



Research article

Antisoma Application: A Fully Integrated V-Like Antibodies Platform

Louis Papageorgiou^{1,2} and Dimitrios Vlachakis^{1,*}

¹ Computational Biology & Medicine Group, Biomedical Research Foundation, Academy of Athens, Soranou Efessiou 4, Athens 11527, Greece

² Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, University Campus, Athens, 15784, Greece

* **Correspondence:** Email: dvlachakis@bioacademy.gr; Tel: +30-210-6597647; Fax: +30-210-6597545.

Abstract: Antibodies, also called immunoglobulins, are large Y-shaped proteins which are produced and used by the immune system in order to identify and neutralize foreign substances called antigens. The study and research of antibodies is a critical issue in science and has led in remarkable results with medical applications such as diseases diagnosis, therapies and antibodies related drugs. Due to the importance of antibodies, a comprehensive database of full length protein sequences or crystal structures of immunoglobulin (IG) and T cell receptor (TR) V-like protein sequences from human and other vertebrate species has been created. In this article, we study in detail the unique numbering system for immunoglobulins of IMGT, we correlate the database antibody entries and we provide a more in detailed approach into the data. Finally, due to the deviation of the expected from the real data, a user friendly application has been created. This powerful and flexible tool manipulates the antibodies V-like protein sequences with the approach described below and gives the opportunity to the user to extract and study the results and its physicochemical properties. The need for new therapeutic targets and the development of more potent drugs is enormous, despite the efforts and the investments all these years for the development of novel next generation drugs, the discovery and the characterization of new therapeutic targets using immunoglobulins or antibodies. The final outcome from this project will support the development of novel therapies. Antisoma can be freely downloaded from <http://www.dimvl.com/antisoma>.

Keywords: antisoma application; antibodies/immunoglobulins database; V-like antibodies; antibodies numbering system; complementarity-determining regions; fragment region; physicochemical properties

1. Introduction

1.1. *Brief history of antibodies*

In 1890 Emil von Behring and Kitasato described an activity in serum of toxin-immunized animals that neutralized toxin. Transfer of immune serum could protect naive animals from diphtheria or tetanus. Bordet found in 1899 that animals could make antibodies against erythrocytes of other species and that these could direct destruction of the cells along with serum “complement” [1]. In 1900 Paul Ehrlich hypothesized that there are side chain receptors on cells that bind to a given pathogen. He speculated that this interaction induces the cell exhibiting the receptor to multiply and produce more copies of the same receptor [2]. Later in 1901–1920 Landsteiner demonstrates the ABO blood group system (Rh antibody was found in 1940). In the 1920s, Michael Heidelberger and Oswald Avery observed that antigens could be precipitated by antibodies and went on to show that antibodies were made of protein. In 1939 Elvin A. Kabat showed that all the three substances (antitoxin, precipitin and agglutinin) were one entity [3]. In the 1960’s Edelman, Porter, and Hilschmann elucidated the primary and secondary structure of antibodies [4]. In 1975 Kohler and Milstein invented monoclonal antibody technology while in 1976 Tonegawa cloned the first antibody gene [5]. Since then, much progress has been made in antibodies as therapeutic carriers [6] but challenges remain and it appears that deeper understanding of antibodies mechanisms is required to be overcome significant.

1.2. *Structure and function of antibodies*

An antibody is a protein produced by the body’s immune system when it detects harmful substances, called antigens. When an intruder, such as a virus, bacteria, or other chemical, enters the body, the immune system springs into action. Each antibody contains four polypeptide chains, two identical heavy chains and two identical light chains. The amino terminal ends of the polypeptide chains show considerable variation in amino acid composition and are referred to as the variable regions to distinguish them from the relatively constant regions [7,8].

