Supporting information for:

Crystal Structure of a Chimeric Antigen Receptor (CAR) scFv Domain Rearrangement Forming a VL-VL Dimer

Jonah Cheung [1], Shagun Wazir [1], David R. Bell [2], James N. Kochenderfer [3], Wayne A. Hendrickson [4], and Philippe Youkharibache [5,*]

[1] New York Structural Biology Center, New York, NY 10027, USA
[2] Advanced Biomedical Computational Science, Frederick National Laboratory for Cancer Research (FNLCR), Frederick, MD 21701, USA
[3] Surgery Branch, Center for Cancer Research, National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD 20892, USA
[4] Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA
[5] Cancer Data Science Lab, Center for Cancer Research, National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD 20892, USA
* Correspondence: philippe.youkharibache@nih.gov

Supplemental files (File S1–S4, and Table S1) can be found in the Zenodo repository:

https://doi.org/10.5281/zenodo.7809232

Any structure file (PDB or iCn3D PNG) in the repository can be accessed as an iCn3D URL from the Zenodo API record https://zenodo.org/api/records/7809232 using the syntax: www.ncbi.nlm.nih.gov/Structure/icn3d/full.html?type=*pdb*&url=https://zenodo.org/api/files/*bucket*/*key* , where:

- *type=pdb or type=icn3dpng*,
- *bucket: 4c8a1713-cd80-4ee5-9e3c-93782c4b3193*
- *key: filename*

See examples below. Structures can be analyzed further in iCn3D. Details can be found in https://github.com/ncbi/icn3d.

File S1. List of PDB structures containing VL-VL quaternary interfaces (PDB accessed on 10/11/2021).

- See **Excel** file: **PDB-Igs-VLVL-asu+bu-with-over-10contacts.xls**
    - The file represents PDB structures that contain contacting VL domains with more than 10 arbitrary contacts in either the first biological Unit (BU) or asymmetric unit (ASU). Both BU and ASU are used as both may contain valid dimers, as in the the inverted VLVL dimer structure (7JO8)
    - Some **iCn3D links** are provided in the excel file to look and analyze a number of structures. Some links compare 7JO8 to 1REI as in Figure 2, or other VLVL parallel vs antiparallel dimer such as 1LVE vs 5LVE.
        - Where no link is specified, PDBids can be directly loaded in iCn3D using the link https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html

File S2. Alphafold2 models of 47G4 (Hu19) scFv dimers show a variety of swapped VH-VL structures.

Alphafold2 dimers models of scFvs formed from VH and VL domains tend to swap to form diabodies with long linkers as for short linkers. The 47G4 scFv used in this study uses a VL-to-VH scFv construct with a long linker (see sequence in Figure 1)

See Supplement files for:

- The 5 best Alphafold2 (version 2.2) models of a predicted 47G4 (Hu19) scFv dimer, using the sequence in Figure 1.
    - In PDB format. The 5 models are ranked by AF2 in **supplement files**:
        - **Hu19-scFv-dimer-AF2-predictions-ranked_0…4.pdb**

Files can be directly visualized and analyzed in iCn3D through the link such as: https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html?type=pdb&url=https://zenodo.org/api/files/4c8a1713-cd80-4ee5-9e3c-93782c4b3193/Hu19-scFv-dimer-AF2-predictions-ranked_0.pdb

    - The 5 models can be compared by superimposition of the VL domain of chain A: in iCn3D using an iCn3D PNG image file. To see the models and analyze them in 3D, invoke iCn3D with the link: https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html ; open "iCn3D PNG image file" using the **supplement file:**
        - **5models_Hu19_scFv_AF2_icn3d_loadable.png**

This file can be directly visualized and analyzed in iCn3D through the link: https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html?type=icn3dpng&url=https://zenodo.org/api/files/4c8a1713-cd80-4ee5-9e3c-93782c4b3193/5models_Hu19_scFv_AF2_icn3d_loadable.png

Files S3. Alphafold2 models for 47G4 (Hu19) VL-VL dimers

- In PDB format. The 25 models are ranked by AF2 in supplement files:
  - **Hu19-VLVL-AF2-ranked_0…24.pdb**
  - Only one out of 25 models offers an inverted interface but offers lower confidence. All other present the canonical interface (see Figure 2).

Files can be directly visualized and analyzed in iCn3D through URLs as described:
https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html?type=pdb&url=https://zenodo.org/api/files/4c8a1713-cd80-4ee5-9e3c-93782c4b3193/Hu19-VLVL-AF2-ranked_0.pdb
https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html?type=pdb&url=https://zenodo.org/api/files/4c8a1713-cd80-4ee5-9e3c-93782c4b3193/Hu19-VLVL-AF2-ranked_1.pdb
.
https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html?type=pdb&url=https://zenodo.org/api/files/4c8a1713-cd80-4ee5-9e3c-93782c4b3193/Hu19-VLVL-AF2-ranked_24.pdb

Files S4. VLVL based diabody models: canonical and inverted

- A model predicted by Alphafold2 version 2.0 for the 47G4 (Hu19) scFv dimer, using the sequence in Figure 1. Supplement file in PDB format:
  - **Hu19-VLVL-diabody-model.pdb**
- Model building of a scFv dimer using the observed VLVL dimer (7JO8)
  - **Hu19-inverted VLVL-diabody-model.pdb**

Files can be directly visualized and analyzed in iCn3D through URLs as described above.

