UNIVERSITY OF CALIFORNIA
SANTA CRUZ

ANTIGENICITY AND CONFORMATIONAL EPITOPES OF THE HUMAN ASTROVIRUS CAPSID

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY

by

Lena Meyer

September 2021

The Dissertation of Lena Meyer is approved:

_____
Professor Rebecca DuBois, thesis advisor

_____
Professor Seth Rubin, chair

_____
Professor Phillip Berman

_____
Peter F. Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures ..... **ix**

List of Tables ..... **xii**

Abstract ..... **xiv**

Acknowledgements ..... **xv**

Significance ..... **xvii**

Chapter 1: Human Astrovirus 1 – 8 Seroprevalence Evaluation in a United States Adult Population **1**

1.1 Acknowledgements 1

1.2 Abstract 2

1.3 Introduction 2

1.4 Materials and Methods 9

1.4.1 Human astrovirus capsid sequences used for sequence identity studies 9

1.4.2 Reagents and supplies 9

1.4.3 Expression and purification of human astrovirus spikes 1-8 10

1.4.4 Expression and purification of the SARS-CoV-2 RBD 11

1.4.5 Expression and purification of recombinant mAb 3B4 11

1.4.6 Generation of rabbit polyclonal serum to human astrovirus 1 13

1.4.7 Human plasma samples 13

1.4.8 Biolayer interferometry immunosorbent assay 13

1.5 Results 14

1.6 Discussion 23

1.7 Keywords 26

1.8 References 26

Chapter 2: A Simplified Workflow for Monoclonal Antibody Sequencing 31

2.1 Acknowledgements 31

2.2 Abstract 32

2.3 Introduction 33

2.4 Results 36

2.4.1 Monoclonal antibody sequencing strategy 36

2.4.2 Primer selection 41

2.4.3 Sequencing results 43

2.4.4 Verification of antigen binding: Comparison of chimeric mAb 2D9 and mouse mAb 2D9 48

2.4.5 Proof-of-concept: RT-PCR amplification of RNA from chimeric antibodies expressed in a human cell line 50

2.5 Discussion 53

2.6 Conclusions 55

2.7 Materials and Methods 55

2.7.1 Hybridoma production and total RNA extraction 55

2.7.2 Reverse transcription to synthesize cDNA of antibody variable regions 56

2.7.3 PCR amplification of antibody variable regions 59

2.7.4 Gel extraction and sequencing or PCR clean-up, blunt-end cloning, miniprep, and sequencing of antibody variable regions 59

2.7.5 Expression and purification of the Spike 8 antigen 61

2.7.6 Expression and purification of chimeric mAb 2D9 62

2.7.7 Total RNA extraction from transfected HEK 293F cells 63

2.7.8 Purification of mouse mAb 2D9 63

2.7.9 SDS-PAGE gel comparing chimeric mAb 2D9 and mouse mAb 2D9 63

2.7.10 ELISAs comparing Spike 8 binding by chimeric mAb 2D9 and mouse mAb 2D9 64

2.8 Software availability 65

2.9 Abbreviations 65

2.10 Keywords 66

2.11 References 66

Chapter 3: Structures of Two Human Astrovirus Capsid/Neutralizing Antibody Complexes Reveal Distinct Epitopes and Antibody Inhibition of Virus Attachment to Host Cells 71

3.1 Acknowledgements 71

3.2 Abstract 73

3.3 Importance 74

3.4 Introduction 74

3.5 Results 78

3.5.1 Neutralizing antibodies 3E8 and 2D9 have high affinity for the HAstV8 capsid spike 78

3.5.2 scFv 3E8 binds Spike 8 at a known epitope 80

3.5.3 scFv 2D9 binds Spike 8 at a novel epitope 84

3.5.4 Antibodies 3E8, 2D9, and PL-2 target epitopes near the P site 88

3.5.5 Antibodies 3E8 and 2D9 block attachment of HAstV8 to Caco-2 cells 89

3.5.6 Antibodies 3E8 and 2D9 block attachment of GFP-Spike 8 to Caco-2 cells 91

3.6 Discussion 92

3.7 Materials and Methods 95

3.7.1 Expression and purification of scFv 3E8 95

3.7.2 Expression and purification of scFv 2D9 96

3.7.3 Expression and purification of the Spike 8 antigen 97

3.7.4 Biolayer interferometry affinity determination of scFv 2D9 and scFv 3E8 for Spike 8 98

3.7.5 Formation and structure determination of the scFv 3E8/Spike 8 complex 100

3.7.6 Formation and structure determination of the scFv 2D9/Spike 8 complex 101

3.7.7 HAstV8 infectivity neutralization assay 102

3.7.8 HAstV8 attachment inhibition assay 102

3.7.9 HAstV8 detachment assay 103

3.7.10 Expression and purification of chimeric Fabs 3E8, 2D9, and 3B4 104

3.7.11 GFP-Spike 8 attachment inhibition assay by fluorescence microscopy 105

3.7.12 Data availability 106

3.8 Abbreviations 107

3.9 References 107

Appendix 1: Crystallization Attempts of 3B4/Spike 1 Complexes 114

A1.1 Introduction 114

A1.2 Crystallization attempts of scFv 3B4/Spike 1 115

A1.3 Crystallization attempts of papain-digested chimeric Fab 3B4/Spike 1..... 121

A1.4 Crystallization attempts of chimeric Fab 3B4 Strep/Spike 1..... 124

A1.5 Crystallization attempts of thrombin-digested chimeric Fab 3B4/Spike 1.. 124

Appendix 2: Variable Region Sequencing of Anti-VA1 Antibodies 2A2 and 7C8..... 128

A2.1 Introduction..... 128

A2.2 RT-PCR on 2A2 and 7C8 128

A2.3 Sequencing results of 2A2 and 7C8..... 130

List of Figures

Chapter 1: Human Astrovirus 1 – 8 Seroprevalence Evaluation in a United States Adult Population 1

Figure 1.1: The human astrovirus capsid. 4

Figure 1.2: Purification of recombinant human astrovirus 1-8 capsid spikes (Spike 1 – Spike 8). 15

Figure 1.3: Validation of BLI-ISA for the detection of plasma IgG antibodies to human astrovirus capsid spikes. 16

Figure 1.4: Human plasma IgG reactivity to human astrovirus spike proteins. 20

Chapter 2: A Simplified Workflow for Monoclonal Antibody Sequencing 31

Figure 2.1: Schematic for cDNA synthesis by template-switching. 39

Figure 2.2: Comparison of primer sets for RT-PCR amplification of variable regions from 5 hybridoma mRNA samples. 41

Figure 2.3: Protein sequence comparison of variable regions for 3H4 kappa and 3H4 lambda. 48

Figure 2.4: Comparison of chimeric mAb 2D9 and mouse mAb 2D9. 49

Figure 2.5: RT-PCR amplification of chimeric antibody variable regions. 52

Chapter 3: Structures of Two Human Astrovirus Capsid/Neutralizing Antibody Complexes Reveal Distinct Epitopes and Antibody Inhibition of Virus Attachment to Host Cells 71

Figure 3.1: scFv 3E8 and scFv 2D9 bind Spike 8 with high affinity. 79

Figure 3.2: Intermolecular interactions at the scFv 3E8/Spike 8 interface. 81

Figure 3.3: Intermolecular interactions at the scFv 2D9/Spike 8 interface. 85

Figure 3.4: Comparison of epitopes for 3E8/Spike 8, 2D9/Spike 8, and PL-2/Spike 2 in relation to the P site. 87

Figure 3.5: Antibodies 3E8 and 2D9 reduce virus infectivity, block virus attachment, and detach the virus from Caco-2 cells. 90

Figure 3.6: Antibodies 3E8 and 2D9 block GFP-Spike 8 attachment to Caco-2 cells. 92

Appendix 1: Crystallization Attempts of 3B4/Spike 1 Complexes 114

Figure A1.1: scFv 3B4/Spike 1 Prep 5 complex preparation from 7/16/19 – 7/17/19. 116

Figure A1.2: scFv 3B4/Spike 1 crystals in sodium malonate pH 7.0. 120

Figure A1.3: scFv 3B4/Spike 1 crystals in sodium malonate pH 7.0 one year after setup. 121

Figure A1.4: Chimeric Fab 3B4/Spike 1 Prep 1 complex preparation from 10/30/19 – 11/1/19. 123

Figure A1.5: Chimeric Fab 3B4/Spike 1 Prep 1 complex preparation from 10/11/20 – 10/15/20. 126

Appendix 2: Variable Region Sequencing of Anti-VA1 Antibodies 2A2 and 7C8..... 128

Figure A2.1: RT-PCR of kappa, lambda, and heavy chains using 2A2 Tube 1 and 7C8 Tube 1. 129

Figure A2.2: RT-PCR of kappa and heavy chains using 2A2 Tube 1 and Tube 2. 130

Figure A2.3: Sequence alignment of 2A2 kappa from Tube 1 (first three sequences) and Tube 2 (last sequence). 131

Figure A2.4: Sequence alignment of 2A2 heavy from Tube 2. 131

Figure A2.5: Sequence alignment of 7C8 kappa from Tube 1. 131

Figure A2.6: Sequence alignment of 7C8 heavy from Tube 1. 132

List of Tables

Chapter 1: Human Astrovirus 1 – 8 Seroprevalence Evaluation in a United States Adult Population 1

Table 1.1: Sequence identity matrix across capsid core domains from human astrovirus serotypes 1–8. 5

Table 1.2: Sequence identity matrix across capsid spike domains from human astrovirus serotypes 1 – 8. 5

Table 1.3: Summary of previous human astrovirus serological surveys. 7

Table 1.4: Heat map of individual human plasma IgG reactivity to human astrovirus spike antigens. 21

Chapter 2: A Simplified Workflow for Monoclonal Antibody Sequencing 31

Table 2.1: Mouse IgG reverse transcription primers. 37

Table 2.2: Mouse IgG PCR primers. 37

Table 2.3: Results of sequencing RT-PCR products directly and following blunt-end cloning. 44

Table 2.4: Percent identity to IgBLAST and IMGT reference sequences. 46

Table 2.5: Human IgG reverse transcription primers. 51

Table 2.6: Human IgG PCR primers. 51

Table 2.7: Kits and antibodies. 56

Chapter 3: Structures of Two Human Astrovirus Capsid/Neutralizing Antibody Complexes Reveal Distinct Epitopes and Antibody Inhibition of Virus Attachment to Host Cells 71

Table 3.1: scFv 3E8 and scFv 2D9 bind Spike 8 with high affinity. 79

Table 3.2: Data collection and refinement statistics for scFv 3E8/Spike8 and scFv 2D9/Spike 8. 82

Appendix 1: Crystallization Attempts of 3B4/Spike 1 Complexes 114

No tables

Appendix 2: Variable Region Sequencing of Anti-VA1 Antibodies 2A2 and 7C8 128

No tables

Abstract

Antigenicity and Conformational Epitopes of the Human Astrovirus Capsid

Lena Meyer

Human astrovirus is an important cause of viral gastroenteritis worldwide. Young children, the elderly, and the immunocompromised are especially at risk. Few studies have been conducted examining human astrovirus incidence and seroprevalence to properly assess disease burden and significance. Furthermore, no vaccines or antiviral therapies exist to address human astrovirus infection, but several lines of evidence suggest that a protective antibody response mounted against a childhood infection can prevent future reinfection in adulthood. Therefore, we expect that vaccines eliciting such an antibody response will protect vulnerable individuals from disease. However, successful design of vaccine antigens requires knowledge of how neutralizing antibodies target sites of importance. In this thesis, I discuss my work on the antigenicity and conformational epitopes of the human astrovirus capsid. I begin with our seroprevalence evaluation of all eight classical human astrovirus serotypes in the United States, followed by a review of our simplified workflow for monoclonal antibody sequencing. Finally, I present a structural and mechanistic characterization of two neutralizing monoclonal antibodies which target the human astrovirus capsid spike at two separate, non-overlapping epitopes and both block virus attachment to host cells. These studies provide a basis for the development of therapies to prevent and treat human astrovirus disease.

Acknowledgements

I would first like to thank my advisor, Rebecca, for all the help and support these six years. I decided to join the DuBois lab because I believed that from you, I would receive the best care and the guidance to become a real scientist like I've always wanted to be. Thank you. Thank you also to my thesis committee members, Seth and Phil, for the many suggestions to make my research stronger. Thank you to the wider PBSE community, especially my cohort Tori and Terren, my roommates Ariel and Christina, and the holiday party planning committee. To our collaborators in the Arias lab, thank you for making all our projects more well-rounded. Thank you to DuBois labmates past and present: Walter and Stas for showing me the ropes, Natasha for the cookies and *E. coli* prep advice, Ed for the scurvy prevention, Jacob for the epic music, Nick and Ana for the birthday cards and hiking, Josh for spelling with amino acids and testing out my protocols, John for the endless crystallography advice, Sarah for the succulents and for taking over the astrovirus antibodies study, Maria for the infectious laugh, Santanu for the quiet humor and microscopy expertise, and especially to Carolina, Kevin, and Jordan – thank you for your work to support my projects and for allowing me to learn how to mentor on you.

To my science teachers in school and at university, especially Mrs. Wynbeek, and to my undergraduate advisors, John and Dylan, thank you for laying the foundation on which I could build the skills and knowledge I acquired during my Ph.D. studies. To my many doctors, thank you for coming up with treatment plans that allowed me to focus on science instead of health issues.

To my parents, Uli and Wolfgang, thank you for showing me your love, always believing in me, and supporting my path. I want to be just like you in how you enjoy life together and welcome everyone to our home. To my siblings, Max, Nick, and Lilly, thank you for your admiration and for being the first I could practice leadership skills on. I'm so happy we are all great friends as adults. To my grandparents, aunts, uncles, and cousins in Germany, thank you for the excitement from afar at any small success. Alison, thank you for always being by my side my whole life and for understanding every part of the grad school journey. To other friends – Lauren, Emily, Celine, Theresa, Kristof, Mali'o, Nikki – thank you for providing the balance in my work/life equation. To my almost in-laws, Karl and Leslie, thank you for welcoming me into your family so wholeheartedly. And to my fiancé, Edward, thank you for being the best partner, friend, supporter, and cook. I am so thankful that undertaking this Ph.D. also led me to you.

Significance

Human astrovirus is an important cause of viral gastroenteritis worldwide. Young children, the elderly, and the immunocompromised are especially at risk for suffering from severe or chronic disease, particularly in developing countries. In addition, astroviruses infect birds and a wide range of other mammals, indicating the potential for a zoonotic disease transmission and the rise of new astrovirus strains that could threaten human health. No vaccines or antiviral therapies exist to address human astrovirus infection. In addition, few studies have been conducted that examine human astrovirus incidence and seroprevalence. Although evidence indicates that antibodies play an important role in protecting healthy adults from reinfection, few neutralizing monoclonal anti-astrovirus antibodies have been thoroughly studied to understand the epitopes on the virus surface to which antibodies bind and the antibody mechanisms of neutralization. To develop an effective subunit vaccine that broadly protects against diverse human astrovirus serotypes, we must understand how neutralizing antibodies target the human astrovirus capsid at the molecular level.

In this thesis, I begin with a discussion of a human astrovirus seroprevalence evaluation in a United States adult population, in which all eight classical human astrovirus serotypes were assessed for the first time. Our study, the first of its kind in implementing the recently developed biolayer interferometry immunosorbent assay (BLI-ISA) as a serosurveillance technique, establishes the seroprevalence rates of IgG antibodies as 73 % for human astrovirus serotype 1, 62 % for serotype 3, 52 % for serotype 4, 29 % for serotype 5, 27 % for serotype 8, 22 % for serotype 2, 8 % for

serotype 6, and 8 % for serotype 7. Importantly, seroprevalence rates towards the capsid spike protein established by BLI-ISA correlate with previously determined neutralizing antibody rates towards the entire virus.

In Chapter 2, I pivot to a discussion of a simplified, cost-effective, and rapidly executed workflow for monoclonal antibody sequencing that we developed to determine the variable region sequences of five novel neutralizing antibodies targeting the human astrovirus capsid spikes from three different serotypes. This method amplifies IgG antibody variable regions from hybridoma RNA by a specialized RT-PCR followed by blunt-end cloning and Sanger sequencing. Reverse transcription occurs with a primer specific to the antibody heavy or light chain constant region and a template-switch oligonucleotide, which creates a custom sequence at the 5' end of the antibody cDNA. Subsequent PCR amplification is accomplished with a primer specific for the template-switch oligonucleotide sequence and a nested primer to the respective constant region. This method enabled sequence determination of all five antibodies with 100 % accuracy.

Successful sequence determination of the antibodies in Chapter 2 permitted recombinant expression and further study of the targeted epitopes on the human astrovirus capsid spike surface and examination of the mechanisms of antibody neutralization. In Chapter 3, I discuss the structure-based discovery that two of these antibodies, 3E8 and 2D9, target the human astrovirus serotype 8 spike at two separate, non-overlapping epitopes. In addition, we show that both antibodies neutralize the human astrovirus by blocking virus attachment to host cells. We point to conserved

amino acids near or within the antibody epitopes which may comprise a receptor-binding site. The findings reported in this thesis underscore the importance of human astrovirus to disease burden and lay the foundation for the development of therapies that prevent and treat human astrovirus gastroenteritis.

Chapter 1: Human Astrovirus 1 – 8 Seroprevalence Evaluation in a United States Adult Population

1.1 Acknowledgements

This chapter consists of a manuscript reprint of our human astrovirus seroprevalence study published in 2021 in *Viruses*: Meyer, L.; Delgado-Cunningham, K.; Lorig-Roach, N.; Ford, J.; DuBois, R.M. Human Astrovirus 1–8 Seroprevalence Evaluation in a United States Adult Population. *Viruses* 2021, 13, 979. https://doi.org/10.3390/v13060979

My specific contributions to this project include expression and purification of human astrovirus capsid spikes 1 – 8 and recombinant mAb 3B4 as well as execution of all BLI-ISA experiments. I also gave assistance with data analysis of the experiments. I made Figure 1.3B and 1.3C and generated Tables 1.1 – 1.3. I wrote the first draft and further edited the manuscript.

Kevin Delgado-Cunningham and Jordan Ford assisted with the expression and purification of human astrovirus capsid spikes 1 – 8. Kevin Delgado-Cunningham also made Figures 1.1 and 1.2 and wrote the methods section related to the spike expressions. Nicholas Lorig-Roach performed data analysis of all BLI-ISA experiments, expressed and purified the SARS-CoV-2 RBD, and generated Figure 1.4 and Table 1.4. Rebecca M. DuBois conceptualized the overall goals for the study, made Figure 1.3A, acquired funding, supervised the work related to the project, and reviewed and edited the manuscript. This research was funded by the National Institutes of

Health under NIH grant R01 AI144090. We also acknowledge Dr. Dorsey Bass for the donation of rabbit polyclonal antibody serum to HAstV1.

I thank my coauthors for their success in accomplishing a quick turnaround on this project and for their permission to reproduce our manuscript here.

1.2 Abstract

Human astroviruses are an important cause of viral gastroenteritis globally, yet few studies have investigated the serostatus of adults to establish rates of previous infection. Here, we applied biolayer interferometry immunosorbent assay (BLI-ISA), a recently developed serosurveillance technique, to measure the presence of blood plasma IgG antibodies directed towards the human astrovirus capsid spikes from serotypes 1-8 in a cross-sectional sample of a United States adult population. The seroprevalence rates of IgG antibodies were 73 % for human astrovirus serotype 1, 62 % for serotype 3, 52 % for serotype 4, 29 % for serotype 5, 27 % for serotype 8, 22 % for serotype 2, 8 % for serotype 6, and 8 % for serotype 7. Notably, seroprevalence rates for capsid spike antigens correlate with neutralizing antibody rates determined previously. This work is the first seroprevalence study evaluating all eight classical human astrovirus serotypes.

1.3 Introduction

Astroviruses are a diverse family of small, nonenveloped, positive-sense RNA viruses that infect mammalian and avian species [1.1]. Astrovirus infection is linked to a variety of disease manifestations, growth defects, and mortality in poultry [1.2]. In

mammals, astrovirus infection mainly causes viral gastroenteritis but can also be asymptomatic [1.3, 1.4] or cause neurological syndromes and encephalitis in rare cases [1.5-1.8]. Human astroviruses are classified into three clades: classical serotypes 1-8, where serotype 1 is the most prevalent globally [1.9-1.11], as well as the emerging serotypes MLB and VA [1.5, 1.8].

Human astroviruses (HAstVs) are a leading worldwide cause of viral gastroenteritis but are among the most poorly characterized enteric viruses [1.12]. Young children, the elderly, and the immunocompromised are most threatened by astrovirus infection, especially in developing countries [1.13-1.20]. Worldwide, human astrovirus infection accounts for approximately 2 to 9 % of all acute non-bacterial gastroenteritis in healthy children [1.21]. The United States sees an estimated 3.9 million cases of astrovirus gastroenteritis each year [1.22]. However, immunofluorescence studies have demonstrated that about 75 % of healthy adults have anti-astrovirus antibodies targeting at least one of the eight classical serotypes [1.23], and another investigation showed that seroprevalence of neutralizing antibodies increases with age [1.10]. These findings suggest that cases of astrovirus gastroenteritis may be undercounted and that astrovirus disease may actually be a common childhood infection.

Several lines of evidence highlight the importance of anti-astrovirus antibodies developed in childhood in preventing reinfection in adulthood. Firstly, astrovirus infection is rare in adults [1.23]. Although approximately 75 % of adults have anti-astrovirus antibodies, clinical investigations in healthy volunteers determined that more

severe disease is associated with seronegativity for anti-astrovirus antibodies [1.24, 1.25]. Finally, immunoglobulin replacement therapy resolved a persistent human astrovirus infection in an immunocompromised patient [1.26]. These studies indicate that a vaccine and therapeutic antibodies could be developed to prevent and/or treat human astrovirus gastroenteritis.

Previous structural studies have defined the structural domains of the human astrovirus capsid (Figure 1.1A and 1.1B) [1.27-1.30]. The capsid core domain, which forms the structural icosahedral shell encapsulating the viral genome, remains highly conserved across all eight classical serotypes, ranging from 83.3 – 97.0 % sequence identity between any two serotypes (Table 1.1). In contrast, the capsid spike domain, which forms dimeric protrusions on the surface of the viral particle, is quite variable, with a spread of 41.4 – 75.7 % sequence identity between any two serotypes (Table 1.2). Prior work using enzyme-linked immunosorbent assays showed that both the core and spike domains are antigenic [1.28].

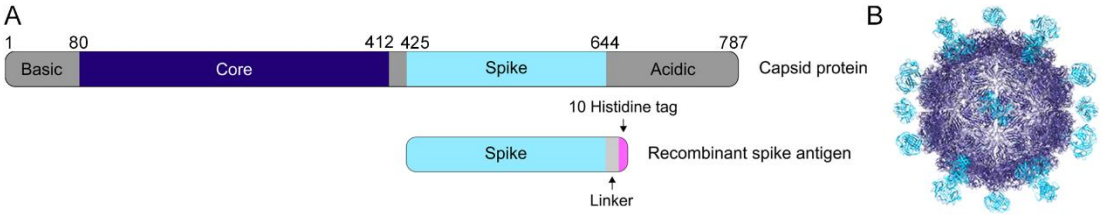

Figure 1.1: The human astrovirus capsid.

(A) Schematic of the human astrovirus capsid protein and the recombinant spike antigen used in this study. **(B)** The mature human astrovirus particle, with the spike in cyan and the core in dark blue (adapted from [1.28]).

Table 1.1: Sequence identity matrix across capsid core domains from human astrovirus serotypes 1–8.

	Core 1	**Core 2**	**Core 3**	**Core 4**	**Core 5**	**Core 6**	**Core 7**	**Core 8**
Core 1	100 %	86.7 %	91.5 %	86.1 %	90.0 %	90.9 %	87.9 %	86.7 %
Core 2	-	100 %	86.7 %	87.9 %	83.3 %	83.9 %	85.2 %	87.9 %
Core 3	-	-	100 %	87.3 %	90.6 %	91.2 %	92.8 %	88.2 %
Core 4	-	-	-	100 %	84.6 %	86.1 %	84.9 %	97.0 %
Core 5	-	-	-	-	100 %	90.3 %	88.8 %	85.8 %
Core 6	-	-	-	-	-	100 %	87.9 %	87.6 %
Core 7	-	-	-	-	-	-	100 %	85.8 %
Core 8	-	-	-	-	-	-	-	100 %

Table 1.2: Sequence identity matrix across capsid spike domains from human astrovirus serotypes 1 – 8.

	Spike 1	**Spike 2**	**Spike 3**	**Spike 4**	**Spike 5**	**Spike 6**	**Spike 7**	**Spike 8**
Spike 1	100 %	50.2 %	61.2 %	41.4 %	47.0 %	53.1 %	58.9 %	53.0 %
Spike 2	-	100 %	46.7 %	47.9 %	41.8 %	47.0 %	44.9 %	49.8 %
Spike 3	-	-	100 %	46.3 %	48.6 %	54.7 %	75.7 %	55.1 %
Spike 4	-	-	-	100 %	46.0 %	47.0 %	43.5 %	46.1 %
Spike 5	-	-	-	-	100 %	63.3 %	52.8 %	52.1 %
Spike 6	-	-	-	-	-	100 %	56.6 %	57.3 %
Spike 7	-	-	-	-	-	-	100 %	53.3 %
Spike 8	-	-	-	-	-	-	-	100 %

Serosurveillance investigations bring the frequency of previous human astrovirus infection in a particular cohort into focus, but only a few have been performed on specific astrovirus subsets (classical serotypes 1-7, MLB1, HMOAstV-C (VA1), or turkey astrovirus 2) or in specific populations (turkey growers and abattoir workers) (Table 1.3) [1.10, 1.23, 1.31-1.36]. Many of the studies focusing on classical human astrovirus were conducted more than 20 years ago. Importantly, no previous serological survey has evaluated seropositivity to all eight classical human astrovirus serotypes. In addition, these studies used a variety of human astrovirus antigens to determine seroprevalence, including mature human astrovirus particles, the full capsid protein, or capsid protein fragments (Table 1.3). However, studies that evaluate reactivity to the whole capsid are limited because of the possibility of cross-reactive antibodies that may recognize the conserved core domains of several serotypes. The use of the more variable capsid spike domain as the antigen provides an opportunity to confidently evaluate serotype reactivity across all eight serotypes. Here, we provide an updated report on the seroprevalence of each of the eight classical serotypes in a United States adult population as determined by biolayer interferometry immunosorbent assay (BLI-ISA) using recombinant human astrovirus 1-8 capsid spikes (Spike 1 - Spike 8) as the antigens [1.37].

Table 1.3: Summary of previous human astrovirus serological surveys.

Study Year	Seroprevalence	Size	Antigen
Koopmans et al. [1.10] 1998	Ages 0-79: Human astrovirus 1: 91 % Human astrovirus 2: 31 % Human astrovirus 3: 69 % Human astrovirus 4: 56 % Human astrovirus 5: 36 % Human astrovirus 6: 16 % Human astrovirus 7: 10 %	242 people ages 0-79	Human astrovirus particles cultured in Caco-2 cells
Mitchell et al. [1.31] 1999	By age 9: Human astrovirus 1: 94 % Human astrovirus 3: 42 %	393 children	Baculovirus-expressed recombinant capsid proteins
Kriston et al. [1.32] 1996	By age 5: Human astrovirus 1: 90 % Human astrovirus 6: 10-30 %	273 children and hospital staff	Baculovirus-expressed recombinant capsid proteins
Kobayashi et al. [1.33] 1999	By age 3: Human astrovirus 1: ~100 % Human astrovirus 2: ~100 %	170 children	Unclear: Article in Japanese

Study Year	Seroprevalence	Size	Antigen
Kurtz et al. [1.23] 1978	Presence of "astrovirus antibody" By 6 months: 45 % By age 1: 7 % By age 2: 25 % By age 4: 71 % By age 10: 75 % Young adult cohort: 77 %	87 children ages 0-10 70 young adults ages 17-30	Human astrovirus particles cultured in HEK cells
Holtz et al. [1.34] 2014	By age 17: MLB1: 100 %	395 people ages 0-95	Baculovirus-expressed recombinant capsid protein
Burbelo et al. [1.35] 2011	HMOAstV-C (VA1): 36 % by age 10 65 % in adults	103 children and 106 adults	Crude protein extracts containing N- and C-terminal capsid fragments
Meliopoulos et al. [1.36] 2014	By adulthood: Human astrovirus 1: 81 % Turkey astrovirus 2: 26 %	160 turkey growers, turkey meat processing plant workers, and unexposed workers	Baculovirus-expressed recombinant capsid proteins

1.4 Materials and Methods

1.4.1 Human astrovirus capsid sequences used for sequence identity studies

The accession numbers of sequences used for sequence alignments of the human astrovirus capsid core are as follows for serotypes 1 to 8 (Core- accession number: residues (depository)): Core 1- AAC34717.1: 80-410 (Genbank), Core 2- AZB52195.1: 80-409 (Genbank), Core 3- Q9WFZ0: 82-412 (Uniprot), Core 4- Q3ZN05: 80-410 (Uniprot), Core 5- Q4TWH7: 79-409 (Uniprot), Core 6- AZB52207.1: 80-410 (Genbank), Core 7- Q96818: 81-411 (Uniprot), and Core 8- Q9IFX1: 80-410 (Uniprot). The accession numbers of sequences used for sequence alignments of the human astrovirus capsid spike are as follows for serotypes 1 to 8 (Spike- accession number: residues (depository)): Spike 1- AAC34717.1: 430-644 (Genbank), Spike 2- KY964327.1: 12-228 (Genbank), Spike 3- Q9WFZ0: 432-645 (Uniprot), Spike 4- Q3ZN05: 430-644 (Uniprot), Spike 5- Q4TWH7: 429-641 (Uniprot), Spike 6- AZB52207.1: 430-642 (Genbank), Spike 7- Q96818: 431-644 (Uniprot), and Spike 8- Q9IFX1: 431-644 (Uniprot).

