DATE: 08/10/22

WEBLEM 1

Introduction to Immunoglobins and its structural features

IMMUNOGLOBULIN (IG)

Immunoglobulins are glycoprotein molecules that are produced by plasma cells in response to an immunogen and which function as antibodies. The immunoglobulins derive their name from the finding that they migrate with globular proteins when antibody-containing serum is placed in an electrical field.

GENERAL FUNCTIONS OF IMMUNOGLOBULINS

a. Antigen Binding:

Immunoglobulins bind specifically to one or a few closely related antigens. Each immunoglobulin actually binds to a specific antigenic determinant. Antigen binding by antibodies is the primary function of antibodies and can result in protection of the host. The valency of antibody refers to the number of antigenic determinants that an individual antibody molecule can bind. The valency of all antibodies is at least two and in some instances more.

b. Effector Functions:

Frequently the binding of an antibody to an antigen has no direct biological effect. Rather, the significant biological effects are a consequence of secondary "effector functions" of antibodies. The immunoglobulins mediate a variety of these effector functions. Usually the ability to carry out a particular effector function requires that the antibody bind to its antigen. Not every immunoglobulin will mediate all effector functions. Such effector functions include:

i. Fixation of complement - This results in lysis of cells and release of biologically active molecules.

ii. Binding to various cell types - Phagocytic cells, lymphocytes, platelets, mast cells, and basophils have receptors that bind immunoglobulins. This binding can activate the cells to perform some function. Some immunoglobulins also bind to receptors on placental trophoblasts, which results in transfer of the immunoglobulin across the placenta. As a result, the transferred maternal antibodies provide immunity to the fetus and newborn

BASIC STRUCTURE OF IMMUNOGLOBULINS

The basic structure of the immunoglobulins is illustrated in figure 2. Although different immunoglobulins can differ structurally, they all are built from the same basic units.

a. Heavy and Light Chains

All immunoglobulins have a four chain structure as their basic unit. They are composed of two identical light chains (23kD) and two identical heavy chains (50-70kD)

b. Disulfide bonds
   i. Inter-chain disulfide bonds - The heavy and light chains and the two heavy chains are held together by inter-chain disulfide bonds and by non-covalent interactions The number of inter-chain disulfide bonds varies among different immunoglobulin molecules.
   ii. Intra-chain disulfide binds - Within each of the polypeptide chains there are also intra-chain disulfide bonds. C. Variable (V) and Constant (C) Regions

c. Variable and Constant Region

When the amino acid sequences of many different heavy chains and light chains were compared, it became clear that both the heavy and light chain could be divided into two regions based on variability in the amino acid sequences. These are the:

i. Light Chain – VL (110 amino acids) and CL (110 amino acids)
ii. Heavy Chain – VH (110 amino acids) and CH (330-440 amino acids

d. Hinge Region

This is the region at which the arms of the antibody molecule forms a Y. It is called the hinge region because there is some flexibility in the molecule at this point.

e. Domains

Three dimensional images of the immunoglobulin molecule show that it is not straight as depicted in figure 2A. Rather, it is folded into globular regions each of which contains an intra-chain disulfide bond. These regions are called domains.

i. Light Chains Domains: VL and CL
ii. Heavy Chain Domains – VH, CH1 – CH3 (or CH4)

f. Oligosaccharides

Carbohydrates are attached to the CH2 domain in most immunoglobulins. However, in some cases carbohydrates may also be attached at other locations.

STRUCTURE OF THE VARIABLE REGION

a. Hypervariable (HVR) or complementary determining regionst(CDR)

Comparisons of the amino acid sequences of the variable regions of immunoglobulins show that most of the variability resides in three regions called the hypervariable regions or the complementarity determining regions as illustrated in figure 3. Antibodies with different specificities (i.e. different combining sites) have different complementarity determining regions while antibodies of the exact same specificity have identical complementarity determining regions (i.e. CDR is the antibody combining site). Complementarity determining regions are found in both the H and the L chains.

b. Framework regions

The regions between the complementarity determining regions in the variable region are called the framework regions. Based on similarities and differences in the framework regions the immunoglobulin heavy and light chain variable regions can be divided into groups and subgroups. These represent the products of different variable region genes.

IMMUNOGLOBULIN FRAGMENTS: STRUCTURE/FUNCTION RELATIONSHIPS

Immunoglobulin fragments produced by proteolytic digestion have proven very useful in elucidating structure/function relationships in immunoglobulins.

a. Fab

Digestion with papain breaks the immunoglobulin molecule in the hinge region before the H-H inter-chain disulfide bond. This results in the formation of two identical fragments that contain the light chain and the VH and CH1 domains of the heavy chain.

