



# Transitions of CDR-L3 Loop Canonical Cluster Conformations on the Micro-to-Millisecond Timescale

Monica L. Fernández-Quintero, Barbara A. Math, Johannes R. Loeffler and Klaus R. Liedl\*

Center for Molecular Biosciences Innsbruck (CMBI), Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, Innsbruck, Austria

Sequence and structural diversity of antibodies are concentrated on six hypervariable loops, also known as the complementarity determining regions (CDRs). Five of six antibody CDR loops presumably adopt a so-called canonical structure out of a limited number of conformations. However, here we show for four antibody CDR-L3 loops differing in length and sequence, that each loop undergoes conformational transitions between different canonical structures. By extensive sampling in combination with Markov-state models we reconstruct the kinetics and probabilities of the transitions between canonical structures. Additionally, for these four CDR-L3 loops, we identify all relevant conformations in solution. Thereby we extend the model of static canonical structures to a dynamic conformational ensemble as a new paradigm in the field of antibody structure design.

**Keywords:** canonical structures, CDR-L3 loop, molecular dynamics simulations, markov-state models, conformational ensemble, antibody structure design

## OPEN ACCESS

### Edited by:

Deborah K. Dunn-Walters,  
University of Surrey, United Kingdom

### Reviewed by:

Guy Georges,  
Roche Innovation Center Munich,  
Germany  
Charlotte Deane,  
University of Oxford, United Kingdom

### \*Correspondence:

Klaus R. Liedl  
klaus.liedl@uibk.ac.at

### Specialty section:

This article was submitted to  
B Cell Biology,  
a section of the journal  
Frontiers in Immunology

**Received:** 30 August 2019

**Accepted:** 25 October 2019

**Published:** 19 November 2019

### Citation:

Fernández-Quintero ML, Math BA,  
Loeffler JR and Liedl KR (2019)  
Transitions of CDR-L3 Loop Canonical  
Cluster Conformations on the  
Micro-to-Millisecond Timescale.  
*Front. Immunol.* 10:2652.  
doi: 10.3389/fimmu.2019.02652

## INTRODUCTION

Antibodies have become key players as therapeutic agents and therefore the understanding of the antigen-binding process is crucial (1, 2). The antibody binding site consists of six hypervariable loops, each three on the variable domains of the heavy ( $V_H$ ) and the light chain ( $V_L$ ) that shape the antigen binding site, the paratope (2–5). Five of the six antibody CDR loops can adopt a limited number of main-chain conformations known as canonical structures, except of the CDR-H3 loop (6–8). The CDR-H3 loop, due to its high diversity in length, sequence and structure and its ability to adopt various different conformations during the V(D)J recombination and somatic hypermutation, remains challenging to predict accurately (9–13). Together with the CDR-H3 loop the CDR-L3 loop is situated in the center of the paratope and contributes to antigen recognition (14). The CDR-L3 loop is similarly diverse, however without the contribution of a D gene the degree of variability is less (15). The CDR-L3 loop reveals a diversity of length and sequence composition due to the recombination of two gene segments  $V_L$  and  $J_L$ . The  $V_L$  segment codes for the residues 1–95, including the first two CDR loops, while the CDR-L3 loop is encoded by the end of the  $V_L$  and the beginning of the  $J_L$  segment (16). The most prominent CDR-L3 loop length consists of nine residues and can adopt six possible canonical clusters. The rarest CDR-L3 loops contain 7, 12 and 13 residues and have only one canonical cluster (17, 18). However, even due to the increase in the number of crystal structures and consequentially also in canonical structures, the relative populations of the canonical clusters are expected to stay the same (15). There are two types of light chains, kappa and lambda. The genes encoding the two light chains are located on separate

chromosomes. Kappa gene segments are encoded on chromosome 2 (52 V genes and 5 J genes) whereas lambda gene segments are encoded on chromosome 22 (30 V genes and 7 J genes) (16, 19–22). Depending on the type of light chain, antibodies reveal differences in conformational flexibility, half-life, and specificity (23). Various studies focused on classifying antibody structures and correlated it with their locus and sequence to improve antibody structure prediction and design (24–27). Additionally there exist several numbering systems for antibodies that are similar in the framework region, but differ around the CDRs (6, 28–30). The PyIgClassify database classifies conformational clusters by determining the CDR sequences and lengths using the IMGT nomenclature (28) and calculating the dihedral angles  $\omega$ ,  $\phi$ , and  $\psi$  of the residues in each CDR (27). We analyzed the conformational diversity of the CDR-L3 loop to identify transition probabilities and timescales between canonical CDR-L3 loop conformations of same length and to characterize the CDR-L3 loop in solution. We focused on the CDR-L3 loop, because it reveals a diversity in sequence and structure comparable to the CDR-H3 loop.

## METHODS

A previously published method characterizing the CDR-H3 loop ensemble in solution (31, 32) was used to investigate the conformational diversity of CDR-L3 loops. Experimental structure information was available for all considered antibody fragments (Fvs). The starting structures for simulations were prepared in MOE (Molecular Operating Environment, Chemical Computing Group, version 2018.01) using the Protonate3D tool (33, 34). To neutralize the charges we used the uniform background charge (35–37). Using the tleap tool of the AmberTools16 (35, 36) package, the crystal structures were soaked with cubic water boxes of TIP3P water molecules with a minimum wall distance of 10 Å to the protein (38). For all crystal structures parameters of the AMBER force field 14SB were used (39). The antibody fragments were carefully equilibrated using a multistep equilibration protocol (40).

## Metadynamics Simulations

To enhance the sampling of the conformational space well-tempered metadynamics (41–43) simulations were performed in GROMACS (44, 45) with the PLUMED 2 implementation (46). As collective variables, we used a linear combination of sine and cosine of the  $\psi$  torsion angles of the CDR-H3 and CDR-L3 loop calculated with functions MATHEVAL and COMBINE implemented in PLUMED 2 (46). As discussed previously, the  $\psi$  torsion angle captures conformational transitions comprehensively (38, 39). The decision to include the CDR-L3 and CDR-H3 loop  $\psi$  torsion angles is based on the structural correlation of the CDR-L3 and CDR-H3 loop and the observed improved sampling efficiency (47). The simulations

were performed at 300 K in an NpT ensemble. We used a Gaussian height of 10.0 kcal/mol. Gaussian deposition occurred every 1,000 steps and a biasfactor of 10 was used. 1  $\mu$ s metadynamics simulations were performed for each available antibody fragment crystal structure. We applied an average linkage hierarchical clustering algorithm with a distance cut-off criterion of 1.2 Å on the resulting trajectories in cpptraj (36, 48) to obtain a large number of clusters.

The cluster representatives for the antibody fragments were equilibrated and simulated for 100 ns using the AMBER16 (35) simulation package.

## Molecular Dynamics Simulations

Molecular dynamics simulations were performed in an NpT ensemble using pmemd.cuda (49). Bonds involving hydrogen atoms were restrained by applying the SHAKE algorithm (50), allowing a time step of 2.0 fs. Atmospheric pressure of the system was preserved by weak coupling to an external bath using the Berendsen algorithm (51). The Langevin thermostat (52) was used to maintain the temperature during simulations at 300 K.

An in-house python hierarchical clustering script using pytraj (36, 53, 54) was used to directly calculate the transitions between the CDR-L3 loop cluster representatives within one simulation. To obtain a representative ensemble in solution and to account for different inherent CDR-L3 loop flexibilities the distance cut-off was chosen for each antibody individually. This clustering is only used to visualize the frequency of transitions, but it is not used for any further analyses. Within these resulting clusters most of the canonical conformation median crystal structures are found. Depending on the CDR-L3 loop length a different number of canonical clusters are available and the median crystal structure information for each loop length was extracted from the PyIgClassify database (27).