The variable domains of heavy and light chains comprise the antigen-binding sites of the immunoglobulin molecules [9]. Each immunoglobulin monomer contains two antigen-binding sites and is said to be bivalent. The hinge region is the area of the heavy chains between the first and second constant region domains and is held together by disulfide bonds. This flexible hinge (found in

IgG, IgA and IgD, but not IgM or IgE) region allows the distance between the two antigen-binding sites to vary.[10] The variable domain is further subdivided into hypervariable and variable regions. Hypervariable regions have a high ratio of different amino acids in a given position, relative to the most common amino acid in that position. Within hypervariable regions of light and heavy chains, three complementarity-determining regions (CDR) exist—CDR 1, 2 and 3. Also, four fragment regions (FR) which have more stable amino acids sequences separate the CDR regions [11]. CDR regions allow antibodies to recognize an equally diverse number of antigens [12]. The CDRs are responsible for the structural interaction between the antibody V-domains and the antigen shape and size. Antibodies bind antigens through CDRs weak chemical interactions, including hydrogen bonds, electrostatic surface interactions, Van der Waals forces, and hydrophobic interactions [13]. This bonding is essentially non-covalent. In essence the CDRs dictate the antibody specificity and affinity for a specific antigen.

1.3. *Use of antibodies*

There is a wide use of the antibodies in science [14,15]. Immune molecules such as monoclonal antibodies are used as therapeutics in several disease conditions [16,17]. Also, the immune serum, a serum containing naturally or artificially produced antibodies to a given antigen, obtained from human or animal sources and it is used against pathogens which are capable of evading the immune system in the unstimulated state but which are not robust enough to evade the stimulated immune system [18]. Trends in antibody based diagnosis show advances in assay specificity, detection technologies and sensitivity. Sensitivity and specificity are ensured depending on whether or not the antigen to be quantified competes with labelled antigen for a limited number of antibody binding sites. Antibodies are also used for the treatment of people with cancer [19–22] with a new class of highly potent biopharmaceutical drugs designed as a targeted therapy[23], called Antibody-drug conjugates (ADCs) [22,24]. Antibody-drug conjugates have revolutionized the field of cancer chemotherapy. Unlike conventional treatments that damage healthy tissues upon dose escalation, ADCs utilize monoclonal antibodies (mAbs) to specifically bind tumour-associated target antigens [25] and deliver a highly potent cytotoxic agent [26].

Variable domains are of major importance in the analysis and the use of antibodies [9]. ADCs for example are complex molecules composed of an antibody (a whole mAb or an antibody fragment such as a single-chain variable fragment [scFv]) linked, via a stable, chemical, linker with labile bonds, to a biological active cytotoxic (anticancer) payload or drug. The variable domains of an antibody consist of a β -sheet framework with hypervariable regions (or complementarity-determining regions [CDRs]) which fashion the antigen-binding site. Due to the big variability of those domains there is an urgent need to establish a set of rules which separate the FR and CDR regions.

1.4. *Numbering system*

When studying the antibodies in more detail there comes to the need to distinguish their regions. For this reason a variety of numbering systems has been used. The Kabat numbering system is a widely adopted standard for numbering the residues in an antibody in a consistent manner. The Kabat Databas initially started in 1970 to determine the combining site of antibodies based on the available amino acid sequences at that time [27]. Bence Jones proteins, mostly from human, were aligned, using the now-known Kabat numbering system, and a quantitative measure, variability, was calculated for every position [27,28]. Other numbering systems are the Chothia Numbering Scheme, the Martin Numbering System and the IMGT Numbering System [29,30].

The IMGT unique numbering relies on the high conservation of the structure of the variable region [31]. This numbering, is set up after aligning more than 5000 sequences, takes into account and combines the definition of the framework (FR) and complementarity determining regions (CDR), structural data from X-ray diffraction studies, and the characterization of the hypervariable loops. The delimitations of the FR-IMGT and CDR-IMGT regions have been defined, and correspondence between the IMGT numbering and the other numberings has been established. By facilitating comparisons between the sequences and the description of alleles and mutation, the IMGT unique numbering represents a big step forward in the analysis of the immunoglobulin sequences of all species [32]. Moreover, it gives insight into the structural configuration of the variable domain and opens interesting views on the evolution of the sequences of the V-set, since this numbering has been applied with success to all the sequences belonging to the V-set of the immunoglobulin super-family, including non-rearranging sequences in vertebrates and in invertebrates [33,34].

