Brief Report

Stereotyped patterns of B-cell receptor in splenic marginal zone lymphoma

Silvia Zibellini,[1] Daniela Capello,[2] Francesco Forconi,[3] Paolo Marcatili,[4] Davide Rossi,[2] Sara Rattotti,[1] Silvia Franceschetti,[2] Elisa Sozzi,[3] Emanuele Cencini,[3] Roberto Marasca,[5] Luca Baldini,[6] Alessandra Tucci,[7] Francesco Bertoni,[8] Francesco Passamonti,[1] Ester Orlandi,[1] Marzia Varettoni,[1] Michele Merli,[1] Silvia Rizzi,[1] Valter Gattei,[9] Anna Tramontano,[4] Marco Paulli,[9] Gianluca Gaidano,[2] and Luca Arcaini[1]

[1]Division of Hematology, Department of Oncohematology, Fondazione IRCCS Policlinico San Matteo, University of Pavia, Italy; [2]Division of Hematology, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; [3]Division of Hematology, University of Siena, Siena, Italy; [4]Department of Biochemical Sciences and Istituto Pasteur Fondazione Cenci Bolognetti, Sapienza University, Roma, Italy; [5]Division of Hematology, University of Modena and Reggio Emilia, Modena, Italy; [6]Division of Hematology, Fondazione IRCCS Ospedale Maggiore Policlinico, University of Milano, Milano, Italy; [7]Division of Hematology, Spedali Civili, Brescia, Italy; [8]Laboratory of Experimental Oncology and Lymphoma Unit, Oncology Institute of Southern Switzerland (IOSI), Bellinzona, Switzerland; [9]Clinical and Experimental Onco-Hematology Unit, Centro di Riferimento Oncologico, IRCCS, Aviano; [10]Department of Pathology, Fondazione IRCCS Policlinico San Matteo, University of Pavia, Italy

ABSTRACT

Antigen stimulation may be important for splenic marginal zone lymphoma pathogenesis. To address this hypothesis, the occurrence of stereotyped B-cell receptors was investigated in 133 SMZL (26 HCV+) compared with 4,414 HCDR3 sequences from public databases. Sixteen SMZL (12%) showed stereotyped BCR; 7 of 86 (8%) SMZL sequences retrieved from public databases also belonged to stereotyped HCDR3 subsets. Three categories of subsets were identified: i) "*SMZL-specific subsets*" (n=5), composed only of 12 SMZL (9 HCV- from our series); ii) "*Non-Hodgkin's lymphoma-like subsets*" (n=5), comprising 5 SMZL (4 from our series) clustering with other indolent lymphomas; iii) "*CLL-like subsets*" (n=6), comprising 6 SMZL (3 from our series) that belonged to known CLL subsets (n=4) or clustered with public CLL sequences. Immunoglobulin 3D modeling of 3 subsets revealed similarities in antigen binding regions not limited to HCDR3. Overall, data suggest that the pathogenesis of splenic marginal zone lymphoma may involve also HCV-unrelated epitopes or an antigenic trigger common to other indolent lymphomas.

Key words: splenic marginal zone lymphoma, B cell receptor, immunoglobulin genes, hepatitis C virus.

Introduction

Splenic marginal zone lymphoma (SMZL) is recognized by the WHO classification as an individual entity.[1,2] Based on the association with HCV infection and autoimmunity,[3] it is conceivable that antigen stimulation may play a role in SMZL.[4,5] Though several studies investigated usage of immunoglobulin heavy chain variable (IGHV) genes in SMZL,[5-9] and two SMZL cases with highly similar HCDR3 have been previously reported,[10] at present no definitive immunogenetic clues for antigen stimulation have emerged.

In other B-cell malignancies, namely chronic lymphocytic leukemia (CLL), non-random combinations of specific IGHV-D-J genes lead to stereotyped complementarity determining region 3 (HCDR3) of the B-cell receptor (BCR) in a significant fraction of cases.[11,12] Occurrence of stereotyped HCDR3 in CLL is regarded as an indication of antigen stimulation during disease development.[13]

The purpose of the present study was to investigate the occurrence and patterns of stereotyped HCDR3 in a large panel of SMZL.