Antigen binding - These fragments were called the Fab fragments because they contained the antigen binding sites of the antibody. Each Fab fragment is monovalent whereas the original molecule was divalent. The combining site of the antibody is created by both VH and VL. An antibody is able to bind a particular

antigenic determinant because it has a particular combination of VH and VL. Different combinations of a VH and VL result in antibodies that can bind a different antigenic determinants.

b. Fc

Digestion with papain also produces a fragment that contains the remainder of the two heavy chains each containing a CH2 and CH3 domain. This fragment was called Fc because it was easily crystallized.

Effector functions - The effector functions of immunoglobulins are mediated by this part of the molecule. Different functions are mediated by the different domains in this fragment (figure 5). Normally the ability of an antibody to carry out an effector function requires the prior binding of an antigen; however, there are exceptions to this rule.

c. F(ab')2

Treatment of immunoglobulins with pepsin results in cleavage of the heavy chain after the H-H inter-chain disulfide bonds resulting in a fragment that contains both antigen binding sites (figure 6). This fragment was called F(ab')2 because it is divalent. The Fc region of the molecule is digested into small peptides by pepsin. The F(ab')2 binds antigen but it does not mediate the effector functions of antibodies.

HUMAN IMUNOCLOBULIN CLASSES, SUBCLASSES, TYPES AND SUBTYPES

a. Immunoglobulin classes

The immunoglobulins can be divided into five different classes, based on differences in the amino acid sequences in the constant region of the heavy chains. All immunoglobulins within a given class will have very similar heavy chain constant regions. These differences can be detected by sequence studies or more commonly by serological means (i.e. by the use of antibodies directed to these differences).

i. IgG – Gamma Heavy
ii. IgM – Mu Heavy
iii. IgA – Alpha Heavy
iv. IgD – Delta Heavy
v. IgE – Epsilon Heavy
vi. Immunoglobulin Subclasses

The classes of immunoglobulins can de divided into subclasses based on small differences in the amino acid sequences in the constant region of the heavy chains. All immunoglobulins within a subclass will have very similar heavy chain constant region amino acid sequences. Again these differences are most commonly detected by serological means.

a. IgG Subclasses

a)IgG1 - Gamma 1 heavy chains

b)IgG2 - Gamma 2 heavy chains

c)IgG3 - Gamma 3 heavy chains

d)IgG4 - Gamma 4 heavy chains

b. IgA Subclasses

a)IgA1 - Alpha 1 heavy chains

b)IgA2 - Alpha 2 heavy chains

c. Immunoglobulin Types

Immunoglobulins can also be classified by the type of light chain that they have. Light chain types are based on differences in the amino acid sequence in the constant region of the light chain. These differences are detected by serological means.

1.Kappa light chains

2.Lambda light chains

d. Immunoglobulin Subtypes

The light chains can also be divided into subtypes based on differences in the amino acid sequences in the constant region of the light chain.

1. Lambda subtypes

a) Lambda 1

b) Lambda 2

c) Lambda 3

d) Lambda 4

e. Nomenclature

Immunoglobulins are named based on the class, or subclass of the heavy chain and type or subtype of light chain. Unless it is stated precisely, you should assume that all subclass, types and subtypes are present. IgG means that all subclasses and types are present.

g. Heterogeneity

Immunoglobulins considered as a population of molecules are normally very heterogeneous because they are composed of different classes and subclasses each of which has different types and subtypes of light chains. In addition, different immunoglobulin molecules can have different antigen binding properties because of different VH and VL regions.

REFERENCES:

1. IMMUNOGLOBULINS - STRUCTURE AND FUNCTION. (n.d.). Www.microbiologybook.org. https://www.microbiologybook.org/mobile/m.immuno-4.htm

DATE: 08/10/22

WEBLEM 1a

UniProt Database

(URL: https://www.uniprot.org/)

AIM:

To browse, search & retrieve immunoglobulin sequence using UniProt database.

