



SOFTWARE

Open Access

# IMGT/Collier-de-Perles: a two-dimensional visualization tool for amino acid domain sequences

Dimitrios Vlachakis<sup>1,2</sup>, Christos Feidakis<sup>1</sup>, Vasileios Megalooikonomou<sup>2</sup> and Sophia Kossida<sup>1\*</sup>

\* Correspondence:

skossida@bioacademy.gr

<sup>1</sup>Bioinformatics & Medical Informatics Team, Biomedical Research Foundation, Academy of Athens, Soranou Efessiou 4, Athens 11527, Greece

Full list of author information is available at the end of the article

## Abstract

IMGT/Collier-de-Perles is a tool that allows the user to analyze and draw two-dimensional graphical representations (or IMGT Collier de Perles) of protein domains (e.g., hydrophathy plots). The IMGT/Collier-de-Perles specializes in the area of immunoglobulins (IG) or antibodies, T cell receptors (TR) and major histocompatibility (MH) of human and other vertebrate species as well as other proteins of the immunoglobulin superfamily (IgSF) and of the major histocompatibility superfamily (MhSF) and related proteins of the immune system of vertebrates and invertebrates.

**Keywords:** IMGT/Collier-de-Perles, IMGT Collier de Perles, Graphical representation, Protein domains

## Introduction

Amino acids can be defined and classified in a number of ways, depending on the perspective they are being examined from each time. Thereby, they can be categorized according to the functional groups of their side chains, which determine their physico-chemical characteristics [1].

Taking into account the importance of proteins, made of amino acids, as a structural component of all living organisms, the significance of a method or tool that could seamlessly manipulate this data would be extremely practical. Indeed, scientists have been using computational tools that enable them to compare and examine amino acid sequences in a number of ways.

Among the different classes of amino acid properties, hydrophobicity determines how strongly an amino acid is attracted to or repelled by water. A series of different hydrophobicity scales have been developed [2-8]. The higher the index value is in a scale, the more hydrophobic is the amino acid. Differences between the scales mainly depend on the method or on the algorithm used to measure or to define hydrophobicity [6,9-12]. Hydrophobicity scales are commonly used to predict the leader region (or signal peptides) or the transmembrane region of proteins. When measuring sequential amino acids of a protein, fluctuations in value indicate protein hydrophobic regions potentially located inside the membrane lipid layer [13] or contributing to the hydrophobic core of a protein [2]. Hydrophathy and other amino acid properties are keys for a better understanding of protein interactions and domain structures.

### **The IMGT/Collier-de-Perles tool**

The IMGT/Collier-de-Perles [14] tool was created by LIGM (Université Montpellier 2, CNRS) and is part of IMGT<sup>®</sup>, the international ImMunoGeneTics information system<sup>®</sup> [15,16] (IMGT<sup>®</sup>, <http://www.imgt.org>), which is acknowledged as the global reference in immunogenetics and immunoinformatics.

IMGT/Collier-de-Perles can provide upon selection three types of displays: the hydropathy plot with 3 classes (hydrophobic, neutral, hydrophilic), the volume plot with 5 classes, and the physicochemical plot, which is the most informative one, with eleven IMGT physicochemical classes (which were defined taking into account hydropathy, volume and chemical characteristics properties) [1,17] (Figure 1A). Eleven IMGT physicochemical classes of the 20 common amino acids have been defined by the physicochemical properties of their side chains [17]. These standardized classes are used in the IMGT/Collier-de-Perles tool.

IMGT Colliers de Perles can currently be drawn for three domain types: variable (V) domain and constant (C) domain of immunoglobulins (IG) or antibodies and T cell receptors (TR) and immunoglobulin superfamilies (IgSF) proteins other than IG and TR, and groove domain (G) of the major histocompatibility (MH) and MH superfamily (MhSF) other than MH [18-21]. In order for an IMGT Collier de Perles to be created, each sequence has to be gapped according to the IMGT unique numbering [22-25], using IMGT/DomainGapAlign [26-28].

IMGT/DomainGapAlign allows the creation of gaps in the user's V, C or G domain amino acid sequence, by aligning the user sequence to the corresponding IMGT domain reference directory and identifies, for the IG or TR V domain, the closest germline V-REGION and J-REGION, and for all other cases (V domain of IgSF other than IG or TR, C domain and G domain) the closest V, C or G domain of the reference gene and/or allele, respectively, and finally obtains the IMGT Collier de Perles. Amino acids which differ from the closest reference sequence are highlighted in the IMGT Collier de Perles (pink border, online) and the IMGT amino acid change characteristics detailed in accompanying tables [26-28].