Design and Methods

Patients

The study was based on a multi-institutional series of 133 SMZL cases. Diagnosis of SMZL was based on criteria proposed by the WHO classification[1] and by Matutes *et al.*[2] (with spleen histology in 40

patients and with bone marrow histology coupled with flow cytometry in 93). Patients provided informed consent in accordance with local institutional review board requirements (IRB) and the Declaration of Helsinki. The study was approved by the IRB of the Fondazione IRCCS Policlinico San Matteo, Pavia, Italy.

Analysis of IGHV-D-J, IGKV-J and IGLV-J rearrangements in SMZL

Mononuclear cells were obtained from bone marrow (n=62 cases), peripheral blood (n=44), spleen tissue (n=25) or locoregional lymph node (n=2). IGHV-D-J, IGKV-J and IGLV-J rearrangements were amplified and directly sequenced from cDNA as previously reported.[9,14] Sequences were aligned to ImMunoGeneTics sequence directory using the IMGT V-QUEST analysis software[15] to obtain IGV, IGD, IGJ gene usage, mutational profile, functional status and CDR3 amino acid (AA) sequence.

Analysis of IGHV-D-J rearrangements and identification of SMZL with stereotyped HCDR3

Using the multiple sequence alignment software ClustalX (2.0),[16] HCDR3 AA sequences from 133 SMZL were aligned to each other and to 4,414 sequences derived from B-cell neoplasia entities and retrieved from in-house and public databases (*Online Supplementary Table S1*), regardless of usage of different IGHV genes. Subsets of stereotyped HCDR3 were defined according to the criteria proposed by Messmer[17] and Stamatopoulos.[11] AA differences at the same HCDR3 position in cases within a subset were also evaluated (volume, hydropathy index, and chemical characteristics). Subsets not previously reported have been assigned a number preceded by N ("Novel"). A difference in HCDR3 length of greater than 3 AA was not allowed in the same cluster. Nomenclature of previously reported subsets was made according to Stamatopoulos[11] and Murray.[12] Subsets composed of 3 or more cases have been defined as "confirmed".[11] Subsets composed of 2 cases were defined as "provisional" and were characterized by identical IGHV/IGHD/IGHJ or IGHD/IGHJ genes, shared junctional residues and/or restricted light-chain CDR3.[18]

Antibody modeling

The PIGS web-server[19] was used to model the immunoglobulins (Igs) for which both light and heavy chains were sequenced. In 3 cases (cases 3926, 4625 and PV3) in which no suitable template for the H3 region was found, we used the Modeller[20] *ab-initio* loop modeling procedure to build 10 different H3 loop conformations for each Ig and then, among these, we selected the centroid. Surface potentials were derived using the APBS software[21] with standard parameters. All pictures where generated using the PyMol software (*http://www.pymol.org*).

Analysis of clinical data

The χ^2 test and Fisher's exact test were used to compare groups. Overall survival (OS) and progression free survival (PFS) were defined according to current criteria.[22] Survival analysis was performed by the Kaplan-Meier method, using Gehan–Wilcoxon statistics to test for significant associations. Cox's analysis was used to build a multivariate model. Analyses were carried out using Stata SE 9 (StataCorp LP), Statistica 8 (StatSoft Inc.) and Microsoft Excel 2000.

Results and Discussion

The clinical features of the 133 SMZL are summarized in the *Online Supplementary Table S2*. Median age at diag-

Table 1. *IGHV*, *IGHD* and *IGHJ* gene usage in 133 SMZL patients according to IGHV identity.

IGHV gene	N. of sequences	Identity <98% (n=93)	Identity ≥98% (n=40)
1-18	3	1	2
1-2	26	15	11
1-3	1	1	0
1-69	10	7	3
2-70	1	1	0
3-11	2	2	0
3-21	4	0	4
3-23	24	22	2
3-30	8	6	2
3-30-3	1	1	0
3-33	5	3	2
3-48	5	4	1
3-49	1	1	0
3-66	2	2	0
3-7	5	4	1
3-72	1	1	0
3-73	1	1	0
3-74	6	4	2
3-9	1	1	0
4-30-4	1	0	1
4-34	10	5	5
4-39	2	2	0
4-4	2	2	0
4-59	5	3	2
4-61	3	2	1
5-51	2	2	0
5-a	1	0	1
IGHD gene			
1-1	1	1	0
1-26	3	2	1
1-7	2	1	1
2-15	3	2	1
2-2	7	5	2
2-21	3	2	1
2-8	5	2	3
3-10	6	5	1
3-16	4	1	3
3-22	15	13	2
3-3	24	15	9
3-9	5	5	0
4-17	3	2	1
4-23	5	5	0
4-4	1	1	0
5-12	4	2	2
5-24	3	2	1
5-5	4	1	3
6-13	5	3	2
6-19	7	5	2
6-25	1	0	1
6-6	5	2	3
NA	17	16	1
IGHJ gene			
1	2	1	1
3	19	15	4
4	50	36	14
5	26	13	13
6	36	28	8

