Rapid mass spectrometry-based sequencing of an anti-SARS-CoV-2 spike S2 antibody using multiple proteases

Jackelinne Yuka Hayashi [1], Alison Felipe Alencar Chaves [2], Daniela Luz Hessel da Cunha[3], Isaías Glezer [1], Solange Maria de Toledo Serrano[2], Alexandre Keiji Tashima [1]

[1.] UNIFESP, Federal University of São Paulo, São Paulo, Brazil ;
[2.] CeTICS, Laboratory of Applied Toxinology, Center of Toxins, Immune-Response and Cell Signaling , Instituto Butantan, São Paulo, Brazil;
[3.] IBu, Laboratory of Bacteriology, Butantan Institute, São Paulo, Brazil;

Antibodies are the secreted forms of the B-cell receptor. Antibodies, composed of two heavy and two light chains, with molecular weights of approximately 150 kDa. Two heavy chains (50 kDa each) are linked by disulfide bonds, and each heavy chain is also linked to a light chain (25 kDa each) by disulfide bonds. Each chain has a constant and a hypervariable CDR region. The latter is responsible for identifying and neutralizing antigens, exhibiting high sequence variability due to somatic recombination, hypermutation, and heavy/light chain pairing. In this context, antibody sequence information is crucial for the study, production, and validation of these important immunological agents. Antibodies are typically sequenced using mRNAs of the paired heavy and light chains. However, antibody-producing cells are not always readily available, and mass spectrometry-based methods have shown to be efficient alternatives to obtain their complete sequences. Therefore, we propose a fast bottom-up LC-MS/MS workflow for antibody sequencing. Monoclonal anti-S2 IgG1 antibody from MPbio (0872040-CF) was denaturated with Rapigest, reduced with DTT and alkylated with IAA. The antibody sample was split into five aliquots and each aliquot was digested with a different protease: trypsin/Lys-C, Glu-C, Asp-N, chymotrypsin (overnight) and thermolysin (15 minutes). Enzymatic digestions were conducted in the appropriate buffers and stopped using trifluoroacetic acid. Samples were desalted using reversed phase C18 tips (TopTips, PolyLC Inc) and peptides were quantified using Pierce Quantitative Colorimetric Assay (Thermo Fisher Scientific). Each digested sample was injected on a Vanquish Core HPLC system coupled with an Orbitrap Exploris 480 Mass Spectrometer (Thermo Fisher Scientific). Raw data were analyzed using Peaks Studio X+ (Bioinformatics Solutions Inc.) using de novo analysis and iterative database search (Taxon ID: 10090, Mus musculus). CDR sequences were completed by manually overlapping high quality de novo spectra. Sequence homology was evaluated using Basic Local Alignment Search Tool (BLAST, NIH) and IgBLAST tools. Antibody 3D structure was in silico predicted using ImmuneBuilder. Using this methodology, we assembled the heavy and light chains of the monoclonal antibody with 100% coverage. Overlapping trypsin/LysC, Glu-C and thermolysin spectra were key for the CDRs sequencing. In this study, we establish a fast method for a monoclonal antibody sequencing using only 5 LC-MS/MS runs and HCD fragmentation. Plasmids of the light and heavy chains were produced, and the monoclonal antibody will be expressed in HEK293T cells. The antibody neutralization activity against S2/Spike will be evaluated using Western blot and ELISA assays.

Agradecimentos: Monoclonal antibody anti-S2 was kindly donated by MP Biomedicals. This study was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (2022/13850-1 to J.Y.H. and 2021/05975-6 to A.K.T).