Separately, a time-lagged independent component analysis (tICA) was performed using the python library PyEMMA 2 employing a lag time of 10 ns (55). Additionally, PyEMMA 2 was chosen to calculate a Markov-state model (56) to reconstruct the thermodynamics and kinetics, using the k-means clustering algorithm (57) to define microstates and the PCCA+ clustering algorithm (58) to coarse grain the microstates to macrostates. The sampling efficiency and the reliability of the Markov-state model (e.g., defining optimal feature mappings) can be evaluated with the Chapman-Kolmogorov test (59, 60), by using the variational approach for Markov processes (61) and by taking into account the fraction of states used, as the network states must be fully connected to calculate probabilities of transitions and the relative equilibrium probabilities. To build the Markov-state model we used the backbone torsions of the CDR-L3 loop, defined 150 microstates using the k-means clustering algorithm and applied a lag time of 10 ns.

## RESULTS

The first antibody variable fragment (Fv) studied is the house dust mite allergen binding antibody. Der p 1 and Der f 1 are potent allergens, produced by house dust mites, and cause allergic sensitization and asthma. The PDB structures

**Abbreviations:** CDR, Complementary determining region; Fv, Antibody variable fragment; MD, Molecular dynamics; PCCA, Perron cluster cluster analysis; RMSD, Root mean square deviation; tICA, Time-lagged independent component analysis; V<sub>H</sub>, Heavy chain; V<sub>L</sub>, Light chain.

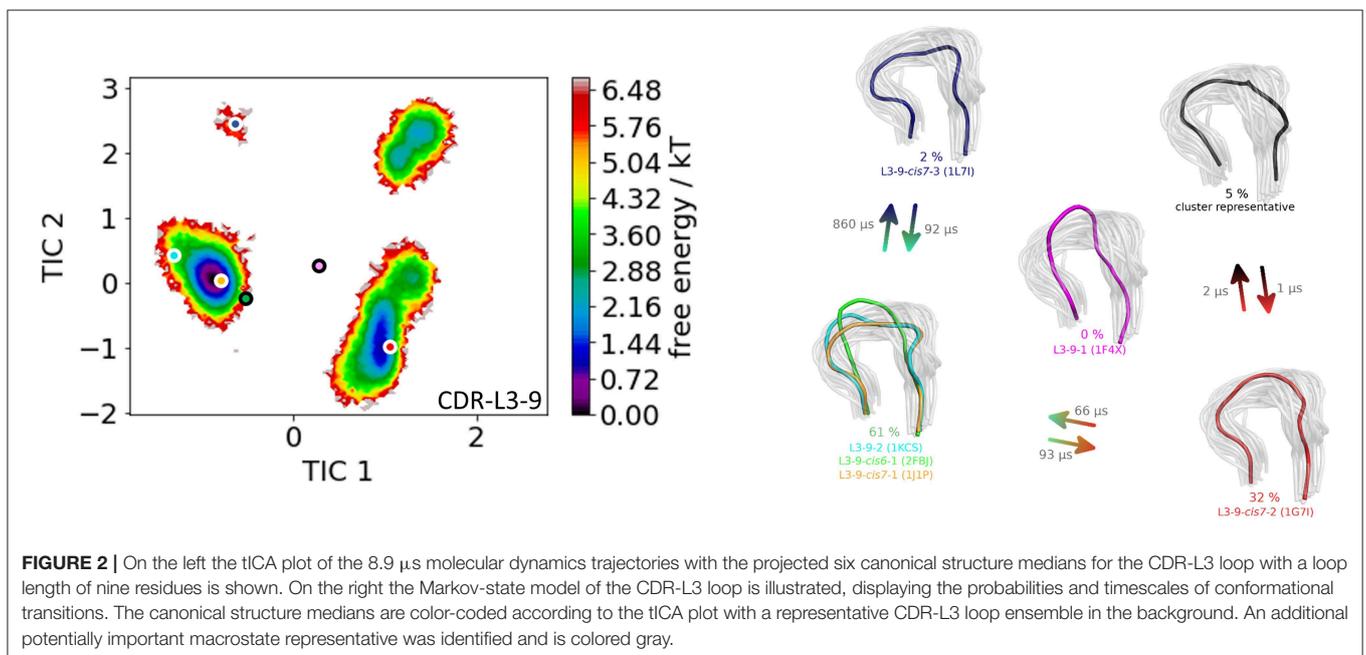
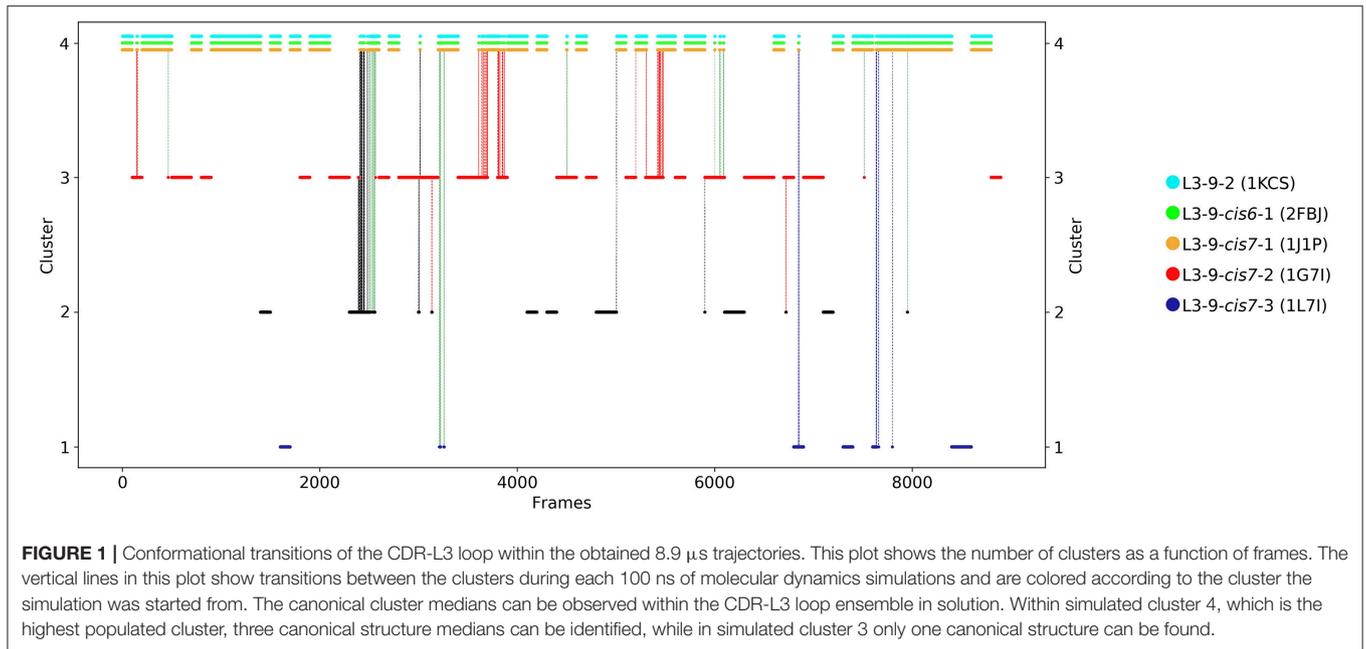
3RVW (crystallized with antigen) and 3RVT (crystallized without antigen) were simulated without the antigen present (62). The CDR-L3 loop length of this house dust mite allergen binding antibody is nine residues. The crystal structures of the 3RVT and 3RVW were originally assigned to the L3-9-cis7-1 cluster containing 1,554 crystal structures, which is the highest populated canonical cluster with the CDR-L3 loop length of nine residues. The PDB accession code of this canonical cluster median is 1J1P, which is colored-coded orange in all following pictures. Besides the characterization of the CDR-H3 loop as conformational ensemble this approach allows to describe the CDR-L3 loop ensemble in solution. As described in the methods section the resulting 89 cluster representatives of the metadynamics simulations were simulated for each 100 ns molecular dynamics simulations. The resulting 8.9  $\mu$ s trajectories were clustered using a hierarchical clustering algorithm with a distance cut-off of 2.4 Å. **Figure 1** shows the conformational transitions observed within the 89 molecular dynamics simulations of 100 ns each. Cluster 4 is the highest populated cluster in which we found three of the six available canonical structure medians (L3-9-2, L3-9-cis6-1, and L3-9-cis7-1) of the CDR-L3 loop with residue length 9. The canonical structure median identified within cluster 3 belongs to the canonical cluster L3-9-cis7-2. The canonical cluster median PDB 1L7I of the L3-9-cis7-3 was found in the very low populated first cluster. The PDB 1F4X belongs to the L3-9-1 canonical structure and is not found in the simulations. **Figure 1** shows various conformational transitions between the four clusters which means that we observe conformational transitions between the canonical structures of the CDR-L3 loop. To identify transition kinetics of the CDR-L3 loop ensemble in solution we calculated a Markov-state model based on a tICA by using the backbone torsions of the CDR-L3 loop (**Figure 2**). **Figure 2** clearly confirms the results of the cluster analysis in **Figure 1**. Combined with a fully connected Markov-state model we identified four macrostates, in which five of the six canonical structures are present. Surprisingly, we even find three canonical structures, including the assigned median canonical structure of the L3-9-cis7-1 canonical cluster 1J1P in the same global minimum in solution. The transitions between the two highest populated macrostates occur in the low microsecond timescale, while the conformational transitions to the least probable macrostate, in which the canonical cluster median of the L3-9-1 1L7I is sampled, occur in the micro-to-millisecond timescale. The canonical cluster median 1F4X, colored in magenta, was not observed.