NA: not assignable.

Table 2. SMZL-specific stereotyped HCDR3 subsets formed by HCV-negative SMZL cases from our cohort (complete dataset of stereotyped HCDR3 subsets in SMZL are reported in the *Online Supplementary Table S3*).

| SUBSET	Average intra-subset identity (%)	SMZL case	HCDR3 AA sequence (IMGT)		IGHV reading frame	IGHD	IGHD	IGHJ	% IDENTITY to germline *IGHV* gene	IGLV	IGLJ
N1	72.5										
		3926	C	A R G P R I T I I G V V I - R G R G G G F D Y	W	IGHV1-2*04	IGHD3-3*01	3	IGHJ4*02	96,18	IGKV3-20*01	IGKJ4*01
		4625	C	A R G P R I S M I G V V I - N G R G G A F D Y	W	IGHV1-2*04	IGHD3-3*02	3	IGHJ4*02	96,72	IGKV3-15*01	IGKJ3*01
		PV3	C	A R G D R I T I F G V V I G N S R H G G F D Y	W	IGHV1-2*04	IGHD3-3*01	3	IGHJ4*02	98,61	IGKV3-11*01	IGJK5*01
				* * * * * : : : * * * * . . * * . * * *								
N3	75										
		PV25	C	A K D Y D Y L Y Y Y Y Y Y M D V	W	IGHV3-23*04	IGHD5-12*01	3	IGHJ6*03	95,83	IGLV1-40*01	IGLJ2/3
		PV52	C	A K D Q D D D Y Y Y Y Y G M D V	W	IGHV3-23*04	na		IGHJ6*02	95,83	IGLV1-40*01	IGLJ2/3
				* * * * * * * * * * * *								
N4	70.6										
		PV19	C	A K D H N C S T Y Y Y Y Y Y Y D G	W	IGHV3-23*04	IGHD3-22*01	2	IGHJ6*03	95,83	IGLV1-44*01	IGLJ3*02
		PV72	C	A K D H N G N D Y Y Y Y Y G C D G	W	IGHV3-23*04	IGHD3-22*01	2	IGHJ6*04	95,14	IGLV1-47*01	IGLJ1*01
				* * * * * . * * * * * * *								
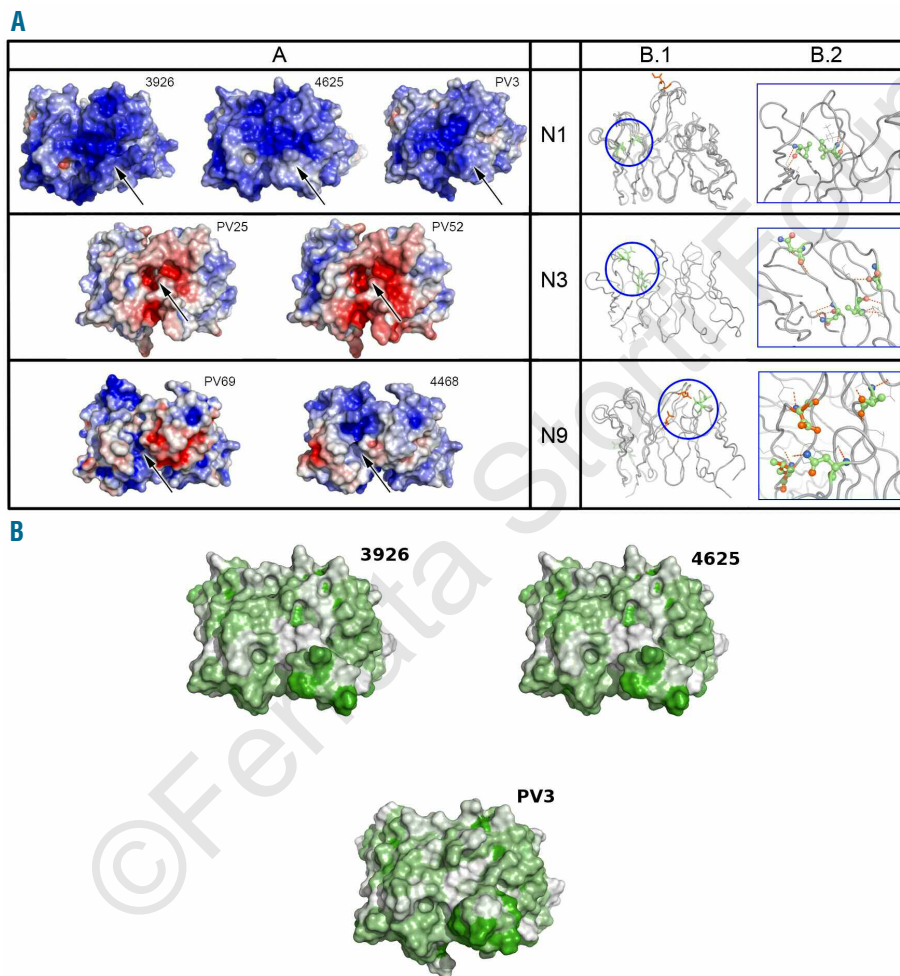