1.5. *The project*

The major goal of this study is to develop a free novel antibodies application which will offer significant information related to antibodies V-domain. Moreover, the generated sample set from each antibodies search could be filtered based on IMGT numbering system. Furthermore, the ability to extract important physiochemical properties information from the antibodies' regions contained in V-domains (FR and CDR regions) [35,36] is also provided.

2. **Methods and Development**

This application was developed in four steps as shown in the figure 1. This approach was chosen in order to reach the final goal in a clearer and more stable way. The success of this application is based on the proper function of each part separately.

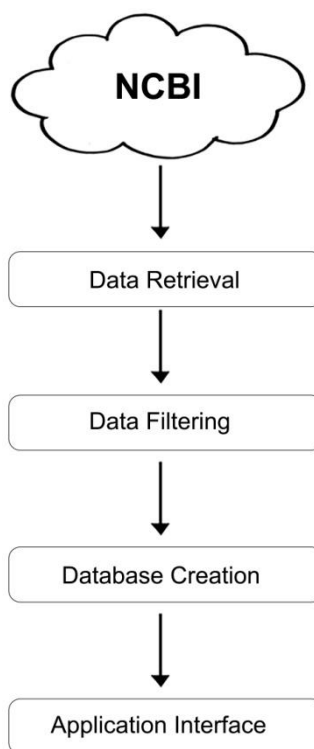


Figure 1. Data flow chart.

The procedure followed in this approach in four basic steps.

2.1. *Data retrieval*

The dataset used for the Antisoma application has been retrieved from the NCBI databases GenBank [37], Uniprot/Swiss-prot [38] and Protein Data Bank [38]. Moreover the dataset correlated with the IMGT/LIGM-DB [39]. For the purposes of this application, amino acid sequences were retrieved in fasta format of full length sequences or 3D solved sequences from 9 different animal species (including Homo sapiens and Mus musculus), as well as some unidentified ones. Along with the fasta format of the IG V-domain sequences, the current IMGT/LIGM-DB annotation [39] metadata was also retrieved, concerning the demarcation of boundaries of the CDR and FR regions within the amino acid sequences. The total number of entries sums up to 6.683 sequences.

2.2. *Data filtering*

In the data filtering step, the IMGT numbering system was used. For the database creation procedure the extracted protein sequences from NCBI databases have been correlated with the IMGT/LIGM-DB data in order to create a database that can answer queries about the CDR and FR regions on each sequence. In addition, the lengths of all regions for each sequence have been tested, against the length limitation given by the IMGT unique numbering system. The expected length for

each V-Domain region is given in the Table 1. As a result, it has been found that only 4.059 out of the entries correspond to the IMGT unique numbering system. Each sequence is labelled as “Rejected” with “Yes” or “No”.

Table 1. Unique numbering system for immunoglobulins of IMGT.

Region (V-like)	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4
Length (amino acids)	25–26	5–12	16–17	0–10	36–39	> 2	10–12

2.3. Database creation

A database was created using the Matlab Stores technologies. More specifically, all V-domain-like protein antibodies sequences were organized in a structure array using element by element organization. Element by element organization allows easier access to all information related to a single antibody entry including species name, protein sequence, sub-domains and sub-regions, and support simple indexing to access the data. Furthermore, element by element organization is the best structure array choice for most sequence processing applications and it can be the best for applications wherein you will routinely need to access corresponding subsets of structure fields. Moreover a structure array is a low memory data type that groups related data using data containers called fields. Each field can contain data of varying types and sizes.

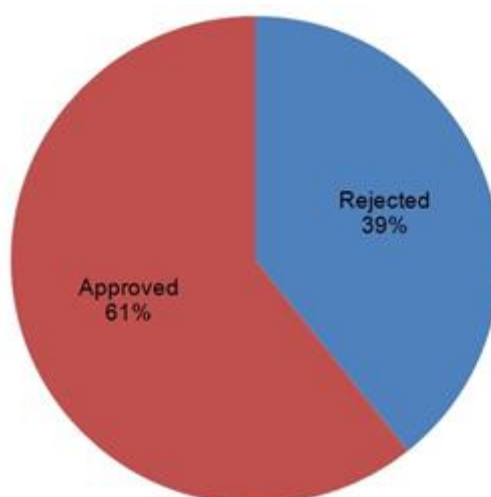


Figure 2. The percentage of the amino acid sequences which have been rejected by applying the length limitation provided by the IMGT numbering system.