The resulting IMGT Colliers de Perles (Figure 1B) help us determine which amino acids are important for the 3D structural configuration and, for the IG and TR V domain, delineate the standardized framework regions (FR-IMGT) (formed by the nine antiparallel beta strands) and complementarity determining regions (CDR-IMGT) (formed by the three loops binding the antigen). The length of the strands, loops and turns in IMGT Colliers de Perles provide critical information in the characterization of each V, C or G domain [21,25].

### **How can we use IMGT Colliers de Perles?**

The first among numerous features of the IMGT Collier de Perles is based on the way that domains of the antibody or T cell receptor are characterized. Each domain is described by the length of its loops, strands and turns, and helix (for G domain) [21,25]. This way, the usual but confusing distinction made in the literature and generalist databases, between C1, C2 and I1 sets is often inappropriate upon the absence of structural data and can be ignored [23].

Unlike generalist databases such as UniProt/Swiss-Prot, IMGT standardization defines the different domains by comparing amino acid or cDNA sequences with genomic

**A**

'Volume' classes in Å <sup>3</sup>	Hydrophobic			'Hydropathy' classes		
	Neutral			Hydrophilic		
Very large 189-228	F	L	W	Y		
Large 162-174	I	L	M			
Medium 138-154	V					
Small 108-117			C			
Very small 60-90	A			T	S	

Aliphatic     Sulfur     Hydroxyl     Basic     Acidic     Amide

Nonpolar
Uncharged     Charged     Uncharged

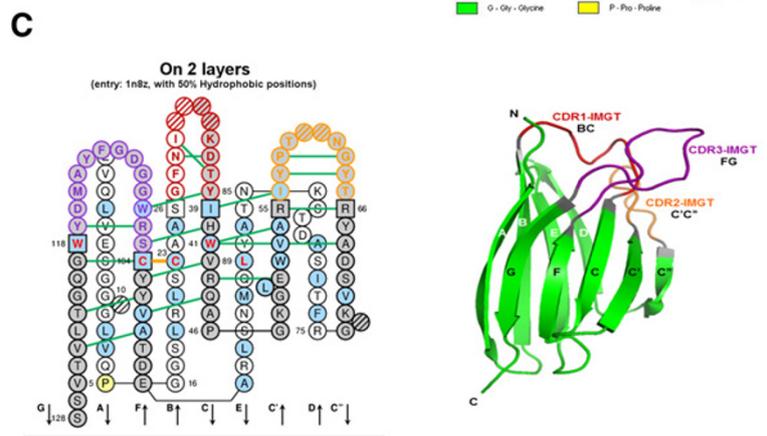
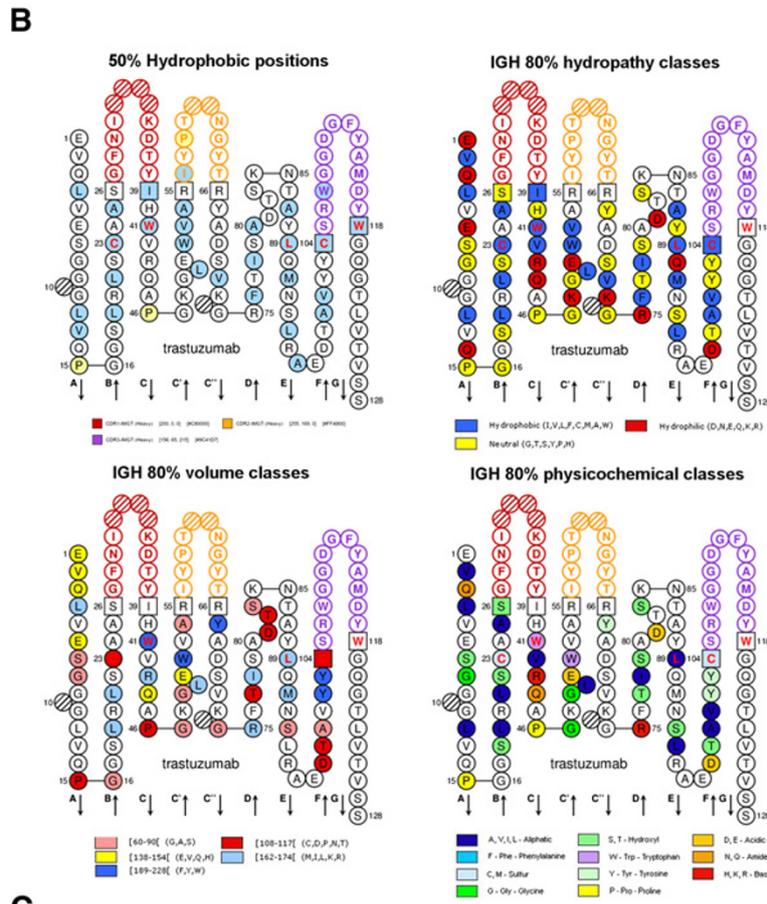


Figure 1 (See legend on next page.)

