Igs Expressed by Chronic Lymphocytic Leukemia B Cells Show Limited Binding-Site Structure Variability

Paolo Marcatili, Fabio Ghiotto, Claudya Tenca, Anna Chailyan, Andrea N. Mazzarello, Xiao-Jie Yan, Monica Colombo, Emilia Albesiano, Davide Bagnara, Giovanna Cutrona, Fortunato Morabito, Silvia Bruno, Manlio Ferrari, Nicholas Chiorazzi, Anna Tramontano and Franco Fais

*J Immunol* 2013; 190:5771-5778; Prepublished online 1 May 2013;
doi: 10.4049/jimmunol.1300321
http://www.jimmunol.org/content/190/11/5771

Supplementary Material

http://www.jimmunol.org/content/suppl/2013/05/01/jimmunol.1300321.DC1

References

This article cites 47 articles, 25 of which you can access for free at:
http://www.jimmunol.org/content/190/11/5771.full#ref-list-1

Subscription

Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

Permissions

Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts

Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Igs Expressed by Chronic Lymphocytic Leukemia B Cells Show Limited Binding-Site Structure Variability

Paolo Marcatili,*,† Fabio Ghiotto,‡,† Claudia Tenca,§ Anna Chailyan,* Andrea N. Mazzarello,*,† Xiao-Jie Yan,*,† Monica Colombo,§ Emilia Albesiano,‡ Davide Bagnara,‡ Giovanna Cutrera,§ Fortunato Morabito,§ Silvia Bruno,‡ Manlio Ferrari,§ Nicholas Chiorazzi,‡∥||,** Anna Tramontano,*,†‡ and Franco Fais†‡

Ag selection has been suggested to play a role in chronic lymphocytic leukemia (CLL) pathogenesis, but no large-scale analysis has been performed so far on the structure of the Ag-binding sites (ABSs) of leukemic cell Igs. We sequenced both H and L chain V(D)J rearrangements from 366 CLL patients and modeled their three-dimensional structures. The resulting ABS structures were clustered into a small number of discrete sets, each containing ABSs with similar shapes and physicochemical properties. This structural classification correlates well with other known prognostic factors such as Ig mutation status and recurrent (stereotyped) receptors, but it shows a better prognostic value, at least in the case of one structural cluster for which clinical data were available. These findings suggest, for the first time, to our knowledge, on the basis of a structural analysis of the AB-binding sites, that selection by a finite quota of antigenic structures operates on most CLL cases, whether mutated or unmutated.

The Journal of Immunology, 2013, 190: 5771–5778.

Chronic lymphocytic leukemia (CLL) is the most frequent form of leukemia in the western world and is characterized by a clonal expansion of neoplastic mature B lymphocytes. CLL pathogenesis is still unclear, but it appears that many factors contribute to the evolution and expansion of the neoplastic clones (1).

Analyses of the sequences of H and L chain variable regions of Igs expressed on the surface of leukemic cells showed that the IGHV and IGL/KV regions undergo somatic hypermutation in ∼50% of leukemic clones (2–4), and that patients with mutated IGHV genes generally have a more indolent clinical course than patients with unmutated IGHV genes (5, 6).

In addition, clones from different CLL patients express Igs that contain remarkably similar IGHV amino acid sequences (7–11). The extent of these recurrent rearrangements, termed “stereotyped Igs,” in the CLL repertoire has been recently appreciated through the analysis of thousands of CLL H chain IGVs: ∼30% of CLL cases fall within 1 of >300 subgroups of stereotyped Igs (stereotyped subsets) (12–14).

Ag–Ig interaction might play a crucial role in CLL pathogenesis as well. It is still unclear, however, whether the role of Ags is crucial in all CLL cases or is restricted to only CLLs with stereotyped Igs, which are mostly unmutated (10, 16, 17).

The definition of stereotyped Igs is mainly based on the HCDR3 amino acid composition and length. These play a crucial role in the Ag–Ig interaction, but the shape of the whole Ag-binding site (ABS) obviously also depends on other Ig regions, and it is important to analyze the whole binding site at the structural level.