Finally, in each antibody entry more critical information about its hyper variable regions has been added such as the fragment and complementary determine regions after filtering results using the antibodies numbering system rules. In Figure 2 the final total percentage of the antibody protein sequences which have been rejected by applying the length limitation in each sub-region of the hyper variable domain provided by the IMGT numbering system is presented.

2.4. Antisoma application interface

Antisoma application has been developed in MATLAB language [40] in the MATLAB R2014a application under a 64-bit Windows 7 operating system and it is compatible with all the 32-bit and 64-bit Windows versions. It comes with an executable file which has been exported from the MATLAB environment under the MATLAB license which is distributed from The MathWorks, Inc. and initializes the application in a Windows-running machine with a few easy steps. Antisoma application platform is a menu-driven interface as well as a tab step by step layout which is separated in five main categories “Species”, “Filter Options”, “Data Set Table”, “Biostatistics Tools” and “Biostatistics Features” (Figure 3). It provides the user with a process window to select, filter and save V-domain antibodies sequences from several species, based on the selected antibodies numbering system regions as well as to monitor physicochemical calculations of the selected region of interest (FR and CDR regions).

The screenshot displays the Antisoma application interface with the following components:

- Species:** A list of species including *Anarhichas minor*, *Homo sapiens*, *Mus musculus domesticus*, *Mus musculus*, *Mus sp.*, *Oryzotagus cuniculus*, *Rattus norvegicus*, *Sus scrofa*, and *unidentified*.
- Filter Options:** Checkboxes for FR1, CDR1, FR2, CDR2, FR3, CDR3, and FR4. There are also checkboxes for 'FR1-IMGT' through 'FR4-IMGT' and buttons for 'IGV-Like' and 'IGV-Like*'. A 'Search' button is located below.
- Data Set Table:** A table with columns: NAME, FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4, SQ, REJECT_FR1, and REJECT_CDR1. It contains 12 rows of antibody sequence data.
- Biostatistics Tools:** A section with 'Import Options' (A1, A2, B, C) and buttons for 'Export text file', 'Export xls file', and 'Export mat file'.
- Biostatistics features:** A table showing physicochemical calculations for the selected regions. The columns are FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4, and IG SEQUENCE. The rows include Carbons Count, Hydrogens Count, Nitrogens Count, Oxygens Count, Sulfur Count, Molecular Volume, Hydrophobicity, Charge, Isoelectric, and Sequence Length.

	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4	IG SEQUENCE
Carbons Count	98	57	99	15	161	69	42	545.8750
Hydrogens Count	178	85	147	21	249	98	67	838
Nitrogens Count	30	15	25	3	45	20	13	147.5000
Oxygens Count	37	17	26	6	60	21	14	171
Sulfur Count	1	0	0	0	1	0	0	3.6250
Molecular Volume	2.8158e+03	1.4829e+03	2.5478e+03	371.2000	4.3844e+03	1.7938e+03	1.1378e+03	1.4635e+04
Hydrophobicity	10.7000	-2.2750	-10.8500	-0.2750	-14.0250	-5.8250	-0.2750	-25.8375
Charge	0.9142	1.1268	0.1268	-0.0393	-1.0865	0.2934	-0.0375	1.7103
Isoelectric	8.5498	9.3007	7.5349	6.0977	4.3703	7.7062	6.3369	8.4532
Sequence Length	25	11	17	3	36	13	16	113.3750

Figure 3. A whole view of the graphical interface of the application which designed in order to provide the results of the approach described in this article.