INTRODUCTION:

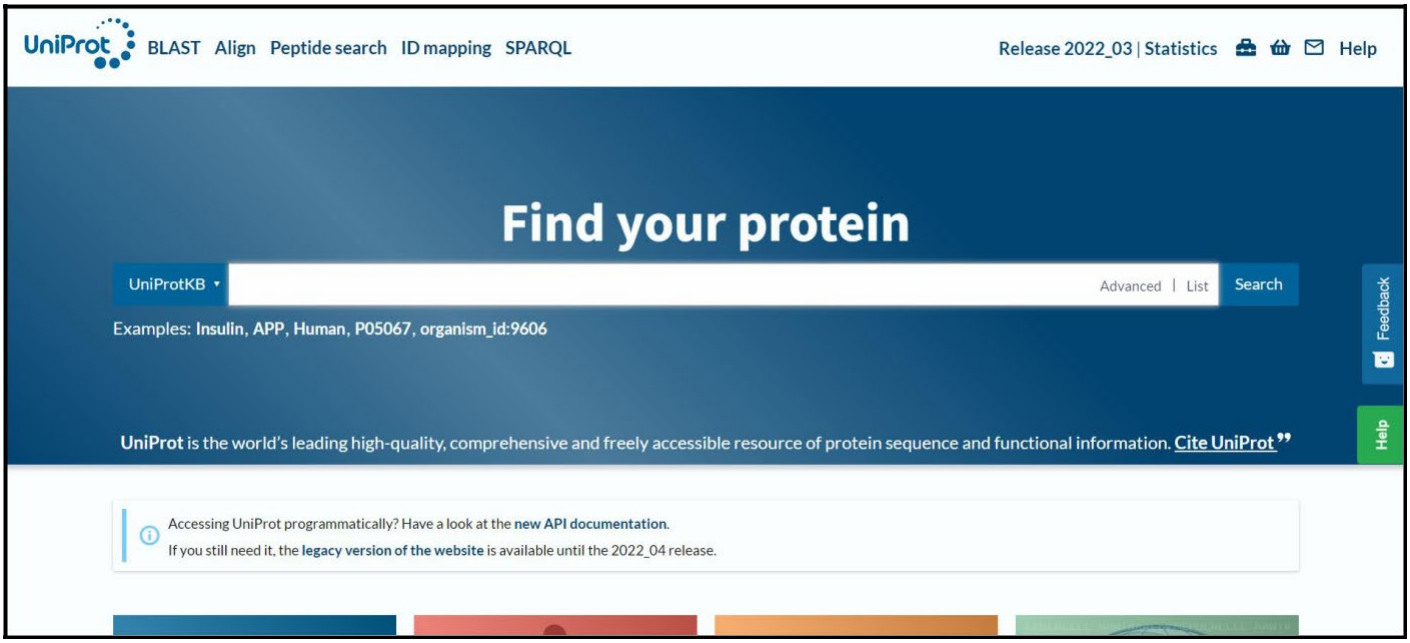
UniProt is a long-standing collection of databases that enable scientists to navigate the vast amount of sequence and functional information available for proteins. The UniProt Knowledgebase (UniProtKB) is the central resource that combines UniProtKB/Swiss-Prot and UniProtKB/TrEMBL. UniProtKB/Swiss-Prot contains over 550 000 sequences that have been created by our expert biocuration team. For these entries experimental information has been extracted from the literature and organized and summarized, greatly easing scientists access to protein information. UniProtKB/TrEMBL provides a further 60 million sequences that have been largely derived from high throughput sequencing of DNA. These entries are annotated by our rule based automatic annotation systems. We also provide a series of UniRef databases that provide sequence sets trimmed at various levels of sequence identity. Finally we provide the UniProt Archive (UniParc) that provides a complete set of known sequences, including historical obsolete sequences

The UniProt Knowledgebase (UniProtKB) is the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation. In addition to capturing the core data mandatory for each UniProtKB entry (mainly, the amino acid sequence, protein name or description, taxonomic data and citation information), as much annotation information as possible is added. This includes widely accepted biological ontologies, classifications and cross-references, and clear indications of the quality of annotation in the form of evidence attribution of experimental and computational data.