1.4.2 Reagents and supplies

Anti-Penta-His (HIS1K) sensor tips (Sartorius FortéBio 18-5120), Benzonase (Merck Millipore 71205), bovine serum albumin (BSA) (Fisher BP1600), CD OptiCHO expression medium (Gibco 12681-029), CHO CD EfficientFeed A (Gibco A10234-01), ChonBlock (Chondrex 9068), 4 nm Colloidal Gold-AffiniPure Goat Anti-Human IgG Fcγ fragment specific (Jackson ImmunoResearch 109-185-170) rehydrated

in 1 mL deionized water per the manufacturer's instructions, 4 nm Colloidal Gold-AffiniPure Goat Anti-Rabbit IgG (H+L) (Jackson ImmunoResearch 111-185-144) rehydrated in 1 mL deionized water per the manufacturer's instructions, glucose (Sigma Aldrich G8769), 1 mL HiTrap Protein A HP column (GE Healthcare 17-0402-03), HT supplement (Gibco 11067-030), L-glutamine (Gibco 35050-061), phosphate buffered saline (PBS) tablets (Sigma P4417), Pierce IgG Elution Buffer (Thermo Fisher Scientific 21004), Pierce Normal Rabbit Serum Control (Thermo Fisher Scientific 31884), Pierce Protein A IgG Binding Buffer (Thermo Fisher Scientific 21001), Pluronic F68 (Gibco 24040-032), protease inhibitor cocktail set V EDTA-free (Merck Millipore 539137), sodium butyrate (Sigma-Aldrich B5887), Superdex 75 10/300 and 16/600 (Cytiva), TALON metal affinity resin (GE Healthcare 28-9574-99), tilted-bottom (TW384) microplates (Sartorius FortéBio 18-5080), Tween-20 (Fisher BP337), yeastolate (BD 292804).

1.4.3 Expression and purification of human astrovirus spikes 1-8

Synthetic cDNA encoding Spike 1 (residues 429-645, accession AAC34717.1 (Genbank)), Spike 2 (residues 12-228, accession KY964327.1 (Genbank), a partial HAstV2 capsid sequence), Spike 3 (residues 431-647, accession Q9WFZ0 (Uniprot)), Spike 4 (residues 429-646, accession Q3ZN05 (Uniprot)), Spike 5 (residues 428-644, accession Q4TWH7 (Uniprot)), Spike 6 (residues 429-644, accession AZB52207.1 (Genbank)), Spike 7 (residues 430-646, accession Q96818 (Uniprot)), or Spike 8 (residues 429-647, accession Q9IFX1 (Uniprot)) was cloned separately into pET52b

(EMD Millipore) in frame with a C-terminal linker, thrombin cleavage site, and 10-Histidine tag. The plasmids were verified by DNA sequencing. Spike plasmids were transformed into *E. coli* strain BL21 (DE3). Expression was induced with 1 mM IPTG and carried out at 18 °C for 16 hours. Cell pellets were resuspended in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole (Buffer A) containing 2 mM $MgCl_2$, 0.0125 U/μL benzonase, and 1x protease inhibitor cocktail set V EDTA-free. Cells were lysed using ultrasonication, and spikes were batch purified using TALON metal affinity resin. Elutions were carried out using Buffer A with 500 mM imidazole. Spikes were dialyzed into 10 mM Tris-HCl pH 8.0, 150 mM NaCl (TBS) with 1 mM DTT and purified using size exclusion chromatography in TBS with a Superdex 75 16/600 (Spikes 1, 2, 5, 7, and 8) or Superdex 75 10/300 column (Spikes 3, 4, and 6).

1.4.4 Expression and purification of the SARS-CoV-2 RBD

The SARS-CoV-2 receptor-binding domain (RBD) in frame with a 10-Histidine tag and AviTag at the C-terminus was expressed and purified as described in [1.37].

1.4.5 Expression and purification of recombinant mAb 3B4

Mouse hybridoma cells producing mAb 3B4 were generated as reported in [1.38]. The amino acid sequences of the mAb 3B4 variable regions were identified as described in [1.39], allowing for recombinant antibody expression. Synthetic cDNA encoding the 3B4 kappa and heavy chain variable regions (Integrated DNA Technologies) was cloned by Gibson assembly into the two pCMV-VRC01 antibody

backbone vectors for light and for heavy chains, which contain the constant regions of VRC01, a human anti-HIV antibody targeting the gp120 protein [1.40]. Cloned sequences were in frame with an N-terminal secretion signal sequence. The resulting expression plasmids, pCMV-VRC01_3B4_kappa and pCMV-VRC01_3B4_heavy, contain the variable regions from the original mouse antibody 3B4 and the constant regions from a human IgG1 antibody under control of the human cytomegalovirus promoter. The plasmids were verified by DNA sequencing. The expression plasmids were used in a 1:1 ratio to electroporate Chinese Hamster Ovary suspension (CHO-S) cells using the MaxCyte system. Recombinant mAb 3B4 was expressed for 8 days by CHO-S cells growing in CD OptiCHO expression medium supplemented with 1 mM sodium butyrate, 8 mM L-glutamine, 1x HT supplement, and 0.1 % Pluronic F68 at 32 °C with 125 rpm shaking. Every 24 hours, cells were fed with CHO CD EfficientFeed A supplemented with 7 mM L-glutamine, 5.5 % glucose, and 23.4 g/L yeastolate. After 8 days, cells were pelleted and medium containing secreted mAb 3B4 was diluted 1:1 with Pierce Protein A IgG Binding Buffer and 0.22-μM filtered. The sample was loaded onto a pre-equilibrated 1 mL HiTrap Protein A HP column. The column was washed with Pierce Protein A IgG Binding Buffer and elution was accomplished with low pH Pierce IgG Elution Buffer. Acidic elutions were neutralized by 2 M Tris pH 8.0 to 10 % of the final volume.

1.4.6 Generation of rabbit polyclonal serum to human astrovirus 1

Rabbit polyclonal serum to human astrovirus 1 was generated as described in [1.41].

1.4.7 Human plasma samples

Human plasma samples were obtained from Discovery Life Sciences. These 63 de-identified samples were collected from adults ages 19-78 (31 male, 32 female) between 2012 – 2016 (before the emergence of SARS-CoV-2). Plasma samples were heated at 56 °C for 1 hour before use.

1.4.8 Biolayer interferometry immunosorbent assay

BLI-ISA studies were performed on an Octet RED384 instrument at 22 °C with shaking at 1000 rpm. BLI-ISA assay buffer consists of PBS pH 7.4, 2 % BSA, and 0.1 % Tween-20, which was 0.22-μm filtered. Before use, Anti-Penta-His (HIS1K) biosensors were placed into the wells of a biosensor-holding plate and pre-hydrated in BLI-ISA buffer for at least 10 minutes. Tilted-bottom 384-well microplates were loaded with 48 μL per well. The assay plate was prepared as follows: column 1 (BLI-ISA buffer), column 2 (10 nM His-tagged human astrovirus spikes 1-8 or 70 nM His-tagged SARS-CoV-2 RBD in BLI-ISA buffer), column 3 (a 1:4 dilution of ChonBlock in BLI-ISA buffer), column 4 (a 1:8 dilution of plasma samples: 6 μL plasma + 42 μL ChonBlock/BLI-ISA buffer), column 5 (BLI-ISA buffer), and column 6 (4 nm Colloidal Gold-AffiniPure Goat Anti-Human IgG or 4 nm Colloidal Gold-AffiniPure Goat Anti-Rabbit IgG secondary antibody diluted 1:10 in BLI-ISA buffer).

The BLI-ISA method was set as follows, with minimal adaptation from [1.37]. Baseline 1 (60 s) in column 1 (Equilibration), Loading (180 s) in column 2 (Antigen Loading: human astrovirus spikes 1-8 or SARS-CoV-2 RBD), Baseline 2 (60 s) in column 3 (Wash), Association 1 (600 s) in column 4 (Total Antibody Binding), Baseline 3 (60 s) in column 5 (Wash), and Association 2 (180 s) in column 6 (Detection: anti-human IgG-gold or anti-rabbit IgG-gold). Loading of 10 nM spikes or 70 nM SARS-CoV-2 RBD over 180 s onto Anti-Penta-His sensor tips resulted in a wavelength shift loading signal of ~0.55 nm.

Analysis of BLI-ISA data was performed and automated as previously described in [1.37].

1.5 Results

To begin, we expressed recombinant Spike 1 – Spike 8 antigens in *E. coli* and purified them by affinity chromatography and size exclusion chromatography. Purity was verified by SDS-PAGE (Figure 1.2A). Comparison of elution volumes for spike antigens to elution volumes for gel filtration standards confirmed that the spikes form dimers in solution (Figure 1.2B and 1.2C), consistent with structural studies [1.27, 1.28, 1.30].

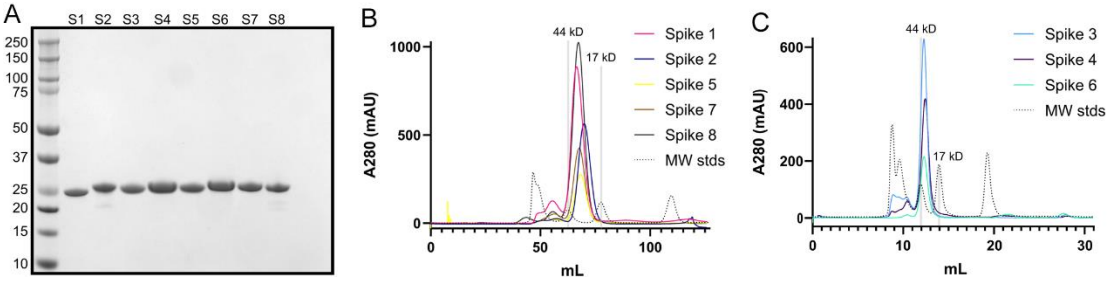

Figure 1.2: Purification of recombinant human astrovirus 1-8 capsid spikes (Spike 1 – Spike 8).

(A) Reducing SDS-PAGE of purified recombinant Spikes 1-8 (S1-S8). **(B)** Size exclusion chromatography column (Superdex 75 16/600) traces of Spikes 1, 2, 5, 7, and 8. **(C)** Size-exclusion chromatography column (Superdex 75 10/300) traces of Spikes 3, 4, and 6. Size-exclusion chromatography confirmed that the purified spikes are folded and form dimers in solution.

We then used the recombinant spike antigens in a biolayer interferometry immunosorbent assay (BLI-ISA). BLI-ISA is a recently developed method for rapid and semi-quantitative measurement of antibodies in blood plasma directed towards a particular antigen [1.37]. In a simple dip-and-read format, the binding of biomolecules to fiber-optic biosensors is measured in real time through a shift in the wavelength of the reflected light upon binding, with results available in 20 minutes. After an initial equilibration, the biosensors are loaded with the antigen of interest. Next, the antigen-loaded biosensors are washed and placed into diluted plasma to allow binding of antibodies to the antigen. After another wash, the antigen/antibody-coated biosensors are dipped into wells containing isotype-specific binding reagents, such as colloidal gold-conjugated anti-human IgG, and a wavelength shift detection signal is measured (Figure 1.3A). BLI-ISA was originally established as a tool for serosurveillance of

SARS-CoV-2 [1.37]. During its development, known SARS-CoV-2 seropositive and seronegative plasma samples were first validated by a standard dilution-series enzyme-linked immunosorbent assay (ELISA) starting at a 1:50 dilution. Importantly, trends observed in subsequent 1:8 single-dilution BLI-ISA measurements of those samples correlated with area-under-the-curve calculations from the dilution-series ELISA, indicating that BLI-ISA can identify the variation in antibody levels in seropositive samples at this plasma dilution.

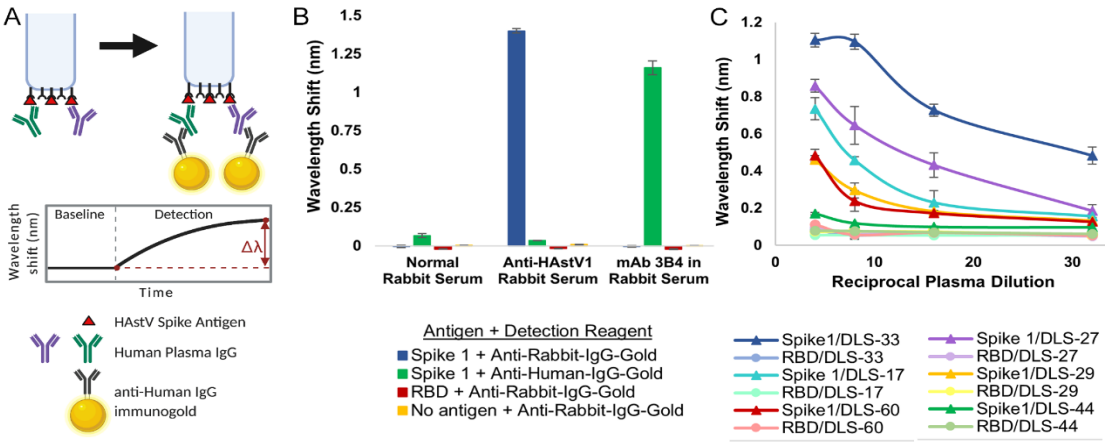

Figure 1.3: Validation of BLI-ISA for the detection of plasma IgG antibodies to human astrovirus capsid spikes.

(A) BLI-ISA schematic of the IgG antibody detection step. This image was created with BioRender.com. **(B)** BLI-ISA detection of control sera. Spike 1, SARS-CoV-2 RBD, or no antigen were loaded onto biosensors and placed into 1:100 normal rabbit serum, 1:100 anti-human astrovirus 1 rabbit serum, or 25 nM mAb 3B4 in 1:100 normal rabbit serum. Bound antibodies were detected with anti-human-IgG-gold or anti-rabbit-IgG-gold. **(C)** Dilution-series BLI-ISA using Spike 1 (triangles) and RBD (circles) as antigens and representative strong astrovirus 1 seropositive (DLS-33, blue), moderate astrovirus 1 seropositive (DLS-17, cyan, and DLS-27, purple), weak astrovirus 1 seropositive (DLS-29, yellow, and DLS-60, red), and astrovirus 1 seronegative (DLS-44, green) plasma samples.

To confirm whether BLI-ISA can be applied to human astrovirus serosurveillance, we assessed Spike 1 antigen binding with commercially available normal rabbit serum and rabbit serum positive for human astrovirus serotype 1 [1.41]. As a positive control, monoclonal antibody (mAb) 3B4, a recombinant mouse-human chimeric antibody containing mouse variable regions that target Spike 1, was added to normal rabbit serum (Figure 1.3B). We also tested the specificity of our detection reagents, colloidal gold-conjugated anti-human IgG and colloidal gold-conjugated anti-rabbit IgG, in a 1:10 dilution. We loaded either Spike 1, the SARS-CoV-2 receptor-binding domain (RBD) negative control antigen, or no antigen onto the biosensor tips and dipped them into the indicated sera. Background signals were observed in all samples with a 1:100 dilution of normal rabbit serum, indicating a lack of anti-Spike 1 and anti-RBD antibodies in the serum, as well as low non-specific binding of the serum to the empty biosensor tip. When Spike 1 was loaded onto the biosensor and dipped into a 1:100 dilution of anti-human astrovirus 1 rabbit serum, a robust signal emerged when detected with the anti-rabbit-IgG-gold reagent (Figure 1.3B, blue) but not the anti-human-IgG-gold reagent, indicating the high specificity of the anti-rabbit-IgG-gold reagent. Background signal was observed for RBD- and no antigen-loaded biosensors. Finally, we added 25 nM mAb 3B4 to a 1:100 dilution of normal rabbit serum into which we dipped Spike 1-loaded biosensors. We noted a strong signal when the sample was detected with the anti-human-IgG-gold reagent (Figure 1.3B, green) but not the anti-rabbit-IgG-gold reagent, showing the precision of the anti-human-IgG-

gold reagent. Background signal was exhibited for RBD- and no antigen-loaded biosensors also in this case.

Next, we conducted dilution-series experiments to assess the optimal human plasma dilution to use in BLI-ISA serological assays for human astrovirus. Dzimianski et al. established a 1:8 plasma dilution as optimal after recent SARS-CoV-2 exposure, providing an estimate of the dilution range for serosurveillance studies using other antigens [1.37]. Sixty-three human plasma samples were collected between 2012 - 2016 (pre-SARS-CoV-2) from a randomized cross-sectional sample of the population in the United States, adults ages 19-78. We performed dilution-series BLI-ISA on a representative subset of these plasma samples using Spike 1 as the antigen (Figure 1.3C). As a control for non-specific binding, the SARS-CoV-2 RBD was loaded onto the biosensors instead of Spike 1. The selected plasma samples include a strongly astrovirus 1-seropositive sample (DLS-33), two moderately astrovirus 1-seropositive samples (DLS-17 and DLS-27), two weakly astrovirus 1-seropositive samples (DLS-29 and DLS-60), and an astrovirus 1-seronegative sample (DLS-44) to represent the diversity of signals in a single-dilution assay. Each point was performed in duplicate. We loaded either Spike 1 or RBD onto the biosensors and then dipped into 1:4, 1:8, 1:16, or 1:32 dilutions of the indicated plasma. Detection signal at the 1:8 dilution compared to the 1:16 and 1:32 dilutions showed dose-dependent improvement for weak seropositives DLS-29 and DLS-60 as well as moderate seropositives DLS-17 and DLS-27 while maintaining low signal for the DLS-44 seronegative as well as the RBD-loaded background samples. While even higher signal was observed for these weakly

and moderately seropositive samples at the 1:4 dilution compared to the 1:8 dilution, strong seropositive DLS-33 appears to reach signal saturation at the 1:8 dilution, with little signal improvement at the 1:4 dilution. In addition, seronegative sample DLS-44 and several of the RBD-loaded control samples showed some non-specific signal at the 1:4 dilution but remained at background levels with the 1:8, 1:16, and 1:32 dilutions. Thus, we identified the 1:8 plasma dilution as optimal to maximize the dynamic range of the assay.

After validation of the control samples and confirmation of the optimal 1:8 plasma dilution, we used BLI-ISA to determine the seroprevalence of antibodies in the 63 human plasma samples to HAstV spikes from serotypes 1 to 8 (Figure 1.4 and Table 1.4). This investigation using BLI-ISA to assess the unknown serostatus of individuals is the first of its kind. Signals for each plasma sample to SARS-CoV-2 RBD were measured to establish each sample's background detection value. To reduce inter-sample variability and prevent false positives from high-background samples, each sample's spike reactivity signal was subtracted by its own RBD reactivity signal to generate a background-corrected detection value. The variability in plasma reactivity to RBD was used to estimate a seropositivity cut-off of four times the standard deviation of background binding to RBD. Overall, the percentage of samples containing IgG antibodies targeting the spike protein was highest for human astrovirus serotype 1 (46/63, 73 %), followed by serotype 3 (39/63, 62 %), serotype 4 (33/63, 52 %), serotype 5 (18/63, 29 %), serotype 8 (17/63, 27 %), serotype 2 (14/63, 22 %), serotype 6 (5/63, 8 %), and serotype 7 (5/63, 8 %), which is in accordance with a

comprehensive human astrovirus serosurveillance study conducted by Koopmans et al. in the Netherlands 23 years ago (Table 1.3) [1.10].

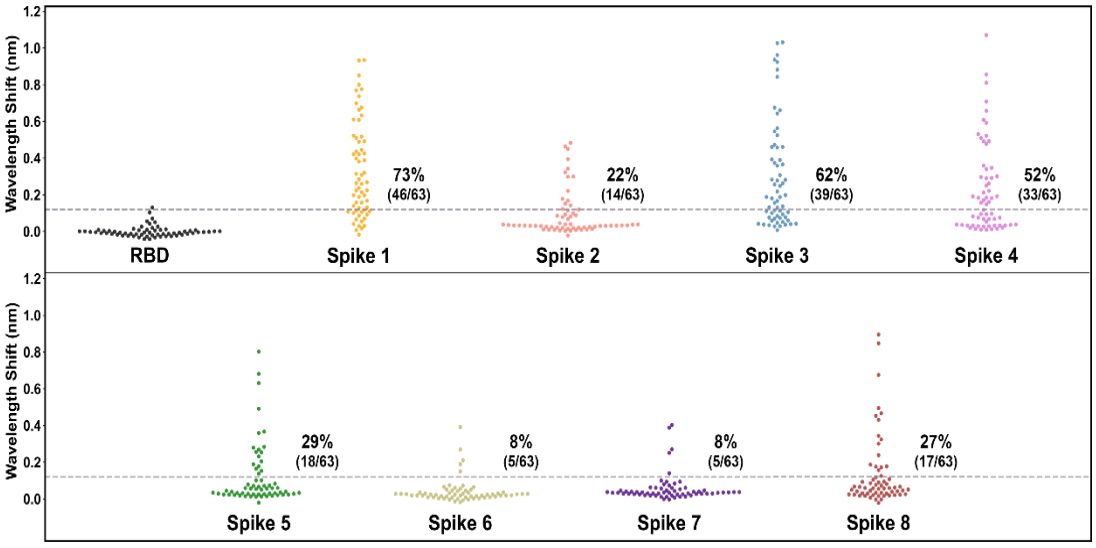

Figure 1.4: Human plasma IgG reactivity to human astrovirus spike proteins.

Data are displayed as swarm plots, with each dot representing the signal for one individual for the indicated antigen. For samples with the RBD antigen, each sample's RBD reactivity signal was subtracted by the mean RBD reactivity signal (0.07 nm) to center samples around zero. The dashed line indicates four standard deviations above zero for the RBD samples (0.12 nm), and the printed percentages denote the number of samples above the line. For samples with the spike antigens, each sample's spike reactivity signal was subtracted by its own RBD reactivity signal to generate a background-corrected signal.

Table 1.4: Heat map of individual human plasma IgG reactivity to human astrovirus spike antigens.[1]

		Controls		Human Astrovirus Spike Antigens (Background-Corrected)							
	ID	No Antigen	RBD	Spike 1	Spike 2	Spike 3	Spike 4	Spike 5	Spike 6	Spike 7	Spike 8
1	DLS-18	0.13	0.06	0.11	0.04	1.03	0.03	0.02	0.00	0.05	0.03
2	DLS-14	0.10	0.05	0.20	0.03	0.88	0.15	0.03	0.03	0.09	0.07
3	DLS-55	0.07	0.06	0.32	0.15	0.92	0.04	0.04	0.06	0.10	0.18
4	DLS-44	0.05	0.04	0.02	0.03	0.46	0.15	0.02	0.05	0.04	0.17
5	DLS-36	0.14	0.07	0.01	0.09	0.54	0.03	0.01	0.01	0.03	0.06
6	DLS-51	0.08	0.07	0.16	0.02	0.66	0.01	0.02	0.01	0.04	0.02
7	DLS-06	0.11	0.06	0.11	0.04	0.67	0.53	0.06	0.03	0.05	0.09
8	DLS-01	0.14	0.08	0.03	0.11	0.96	0.59	0.03	0.04	0.06	0.32
9	DLS-30	0.13	0.09	0.43	0.02	0.31	0.09	0.26	0.03	0.07	0.07
10	DLS-02	0.20	0.10	0.52	0.03	0.47	0.09	0.27	0.05	0.05	0.10
11	DLS-41	0.10	0.07	0.68	0.03	0.52	0.22	0.28	0.19	0.08	0.08
12	DLS-20	0.09	0.05	0.44	0.03	0.46	0.03	0.03	0.02	0.04	0.03
13	DLS-05	0.37	0.20	0.42	-0.02	0.56	0.07	0.16	0.00	0.01	-0.02
14	DLS-54	0.07	0.07	0.31	0.01	0.37	0.03	0.03	0.03	0.09	0.02
15	DLS-21	0.22	0.17	0.27	0.02	0.46	0.03	-0.02	0.00	0.04	0.12
16	DLS-49	0.09	0.06	0.78	0.02	0.84	0.02	0.06	0.02	0.25	0.06
17	DLS-35	0.09	0.07	0.63	0.14	0.64	0.19	0.01	0.01	0.02	0.01
18	DLS-24	0.12	0.06	0.11	0.03	1.03	0.07	0.19	0.06	0.27	0.89
19	DLS-19	0.08	0.08	0.12	0.00	0.18	0.01	0.14	0.00	0.39	0.47
20	DLS-28	0.06	0.03	0.19	0.03	0.28	0.04	0.02	0.01	0.02	0.67
21	DLS-42	0.07	0.06	0.26	0.03	0.13	0.01	0.02	0.01	0.04	0.43
22	DLS-07	0.09	0.05	0.26	0.01	0.08	0.10	0.68	0.05	0.03	0.05
23	DLS-25	0.10	0.06	0.85	0.46	0.05	0.12	0.63	0.03	0.03	0.09
24	DLS-27	0.11	0.08	0.77	0.01	0.06	0.08	0.02	-0.01	0.01	0.34
25	DLS-34	0.10	0.07	0.70	0.01	0.14	0.03	0.03	0.00	0.02	0.45
26	DLS-23	0.09	0.05	0.93	0.09	0.28	0.07	0.36	0.03	0.04	0.04
27	DLS-33	0.10	0.10	0.93	0.03	0.08	0.01	0.03	-0.02	0.00	0.00
28	DLS-09	0.29	0.12	0.80	0.07	0.10	0.17	0.05	0.07	0.02	0.02
29	DLS-53	0.08	0.06	0.74	0.03	0.04	0.29	0.10	0.06	0.04	0.08
30	DLS-61	0.08	0.07	0.49	0.01	0.03	0.02	0.06	0.01	0.03	0.01
31	DLS-17	0.09	0.06	0.61	0.03	0.19	0.03	0.02	0.01	0.04	0.03
32	DLS-04	0.09	0.04	0.49	0.03	0.25	0.05	0.07	0.02	0.06	0.06
33	DLS-46	0.08	0.06	0.39	0.02	0.17	0.03	0.03	0.02	0.03	0.02
34	DLS-12	0.06	0.03	0.38	0.03	0.11	0.03	0.25	0.03	0.03	0.05
35	DLS-31	0.16	0.14	0.40	0.01	0.07	0.18	0.08	-0.01	0.01	0.00

		Controls		Human Astrovirus Spike Antigens (Background-Corrected)							
	ID	No Antigen	RBD	Spike 1	Spike 2	Spike 3	Spike 4	Spike 5	Spike 6	Spike 7	Spike 8
36	DLS-40	0.09	0.05	0.31	0.02	0.06	0.21	0.07	0.03	0.03	0.03
37	DLS-26	0.12	0.08	0.61	0.04	0.20	0.26	0.06	0.00	0.02	0.17
38	DLS-45	0.10	0.07	0.51	0.12	0.28	0.25	0.07	0.02	0.06	0.30
39	DLS-39	0.13	0.07	0.45	0.17	0.10	0.36	0.02	0.02	0.06	0.06
40	DLS-47	0.09	0.05	0.66	0.01	0.04	0.52	0.03	0.02	0.03	0.02
41	DLS-59	0.08	0.07	0.21	0.02	0.09	0.49	0.03	0.03	0.40	0.02
42	DLS-32	0.11	0.07	0.07	0.00	0.13	0.49	0.01	0.02	0.01	0.07
43	DLS-37	0.11	0.07	0.09	0.03	0.36	0.35	0.05	0.02	0.04	0.24
44	DLS-43	0.10	0.06	0.09	0.03	0.36	0.30	0.03	0.02	0.03	0.04
45	DLS-15	0.08	0.04	0.22	0.08	0.25	0.34	0.23	0.21	0.14	0.16
46	DLS-58	0.08	0.06	0.28	0.22	0.39	0.48	0.04	0.03	0.04	0.12
47	DLS-11	0.08	0.04	0.07	0.45	0.03	0.19	0.18	0.02	0.03	0.06
48	DLS-62	0.13	0.12	0.10	0.32	0.26	0.07	0.03	0.05	0.08	0.08
49	DLS-57	0.07	0.06	0.04	0.04	0.04	0.16	0.04	0.02	0.03	0.19
50	DLS-13	0.09	0.04	0.03	0.12	0.19	0.07	0.15	0.01	0.04	0.05
51	DLS-48	0.09	0.06	0.06	0.09	0.22	0.18	0.07	0.02	0.03	0.02
52	DLS-56	0.07	0.08	0.14	0.34	0.93	0.71	0.80	0.27	0.01	0.11
53	DLS-63	0.07	0.08	0.11	0.48	0.39	0.51	0.49	0.15	0.03	0.04
54	DLS-08	0.08	0.05	0.52	0.30	0.11	1.07	0.28	0.39	0.03	0.85
55	DLS-52	0.22	0.11	-0.02	0.39	0.01	0.81	0.01	0.00	0.00	0.49
56	DLS-50	0.10	0.07	0.42	0.18	0.12	0.66	0.37	0.05	0.04	0.09
57	DLS-03	0.08	0.04	0.17	0.04	0.04	0.85	0.20	0.07	0.03	0.05
58	DLS-22	0.11	0.06	0.16	0.30	0.07	0.61	0.08	0.03	0.06	0.04
59	DLS-29	0.08	0.05	0.20	0.04	0.04	0.18	0.06	0.02	0.03	0.02
60	DLS-10	0.10	0.05	0.22	0.10	0.05	0.30	0.10	0.04	0.03	0.06
61	DLS-16	0.09	0.04	0.13	0.08	0.08	0.29	0.03	0.02	0.03	0.03
62	DLS-38	0.16	0.08	0.24	0.01	0.15	0.03	0.01	0.01	0.02	0.02
63	DLS-60	0.05	0.04	0.13	0.01	0.03	0.01	0.04	0.03	0.02	0.02
64	PBS	0.06	0.05	0.01	0.02	0.02	0.02	0.02	0.03	0.02	0.02

[1]For samples with no antigen and the RBD antigen, the true reactivity signals are reported. For the spike samples, each sample's spike reactivity signal was subtracted by its own RBD reactivity signal to generate a background-corrected signal. To aid the visualization of patterns in human astrovirus reactivity, spike reactivity values were hierarchically clustered using a SciPy python library function wherein distance was calculated with the Ward variance minimization algorithm (scipy.cluster.hierarchy.linkage(spike-data, method = 'ward')). After initial clustering, minor additional sorting was performed and reactivity values were colored according to their magnitude to produce the heatmap.