As a second antibody Fv fragment to characterize the CDR-L3 loop ensemble in solution, the antibody binding to lymphocyte function—associated antigen-1 integrin (LFA-1 integrin) was analyzed. LFA-1 integrin plays a vital role in adhesive interactions with both endothelial cells and antigen-presenting cells (63). Again, the two crystal structures 3HI6 (crystallized with antigen) and 3HI5 (crystallized without antigen) were simulated without antigen present. The CDR-L3 loop contains eight residues and the crystal structures were assigned to the canonical CDR-L3 loop cluster L3-8-1. **Figure 3** shows the clustering transitions of the obtained 8.6  $\mu$ s molecular dynamics simulations with a distance cut-off of 1.4 Å. Within the highest populated cluster

4 the assigned canonical cluster median crystal structure 3CMO is present. The canonical cluster median of the cluster L3-8-2 (PDB 1KEG) is present within cluster 3. Within the least probable cluster 2 the rarest occurring canonical cluster for this loop length L3-8-cis6-1 consisting of only four crystal structures can be found. Besides the sampling of all available canonical conformations of the CDR-L3 loop with eight residues we also observe in **Figure 3** another possible CDR-L3 loop conformation in solution. To identify the kinetic and thermodynamic role of the sampled conformations again a tICA in combination with a Markov-state model was performed (**Figure 4**). **Figure 4** shows the probabilities and transition kinetics of the CDR-L3 loop ensemble in solution. The three available canonical median structures of the CDR-L3 loop are color-coded according to the clustering in **Figure 3**. Again, the assigned canonical cluster median structure 3CMO is present in the highest populated macrostate of the free energy landscape, which is in line with the hierarchical clustering in **Figure 3**. The transitions between the canonical cluster medians 3CMO and 1KEG occur in the nano-to-microsecond timescale, while the conformational transitions to the least probable macrostate show high microsecond timescales, in which the third canonical structure 1E6O was found.

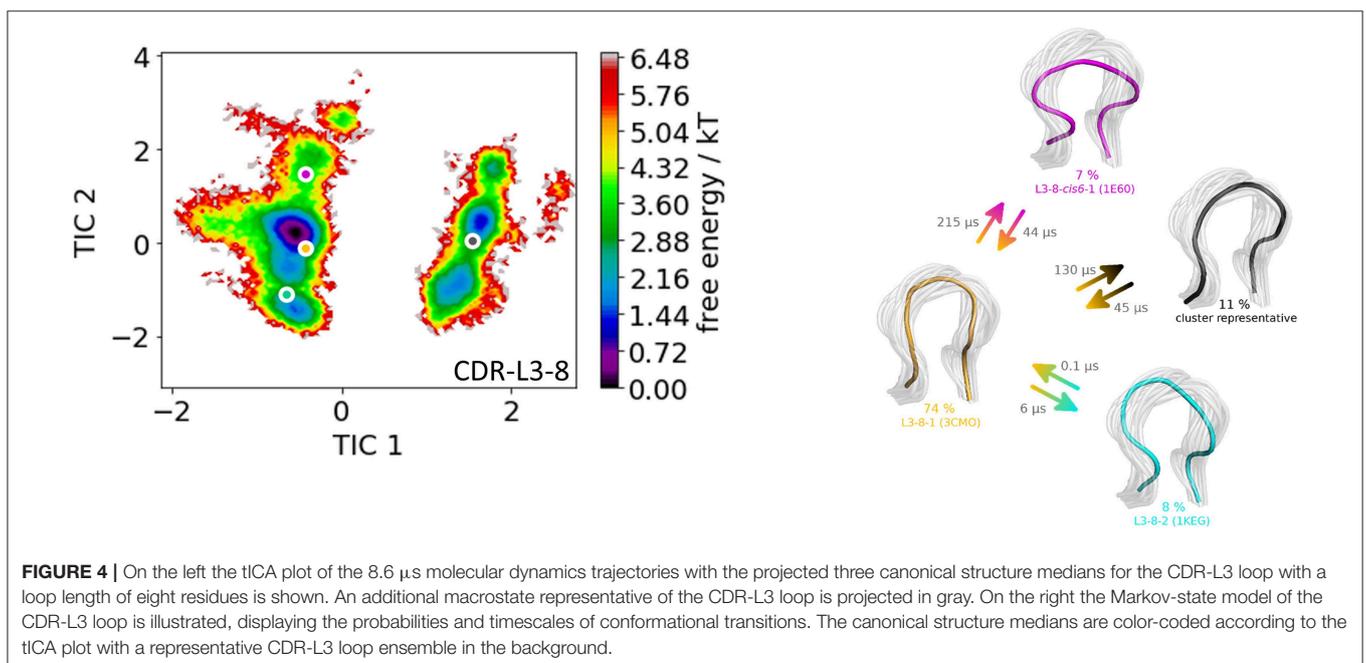
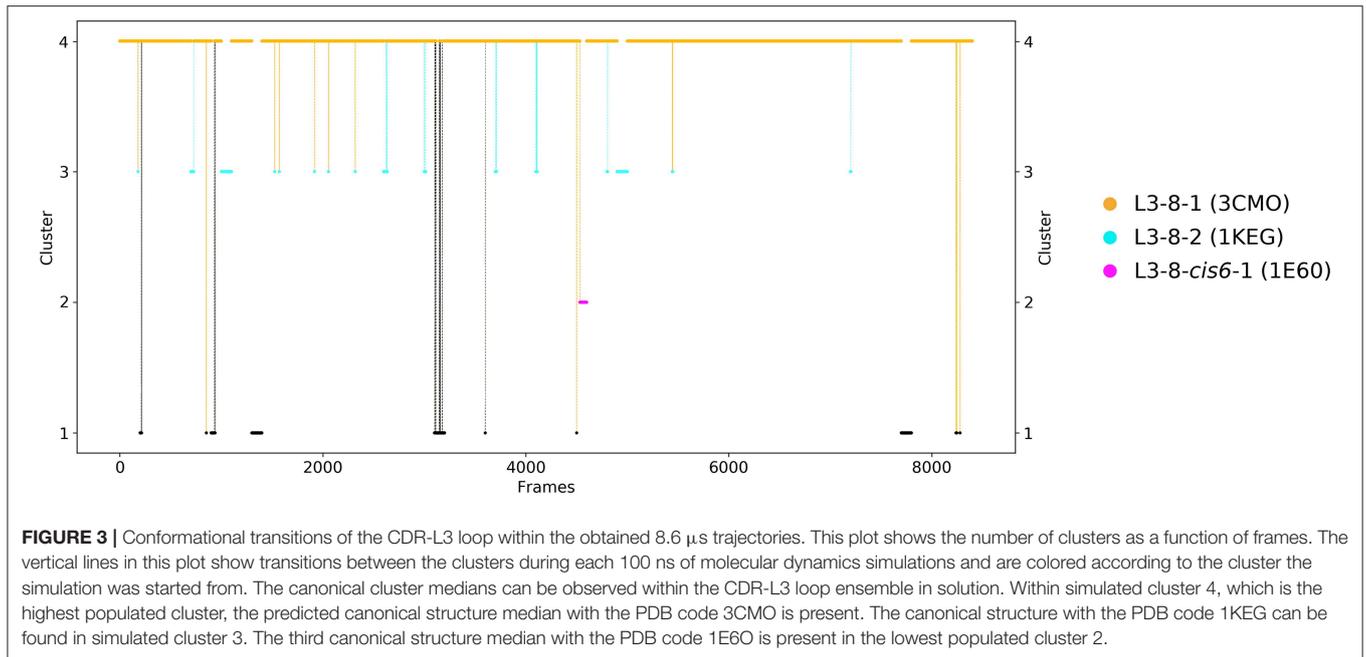
The third analyzed antibody Fv fragment is binding interleukin-13 (IL-13), which is a member of the growth-hormone-like cytokine family and plays a central role in the development of asthma (64, 65). Again, two crystal structures (3G6D and 3G6A) were available and simulated without antigen present. The CDR-L3 loop length of this IL-13 binding antibody is 10 residues. For this IL-13 binding antibody, because of its length, sequence composition and type of light chain ( $\lambda$ ) no canonical cluster could be assigned by sequence comparison. We compared the resulting CDR-L3 loop ensemble in solution to the available three canonical cluster medians of the same length. **Figure 5** shows the results of the hierarchical clustering of 8.5  $\mu$ s molecular dynamics trajectories of the CDR-L3 loop ensemble using a distance cut-off of 1.6 Å. Within the low populated cluster 4 we find canonical cluster median crystal structures 1JGU and 3B5G of the canonical clusters L3-10-cis7,8-1 and the L3-10-1, respectively. Within the least populated cluster 1 we were also able to identify the third canonical cluster median 1I7Z of the canonical cluster L3-10-cis8-1. Besides sampling transitions between the canonical clusters, we observed the highly populated clusters 2 and 3 showing various conformational transitions. To retain the kinetics and state probabilities a Markov-state model was performed to identify the dominant CDR-L3 loop solution structures. **Figure 6** displays the free energy surface with the projected canonical cluster representatives, color-coded according to **Figure 5**. Besides the local shallow side minima, in which the canonical cluster median structures are lying, **Figure 5** shows a broad free energy surface indicating the existence of other more probable and dominant CDR-L3 loop conformations in solution. The transitions between the four macrostates of this antibody occur in the nano-to-microsecond timescale, in which we again observed transitions between canonical structures.