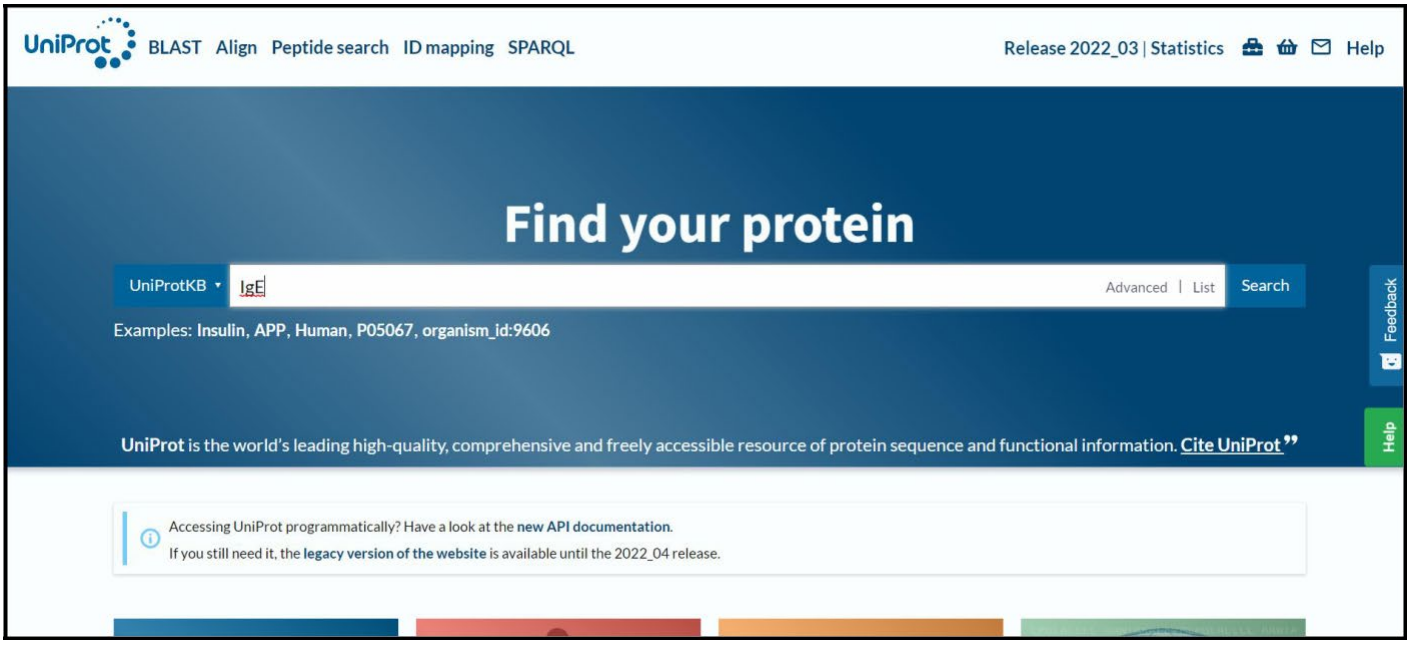
The UniProt Knowledgebase consists of two sections: a section containing manually-annotated records with information extracted from literature and curator-evaluated computational analysis, and a section with computationally analyzed records that await full manual annotation. For the sake of continuity and name recognition, the two sections are referred to as "UniProtKB/Swiss-Prot" (reviewed, manually annotated) and "UniProtKB/TrEMBL" (unreviewed, automatically annotated), respectively.

UniProtKB/TrEMBL (unreviewed) contains protein sequences associated with computationally generated annotation and large-scale functional characterization. UniProtKB/Swiss-Prot (reviewed) is a high quality manually annotated and non-redundant protein sequence database, which brings together experimental results, computed features and scientific conclusion

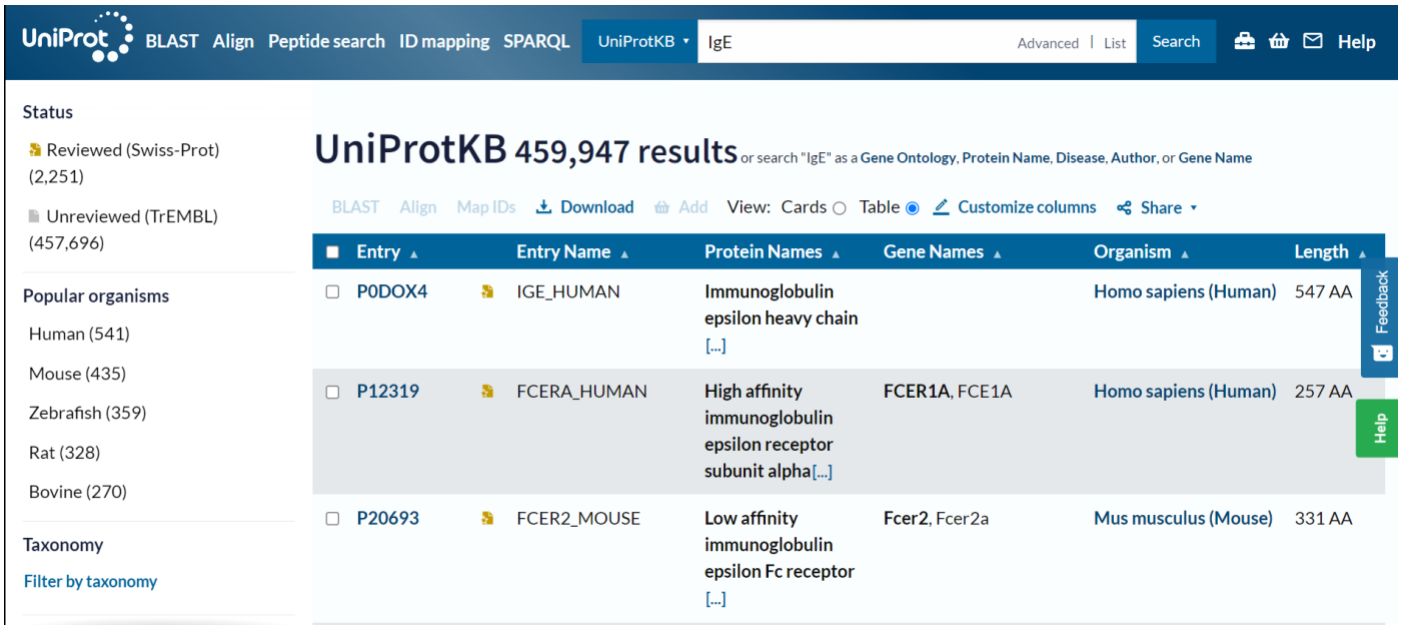
Steps to browse, search & retrieve immunoglobulin sequence using UniProt database:



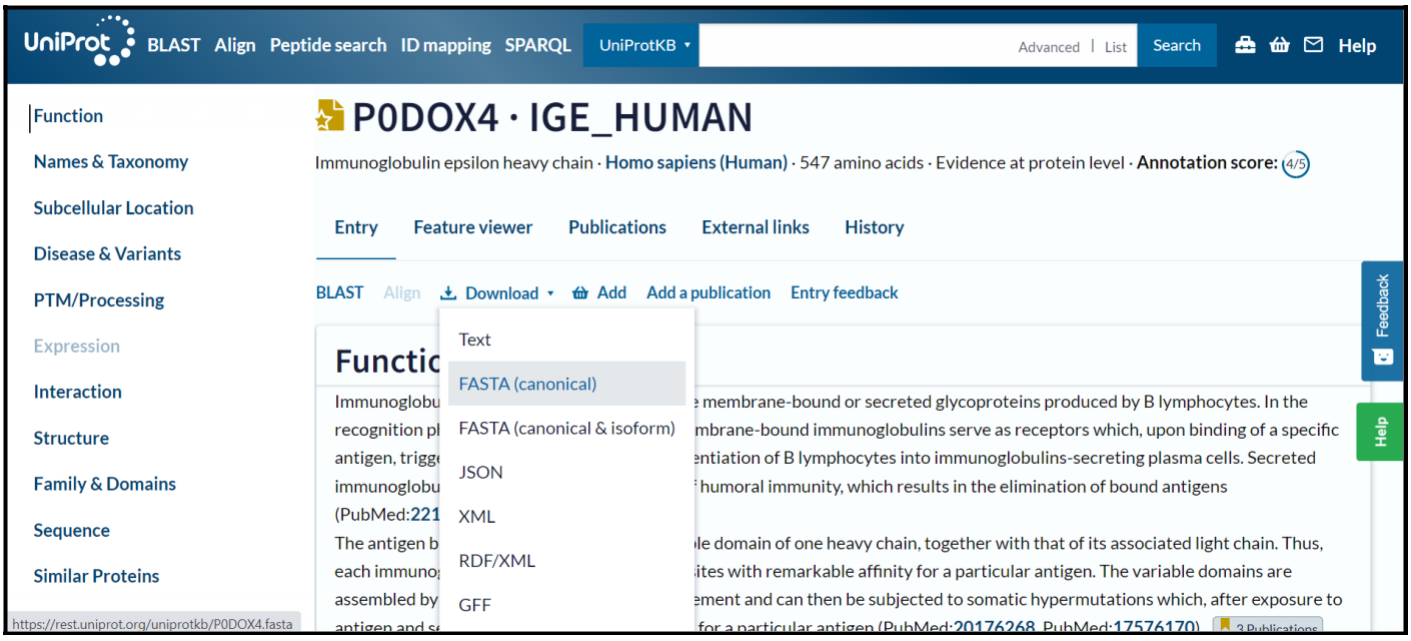
STEP 1: Open homepage of UniProt datbase