In their seminal work, Wu and Kabat (18) identified three sequence portions on each Ig VH, and VK or VL domain, the so-called hypervariable regions, with an extremely variable amino acid composition in comparison with the other less variable parts. They correctly predicted such regions to assume a loop conformation and to be responsible for the selective binding of the Ag, and named them “complementary determining regions” (CDRs) in contrast with the surrounding “framework” regions. The work of Chothia and Lesk (19) and of some of us (20, 21) extended the analysis and pointed out that five of six CDRs (LCDR1–3 and HCDR1–2) and a portion of the sixth loop (HCDR3), although presenting a very frequent sequence repertoire, usually adopt a limited set of backbone conformations referred to as canonical structures determined by the nature of relatively few residues that are primarily responsible for their main-chain conformations. These residues are found both within the hypervariable regions and in the conserved β-sheet framework (21).
These studies made it possible to develop ad hoc modeling techniques (22) to build Ab models accurate enough for theoretical and practical studies, such as docking (23), engineering (24), and comparison (25).

Taking advantage of the above, we analyze in this article for the first time, to our knowledge, the structures of Igs from CLL patients, with the aim of evaluating whether information on the ABS structure can provide novel insights into the Ag role in CLL pathogenesis. We modeled the structure of Igs derived from a cohort of 366 CLL patients starting from the amino acid sequences of their paired H and L chains, and studied the structural features of their ABS to highlight possible common patterns potentially correlated with the pathological phenotype.

Materials and Methods

Patients, leukemic cells, IGVH sequences, and analyses

After informed consent according to the Declaration of Helsinki, PBMCs were isolated from heparinized venous blood of patients with CLL. CLL diagnosis was based on accepted clinical immunophenotypic criteria (26). Rearranged IGHV-D-J and IGKV-J or IGLV-J paired segments were sequenced from the cDNA of 218 CLL patients as described previously (2, 3); in addition, 148 Ig sequences (IGHV + IGK/LV) were retrieved from GenBank, bringing the total number of analyzed cases to 366. All of the latter were submitted to the database by the research group of Dr. K. Stamatopoulos (Hematology Department and HCT Unit, G. Papanicolaou Hospital, Thessaloniki, Greece), who verified that allelic exclusion was taken into account (personal communication). Only samples with allelic exclusion of both IGHV and IGK/LV were included in the study. Sequences were analyzed using the ImMunoGeneTics Information System (http://www.imgt.org) (27). The mutational status of Ig clones was defined based on both IGHV and IGK/LV. Patients with leukemic cell clones exhibiting <2% mutations in both V segments were labeled as “unmutated CLL cases” (U-CLL), whereas patients with >2% IGHV and/or IGK/LV somatic mutations were defined as “mutated CLL cases” (M-CLL).

We also used a finer mutational classification by dividing the Igs into three classes: heavily mutated (HM) Igs (IGHV and/or IGK/LV percentage of mutation ≥5%); scarcely mutated (SM) Igs (IGHV and/or IGK/LV ≥1% and <3% mutations), and unmutated Igs (IGHV and IGK/LV percentage of mutation <1%). The cutoffs adopted for our three-class partitioning are slightly different from those defined in a previous study on the IGHV gene repertoire in splenic marginal zone lymphoma (28). The classification did not change when the sequences were inspected by running BLAST on the National Center for Biotechnology Information human gene database.

Ig data sets

We built a “test” data set by querying the DIGIT database (29) for all available human Igs for which the paired sequences of the H and L chains were available. After inspecting the Ig description contained in the DIGIT database and the related PubMed entry, we discarded all Igs for which no reference to any published article could be found or not corresponding to an entry in Entrez Nucleotide (http://www.ncbi.nlm.nih.gov/nucleotide/), as well as all Igs already contained in our initial CLL data set. We ended up with 2441 Igs for which complete information on canonical structures, loop length, and mutation rates could be retrieved using the tools provided by DIGIT (29).

Among the 2441 Igs of the “test” data set, 212 were from CLLs, and we labeled them as “test CLL” data set. Among the remaining 2229 Igs ("test without [w/o] CLL"), we also defined a “test AI” data set including the 294 sequences for which the associated PubMed entry contained any of the MeSH terms “autoimmunity” (MeSH tree no. D12.450.192), “autoimmune disease” (MeSH tree no. C20.111), or “autoantibody” (MeSH tree no. D12.776.124.486.485.114.323). All the remaining 1935 Igs were defined as the “test w/o (CLL-AI) data set. All considered Igs included the full-length variable domains.

Ig three-dimensional models

We used the PIGS server, based on the canonical structure method, to derive the sequence alignments of the Ig frameworks and to build the three-dimensional models of all Igs in our data set (22). We could build 342 complete and correctly assembled models. The remaining 24 models were discarded because the modeling procedure returned an incomplete or improperly assembled model: in 14 models, the IGHV-IGL/KV packing was incorrect; in 8 cases, no template was found to model the LCDR2, and in 1 case, the HCDR3 was too long to be properly modeled.

