Supplemental Materials

Somatic mutations modulate autoantibodies against galactose-deficient IgA1 in IgA nephropathy

Zhi qiang Huang[1], Milan Raska[1,2], Tyler Stewart[1], Colin Reily[1], Rodney G. King[1], David K. Crossman[1], Michael R. Crowley[1], Audra Hargett[1], Zhixin Zhang[3,7], Hitoshi Suzuki[1,4], Stacy Hall[1], Robert J. Wyatt[5], Bruce A. Julian[1], Matthew B. Renfrow[1], Ali G. Gharavi[6], Jan Novak[1]

[1]University of Alabama at Birmingham, Birmingham, AL; [2]Palacky University Olomouc, Olomouc, Czech Republic; [3]University of Nebraska Medical Center, Omaha, NE; [4]Juntendo University Faculty of Medicine, Tokyo, Japan; [5]University of Tennessee, Memphis, TN; [6]Columbia University, New York, NY; [7]Current affiliation: West China Second University Hospital, State Key Laboratory of Biotherapy, Sichuan University, Chengdu, China

Chimera Homology Modeling: PDB ID 3KDM (1) was chosen as the homology-modeling template after NCBI BLAST against the PDB databank (2) of full-length aa sequence of IgG VH from a patient with IgAN (#1123) indicated 86% identity to that of 3KDM. Chimeric aa sequences for an IgAN patient's and a healthy control's VH CDR3 sequences were made by using the corresponding VH CDR3 from the patient's and healthy control's IgG VH regions (**Supplemental Table 1**). PDB ID 3KDM heavy chain was used as a template for homology modeling of chimeric IgG sequences for IgAN patients and healthy controls. Modeling was performed using SWISS-MODEL Automatic Modeling Mode online server found at swissmodel.expasy.org (3-5). Models and structures of 3IET and 3KDM were visualized and manipulated using PyMOL v1.6 molecular graphics system (www.pymol.org).

References:

1. Niemi MH, Takkinen K, Amundsen LK, et al. The testosterone binding mechanism of an antibody derived from a naive human scFv library. *J Mol Recognit.* 2011;24(2):209-219
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403-410
3. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics.* 2006;22(2):195-201.
4. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* 2009;37(Database issue):D387-392
5. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* 2003;31(13):3381-3385

Next-Generation Sequencing on Illumina Platforms: Targeted sequencing was performed on the Illumina MiSeq using the latest version of the sequencing reagents and flow cells providing up to 8 Gb of sequence information per flow cell. The PCR product of VH-3 genes, which was generated using a generic Taq Polymerase with previously published primers (1), was ligated to the Illumina adaptors. The libraries were cleaned twice with SPRI beads and quantitated by QPCR using the Kapa kit. The final libraries were checked on the BioAnalyzer. We performed single end 1x500 bp sequencing runs to align the cDNA sequences to the reference sequences.

Data assessment: The MiSeq Reporter converted the bcl image files to fastq files. These raw fastq reads were aligned to the reference genome using BWA version 0.7.12 (2). The percent frequency of the nucleotides encoding the amino-acid sequence YCA/SR in the framework 3-CDR3 in the subset of VH3 sequences of two patients with IgAN (1123 and 1139) in first-freeze cell lines and PBMC of 1139 (recalled in 2015) was then calculated from the aligned reads of these subsets of VH3 sequences. The YCA/SR amino-acid sequence in the framework 3 was found in a small fraction of all amplified and sequenced VH3 reads.

References:

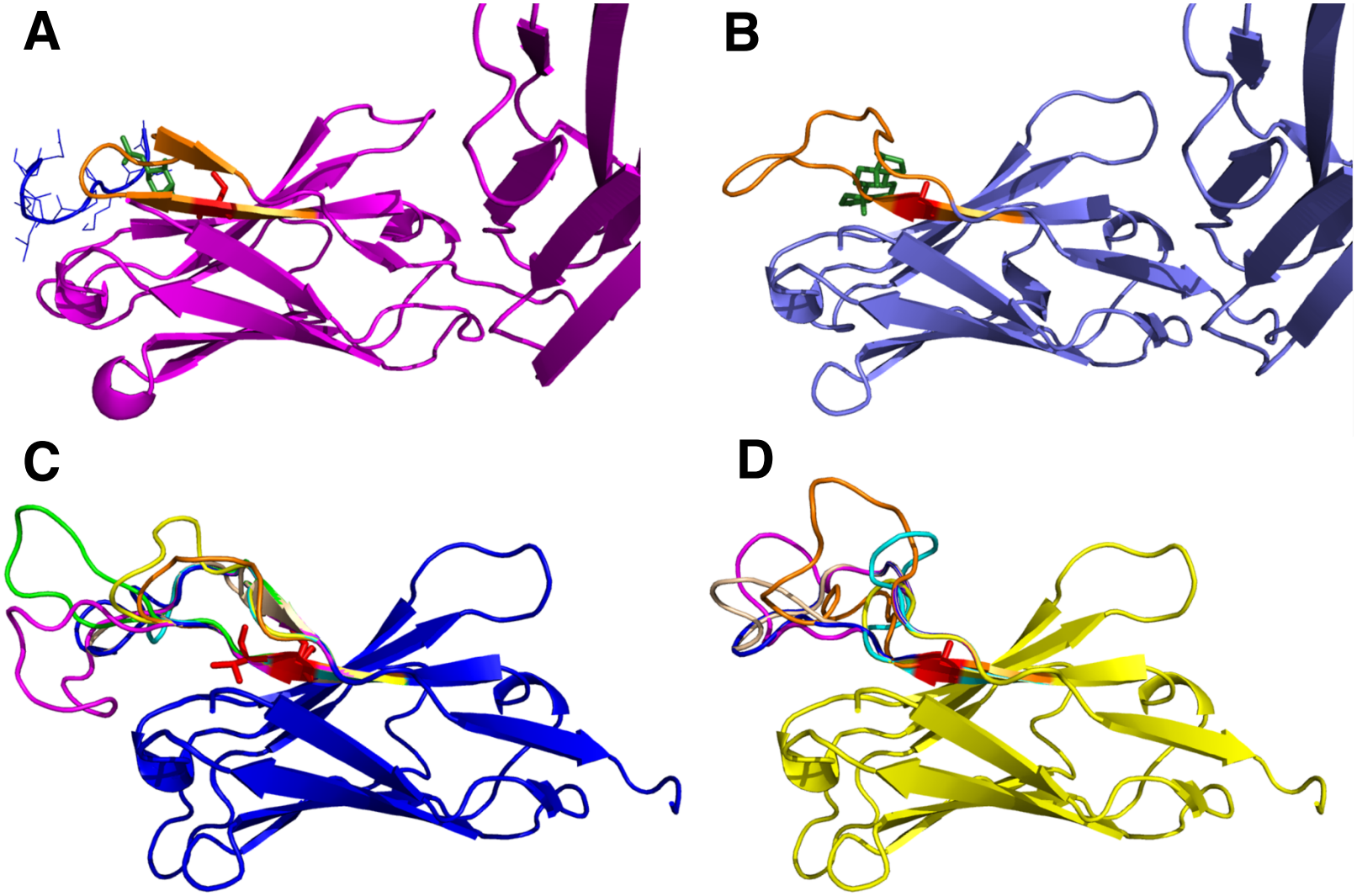
1. Suzuki H, Fan R, Zhang Z, Brown R, Hall S, Julian BA, Chatham WW, Suzuki Y, Wyatt RJ, Moldoveanu Z, Lee JY, Robinson J, Tomana M, Tomino Y, Mestecky J, Novak J: Aberrantly glycosylated IgA1 in IgA nephropathy patients is recognized by IgG antibodies with restricted heterogeneity. *J Clin Invest* 119: 1668-1677, 2009.

2. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics*, 25:1754-60, 2009.

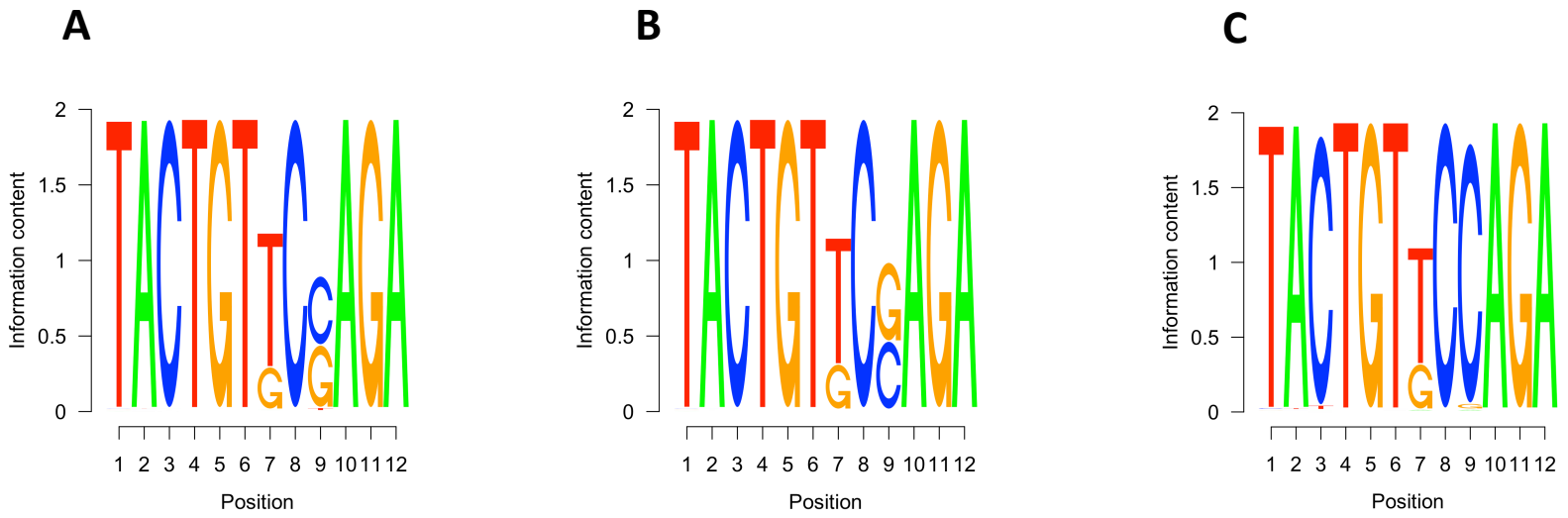
Supplemental Table 1: VH CDR3 sequences used for homology modelling