1.6 Discussion

Here, we evaluate the seroprevalence of plasma antibodies towards human astrovirus serotypes 1-8 in a cross-sectional sample of United States adults. We adapted BLI-ISA, a recently developed method for the rapid and semi-quantitative measurement of plasma antibodies, to measure IgG antibodies against recombinant human astrovirus 1-8 capsid spikes (Spike 1 – Spike 8). We chose Spike 1 – Spike 8 as antigens for their ease of production in *E. coli* and for their variable sequence identity (41.4 – 75.7 %) that should minimize antibody cross-reactivity between spikes from different serotypes. Indeed, most known neutralizing monoclonal antibodies that target the spike are serotype specific [1.38, 1.41, 1.42], suggesting that total antibody levels against individual spikes should show minimal cross-reactivity. Plasma dilution series confirm that BLI-ISA is semi-quantitative, and a plasma dilution of 1:8 was selected as the ideal dilution to maximize the dynamic range of the BLI-ISA signals and minimize background signal using biosensors coated with a negative control antigen, the SARS-CoV-2 RBD, which should be non-reactive for antibodies in plasma collected before 2019 like the ones in this study. Of note, a plasma dilution of 1:8 was also chosen to measure levels of antibodies to the SARS-CoV-2 RBD after recent SARS-CoV-2 exposure during BLI-ISA method development [1.37]. Time since exposure to human astrovirus is unknown for the samples in our study, yet the optimal plasma dilution remained 1:8. Therefore, BLI-ISA is a sensitive method that can detect antibodies elicited by an immune response which occurred potentially decades before plasma sample collection, while requiring very little plasma (6 μL total per sample per antigen).

This study estimates the seroprevalence rates, in order of prevalence, as 73 % for HAstV1, 62 % for HAstV3, 52 % for HAstV4, 29 % for HAstV5, 27 % for HAstV8, 22 % for HAstV2, 8 % for HAstV6, and 8 % for HAstV7. These findings are consistent with previous seroprevalence studies as well as genotypic surveillance studies identifying human astrovirus 1 as the most prevalent serotype/genotype [1.9-1.12]. This result is also consistent with a previous seroprevalence study in identifying human astrovirus serotypes 3, 4, and 5 as having the next highest prevalence after serotype 1 (Table 1.3) [1.10]. We extend previous work by evaluating seroprevalence for human astrovirus serotype 8 for the first time and determine that it is relatively high at 27 %, an unexpected observation given that it was discovered after serotypes 6 and 7, which have relatively low seroprevalence rates of 8 % each. We note that seroprevalence rates in this study are likely modest estimates based upon a signal threshold of four times the standard deviation of the samples with a negative control antigen. Indeed, many individual samples had signals for spikes that were two or three times their own background signal yet did not meet the conservative criteria for seropositivity in this study.

Critically, seroprevalence rates for recombinant human astrovirus capsid spike antigens correlate closely with neutralizing antibody rates determined previously (Table 1.3) [1.10]. This finding is significant because neutralizing antibodies are often a marker of protection from viral infection. While correlates of protection have yet to be defined for human astrovirus infection, there are some studies suggesting that antibodies are protective [1.23-1.26]. The observation that antibodies to the human

astrovirus capsid spike correlate with virus neutralization is not entirely surprising given previous evidence. Specifically, studies of HAstV-neutralizing monoclonal antibodies have identified that their epitopes reside on the spike domain or the capsid fragments that contains the spike domain, VP24 and VP26 [1.38, 1.41-1.43]. Moreover, while both the capsid core and capsid spike domains stimulate antibody production in mice and rabbits, only antibodies to the spike show neutralizing activity [1.28, 1.38, 1.41]. Thus, plasma IgG levels against human astrovirus capsid spikes may provide a useful estimate of neutralizing antibody levels.

Finally, we present Table 1.4 as a heatmap of antibody levels of each individual for human astrovirus 1 - 8 capsid spikes. Individuals were clustered using the Ward distance minimization algorithm to identify trends in seropositivity. We find that every individual has reactivity to at least one serotype of spike and, on average, individuals are seropositive for about 3 human astrovirus serotypes. While no sample had reactivity to all 8 spikes, one sample had reactivity to 7 spikes (DLS-15). A few samples were only weakly to moderately seropositive to only one or two spikes (DLS-13, -57, -60). We examined the data for correlations in sample positivity between any two serotypes, indicating potential antibody cross-reactivity, but did not identify any trends. Interestingly, clustering revealed a trend in individuals with high seroreactivity for one spike from either HAstV1, -3, or -4 (which have the highest seroprevalence) in that these individuals rarely have high seroreactivity to another spike from these prevalent serotypes. This result might indicate that a strong antibody response to a first infection provides partial protection to an infection by a different serotype, resulting in a weaker

antibody response to the different serotype. However, further research is necessary to explore this possibility.

1.7 Keywords

Human astrovirus, seroprevalence, biolayer interferometry immunosorbent assay

1.8 References

1.1. Mendez, E.; Cf, A. Astroviruses. In *Fields Virology*, 6th edition ed.; Knipe, D.M.; Howley, P.M.; Cohen, J.I.; Griffin, D.E.; Lamb, R.A.; Martin, M.A.; Racaniello, V.R.; B, R., Eds. Lippincott Williams & Wilkins: Philadelphia, PA, 2013; Volume 1, pp 609-628.

1.2. Koci, M.D.; Schultz-Cherry, S. Avian astroviruses. *Avian Pathology : Journal of the W.V.P.A* 2002, 31, 213-227.

1.3. Woode, G.N.; Pohlenz, J.F.; Gourley, N.E.; Fagerland, J.A. Astrovirus and breda virus infections of dome cell epithelium of bovine ileum. *J Clin Microbiol* 1984, 19, 623-630.

1.4. Cortez, V.; Sharp, B.; Yao, J.; Livingston, B.; Vogel, P.; Schultz-Cherry, S. Characterizing a murine model for astrovirus using viral isolates from persistently infected immunocompromised mice. *J Virol* 2019, 93.

1.5. Brown, J.R.; Morfopoulou, S.; Hubb, J.; Emmett, W.A.; Ip, W.; Shah, D.; Brooks, T.; Paine, S.M.; Anderson, G.; Virasami, A., et al. Astrovirus VA1/HMO-C: An increasingly recognized neurotropic pathogen in immunocompromised patients. *Clin Infect Dis* 2015, 60, 881-888.

1.6. Naccache, S.N.; Peggs, K.S.; Mattes, F.M.; Phadke, R.; Garson, J.A.; Grant, P.; Samayoa, E.; Federman, S.; Miller, S.; Lunn, M.P., et al. Diagnosis of neuroinvasive astrovirus infection in an immunocompromised adult with encephalitis by unbiased next-generation sequencing. *Clin Infect Dis* 2015, 60, 919-923.

1.7. Quan, P.L.; Wagner, T.A.; Briese, T.; Torgerson, T.R.; Hornig, M.; Tashmukhamedova, A.; Firth, C.; Palacios, G.; Baisre-De-Leon, A.; Paddock, C.D., et al. Astrovirus encephalitis in boy with X-linked agammaglobulinemia. *Emerg Infect Dis* 2010, 16, 918-925.

1.8. Cordey, S.; Vu, D.L.; Schibler, M.; L'Huillier, A.G.; Brito, F.; Docquier, M.; Posfay-Barbe, K.M.; Petty, T.J.; Turin, L.; Zdobnov, E.M., et al. Astrovirus MLB2, a new gastroenteric virus associated with meningitis and disseminated infection. *Emerg Infect Dis* 2016, 22, 846-853.

1.9. King, A.M.Q.; Lefkowitz, E.; Adams, M.J.; Carstens, E.B. Virus taxonomy: Classification and nomenclature of viruses. In *Ninth Report of the International Committee on Taxonomy of Viruses*., 9th edition ed.; Elsevier Inc.: Philadelphia, PA, 2011.

1.10. Koopmans, M.P.; Bijen, M.H.; Monroe, S.S.; Vinje, J. Age-stratified seroprevalence of neutralizing antibodies to astrovirus types 1 to 7 in humans in the Netherlands. *Clinical and Diagnostic Laboratory Immunology* 1998, 5, 33-37.

1.11. Kurtz, J.B.; Lee, T.W. Human astrovirus serotypes. *Lancet* 1984, 2, 1405.

1.12. Cortez, V.; Meliopoulos, V.A.; Karlsson, E.A.; Hargest, V.; Johnson, C.; Schultz-Cherry, S. Astrovirus biology and pathogenesis. *Annu Rev Virol* 2017, 4, 327-348.

1.13. Dennehy, P.H.; Nelson, S.M.; Spangenberger, S.; Noel, J.S.; Monroe, S.S.; Glass, R.I. A prospective case-control study of the role of astrovirus in acute diarrhea among hospitalized young children. *J Infect Dis* 2001, 184, 10-15.

1.14. Gallimore, C.I.; Taylor, C.; Gennery, A.R.; Cant, A.J.; Galloway, A.; Xerry, J.; Adigwe, J.; Gray, J.J. Contamination of the hospital environment with gastroenteric viruses: Comparison of two pediatric wards over a winter season. *J Clin Microbiol* 2008, 46, 3112-3115.

1.15. Goodgame, R.W. Viral causes of diarrhea. *Gastroenterology Clinics of North America* 2001, 30, 779-795.

1.16. Gray, J.J.; Wreghitt, T.G.; Cubitt, W.D.; Elliot, P.R. An outbreak of gastroenteritis in a home for the elderly associated with astrovirus type 1 and human calicivirus. *J Med Virol* 1987, 23, 377-381.

1.17. Lewis, D.C.; Lightfoot, N.F.; Cubitt, W.D.; Wilson, S.A. Outbreaks of astrovirus type 1 and rotavirus gastroenteritis in a geriatric in-patient population. *The Journal of Hospital Infection* 1989, 14, 9-14.

1.18. Marshall, J.A.; Bruggink, L.D.; Sturge, K.; Subasinghe, N.; Tan, A.; Hogg, G.G. Molecular features of astrovirus associated with a gastroenteritis outbreak in an aged-care centre. *European Journal of Clinical Microbiology & Infectious Diseases* : Official Publication of the European Society of Clinical Microbiology 2007, 26, 67-71.

1.19. Olortegui, M.P.; Rouhani, S.; Yori, P.P.; Salas, M.S.; Trigoso, D.R.; Mondal, D.; Bodhidatta, L.; Platts-Mills, J.; Samie, A.; Kabir, F., et al. Astrovirus infection and diarrhea in 8 countries. *Pediatrics* 2017, 141.

1.20. Walter, J.E.; Mitchell, D.K. Astrovirus infection in children. *Curr Opin Infect Dis* 2003, 16, 247-253.

1.21. Bosch, A.; Pinto, R.M.; Guix, S. Human astroviruses. *Clinical Microbiology Reviews* 2014, 27, 1048-1074.

1.22. Mead, P.S.; Slutsker, L.; Dietz, V.; McCaig, L.F.; Bresee, J.S.; Shapiro, C.; Griffin, P.M.; Tauxe, R.V. Food-related illness and death in the United States. *Emerg Infect Dis* 1999, 5, 607-625.

1.23. Kurtz, J.; Lee, T. Astrovirus gastroenteritis age distribution of antibody. *Med Microbiol Immunol* 1978, 166, 227-230.

1.24. Kurtz, J.B.; Lee, T.W.; Craig, J.W.; Reed, S.E. Astrovirus infection in volunteers. *J Med Virol* 1979, 3, 221-230.

1.25. Mitchell, D.K. Astrovirus gastroenteritis. *The Pediatric Infectious Disease Journal* 2002, 21, 1067-1069.

1.26. Bjorkholm, M.; Celsing, F.; Runarsson, G.; Waldenstrom, J. Successful intravenous immunoglobulin therapy for severe and persistent astrovirus gastroenteritis after fludarabine treatment in a patient with Waldenstrom's macroglobulinemia. *International Journal of Hematology* 1995, 62, 117-120.

1.27. Dong, J.; Dong, L.; Mendez, E.; Tao, Y. Crystal structure of the human astrovirus capsid spike. *Proceedings of the National Academy of Sciences of the United States of America* 2011, 108, 12681-12686.

1.28. York, R.L.; Yousefi, P.A.; Bogdanoff, W.; Haile, S.; Tripathi, S.; DuBois, R.M. Structural, mechanistic, and antigenic characterization of the human astrovirus capsid. *J Virol* 2015, 90, 2254-2263.

1.29. Toh, Y.; Harper, J.; Dryden, K.A.; Yeager, M.; Arias, C.F.; Mendez, E.; Tao, Y.J. Crystal structure of the human astrovirus capsid protein. *J Virol* 2016, 90, 9008-9017.

1.30. Bogdanoff, W.A.; Perez, E.I.; Lopez, T.; Arias, C.F.; DuBois, R.M. Structural basis for escape of human astrovirus from antibody neutralization: Broad implications for rational vaccine design. *J Virol* 2018, 92.

1.31. Mitchell, D.K.; Matson, D.O.; Cubitt, W.D.; Jackson, L.J.; Willcocks, M.M.; Pickering, L.K.; Carter, M.J. Prevalence of antibodies to astrovirus types 1 and 3 in children and adolescents in Norfolk, Virginia. *The Pediatric Infectious Disease Journal* 1999, 18, 249-254.

1.32. Kriston, S.; Willcocks, M.M.; Carter, M.J.; Cubitt, W.D. Seroprevalence of astrovirus types 1 and 6 in London, determined using recombinant virus antigen. *Epidemiol Infect* 1996, 117, 159-164.

1.33. Kobayashi, S.; Kobayashi, M.; Araki, K.; Shinozaki, T.; Yanagawa, Y. [Antibody prevalence against astrovirus according to age groups]. *Kansenshogaku Zasshi* 1999, 73, 578-583.

1.34. Holtz, L.R.; Bauer, I.K.; Jiang, H.; Belshe, R.; Freiden, P.; Schultz-Cherry, S.L.; Wang, D. Seroepidemiology of astrovirus MLB1. *Clin Vaccine Immunol* 2014, 21, 908-911.

1.35. Burbelo, P.D.; Ching, K.H.; Esper, F.; Iadarola, M.J.; Delwart, E.; Lipkin, W.I.; Kapoor, A. Serological studies confirm the novel astrovirus HMOAstV-C as a highly prevalent human infectious agent. *PLoS One* 2011, 6, e22576.

1.36. Meliopoulos, V.A.; Kayali, G.; Burnham, A.; Oshansky, C.M.; Thomas, P.G.; Gray, G.C.; Beck, M.A.; Schultz-Cherry, S. Detection of antibodies against turkey astrovirus in humans. *PLoS One* 2014, 9, e96934.

1.37. Dzimianski, J.V.; Lorig-Roach, N.; O'Rourke, S.M.; Alexander, D.L.; Kimmey, J.M.; DuBois, R.M. Rapid and sensitive detection of SARS-CoV-2 antibodies by biolayer interferometry. *Sci Rep* 2020, 10, 21738.

1.38. Espinosa, R.; López, T.; Bogdanoff, W.A.; Espinoza, M.A.; López, S.; DuBois, R.M.; Arias, C.F. Isolation of neutralizing monoclonal antibodies to human astrovirus and characterization of virus variants that escape neutralization. *J Virol* 2018.

1.39. Meyer, L.; López, T.; Espinosa, R.; Arias, C.F.; Vollmers, C.; DuBois, R.M. A simplified workflow for monoclonal antibody sequencing. *PLoS One* 2019, 14, e0218717.

1.40. Wu, X.; Yang, Z.Y.; Li, Y.; Hogerkorp, C.M.; Schief, W.R.; Seaman, M.S.; Zhou, T.; Schmidt, S.D.; Wu, L.; Xu, L., et al. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 2010, 329, 856-861.

1.41. Bass, D.M.; Upadhyayula, U. Characterization of human serotype 1 astrovirus-neutralizing epitopes. *J Virol* 1997, 71, 8666-8671.

1.42. Sanchez-Fauquier, A.; Carrascosa, A.L.; Carrascosa, J.L.; Otero, A.; Glass, R.I.; Lopez, J.A.; San Martin, C.; Melero, J.A. Characterization of a human astrovirus serotype 2 structural protein (vp26) that contains an epitope involved in virus neutralization. *Virology* 1994, 201, 312-320.

1.43. Bogdanoff, W.A.; Campos, J.; Perez, E.I.; Yin, L.; Alexander, D.L.; DuBois, R.M. Structure of a human astrovirus capsid-antibody complex and mechanistic insights into virus neutralization. *J Virol* 2016, 91.

Chapter 2: A Simplified Workflow for Monoclonal Antibody Sequencing

2.1 Acknowledgements

This chapter consists of a manuscript reprint which discusses our development of a novel antibody sequencing method, published in 2019 in *PLoS One*: Meyer L, López T, Espinosa R, Arias CF, Vollmers C, DuBois RM (2019) A simplified workflow for monoclonal antibody sequencing. *PLoS ONE* 14(6): e0218717. https://doi.org/10.1371/journal.pone.0218717.

My specific contributions to this project include conceptualization of the overall goals for the workflow, administration and troubleshooting of the specialized RT-PCR, blunt-end cloning and miniprep, expression and purification of the Spike 8 antigen, cloning, expression, and purification of chimeric mAb 2D9, total RNA extraction from transfected HEK 293F cells, and execution of the SDS-PAGE gel and ELISA comparing chimeric mAb 2D9 and mouse mAb 2D9. I analyzed and validated antibody sequencing data and wrote software to parse that data. I also made all figures and tables, wrote the first draft, and further edited the manuscript.

Tomás López, Rafaela Espinosa, and Carlos F. Arias contributed resources to this project. Carlos F. Arias also reviewed and edited the manuscript. Christopher Vollmers conceptualized the overall goals for the workflow, designed the specific methodology including primer sequences, and reviewed and edited the manuscript. Rebecca M. DuBois conceptualized the overall goals for the workflow, acquired funding, supervised the work related to the project, and reviewed and edited the

manuscript. This work was supported by the National Institutes of Health under grant 1RO1AI130073-01A1. We also acknowledge Walter Bogdanoff for the purification of mouse mAb 2D9, Marco A. Espinoza for the extraction of hybridoma RNA, Ashley Byrne for assistance with RT-PCR, Kate Mesa for the donation of HEK 293F cells, and Roger Volden for assistance with analysis of IgG subclass sequence conservation.

I thank my coauthors for their work, without which this project would not have been possible, and for their permission to reproduce our manuscript here.

2.2 Abstract

The diversity of antibody variable regions makes cDNA sequencing challenging, and conventional monoclonal antibody cDNA amplification requires the use of degenerate primers. Here, we describe a simplified workflow for amplification of IgG antibody variable regions from hybridoma RNA by a specialized RT-PCR followed by Sanger sequencing. We perform three separate reactions for each hybridoma: one each for kappa, lambda, and heavy chain transcripts. We prime reverse transcription with a primer specific to the respective constant region and use a template-switch oligonucleotide, which creates a custom sequence at the 5’ end of the antibody cDNA. This template-switching circumvents the issue of low sequence homology and the need for degenerate primers. Instead, subsequent PCR amplification of the antibody cDNA molecules requires only two primers: one primer specific for the template-switch oligonucleotide sequence and a nested primer to the respective constant region. We successfully sequenced the variable regions of five mouse monoclonal IgG

antibodies using this method, which enabled us to design chimeric mouse/human antibody expression plasmids for recombinant antibody production in mammalian cell culture expression systems. All five recombinant antibodies bind their respective antigens with high affinity, confirming that the amino acid sequences determined by our method are correct and demonstrating the high success rate of our method. Furthermore, we also designed RT-PCR primers and amplified the variable regions from RNA of cells transfected with chimeric mouse/human antibody expression plasmids, showing that our approach is also applicable to IgG antibodies of human origin. Our monoclonal antibody sequencing method is highly accurate, user-friendly, and very cost-effective.

2.3 Introduction

Recombinant monoclonal antibodies (mAbs) are a multibillion-dollar industry [2.1]. In contrast to monoclonal antibodies generated using traditional hybridoma-based methods and isolated from ascites fluid, recombinant monoclonal antibodies are produced by cloning antibody cDNA or synthetic sequences into expression plasmids and expressing in mammalian cell culture [2.2]. Before the design of recombinant antibody expression plasmids, sequencing of the antibody light and heavy chain variable regions is required. These variable regions determine antigen binding. It is therefore critical to obtain the correct sequence of the variable regions to maintain antibody affinity and specificity. In addition, knowledge of the variable region sequences and subsequent recombinant antibody expression reduces the impact of

hybridoma cell loss and hybridoma instability caused by mutations, chromosome deletions, or environmental factors [2.3].

There are several existing methods to sequence antibody variable regions from hybridoma cells or lymphocytes. Some involve the use of high-throughput RNA-sequencing technologies [2.4-2.6]. These methods prove highly accurate and allow for the analysis of antibody repertoires to great depths [2.6]. However, most labs are not familiar with high-throughput sequencing technologies, which require expertise for the preparation of RNA-seq libraries and for computational analysis. Furthermore, the cost of high-throughput library preparation and sequencing can be substantial, and turn-around time at sequencing cores can be weeks to months.

Other methods to sequence antibody variable regions use PCR and Sanger sequencing [2.7-2.13]. Variable region sequence determination by PCR-based approaches is challenging due to difficulties in designing universal primers that amplify all possible variable region sequences. This problem arises as a result of the inherent low sequence identity in the variable regions themselves as well as in the 5' leader sequence of antibody light and heavy chains, directly upstream of the variable regions [2.14]. Some approaches use sets of degenerate primers targeting the 5' region to overcome this issue [2.7-2.10]. However, these degenerate primers sometimes result in amplification success rates of only 80–90% because of non-specific priming or no priming [2.7, 2.10], meaning that 10–20% of antibody variable regions cannot be sequenced with these methods. An additional risk with degenerate primers is that the variable regions of the parent myeloma cell line can also amplify using these primers

[2.10]. Other approaches use 5' RACE (rapid amplification of 5' cDNA ends) [2.11, 2.12], but mRNA degradation, cDNA purification, and polyA tail addition in between reverse transcription and PCR makes this approach somewhat tedious [2.13]. A technique using non-degenerate primers also exists, but each variable region needs multiple amplification attempts with different sets of primers as well as further sequence validation with mass spectrometry [2.15]. Furthermore, there is a non-negligible risk of introducing primer-derived mutations in these methods.

In addition to nucleic acid-based approaches, there are de novo protein sequencing approaches to determine antibody variable regions by mass spectrometry [2.16-2.18], but these methods do not always lead to a single variable region sequence due to isobaric residues such as isoleucine and leucine [2.19]. A combination of X-ray crystallography and mass spectrometry positively identified variable region sequences [2.20]. However, this method is time-consuming, requires large amounts of purified monoclonal antibody, and is expensive.

Finally, researchers without access to these technologies may employ antibody sequencing services, such as those provided by GenScript, Syd Labs, Fusion Antibodies, or LakePharma [2.21-2.24]. Unfortunately, these services can become prohibitively expensive, costing at least $800 to sequence a single antibody's variable regions. Here, we successfully implemented a robust, simple, and affordable approach to sequence monoclonal antibody variable regions from RNA with a turn-around time of five days at a cost of $70 – $120 per antibody.

2.4 Results

2.4.1 Monoclonal antibody sequencing strategy

To sequence the variable regions of five mouse monoclonal IgG1 antibodies (2D9, 3B4, 3E8, 3H4, and 4B6) [2.25], we extracted total RNA from the hybridoma cell lines expressing these antibodies and applied a modified RT-PCR (reverse transcription polymerase chain reaction) using SMART (switching mechanism at 5' end of RNA transcript) technology [2.26, 2.27]. This technology is based on the intrinsic features of the reverse transcriptase from the Moloney murine leukemia virus (MMLV) and the application of a custom-sequence template-switch oligonucleotide (template-switch oligo) forward primer containing 3 riboguanines (rGrGrG) at its 3' end. To amplify antibody variable regions, we designed the RT-PCR reverse primers to be specific for highly conserved sequences in the constant regions of kappa, lambda, and IgG heavy chains of mouse antibodies (Tables 2.1 and 2.2).

Table 2.1: Mouse IgG reverse transcription primers.

Primer Name	Forward or Reverse	Primer Sequence
Template-switch oligo	Universal forward primer	5’ aagcagtggtatcaacgcagagtacatg$_r$g$_r$g$_r$ 3’
mIGK RT	Reverse primer for kappa chain	5’ ttgtcgttcactgccatcaatc 3’
mIGL RT	Reverse primer for lambda chain	5’ ggggtaccatctaccttccag 3’
mIGHG RT	Reverse primer for heavy chain	5’ agctgggaaggtgtgcacac 3’

Table 2.2: Mouse IgG PCR primers.

Primer Name	Forward or Reverse	Primer Sequence
ISPCR	Universal forward primer	5’ aagcagtggtatcaacgcagag 3’
mIGK PCR	Reverse primer for kappa chain	5’ acattgatgtctttggggtagaag 3’
mIGL PCR	Reverse primer for lambda chain	5’ atcgtacacaccagtgtggc 3’
mIGHG PCR	Reverse primer for heavy chain	5’ gggatccagagttccaggtc 3’

The RT-PCR amplification of antibody variable regions occurs as follows: To begin reverse transcription of a particular variable region, the reverse transcription primer specific for that antibody chain (either kappa, lambda, or heavy; Table 2.1) binds the hybridoma RNA within the constant region sequence at a highly conserved

site. The MMLV reverse transcriptase initiates polymerization (Figure 2.1, Step 1). After the MMLV reverse transcriptase reaches the 5’ end of the RNA template during first strand synthesis, it adds several nucleotides, usually deoxycytosine, to the 3' end of the cDNA transcript (Figure 2.1, Step 2). The reliable addition of these bases by the MMLV reverse transcriptase allows annealing of the template-switch oligo (Table 2.1). When base pairing occurs between the template-switch oligo’s 3’ riboguanines and the cDNA deoxycytosine overhang (Figure 2.1, Step 3), the MMLV reverse transcriptase switches templates and continues polymerization, now using the template-switch oligo as the template rather than the hybridoma RNA, until it reaches the 5’ end of the template-switch oligo (Figure 2.1, Step 4). The final product is a single-stranded cDNA molecule containing an initial universal sequence added by the template-switch oligo followed by the complete 5’ to 3’ sequence of the RNA template region (Figure 2.1, Step 5). This cDNA becomes the template for second-strand synthesis (Figure 2.1, Step 6) and amplification in PCR (Figure 2.1, Step 7) by taking advantage of the added universal sequence. The forward PCR primer (Table 2.2) has the same sequence as the template-switch oligo and therefore binds the universal sequence added to the cDNA transcript during reverse transcription. The reverse PCR primers (Table 2.2) are again specific for each chain’s constant region at a second highly conserved sequence but are nested within the cDNA sequence synthesized during reverse transcription to promote amplification specificity.

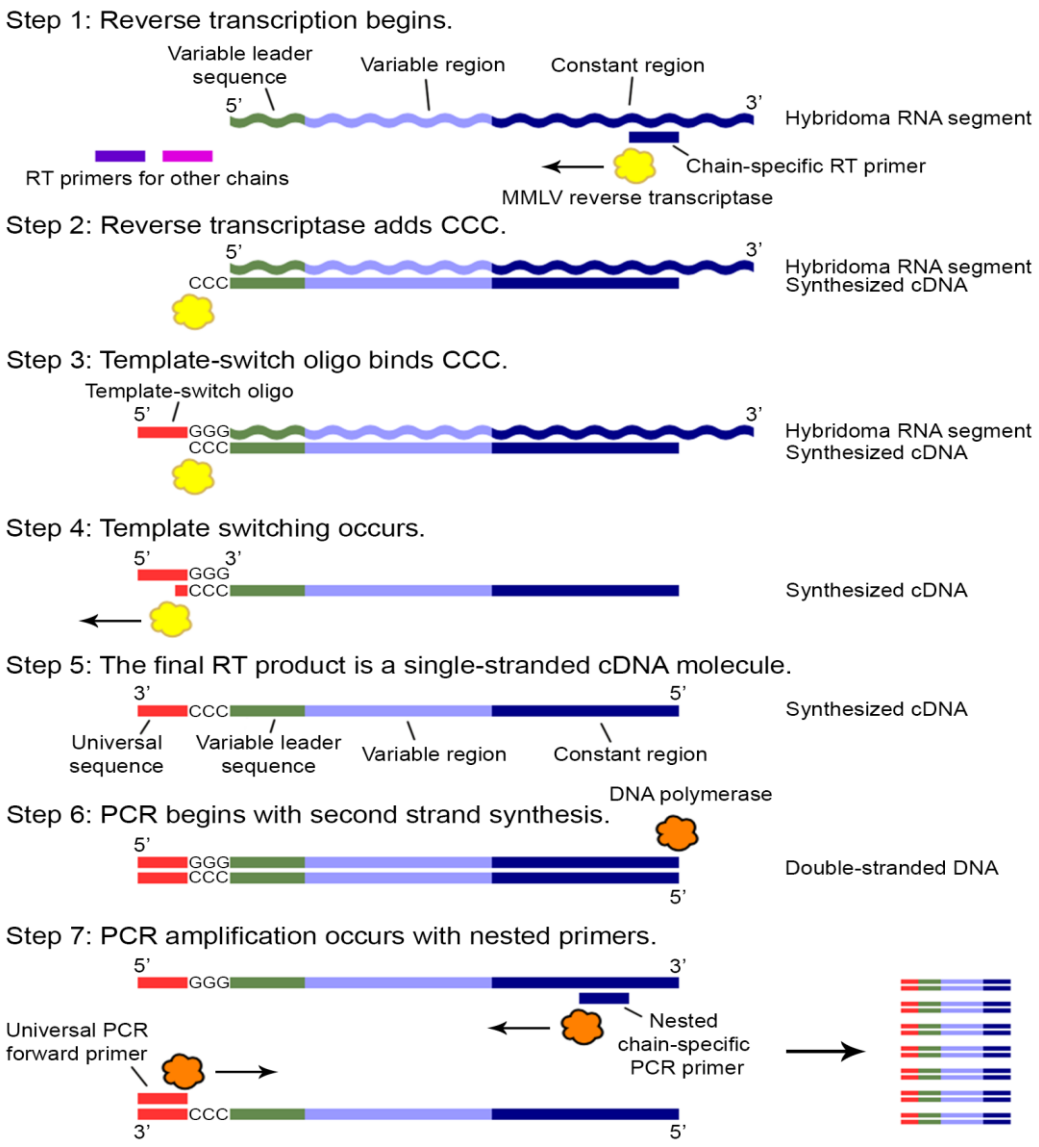

Figure 2.1: Schematic for cDNA synthesis by template-switching.