Antisoma application has been developed in order to provide the information found with this new approach and concerning the antibodies of the species. It provides information about both the filtered and the unfiltered data in a fast and clearly legible way and it is intended to maintain a complete and proper database of immunoglobulins and the information concerning them and also help the relevant scientists in their work. Antisoma elaborates the database that has been created in the previous step. This user-friendly application interacts with the user with a graphical user's interface (GUI) and provides all the information needed. In the first of the five sections the user has the ability to choose between one or more from the available species from which he wants to display the antibodies and the information provided about them. Subsequently, in the second section, under the name "Filter Options", the user has to choose which of the FR and CDR regions of the antibody wants to be extracted from the database. Notice that in this section there are two buttons under the names of "IGV-Like" and "IGV-Like IMGT" which are used for extracting all the hypervariable regions after and before the filtering respectively (Figure 4). The "Data Set Table" section is being used for displaying the results and additionally gives the ability to the user to extract the information in .text, .xls and .mat file by clicking the appropriate buttons (Figure 5). After displaying the results as described in the previous step, a menu appears in the "Biostatistics Tools" section which gives the ability for the user to select between one or more from the immunoglobulins displayed before. By clicking the "General Features" button or the "Specific Features" button, features concerning the selected immunoglobulins based in the molecular operating environment application (MOE) or in 10 critical specific features respectively, in the final section under the name "Biostatistics Features" appears (Figure 6). Also here the ability to extract the information displayed is also provided in the same way as before (Figure 7).

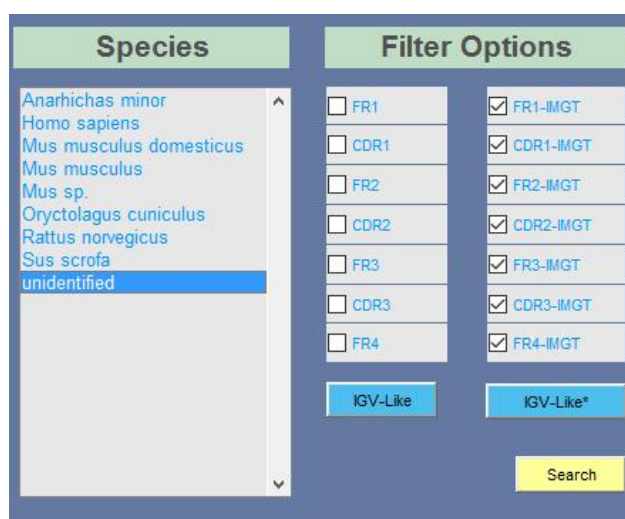


Figure 4. The first two main layouts of the graphical user's interface which gives the opportunity to the user to select which of the species and the FR-CDR regions wants to extract from the Antisoma application database.