(See figure on previous page.)

**Figure 1 Amino acid properties and the IMGT Collier de Perles 2D hydropathy plot.** (A): The 11 IMGT physicochemical classes of the 20 common amino acids [1]. (B): IMGT Colliers de Perles of a V domain of an IG or antibody. IMGT Collier de Perles on one layer, obtained by the IMGT/Collier-de-Perles online tool [14] at the IMGT Home page, <http://www.imgt.org>. Four options are shown for the V domain of the heavy domain (VH) of trastuzumab, a therapeutic antibody: '50% hydrophobic positions' (with that option, an amino acid at a given position is colored in blue if an hydrophobic amino acid is found at that position in 50% or more of the sequences used for establishing the statistical profile [17], 'IGH 80% hydropathy classes', 'IGH 80% volume classes', 'IGH 80% physicochemical classes'. The CDR-IMGT lengths are [8.8.13]. The five conserved amino acid positions of a V domain 23, 41, 89, 104 and 118 (1st-CYS C23, CONSERVED-TRP W41, hydrophobic 89, 2nd-CYS C104, J-PHE or J-TRP 118, here W118) appear with red and bold letters. The six anchors (positions 26 and 39, 55 and 66, 104 and 118) which belong to the framework regions (FR-IMGT) and support the complementarity determining regions (CDR-IMGT) are in squares. (C): IMGT Collier de Perles on 2 layers and with hydrogen bonds (green lines) of the trastuzumab VH, provided by IMGT/3Dstructure-DB [26,34] (PDB code 1n8z). These hydrogen bonds are from the experimental 3D structures and are displayed by the IMGT/Collier-de-Perles tool integrated in IMGT/3Dstructure-DB [26,34]. The 3D structure representation of a V domain with its strands and loops, based on the IMGT Collier de Perles delimitations [21,22,25], is also shown.

sequences, therefore identifying the splicing sites and giving a more accurate delimitation of the existing domains [23].

Another strong feature of the IMGT Collier de Perles is that it produces a standardized graphical representation and allows to visualize and to localize the differences between domains of proteins whatever the species even when 3D data are unavailable [29]. This can be a great tool for molecular engineers. Antibody humanization, particularly, is greatly benefited by an interface that seamlessly displays and compares FR-IMGT and CDR-IMGT among several species [30-33].

IMGT Collier de Perles can also be used to compare a given amino acid sequence against an IMGT reference sequence, in order to facilitate the identification of potential immunogenic residues at certain positions of humanized antibodies or to assess the immunogenicity of therapeutic antibodies (Figure 1B). The reference sequence is in essence a statistical profile created for the human IG heavy, kappa and lambda expressed variable domains sets and is based, as the physicochemical plots, on the description of the 11 IMGT amino acid physicochemical classes (that include hydropathy, volume and chemical characteristics of the 20 common amino acids) [17].

Besides being a 2D visualization tool, IMGT/Collier-de-Perles can also take advantage of 3D structures when those are available [26,34], by displaying IMGT Colliers de Perles on two layers with hydrogen bonds between amino acids of V or C domains of the antibody (Figure 1C). In Figure 1C, the FR-IMGT is made up of 9 strands (arrows) and turns is in green and the 3 CDR-IMGT are in red, orange and purple, respectively (<http://www.imgt.org>, The IMGT Biotechnology page, Antibody engineering, FR-IMGT and CDR-IMGT).

Additional information such as atom contact types and categories can be provided for each amino acid separately, by clicking on each amino acid in the IMGT Collier de Perles.