(Step 1) Primer binding and initiation of polymerization. **(Step 2)** MMLV reverse transcriptase adds deoxycytosines to the cDNA 3' end. **(Step 3)** Template-switch oligo binds the CCC overhang. **(Step 4)** Reverse transcriptase switches templates and continues polymerization using the template-switch oligo as the template. **(Steps 5 – 7)** The single-stranded cDNA product of reverse transcription becomes the template for second-strand synthesis primed by the universal PCR forward primer. Amplification follows using the universal PCR forward primer and nested chain-specific PCR reverse primers. Note that the lengths of the different antibody regions and primers are not drawn to scale.

We set up fifteen total RT-PCR reactions: five hybridoma RNA samples with three RT-PCR reactions each to amplify kappa, lambda, and heavy chain variable regions. We set up kappa and lambda chain amplifications for each antibody because we did not know which light chain was present. We checked these reactions by agarose gel electrophoresis. After optimizing primer design (see Section 2.4.2 Primer selection), 2D9, 3B4, 3E8, and 4B6 showed amplification of kappa chains, 3H4 showed amplification of both kappa and lambda chains, and all five samples showed amplification of heavy chains. Each amplicon is 550 – 600 base pairs in size (Figure 2.2B).

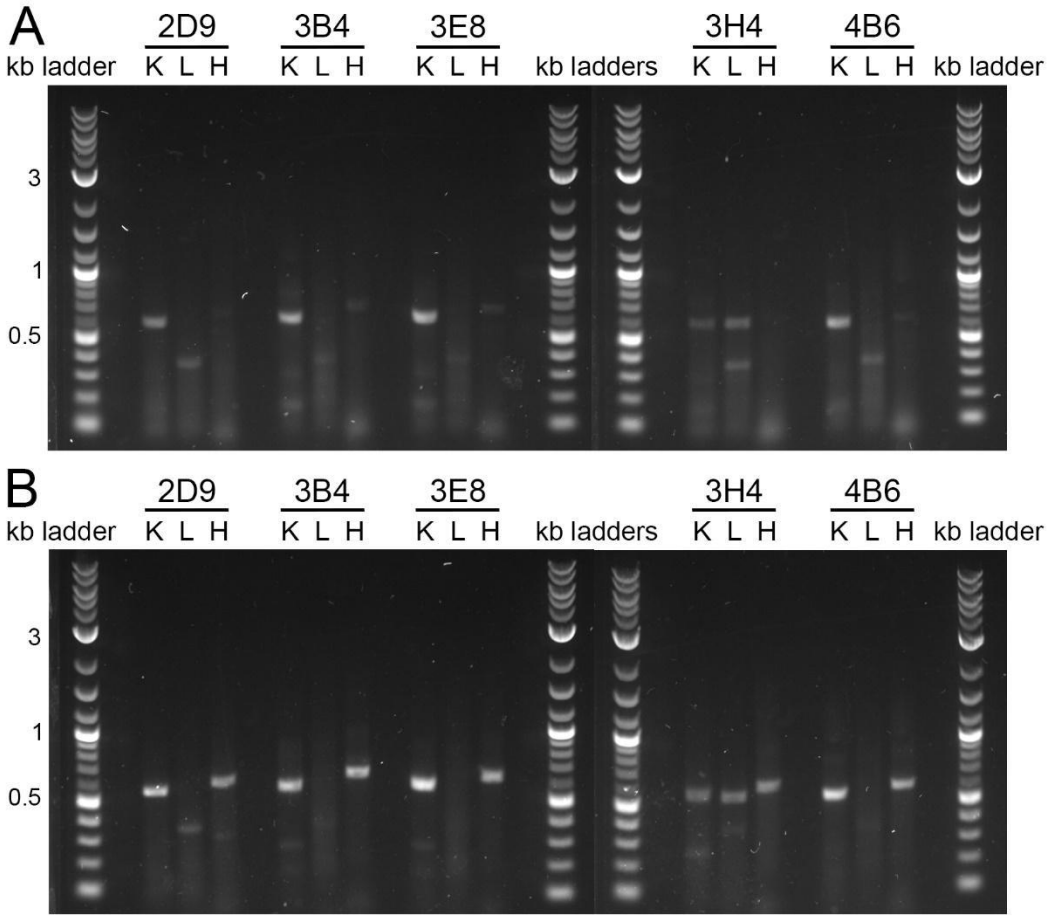

Figure 2.2: Comparison of primer sets for RT-PCR amplification of variable regions from 5 hybridoma mRNA samples.

K = kappa chain, L = lambda chain, H = heavy chain. **(A)** RT-PCR result using the same reverse primers for RT and for PCR. **(B)** RT-PCR result using a set of nested reverse primers for RT and for PCR.

2.4.2 Primer selection

To optimize RT-PCR amplification of the variable regions of mouse kappa, lambda, and IgG heavy chains, we designed and tested multiple sets of primers. We examined two strategies. For both strategies, we used the following primers for the RT step: the template-switch oligo forward primer and the reverse transcription primers

specific for kappa, lambda, or heavy chain constant regions, called mIGK RT, mIGL RT, and mIGHG RT, respectively (Table 2.1). Following reverse transcription, we implemented different strategies for PCR amplification. For the first strategy, we used the same reverse primers for both RT and PCR; therefore, in the PCR step, we used the ISPCR forward primer (Table 2.2, Row 1), which has the same 5' sequence as the template-switch oligo, and again used the mIGK RT, mIGL RT, or mIGHG RT reverse primers (Table 2.1). For the second primer strategy, the reverse PCR primers were nested within the sequence created by reverse transcription to promote specificity for the desired variable region amplicon (Figure 2.1, Step 7). In the PCR step in this strategy, we used the ISPCR forward primer and the nested reverse primers mIGK PCR, mIGL PCR, or mIGHG PCR (Table 2.2).

A comparison of the RT-PCR results obtained using the same reverse primers for RT and for PCR (Figure 2.2A) to the results obtained using nested primers (Figure 2.2B) clearly shows that nested primers produce better amplification of antibody variable regions. In particular, the heavy chains of each antibody did not amplify, or only amplified faintly, using the same reverse primers in RT and PCR but amplified well when choosing nested primers for PCR. In addition, using nested primers increased the intensity of the kappa amplification product for each antibody. Finally, use of the nested primers led to a reduction in intensity or even elimination of a ~350 base pair non-specific lambda chain amplicon. Therefore, we suggest using nested primers for monoclonal antibody sequencing.

We performed a Clustal Omega [2.28] multiple sequence alignment (not shown) of the constant regions from all subclasses of mouse IgG (IgG1, IgG2a, IgG2b, IgG2c, and IgG3) using sequences available on IMGT, the international ImMunoGeneTics information system [2.29]. Based on the alignment, we expect that mIGHG RT and mIGHG PCR can prime from the constant regions of IgG1, IgG2a, IgG2b, and IgG2c antibodies but likely not prime from the constant region of IgG3 antibodies because of five mismatches of the IgG3 constant region to each primer. Therefore, we predict that the primers given in Tables 2.1 and 2.2 may be used to sequence antibodies from the majority of all mouse IgG subclasses.

2.4.3 Sequencing results

Following RT-PCR, we purified the amplicons by agarose gel extraction and directly sequenced by Sanger sequencing to determine the variable region sequences of the light and heavy chains from all five antibodies. We analyzed the sequencing data with a custom Python program available on GitHub. We identified the sequences of one kappa and three heavy chains in this manner (Table 2.3 under Number of Amplicons Sequenced). However, the sequencing data was not clear enough to determine the remaining sequences. Therefore, we turned to a sequencing vector to improve the quality of the DNA sequenced. We purified RT-PCR products with a PCR clean-up kit, blunt-end cloned into a plasmid, transformed into *E. coli*, and sequenced plasmid clones by Sanger sequencing. We obtained clear sequencing data for plasmid clones from each of the variable regions (Table 2.3). This result is an improvement on

the outcome of direct PCR sequencing, which only positively identified four variable regions. We sequenced enough plasmid clones of each light and heavy chain variable region to compare at least three amino acid sequences of each variable region to confirm sequence identity (Table 2.3 under Number of Plasmid Clones Sequenced). We used Clustal Omega for these alignments (not shown).

Table 2.3: Results of sequencing RT-PCR products directly and following blunt-end cloning.

Antibody	Number of Amplicons Sequenced			Number of Plasmid Clones Sequenced		
	Total	Amplicons Containing Light Chain	Amplicons Containing Heavy Chain	Total	Clones Containing Light Chain	Clones Containing Heavy Chain
2D9	6	0	3	20	3	4
3B4	6	0	0	10	4	3
3E8	6	3	0	20	5	3
3H4	6	0	3	15	5	4
4B6	6	0	3	18	3	3

Once we identified a consensus sequence for each of the ten variable regions, we used the IgBLAST tool [2.30], a tool for alignment of immunoglobulin and T cell receptor variable domain sequences, to determine percent identity of our light and

heavy chain variable regions to IgBLAST reference sequences, i.e. the top-matched germline V genes. Table 2.4 shows the results of this query. Across the light and heavy chains we sequenced, percent identity to the reference ranges from 92.1 % to 100 % in the frame regions (FRs) and ranges from 87.5 % to 100 % in the complementarity-determining regions (CDRs). All regions of 3E8 kappa, 4B6 kappa, 2D9 heavy, and 3E8 heavy chains match 100% to the reference sequences. For light chains, average percent identity of all frame regions and CDRs to the references is 99.2 %. For heavy chains, average percent identity of all frame regions and CDRs to the references is 98 %. We repeated this analysis using the IMGT database, which calculated the average percent identity of an antibody's frame regions and CDRs combined instead of the percent identity of the individual regions. We also report these values in Table 2.4. These results support the conclusion that the sequences determined with our method are viable antibodies from hybridoma RNA.

Table 2.4: Percent identity to IgBLAST and IMGT reference sequences.

Antibody	Percent Identity of Light Chain (%)							
	to IgBLAST References							to IMGT
	FR 1	FR 2	FR 3	CDR	CDR	CDR	Total	Total
2D9	100	100	100	94.4	100	100	99.7	96
3B4	100	98	99.1	100	88.9	90	98.2	94
3E8	100	100	100	100	100	100	100	94
3H4	100	96.1	98.1	100	100	95.5	98.3	98
4B6	100	100	100	100	100	100	100	97
Average	100	98.8	99.4	98.9	97.8	97.1	99.2	95.8
Antibody	Percent Identity of Heavy Chain (%)							
	to IgBLAST References							to IMGT
	FR 1	FR 2	FR 3	CDR	CDR	CDR	Total	Total
2D9	100	100	100	100	100	100	100	100
3B4	97.3	100	94.7	100	91.7	---	96.5	96
3E8	100	100	100	100	100	100	100	99
3H4	98.7	96.1	92.1	91.7	90.5	87.5	94.2	93
4B6	100	100	99.1	100	95.2	100	99.3	99
Average	99.2	99.2	97.2	98.3	95.5	96.9	98	97.4

Percent identity of each region is reported as given by IgBLAST or IMGT. The total percent identity was calculated by IgBLAST and IMGT as the number of matches between the query and reference sequences over the length of the aligned sequence multiplied by 100. The averages refer to the average (mean) of all values in that column.

Because amplification of both the kappa and lambda chains occurred for 3H4, as opposed to amplification of only one of the light chains as for the other four antibodies (Figure 2.2B), we purified the amplicons, blunt-end cloned, and sequenced the RT-PCR products for both 3H4 kappa and 3H4 lambda. Sequencing revealed that the 3H4 kappa variable region contains a frameshift mutation in the V-gene/J-gene junction, resulting in an early stop codon. In contrast, the 3H4 lambda variable region has a properly in-frame V-gene/J-gene junction and aligns to the reference sequence. Therefore, 3H4 lambda is likely the correct 3H4 light chain while 3H4 kappa represents an abortive rearrangement and might originate from a hybridoma fusion partner transcript [2.31, 2.32]. An amino acid sequence comparison of 3H4 kappa and 3H4 lambda can be seen in Figure 2.3, in which the alignment and coloring were performed with Jalview [2.33]. The values given in Tables 2.3 and 2.4 are for 3H4 lambda.

Figure 2.3: Protein sequence comparison of variable regions for 3H4 kappa and 3H4 lambda.

Blue = Frame region, Orange = Complementarity-determining region, Red = J region out-of-frame, Green = J region in-frame 3H4 kappa (top) has an early stop codon due to a frameshift mutation. 3H4 lambda (bottom) is full-length.

2.4.4 Verification of antigen binding: Comparison of chimeric mAb 2D9 and mouse mAb 2D9

To validate our antibody sequences, we cloned, expressed, and purified five recombinant antibodies. Using the heavy chain and light chain amino acid sequences of each antibody, we engineered recombinant antibody constructs comprised of the variable regions from the original mouse antibody and the constant regions from the human IgG1 antibody VRC01 [2.34, 2.35]. We refer to these constructs as chimeric monoclonal antibodies in this manuscript.

Using antibody 2D9 as a representative sample, we compared chimeric mAb 2D9 to its corresponding mouse antibody. Chimeric mAb 2D9 was transiently expressed in human embryonic kidney (HEK) 293F cells and purified from the medium by Protein A beads. Mouse mAb 2D9, the original antibody, was isolated from mouse ascites fluid and purified by Protein G beads. An SDS-PAGE gel comparison (Figure 2.4A) shows that both the original mouse mAb 2D9 and chimeric mAb 2D9 express light and heavy chains (reducing lanes) that form an antibody complex of the correct size, ~145 kD (non-reducing lanes).

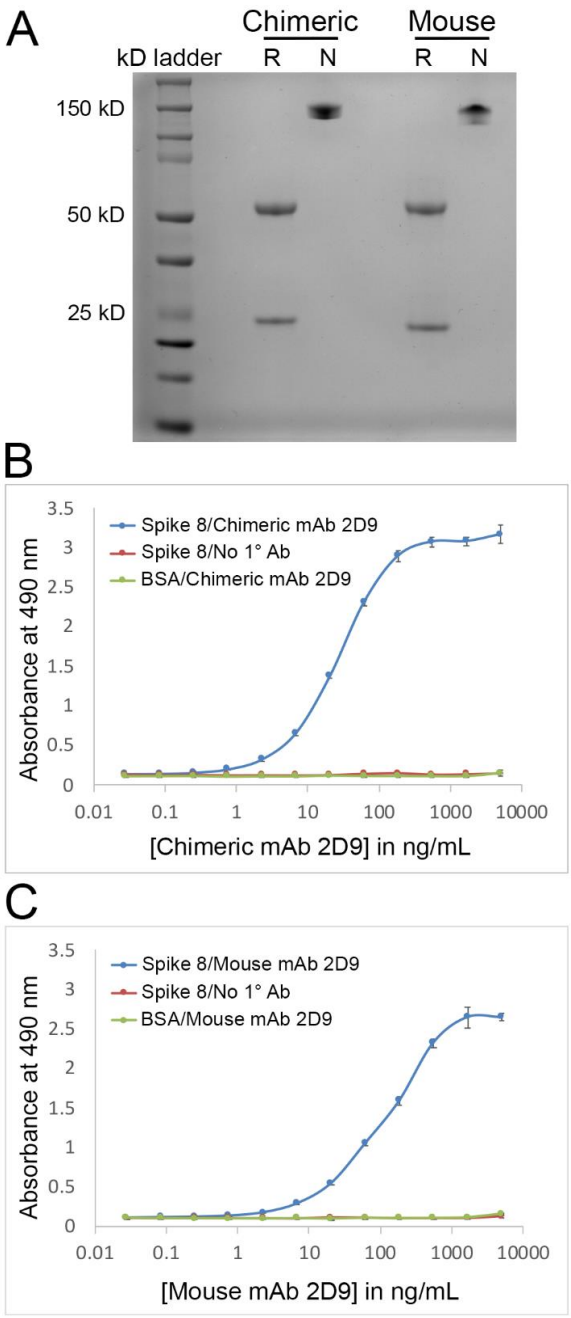

Figure 2.4: Comparison of chimeric mAb 2D9 and mouse mAb 2D9.

R = reducing gel sample, N = non-reducing gel sample **(A)** SDS-PAGE gel comparing chimeric mAb 2D9 (left) to mouse mAb 2D9 (right). A reducing (R) and a non-reducing (N) sample is shown for each mAb. **(B)** Indirect ELISA showing that chimeric mAb 2D9 binds the Spike 8 antigen. **(C)** Indirect ELISA showing that mouse mAb 2D9 binds the Spike 8 antigen.

Next, we tested the ability of the recombinant chimeric mAb 2D9 to bind its antigen, Spike 8, the recombinant capsid spike domain from human astrovirus serotype 8, against which mouse mAb 2D9 was raised [2.25]. By indirect ELISA, we found that both chimeric mAb 2D9 and mouse mAb 2D9 bind Spike 8 (Figure 2.4B and 2.4C). The other four chimeric mAbs also bind the human astrovirus spike against which they were raised. Thus, our simplified method of sequencing mouse IgG variable regions resulted in the correct sequences each time since all five recombinant chimeric antibodies constructed with those sequences retained their ability to bind the antigen used to raise the original mouse antibodies. This result indicates a performance rate of 100 % for our method.

2.4.5 Proof-of-concept: RT-PCR amplification of RNA from chimeric antibodies expressed in a human cell line

Our success in amplifying and sequencing mouse antibody variable regions from hybridoma RNA led us to perform a proof-of-concept experiment and apply the same RT-PCR method, including cycle conditions, to RNA extracted from HEK 293F cells transiently transfected with chimeric mAb 2D9 plasmid constructs. We designed new RT-PCR reverse primers (Tables 2.5 and 2.6), which were this time specific for the constant regions of human IgG antibodies rather than mouse IgG antibodies. Since we now knew that mAb 2D9 contains a kappa chain and a heavy chain and does not have a lambda chain, we designed RT-PCR primers only for human kappa and IgG heavy chains. As shown in Figure 2.5, RT-PCR using the human primers on RNA

extracted from HEK 293F cells was equally as successful as RT-PCR using the mouse primers on hybridoma RNA (Figure 2.2B). Therefore, we conclude that our method can be applied for the sequencing of human IgG variable regions as well as mouse IgG variable regions.

Table 2.5: Human IgG reverse transcription primers.

Primer Name	Forward or Reverse	Primer Sequence
Template-switch oligo	Universal forward primer	As above in Table 2.1
hIGK RT	Reverse primer for kappa chain	5’ gattggagggcgttatccacc 3’
hIGHG RT	Reverse primer for heavy chain	5’ gccgggaaggtgtgcacg 3’

Table 2.6: Human IgG PCR primers.

Primer Name	Forward or Reverse	Primer Sequence
ISPCR	Universal forward primer	As above in Table 2.2
hIGK PCR	Reverse primer for kappa chain	5’ tttggcctctctgggatagaag 3’
hIGHG PCR	Reverse primer for heavy chain	5’ agggcgcctgagttccacg 3’

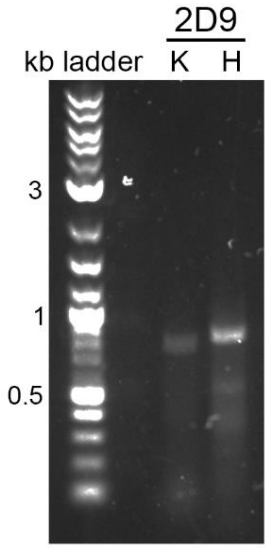

Figure 2.5: RT-PCR amplification of chimeric antibody variable regions.

K = kappa chain, H = heavy chain RT-PCR result with reverse primers designed for human constant regions and using as a template the RNA extracted from HEK 293F cells transiently transfected with chimeric mAb 2D9 constructs.

As for mouse IgG subclasses, we performed a Clustal Omega [2.28] multiple sequence alignment (not shown) of the constant regions from all subclasses of human IgG (IgG1, IgG2, IgG3, and IgG4) using sequences available on IMGT [2.29]. We predict that hIGHG RT and hIGHG PCR can prime from the constant regions of human antibodies of all IgG subclasses because of a high degree of sequence conservation in the priming regions, with a maximum of one mismatch between the primers and a particular IgG subclass' constant regions. Because of this high degree of conservation, we anticipate that our primers may be used to sequence antibodies from all human IgG subclasses.

We further expect that our method may be implemented to determine the variable region sequences of any antibody, provided that the proper constant region reverse primers for the antibody of interest are constructed. This feature makes our approach relevant for many users.

2.5 Discussion

We have developed a straightforward method to sequence any variable region from any antibody, as long as constant region reverse primers for the proper chain are designed. We showed that this method works for mouse IgG1 and human IgG1 antibodies and have provided these primer sequences, which are expected to cover the majority of mouse and human IgG subclasses. We demonstrated that implementation of this method is uncomplicated; it requires only basic laboratory skills and equipment, which also contribute to the method's affordability and rapidity. Compared to at least $800 to use an antibody sequencing service [2.21–2.24], we estimate that it cost us $70 –$120 to sequence one antibody. Sanger sequencing, at approximately $5 a reaction, makes up about 80 % of the cost and causes cost variation based on the number of plasmid clones sequenced. The duration of the method is as low as five days from RNA extraction to variable region sequence determination.

This method is broadly applicable. It can help researchers determine antibody sequences from brand-new hybridoma cell lines and from lymphocytes isolated from patients, thus allowing for production of recombinant antibodies and alternative antibody formats such as antibody-enzyme fusion proteins or bispecific antibodies.

This method can also verify that old hybridoma cell lines still produce the proper antibody. Finally, application of this method ahead of time prevents loss of the antibody sequence if hybridoma cells cease to produce antibody when removed from frozen storage.

The RT-PCR basis of the method allows for sequence determination of multiple antibody samples at once by simply setting up several RT-PCR reactions in parallel. In addition, a small RNA sample is sufficient; each RT-PCR reaction was set up with only 100 ng of cellular RNA. Note that 45 % of plasmid clones sent for sequencing contained DNA encoding antibody variable regions (Table 2.3). This percentage might be improved by purification of the RT-PCR product by gel extraction before blunt-end cloning rather than purification only by a PCR clean-up kit. Alternatively, Gibson assembly may be used to clone the RT-PCR products into a plasmid containing the template-switch oligo and reverse primer sequences.

Another advantage of this method is that it is not necessary to know if a monoclonal antibody contains a kappa or a lambda chain before RT-PCR amplification. Separate RT-PCR reactions can simply be set up for both types of light chain. In our case, for four of five antibodies, the correct light chain was clearly a kappa chain, as the lambda chain did not amplify. In the fifth case, Sanger sequencing distinguished between 3H4 kappa and 3H4 lambda and showed that 3H4 lambda chain is full-length. Thus, this method enables distinction between the two types of light chain.

2.6 Conclusions

In summary, we successfully developed a simplified and affordable workflow to consistently determine antibody variable region sequences from RNA. First, we applied our method to RNA from five hybridomas producing mouse IgG1 monoclonal antibodies. All five recombinant chimeric antibodies engineered with the variable region sequences determined by our method preserved their ability to bind antigen, showing that our method is highly efficient and reliable in retrieving the sequences for a cognate light chain/heavy chain pair. We also showed that this method can be applied to IgG1 antibodies of human origin. Based on sequence conservation, we expect that the given primer sequences can amplify variable regions from the majority of all mouse IgG and human IgG subclasses. Other users may implement this method for antibodies from other classes or other organisms simply by designing reverse primers for the constant regions of their antibody of study.

2.7 Materials and Methods

2.7.1 Hybridoma production and total RNA extraction

The hybridoma cells producing monoclonal antibodies were generated as described in [2.25]. Total RNA was extracted from the hybridoma cells using TRIzol Reagent (Invitrogen, 15596026) according to the manufacturer's instructions.

2.7.2 Reverse transcription to synthesize cDNA of antibody variable regions

The SMARTScribe Reverse Transcriptase kit from Clontech (Table 2.7) was used. Additionally used were the RNA samples (five hybridoma RNA samples from the mouse antibodies or one RNA sample from the chimeric antibody), primers (Tables 2.1 and 2.2 for mouse antibodies or Tables 5 and 6 for chimeric antibodies), 10 mM deoxynucleotide triphosphate mix (dNTPs), H2O, and an 80 U/μL RNAse inhibitor. Note: Make aliquots of the template-switch oligo and store at –80 °C due to its RNA content.

Table 2.7: Kits and antibodies.

Purpose	Manufacturer: Catalog Number	Kit Contents or Antibody Details
Reverse Transcription	Clontech: 639537	5x SMARTScribe buffer, 20 mM DTT, 100 U/μL SMARTScribe Reverse Transcriptase
PCR Clean-Up and Gel Extraction	Macherey-Nagel: 740609.50	Buffers NTI, NT3, and NE, DNA-binding columns
Blunt-End Cloning	Invitrogen: 450245	pCR-Blunt II-TOPO plasmid, Salt solution (1.2 M NaCl, 0.06 M $MgCl_2$), dNTP mix, M13 forward and reverse sequencing primers

Purpose	Manufacturer: Catalog Number	Kit Contents or Antibody Details
Miniprep	Macherey-Nagel: 740588.50	Buffers A1, A2, A3, AW, A4, and AE, RNAse A, DNA-binding columns
Effectene Transfection	Qiagen: 301425	Effectene, Enhancer, Buffer EC
Pierce Protein A Fab Preparation Kit	Thermo Fisher Scientific: 44985	Protein A beads, IgG elution buffer (pH 2.8, amine-based)
Goat Anti-Human IgG Fc Antibody, Horseradish Peroxidase Conjugate, Affinity Purified	Thermo Fisher Scientific: A18817	Horseradish-peroxidase-conjugated polyclonal 2° antibody with specificity for human antibody Fc region Antibody Registry ID: AB_2535594
Peroxidase-Conjugated AffiniPure Goat Anti-Mouse IgG, Fcγ Fragment Specific	Jackson ImmunoResearch: 115-035-071	Horseradish-peroxidase-conjugated polyclonal 2° antibody with specificity for mouse antibody Fc region Antibody Registry ID: AB_2338506

Reverse transcription was executed according to the following protocol: All reactions were kept on ice during setup. For each RNA sample, three cDNA synthesis reactions were set up: one for the kappa chain, one for the lambda chain, and one for the heavy chain. Ideally, only one of the light chains will amplify per antibody. 1. In PCR tubes, Mix #1 was prepared: 2 μL 50 ng/μL RNA, 1 μL 10 μM reverse RT primer based on antibody chain (e.g. mIGK RT, mIGL RT, or mIGHG RT for mouse antibodies), and 1 μL 10 mM dNTPs. For one RNA sample, three tubes of Mix #1 were needed, each containing a different reverse primer. 2. In a 0.5 mL Eppendorf tube, Mix #2 was prepared: 1.95 μL H2O, 2 μL 5x SMARTScribe buffer, 1 μL 20 mM DTT, and 0.3 μL 100 μM template-switch oligo. Volumes given for Mix #2 are for one cDNA synthesis reaction, so scale-up occurred as necessary, i.e. three times the volumes for Mix #2 were prepared per hybridoma RNA sample. One master mix of Mix #2 was prepared for all reactions. 3. Any RNA secondary structure was denatured by incubating the tubes containing Mix #1 at 72 °C for 3 minutes in a thermocycler. 4. During denaturation of Mix #1, the following was added to Mix #2: 0.25 μL 80 U/μL RNAse inhibitor and 0.5 μL 100 U/μL SMARTScribe Reverse Transcriptase per cDNA synthesis reaction. 5. 6 μL of Mix #2 was added to each tube of denatured Mix #1. 6. In the thermocycler, the combined mix was incubated at 42 °C for 60 minutes, then at 70 °C for 5 minutes to stop the reaction. The reactions were held at 4 °C. PCR amplification was done immediately after reverse transcription. No cDNA purification step was necessary.