STEP 2: Entry any immunoglobulin to be searched



STEP 3: Select any one hit whose sequence is be retrieved



STEP 4: Click on download and FASTA(canonical)

```
>sp|P0DOX4|IGE_HUMAN Immunoglobulin epsilon heavy chain OS=Homo sapiens OX=9606 PE=1 SV=1
QVQLVQSGAEVRKPGASVRVSCKASGYTFIDSYVGWIRQAPGHGLEWIHWINPNSGGTNY
APRFQGRVTMTRDASFSTAYMDLRSLRSDDSAVFYCAKSDPFWSDYNFDYSSSEEGTEVT
YTVSGAWTLPSVFPLTRCCKNIPSNATSVTLGCLATGYFPEPVMVTWDTGSLNGTTLPAT
TLTLSGHYATISLLTVSGAWAKQMFTCRVAHTPSSTVDNKTFSVCSRDFTPPTVKILQSS
CDGLGHFPPTIQLCLVSGYTPGTINITWLEDGQVMDVDLSTASTESQGELASTESQLTLS
QKHWLSDRTYTCQVTYQGHTFQDSTKKCADSNPRGVSAYLSRPSPFDLFIRKSPTITCLV
VDLAPSKGTVNLTWSRASGKPVNHSTRKEEKQRNGTLTVTSTLPVGTRDWIEGETYQCRV
THPHLPRALMRSTTKTSGPRAAPEVYAFATPEWPGSRDKRTLACLIQNFMPEDISVQWLH
NEVQLPDARHSTTQPRKTKGSGFFVFSRLEVTRAEWQEKDEFICRAVHEAASPSQTVQRA
VSVNPGK
```

FIG 1. FASTA sequence for immunoglobulin ‘IgE’

RESULT:

Browsing, searching and retrieval of immunoglobulin sequence from UniProt database was demonstrated.

CONCLUSION:

UniProt is a long-standing collection of databases that enable scientists to navigate the vast amount of sequence and functional information available for proteins. It includes amino acid sequence, protein name or description, taxonomic data and citation information, biological ontologies, classifications and cross-references, and clear indications of the quality of annotation in the form of evidence attribution of experimental and computational data.

REFERENCE:

- UniProt. (2019). Uniprot.org. https://www.uniprot.org/

- UniProtKB. (2019). Uniprot.org. https://www.uniprot.org/help/uniprotkb
- UniProt: the universal protein knowledgebase. (2016). *Nucleic Acids Research*, *45*(D1), D158–D169. https://doi.org/10.1093/nar/gkw1099

DATE: 08/10/22

WEBLEM 1b

PDB Database

(URL: https://www.rcsb.org/)

AIM:

To browse, search & retrieve immunoglobulin structure using PDB database.

INTRODUCTION:

RCSB PDB (RCSB.org) is the US data center for the global Protein Data Bank (PDB) archive of 3D structure data for large biological molecules (proteins, DNA, and RNA) essential for research and education in fundamental biology, health, energy, and biotechnology.

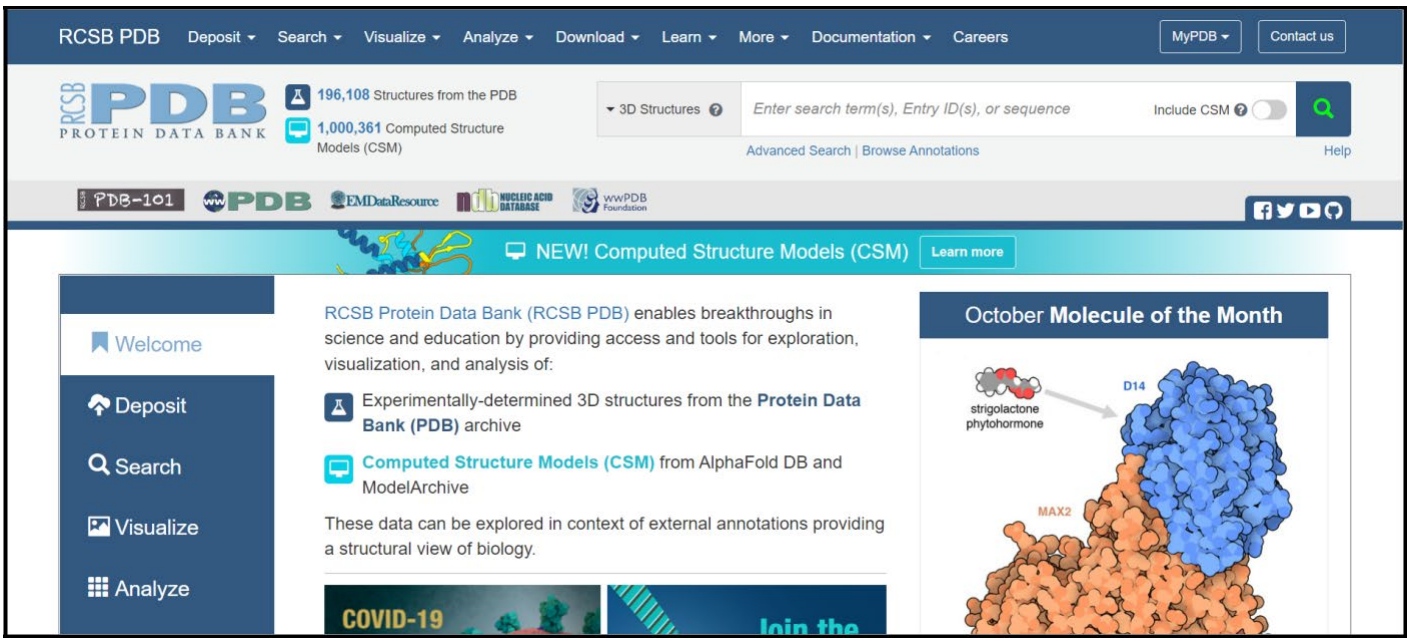
The Protein Data Bank (PDB) was established as the 1st open access digital data resource in all of biology and medicine (Historical Timeline). It is today a leading global resource for experimental data central to scientific discovery. Through an internet information portal and downloadable data archive, PDB provides access to 3D structure data for the molecules of life, found in all organisms on the planet.