- Both models compared: **VLVLDIAB_VLINVDIAB_icn3d_loadable.png**

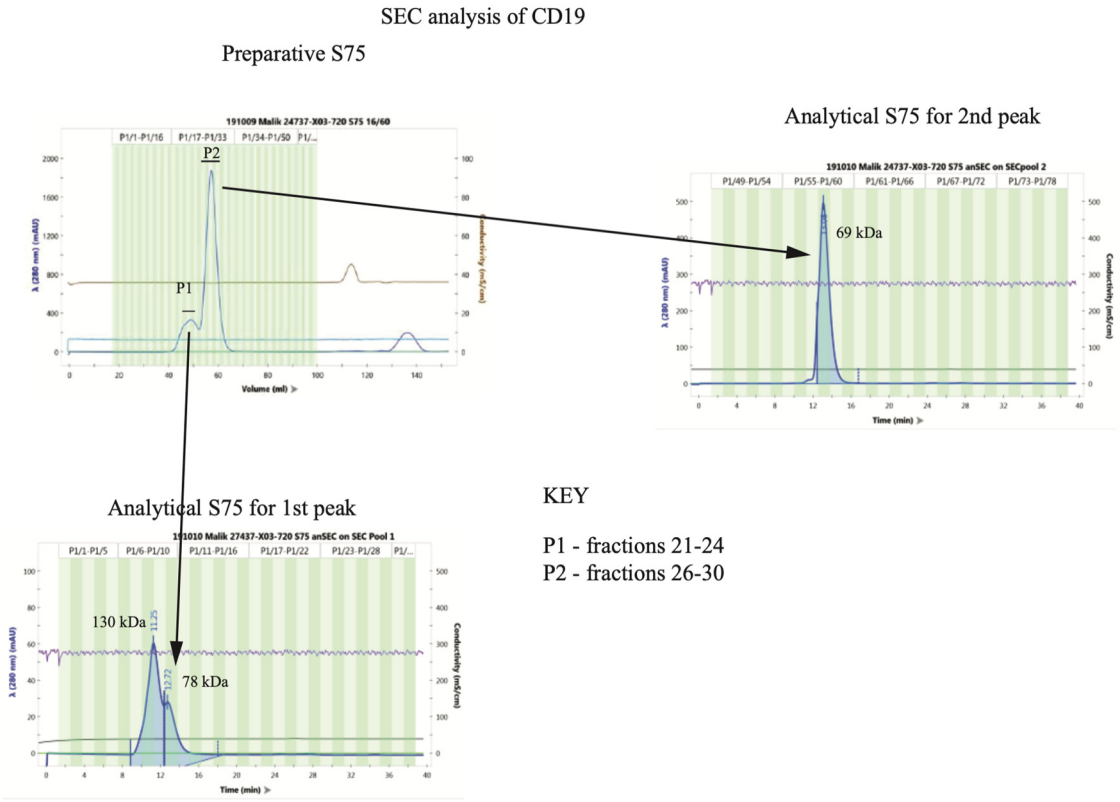
File can be directly visualized and analyzed in iCn3D through the link: https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html?type=icn3dpng&url=https://zenodo.org/api/files/4c8a1713-cd80-4ee5-9e3c-93782c4b3193/VLVLDIAB_VLINVDIAB_icn3d_loadable.png

Figure S1. SEC analysis of IMAC purified extracellular region of the anti-BCMA 47G4-CD828Z CAR

[Credit: Jane Jones, Dominic Esposito, Protein Expression Lab, FNLCR - 10/11/2019]

Protein Name	MW (kDa)	Expression host
CD8ss-CAR(CD19)-tev-His8	33.5	Exi293 (200 ml)

Purification method: IMAC, prep S75, analytical SEC for peaks from prep S75



Notes: Process outline for purification of CD19, predicted molecular weight of 33.5 kDa - 200 mls of culture dialyzed into purification buffer. The sample was then IMAC purified, pooled, and purified across a preparative size exclusion column. Two peaks were analyzed from the preparative SEC on analytical SEC to verify molecular weight. The smaller first peak runs at as a peak at 130 kDa and a trailing shoulder of 78 kDa. This is likely because there is not baseline resolution between the two peaks in the preparative column. The second, and much larger peak is 69 kDa.

Table S1. Germline genes in the inverted VLVL dimer structure (7JO8) *vs.* the canonical Bence Jones VLVL dimer (1REI)

Antibody Variable domain	Germline gene (% identity)
47G4 **VL** (7JO8)	IGKV3-20*01 (100%)
Bence-Jones protein **VL** (1REI)	IGKV1-33*01 (90.5%)
47G4 **VH** (paired to VL in scFv)	IGHV1-69*01 (97%)

[matching using IgBlast: https://www.ncbi.nlm.nih.gov/igblast/igblast.cgi]