Figure 1. (A) Three-dimensional models of immunoglobulins belonging to subsets N1, N3 and N9. Antigen binding regions (panel A) and mutation analysis (panels B.1 and B.2) of modeled antibodies are shown. In all cases, Igs belonging to the same HCDR3 subset display an overall similarity, both in sequence and structure, that is not limited to the HCDR3 region. Samples 3926, 4625 and PV3 in subset N1 (panel A) have a positively charged antigen binding site (Abs) with a protruding and hydrophobic H3 loop (black arrow). Subset N3 (cases PV25 and PV52, panel A) presents nearly identical Abs. A negatively charged pocket is in the center of the binding site (black arrow). Cases PV69 and 4468 (subset N9, panel A), though presenting a less remarkable similarity, share some common features. The Abs is formed by a negatively charged central protruding region surrounded by a positive area (black arrow). Panel B.1: common mutations (i.e. replacement of the same specific residue of Igs in the same group deriving from the same germline) and convergent mutations (i.e. replacement of a residue in an Ig increasing its similarity to another Ig belonging to the same group but deriving from a different germline) are depicted in green and red, respectively, on the superimposed backbones of all the Igs in each group. Blue circled area highlights the most relevant mutations; a close-up of the same region is in panel B.2, where relevant mutations are reported in a ball-and-stick representation mapped on a single representative structure of each group.