The last antibody Fv fragment investigated is the anti-hemagglutinin binding influenza antibody (66). Three crystal



structures were available (PDB codes 1HIM, 1HIN, 1HIL) and simulated without antigen present. This anti-hemagglutinin binding antibody has a CDR-L3 loop length of nine residues and the available crystal structures were assigned to the highest populated canonical cluster L3-9-cis7-1 with the median crystal structure 1J1P. The obtained 12.7  $\mu$ s molecular dynamics trajectories were clustered with a distance cut-off of 1.1 Å and the conformational transitions are shown in **Figure 7**. Within cluster 3 the median crystal structures of the canonical clusters L3-9-cis7-2 (cluster median 1G7I) and the L3-9-1 (cluster median 1F4X) were sampled. The other four available canonical cluster

medians for the CDR-L3 loop length of nine residues were found in cluster 2. According to the hierarchical clustering the highest populated clusters are cluster 4 and cluster 1. **Figure 8** displays the Markov-state model of the CDR-L3 loop and confirms the observations of the clustering, because four canonical cluster crystal structure medians are located in the same local side-minimum. The other two canonical cluster medians 1F4X and 1G7I are situated in very unfavorable regions of another side-minimum. The most probable macrostates in **Figure 8** indicate the existence of various other dominant CDR-L3 loop solution structures.

