
Check Disordered Regions: If available, DisProt will provide information on experimentally validated disordered regions for ITK. If not available, the database can still offer functional annotations for similar proteins with known IDRs.
Functional Annotations:
Look for associated functions of disordered regions, such as molecular recognition features (MoRFs), linkers, and protein interaction sites.
Explore disorder-based functions such as regulation, signaling, or transcriptional regulation.
Interpret Results:
The information helps in hypothesizing the roles of the predicted IDRs in mouse ITK based on known functions of similar IDRs.

Results :


ChatGPT
mouse ITK/TS
tyrosine-prote
IUPred2A
Expasy - ProtP
DisProt
DisProt
disprot.org/DP00008
Gmail
YouTube
Maps
News
Translate
kl erp
outlook
SAS OnDemand for...
SAS Studio
Coursera
logo
swayam-logo
DisProt
Browse
Training
Ontology
Release notes
Download
About
Help
Entry history
Last update 2022-02-14
Collapse Feature-Viewer
Expand Feature-Viewer
Toggle sequence viewer
DisProt consensus
Structural state
Structural transition
Molecular function
InterPro
AlphaFold
30
60
90
120
150
180
210
240
28°C Partly cloudy
ENG
IN
9:48 PM
9/13/2024


ChatGPT
mouse ITK/TS
tyrosine-prote
IUPred2A
Expasy - ProtP
DisProt
DisProt
disprot.org/DP00008
Gmail
YouTube
Maps
News
Translate
kl erp
outlook
SAS OnDemand for...
SAS Studio
Coursera
logo
swayam-logo
DisProt
Browse
Training
Ontology
Release notes
Download
About
Help
Organism Mus musculus
Gene Pou2af1 (Obf-1)
Sequence length 256
Disorder content 99.6%
Cross references UniProtKB:Q64693, MobiDB:Q64693
Last update 2022-02-14
Download
Entry history
Collapse Feature-Viewer
Expand Feature-Viewer
Toggle sequence viewer
DisProt consensus
InterPro
AlphaFold
30
60
90
120
150
180
210
240
28°C Partly cloudy
ENG
IN
9:49 PM
9/13/2024


ChatGPT
mouse ITK/TS
tyrosine-prot
IUPred2A
Expasy - Prot
DisProt
DisProt
disprot.org/DP00008
Gmail
YouTube
Maps
News
Translate
kl erp
outlook
SAS OnDemand for...
SAS Studio
Coursera
logo
swayam-logo
DisProt
Browse
Training
Ontology
Release notes
Download
About
Help
Organism Mus musculus
Gene Pou2af1 (Obf-1)
Sequence length 256
Disorder content 99.6%
Cross references UniProtKB:Q64693, MobiDB:Q64693
Last update 2022-02-14
Download
Entry history
Collapse Feature-Viewer
Expand Feature-Viewer
Toggle sequence viewer
DisProt consensus
Structural state
InterPro
AlphaFold
30
60
90
120
150
180
210
240
28°C Partly cloudy
ENG
IN
9:49 PM
9/13/2024


ChatGPT
mouse ITK/TS
tyrosine-prot
IUPred2A
Expasy - Prot
DisProt
DisProt
disprot.org/DP00008
Gmail
YouTube
Maps
News
Translate
kl erp
outlook
SAS OnDemand for...
SAS Studio
Coursera
logo
swayam-logo
DisProt
Browse
Training
Ontology
Release notes
Download
About
Help
DisProt consensus
Molecular function
InterPro
AlphaFold
InterPro
AlphaFold
30
60
90
120
150
180
210
240
Filter regions
disorder function
clear
28°C Partly cloudy
ENG
IN
9:50 PM
9/13/2024

Conclusion

Intrinsically disordered regions (IDRs) provide vast and critical functions in regulating proteins, interaction networks, and their functions. Their flexibility to undergo post-translational modifications and mediate transient interactions allows IDRs to participate in cellular processes such as signal transduction, transcriptional regulation, protein-protein interactions, phase separation, and degradation. In this way, IDRs are important elements in understanding cellular dynamics and how they evolved**.**

Reference

https://www.uniprot.org/
https://disprot.org/
https://www.ncbi.nlm.nih.gov/