## Conclusion

Based on the IMGT unique numbering concept generated from the NUMEROTATION axiom of IMGT-ONTOLOGY [16], the IMGT/Collier-de-Perles tool is a very friendly tool which allows users to create visual representation of their own amino acid

sequences for V, C and G domains [21,25]. The tool can be used on its own or as an output functionality of IMGT/DomainGapAlign [26-28]. It has also been integrated in IMGT/3Dstructure-DB (<http://www.imgt.org>), the IMGT three-dimensional (3D) structure database [26,34]. Thus the users can compare their own IMGT Colliers de Perles with those provided in the database for analysis of interactions, e.g., those of the V domains of IG and TR in complex with their antigen (query on IG/Ag and TR/pMH) or those of the C or G domains of the FcR in complex with the IG Fc. IMGT Colliers de Perles provide a great help for understanding relations between sequences and structures in the design of therapeutic monoclonal antibodies (antibody engineering and humanization) [30,31,33] with their functional properties (specificity, affinity, immunogenicity, allotype expression [32], etc.), and more generally for characterizing the V, C and G domains of the proteins belonging to the IgSF and MhSF superfamilies of all vertebrates and invertebrates.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

DV and SK conceived, coordinated, supervised and designed the study. DV, CF, VM and SK carried out the study and drafted the manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Thales. Investing in knowledge society through the European Social Fund.

#### Author details

<sup>1</sup>Bioinformatics & Medical Informatics Team, Biomedical Research Foundation, Academy of Athens, Soranou Efessiou 4, Athens 11527, Greece. <sup>2</sup>Computer Engineering and Informatics Department, School of Engineering University of Patras, 26500, Patras, Greece.

Received: 28 January 2013 Accepted: 16 February 2013

Published: 21 February 2013

#### References

1. IMGT classes of the 20 common amino acids. In *IMGT Education > IMGT Aide-mémoire > Amino acids*; <http://www.imgt.org> (11 November 2012, date last accessed).
2. Janin J: Surface and inside volumes in a globular protein. *Nature* 1979, **277**:491–492.
3. Wolfenden R, Andersson L, Cullis P, et al: Affinities of amino acid side chains for solvent water. *Biochemistry* 1981, **20**:849–855.
4. Kyte J, Doolittle RF: A simple method for displaying the hydrophobic character of a protein. *J Mol Biol* 1982, **157**(1):105–132.
5. Rose G, Geselowitz A, Lesser G, et al: Hydrophobicity of amino acid residues in globular proteins. *Science* 1985, **229**:834–838.
6. Engelman DM, Steitz TA, Goldman A: Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu Rev Biophys Chem* 1986, **15**:321–353.
7. Cornette JL, Cease KB, Margalit H, et al: Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *J Mol Biol* 1987, **195**(3):659–685.
8. Wimley WC, White SH: Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* 1996, **3**(10):842–848.
9. Charton M, Charton BJ: The structural dependence of amino acid hydrophobicity parameters. *J Theor Biol* 1982, **99**:629–644.
10. Eisenberg D: Three-dimensional structure of membrane and surface proteins. *Ann Rev Biochem* 1984, **53**:595–623.
11. Rose GD, Wolfenden R: Hydrogen bonding, hydrophobicity, packing, and protein folding. *Annu Rev Biophys Biomol Struct* 1993, **22**:381–415.
12. Biswas KM, DeVido DR, Dorsey JG: Evaluation of methods for measuring amino acid hydrophobicities and interactions. *J Chromatogr A* 2003, **1000**:637–655.
13. Clements JD, Martin RE: Identification of novel membrane proteins by searching for patterns in hydrophathy profiles. *Eur J Biochem* 2002, **269**:2101–2107.
14. Ehrenmann F, Giudicelli V, Duroux P, et al: IMGT/Collier de Perles: IMGT Standardized Representation of Domains (IG, TR, and IgSF Variable and Constant Domains, MH and MhSF Groove Domains). *Cold Spring Harb Protoc* 2011, **6**(0):726–736.