Knowing the 3D structure of a biological macromolecule is essential for understanding its role in human and animal health and disease, its function in plants and food and energy production, and its importance to other topics related to global prosperity and sustainability.

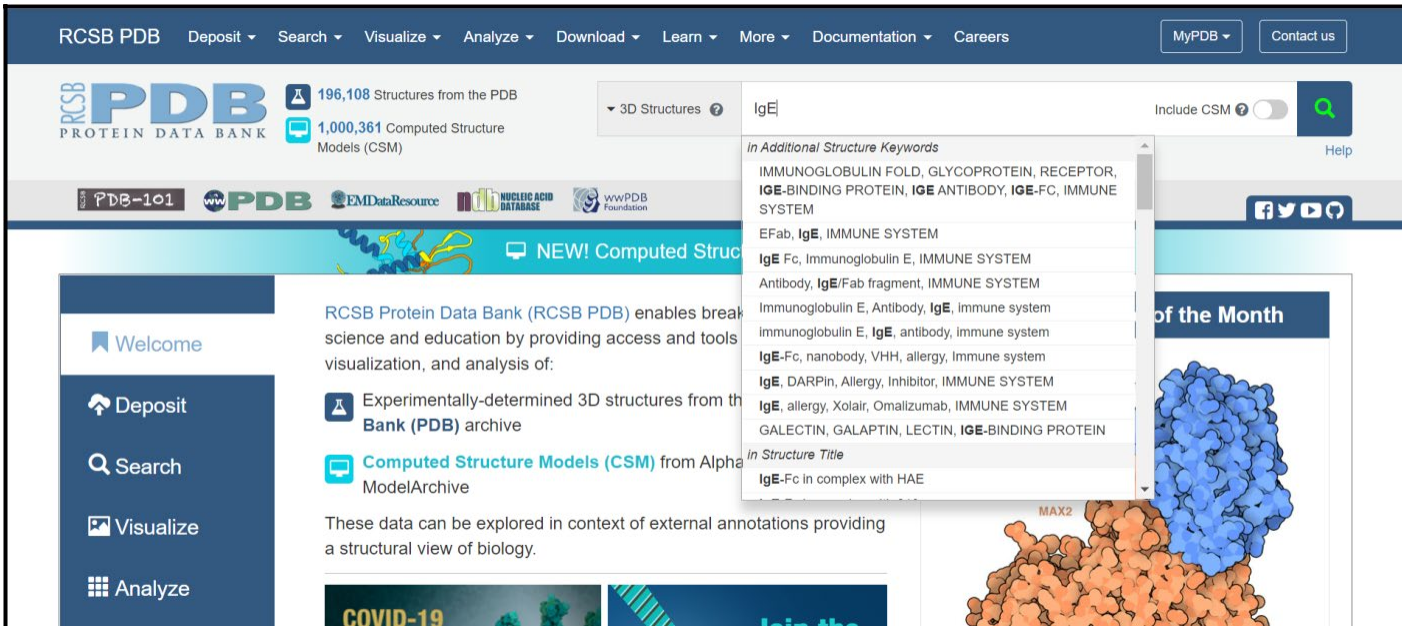
The enormous wealth of 3D structure data stored in the PDB has underpinned significant advances in our understanding of protein architecture, culminating in recent breakthroughs in protein structure prediction accelerated by artificial intelligence approaches and deep or machine learning methods.