Data Set Table												
	NAME	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4	SQ	REJECT_FR1	REJECT_CD...	I
1	AC172685	DIPITRITSIT...	IDLKKINVY	MHWYRQKP...	ISSNEN	VHEGGISEE...	ACWVE		DIPITRITSIT...	NOO	NOO	N
2	AY644517	SSNAKGAQ...	QDATY	IHLKQOEG...	YDVYYTKI	QFESGKKE...	AVWDK		SSNAKGAQ...	NOO	NOO	N
3	AY644518	ERVFORSV...	KSDVY	VHWFRQLE...	LALSKRDV	QWDSVLRG...	ATHGARQHD		ERVFORSV...	NOO	NOO	N
4	AY735449	QLTLVQTEV...	SKDFSKDY	HWYRQKPD...	VLDAPA	LNDLGGKKN...	AGWLLGSS...		QLTLVQTEV...	NOO	NOO	N
5	BC104586	AQRVTQDQ...	TQSHYN	IYWKQLPS...	QGS	SSQINARYG...	ALWPLRWG...		AQRVTQDQ...	NOO	NOO	N
6	BC113229	AQKVTQDQ...	TSWSYYN	IFWYKQLP...	QYS	EHGNARNG...	ALCLNSVGF...		AQKVTQDQ...	NOO	NOO	N
7	BC142026	AQKVTQDQ...	TSWNSYN	IFWYKQLP...	QNS	YSPNARDG...	ARNSQRRW...		AQKVTQDQ...	NOO	NOO	N
8	BC142356	AQKVTQDQ...	TSLSYR	LYWKQLP...	QGS	EATAWKDR...	ALOLLYSV...		AQKVTQDQ...	NOO	NOO	N
9	BC142390	AQKVTQDQ...	VSWSGYTHY	LYWKQLP...	QES	SGPNARNG...	AVWAPPYG...		AQKVTQDQ...	NOO	NOO	N
10	BC142393	AQKVTQDQ...	TNYWSYYN	IFWYKQLP...	QYS	EYGNARNG...	ALWASRWE...		AQKVTQDQ...	NOO	NOO	N
11	BC142412	SQKVTQDQ...	TSWGTYY	LYWKQLP...	QGS	LVTNAKKDR...	ALRDSAVG...		SQKVTQDQ...	NOO	NOO	N
12	BC142414	AQKVIQDQ...	TSWSSYS	IFWYKQLP...	QNS	DGRHSINVE...	ALWELDGL...		AQKVIQDQ...	NOO	NOO	N
13	BC147856	AQQTQVQ...	VSWSMDYYY	IFWYKQLP...	QYS	EYGNARNG...	ALFSRRRD...		AQQTQVQ...	NOO	NOO	N
14	BC147858	AQKVTQDQ...	IGWFTYYS	IFWYKQLP...	QNS	EHGNAGY...	GLRELSKR...		AQKVTQDQ...	NOO	NOO	N

Figure 5. The section of the application which displays the results in detail.

Biostatistics Tools

Import Options:

A1.

A2.

B.

C.

Figure 6. In this optional step it is given the ability to select which biostatistics features which follows the selected immunoglobulins wants the user to display.

Biostatistics features								
	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4	SQ
hydrophobicity	-2.3700	-2.6600	-6.0100	-1.2800	-6.3400	2.0600	0	-1.1055
AM1_dipole	261.3774	68.6648	155.0109	60.7592	384.1212	54.1041	0	227.6622
AM1_E	-1.1204e+06	-3.5478e+05	-7.0416e+05	-2.6124e+05	-1.5731e+06	-2.1577e+05	0	-9.0048e+05
AM1_Eele	-5.0041e+06	-1.6404e+06	-3.3513e+06	-1.0696e+06	-6.8509e+06	-9.6594e+05	0	-3.8149e+06
AM1_HF	-1.8092e+03	-395.4693	-622.8311	-725.5203	-3.0001e+03	-397.2729	0	-1.9328e+03
AM1_HOMO	-248.0941	-80.0486	-145.5876	-51.0281	-312.4930	-40.8330	0	-193.7955
AM1_IP	248.0941	80.0486	145.5876	51.0281	312.4930	40.8330	0	193.7955
AM1_LUMO	-3.8909	-3.0620	-8.0307	4.7352	15.0247	3.4402	0	6.4263
apol	507.3297	168.1301	332.5368	101.2424	669.1180	96.6497	0	382.4352
ASA	7.9003e+03	2.5439e+03	4.9485e+03	1.7100e+03	1.0765e+04	1.5003e+03	0	6.2507e+03
ASA+	4.7378e+03	1.6668e+03	3.2170e+03	991.9411	5.7662e+03	684.1055	0	3.5004e+03
ASA-	2.6651e+03	877.1462	1.6579e+03	718.0631	3.7979e+03	573.4074	0	2.2153e+03
ASA_H	3.3288e+03	917.4720	1.9321e+03	461.9567	4.3268e+03	680.4410	0	2.4421e+03

Figure 7. The section of the application which displays the biostatistics information related to the selected immunoglobulins.