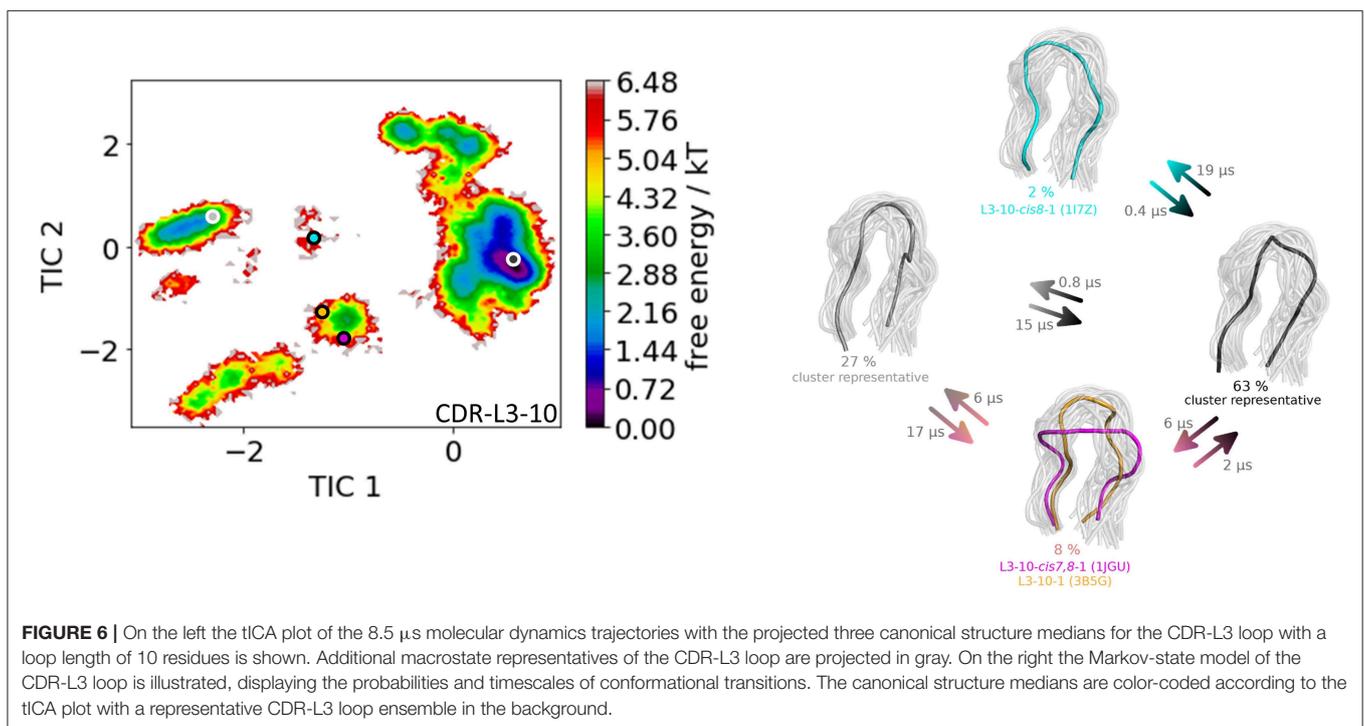
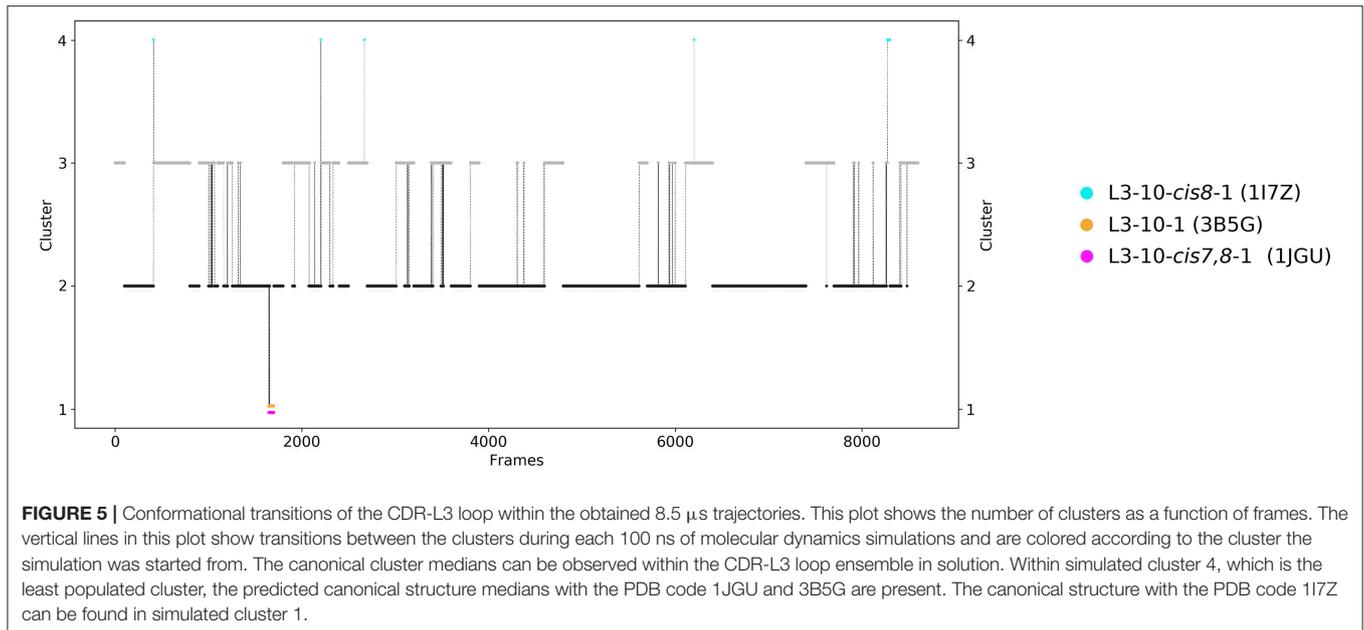


## DISCUSSION

This present study characterizes the conformational ensemble of the CDR-L3 loop and investigates conformational transitions between different canonical clusters of same length. Structural description of the CDR loops, especially the CDR-H3 and the CDR-L3 loops, are known to be a major challenge for *in silico* development of antibody biotherapeutics because of their diversity in length, sequence and structure (67). Another study focused on characterizing the stability of antigen-binding fragments in dependency of different heavy and

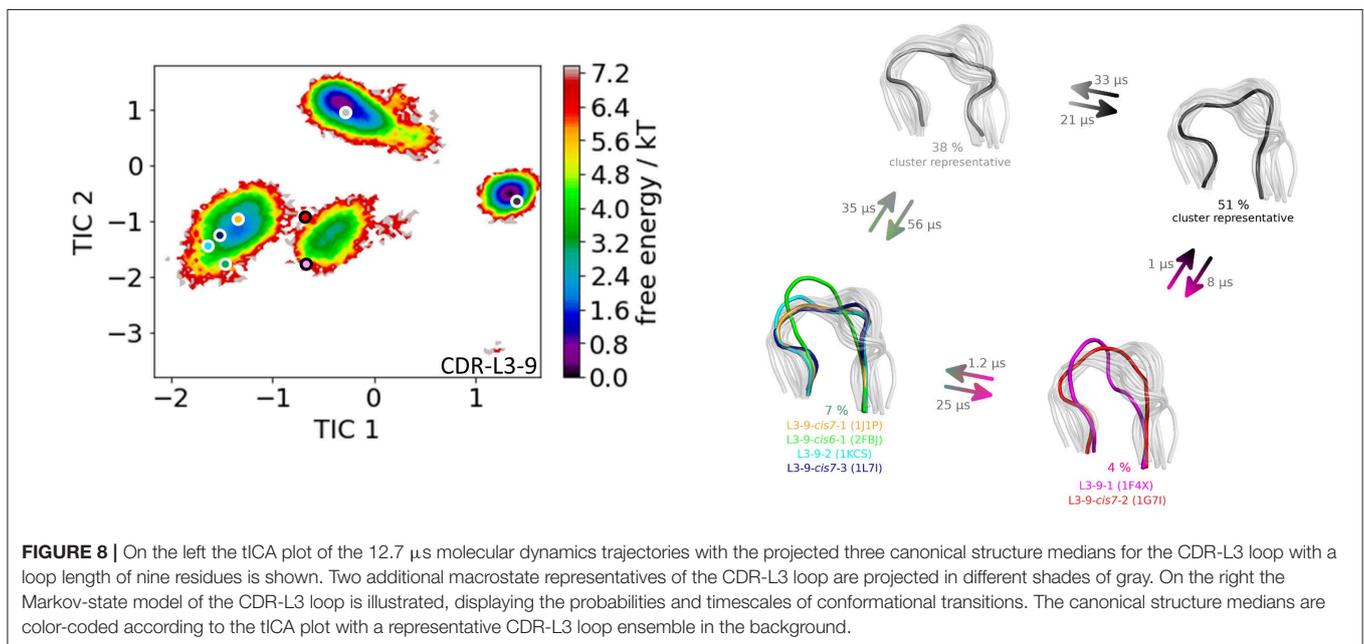
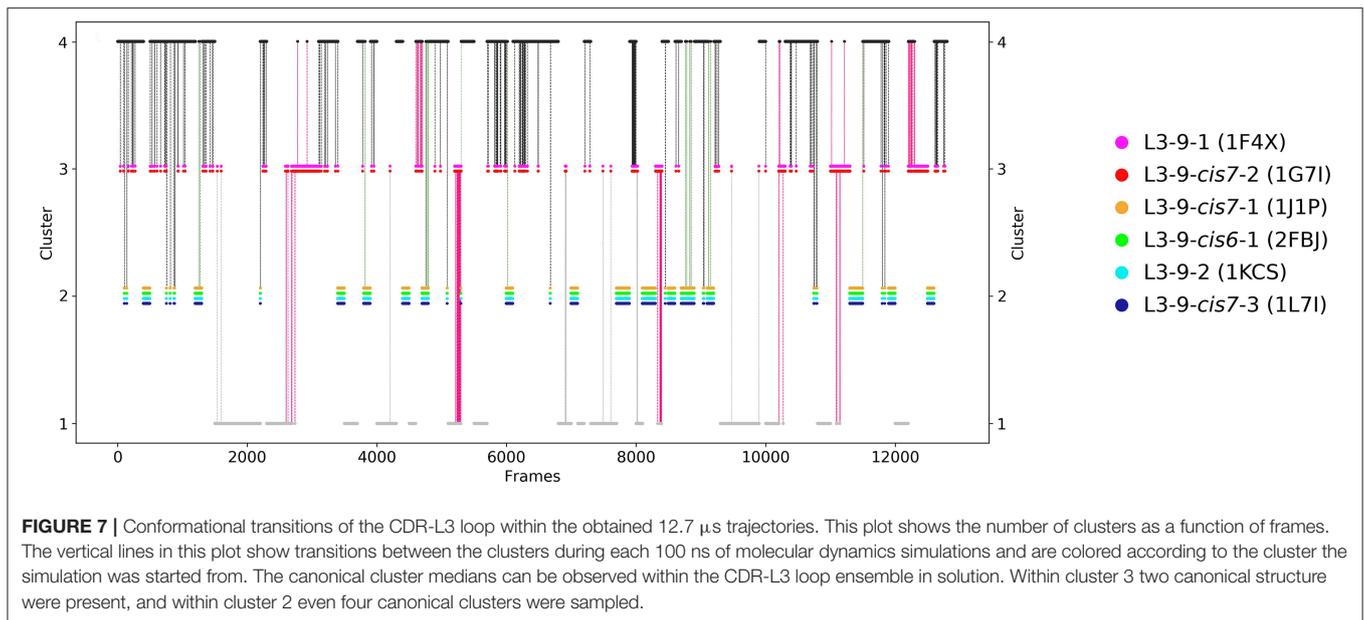
light chain pairings and the respective effect on the CDR loop conformational variability. The concept of canonical structures was supported by this investigation, suggesting that the structural repertoire could be diversified by extending beyond the human germline usage (68). The concept of conformational diversity of antibodies and the ability of the same antibody to adopt various conformations was proposed by Pauling and Landsteiner and demonstrated by Milstein and Foote (69–72).