2.7.3 PCR amplification of antibody variable regions

1. PCR reactions for each cDNA synthesis were set up: 10 μL 5x PCR buffer, 1 μL 10 mM dNTPs, 3 μL synthesized cDNA from the RT reaction, 2.5 μL 10 μM universal forward primer ISPCR, 2.5 μL 10 μM reverse PCR primer based on antibody chain (e.g. mIGK PCR, mIGL PCR, or mIGHG PCR for mouse antibodies), 30.5 μL H2O, and 0.5 μL 2 U/μL Phusion polymerase (or other high-fidelity polymerase). 2. A touch-down/step-down PCR was performed according to the following thermocycler conditions: 98 °C for 30 seconds; 10 cycles of 98 °C for 15 seconds, 63–57.5 °C for 30 seconds (decreasing the temperature by 0.5 °C each cycle), and 72 °C for 30 seconds; 15 cycles of 98 °C for 15 seconds, 56 °C for 30 seconds, and 72 °C for 30 seconds; followed by 72 °C for 7 minutes; and holding at 4 °C. 3. 5 μL of each RT-PCR reaction was run on a 1 % agarose gel in TAE buffer at 90 V. The amplified mouse antibody products appeared between 550–600 base pairs. The amplified human antibody products appeared between 750–850 base pairs. The Quick Load Purple 2-Log DNA Ladder (NEB, N0550S) was used as the standard.

2.7.4 Gel extraction and sequencing or PCR clean-up, blunt-end cloning, miniprep, and sequencing of antibody variable regions

The total volume of each RT-PCR reaction was run on a 1 % agarose gel in TAE buffer at 90 V. Bands of interest were excised, and DNA was extracted from the gel using Macherey-Nagel's PCR Clean-Up and Gel Extraction kit (Table 2.7). The

extracted DNA was Sanger sequenced by Sequetech Corporation using the ISPCR primer (Table 2.2).

Alternatively, the RT-PCR reactions were PCR-cleaned using Macherey-Nagel's PCR Clean-Up and Gel Extraction kit (Table 2.7). 2 μL of each PCR-cleaned product was blunt-end cloned into the pCR-Blunt-II-TOPO vector according to the blunt-end cloning kit manual (Table 2.7). Next, 3 μL of each TOPO cloning reaction was transformed into chemically competent *E. coli*. 100 μL of each transformation was spread on LB plates containing 50 μg/mL kanamycin and incubated at 37 °C overnight. After obtaining colonies, 5–10 colonies per antibody chain were inoculated in 5 mL LB/kanamycin medium and grown at 37 °C with 250 rpm shaking overnight. These cultures were miniprepped using Macherey-Nagel's miniprep kit (Table 2.7) and the resulting plasmid DNA was Sanger sequenced by Sequetech Corporation using the M13 forward primer.

A custom-written Python program was used to identify amplicons and plasmid clones containing antibody variable region sequences. The sequences originating from the same *E. coli* transformation, corresponding to a specific chain from a specific antibody, were then aligned using Clustal Omega to check for sequence consensus of the antibody chain. The final DNA sequence for each antibody variable region was submitted to IgBLAST with default parameters and mouse selected as the organism for query sequence to determine percent identity to the IgBLAST reference sequences for light and for heavy chains. The DNA sequences were also submitted to

IMGT/BlastSearch using default parameters to determine percent identity to the IMGT reference sequences.

2.7.5 Expression and purification of the Spike 8 antigen

A synthetic gene codon-optimized for *E. coli* expression encoding the human astrovirus serotype 8 capsid spike protein amino acids 424 to 648 (Spike 8, UniProtKB entry Q9IFX1) was purchased from Integrated DNA Technologies. To make the Spike 8 expression plasmid, the gene was cloned into pET52b (Addgene) in-frame with a C-terminal thrombin cleavage site and a 10-histidine purification tag. The plasmid was verified by DNA sequencing. Next, the plasmid was transformed into *E. coli* strain BL21(DE3). Cultures were inoculated and grown in LB/ampicillin medium. At an optical density of 0.6, protein production was induced with 1 mM isopropyl-D-thiogalactopyranoside (IPTG) at 18 °C for 18 hours. *E. coli* cells were lysed by ultrasonication in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, and 20 mM imidazole (Buffer A) containing 2 mM MgCl2, 0.0125 U/µL benzonase (Merck Millipore, 71205), and 1x protease inhibitor cocktail set V EDTA-free (Merck Millipore, 539137). Proteins were batch purified from soluble lysates with TALON metal affinity resin (GE Healthcare, 28-9574-99) and eluted with Buffer A containing 500 mM imidazole. Proteins were dialyzed overnight into 10 mM Tris-HCl pH 8.0, 150 mM NaCl (TBS) and further purified in TBS by size exclusion chromatography on a Superdex 75 column.

2.7.6 Expression and purification of chimeric mAb 2D9

A synthetic gene codon-optimized for insect cell expression containing the 2D9 kappa and heavy chain variable regions connected by a linker was ordered from Integrated DNA Technologies. The kappa and heavy chain variable regions were amplified separately and individually cloned by Gibson assembly into the two pCMV-VRC01 antibody backbone vectors for light and for heavy chains, which contain the constant regions of VRC01, a human anti-HIV antibody targeting the gp120 protein [2.34-2.36]. The resulting expression plasmids, pCMV-VRC01_2D9_kappa and pCMV-VRC01_2D9_heavy, contain the variable regions from the original mouse antibody 2D9 and the constant regions from a human IgG1 antibody under control of the human cytomegalovirus promoter. This same cloning procedure was accomplished with the four remaining antibodies. All plasmids were verified by DNA sequencing. Using the Effectene Transfection kit (Table 2.7), 2 μg of each of the 2D9 constructs was transiently co-transfected into HEK 293F cells (Thermo Scientific, R79007) obtained from a neighboring laboratory. The HEK 293F cells were seeded the day before at 0.5 x 10^6 cells/mL in 10 mL FreeStyle 293 medium (Gibco, 12338018). After 8 days of incubation at 37 °C with 5 % CO2, chimeric mAb 2D9 was purified from the HEK 293F cell medium with Protein A beads (Table 2.7). Chimeric mAb 2D9 was eluted with IgG elution buffer (pH 2.8, amine-based), and the elution was immediately neutralized with 2.0 M Tris pH 8.0.

2.7.7 Total RNA extraction from transfected HEK 293F cells

The extraction of RNA from HEK 293F cells transiently transfected with expression plasmids for chimeric mAb 2D9 was done according to the manufacturer's protocol for using TRIzol Reagent (Invitrogen, 15596026) to extract total RNA from cells grown in a monolayer. 3 mL TRIzol Reagent was used per T75 flask seeded 8 days prior with 10 mL of 0.5 x 10^6 HEK 293F cells/mL.

2.7.8 Purification of mouse mAb 2D9

Mouse mAb 2D9 was purified from mouse ascites fluid with Protein G beads (Thermo Scientific, 20398). Mouse mAb 2D9 was eluted with IgG elution buffer (pH 2.8, amine-based), and the elution was immediately neutralized with 2.0 M Tris pH 8.0.

2.7.9 SDS-PAGE gel comparing chimeric mAb 2D9 and mouse mAb 2D9

Each SDS-PAGE gel sample was prepared with 4 μg protein. For each monoclonal antibody, a reducing sample and a non-reducing sample was prepared. For the reducing samples, purified protein was mixed with 5x SDS-PAGE loading dye and boiled at 100 °C for 7 minutes. For the non-reducing samples, purified protein was mixed with 5x loading dye not containing β-mercaptoethanol and not boiled. The protein samples were loaded on a pre-cast 4-12 % Bis-Tris gel (Novex Life Technologies, NP0321BOX) and run in MES-SDS buffer at 140 V. The gel was stained by Coomassie Blue. The Precision Plus Protein Dual Color Standard (Bio-Rad, 1610374) was used as the standard.

2.7.10 ELISAs comparing Spike 8 binding by chimeric mAb 2D9 and mouse mAb 2D9

Each point was performed in triplicate. 150 µL per well of Spike 8 at 5 µg/mL in phosphate-buffered saline (PBS) was incubated overnight at room temperature on two 96-well ELISA microtiter plates. As a control, 150 µL per well of 5 µg/mL bovine serum albumin (BSA) in PBS was also incubated overnight. The plates were then washed three times with PBS containing 0.05 % Tween 20 (PBST). The wells were blocked by adding 150 µL of 5 % BSA in PBS to each well and incubating at room temperature for 1 hour followed by three PBST washes. Chimeric mAb 2D9 and mouse mAb 2D9, the primary antibodies, were diluted to 5 µg/ml with 1 % BSA in PBS. 150 µL of chimeric mAb 2D9 was added to wells in the first column of one ELISA plate and serially diluted 1:3 with 1 % BSA in PBS. This serial dilution was repeated on the other ELISA plate with mouse mAb 2D9. As a control, three rows of 5 µg/mL Spike 8 on each plate were left without primary antibody; 150 µL of 1 % BSA in PBS was added instead to the first wells of these rows and serially diluted 1:3. The plates were incubated for 1 hour at room temperature and then washed three times with PBST.

For the ELISA in which the primary antibody was chimeric mAb 2D9, the plate was incubated for 1 hour at room temperature with 150 µL per well of secondary antibody, a goat anti-human IgG Fc antibody conjugated to horseradish peroxidase (HRP) (Table 2.7) diluted 1:20,000 with 1 % BSA in PBS. For the ELISA in which the primary antibody was mouse mAb 2D9, the plate was incubated for 1 hour at room temperature with 150 µL per well of secondary antibody, a goat anti-mouse IgG Fc antibody conjugated to HRP (Table 2.7) diluted 1:8,500 with 1 % BSA in PBS. Then

the plates were washed three times with PBST and developed by adding 150 µL of 0.4 mg/mL horseradish peroxidase substrate o-phenylenediamine dihydrochloride (OPD) (Thermo Scientific, 34006) in 0.05 M phosphate-citrate buffer (pH 5.0) with 0.015 % hydrogen peroxide for 10 minutes at room temperature. The reactions were stopped by incubation with 150 µL of 2 N sulfuric acid for 10 minutes at room temperature. The absorbance was measured at 490 nm.

2.8 Software availability

The custom Python program written to analyze Sanger sequencing data is available at: https://github.com/Lena-Meyer/findAntibodies.

2.9 Abbreviations

BSA: Bovine serum albumin; CDR: Complementarity-determining region; ELISA: Enzyme-linked immunosorbent assay; HEK 293F: Human embryonic kidney cell line; IgG1: Immunoglobulin G subclass 1; mAb: Monoclonal antibody; MMLV: Moloney murine leukemia virus; 5’ RACE: Rapid amplification of 5’ cDNA end; RT-PCR: Reverse transcription polymerase chain reaction; SMART: Switching mechanism at 5’ end of RNA transcript; Spike 8: Recombinant capsid spike domain from human astrovirus serotype 8

2.10 Keywords

Monoclonal antibody sequencing, Hybridoma antibody sequencing, Variable region sequencing

2.11 References

2.1. Ecker DM, Jones SD, Levine HL. The therapeutic monoclonal antibody market. *MAbs*. 2015;7(1):9–14. pmid:25529996.

2.2. Frenzel A, Hust M, Schirrmann T. Expression of recombinant antibodies. *Front Immunol* 2013;4:217. Epub 07/29. pmid:23908655.

2.3. Coco-Martin JM, Oberink JW, van der Velden-de Groot TAM, Beuvery EC. Methods for studying the stability of antibody expression by hybridoma cells in homogeneous continuous culture systems. *Analytica Chimica Acta* 1991;249(1):257–62. https://doi.org/10.1016/0003-2670(91)87031-2.

2.4. Kuniyoshi Y, Maehara K, Iwasaki T, Hayashi M, Semba Y, Fujita M, et al. Identification of Immunoglobulin Gene Sequences from a Small Read Number of mRNA-Seq Using Hybridomas. *PLoS One* 2016;11(10):e0165473. Epub 10/27. pmid:27788226.

2.5. Schanz M, Liechti T, Zagordi O, Miho E, Reddy ST, Günthard HF, et al. High-throughput sequencing of human immunoglobulin variable regions with subtype identification. *PLoS One* 2014;9(11):e111726. Epub 11/03. pmid:25364977.

2.6. Henry KA. Next-Generation DNA Sequencing of VH/VL Repertoires: A Primer and Guide to Applications in Single-Domain Antibody Discovery. *Methods Mol Biol* 2018;1701:425–46. pmid:29116520.

2.7. Wang Z, Raifu M, Howard M, Smith L, Hansen D, Goldsby R, et al. Universal PCR amplification of mouse immunoglobulin gene variable regions: The design of degenerate primers and an assessment of the effect of DNA polymerase 3' to 5' exonuclease activity. *J Immunol Methods* 2000;233(1–2):167–77. pmid:10648866.

2.8. Krebber A, Bornhauser S, Burmester J, Honegger A, Willuda J, Bosshard HR, et al. Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J Immunol Methods* 1997;201(1):35–55. pmid:9032408.

2.9. Ihle Ø, Beckstrøm KJ, Michaelsen TE. Cloning, sequencing and expression of immunoglobulin variable regions of murine monoclonal antibodies specific for the P1.7 and P1.16 PorA protein loops of Neisseria meningitidis. *Scand J Immunol* 2003;57(5):453–62. pmid:12753502.

2.10. Koren S, Kosmač M, Colja Venturini A, Montanič S, Čurin Šerbec V. Antibody variable-region sequencing as a method for hybridoma cell-line authentication. *Appl Microbiol Biotechnol* 2008;78(6):1071–8. Epub 2008/03/05. pmid:18320189.

2.11. Ruberti F, Cattaneo A, Bradbury A. The use of the RACE method to clone hybridoma cDNA when V region primers fail. *J Immunol Methods* 1994;173(1):33–9. pmid:8034983.

2.12. Doenecke A, Winnacker EL, Hallek M. Rapid amplification of cDNA ends (RACE) improves the PCR-based isolation of immunoglobulin variable region genes from murine and human lymphoma cells and cell lines. *Leukemia* 1997;11(10):1787–92. pmid:9324302.

2.13. Kosmač M, Koren S, Škrlj N, Dolinar M, Čurin Šerbec V. Antibody Variable-Region Sequencing for Hybridoma Authentication. In: Ovesen K, Matthiesen U, editors. *DNA Fingerprinting, Sequencing and Chips*. DNA properties and modifications, functions and interactions, recombination and applications series. New York, NY: Nova Biomedical Books; 2009.

2.14. Dübel S, Breitling F, Fuchs P, Zewe M, Gotter S, Welschof M, et al. Isolation of IgG antibody Fv-DNA from various mouse and rat hybridoma cell lines using the polymerase chain reaction with a simple set of primers. *J Immunol Methods* 1994;175(1):89–95. pmid:7930642.

2.15. Babrak L, McGarvey JA, Stanker LH, Hnasko R. Identification and verification of hybridoma-derived monoclonal antibody variable region sequences using recombinant DNA technology and mass spectrometry. *Mol Immunol* 2017;90:287–94. Epub 09/01. pmid:28865256.

2.16. Bandeira N, Pham V, Pevzner P, Arnott D, Lill JR. Automated de novo protein sequencing of monoclonal antibodies. *Nat Biotechnol* 2008;26(12):1336–8. pmid:19060866.

2.17. Castellana NE, McCutcheon K, Pham VC, Harden K, Nguyen A, Young J, et al. Resurrection of a clinical antibody: Template proteogenomic de novo proteomic sequencing and reverse engineering of an anti-lymphotoxin-alpha antibody. *Proteomics* 2011;11(3):395–405. Epub 2011/01/27. pmid:21268269.

2.18. Sen KI, Tang WH, Nayak S, Kil YJ, Bern M, Ozoglu B, et al. Automated Antibody De Novo Sequencing and Its Utility in Biopharmaceutical Discovery. *J Am Soc Mass Spectrom* 2017;28(5):803–10. Epub 01/19. pmid:28105549.

2.19. Rickert KW, Grinberg L, Woods RM, Wilson S, Bowen MA, Baca M. Combining phage display with de novo protein sequencing for reverse engineering of monoclonal antibodies. *MAbs* 2016;8(3):501–12. pmid:26852694.

2.20. Bogdanoff WA, Morgenstern D, Bern M, Ueberheide BM, Sanchez-Fauquier A, DuBois RM. De Novo Sequencing and Resurrection of a Human Astrovirus-Neutralizing Antibody. *ACS Infect Dis* 2016;2(5):313–21. Epub 03/14. pmid:27213181.

2.21. GenScript. Antibody Sequencing Services [cited 2018 October 8, 2018]. Available from: https://www.genscript.com/mAb-sequencing.html.

2.22. SydLabs. Antibody Sequencing, Hybridoma Sequencing [October 8, 2018]. Available from: http://www.sydlabs.com/antibody-sequencing-hybridoma-sequencing-
p46.htm?gclid=CjwKCAiAlfnUBRBQEiwAWpPA6SA4TjEWIjBFlhBzTSsGBtjBIpzn8MkwJdNUd7I9Sfs1BbUKS80fKRoClIIQAvD_BwE.

2.23. FusionAntibodies. Antibody Sequencing Service [October 8, 2018]. Available from: https://www.fusionantibodies.com/services/antibody-sequencing?gclid=CjwKCAiAlfnUBRBQEiwAWpPA6a2FV_l06ik77zIhRCmxBb_kGg8ES9BQ6wnA2Qb-z8f2tUT5l7CGlhoCzXcQAvD_BwE.

2.24. LakePharma. Antibody Variable Region Sequencing and Cloning [October 8, 2018]. Available from: https://lakepharma.com/services/antibody-discovery/antibody-variable-region-sequencing-and-cloning.

2.25. Espinosa R, López T, Bogdanoff WA, Espinoza MA, López S, DuBois RM, et al. Isolation of neutralizing monoclonal antibodies to human astrovirus and characterization of virus variants that escape neutralization. *J Virol* 2018. Epub 10/24. pmid:30355681.

2.26. Takara. SMART Technology Overview [October 8, 2018]. Available from: https://www.takarabio.com/learning-centers/next-generation-sequencing/technology-and-application-overviews/smart-technology.

2.27. Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* 2001;30(4):892–7. pmid:11314272.

2.28. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011;7:539. Epub 10/11. pmid:21988835.

2.29. Lefranc MP. IMGT, The International ImMunoGeneTics Information System, http://imgt.cines.fr. Methods Mol Biol. 2004;248:27–49. pmid:14970490.

2.30. Ye J, Ma N, Madden TL, Ostell JM. IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res. 2013;41(Web Server issue):W34–40. Epub 05/13. pmid:23671333.

2.31. Carroll WL, Mendel E, Levy S. Hybridoma fusion cell lines contain an aberrant kappa transcript. *Mol Immunol* 1988;25(10):991–5. pmid:3146025.

2.32. Ding G, Chen X, Zhu J, Cao B. Identification of two aberrant transcripts derived from a hybridoma with amplification of functional immunoglobulin variable genes. *Cell Mol Immunol* 2010;7(5):349–54. Epub 07/26. pmid:20657605.

2.33. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009;25(9):1189–91. Epub 01/16. pmid:19151095.

2.34. Zhou T, Georgiev I, Wu X, Yang ZY, Dai K, Finzi A, et al. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* 2010;329(5993):811–7. Epub 07/08. pmid:20616231.

2.35. Wu X, Zhang Z, Schramm CA, Joyce MG, Kwon YD, Zhou T, et al. Maturation and Diversity of the VRC01-Antibody Lineage over 15 Years of Chronic HIV-1 Infection. *Cell* 2015;161(3):470–85. Epub 2015/04/09. pmid:25865483; PubMed Central PMCID: PMC4706178.

2.36. Barouch DH, Yang ZY, Kong WP, Korioth-Schmitz B, Sumida SM, Truitt DM, et al. A human T-cell leukemia virus type 1 regulatory element enhances the immunogenicity of human immunodeficiency virus type 1 DNA vaccines in mice and nonhuman primates. *J Virol* 2005;79(14):8828–34. pmid:15994776.

Chapter 3: Structures of Two Human Astrovirus Capsid/Neutralizing Antibody Complexes Reveal Distinct Epitopes and Antibody Inhibition of Virus Attachment to Host Cells

3.1 Acknowledgements

This chapter consists of a manuscript which focuses on two antibodies, 3E8 and 2D9, that target the human astrovirus capsid spike and block viral attachment to the host cell receptor. This manuscript has been submitted under my married name, Ricemeyer: Ricemeyer, L.; Aguilar-Hernández, N.; López, T.; Espinosa, R.; Lanning, S.; Mukherjee, S.; Cuellar, C.; López, S.; Arias, C.F.; DuBois, R.M. "Structures of Two Human Astrovirus Capsid/Neutralizing Antibody Complexes Reveal Distinct Epitopes and Antibody Inhibition of Virus Attachment to Host Cells." Submitted.

My specific contributions to this project include expression and purification of scFvs 2D9 and 3E8, determination of dissociation constants of the 3E8/Spike 8 and 2D9/Spike 8 interactions, and the BLI competition assay. I formed the scFv 3E8/Spike 8 and scFv 2D9/Spike 8 complexes, purified them, subjected them to crystal trials, and optimized crystals. I collected diffraction data on both complexes and performed data processing and refinement to obtain the structures. I made Figures 3.1 – 3.4 and all tables and gave assistance on Figures 3.5 and 3.6. I wrote the first draft and further edited the manuscript.

Nayeli Aguilar-Hernández, Tomás López, Rafaela Espinosa, and Susana López performed viral infectivity and attachment assays with mouse mAbs 3E8 and 2D9 and also made Figure 3.5. Sarah Lanning and Santanu Mukherjee performed microscopy experiments to visualize Spike 8 attachment to cells and inhibition of attachment by chimeric Fabs 3E8 and 2D9 and also made Figure 3.6. Carolina Cuellar expressed and purified Spike 8 and assisted with the scFv 2D9 expression and purification. Carlos F. Arias acquired funding and reviewed and edited the manuscript. Rebecca M. DuBois conceptualized the overall goals for the study, acquired funding, supervised the work related to the project, and reviewed and edited the manuscript.

We thank Jose Luis Martínez and Joaquín Moreno-Contreras for their assistance. This research was funded by NIH grant R01 AI144090 R.M.D. and C.F.A.. This work was partially supported by M0037-Fordecyt grant 302965 from the National Council for Science and Technology-Mexico (CONACyT) to S. López. L.R. was supported by an ARCS Foundation Fellowship. S. Lanning was supported by the NIH training grant T32 GM133391. C.C. was supported by an NIH Maximizing Access to Research Careers (MARC) Award and by a Koret Undergraduate Research Scholarship. Funding for the purchase of the Octet RED384 instrument was supported by the NIH S10 shared instrumentation grant 1S10OD027012-01. Beamline 8.3.1 of the Advanced Light Source, a U.S. DOE Office of Science User Facility under Contract No. DE-AC02-05CH11231, is supported in part by the ALS-ENABLE program funded by the National Institutes of Health, National Institute of General Medical Sciences, grant P30 GM124169-01. This research used resources of the Advanced Photon

Source, a U.S. Department of Energy (DOE) Office of Science User Facility, operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. We received technical support from Benjamin Abrams, UCSC Life Sciences Microscopy Center. RRID: SCR_021135.

I thank my many coauthors for their separate contributions to make this project a well-rounded study and for allowing its reprint here.

3.2 Abstract

Human astrovirus is an important cause of viral gastroenteritis worldwide. Young children, the elderly, and the immunocompromised are especially at risk for contracting severe disease. However, no vaccines or antiviral therapies exist to combat human astrovirus infection. Evidence points to the importance of antibodies in enabling protection of healthy adults from reinfection. To develop an effective subunit vaccine that broadly protects against diverse astrovirus serotypes, we must understand how neutralizing antibodies target the capsid surface at the molecular level. Here, we report the structures of the human astrovirus capsid spike domain bound to two neutralizing monoclonal antibodies. These antibodies bind two distinct conformational epitopes on the spike surface. We add to existing evidence that the human astrovirus capsid spike contains a receptor-binding domain and demonstrate that both antibodies neutralize human astrovirus by blocking virus attachment to host cells. We identify patches of conserved amino acids that overlap or border the antibody epitopes and may constitute

a receptor-binding site. Our findings provide a basis to develop therapies that prevent and treat human astrovirus gastroenteritis.

3.3 Importance

Human astroviruses infect nearly every person in the world during childhood and cause diarrhea, vomiting, and fever. Despite the prevalence of this virus, little is known about how antibodies block virus infection. Here, we determined crystal structures of the astrovirus capsid protein in complex with two virus-neutralizing antibodies. We show that the antibodies bind two distinct sites on the capsid spike domain. However, both antibodies block virus attachment to human cells. Importantly, our findings support the use of the human astrovirus capsid spike as an antigen in a subunit-based vaccine to prevent astrovirus disease.

3.4 Introduction

Astroviruses are a distinct family of small, non-enveloped RNA viruses that infect mammalian and avian species [3.1]. Members of the *Avastrovirus* genus have been associated with a variety of disease manifestations, growth defects, and mortality in poultry [3.2]. Members of the *Mamastrovirus* genus cause infections in humans and a wide range of other mammals, indicating the potential for a zoonotic disease transmission and the emergence of new astrovirus strains that could threaten human health [3.3]. Within *Mamastrovirus,* human astroviruses (HAstVs) are classified into eight canonical serotypes (HAstV1-8), where serotype 1 is the most prevalent globally [3.4-3.7]. HAstV is a leading worldwide cause of viral gastroenteritis but remains one

of the most inadequately understood enteric viruses [3.8]. In particular, young children, the elderly, and the immunocompromised are at risk for astrovirus infection, especially in developing countries [3.9-3.16]. The United States alone reports approximately 3.9 million cases of HAstV gastroenteritis each year [3.17], and children under 12 months may require hospitalization [3.18]. While HAstV infection accounts for an estimated 2 to 9 % of all acute non-bacterial gastroenteritis in healthy children worldwide [3.19], studies found that HAstV also causes persistent infections that spread easily in pediatric oncology wards [3.20, 3.21]. In addition to the canonical serotypes causing gastroenteritis, highly divergent MLB and VA clades have recently been associated with neurological complications such as encephalitis in immunocompromised and immunocompetent individuals, demonstrating that astroviruses not only infect cells in the gastrointestinal tract but also have systemic potential [3.22-3.25]. However, no vaccine or antiviral therapies preventing or treating human astrovirus infection have been developed.

A growing body of evidence underlines the importance of antibodies in protecting healthy adults from infection. Firstly, human astrovirus infection is rare in adults, indicating that a protective adaptive immune response conveying lifelong immunity develops during childhood [3.26]. In fact, greater than 75 % of healthy adults have anti-HAstV antibodies targeting at least one of the eight classical serotypes [3.7, 3.26], and seroprevalence rates increase with age [3.5]. Secondly, clinical studies determined that more severe disease is correlated with a lack of anti-HAstV antibodies in healthy volunteers [3.27, 3.28]. Finally, immunoglobulin replacement therapy

facilitated the recovery of an immunocompromised patient from persistent HAstV infection [3.29]. Together, these findings suggest that the incidence of HAstV gastroenteritis is likely underappreciated. In addition, these observations reveal that the adaptive immune response plays a crucial role in shielding an individual from HAstV disease. Accordingly, we anticipate that a vaccine eliciting protective antibodies will reduce HAstV infection in vulnerable populations. However, rational design of subunit vaccine immunogens or antiviral therapies relies on an understanding of the sites at which neutralizing antibodies bind human astrovirus and on insight into viral defenses against antibody neutralization.

HAstV particles contain a 6–7 kb, positive-sense, single-stranded RNA genome surrounded by a ~35 nm non-enveloped capsid protein shell. The genome's three open reading frames (ORFs) encode non-structural polyproteins (ORF1a and ORF1b) and the multi-domain capsid protein (ORF2) [3.30, 3.31]. This capsid protein contains a highly basic N-terminal region, a core domain, a spike domain, and a C-terminal acidic region [3.32]. During maturation, HAstV capsid proteins undergo a series of intra- and extracellular proteolytic cleavages that are required for infectivity [3.33-3.36]. Our lab and others have solved the crystal structure of the capsid core domain, which forms the T = 3 icosahedral shell that encapsidates the viral RNA genome [3.37, 3.38]. Our lab and others have also solved the crystal structure of the capsid spike domain, which forms the thirty dimeric spike projections on the surface of the mature virus particle [3.37, 3.39, 3.40].

While structural characterization of the human astrovirus capsid has made these advancements, further study is necessary to truly understand the functional sites on the HAstV capsid, including the location of the receptor-binding site(s), the epitopes where neutralizing antibodies bind, and the identity of the unknown host cell receptor(s). Only one neutralizing epitope, located on the capsid spike domain, has been defined by X-ray crystallography of an antibody/spike complex [3.40]. This study provided evidence that the spike is a receptor-binding domain and that this antibody neutralizes HAstV2 by obstructing a receptor-binding site on the spike. Further study has shown that while both the core and spike domains of the HAstV capsid are immunogenic, only the spike domain elicits antibodies that neutralize virus infectivity [3.41]. Indeed, the neutralizing monoclonal antibodies against HAstV for which neutralization mechanisms have been described all target the spike domain [3.41-3.43].

To better understand the HAstV antigenic structure and the mechanistic basis of virus neutralization, in this work we report the structures of the HAstV8 capsid spike (Spike 8) bound to two recently discovered HAstV8-neutralizing monoclonal antibodies, 3E8 and 2D9, at resolutions of 2.05 Å and 2.65 Å. These structures reveal two separate, non-overlapping epitopes that each contain conserved amino acids. We determine that both 3E8 and 2D9 neutralize HAstV8 by blocking attachment to cells. Our studies lay a foundation for the mechanistic understanding of HAstV infectivity and provide a blueprint for the design of vaccine immunogens that induce neutralizing antibodies against all classical HAstV serotypes.