3. Discussion

Antisoma has been created with the aim to filter, represent and export data obtained from the NCBI databases. The Database can be represented in its original form or filtered in a way that corresponds to the IMGT unique numbering system [33]. We believe that when it comes to immunoglobulin comparison it is very important that the regions of the IGs in question should be clearly defined and well documented. This way we provide the user with noise-free data concerning the standard lengths of immunoglobulin variable regions and domains. Also, we provide a way to examine the outlying IGs that do not consort with the IMGT unique numbering and get an undistracted view on them. Moreover, the user has access to a large collection of biostatistical metadata for selected IG sequences. All those features, described in detail on previous sections of this article, are represented in a user-friendly, easy-to-install application. Antisoma provides scientists and researchers with normalized data and state-of-the-art biostatistical metadata to help them produce more precise and efficient scientific results.

4. Conclusion

Better understanding of the detailed variable domains of antibodies will not only benefit the design of monoclonal antibodies and antibodies drug conjugates, but will also pave the way for the full understanding of its role in immune defence. In the presented work, a consolidated platform with correlated information about antibodies hypervariable regions and especially CDRs has been included. In the present application comprehensive information has been contained about all the available antibodies, including sequence, domains, regions and physicochemical properties. This application will be an important tool to support the development of novel therapies. Future goal of the application is the addition of more features such as the inclusion of comprehensive V-domain 3D structural analysis and the discovery of normalized representative forms of their conserved regions. Moreover, a future plan will be to separate the existing antibodies in groups in order to extract critical information related to antibody engineering, ADCs efficient design, modification of the antigens binding affinity and epitope identification of a given antibody.

Conflict of Interest

The authors declare that there is no conflict of interest.

References

1. Llewelyn MB, Hawkins RE, Russell SJ (1992) Discovery of antibodies. *BMJ* 305: 1269-1272.

2. Silverstein AM (1999) Paul Ehrlich's passion: the origins of his receptor immunology. *Cell Immunol* 194: 213-221.
3. Kabat EA, Furth J (1940) Chemical and Immunological Studies on the Agent Producing Leukosis and Sarcoma of Fowls. *J Exp Med* 71: 55-70.
4. Ribatti D (2015) Edelman's view on the discovery of antibodies. *Immunol Lett* 164: 72-75.
5. Liu JK (2014) The history of monoclonal antibody development - Progress, remaining challenges and future innovations. *Ann Med Surg (Lond.)* 3: 113-116.
6. Scott AM, Wolchok JD, Old LJ (2012) Antibody therapy of cancer. *Nat Rev Cancer* 12: 278-287.
7. Novotný J, Bruccoleri R, Newell J, et al. (1983) Molecular anatomy of the antibody binding site. *J Biol Chem* 258: 14433-14437.
8. Coleclough C (1983) Chance, necessity and antibody gene dynamics. *Nature* 303: 23-26.
9. Chothia C, Novotný J, Bruccoleri R, et al. (1985) Domain association in immunoglobulin molecules. The packing of variable domains. *J Mol Biol* 186: 651-663.
10. Maltezos A, Platis D, Vlachakis D, et al. (2014) Design, synthesis and application of benzylsulfonate biomimetic affinity adsorbents for monoclonal antibody purification from transgenic corn. *J Mol Recognit* 27: 19-31.
11. Chothia C, Lesk AM, Tramontano A, et al. (1989) Conformations of immunoglobulin hypervariable regions. *Nature* 342: 877-883.
12. Kovtun YV, Goldmacher VS (2007) Cell killing by antibody-drug conjugates. *Cancer Lett* 255: 232-240.
13. Sela-Culang I, Kunik V, Ofran Y (2013) The structural basis of antibody-antigen recognition. *Front Immunol* 4: 302.
14. Shirai H, Prades C, Vita R, et al. (2014) Antibody informatics for drug discovery. *Biochim Biophys Acta* 1844: 2002-2015.
15. Winter G, Milstein C (1991) Man-made antibodies. *Nature* 349: 293-299.
16. Fajardo-Ramirez OR, Ascacio-Martinez JA, Licea-Navarro AF, et al. (2015) Technological Evolution in the Development of Therapeutic Antibodies. *Rev Invest Clin* 67: 158-169.
17. Wang J, Bu D, Zhu X (2007) Immunoglobulin variable region gene analysis to the autoantibody-secreting B cells from tumors in association with paraneoplastic autoimmune multiorgan syndrome. *Int J Dermatol* 46: 1146-1154.
18. Peters C, Brown S (2015) Antibody-drug conjugates as novel anti-cancer chemotherapeutics. *Biosci Rep* 35: e00225.
19. Zhang N, Deng H, Fan X, et al. (2015) Dysfunctional antibodies in the tumor microenvironment associate with impaired anticancer immunity. *Clin Cancer Res* 21: 5380-5390.
20. Campa MJ, Gottlin EB, Bushey RT, et al. (2015) Complement Factor H Antibodies from Lung Cancer Patients Induce Complement Dependent Lysis of Tumor Cells, Suggesting a Novel Immunotherapeutic Strategy. *Cancer Immunol Res* 3: 1325-1332.