RCSB PDB (Research Collaboratory for Structural Bioinformatics PDB) operates the US data center for the global PDB archive, and makes PDB data available at no charge to all data consumers without limitations on usage.

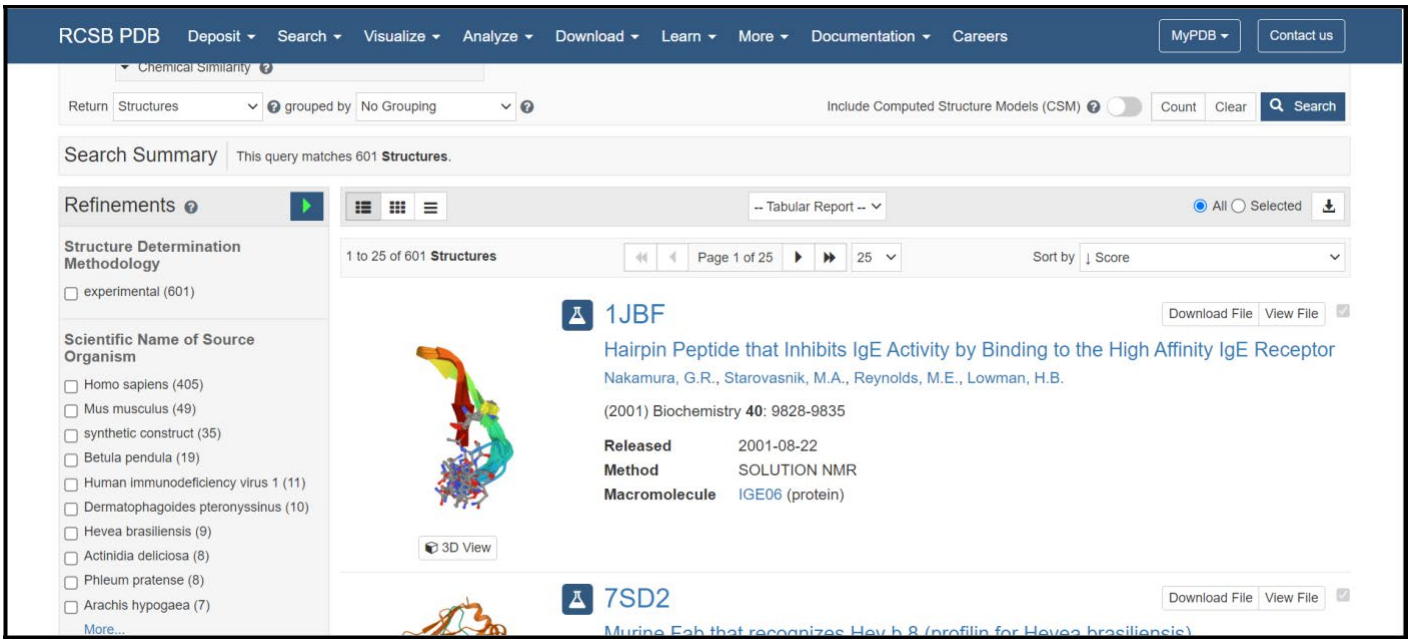
Steps to browse, search & retrieve immunoglobulin structure using PDB database:



STEP 1: Open homepage of PDB database



STEP 2: Search for immunoglobulin of interest



STEP 3: Out of all hits select immunoglobin structure of interest

STEP 4: Click on download and then PDB format

RESULT:

Browsing, searching and retrieval of immunoglobulin structure from PDB database was demonstrated.

CONCLUSION:

Protein Data Bank (PDB) archive of 3D structure data for large biological molecules (proteins, DNA, and RNA) essential for research and education in fundamental biology, health, energy, and biotechnology. Knowing the 3D structure of a biological macromolecule is essential for understanding its role in human and animal health and disease, its function in plants and food and energy production, and its importance to other topics related to global prosperity and sustainability.

REFERENCES:

1. *RCSB PDB: Homepage*. (2010). Rcsb.org. https://www.rcsb.org/
2. Bank, R. P. D. (n.d.). *RCSB PDB: About RCSB PDB: Enabling Breakthroughs in Scientific and Biomedical Research and Education*. Www.rcsb.org. https://www.rcsb.org/pages/about-us/index