Structural analysis and clustering

Structural superpositions were performed using the LGA package (30). The loop coordinates as defined by Al-Lazikani et al. (31) plus the two residues flanking the N and C termini of each loop were used for the superposition of the ABSs. Two residues were considered as corresponding to each other if the distance of their Cα after superposition was <8 Å. As described in the text, we used the clustering software of the best cluster separations. The tested distance metrics were root-mean-square deviation, global distance test, and template modeling (TM) score distance matrices for the superimposed structures (33). For each distance matrix, we performed agglomerative hierarchical clustering using the agnes function (Maechler & Rousseeuw, http://cran.r-project.org/web/packages/cluster/) and divisive hierarchical clustering (diana method, Maechler & Rousseeuw, http://cran.r-project.org/web/packages/cluster/) of the R package with a number of clusters ranging from 10 to 50. The linkage functions used in our analysis were complete, single, average, and Ward’s method (34).

The best average silhouette value (0.146) was obtained using TM score as metric and a divisive clustering scheme with 21 clusters. Ig images were generated using the Pymol software (W. L. DeLano, The PyMOL Molecular Graphics System, 2002 HYPERLINK, http://www.pymol.org). Solvent-accessible surface electrostatic potentials were calculated using Adaptive Poisson-Boltzmann Solver (35).

Clustering analysis

Clustering results were compared with the IGHV and IGLV mutational status. As described earlier, we used two different classifications to represent the mutation level of Igs, namely, the two-class partition with a 2% cutoff for defining mutated and unmutated groups, and the three-class partition that divides CLLs in HM, SM, and unmutated samples. For each cluster and for the two classifications, we computed the probability that an equal or higher number of Igs belonging to the same class could be found by chance in a randomly extracted subset of the same size (hypergeometric distribution). The two (Hamron–Hildrum)–Hilbert–Munkres-Munkres (O(n^3)) was used to perform the clustering. We assigned the smaller value between the two/three probability values to each cluster. The graphical representation of the cluster results was generated using the R package tool A2R.

CLL specificity in structural clusters

To test whether the structural clusters were describing specific features of CLL Igs rather than general Ig characteristics, we built, for each cluster containing more than five Igs, a sequence-based hidden Markov model (HMM) (36) including the H and L chains of all Igs in the cluster. To this end, we used the HMMER package with default parameters.

Each HMM was used to score each Ig in the test data set and, for each Ig, the largest score was recorded. The same procedure was applied to the Igs in each of the clusters used to build the HMMs. For the sequence-based clustering, IGHV-IGLV/KI paired sequences of all Igs of our CLL data set were clustered using the cd-hit software (37) with a sequence identity threshold of 80%. The statistical difference between the Igs in the test data set marked as CLL with all others was computed using the R implementation of the Wilcoxon Mann–Whitney U test.

Results

Sample description and stereotyped receptor frequency

We sequenced the VH and VKL regions of a cohort of 366 IgM+ CLL patients, 61.7% (226/366) of which expressed IgK and 38.3% (140/366) IgL isotype. According to the two-class classification described in Materials and Methods, the cohort comprised 47.3% (173/366) U-CLL samples, 63.6% (110/173) of which expressed κ-isotype L chains and 36.4% (63/173) the λ-isotype L chains. Of the remaining 52.7% (193/366) M-CLL samples, 60.1% (116/193) expressed the κ-isotype L chains and 39.9% (77/193) the λ-isotype L chains. Of the 366 CLL patients, 13.7% (50/366) expressed a stereotyped BCR: 48 of these belonged to 24 different previously described CLL subsets (12, 14), whereas 2 did not, thus defining a novel stereotyped subset (Supplemental Table I). The most represented CLL subsets were subsets 1, 2, and 6,
respectively, representing 19.2% (n = 10), 19.2% (n = 10), and 11.9% (n = 7) of all stereotyped receptors in the cohort.

Analysis of the Ig three-dimensional models
The definition of stereotyped BCRs is based on sequence information from the H chain only (12, 14). We conjectured that adding information on the L chain and focusing on structural features of the ABS would be more informative.