IgAN Patients' VH CDR3	
VH-1123	YCSKVCRPWNYRRPYYYGMDVW
VH-1023	YCSRDLAAFCSGGNCHSVAIDFW
VH-1125	YCSRDRYYCSGGAFDYW
VH-1139	YCSRKTSYPPTVGEVRGTSYYYGMDVW
VH-2047	YCSKTKFKGYSGFHYW
VH-3061	YCSRDRYGLFDYW
VH-3081	YCATGDYFGSGTYPIGAFDTW
	Healthy Controls' VH CDR3
VH-3064	YCARDVNITATEYYFDYW
VH-3066	YCARDLDLW
VH-3070	YCASEGHLDYGGNSDAFDIW
VH-8043	YCARGNDDYFDYW
VH-9017	YCARVQRYDSTGYYPLGYLDLW
VH-9035	YCAREWYSYLWDSSYYFDYW

Sequences are from: Suzuki H, Fan R, Zhang Z, Brown R, Hall S, Julian BA, Chatham WW, Suzuki Y, Wyatt RJ, Moldoveanu Z, Lee JY, Robinson J, Tomana M, Tomino Y, Mestecky J, Novak J: Aberrantly glycosylated IgA1 in IgA nephropathy patients is recognized by IgG antibodies with restricted heterogeneity. *J Clin Invest* 119: 1668-1677, 2009.



Supplemental Figure 1. Overview of chimeric VH CDR3 models for IgAN patients and healthy controls. A: Crystal structure of PDB 3IET, a mouse monoclonal antibody against Tn antigen (green) with S3 (red) in VH CDR3 (orange). **B:** Crystal structure of PDB 3KDM, a human monoclonal antibody against testosterone with A3 (red) in VH CDR3 (orange), was used for chimeric homology modeling of the template and backbone (orange CDR3 loop replaced with modeled CDR3 loops in chimeric models). **C:** Overlay of chimeric models of VH CDR3 of IgG from seven IgAN patients (each VH CDR3 sequence colored differently for clarity) using sequences from Supplemental Table 1. Five of seven IgAN patient models demonstrated inclusion of additional β-sheet of which all had S3 (red) in VH CDR3. Of the two models that lacked a β-sheet (patients #3061 and #3081), one did not have S3 but instead had T4 (red) and has been shown to have weaker binding to Gd-IgA1 (Suzuki et al., *J Clin Invest* 119: 1668-1677, 2009). All models demonstrated increased polar contacts across the CDR3 stem structure. **D:** Overlay of chimeric models of VH CDR3 of IgGs from six healthy controls (each VH CDR3 sequence colored differently for clarity) using sequences from Supplemental Table 1. All sequences contain A3 (red). No increased polar contact or introduction of β-sheets was observed.



Supplemental Figure 2. Sequence logos of the nucleotides representing the YCS_3R sequence in VH-3 CDR3. The size of the nucleotide is representative of the frequency of the nucleotide sequenced at that position. VH-3 and IgG-C primers were used to amplify VH-3 genes using cDNA from: **A.** first-freeze cells from patient with IgAN 1123; **B.** first freeze-cells from patient with IgAN 1139; and **C.** PBMC from patient with IgAN 1139 recalled in 2015. Next-Generation Sequencing pilot data showed that cDNA sequences corresponding to the VH-3 CDR3 region with YCSR sequence have a majority of the nucleotides encoding S (nucleotides TCC or TCG) instead of the germline-sequence-encoded A (nucleotides GCG) in the early clones of EBV-immortalized cells ("first freeze") of two patients with IgAN (1123 and 1139) as well as in primary cells (PBMC) of a recalled patient with IgAN (1139). For PBMC of patient with IgAN 1139, when we compared frequency of amino-acid sequence C(X) encoded in all productive framework 3-CDR3 sequences of VH-3, CS was encoded in 1.2% of VH-3 sequences whereas CA was encoded in 74% of VH-3 and 92% of IGHV3-21 sequences, supporting the conclusion that CA amino-acid sequence is germline-encoded. Together, these data indicate that S_3 in YCS_3R amino-acid sequence in VH CDR3 is not an artifact of EBV immortalization but occurs naturally. Analysis of IgG from serum of a patient with IgAN 1139 by high-resolution mass spectrometry further supported this conclusion (data not shown).