15. Lefranc M-P, Giudicelli V, Ginestoux C, et al: **IMGT<sup>®</sup>**, the international ImMunoGeneTics information system<sup>®</sup>. *Nucl Acids Res* 2009, **37**:D1006–D1012.
16. Lefranc M-P: **IMGT, the International ImMunoGeneTics Information System**. *Cold Spring Harb Protoc* 2011, **6**:595–603.
17. Pommié C, Levadoux S, Sabatier R, et al: **IMGT standardized criteria for statistical analysis of immunoglobulin V-REGION amino acid properties**. *J Mol Recognit* 2004, **17**:17–32.
18. Ruiz M, Lefranc M-P: **IMGT gene identification and Colliers de Perles of human immunoglobulin with known 3D structures**. *Immunogenetics* 2002, **53**:857–883.
19. Kaas Q, Lefranc M-P: **IMGT Colliers de Perles: standardized sequence-structure representations of the IgSF and MhcSF superfamily domains**. *Curr Bioinforma* 2007, **2**:21–30.
20. Kaas Q, Ehrenmann F, Lefranc M-P: **IG, TR, MHC, IgSF and MhcSF: what do we learn from the IMGT Colliers de Perles?** *Brief Funct Genomic Proteomic* 2007, **6**:253–264.
21. Lefranc M-P: **IMGT Collier de Perles for the Variable (V), Constant (C), and Groove (G) Domains of IG, TR, MH, IgSF, and MhSF**. *Cold Spring Harb Protoc* 2011, **6**:643–651.
22. Lefranc M-P, Pommié C, Ruiz M, et al: **IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains**. *Dev Comp Immunol* 2003, **27**:55–77.
23. Lefranc M-P, Pommié C, Kaas Q, et al: **IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains**. *Dev Comp Immunol* 2005, **29**:185–203.
24. Lefranc M-P, Duprat E, Kaas Q, et al: **IMGT unique numbering for MHC groove G-DOMAIN and MHC superfamily (MhcSF) G-LIKE-DOMAIN**. *Dev Comp Immunol* 2005, **29**:917–938.
25. Lefranc M-P: **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, **6**:633–642.
26. Ehrenmann F, Kaas Q, Lefranc M-P: **IMGT/3Dstructure-DB and IMGT/ DomainGapAlign: A database and a tool for immunoglobulins or antibodies, T cell receptors, MHC, IgSF and MhcSF**. *Nucl Acids Res* 2010, **38**:D301–D307.
27. Ehrenmann F, Lefranc M-P: **IMGT/DomainGapAlign: IMGT Standardized Analysis of Amino Acid Sequences of Variable, Constant, and Groove Domains (IG, TR, MH, IgSF, MhSF)**. *Cold Spring Harb Protoc* 2011, **6**:737–749.
28. Ehrenmann F, Lefranc M-P: **IMGT/DomainGapAlign: the IMGT<sup>®</sup> tool for the analysis of IG, TR, MHC, IgSF and MhcSF domain amino acid polymorphism**. In *Immunogenetics, chap 33, Methods Mol Biol, Volume 882*. Edited by Tait B, Christiansen F. New York, USA: Humana Press, Springer; 2012:605–633.
29. Garapati VP, Lefranc M-P: **IMGT Colliers de Perles and IgSF domain standardization for T cell costimulatory activatory (CD28, ICOS) and inhibitory (CTLA4, PDCD1 and BTLA) receptors**. *Dev Comp Immunol* 2007, **31**:1050–1072.
30. Lefranc M-P: **Antibody databases and tools: The IMGT<sup>®</sup> experience**. In *Therapeutic monoclonal antibodies: from Bench to Clinic, Volume chap 4*. Edited by An Z. Hoboken, New Jersey, USA: John Wiley & Sons, Inc; 2009:91–114.
31. Ehrenmann F, Duroux P, Giudicelli V, et al: **Standardized sequence and structure analysis of antibody using IMGT<sup>®</sup>**. In *Antibody engineering, Vol. 2, Volume chap. 2*. Edited by Kontermann R, Dübel S. Berlin Heidelberg: Springer-Verlag; 2010:11–31.
32. Lefranc M-P, Lefranc G: **Human Gm, Km and Am allotypes and their molecular characterization: a remarkable demonstration of polymorphism**. In *Immunogenetics, chap 34, Methods Mol Biol, Volume 882*. Edited by Tait B, Christiansen F. New York, USA: Humana Press, Springer; 2012:635–680.
33. Lefranc M-P, Ehrenmann F, Ginestoux C, et al: **Use of IMGT<sup>®</sup> databases and tools for antibody engineering and humanization**. In: P. Chames (Ed.), *Antibody engineering, chap 1*, Humana Press, Springer, New York, USA. *Methods Mol Biol* 2012, **907**:3–37.
34. Ehrenmann F, Lefranc M-P: **IMGT/3Dstructure-DB: Querying the IMGT Database for 3D Structures in Immunology and Immunoinformatics (IG or Antibodies, TR, MH, RPI, and FPIA)**. *Cold Spring Harb Protoc* 2011, **6**:750–761.

doi:10.1186/1742-4682-10-14

**Cite this article as:** Vlachakis et al: **IMGT/Collier-de-Perles: a two-dimensional visualization tool for amino acid domain sequences**. *Theoretical Biology and Medical Modelling* 2013 **10**:14.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at  
www.biomedcentral.com/submit