The idea of having ensemble of pre-existing conformations out of which the functional ones are selected was supported by



population shift models originating from the Monod-Wyman-Changeux model (73–77). This new view on proteins, i.e., that one sequence can show high structural diversity, facilitated the understanding and evolution of new functions and structures (71). Proper characterization of the CDR loops, especially the loops which are mainly involved in the binding process, is crucial to understand protein-protein interactions and antigen binding. Various studies focused on classifying the CDR loops according to their loop length and sequence composition based on strong experimental structural information (6, 8, 27). We

used this experimental support to characterize the CDR-L3 loop ensemble in solution. Four different antibodies with distinct CDR-loop lengths, sequence compositions and types of light chains were used to identify functional solution structures within this ensemble of pre-existing conformations. **Figure 1** shows the results of the hierarchical clustering of the first analyzed antibody with the most prominent CDR-L3 loop length of nine residues and displays a high conformational diversity with various transitions between the four observed clusters. Comparison of this result with the six available



canonical cluster median crystal structures clearly showed that within one simulated cluster we were able to sample several canonical cluster representatives. Within the highest populated simulated cluster, the assigned canonical cluster representative of L3-9-cis7-1 (cluster median 1J1P) was present. Taking the crystal structure populations into account the L3-9-cis7-1 is the most abundant canonical cluster for all CDR-L3 loop lengths. To compare the populations observed in the PDB with our conformational ensemble in solution we calculated a Markov-state model of the CDR-L3 loop (**Figure 2**) and found two additional canonical cluster representatives close to the same global minimum of the L3-9-cis7-1 median. The representative of the L3-9-cis7-2 canonical cluster (cluster median 1G7I) is

situated in another local side-minimum and displays transition kinetics to the most probable macrostate in the microsecond timescale. Astonishingly, we were also able to sample the transition to the canonical cluster representative of the L3-9-cis7-3 cluster (cluster median 1L7I) in the high micro-to-millisecond timescale. Besides the sampling of conformational transitions between different available canonical clusters we identified an additional macrostate representative which could be an important conformation in solution. The second antibody analyzed has a CDR-L3 loop length of eight residues. Up to now only three canonical clusters could be classified for this length. Again, **Figure 3** shows the conformational transitions, as result of the hierarchical clustering, and within the highest

populated cluster we identified the assigned canonical cluster L3-8-1 (representative structure 3CMO). With a Markov-state model (Figure 4) we were able to calculate the populations and probabilities of our resulting CDR-L3 loop ensemble and in line with the observations of the first investigated antibody we identified the assigned canonical cluster representative as dominant solution structure. Additionally, we were able to sample transitions between all three canonical clusters in the microsecond timescale. Another potentially important solution structure within this ensemble was identified and is colored gray. The third studied antibody has a CDR-L3 loop length of ten residues and in this case no canonical cluster could be assigned. We compared our hierarchical clustering results (Figure 5) with the three available canonical cluster representatives, which we find within the lowest populated clusters. Besides sampling of available canonical cluster medians, we also identified two highly populated clusters being potentially relevant solution structures. The Markov-state model in Figure 6 reconstructs the kinetics and thermodynamics of the CDR-L3 loop ensemble and identifies a broad and shallow global minimum in which the dominant solution structure is present. The shallow free energy surface observed for this antibody indicates a higher conformational diversity of the CDR-L3 loop most likely originating from the lambda light chain (15). Figure 7 displays the conformational transitions of the last investigated antibody CDR-L3 loop with the length of nine amino acid residues. For this prominent and most common CDR-L3 loop length six canonical clusters were available and compared with our conformational ensemble. Four canonical cluster representatives are sampled within the second highest populated cluster 2 in our simulation. The other two canonical cluster medians were identified in simulated cluster 3. In line with the results in Figure 6, where the canonical cluster representatives are situated in local shallow side-minima, other more probable solution structures dominate in the Markov-state model in Figure 8.

For structure design our results imply, that for a given CDR-L3 loop sequence several canonical structures have to be considered. Our results also indicate that there are dominant CDR-L3 loop structures in solution, that are not apparent from X-ray analysis most likely due to crystal packing effects (31, 32). Further extensive studies of possible solution structures would be needed to decide, whether these dominant structures in solution also can be classified in new canonical structures. It is also evident,

that some of the canonical structures indeed belong to the same kinetic minimum in solution (cf. SI Table 1) and thus might be combined.

## CONCLUSION

We characterized the CDR-L3 loop ensemble in solution for different loop lengths and types of light chains. For four antibodies we were able to structurally, thermodynamically and kinetically profile the conformational space of the CDR-L3 loop in solution. Comparison of the resulting the CDR-L3 loop ensemble with the available canonical structures allowed us to calculate transition kinetics between different canonical clusters. Additionally, we identified all relevant conformations in solution. Our results clearly indicate that the static model of canonical structures should be extended to the description of the CDR-L3 loop as conformational ensemble. These findings have broad implications in the field of antibody structure design, antibody docking and might play a key role in the development of biotherapeutics as they provide a new paradigm in the understanding of CDR-L3 loop conformations and their dynamics.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/Supplementary Material.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

## FUNDING

This work was supported by the Austrian Science Fund (FWF) via the grant P30565.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2019.02652/full#supplementary-material>