21. Zitvogel L, Kroemer G (2015) Cancer: Antibodies regulate antitumour immunity. *Nature* 521: 35-37.
22. Papageorgiou L, Cuong NT, Vlachakis D (2016) Antibodies as stratagems against cancer. *Mol Biosyst* 12: 2047-2055.
23. Kumar S, King L E, Clark T H, et al. (2015) Antibody-drug conjugates nonclinical support: from early to late nonclinical bioanalysis using ligand-binding assays. *Bioanalysis* 7: 1605-1617.
24. Gorovits B (2015) Bioanalysis of antibody-drug conjugates. *Bioanalysis* 7: 1559-1560.
25. Merten H, Brandl F, Plückthun A, et al. (2015) Antibody-Drug Conjugates for Tumor Targeting-Novel Conjugation Chemistries and the Promise of non-IgG Binding Proteins. *Bioconjug Chem* 26: 2176-2185.
26. Hamilton GS (2015) Antibody-drug conjugates for cancer therapy: The technological and regulatory challenges of developing drug-biologic hybrids. *Biologicals* 43: 318-332.
27. Johnson G, Wu TT (2000) Kabat database and its applications: 30 years after the first variability plot. *Nucleic Acids Res* 28: 214-218.
28. Abhinandan KR, Martin AC (2008) Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. *Mol Immunol* 45: 3832-3839.
29. Lefranc MP (2014) Immunoinformatics of the V, C, and G domains: IMGT(R) definitive system for IG, TR and IgSF, MH, and MhSF. *Methods Mol Biol* 1184: 59-107.
30. Lefranc MP (2014) Antibody Informatics: IMGT, the International ImMunoGeneTics Information System. *Microbiol Spectr*: 363-379.
31. Lefranc MP (2011) IMGT unique numbering for the variable (V), constant (C), and groove (G) domains of IG, TR, MH, IgSF, and MhSF. *Cold Spring Harb Protoc* 2011: 633-642.
32. Lefranc MP, Giudicelli V, Duroux P, et al. (2014) IMGT®, the international ImMunoGeneTics information system® 25 years on. *Nucleic Acids Res* 43: D413-D422.
33. Lefranc MP, Pommi éC, Ruiz M, et al. (2003) IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev Comp Immunol* 27: 55-77.
34. Lefranc MP (2014) Immunoglobulins: 25 years of immunoinformatics and IMGT-ONTOLOGY. *Biomolecules* 4: 1102-1139.
35. Alamyar E, Giudicelli V, Duroux P, et al. (2014) Antibody V and C domain sequence, structure, and interaction analysis with special reference to IMGT(R). *Methods Mol Biol* 1131: 337-381.
36. Tinberg CE, Khare SD, Dou J, et al. (2013) Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* 501: 212-216.
37. Benson DA, Cavanaugh M, Clark K, et al. (2017) GenBank. *Nucleic Acids Res* 45: D37-D42.
38. Apweiler R, Bairoch A, Wu CH, et al. (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45: D158-D169.
39. Giudicelli V, Duroux P, Ginestoux C, et al. (2006) IMGT/LIGM-DB, the IMGT comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. *Nucleic Acids Res* 34: D781-D784.

40. Sobie EA (2011) An introduction to MATLAB. *Sci Signal* 4: tr7.



AIMS Press

© 2017 the authors, licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)