Correlated mutations are found on IGHV in subset N1, no significant correlation is evident on the light chains. Two common mutations (M33I and A71V) are located in close proximity in the models and are likely to play a role in Abs specificity. Two similar mutations (S30BN and S34N) are found in loop H1 of both samples belonging to subset N3 and may influence the H3 position and the Abs shape. Several mutations on the heavy and light chains of both samples in subset N9 introduce polar and charged residues in the Abs. Sequence numbering follows the Kabat-Chothia scheme.[23] Abs are colored according to their electrostatic potential, ranging from −5 k$_B$T/e (red) to + 5 k$_B$T/e (blue). (B) Hydrophobic areas of cases belonging to subset N1. Antigen binding sites of subset N1 immunoglobulins are depicted using the Kyte-Doolittle hydropaticity score.[24] Note the large hydrophobic patch found in the H3 loops of all samples (dark-green region) that is often associated with an interaction area.

nosis was 67 years (range 40-82). HCV serology was positive in 26 of 127 cases (20%). Five-year OS was 81% and 5-year PFS 46%.

Productive IGHV-D-J rearrangements were obtained from all 133 SMZL cases. Use of IGHV, IGHD, and IGHJ genes is reported in Table 1. The IGHV families most frequently used were IGHV3 (n=66, 50%), IGHV1 (n=40, 30%) and IGHV4 (n=23, 17%). The IGHV genes most frequently rearranged were IGHV1-2 (n=26, 20%), IGHV3-23 (n=24, 18%), IGHV4-34 (n=10, 8%), and IGHV1-69 (n=10, 8%), confirming a biased VH gene usage in SMZL as already suggested. Using the 98% identity cut-off value, 93 (70%) sequences were mutated and 40 (30%) sequences were unmutated (15 with 100% identity, 10 with 99-99.9%, 15 with 98-98.9%).

By means of HCDR3 clustering analysis, 16 of 133 (12%) SMZL investigated in this study met the minimal criteria to be included in subsets with stereotyped HCDR3 (Table 2 and *Online Supplementary Table S3*); 7 of 86 (8%) SMZL sequences retrieved from public databases also belonged to stereotyped HCDR3 subsets (*Online Supplementary Table S3*).

On the basis of disease representation within individual subsets, three major patterns of HCDR3 subsets were identified: i) "*SMZL-specific subsets*", that included only cases of SMZL; ii) "*non-Hodgkin's lymphoma (NHL)-like subsets*", that included cases of both SMZL and NHL in the same subset; and iii) "*CLL-like subsets*", that included cases of both SMZL and CLL in the same subset (Table 2 and *Online Supplementary Table S3*).

The 5 SMZL-specific subsets were in all instances novel subsets (confirmed: N1, N2; provisional: N3, N4, N5). Nine HCV-negative SMZL cases from our cohort and 3 SMZL cases from public databases belonged to SMZL-specific subsets. The 5 NHL-like subsets (confirmed: N6, N7, N8; provisional: N9, N10) were composed of 5 SMZL cases (4 from this series) clustering with non-splenic MZL and extranodal indolent lymphomas. Four SMZL (2 from our series) clustered within known CLL subsets (subsets 1, 6, 9, 25) and two (one from our series) formed 2 novel provisional subsets (N11, N12).

Three-dimensional models of Igs were produced for 3 subsets (2 SMZL-specific: N1, N3; 1 NHL-like: N9), after obtaining light chain sequences (*Online Supplementary Table S3*). Details of the modeling are reported in the Figure 1 legend. In all cases, Igs belonging to the same stereotyped group displayed an overall similarity, both in sequence and structure, not limited to HCDR3. In particular, in subset N1, a large hydrophobic area has been found on loop H3 (Figure 1) that is likely to be involved in antigen recognition.

HCV serology was negative in SMZL belonging to SMZL-specific subsets and CLL-like subsets, whereas was positive in 2/3 SMZL belonging to NHL subsets (*P*=0.02; *Online Supplementary Table S4*). Median PFS was 1.7 years (SE 0.5, 95% CI 0.7-4.7) for stereotyped cases and 5.3 for non-stereotyped (SE 0.8, 95% CI 2.6-7.3) (*P*=0.06). Median PFS was 2.7 years (SE 0.45, 95% CI 1.2-ND) for unmutated cases and 4.7 for mutated cases (SE 0.48, 95% CI 2.6-ND) (*P*=0.06). By Cox's regression analysis with mutational status and stereotyped HCDR3 as covariates, both had a significant effect on PFS (mutational status HR=1.8, *P*=0.04; cluster HR=2.1 *P*=0.02).

These data point to the occurrence of SMZL-specific, stereotyped HCDR3 sequences and Igs structures in SMZL. Occurrence of stereotyped HCDR3 in both HCV-negative and HCV-positive cases suggests that the potential role of antigens in SMZL pathogenesis is not restricted to HCV, but may involve also HCV-unrelated epitopes with a common BCR-binding signature. The role of antigen stimulation in SMZL is further reinforced by the observation that a fraction of SMZL share stereotyped HCDR3 with non-splenic MZL and with CLL, suggesting a potential common pathogenetic trigger in different indolent B-cell malignancies.

Authorship and Disclosures

References

1. Swerdlow SH CE, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Lyon: IARC Press, 2008.

2. Matutes E, Oscier D, Montalban C, Berger F, Callet-Bauchu E, Dogan A, et al. Splenic marginal zone lymphoma proposals for a revision of diagnostic, staging and therapeutic criteria. Leukemia. 2008;22(3):487-95.