## REFERENCES

- Reichert JM. Antibodies to watch in 2017. *mAbs*. (2016) 9:167–81. doi: 10.1080/19420862.2016.1269580
- Chames P, Van Regenmortel M, Weiss E, Baty D. Therapeutic antibodies: successes, limitations and hopes for the future. *Br J Pharmacol*. (2009) 157:220–33. doi: 10.1111/j.1476-5381.2009.00190.x
- Alzari PM, Lascombe MB, Poljak RJ. Three-dimensional structure of antibodies. *Annu Rev Immunol*. (1988) 6:555–80. doi: 10.1146/annurev.iy.06.040188.003011
- MacCallum RM, Martin ACR, Thornton JM. Antibody-antigen interactions: contact analysis and binding site topography. *J Mol Biol*. (1996) 262:732–45. doi: 10.1006/jmbi.1996.0548
- Chailyan A, Marcatili P, Tramontano A. The association of heavy and light chain variable domains in antibodies: implications for antigen specificity. *FEBS J*. (2011) 278:2858–66. doi: 10.1111/j.1742-4658.2011.08207.x
- Al-Lazikani B, Lesk AM, Chothia C. Standard conformations for the canonical structures of immunoglobulins. *J Mol Biol*. (1997) 273:927–48. doi: 10.1006/jmbi.1997.1354
- Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol*. (1987) 196:901–17. doi: 10.1016/0022-2836(87)90412-8
- Nowak J, Baker T, Georges G, Kelm S, Klostermann S, Shi J, et al. Length-independent structural similarities enrich the antibody CDR canonical class model. *mAbs*. (2016) 8:751–60. doi: 10.1080/19420862.2016.1158370

9. Regep C, Georges G, Shi J, Popovic B, Deane CM. The H3 loop of antibodies shows unique structural characteristics. *Proteins Struct Funct Bioinforma.* (2017) 85:1311–8. doi: 10.1002/prot.25291
10. Tonegawa S. Somatic generation of antibody diversity. *Nature.* (1983) 302:575–81. doi: 10.1038/302575a0
11. Wabl, Charles MS. Affinity maturation and class switching. *Curr Opin Immunol.* (1996) 8:89–92. doi: 10.1016/S0952-7915(96)80110-5
12. Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM. Conformations of the third hypervariable region in the VH domain of immunoglobulins11Edited by I. A. Wilson. *J Mol Biol.* (1998) 275:269–94. doi: 10.1006/jmbi.1997.1442
13. Shirai H, Kidera A, Nakamura H. H3-rules: identification of CDR-H3 structures in antibodies. *FEBS Lett.* (1999) 455:188–97. doi: 10.1016/S0014-5793(99)00821-2
14. Market E, Papavasiliou FN. V(D)J recombination and the evolution of the adaptive immune system. *PLoS Biol.* (2003) 1:e16. doi: 10.1371/journal.pbio.0000016
15. Townsend CL, Laffy MJ, Wu Y-CB, Silva O'Hare J, Martin V, Kipling D, et al. Significant differences in physicochemical properties of human immunoglobulin kappa and lambda CDR3 regions. *Front Immunol.* (2016) 7:388. doi: 10.3389/fimmu.2016.00388
16. Tomlinson IM, Cox JP, Gherardi E, Lesk AM, Chothia C. The structural repertoire of the human V kappa domain. *EMBO J.* (1995) 14:4628–38. doi: 10.1002/j.1460-2075.1995.tb00142.x
17. Kuroda D, Shirai H, Kobori M, Nakamura H. Systematic classification of CDR-L3 in antibodies: implications of the light chain subtypes and the VL–VH interface. *Proteins Struct Funct Bioinforma.* (2009) 75:139–46. doi: 10.1002/prot.22230
18. Teplyakov A, Obmolova G, Malia TJ, Luo J, Gilliland GL. Structural evidence for a constrained conformation of short CDR-L3 in antibodies. *Proteins Struct Funct Bioinforma.* (2014) 82:1679–83. doi: 10.1002/prot.24522
19. Pallarès N, Fripiat J-P, Giudicelli V, Lefranc M-P. The human immunoglobulin lambda variable (IGLV) genes and joining (IGLJ) segments. *Exp Clin Immunogenet.* (1998) 15:8–18. doi: 10.1159/000019054
20. Barbié V, Lefranc M-P. The human immunoglobulin kappa variable (IGKV) genes and joining (IGKJ) segments. *Exp Clin Immunogenet.* (1998) 15:171–83. doi: 10.1159/000019068
21. Malcolm S, Barton P, Murphy C, Ferguson-Smith MA, Bentley DL, Rabbitts TH. Localization of human immunoglobulin kappa light chain variable region genes to the short arm of chromosome 2 by *in situ* hybridization. *Proc Natl Acad Sci USA.* (1982) 79:4957. doi: 10.1073/pnas.79.16.4957
22. McBride OW, Hieter PA, Hollis GF, Swan D, Otey MC, Leder P. Chromosomal location of human kappa and lambda immunoglobulin light chain constant region genes. *J Exp Med.* (1982) 155:1480–90. doi: 10.1084/jem.155.5.1480
23. Wardemann H, Hammersen J, Nussenzweig MC. Human autoantibody silencing by immunoglobulin light chains. *J Exp Med.* (2004) 200:191. doi: 10.1084/jem.20040818
24. Marcatili P, Rosi A, Tramontano A. PIGS: automatic prediction of antibody structures. *Bioinformatics.* (2008) 24:1953–4. doi: 10.1093/bioinformatics/btn341
25. Weitzner BD, Jeliakov JR, Lyskov S, Marze N, Kuroda D, Frick R, et al. Modeling and docking of antibody structures with Rosetta. *Nat Protoc.* (2017) 12:401. doi: 10.1038/nprot.2016.180
26. Kuroda D, Shirai H, Jacobson MP, Nakamura H. Computer-aided antibody design. *Protein Eng Des Sel.* (2012) 25:507–22. doi: 10.1093/protein/gzs024
27. Adolf-Bryfogle J, Xu Q, North B, Lehmann A, Dunbrack RL. PyIgClassify: a database of antibody CDR structural classifications. *Nucleic Acids Res.* (2015) 43:D432–8. doi: 10.1093/nar/gku1106
28. Lefranc M-P, Giudicelli V, Ginestoux C, Jabado-Michaloud J, Folch G, Bellahcene F, et al. IMGT<sup>®</sup>, the international ImmunoGeneTics information system<sup>®</sup>. *Nucleic Acids Res.* (2008) 37:D1006–12. doi: 10.1093/nar/gkn838
29. Kabat EA, National institutes of health (U.S.), Columbia University. *Sequences of Proteins of Immunological Interest.* Bethesda, MD: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health (1991).
30. North B, Lehmann A, Dunbrack RL Jr. A new clustering of antibody CDR loop conformations. *J Mol Biol.* (2011) 406:228–56. doi: 10.1016/j.jmb.2010.10.030
31. Fernández-Quintero ML, Loeffler JR, Kraml J, Kahler U, Kamenik AS, Liedl KR. Characterizing the diversity of the CDR-H3 loop conformational ensembles in relationship to antibody binding properties. *Front Immunol.* (2019) 9:3065. doi: 10.3389/fimmu.2018.03065
32. Fernández-Quintero ML, Kraml J, Georges G, Liedl KR. CDR-H3 loop ensemble in solution – conformational selection upon antibody binding. *mAbs.* (2019) 11:1077–88. doi: 10.1080/19420862.2019.1618676
33. Labute P. Protonate3D: assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins.* (2009) 75:187–205. doi: 10.1002/prot.22234
34. Molecular Operating Environment (MOE). 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7. (2018).
35. Case DA, Betz RM, Cerutti DS, Cheatham TE III, Darden TA, Duke RE, et al. *AMBER 2016.* San Francisco, CA: University of California, San Francisco (2016)
36. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput.* (2013) 9:3084–95. doi: 10.1021/ct400341p
37. Hub JS, de Groot BL, Grubmüller H, Groenhof G. Quantifying artifacts in ewald simulations of inhomogeneous systems with a net charge. *J Chem Theory Comput.* (2014) 10:381–90. doi: 10.1021/ct400626b
38. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* (1983) 79:926–35. doi: 10.1063/1.445869
39. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput.* (2015) 11:3696–713. doi: 10.1021/acs.jctc.5b00255
40. Wallnofer HG, Liedl KR, Fox T. A challenging system: free energy prediction for factor Xa. *J Comput Chem.* (2011) 32:1743–52. doi: 10.1002/jcc.21758
41. Barducci A, Bussi G, Parrinello M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett.* (2008) 100:020603. doi: 10.1103/PhysRevLett.100.020603
42. Biswas M, Lickert B, Stock G. Metadynamics enhanced markov modeling of protein dynamics. *ACS Pubh.* (2018) 122:5508–16. doi: 10.1021/acs.jpcc.7b11800
43. Barducci A, Bonomi M, Parrinello M. Metadynamics. *Wiley Interdiscip Rev Comput Mol Sci.* (2011) 1:826–43. doi: 10.1002/wcms.31
44. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* (2015) 1–2:19–25. doi: 10.1016/j.softx.2015.06.001
45. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* (2013) 29:845–54. doi: 10.1093/bioinformatics/bt055
46. Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. PLUMED 2: new feathers for an old bird. *Comput Phys Commun.* (2014) 185:604–13. doi: 10.1016/j.cpc.2013.09.018
47. James LC, Tawfik DS. Structure and kinetics of a transient antibody binding intermediate reveal a kinetic discrimination mechanism in antigen recognition. *Proc Natl Acad Sci USA.* (2005) 102:12730. doi: 10.1073/pnas.0500909102
48. Shao J, Tanner SW, Thompson N, Cheatham TE. Clustering molecular dynamics trajectories: 1. characterizing the performance of different clustering algorithms. *J Chem Theory Comput.* (2007) 3:2312–34. doi: 10.1021/ct700119m
49. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J Chem Theory Comput.* (2013) 9:3878–88. doi: 10.1021/ct400314y
50. Miyamoto S, Kollman PA. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem.* (1992) 13:952–62. doi: 10.1002/jcc.540130805
51. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *J Chem Phys.* (1984) 81:3684–90. doi: 10.1063/1.448118