3.5 Results

3.5.1 Neutralizing antibodies 3E8 and 2D9 have high affinity for the HAstV8 capsid spike

Previously, mouse monoclonal antibodies (mAbs) 3E8 and 2D9 were raised against recombinant Spike 8 in mice and shown to neutralize infectivity of HAstV8 in human colon adenocarcinoma cells (Caco-2 cells), the standard cell line for HAstV propagation [3.41]. We first examined the affinities of purified recombinant antibodies 3E8 and 2D9 for purified recombinant Spike 8. To assess and quantify binding, we performed biolayer interferometry (BLI) experiments. We used single-chain variable fragment (scFv) antibody constructs for these kinetic studies to eliminate potential complications of avidity with bivalent antibody and dimeric Spike 8. After a baseline measurement, Spike 8 fused to a 10x histidine tag was loaded onto Anti-Penta-His biosensors, which were subsequently submerged in 1:2 serial dilutions of scFv to measure on-rates, followed by submersion in buffer to measure off-rates. We find that scFv 3E8 and scFv 2D9 bind to Spike 8 in a dose-dependent manner with high-affinity dissociation constants (K_D) of 40.03 ± 2.37 nM and 2.45 ± 0.26 pM, respectively (Figure 3.1A and 3.1B and Table 3.1). From these results, we conclude that immunization of mice with the Spike 8 antigen produces high-affinity HAstV8-neutralizing antibodies. These studies support the development of a subunit vaccine for HAstV using recombinant spikes as immunogens.

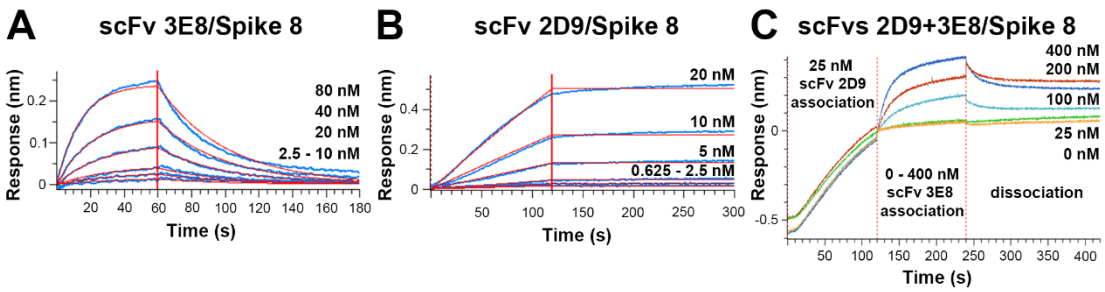

Figure 3.1: scFv 3E8 and scFv 2D9 bind Spike 8 with high affinity.

(A) Affinity determination for the scFv 3E8/Spike 8 interaction. **(B)** Affinity determination for the scFv 2D9/Spike 8 interaction. **(C)** Competition BLI assay showing simultaneous binding of scFvs 3E8 and 2D9. Data shown in 1A and 1B are from one representative experiment of the triplicate assays.

Table 3.1: scFv 3E8 and scFv 2D9 bind Spike 8 with high affinity.

	$K_D \pm \sigma$	R^2	χ^2
scFv 3E8/Spike 8	40.03 ± 2.37 nM	0.9900	0.4435
scFv 2D9/Spike 8	2.45 ± 0.26 pM	0.9968	0.8963

Next, we performed a competition BLI experiment to determine whether 3E8 and 2D9 compete for the same target on Spike 8 or whether they can bind Spike 8 simultaneously. We loaded Spike 8 onto the biosensors and first allowed scFv 2D9, which has the stronger affinity, to bind. Association of the higher-affinity scFv 2D9 first ensured that it would not be displaced upon association of the weaker-affinity scFv 3E8. We then dipped into dilutions of scFv 3E8 and observed an additional association signal, suggesting that 3E8 and 2D9 can bind Spike 8 at the same time, therefore at different epitopes (Figure 3.1C).

This result is in accordance with escape mutation studies done previously, in which HAstV8 escape mutants were generated using mouse mAbs 3E8 and 2D9 [3.41].

Virus was cultured with low concentrations of one mouse mAb. Under selective pressure, the virus mutated to escape mAb neutralization. After sequencing, the viral variant able to escape neutralization by mouse mAb 3E8 contained the mutation Y464H in the Spike 8 sequence. In contrast, the viral variant that can escape neutralization by mouse mAb 2D9 contained the mutation D597Y in the Spike 8 sequence. Both of these mutations map to the surface of Spike 8, but are ~30 Å distant from each other in space, hinting at non-overlapping epitopes. In addition, mAb 3E8 still effectively neutralized the mAb 2D9 viral escape mutant and vice versa. These results suggest that the Spike 8 epitopes recognized by mAb 3E8 and mAb 2D9 are situated in different locations on the Spike 8 surface but do not provide a complete picture of the epitope footprint(s) present on Spike 8.

3.5.2 scFv 3E8 binds Spike 8 at a known epitope

Previously, we had solved the structure of another neutralizing antibody, PL-2, in complex with the HAstV2 capsid spike (Spike 2) [3.40]. This PL-2 antibody mainly contacts Spike 2 Loop 1, and the epitope contains residues directly adjacent to 3E8 escape mutant Y464. Therefore, PL-2 and 3E8 may target a similar epitope. To define the mAb 3E8 binding site on Spike 8, we co-crystallized and solved the structure of the scFv 3E8/Spike 8 complex to 2.05 Å resolution (Figure 3.2 and Table 3.2). The structure confirms size exclusion chromatography data demonstrating that scFv 3E8 binds to Spike 8 in a 2:2 ratio, whereby one scFv 3E8 molecule makes contact with one protomer of the Spike 8 dimer, and a second scFv 3E8 molecule binds the same epitope

on the other Spike 8 protomer (Figure 3.2). We observed a conformation-dependent tertiary epitope on each Spike 8 protomer that overlaps with the PL-2 epitope (discussed further below). Four segments of amino acids from one linear Spike 8 molecule comprise the three-dimensional 3E8-targeted epitope.

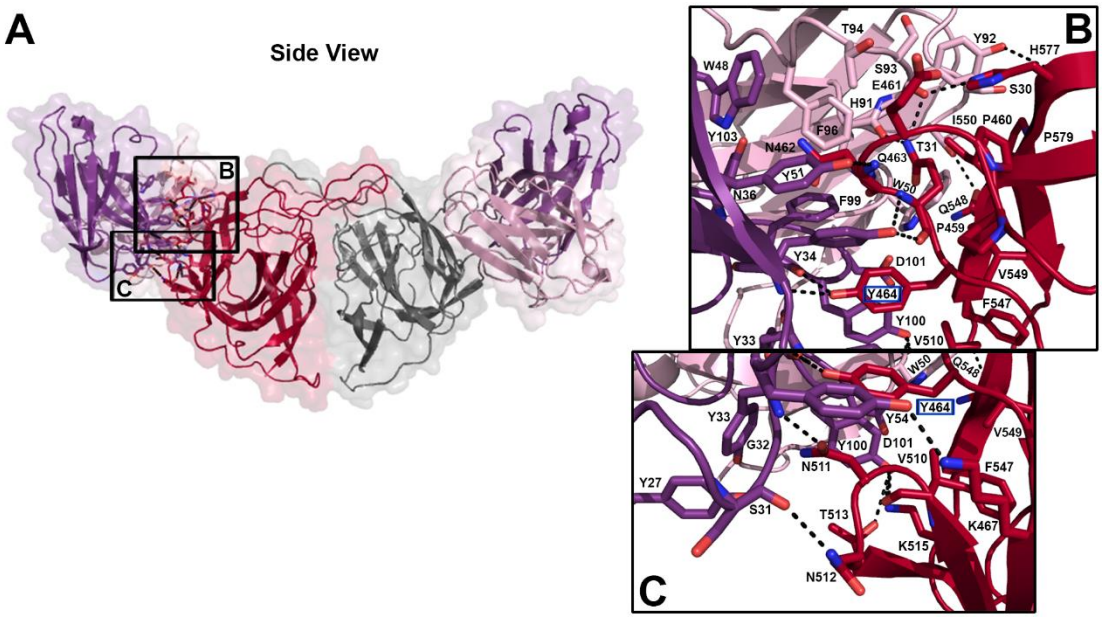

Figure 3.2: Intermolecular interactions at the scFv 3E8/Spike 8 interface.

(A) Semitransparent surface representation of the scFv 3E8/Spike 8 complex. The Spike 8 dimer is colored berry and grey, scFv 3E8 kappa chain is colored pale purple, and scFv 3E8 heavy chain is colored dark purple. Primary regions of interaction are highlighted by black boxes with labels that correspond to the zoomed-in views in panels B and C. Dashed lines represent hydrogen bonds. Escape mutant residue Y464 is indicated by a blue box around the residue name. **(B)** Interactions of 3E8 CDRs K3, H1, H2, and H3 with Spike 8 Loop 1. Interactions of 3E8 CDRs K1, K2, K3, and H3 with the Spike 8 β8 strand. Interactions of CDRs K1 and K3 with the Spike 8 β9 strand. **(C)** Interactions of 3E8 CDRs H1 and H3 with Spike 8 β5-β6 strands.

Table 3.2: Data collection and refinement statistics for scFv 3E8/Spike8 and scFv 2D9/Spike 8.

Characteristic	scFv 3E8/Spike 8	scFv 2D9/Spike 8
PDB code	7RK1	7RK2
Data collection		
Space group	P 1 21 1	P 21 21 21
Cell dimensions		
a, b, c (Å)	48.83, 80.43, 117.21	49.59, 97.34, 208.31
α, β, γ (°)	90.00, 90.67, 90.00	90.00, 90.00, 90.00
Resolution (Å)	80.43 – 2.05 (2.11 – 2.05)	208.31 – 2.65 (2.78 – 2.65)
R_{sym} or R_{merge}	0.184 (1.268)	0.451 (3.877)
I/σI	7.2 (2.2)	16.0 (2.0)
Completeness (%)	97.9 (94.8)	100 (99.8)
Redundancy	6.1 (5.3)	76.5 (66.6)
$CC_{1/2}$	0.992 (0.602)	0.998 (0.692)
Refinement		
Number of reflections	55,630	30,211
Resolution (Å)	66.32 – 2.05	104.15 – 2.65
R_{work}/R_{free}	0.223/0.238	0.218/0.247
Number of atoms	7378	7105
Protein	7043	6955
Ligands	0	0
Water	335	150
B factors	32.18	40.05
Protein	32.13	40.15
Ligands	N/A	N/A
Water	33.32	35.52
RMSD		
Bond lengths (Å)	0.012	0.013
Bond angles (°)	1.70	1.49
Ramachandran statistics		
Favored (%)	97.5	97.5
Allowed (%)	2.5	2.5
Outliers (%)	0	0

Data from one crystal were used for each structure determination. Values in parentheses are for the highest-resolution shell.

All six complementarity-determining regions (CDRs) from the kappa (K) and heavy (H) chains of each scFv 3E8 interact with one protomer of Spike 8 (Figure 3.2). The interface is defined by several contact points: (1) scFv 3E8 CDR H2 contacts Spike 8 Loop 1 residues P459, P460, E461, N462, and K467, while scFv 3E8 CDR H1 contacts N462, Q463, and Y464. CDRs K3 and H3 also contact Spike 8 Q463 and Y464, respectively (Figure 3.2B). (2) scFv 3E8 CDR H3 contacts Spike 8 β5-β6 residues V510, N511, T513, and K515, and CDR H1 contacts N511 and N512 (Figure 3.2C). (3) scFv 3E8 CDR H3 contacts Spike 8 β8 residue F547, CDRs K1 and K2 contact Q548, and CDRs K2 and K3 contact V549 and I550 (Figure 3.2B). (4) scFv 3E8 CDR K3 contacts Spike 8 β9 residue H577, and CDRs K1 and K3 contact P579 (Figure 3.2B). scFv 3E8 binds Spike 8 through a large network of hydrogen bonds, electrostatic bonds, and van der Waals interactions. Many of the amino acids in the 3E8 epitope are unique to serotype 8, consistent with virus neutralization data demonstrating that 3E8 is serotype specific [3.41]. However, three residues, K515, F547, and V549, are strictly conserved across all eight serotypes, while P460 and K467 are mostly conserved. Spike 8 residue Y464, the location of the Y464H escape mutant, is buried by scFv 3E8 binding (Figure 3.2B and 3.2C). Compared to the wild-type tyrosine, the side chain of the histidine mutant may be too short to form hydrogen bonds with the backbone nitrogen of scFv 3E8 CDR H1 residue Y34 and/or the backbone carbonyl oxygen of scFv 3E8 CDR H3 residue F99, thereby affecting antibody binding. Moreover, Y464 makes pi-stacking interactions with scFv 3E8 CDR H1, H2, and H3 residues Y34, Y54, F99, and Y100, all within 3 to 5 Å, and mutation to histidine may

diminish these interactions. The finding that Spike 8 residue Y464 is part of the epitope signifies that the escape mutation studies accurately pointed to the approximate location of the epitope. However, only the structure of the entire complex revealed the full extent of the scFv 3E8/Spike 8 binding interface.

3.5.3 scFv 2D9 binds Spike 8 at a novel epitope

To fully characterize the mAb 2D9 binding site on Spike 8 and compare it to the 3E8 and PL-2 binding site, we next co-crystallized and solved the structure of the scFv 2D9/Spike 8 complex to 2.65-Å resolution (Figure 3.3 and Table 3.2). The structure confirms size exclusion chromatography data demonstrating that scFv 2D9 also binds to Spike 8 in a 2:2 ratio, the same ratio as for scFv 3E8/Spike 8 and scFv PL-2/Spike 2. In contrast to the 3E8 tertiary epitope, we find that scFv 2D9 binds to a conformation-dependent quaternary epitope across the Spike 8 dimer interface that is fully distinct from the epitope targeted by 3E8 and PL-2. Five patches of amino acids from two linear Spike 8 molecules come together in the three-dimensional structure.

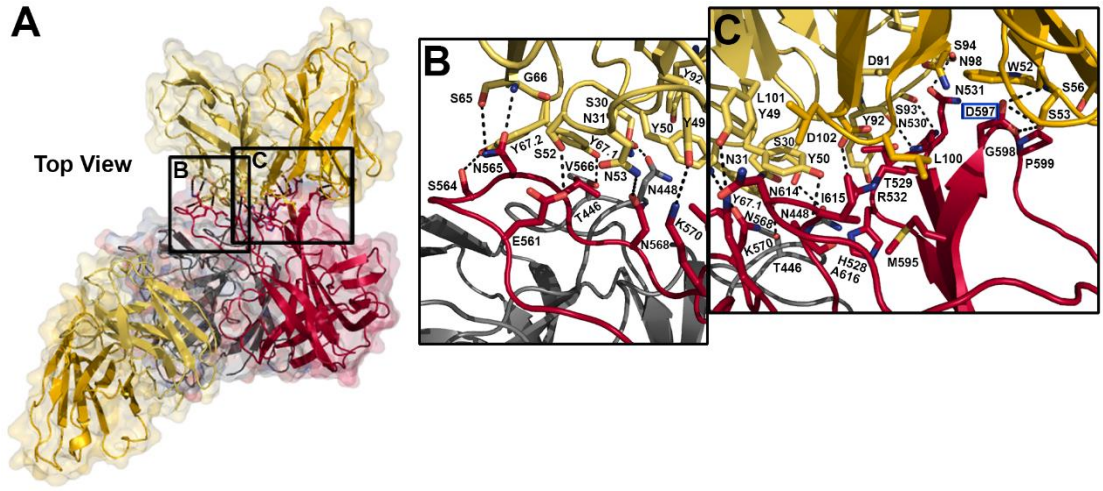

Figure 3.3: Intermolecular interactions at the scFv 2D9/Spike 8 interface.

(A) Semitransparent surface representation of the scFv 2D9/Spike 8 complex. The Spike 8 dimer is colored berry and grey, scFv 2D9 kappa chain is colored pale yellow, and scFv 2D9 heavy chain is colored golden yellow. Primary regions of interaction are highlighted by black boxes with labels that correspond to the zoomed-in views in panels B and C. Dashed lines represent hydrogen bonds. Escape mutant residue D597 is indicated by a blue box around the residue name. **(B)** Interactions of 2D9 CDRs K1 and K2 with Spike 8 Loop 1 (grey protomer) and Spike 8 Loop 3 (berry protomer). **(C)** Interactions of 2D9 CDRs K2, K3, H2, and H3 with Spike 8 β6-β7 and β10 strands as well as Loop 4 (all on the berry protomer).

Five complementarity-determining regions (CDRs) from the kappa (K) and heavy (H) chains of each scFv 2D9 contact Spike 8 (Figure 3.3). The interface is characterized by several interactions: (1) scFv 2D9 kappa chain residue Y67 near CDR K1, which exhibits electron density for two conformations in the structure, contacts T446 with Orientation 1 (Y67.1) in Loop 1 of one Spike 8 protomer. In addition, scFv 2D9 CDR K1 contacts N448 on that protomer (Figure 3.3B). The remaining interactions take place on the other Spike 8 protomer, indicating a quaternary epitope: (2) scFv 2D9 CDR K3 contacts Spike 8 residues H528, T529, N530, and N531 in the

β6-β7 junction. CDR H3 also contacts T529 and R532. CDR K2 contacts R532 (Figure 3.3C). (3) scFv 2D9 CDR K2 contacts Spike 8 E561, scFv 2D9 kappa chain Y67 Orientation 2 (Y67.2) contacts Spike 8 S564, and scFv 2D9 kappa chain S65 and G66 contact N565. scFv 2D9 kappa chain Y67.1 and CDR K1 bury Spike 8 V566. CDR K2 contacts Spike 8 N568 and K570. These Spike 8 residues are all in Loop 3 (Figure 3.3B). (4) scFv 2D9 CDR H3 contacts Spike 8 M595 in the β10 strand. scFv 2D9 CDR H2 contacts Spike 8 β10 residues D597, G598, and P599 (Figure 3.3C). (5) scFv 2D9 CDR K2 contacts Spike 8 Loop 4 residues N614 and A616. CDR H3 buries Spike 8 I615 (Figure 3.3C). Similarly to scFv 3E8, scFv 2D9 relies on a large system of hydrogen bonds, electrostatic bonds, and van der Waals interactions to bind to a quaternary epitope on the surface of the Spike 8 dimer. Notably, Spike 8 residue D597, which mutated to Y during escape mutant studies, is buried by scFv 2D9 binding (Figure 3.3C). The mutation of aspartate to the larger tyrosine likely caused steric clashing with scFv 2D9 CDR H2 residue S53, completely prohibiting antibody binding.

While some amino acids in the epitope bound by scFv 2D9 are unique to HAstV8, the epitope also overlaps or abuts several patches of amino acids that are highly conserved between serotypes 1 – 8, suggesting that scFv 2D9 may block an important functional site on the Spike 8 surface. Interestingly, the epitope includes all four amino acids (T529, N530, N531, and R532) of an exposed, flexible β-turn knob which is very well-conserved among all serotypes and hypothesized to interact with other viral or host proteins (Figure 3.3C and Figure 3.4C) [3.39]. Critically, there is no overlap in the Spike 8 residues targeted by scFv 3E8 and scFv 2D9, indicating two

separate and distinct epitopes on the human astrovirus capsid spike surface (Figure 3.4A and 3.4B).

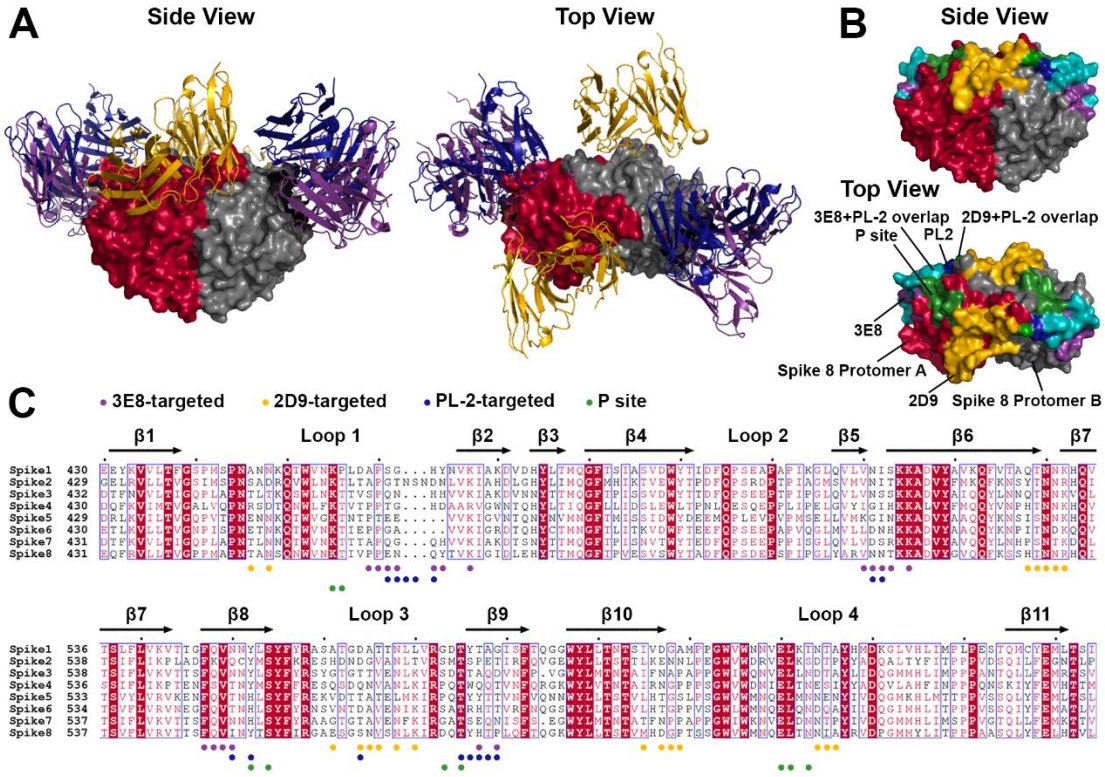

Figure 3.4: Comparison of epitopes for 3E8/Spike 8, 2D9/Spike 8, and PL-2/Spike 2 in relation to the P site.

(A) Structure of scFv 3E8 (purple), scFv 2D9 (yellow), and PL-2 (blue) bound to spike (berry and grey) viewed from the side and from the top. **(B)** Epitope footprints on spike. The spike dimer and the footprints for 3E8, 2D9, and PL-2 are colored as in (A). The eight-residue 3E8+PL-2 overlap is colored teal. The single-residue 2D9+PL-2 overlap is colored grass green. The P site is colored forest green. **(C)** Sequence alignment of the capsid spike domains of HAstV1-8. Conserved, strongly similar, weakly similar, and non-conserved amino acids are colored on a gradient where darker color represents more conserved. Alignments were done with Clustal Omega [3.44], and mapping of conservation onto the structure was performed with the online ESPript server [3.45]. Dots below the sequence alignment indicate residues in the epitopes targeted by 3E8 (purple), 2D9 (yellow), and PL-2 (blue) as well as residues in the P site (green).

3.5.4 Antibodies 3E8, 2D9, and PL-2 target epitopes near the P site

The human astrovirus host cell receptor is unidentified. In addition, the receptor-binding site(s) on the HAstV capsid surface is unknown. Because 3E8, 2D9, and PL-2 all inhibit virus attachment to cells (discussed further below) [3.40], they likely each prevent virus access to the receptor. Several putative receptor-binding sites on the spike surface have been proposed [3.39]. Among them is the P site, located in a shallow groove on top of the spike domain. The P site has many hydrophilic residues with side chains exposed to solvent, making it highly accessible to potential cell receptors without steric hindrance. While Spike 8 residues in the 3E8, 2D9, and PL-2 epitopes mostly do not overlap with residues in the P site, the residues in the epitopes appear to frame the P sites on each Spike 8 protomer (Figure 3.4B). If the P site plays a role in virus attachment to the host cell receptor, its placement between the epitopes may lead to a loss of virus access to the receptor when any of these three antibodies bind.

Figure 3.4C presents a sequence alignment of the spikes from all eight classical HAstV serotypes, with the residues in the P site as well as in the epitopes targeted by 3E8, 2D9, and PL-2 indicated. Eight residues are located in the epitopes targeted by both 3E8 and PL-2 (although N511 is the only one with sequence identity between Spike 2 and Spike 8), which reveals the particular vulnerability to antibody binding of this region across HAstV serotypes. While there is no overlap between the 3E8 and 2D9 epitopes, S564 (D566 in Spike 2) is found in both the 2D9 and PL-2 epitopes. Sequence conservation may indicate that 3E8, 2D9, and PL-2 obstruct a conserved

functional site on the virus surface directly or by proximity. In the next section, we demonstrate that 3E8 and 2D9 neutralize HAstV8 by the same mechanism despite their fully separate epitopes.

3.5.5 Antibodies 3E8 and 2D9 block attachment of HAstV8 to Caco-2 cells

First, we confirmed that mAbs 3E8 and 2D9 reduce virus infectivity. Purified HAstV8 particles were pre-incubated with serial dilutions of the ascites fluids of either 3E8 or 2D9, added to Caco-2 cell monolayers, and washed. The next day, infected cells were detected by an immunoperoxidase focus-forming assay. Virus infectivity decreased with increasing amounts of antibody (Figure 3.5A). To determine the neutralization mechanisms of 3E8 and 2D9, we next explored if the antibodies could block attachment of the virus to the surface of Caco-2 cells. Purified HAstV8 was pre-incubated with serial dilutions of the ascites fluids of either 3E8 or 2D9, and the virus-antibody mix was added to Caco-2 cell monolayers on ice. At this temperature, the virus is able to attach to the cell surface but does not enter the cell. The unbound virus was removed with washing, total RNA was extracted, and the viral genome was quantified by RT-qPCR. As can be seen in Figure 3.5B, both antibodies were able to efficiently impede HAstV8 attachment, reducing binding of the virus by more than 90 % at a 1:2500 dilution of the ascites fluids, as compared to virus bound in the absence of antibody. As a negative control, we added mAb 3B4, specific for HAstV1 and previously shown not to neutralize HAstV8 infectivity; no reduction of virus

binding was observed with this antibody. This result shows a direct correlation between the inhibition of virus infectivity and of virus binding.

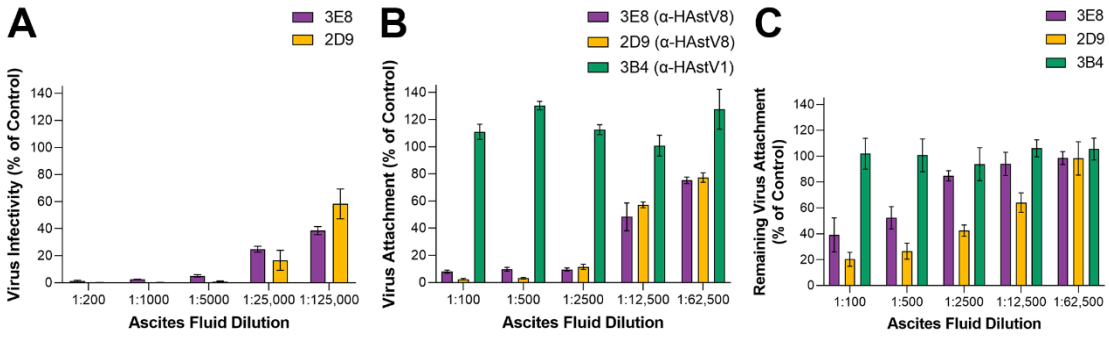

Figure 3.5: Antibodies 3E8 and 2D9 reduce virus infectivity, block virus attachment, and detach the virus from Caco-2 cells.

(A) 3E8 and 2D9 reduce virus infectivity. **(B)** 3E8 and 2D9 block virus attachment when HAstV8 is pre-incubated with the antibodies. **(C)** 3E8 and 2D9 detach HAstV8 from cells when the antibodies are added after the virus has attached.

To confirm that mAbs 3E8 and 2D9 inhibit virus attachment directly and not through aggregation of virus particles, the ability of the antibodies to detach the virus once it has bound to the cell surface was evaluated. For this assay, purified HAstV8 was added to Caco-2 cells and incubated on ice. The unbound virus was removed with washing, and then serial 1:5 dilutions of the ascites fluids of 3E8, 2D9, and 3B4 were added to the cells and further incubated on ice. After this incubation, the cells were washed again, and the viral RNA was quantified as described above. As can be seen in Figure 3.5C, both mAbs 2D9 and 3E8, although 2D9 more efficiently, were able to detach the virus from the cell surface, presumably by competition of the antibody with

the virus-receptor interaction. Negative control 3B4 was unable to detach the virus. Altogether, these results indicate that antibodies 3E8 and 2D9 neutralize virus infectivity by inhibiting the attachment of the virus to the cell surface.

3.5.6 Antibodies 3E8 and 2D9 block attachment of GFP-Spike 8 to Caco-2 cells

Finally, to test if the HAstV capsid spike alone is sufficient for cell attachment and to show that antibodies block this attachment, we examined the capacity of a recombinantly expressed fluorescent GFP-Spike 8 fusion protein to attach to Caco-2 cells and the ability of recombinant antigen binding fragments (Fabs) 3E8 and 2D9 to prevent this attachment. Caco-2 cells grown on glass coverslips were incubated with GFP alone, GFP-Spike 8 alone, or with GFP-Spike 8 and Fab 3E8, 2D9, or 3B4. Cells were washed, fixed, and imaged by fluorescence microscopy. As shown in Figure 3.6, Caco-2 cells incubated with PBS alone or GFP alone had little to no green signal. In contrast, Caco-2 cells incubated with GFP-Spike 8 had bright green punctate patterns that appear at or near the cell membranes. Furthermore, cells incubated with GFP-Spike 8 and Fab 3E8 or Fab 2D9 had little to no puncta, suggesting specific blocking of Spike 8 attachment to cells by the Fabs. The negative control Fab 3B4 did not block attachment. These data reveal that the human astrovirus capsid spike alone is sufficient for cell attachment and that antibodies 3E8 and 2D9 are able to preclude virus attachment by specifically blocking the spike.