We used the atomic coordinates of the ABS obtained by structural modeling and quantified their structural similarity. As described in Materials and Methods, we could build reliable models for 342 of the 366 Igs in our CLL data set. Their ABS structures were clustered as described, leading to the definition of 21 well-separated clusters. The most populated cluster (cluster 2) contained 28% of all Igs, followed by cluster 5 (13%), cluster 9 (8.5%), and cluster 1 (7.6%). Altogether, 323 of the 342 modeled Igs (94.5%) fell in clusters containing at least 5 Ig clones. Only 19 Igs (5.5%) were distributed in smaller clusters (Supplemental Table I). Fig. 1 illustrates all 15 clusters containing ≥5 Igs. Among these, seven (clusters 1, 2, 5, 6, 16, and 21) contained Igs with only κ L chains and eight (clusters 3, 4, 7, 8, 9, 10, 19, and 20) only λ L chains. The genetic features of the samples belonging to the 21 clusters are listed in Supplemental Table I. All of the following analyses were performed on the 15 clusters formed by >5 Igs.

We analyzed the correlation between the structural clusters and the mutational status of the Igs. Interestingly, we found that the 172 M-CLL samples and 151 U-CLL Igs segregated with a significant

![FIGURE 1. Structural clustering and mutational status of CLL Igs. (A) Hierarchical divisive clustering of the 342 modeled CLL Igs grouped according to the structural similarity (TM score) of their ABSs. Clusters with five or more samples are shown. On the left, the silhouette value corresponding to different number of clusters is shown. The optimal cut (corresponding to 21 clusters) is reported as a blue dot. (B) Two-class (upper bars, unmuted [U]; mutated [M]) and three-class (lower bars, U, SM, HM) description of the mutational status of all samples. (C) Statistical analysis of the structural clusters. The number of samples, according to their two-class (upper table) and three-class (lower table) description, is reported together with the probability that the enrichment of Igs with the same mutational status observed in a cluster is due to chance alone. The p values <0.05 are reported in bold.](http://www.jimmunol.org/)

Downloaded from http://www.jimmunol.org/ by guest on May 31, 2017
overrepresentation of either mutated or unmutated Igs in 5 of the clusters (clusters 1, 2, 3, 6, and 9), accounting for 53% of the cases (180/342; Fig. 1). If three different intervals of mutation are used instead of two (HM, SM, unmutated), 9 of the 15 clusters were significantly enriched in samples belonging to 1 of the 3 groups (clusters 1, 2, 3, 4, 6, 9, 18, 19, and 20), accounting for 224 of 342 (65%) modeled Igs (Fig. 1).

We also mapped the hydrophobicity and electrostatic potential of the ABS surface of our models, and these properties also turned out to be very similar in Igs within a cluster (Figs. 2–4). As an example, the structure of all Igs belonging to one cluster (cluster 19) is shown in Fig. 2. As can be appreciated from Fig. 2, remarkable similarities can be observed in terms of conformation, hydrophobicity, and electrostatic potential of the ABS surface. Fig. 3 shows, for each of the clusters, one representative Ig structure. Conserved hydrophobic patches can be identified in some clusters, located either in the center of the ABS (clusters 7 and 9), near H3 (clusters 2, 6, 12, 18), or near the H2 loop (clusters 1 and 3; Fig. 3). Based on the classifications adopted in the literature (38, 39), a summary of the ABS characteristics are reported in Table I. It is apparent that these characteristics, even if very difficult to quantify in an objective way in protein models, can provide an overview of the main differences and similarities between the ABSs belonging to different clusters and to the same cluster, respectively. It is to be expected that they are related to the nature of the respective Ags. For example, Abs with deep pockets, grooves, and flat sites are often specific for small molecules, peptides, and proteins, respectively. Interestingly, Igs belonging to the same ABS cluster, thus sharing high structural similarity, do not necessarily show a high level of sequence similarity, as demonstrated by the examples shown in Fig. 4.

The length of some hypervariable loops belonging to some clusters also shows a bias (Table I). For instance, cluster 2 includes...
To verify whether the structural clustering captures, in whole or in part, features that are specific for CLL Igs, we built HMMs for each of the clusters using the sequences of its members and used them to classify the sequences of the independent Ig data sets described in Materials and Methods that are the "test" data set composed by Igs from CLL patients not present in our original data set ($n = 212$) and the "test w/o CLL" data set including non-CLL Igs ($n = 2229$).