52. Adelman SA, Doll JD. Generalized Langevin equation approach for atom/solid-surface scattering: general formulation for classical scattering off harmonic solids. *J Chem Phys.* (1976) 64:2375–88. doi: 10.1063/1.432526
53. Nguyen H, Roe DR, Swails J, Case DA. PYTRAJ v1.0.0.dev1: *Interactive Data Analysis for Molecular Dynamics Simulations.* (2016). doi: 10.5281/zenodo.44612
54. Millman KJ, Aivazis M. Python for scientists and engineers. *Comput Sci Eng.* (2011) 13:9–12. doi: 10.1109/MCSE.2011.36
55. Scherer MK, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Plattner N, et al. PyEMMA 2: a software package for estimation, validation, and analysis of markov models. *J Chem Theory Comput.* (2015) 11:5525–42. doi: 10.1021/acs.jctc.5b00743
56. Chodera JD, Noé F. Markov state models of biomolecular conformational dynamics. *Curr Opin Struct Biol.* (2014) 25:135–44. doi: 10.1016/j.sbi.2014.04.002
57. Likas A, Vlassis N, Verbeek J. The global k-means clustering algorithm. *Biometrics.* (2003) 36:451–61. doi: 10.1016/S0031-3203(02)00060-2
58. Röblitz S, Weber M. Fuzzy spectral clustering by PCCA+: application to Markov state models and data classification. *Adv Data Anal Classif.* (2013) 7:147–79. doi: 10.1007/s11634-013-0134-6
59. Karush J. On the chapman-kolmogorov equation. *Ann Math Stat.* (1961) 32:1333–7. doi: 10.1214/aoms/1177704871
60. Miroshin RN. Special solutions of the Chapman–Kolmogorov equation for multidimensional-state Markov processes with continuous time. *Vestn St Petersburg Univ Math.* (2016) 49:122–9. doi: 10.3103/S1063454116020114
61. Wu H, Noé F. Variational approach for learning Markov processes from time series data. *J Nonlinear Sci.* (2019) 1–44. doi: 10.1007/s00332-019-09567-y
62. Chruszcz M, Pomés A, Glesner J, Vailes LD, Osinski T, Porebski PJ, et al. Molecular determinants for antibody binding on group 1 house dust mite allergens. *J Biol Chem.* (2012) 287:7388–98. doi: 10.1074/jbc.M111.311159
63. Zhang H, Liu J-H, Yang W, Springer T, Shimaoka M, Wang J-H. Structural basis of activation-dependent binding of ligand-mimetic antibody AL-57 to integrin LFA-1. *Proc Natl Acad Sci USA.* (2009) 106:18345–50. doi: 10.1073/pnas.0909301106
64. Grünig G, Corry DB, Reibman J, Wills-Karp M. Interleukin 13 and the evolution of asthma therapy. *Am J Clin Exp Immunol.* (2012) 1:20–7. Available online at: <http://www.ajcei.us/files/AJCEI1201001.pdf>
65. Teplyakov A, Obmolova G, Wu S-J, Luo J, Kang J, O’Neil K, et al. Epitope mapping of anti-interleukin-13 neutralizing antibody CNTO607. *J Mol Biol.* (2009) 389:115–23. doi: 10.1016/j.jmb.2009.03.076
66. Rini JM, Schulze-Gahmen U, Wilson IA. Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science.* (1992) 255:959–65. doi: 10.1126/science.1546293
67. Eigenbrot C, Gonzalez T, Mayeda J, Carter P, Werther W, Hotaling T, et al. X-ray structures of fragments from binding and non-binding versions of a humanized anti-CD18 antibody: structural indications of the key role of VH residues 59 to 65. *Proteins Struct Funct Bioinforma.* (1994) 18:49–62. doi: 10.1002/prot.340180107
68. Teplyakov A, Obmolova G, Malia TJ, Luo J, Muzammil S, Sweet R, et al. Structural diversity in a human antibody germline library. *mAbs.* (2016) 8:1045–63. doi: 10.1080/19420862.2016.1190060
69. Pauling L. A theory of the structure and process of formation of antibodies\*. *J Am Chem Soc.* (1940) 62:2643–57. doi: 10.1021/ja01867a018
70. Foote J, Milstein C. Conformational isomerism and the diversity of antibodies. *Proc Natl Acad Sci USA.* (1994) 91:10370–4. doi: 10.1073/pnas.91.22.10370
71. James LC, Tawfik DS. Conformational diversity and protein evolution – a 60-year-old hypothesis revisited. *Trends Biochem Sci.* (2003) 28:361–8. doi: 10.1016/S0968-0004(03)00135-X
72. James LC, Roversi P, Tawfik DS. Antibody multispecificity mediated by conformational diversity. *Science.* (2003) 299:1362–7. doi: 10.1126/science.1079731
73. Monod J, Wyman J, Changeux J-P. On the nature of allosteric transitions: a plausible model. *J Mol Biol.* (1965) 12:88–118. doi: 10.1016/S0022-2836(65)80285-6
74. Csermely P, Palotai R, Nussinov R. Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem Sci.* (2010) 35:539–46. doi: 10.1016/j.tibs.2010.04.009
75. Ma B, Kumar S, Tsai C-J, Nussinov R. Folding funnels and binding mechanisms. *Protein Eng Des Sel.* (1999) 12:713–20. doi: 10.1093/protein/12.9.713
76. Tsai C-J, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels, and protein function. *Protein Sci.* (1999) 8:1181–90. doi: 10.1110/ps.8.6.1181
77. Wedemayer GJ, Patten PA, Wang LH, Schultz PG, Stevens RC. Structural insights into the evolution of an antibody combining site. *Science.* (1997) 276:1665. doi: 10.1126/science.276.5319.1665

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Fernández-Quintero, Math, Loeffler and Liedl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.