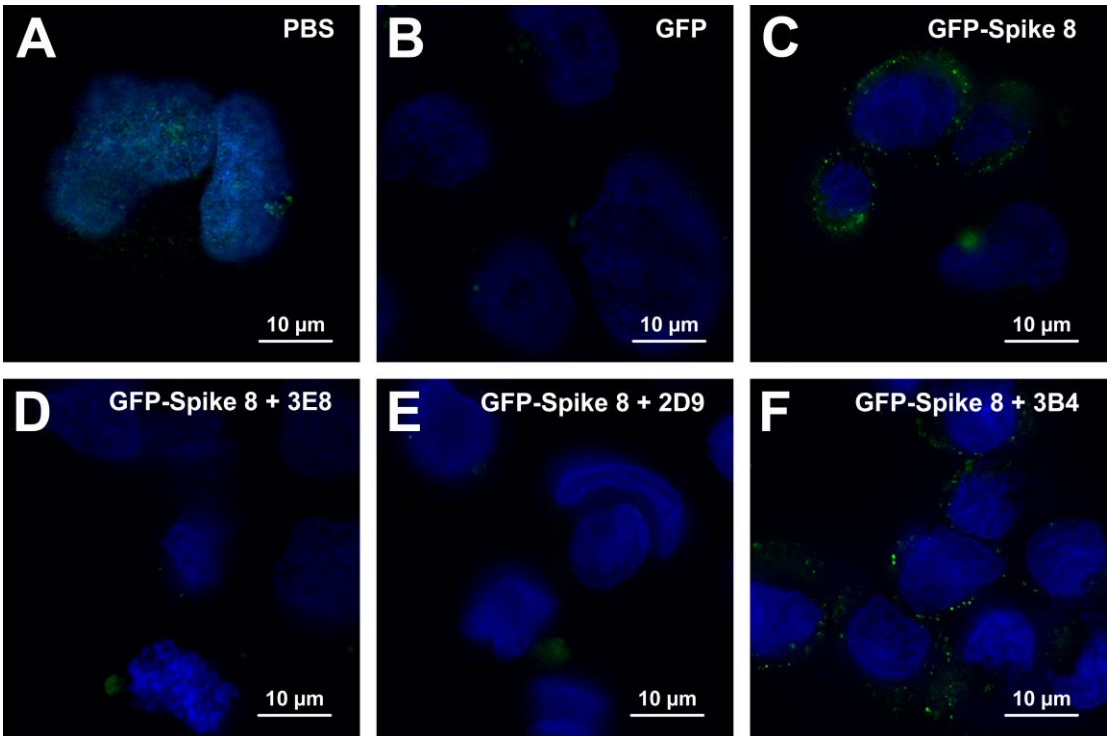

Figure 3.6: Antibodies 3E8 and 2D9 block GFP-Spike 8 attachment to Caco-2 cells.

GFP fluorescence is green and Hoechst DNA stain is blue. **(A)** Caco-2 cells incubated with PBS show slight autofluorescence. **(B)** Caco-2 cells incubated with GFP alone show little to no non-specific GFP binding. **(C)** Caco-2 cells incubated with GFP-Spike 8 show bright punctate patterns at or near the cell membranes, suggesting spike-specific attachment. **(D)** Fab 3E8 blocks GFP-Spike 8 attachment to Caco-2 cells. **(E)** Fab 2D9 blocks GFP-Spike 8 attachment to Caco-2 cells. **(F)** Fab 3B4 does not block GFP-Spike 8 attachment to Caco-2 cells.

3.6 Discussion

Our findings uncover several key insights into the biology of human astrovirus and the antibody response to the HAstV capsid antigen. First, we have shown that immunization with recombinant capsid spike elicits high-affinity antibodies that neutralize the virus, making the spike a candidate antigen for a subunit vaccine. We

have also discovered that there are at least two distinct epitopes on the HAstV capsid spike surface. Antibodies targeting either one of these epitopes neutralize the virus by inhibiting virus attachment to cells. There may be several possible explanations for this phenomenon. The receptor-binding site may be in between the two epitopes, such as in the proposed P site, and the binding of an antibody to either epitope sterically hinders viral access to the receptor. The receptor-binding site may also be very large and thus contain residues in both epitopes. Another possibility is that there may be two (or more) essential receptor-binding sites, and obstructing one site is sufficient to prevent virus attachment to cells. Alternatively, one or both antibodies may bind at an allosteric site, triggering a structural change in another region of the spike that disrupts access to the receptor. However, root-mean-square deviations between Spike 8 with and without antibodies are low at 0.241 Å when 3E8 is bound and 0.325 Å when 2D9 is bound, suggesting that the structural integrity of the spike is preserved upon antibody binding. Finally, the receptor-binding site may be in different places on spikes from different serotypes, but the observation that antibodies can target the spikes from two serotypes at the same epitope provides evidence against this argument. Efforts in our labs to discover the identity of the receptor and uncover the mechanism by which HAstV gains entry into cells are ongoing.

Secondly, the significant overlap between the epitopes targeted by PL-2 on Spike 2 and 3E8 on Spike 8 shows that antibodies can target the same epitope on spikes from two HAstV serotypes and neutralize the virus by the same mechanism, blocking attachment to cells. This epitope intersection indicates that functional regions on the

spike surface are likely conserved across HAstV serotypes, despite a high degree of sequence variability in the spike, which has 41 – 76 % sequence identity between all eight classical serotypes [3.7].

Finally, our structures of scFv 3E8/Spike 8 and scFv 2D9/Spike 8 illustrate that the immune response can target the spike of a single serotype at different epitopes. This targeting mechanism reveals that spikes used as immunogens elicit a neutralizing polyclonal response to human astrovirus. A polyclonal response is beneficial in generating long-term protection against infection because it minimizes virus neutralization escape by mutations. Indeed, the HAstV8 escape mutant to 3E8 is still susceptible to neutralization by 2D9, and vice versa [3.41]. Therefore, immunization with spikes presents a robust strategy for HAstV subunit vaccine immunogen design. Importantly, these results also indicate that obstructing any one of multiple sites on the spike surface can prevent virus infectivity. Consequently, HAstV has several weak points that may be exploited individually or in combination in an antiviral therapy.

In conclusion, this work has increased our understanding of the antigenic structure of the human astrovirus capsid spike, the basis of virus neutralization by antibodies targeting the spike, and the functional sites on spike relevant for virus attachment to the host cell, thus supporting the development of vaccines and therapies against astrovirus gastroenteritis.

3.7 Materials and Methods

3.7.1 Expression and purification of scFv 3E8

Mouse hybridoma cells producing mAb 3E8 were generated as reported in [3.41]. The amino acid sequences of the mAb 3E8 variable regions were identified as described in [3.46], allowing for recombinant antibody expression. A synthetic gene codon-optimized for *Drosophila melanogaster* containing the 3E8 kappa chain variable region connected to the 3E8 heavy chain variable region by a $GGS(GGGGS)_3$ linker was purchased from Integrated DNA Technologies. This gene was cloned into the pCMV-VRC01 vector by Gibson assembly in frame with an N-terminal secretion signal sequence and a C-terminal thrombin cleavage site followed by a Twin-Strep-tag. The resulting pCMV-VRC01_scFv_3E8 expression plasmid was used to electroporate Chinese Hamster Ovary suspension (CHO-S) cells using the MaxCyte system. The scFv 3E8 was expressed for 8 days by CHO-S cells growing in CD OptiCHO expression medium (Gibco) supplemented with 1 mM sodium butyrate (Sigma-Aldrich), 8 mM L-glutamine (Gibco), 1x HT supplement (Gibco), and 0.1 % Pluronic F68 (Gibco) at 32 °C with 125 rpm shaking. Every 24 hours, cells were fed with CHO CD EfficientFeed A (Gibco) supplemented with 7 mM L-glutamine (Gibco), 5.5 % glucose (Sigma Aldrich), and 23.4 g/L yeastolate (BD). After 8 days, cells were pelleted and medium containing secreted scFv 3E8 was 0.22-μM filtered, buffered to Strep wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA), and supplemented with BioLock (IBA Lifesciences) to mask free biotin in the medium. The sample was 0.22-μM filtered again, affinity purified on two tandem 5 mL StrepTrap

HP columns (GE), and eluted in Strep elution buffer (Strep wash buffer with 2.5 mM desthiobiotin). The scFv 3E8 was dialyzed overnight into 10 mM Tris-HCl pH 7.5, 150 mM NaCl (TBS pH 7.5).

3.7.2 Expression and purification of scFv 2D9

Mouse hybridoma cells producing mAb 2D9 were generated as reported in [3.41]. The amino acid sequences of the mAb 2D9 variable regions were identified as described in [3.46], allowing for recombinant antibody expression. A synthetic gene codon-optimized for *Drosophila melanogaster* containing the 2D9 kappa chain variable region connected to the 2D9 heavy chain variable region by a $GGS(GGGGS)_3$ linker and flanked by BglII and NheI restriction sites was purchased from Integrated DNA Technologies. This gene was cloned into the pMT_puro_BiP vector by restriction digest in frame with an N-terminal BiP secretion signal sequence and a C-terminal thrombin cleavage site followed by a Twin-Strep-tag. The plasmid contains a metallothionein promoter for induction of gene expression as well as a puromycin-resistance gene. The resulting pMT_puro_BiP_scFv_2D9 expression plasmid was used to obtain stably transfected *D. melanogaster* Schneider 2 (S2) cells by transfection with FuGENE HD (Promega) followed by selection with 5 μg/mL puromycin. The S2 cells were grown in Schneider's S2 medium (Gibco) with 10 % heat-inactivated FBS and 1x pen/strep (Gibco) at the selection stage. Cells were then adapted to serum-free, antibiotic-free ESF 921 medium (Expression Systems) for expression. The stable S2 cells were grown in shaker flasks to 3.0×10^6 cells/mL. Expression of scFv 2D9 was

induced with 500 μM $CuCl_2$ and cells were incubated at 27 °C with 125 rpm shaking. After 5 days, cells were pelleted and medium containing secreted scFv 2D9 was 0.22-μM filtered, buffered to Strep wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA) and concentrated 200-fold by tangential flow filtration. After supplementation with BioLock (IBA Lifesciences) to mask free biotin in the medium, the sample was 0.22-μM filtered again, affinity purified on two tandem 5 mL StrepTrap HP columns (GE), and eluted in Strep elution buffer (Strep wash buffer with 2.5 mM desthiobiotin). The scFv 2D9 was dialyzed overnight into 10 mM Tris-HCl pH 8.0, 150 mM NaCl (TBS pH 8.0).

3.7.3 Expression and purification of the Spike 8 antigen

A synthetic gene codon-optimized for *E. coli* expression encoding the HAstV serotype 8 capsid spike protein amino acids 429 to 647 (Spike 8, UniProtKB entry Q9IFX1) was purchased from Integrated DNA Technologies. To make the Spike 8 expression plasmid, the gene was cloned into pET52b (Addgene) in-frame with a C-terminal thrombin cleavage site and a 10-histidine purification tag. The plasmid was verified by DNA sequencing. Next, the plasmid was transformed into *E. coli* strain BL21(DE3). Cultures were inoculated and grown in LB/ampicillin medium. At an optical density of 0.6, protein production was induced with 1 mM isopropyl-D-thiogalactopyranoside (IPTG) at 18 °C for 18 hours. *E. coli* cells were lysed by ultrasonication in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, and 20 mM imidazole (Buffer A) containing 2 mM $MgCl_2$, 0.0125 U/μL benzonase (Merck Millipore), and

1x protease inhibitor cocktail set V EDTA-free (Merck Millipore). The protein was batch purified from soluble lysates with TALON metal affinity resin (GE Healthcare) and eluted with Buffer A containing 500 mM imidazole. The protein was dialyzed overnight into 10 mM Tris-HCl pH 8.0, 150 mM NaCl (TBS pH 8.0) and further purified by size exclusion chromatography on a Superdex 75 column in TBS.

3.7.4 Biolayer interferometry affinity determination of scFv 2D9 and scFv 3E8 for Spike 8

His-tagged Spike 8 was diluted to 0.5 μg/mL in BLI blocking buffer (PBS pH 7.4, 2 % bovine serum albumin (BSA), 0.09 % Tween-20). The scFv 3E8 was diluted to 80 nM (2.15 μg/mL) in BLI blocking buffer and five serial 1:2 dilutions were prepared. The scFv 2D9 was diluted to 20 nM (0.592 μg/mL) in BLI blocking buffer and five serial 1:2 dilutions were prepared. Using the 8-channel setting on an Octet® RED384 instrument (FortéBio), pre-equilibrated Anti-Penta-His (HIS1K) sensor tips were dipped into the following solutions at 22 °C with shaking at 1000 rpm: First Baseline: BLI blocking buffer for 60 seconds, Loading: 0.5 μg/mL Spike 8 for 180 seconds, Second Baseline: BLI blocking buffer for 60 seconds, Association: 6 different concentrations of the scFv analyte in a serial 1:2 dilution for 60 seconds (scFv 3E8) or 120 seconds (scFv 2D9), and Dissociation: BLI blocking buffer for 120 seconds (scFv 3E8) or 180 seconds (scFv 2D9). A reference sample control was included in which 0 nM scFv was tested. A reference sensor control was also included in which 80 nM scFv 3E8 or 20 nM scFv 2D9 was associated to sensor tips that had not been pre-loaded

with Spike 8 to test for non-specific scFv binding to the sensors. The sensors were dipped into BLI blocking buffer in these controls.

Each affinity determination experiment was performed in triplicate as independent assays. Data were processed separately and fit using the Octet® Data Analysis software v.7 (FortéBio). Before fitting, all datasets were reference-subtracted, aligned to the baseline, and aligned for inter-step correction through their respective dissociation steps as per the manufacturer's instructions. For each experiment, six different scFv analyte concentrations were used to fit association and dissociation globally using a 1:1 binding model. Ultimately, the goodness of fit was determined using R^2 and χ^2 values according to the manufacturer's guidelines. The reported K_D, R^2, and χ^2 values were averaged manually from the triplicate assays, generating the reported standard deviation.

For the BLI competition assay, His-tagged Spike 8 was diluted to 0.5 μg/mL in BLI blocking buffer. The scFv 2D9 was diluted to 25 nM (0.74 μg/mL) in BLI blocking buffer. The scFv 3E8 was diluted to 400 nM (12 μg/mL) in BLI blocking buffer and used to prepare two serial 1:2 dilutions followed by one serial 1:4 dilution to obtain concentrations of 400, 200, 100, and 25 nM scFv 3E8. 0 nM scFv 3E8 was also tested. Pre-equilibrated Anti-Penta-His (HIS1K) sensor tips were dipped into the following solutions at 22 °C with shaking at 1000 rpm: First Baseline: BLI blocking buffer for 60 seconds, Loading: 0.5 μg/mL Spike 8 for 180 seconds, Second Baseline: BLI blocking buffer for 60 seconds, scFv 2D9 Association: 25 nM scFv 2D9 for 120 seconds, scFv 3E8 Association: 5 different concentrations of scFv 3E8 for 120 seconds, and

Dissociation: BLI blocking buffer for 180 seconds. A reference sensor control was included in which 25 nM scFv 2D9 followed by 400 nM scFv 3E8 were associated to sensor tips that had not been pre-loaded with Spike 8 to test for non-specific scFv binding to the sensors. The sensors were dipped into BLI blocking buffer in this control. The dataset was reference-subtracted and aligned to the scFv 3E8 association step to generate Figure 3.1C.

3.7.5 Formation and structure determination of the scFv 3E8/Spike 8 complex

Complex formation was performed by incubation of 3 molar excess scFv 3E8 with Spike 8 overnight at 4 °C in TBS pH 7.2. Simultaneously, the purification tags from both scFv 3E8 and Spike 8 were removed by digestion of the complex with thrombin protease. The scFv 3E8/Spike 8 complex was purified by size exclusion chromatography on a Superdex 200 column in TBS pH 7.2. The complex coeluted at an apparent molecular mass of ~105 kD compared to gel filtration standards, consistent with a 2:2 (scFv 3E8:Spike 8) complex in solution (data not shown). The purified complex was concentrated to 6.33 mg/mL. Hanging drops (2 μL) were formed by a 1:1 addition of concentrated protein complex and a well solution of 0.2 M lithium citrate tribasic and 18 % PEG 3350. Crystals were grown by hanging drop vapor diffusion at 22 °C. Crystals were transferred into a cryoprotectant solution of 0.2 M lithium citrate tribasic, 18.9 % PEG 3350, and 25 % glycerol and flash frozen in liquid nitrogen. Diffraction data from a single crystal were collected at cryogenic temperature at the Advanced Photon Source on beamline 23-ID-D using a wavelength of 1.033184 Å. The

data were processed with Mosflm (ccp4i) and scaled with Aimless (ccp4i). The structure was solved by molecular replacement using Phenix.

3.7.6 Formation and structure determination of the scFv 2D9/Spike 8 complex

Complex formation was performed by incubation of 3.5 molar excess scFv 2D9 with Spike 8 overnight at 4 °C in TBS pH 8.0. Simultaneously, the purification tags from both scFv 2D9 and Spike 8 were removed by digestion of the complex with thrombin protease. The scFv 2D9/Spike 8 complex was purified by size exclusion chromatography on a Superdex 200 column in TBS pH 8.0. The complex coeluted at an apparent molecular mass of ~105 kD compared to gel filtration standards, consistent with a 2:2 (scFv 2D9:Spike 8) complex in solution (data not shown). The purified complex was concentrated to 3 mg/mL. Hanging drops (2 μL) were formed by a 1:1 addition of concentrated protein complex and a well solution of 0.1 M ammonium acetate, 0.1 M Bis-Tris pH 5.5, and 16 % PEG 10,000. Crystals were grown by hanging drop vapor diffusion at 22 °C. Crystals were transferred into a cryoprotectant solution of 0.1 M ammonium acetate, 0.1 M Bis-Tris pH 5.5, 16.8 % PEG 10,000, and 25 % ethylene glycol and flash frozen in liquid nitrogen. Six diffraction datasets from a single crystal were collected at cryogenic temperature at the Advanced Light Source on Beamline 8.3.1 using a wavelength of 1.115830 Å. The six datasets were processed separately with XDS and scaled together with Aimless (ccp4i). The structure was solved by molecular replacement using Phenix.

3.7.7 HAstV8 infectivity neutralization assay

Mouse mAbs 3E8 and 2D9 were raised in mice, hybridomas were generated, and ascites fluids were collected as described in [3.41]. Serial 1:5 dilutions of the ascites fluids for 3E8 or 2D9 were pre-incubated with HAstV8 particles (MOI = 30) purified as reported previously [3.47] for 1 h at room temperature. To determine the titer of the non-neutralized infectious virus, confluent Caco-2 cell monolayers grown in 96-well plates were incubated with the virus-mAb complex for 1 h at 37 °C. After this time, the cells were washed twice with PBS and were further incubated for 16 h at 37 °C. The infected cells were detected by an immunoperoxidase focus-forming assay as described previously [3.48]. The remaining percent infectivity was determined by comparing the titer of infectious virus in the virus-mAb mix as compared to that obtained in the absence of antibodies, which was considered as 100 %. Each point was performed in duplicate.

3.7.8 HAstV8 attachment inhibition assay

Serial 1:5 dilutions of the ascites fluids for 3E8 or 2D9 were pre-incubated with HAstV8 particles (MOI = 30) for 1 h at room temperature. Caco-2 cell monolayers grown in 48-well plates were washed once with PBS, and blocking solution (1 % BSA in PBS) was added for 45 min at room temperature followed by a 15 min incubation on ice. The cells were then washed once with ice cold PBS and incubated with the virus-antibody complex for 1 h on ice. As a negative control, mAb 3B4 against HAstV1 was used. The unbound virus was washed three times with cold PBS, and the total RNA

was extracted with TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Viral RNA or cellular 18S RNA was reverse transcribed using MMLV reverse transcriptase (Invitrogen). RT-qPCR was performed with the pre-mixed reagent Real Q Plus Master Mix Green (Ampliqon), and the PCR reaction was carried out in an ABI Prism 7500 detection system (Applied Biosystems). The primers used to detect HAstV8 RNA were forward primer 5´ atgaattattttgatactgaggaagattacttggaa 3´ and reverse primer 5´ ctttcttgagaaatagataccaaagtacttcag 3´ (ORF 1b). For normalization, 18S ribosomal cellular RNA was amplified and quantified using forward primer 5´ cgaaagcatttgccaagaat 3´ and reverse primer 5´ gcatcgtttatggtcggaac 3´. The arithmetic means and standard deviations from six (3E8 and 2D9) or two (3B4) independent experiments performed in duplicate are shown.

3.7.9 HAstV8 detachment assay

Confluent Caco-2 cell monolayers in 48-well plates were blocked with 1 % BSA in PBS for 45 min at room temperature followed by a 15 min incubation on ice. Purified HAstV8 viral particles were added at an MOI of 30 and incubated for 1 h on ice to allow binding of the virus to the cell surface. The unbound virus was subsequently removed by washing three times with cold PBS. Serial 1:5 dilutions of the indicated ascites fluids of 3E8 and 2D9 were added to the cells and incubated for 1 h on ice. After this incubation, the antibody and detached virus were removed with cold PBS, and RNA extraction and RT-qPCR quantification were performed as described above. As negative control, mAb 3B4 against HAstV1 was used. The arithmetic means and

standard deviations from two (3E8 and 2D9) or one (3B4) independent experiments performed in triplicate are shown.

3.7.10 Expression and purification of chimeric Fabs 3E8, 2D9, and 3B4

Synthetic cDNA encoding the kappa and heavy chain variable regions of each antibody was cloned by Gibson assembly into the pCMV-VRC01 antibody vectors for light and heavy chains, in place of the variable regions of antibody VRC01, a human anti-HIV antibody targeting the gp120 protein [3.49]. For the Fab heavy chain, only the variable region followed by the constant heavy 1 region ending with residues "DKKVEPKSC" was included, followed by an AS linker, C-terminal thrombin cleavage site, and a Twin-Strep-tag. Sequences were in-frame with the N-terminal signal sequence. The resulting chimeric Fab expression plasmids, pCMV-Fab_kappa and pCMV-Fab_heavy_VH+CH1, where Fab is 3E8, 2D9, or 3B4, contain the variable regions from the original mouse antibodies and the constant regions from the human IgG1 antibody under the control of the human cytomegalovirus promoter. The plasmids were verified by DNA sequencing. The expression plasmids were used in a ratio of 3:2 kappa chain: heavy VH+CH1 chain to electroporate Chinese Hamster Ovary suspension (CHO-S) cells using the MaxCyte system. Recombinant chimeric Fabs were expressed for 9 days (Fabs 3E8 and 3B4) or 7 days (Fab 2D9) by CHO-S cells growing in CD OptiCHO expression medium supplemented with 1 mM sodium butyrate, 8 mM L-glutamine, 1x HT supplement, and 0.1 % Pluronic F68 at 32 °C with 125 rpm shaking. Every 24 h, cells were fed with CHO CD EfficientFeed A

supplemented with 7 mM L-glutamine, 5.5 % glucose, and 23.4 g/L yeastolate. After 9 days, cells were pelleted and medium containing secreted Fabs was 0.22-μM filtered, buffered to Strep wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA), and supplemented with BioLock (IBA Lifesciences) to mask free biotin in the medium. The samples were 0.22-μM filtered again, each affinity purified on two tandem 5 mL StrepTrap HP columns (GE), and eluted in Strep elution buffer (Strep wash buffer with 2.5 mM desthiobiotin). Fabs 3E8 and 3B4 were dialyzed overnight into 10 mM Tris-HCl pH 7.2, 150 mM NaCl (TBS pH 7.2). Fab 2D9 was dialyzed overnight into 1x PBS pH 7.4 (Sigma-Aldrich).

3.7.11 GFP-Spike 8 attachment inhibition assay by fluorescence microscopy

GFP-Spike 8 was expressed and purified as described previously [3.40]. 24-well plates containing fibronectin-treated glass coverslips were seeded with 100,000 Caco-2 cells per well and allowed to adhere overnight. 250 μL samples containing 400 nM GFP-Spike 8 were incubated with three molar excess of antigen-binding fragments (Fabs) of Spike 8-specific monoclonal antibodies 3E8 and 2D9 and the Fab of Spike 1-specific monoclonal antibody 3B4 for 1 h at room temperature in Dulbecco's phosphate-buffered saline (DPBS) (Gibco). DPBS alone and 400 nM GFP samples were used as controls for autofluorescence and non-specific binding respectively. Media was aspirated from the cell monolayer and Spike-Fab mixtures were then added to Caco-2 cells and incubated at 4 °C for 1 h. Protein mixtures were removed, and cells were washed with DPBS and fixed with 2 % paraformaldehyde (ThermoFisher) in

DPBS for 15 min. Cells were washed with DPBS and then stained with Hoechst 33342 dye (ThermoFisher) in DPBS for 30 minutes. Coverslips were washed with PBS, dried, and mounted in Vectashield mounting media on glass microscope slides.

Z-stack images were acquired by using identical acquisition parameters with a Zeiss Axio Imager equipped with an AxioCam 506 monochrome camera using an oil-immersion 100X /1.4 n.a. plan apo objective lens. Z-stack images contained 9 slices at 0.24 µm intervals with GFP and Hoescht channels exposed for 1600 and 95 ms, respectively. GFP signal was collected with a Zeiss Fset38 filter cube and Hoechst signal was collected with a Zeiss Fset49 filter cube. After acquisition, images were deconvolved with AutoQuant X 3D deconvolution software (Media Cybernetics Version X3.1.3) for 10 iterations using the Z montage option. After deconvolution, a median filter with a 2-pixel kernel size was applied to the GFP channel of all images to reduce noise. Linear histogram adjustments were made to the GFP channel using FIJI [3.50] such that minimum values were 1300 and maximum values were 7500 to help reduce background fluorescence. Single Z stack slices from representative images were then converted to an RGB image. Images were cropped to identical sizes in FIJI to select representative cells.

3.7.12 Data availability

Coordinates and structure factors for the scFv 3E8/Spike 8 structure and the scFv 2D9/Spike 8 structure have been deposited in the Protein Data Bank under

accession codes 7RK1 and 7RK2, respectively. All other data has been made available in this manuscript.

3.8 Abbreviations

BLI: Biolayer interferometry; BSA: Bovine serum albumin; Caco-2 cells: Human colon adenocarcinoma cells; CDR: Complementarity-determining region; CHO cells: Chinese hamster ovary cells; Fab: Fragment antigen binding; GFP-Spike 8: Fusion protein construct of GFP and the human astrovirus serotype 8 capsid spike; HAstV: Human astrovirus; K_D: Dissociation constant; mAb: Monoclonal antibody; MOI: Multiplicity of infection; ORF: Open reading frame; PBS: Phosphate-buffered saline; S2 cells: *D. melanogaster* Schneider 2 cells; scFv: Single-chain variable fragment; Spike 2: Human astrovirus serotype 2 capsid spike; Spike 8: Human astrovirus serotype 8 capsid spike

3.9 References

3.1. Mendez, E.; Arias, C. F., Astroviruses. In *Fields Virology*, 2 ed.; Howley, P. M., Ed. Lippincott Williams & Wilkins: 2007; pp 982-1000.

3.2. Koci, M. D.; Schultz-Cherry, S., Avian astroviruses. *Avian Pathol* 2002, 31 (3), 213-27.

3.3. De Benedictis, P.; Schultz-Cherry, S.; Burnham, A.; Cattoli, G., Astrovirus infections in humans and animals - molecular biology, genetic diversity, and interspecies transmissions. *Infect Genet Evol* 2011, 11 (7), 1529-44.

3.4. King, A. M. Q.; Lefkowitz, E.; Adams, M. J.; Carstens, E. B., Virus taxonomy: classification and nomenclature of viruses. In *Ninth Report of the International Committee on Taxonomy of Viruses*, 9th edition ed.; Elsevier Inc: Philadelphia, PA, 2011.

3.5. Koopmans, M. P.; Bijen, M. H.; Monroe, S. S.; Vinje, J., Age-stratified seroprevalence of neutralizing antibodies to astrovirus types 1 to 7 in humans in The Netherlands. *Clin Diagn Lab Immunol* 1998, 5 (1), 33-7.

3.6. Kurtz, J. B.; Lee, T. W., Human astrovirus serotypes. *Lancet* 1984, 2 (8416), 1405.

3.7. Meyer, L.; Delgado-Cunningham, K.; Lorig-Roach, N.; Ford, J.; DuBois, R. M., Human Astrovirus 1-8 Seroprevalence Evaluation in a United States Adult Population. *Viruses* 2021, 13 (6).

3.8. Cortez, V.; Meliopoulos, V. A.; Karlsson, E. A.; Hargest, V.; Johnson, C.; Schultz-Cherry, S., Astrovirus Biology and Pathogenesis. *Annu Rev Virol* 2017, 4 (1), 327-348.

3.9. Walter, J. E.; Mitchell, D. K., Astrovirus infection in children. *Curr Opin Infect Dis* 2003, 16 (3), 247-53.

3.10. Goodgame, R. W., Viral causes of diarrhea. *Gastroenterol Clin North Am* 2001, 30 (3), 779-95.

3.11. Gray, J. J.; Wreghitt, T. G.; Cubitt, W. D.; Elliot, P. R., An outbreak of gastroenteritis in a home for the elderly associated with astrovirus type 1 and human calicivirus. *J Med Virol* 1987, 23 (4), 377-81.

3.12. Lewis, D. C.; Lightfoot, N. F.; Cubitt, W. D.; Wilson, S. A., Outbreaks of astrovirus type 1 and rotavirus gastroenteritis in a geriatric in-patient population. *J Hosp Infect* 1989, 14 (1), 9-14.

3.13. Marshall, J. A.; Bruggink, L. D.; Sturge, K.; Subasinghe, N.; Tan, A.; Hogg, G. G., Molecular features of astrovirus associated with a gastroenteritis outbreak in an aged-care centre. *Eur J Clin Microbiol Infect Dis* 2007, 26 (1), 67-71.