The resulting similarity score distributions showed that members of the "test CLL" had a significantly higher score than those from the "test w/o CLL" Igs ($p = 0.0014$). We also repeated the procedure on our "test AI" and "test w/o (CLL-AI)" data sets, and observed that the scores for the "test AI" samples were rather similar to those of CLL Igs. Accordingly, when the method was applied to the "test CLL" and "test w/o (CLL-AI)" Igs (the remaining non-CLL/nonautoreactive Igs, $n = 1935$), the difference between the scores of the former and latter data sets was even more pronounced ($p = 7.1 \times 10^{-5}$; Fig. 6A). These results indicate that our structural clustering well reflects the properties of CLL Abs, and that these are somewhat similar to those of autoreactive Ig-binding sites, as it has been suggested before by studies showing that CLLs originate from self-reactive B cell precursors (40,41). However, it was also observed that although U-CLLs expressed autoreactive Abs, most M-CLLs did not (41). In line with this latter observation, when scoring "test AI" Ig sequences expressed autoreactive Abs, most M-CLLs did not (41). However, it was also observed that although U-CLLs expressed autoreactive Abs, most M-CLLs did not (41).

It has been reported that the IGHV gene mutational status correlates with the overall survival of affected patients. This is also the case, on average, for Igs in our CLL patients for which clinical information is available (Fig. 5A). Remarkably, a different pattern is observed for Igs belonging to cluster 2, for which there are sufficient clinical data and a balanced distribution of mutated and nonmutated Igs to carry out a statistically significant analysis. Mutated and unmutated Igs belonging to this cluster have almost overlapping overall survival curves and very similar median of the survival time (174 versus 172 mo; Fig. 5B). We compared this result with that obtained by selecting at random 10,000 times the same number of Igs as in cluster 2 from our data set, whereas keeping the same ratio of mutated and unmutated samples. In only 6.6% of the cases was the difference in the median survival time between patients with mutated and nonmutated Igs in the random sampling smaller than or equal to that observed in cluster 2. This implies that the survival of CLL patients is not necessarily only correlated to the mutational status of their Igs, but, as in the case of the cluster 2 samples, might be related to the ABS properties and, therefore, most likely to the recognized Ag.

The BCRR stereotypy of the obtained structural clusters is remarkably biased, with cluster 4 containing all the 12 subset 2 BCRR samples, cluster 1 all subset 6, cluster 2 most of the subset 1 (7/10), and cluster 7 containing both of the 2 novel stereotyped BCRRs identified in this study for the first time, to our knowledge (Supplemental Table I).

**CLL specificity of the structural clusters**

To verify whether the structural clustering captures, in whole or in part, features that are specific for CLL Igs, we built HMMs for each of the clusters using the sequences of its members and used them to classify the sequences of the independent Ig data sets described in Materials and Methods that are the "test" data set composed by Igs from CLL patients not present in our original data set ($n = 212$) and the "test w/o CLL" data set including non-CLL Igs ($n = 2229$).
captures additional relevant information that is not, or perhaps only very weakly, correlated with germline usage, such as the HCDR3 structure and the VL/VH packing.

**Discussion**

It has been a matter of discussion for several years whether the B lymphocyte clones that accumulate in CLL patients display widely distributed Ag specificity among the billions of self and nonself antigenic epitopes encountered by the immune system, or alternatively, if they express a restricted set of ABSs. Support to the first hypothesis is provided by the observation that stereotyped IgV rearrangements exist (2, 7–9), even though they are mostly found in U-CLL cases (10, 16, 17). Furthermore, a classification based on stereotyped receptors accounts for only a fraction of CLL patients, and these are distributed in quite a large number of subsets (300). This is likely due to the fact that the definition of stereotyped BCRs is mostly based on the HCDR3 amino acid sequence composition and length, which only partially contributes to the shape of the ABS of an Ig.

In this study, we exploited the availability of our accurate protocol for modeling the structure of Igs and of our collection of paired VL/VH sequences of Igs to investigate whether a better understanding of the CLL Igs could be obtained by taking into account the shape of their complete binding site.

To this end, we built models of all the complete Igs from CLL patients obtained by us and retrieved from public sources, and clustered them on the basis of the structural similarity of their binding sites. Interestingly, the samples, both U-CLL and M-CLL, could be partitioned in a limited number of clusters. Notably, this is not the case if the clustering is done on the basis of sequence similarity. Most members of the clusters shared interesting properties other than their structural similarity, such as the type of L chain (κ or λ), BCR stereotypes, and mutational status. In some instances, members of the same cluster display remarkable homogeneity also in terms of the IGHV-IGK/LV usage and H/L chains pairing.

For example, Igs belonging to cluster 1 all carry the IGHV1-69 gene combined with IGKV L chains, and cluster 3 contains all
unmutated Igs that use the IGHV1-69 gene combined with L chains that almost exclusively use genes of the IGLV3 family.