3. Arcaini L, Lazzarino M, Colombo N, Burcheri S, Boveri E, Paulli M, et al. Splenic marginal zone lymphoma: a prognostic model for clinical use. Blood. 2006; 107(12):4643-9.

4. Saadoun D, Suarez F, Lefrere F, Valensi F, Mariette X, Aouba A, et al. Splenic lymphoma with villous lymphocytes, associated with type II cryoglobulinemia and HCV infection: a new entity? Blood. 2005; 105(1):74-6.

5. Stamatopoulos K, Belessi C, Papadaki T, Kalagiakou E, Stavroyianni N, Douka V, et al. Immunoglobulin heavy- and light-chain repertoire in splenic marginal zone lymphoma. Mol Med. 2004;10(7-12):89-95.

6. Algara P, Mateo MS, Sanchez-Beato M, Mollejo M, Navas IC, Romero L, et al. Analysis of the IgV(H) somatic mutations in splenic marginal zone lymphoma defines a group of unmutated cases with frequent 7q deletion and adverse clinical course. Blood. 2002;99(4):1299-304.

7. Zhu D, Orchard J, Oscier DG, Wright DH, Stevenson FK. V(H) gene analysis of splenic marginal zone lymphomas reveals diversity in mutational status and initiation of somatic mutation in vivo. Blood. 2002; 100(7):2659-61.

8. Tierens A, Delabie J, Malecka A, Wang J, Gruszka-Westwood A, Catovsky D, et al. Splenic marginal zone lymphoma with villous lymphocytes shows on-going immunoglobulin gene mutations. Am J Pathol. 2003;162(2):681-9.

9. Arcaini L, Zibellini S, Passamonti F, Rattotti S, Lucioni M, Invernizzi R, et al. Splenic marginal zone lymphoma: Clinical clustering of immunoglobulin heavy chain repertoires. Blood Cells Mol Dis. 2009;42(3):286-91.

10. Traverse-Glehen A, Davi F, Ben Simon E, Callet-Bauchu E, Felman P, Baseggio L, et al. Analysis of VH genes in marginal zone lymphoma reveals marked heterogeneity between splenic and nodal tumors and suggests the existence of clonal selection. Haematologica. 2005;90(4):470-8.

11. Stamatopoulos K, Belessi C, Moreno C, Boudjograh M, Guida G, Smilevska T, et al. Over 20% of patients with chronic lymphocytic leukemia carry stereotyped recep-

tors: Pathogenetic implications and clinical correlations. Blood. 2007;109(1):259-70.

12. Murray F, Darzentas N, Hadzidimitriou A, Tobin G, Boudjogra M, Scielzo C, et al. Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. Blood. 2008;111(3):1524-33.

13. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. N Engl J Med. 2005; 352(8):804-15.

14. Forconi F, Sahota SS, Raspadori D, Mockridge CI, Lauria F, Stevenson FK. Tumor cells of hairy cell leukemia express multiple clonally related immunoglobulin isotypes via RNA splicing. Blood. 2001; 98(4):1174-81.

15. Lefranc MP. IMGT, the International ImMunoGeneTics Information System for Immunoinformatics: methods for querying IMGT databases, tools, and web resources in the context of immunoinformatics. Mol Biotechnol. 2008;40(1):101-11.

16. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23(21):2947-8.

17. Messmer BT, Albesiano E, Efremov DG, Ghiotto F, Allen SL, Kolitz J, et al. Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. J Exp Med. 2004;200(4):519-25.

18. Rossi D, Spina V, Cerri M, Rasi S, Deambrogi C, De Paoli L, et al. Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. Clin Cancer Res. 2009;15(13):4415-22.

19. Marcatili P, Rosi A, Tramontano A. PIGS: automatic prediction of antibody structures. Bioinformatics. 2008;24(17):1953-4.

20. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol. 1993;234(3):779-815.

21. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA. 2001;98(18):10037-41.

22. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25(5):579-86.

23. Al-Lazikani B, Lesk AM, Chothia C. Standard conformations for the canonical structures of immunoglobulins. J Mol Biol. 1997;273(4):927-48.

24. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157(1):105-32.