Cluster 4 includes almost half of all SM cases of our CLL cohort and all subset 2 stereotyped cases of the data set. These CLLs use the IGHV3-21/IGLV3-21 genes and are known to have unfavorable clinical outcome regardless of their mutational status (42). Interestingly, cluster 4 includes IGHV3-21, but also IGHV3-48 and IGHV3-11 CLL cases. In a previous study (43), a very large data set of HCDR3 amino acid sequences was used to cluster CLL cases based on their HCDR3 sequences, and the IGHV3-21 gene predominated in a cluster where also CLLs using the IGHV3-48 and IGHV3-11 genes were present. This was interpreted as suggestive of the presence of some functional constraint that we can now also relate to the structure of the binding site. Interestingly, the composition of our cluster 4 indicates that the IGHV3-21, IGHV3-48, and IGHV3-11 CLL genes generate a structurally similar binding site (almost) only when paired with the IGLV3-21 L chain gene.

We used our structural clustering results to generate statistical models of their members in the form of HMMs. The generated HMMs have discriminative power in that they are able to identify CLL Igs in a large data set not including the Igs used for clustering and, remarkably, are also able to separate AI Abs from non-AI ones. Also in this case, no discriminative power could be achieved using a sequence-based classification.

The ability of the structure-based HMMs to identify common features among AI Abs can, in principle, be due to a bias in the data set because sequenced AI Abs react to a subset of specific Ags. However, it is very likely that this potential bias is not, or is not solely, responsible for our finding given the fact that in several cases auto-Ags have been proved to be reactive with CLL clones (44).

Our results strongly suggest that specific features of the CLL Igs reside in the overall atomic structure of their binding site and therefore provide support to the hypothesis that a finite number of antigenic structures may be involved in CLL pathogenesis.

Our data also show that the features of the ABS are partially captured by the stereotype subset classification, even though the latter relies only on the amino acid sequence of CDR3. For example, our cluster 1 contains all subset 6 stereotyped Igs, cluster 4 all subset 2 stereotyped cases, cluster 2 mostly includes subset 1 stereotyped BCRs (7/10), and cluster 7 contains the 2 novel stereotyped BCRs identified for the first time, to our knowledge, in this study.

The possibility of clustering CLL Igs on the basis of their functional properties as determined by the structure of their binding site and potentially by their recognized Ag raises the interesting question of whether there is any correlation with the clinical outcome of the patients. This might be expected because some BCR stereotype subsets (14, 17, 45–48), as well as the mutational status (5, 6), are known to correlate with the patient prognosis.

Clinical data are available for only a limited set of patients; however, the analysis of our most populated cluster (cluster 2) shows very promising results.

This cluster mostly contains M-CLL cases and a smaller fraction of U-CLL cases. As mentioned earlier, M-CLLs have generally longer life expectancy, less need for therapy, and better response to treatment when compared with U-CLLs. However, the U-CLL cases within cluster 2 display a clinical outcome much more favorable than what is generally observed for U-CLL patients. Should this result be confirmed on a larger cohort of patients, it would indicate that clinical outcome is correlated to the structure of the binding site and only indirectly linked to the IGHV mutation status. In this view, cluster subdivision would be of help in better classifying the numerous outliers observed using the more simplistic U-CLL and M-CLL classification.

This is the first time, to our knowledge, that Igs from CLL patients were analyzed from a structural point of view, and we believe that our results point to the relevance of using this approach on a larger scale, which can now be easily handled by current methodologies for modeling and structural analysis. This is a perspective study and the clinical data are still rather sparse. Clearly we will continue to monitor the clinical outcome of the patients enrolled in this study, as well as repeat the analysis on samples that will become available in the future, because we strongly believe that our strategy has the potential to lead to advances in understanding the nature of CLL and in managing patients.

The correlation between ABS structure and clinical outcome, if confirmed, may provide novel tools for a more robust prognostic stratification of CLL, also thanks to the fact that sequencing of the IgVL can easily become a standard laboratory test, as it is already the case for IgVH sequencing, and the modeling and clustering protocols are very well defined and available.

Currently, it is essentially impossible to identify the Ag given the structure of the cognate Ab-binding site, but this might change in the future and hopefully we might also be able to gain insight in the nature of the Ags associated with CLL pathogenesis, which would obviously have important applications for therapy.
Disclosures

The authors have no financial conflicts of interest.

References