3.14. Dennehy, P. H.; Nelson, S. M.; Spangenberger, S.; Noel, J. S.; Monroe, S. S.; Glass, R. I., A prospective case-control study of the role of astrovirus in acute diarrhea among hospitalized young children. *J Infect Dis* 2001, 184 (1), 10-5.

3.15. Gallimore, C. I.; Taylor, C.; Gennery, A. R.; Cant, A. J.; Galloway, A.; Xerry, J.; Adigwe, J.; Gray, J. J., Contamination of the hospital environment with gastroenteric viruses: comparison of two pediatric wards over a winter season. *J Clin Microbiol* 2008, 46 (9), 3112-5.

3.16. Olortegui, M. P.; Rouhani, S.; Yori, P. P.; Salas, M. S.; Trigoso, D. R.; Mondal, D.; Bodhidatta, L.; Platts-Mills, J.; Samie, A.; Kabir, F.; Lima, A.; Babji, S.; Shrestha, S. K.; Mason, C. J.; Kalam, A.; Bessong, P.; Ahmed, T.; Mduma, E.; Bhutta, Z. A.; Lima, I.; Ramdass, R.; Moulton, L. H.; Lang, D.; George, A.; Zaidi, A. K. M.; Kang, G.; Houpt, E. R.; Kosek, M. N.; Network, M.-E., Astrovirus Infection and Diarrhea in 8 Countries. *Pediatrics* 2018, 141 (1).

3.17. Mead, P. S.; Slutsker, L.; Dietz, V.; McCaig, L. F.; Bresee, J. S.; Shapiro, C.; Griffin, P. M.; Tauxe, R. V., Food-related illness and death in the United States. *Emerg Infect Dis* 1999, 5 (5), 607-25.

3.18. Palombo, E. A.; Bishop, R. F., Annual incidence, serotype distribution, and genetic diversity of human astrovirus isolates from hospitalized children in Melbourne, Australia. *J Clin Microbiol* 1996, 34 (7), 1750-3.

3.19. Bosch, A.; Pinto, R. M.; Guix, S., Human astroviruses. *Clin Microbiol Rev* 2014, 27 (4), 1048-74.

3.20. Cortez, V.; Freiden, P.; Gu, Z.; Adderson, E.; Hayden, R.; Schultz-Cherry, S., Persistent Infections with Diverse Co-Circulating Astroviruses in Pediatric Oncology Patients, Memphis, Tennessee, USA. *Emerg Infect Dis* 2017, 23 (2), 288-290.

3.21. van der Doef, H. P.; Bathoorn, E.; van der Linden, M. P.; Wolfs, T. F.; Minderhoud, A. L.; Bierings, M. B.; Wensing, A. M.; Lindemans, C. A., Astrovirus outbreak at a pediatric hematology and hematopoietic stem cell transplant unit despite strict hygiene rules. *Bone Marrow Transplant* 2016, 51 (5), 747-50.

3.22. Brown, J. R.; Morfopoulou, S.; Hubb, J.; Emmett, W. A.; Ip, W.; Shah, D.; Brooks, T.; Paine, S. M.; Anderson, G.; Virasami, A.; Tong, C. Y.; Clark, D. A.; Plagnol, V.; Jacques, T. S.; Qasim, W.; Hubank, M.; Breuer, J., Astrovirus VA1/HMO-C: an increasingly recognized neurotropic pathogen in immunocompromised patients. *Clin Infect Dis* 2015, 60 (6), 881-8.

3.23. Naccache, S. N.; Peggs, K. S.; Mattes, F. M.; Phadke, R.; Garson, J. A.; Grant, P.; Samayoa, E.; Federman, S.; Miller, S.; Lunn, M. P.; Gant, V.; Chiu, C. Y., Diagnosis of neuroinvasive astrovirus infection in an immunocompromised adult with encephalitis by unbiased next-generation sequencing. *Clin Infect Dis* 2015, 60 (6), 919-23.

3.24. Quan, P. L.; Wagner, T. A.; Briese, T.; Torgerson, T. R.; Hornig, M.; Tashmukhamedova, A.; Firth, C.; Palacios, G.; Baisre-De-Leon, A.; Paddock, C. D.; Hutchison, S. K.; Egholm, M.; Zaki, S. R.; Goldman, J. E.; Ochs, H. D.; Lipkin, W. I., Astrovirus encephalitis in boy with X-linked agammaglobulinemia. *Emerg Infect Dis* 2010, 16 (6), 918-25.

3.25. Cordey, S.; Vu, D. L.; Schibler, M.; L'Huillier, A. G.; Brito, F.; Docquier, M.; Posfay-Barbe, K. M.; Petty, T. J.; Turin, L.; Zdobnov, E. M.; Kaiser, L., Astrovirus MLB2, a New Gastroenteric Virus Associated with Meningitis and Disseminated Infection. *Emerg Infect Dis* 2016, 22 (5), 846-53.

3.26. Kurtz, J.; Lee, T., Astrovirus gastroenteritis age distribution of antibody. *Med Microbiol Immunol* 1978, 166 (1-4), 227-30.

3.27. Kurtz, J. B.; Lee, T. W.; Craig, J. W.; Reed, S. E., Astrovirus infection in volunteers. *J Med Virol* 1979, 3 (3), 221-30.

3.28. Mitchell, D. K., Astrovirus gastroenteritis. *Pediatr Infect Dis* J 2002, 21 (11), 1067-9.

3.29. Bjorkholm, M.; Celsing, F.; Runarsson, G.; Waldenstrom, J., Successful intravenous immunoglobulin therapy for severe and persistent astrovirus gastroenteritis after fludarabine treatment in a patient with Waldenstrom's macroglobulinemia. *Int J Hematol* 1995, 62 (2), 117-20.

3.30. Lewis, T. L.; Greenberg, H. B.; Herrmann, J. E.; Smith, L. S.; Matsui, S. M., Analysis of astrovirus serotype 1 RNA, identification of the viral RNA-dependent RNA polymerase motif, and expression of a viral structural protein. *J Virol* 1994, 68 (1), 77-83.

3.31. Jiang, B.; Monroe, S. S.; Koonin, E. V.; Stine, S. E.; Glass, R. I., RNA sequence of astrovirus: distinctive genomic organization and a putative retrovirus-like ribosomal frameshifting signal that directs the viral replicase synthesis. *Proc Natl Acad Sci U S A* 1993, 90 (22), 10539-43.

3.32. Krishna, N. K., Identification of structural domains involved in astrovirus capsid biology. *Viral Immunol* 2005, 18 (1), 17-26.

3.33. Dryden, K. A.; Tihova, M.; Nowotny, N.; Matsui, S. M.; Mendez, E.; Yeager, M., Immature and Mature Human Astrovirus: Structure, Conformational Changes, and Similarities to Hepatitis E Virus. *J Mol Biol* 2012.

3.34. Mendez, E.; Fernandez-Luna, T.; Lopez, S.; Mendez-Toss, M.; Arias, C. F., Proteolytic processing of a serotype 8 human astrovirus ORF2 polyprotein. *J Virol* 2002, 76 (16), 7996-8002.

3.35. Mendez, E.; Salas-Ocampo, E.; Arias, C. F., Caspases mediate processing of the capsid precursor and cell release of human astroviruses. *J Virol* 2004, 78 (16), 8601-8.

3.36. Bass, D. M.; Qiu, S., Proteolytic processing of the astrovirus capsid. *J Virol* 2000, 74 (4), 1810-4.

3.37. York, R. L.; Yousefi, P. A.; Bogdanoff, W.; Haile, S.; Tripathi, S.; DuBois, R. M., Structural, Mechanistic, and Antigenic Characterization of the Human Astrovirus Capsid. *J Virol* 2015, 90 (5), 2254-63.

3.38. Toh, Y.; Harper, J.; Dryden, K. A.; Yeager, M.; Arias, C. F.; Méndez, E.; Tao, Y. J., Crystal Structure of the Human Astrovirus Capsid Protein. *J Virol* 2016, 90 (20), 9008-17.

3.39. Dong, J.; Dong, L.; Mendez, E.; Tao, Y., Crystal structure of the human astrovirus capsid spike. *Proceedings of the National Academy of Sciences of the United States of America* 2011, 108 (31), 12681-6.

3.40. Bogdanoff, W. A.; Campos, J.; Perez, E. I.; Yin, L.; Alexander, D. L.; DuBois, R. M., Structure of a Human Astrovirus Capsid-Antibody Complex and Mechanistic Insights into Virus Neutralization. *J Virol* 2017, 91 (2).

3.41. Espinosa, R.; López, T.; Bogdanoff, W. A.; Espinoza, M. A.; López, S.; DuBois, R. M.; Arias, C. F., Isolation of neutralizing monoclonal antibodies to human astrovirus and characterization of virus variants that escape neutralization. *J Virol* 2018.

3.42. Sanchez-Fauquier, A.; Carrascosa, A. L.; Carrascosa, J. L.; Otero, A.; Glass, R. I.; Lopez, J. A.; San Martin, C.; Melero, J. A., Characterization of a human astrovirus serotype 2 structural protein (VP26) that contains an epitope involved in virus neutralization. *Virology* 1994, 201 (2), 312-20.

3.43. Bass, D. M.; Upadhyayula, U., Characterization of human serotype 1 astrovirus-neutralizing epitopes. *J Virol* 1997, 71 (11), 8666-71.

3.44. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011, 7, 539.

3.45. Robert, X.; Gouet, P., Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 2014, 42, Issue W1, W320-4.

3.46. Meyer, L.; López, T.; Espinosa, R.; Arias, C. F.; Vollmers, C.; DuBois, R. M., A simplified workflow for monoclonal antibody sequencing. *PLoS One* 2019, 14 (6), e0218717.

3.47. Aguilar-Hernández, N.; López, S.; Arias, C. F., Minimal capsid composition of infectious human astrovirus. *Virology* 2018, 521, 58-61.

3.48. Bogdanoff, W. A.; Perez, E. I.; López, T.; Arias, C. F.; DuBois, R. M., Structural Basis for Escape of Human Astrovirus from Antibody Neutralization: Broad Implications for Rational Vaccine Design. *J Virol* 2018, 92 (1).

3.49. Wu, X.; Yang, Z. Y.; Li, Y.; Hogerkorp, C. M.; Schief, W. R.; Seaman, M. S.; Zhou, T.; Schmidt, S. D.; Wu, L.; Xu, L.; Longo, N. S.; McKee, K.; O'Dell, S.; Louder, M. K.; Wycuff, D. L.; Feng, Y.; Nason, M.; Doria-Rose, N.; Connors, M.; Kwong, P. D.; Roederer, M.; Wyatt, R. T.; Nabel, G. J.; Mascola, J. R., Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 2010, 329 (5993), 856-61.

3.50. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; Tinevez, J. Y.; White, D. J.; Hartenstein, V.; Eliceiri, K.; Tomancak, P.; Cardona, A., Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012, 9 (7), 676-82.

Appendix 1: Crystallization Attempts of 3B4/Spike 1 Complexes

A1.1 Introduction

3B4 is a mouse monoclonal antibody targeting the human astrovirus serotype 1 spike. It is part of the Arias lab's panel of five novel antibodies and is predicted to bind the top of the spike based on the location of escape mutant S560P in the middle of the long Loop 3. Size exclusion chromatography traces of complexes with scFv and with chimeric Fab possibly indicate that only one 3B4 binds the Spike 1 dimer, in contrast to the three known antibody/spike complexes PL-2/Spike 2, 2D9/Spike 8, and 3E8/Spike 8, in which two antibodies bind the spike dimer.

Starting in December 2018 and continuing through October 2020, I attempted crystallization of the following complexes in this order:

- scFv 3B4/Spike 1, in which S2-expressed scFv 3B4 2x Strep had been thrombin-digested to remove the 2x Strep tag
- chimeric Fab 3B4/Spike 1, in which chimeric Fab 3B4 had been generated from CHO-S-expressed chimeric mAb 3B4 by a papain digest
- chimeric Fab 3B4 Strep/Spike 1, in which CHO-S-expressed chimeric Fab 3B4 Strep expressed well but was difficult to purify because the Strep tag was too close to the Fab
- chimeric Fab 3B4/Spike 1, in which CHO-S-expressed chimeric Fab 3B4 2x Strep had been thrombin-digested to remove the 2x Strep tag

A1.2 Crystallization attempts of scFv 3B4/Spike 1

I spent most of the optimization effort on this project using the scFv 3B4 construct because it was the only one that resulted in real crystals at all when complexed with Spike 1. A representative sizing run and gel for scFv 3B4/Spike 1 Prep 5 from 7/16/19 – 7/17/19 is below. Gels of all scFv 3B4/Spike 1 preparations were run on 12/14/18 (Prep 1), 4/18/19 (Prep 2), 5/8/19 (Prep 3), 5/31/19 (Prep 4), 7/17/19 (Prep 5), and 7/25/19 (Prep 6), with the sizing runs usually occurring the day before. It appears that scFv 3B4 and Spike 1 each run as single discreet bands. After an overnight incubation at 4 °C of Spike 1 with 2.2 molar excess scFv 3B4 and thrombin at 10 U/mg total protein, the scFv 3B4/Spike 1 Prep 5 complex eluted from 74 – 84 mL with a peak at 79 mL on the S200 16/600 column with a large excess scFv 3B4 peak directly following (Figure A1.1A). Based on the complex elution volume and the observation that only 2.2 molar excess scFv 3B4 still led to a large excess scFv 3B4 peak, it seems that only one scFv 3B4 may bind the Spike 1 dimer. scFv 3E8/Spike 8, in which two scFv 3E8 molecules bind the Spike 8 dimer as confirmed by crystallography, eluted from 68 - 82 mL with the peak at 74 mL on the S200 16/600 (1/10/21).

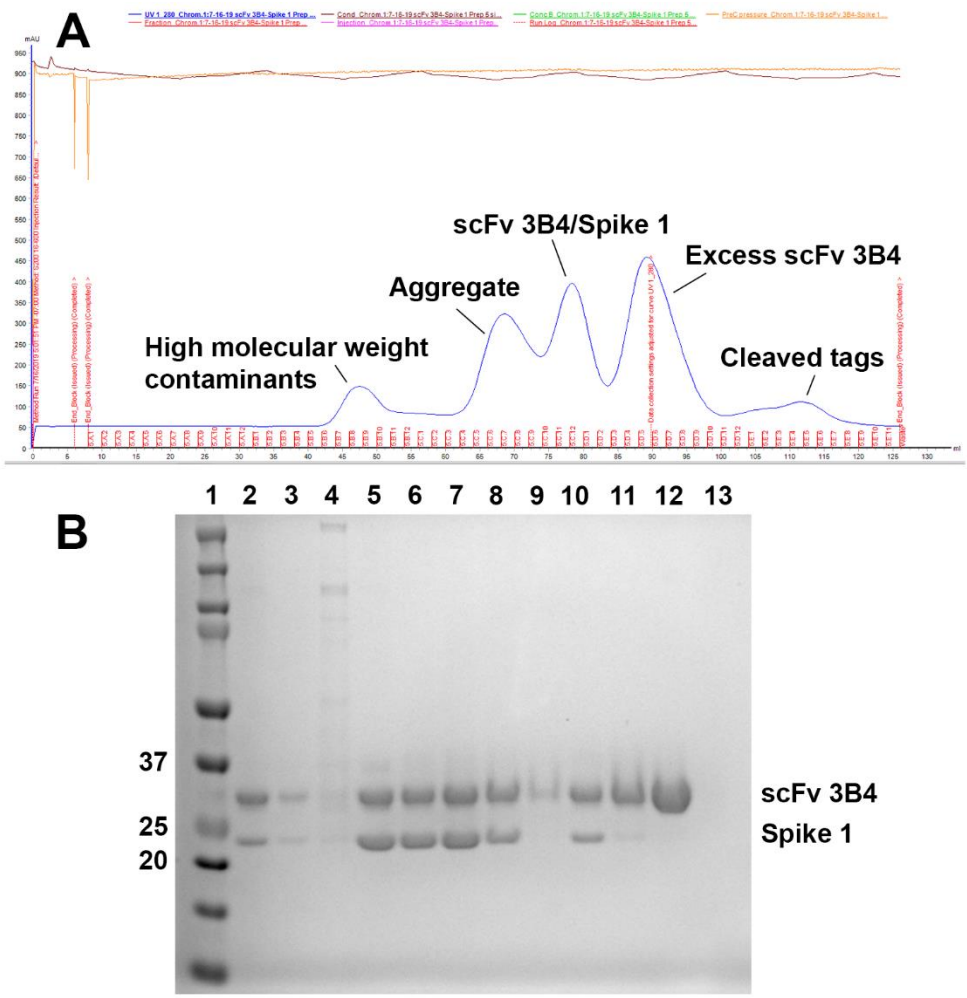

Figure A1.1: scFv 3B4/Spike 1 Prep 5 complex preparation from 7/16/19 – 7/17/19.

(A) Size exclusion chromatography of scFv 3B4/Spike 1 performed 7/16/19. **(B)** Reducing SDS-PAGE gel run 7/17/19 on the size exclusion chromatography of scFv 3B4/Spike 1. Order from left to right: (1) Molecular weight ladder (kD), (2) scFv 3B4/Spike 1 pre-sizing, (3) scFv 3B4/Spike 1 10 kD concentration flow-through, fractions (4) B8, (5) C6, (6) C10, (7) C11, (8) C12, (9) D1, (10) D2, (11) D3, (12) D6, and (13) E5.

Preps 1 – 5 of scFv 3B4/Spike 1 were used in crystallization attempts. I moved on to a chimeric Fab construct before I needed to use Prep 6. Crystallization attempts are summarized in the bullet-pointed list below. Information in this list includes date

set up, protein complex prep and concentration, screens or conditions optimized, and results in an indented bullet point below the setup description.

- 12/17/2018, 1.12 mg/mL, scFv 3B4/Spike 1 Prep 1, all conditions of an MCSG1 screen
    - No hits
- 4/19/2019, 5.06 mg/mL, scFv 3B4/Spike 1 Prep 2, all conditions of an MCSG1 screen
    - No great hits but tried optimizing a few
- 4/23/2019, 5.06 mg/mL, scFv 3B4/Spike 1 Prep 2, setup of Plate 1 A1-B3 with optimizations of 4/19/19 MCSG1 E12, F6, and H9
    - No crystals
- 4/26/2019, 2.78 mg/mL, scFv 3B4/Spike 1 Prep 2, all conditions of an MCSG2 screen
    - No hits
- 5/13/2019, 8.11 mg/mL, scFv 3B4/Spike 1 Prep 3, all conditions of an MCSG2 screen
    - Fluffy protein in C10? Tried optimizing
- 5/17/2019, 8.11 mg/mL, scFv 3B4/Spike 1 Prep 3, setup of Plate 1 B4-B6 with optimization of 5/13/19 MCSG2 C10
    - No crystals
- 5/20/2019, 8.11 mg/mL, scFv 3B4/Spike 1 Prep 3, setup of Plate 1 C1-C3 with optimization of 4/23/19 Plate 1 A1 (MCSG1 E12 optimization)

    - Dumbbell-shaped crystals appeared in C2 after 1 month, which I used to seed C3
- 5/31/2019, 5.09 mg/mL, scFv 3B4/Spike 1 Prep 4, all conditions of an MCSG3 screen
    - No hits
- 7/17/2019, 17.72 mg/mL, scFv 3B4/Spike 1 Prep 5, all conditions of MCSG1 – 4 screens, but some of the MCSG4 conditions were running low
    - No hits
- 7/25/2019, 17.72 mg/mL, scFv 3B4/Spike 1 Prep 5, all conditions of Grid Screen Salt HT and Index HT screens
    - No great hits but set up two to optimize anyway
- 7/29/2019, 17.72 mg/mL, scFv 3B4/Spike 1 Prep 5, setup of Plate 1 C4-C5 with optimization of 7/25/19 Grid Screen Salt HT E7 and F6
    - No crystals
- 7/30/2019, 3 – 18 mg/mL, scFv 3B4/Spike 1 Prep 5 dilutions, setup of Plate 1 D1-D6 with a protein concentration screen of Plate 1 C2 condition
    - No crystals but a range from clear to precipitated
- 7/31/2019, 7.82 mg/mL, scFv 3B4/Spike 1 Prep 5 dilutions, setup of Plate 2 A1-A4 and Plate 1 C3 by varying condition of Plate 1 C2
    - No crystals but a range from clear to precipitated
- 7/31/2019, Seeded from Plate 1 C2 into Plate 1 C3

    - Tiny crystals in C3 after seeding which did not diffract after growing for 1 month (September 2019) or 4 months (12/3/19)
- 8/7/2019, Seeded from Plate 1 C2 into Plate 1 C3 again, as well as Plate 1 D5-D6 and Plate 2 A1-A2
    - No crystals, broke Plate 2 A1-A2
- 8/14/2019, 5 – 10 mg/mL, scFv 3B4/Spike 1 Prep 5 dilutions, created a matrix on Plate 2 B1-D6 and Plate 3 A1-C6 with sodium malonate pH 7 concentration vs Prep 5 concentration
    - No crystals until about a year later (August 2020)
- 8/20/2019, Seeded from Plate 1 C2 into Plate 2 C4-D4 and Plate 3 A4, B3, and C3
    - No crystals
- 8/28/2019, Seeded from Plate 1 C2 into Plate 1 C4, Plate 2 B1-B3, C1-C3, and D3, and Plate 3 A3 and B2
    - No crystals, broke Plate 3 B2

As shown from this summary, I made many attempts to crystallize scFv 3B4/Spike 1. The only condition which produced convincing crystals contained 1.5 – 2.0 M sodium malonate pH 7.0 (MCSG1 E12). Dumbbell-shaped crystal clusters first appeared on Plate 1 C2 (1.8 M sodium malonate pH 7.0) in July 2019 about 1-2 months after setup (Figure A1.2A). However, these crystals were hard to reproduce. A first attempt to seed into Plate 1 C3 (1.5 M sodium malonate pH 7.0) seemed successful, but then the crystals disappeared. A second attempt led to a shower of very small but

singular crystals (Figure A1.2B). Unfortunately, these crystals did not diffract after growing for one month (September 2019) or four months (December 2019). Izit dye staining suggested these crystals were real protein crystals.

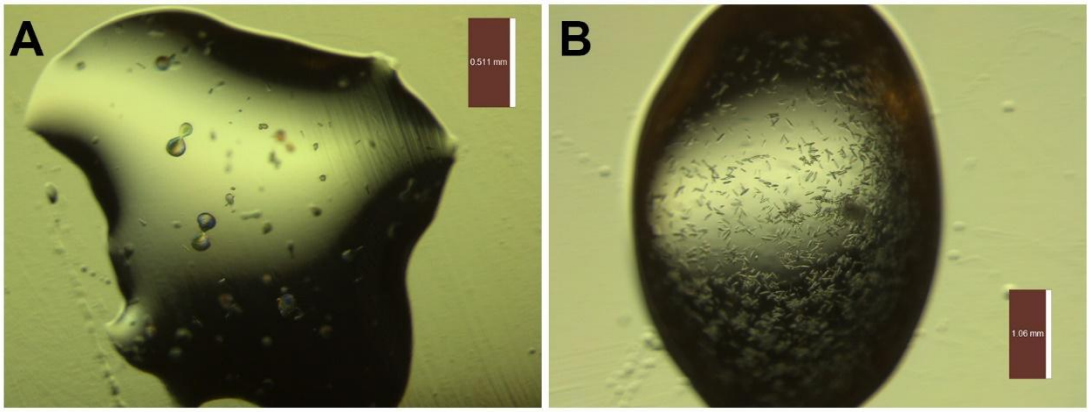

Figure A1.2: scFv 3B4/Spike 1 crystals in sodium malonate pH 7.0.

(A) 7/29/19 scFv 3B4/Spike 1 Plate 1 C2. **(B)** 9/6/19 scFv 3B4/Spike 1 Plate 1 C3.

In August 2020, about a year after setting up the matrix varying concentration of sodium malonate pH 7.0 from 1.5 – 2.0 M and varying concentration of scFv 3B4/Spike 1 from 5 – 10 mg/mL on Plates 2 and 3, I noticed that similar dumbbell-shaped clusters had formed in some of these drops, especially in drops with lower concentrations of sodium malonate pH 7.0 and higher concentrations of protein (Figure A1.3). I had checked the drops frequently in the first few months after setup, but had not checked them in at least six months. I ran a neighboring drop (Plate 3 C4) to check for protein degradation, but both scFv 3B4 and Spike 1 still ran at the expected size. I did not pursue these crystals further because I did not have any freshly purified scFv 3B4/Spike 1 complex and because this same condition resulted in crystals that did not

diffract. Instead, I focused on preparation of complexes with chimeric Fab 3B4 constructs.

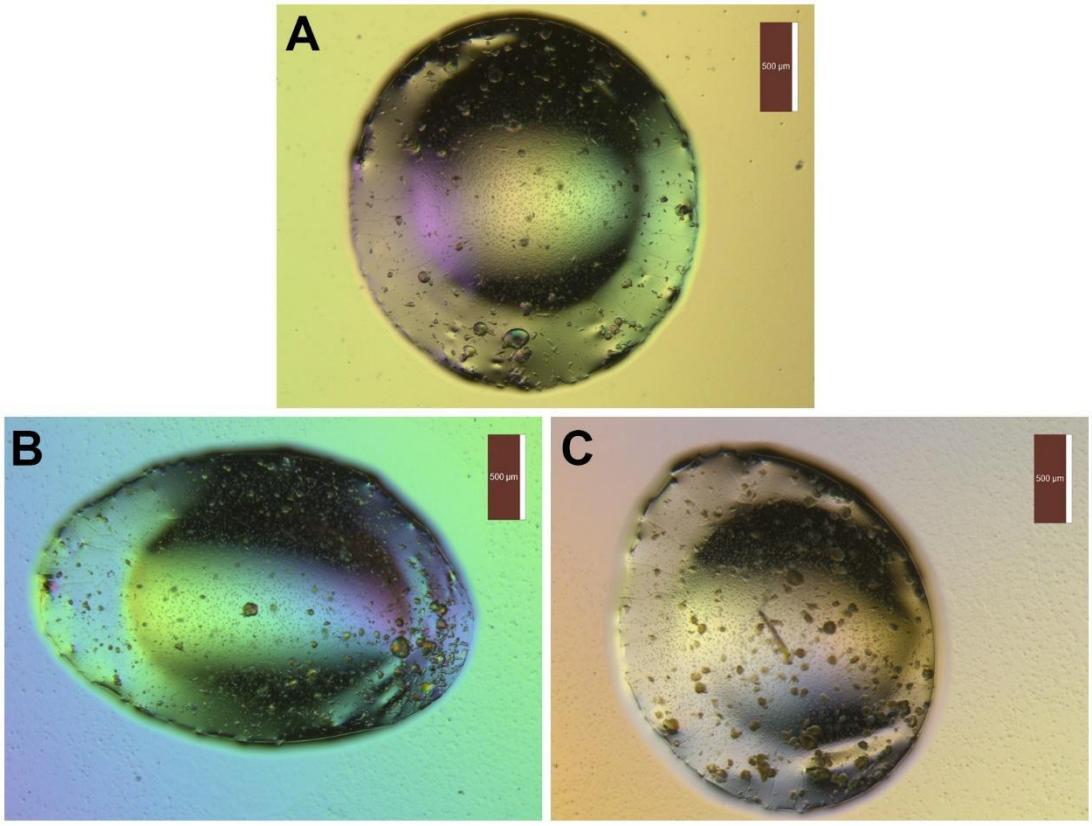

Figure A1.3: scFv 3B4/Spike 1 crystals in sodium malonate pH 7.0 one year after setup.

(A) 8/11/20 scFv 3B4/Spike 1 Plate 3 A1. **(B)** 8/11/20 scFv 3B4/Spike 1 Plate 3 B1. **(C)** 8/11/20 scFv 3B4/Spike 1 Plate 3 C1.

A1.3 Crystallization attempts of papain-digested chimeric Fab 3B4/Spike 1

Following scFv 3B4/Spike 1 crystallization attempts, I moved on to chimeric Fab 3B4/Spike 1 constructs. First, I expressed full-length chimeric mAb 3B4 in CHO-S cells and digested with papain to obtain chimeric Fab 3B4. This process was very

inefficient despite optimization attempts and very long digestion times. I usually obtained ~1 mg of chimeric Fab from ~5 mg starting mAb. In addition, I always saw a secondary, slightly smaller heavy chain band in both the full-length mAb and Fab forms on non-reducing gels, which I did not see with the scFv construct. This secondary band also bound Spike 1, and I could not purify it away. It also showed up in all the other CHO-S-expressed mAb and Fab constructs. Both bands appeared on an anti-His Western for chimeric Fab 3B4 His (with the missing last cysteine) and on an anti-Strep Western for chimeric Fab 3B4 Strep. This heterogeneity likely caused the lack of crystals obtained using any of the chimeric mAb 3B4 or chimeric Fab 3B4 constructs.

There was possibly no second band when I expressed chimeric mAb 3B4 in HEK 293 cells, which I only did once transiently. I say possibly because the gel may also have been overloaded, making it hard to distinguish two bands if present. That gel was run 1/16/18 (Notebook 2 page 103). Expression in HEK 293 cells may be an avenue worth pursuing if the mAb and Fab are partially cleaved by some kind of CHO-S cell protease but not by HEK 293 cells.

I cut off the Spike 1 His tag with thrombin and formed the chimeric Fab 3B4/Spike 1 complex overnight by incubating 2-3x molar excess chimeric Fab 3B4 with Spike 1. The Fab/Spike ratio again seemed to be 1:2 based elution volume and on the excess chimeric Fab 3B4 peak after complex sizing when only incubating with 2x molar excess Fab. Below is the 10/31/19 size exclusion chromatography run of chimeric Fab 3B4/Spike 1 after incubating with 3x molar excess as well as reducing and non-reducing gels (Figure A1.4).

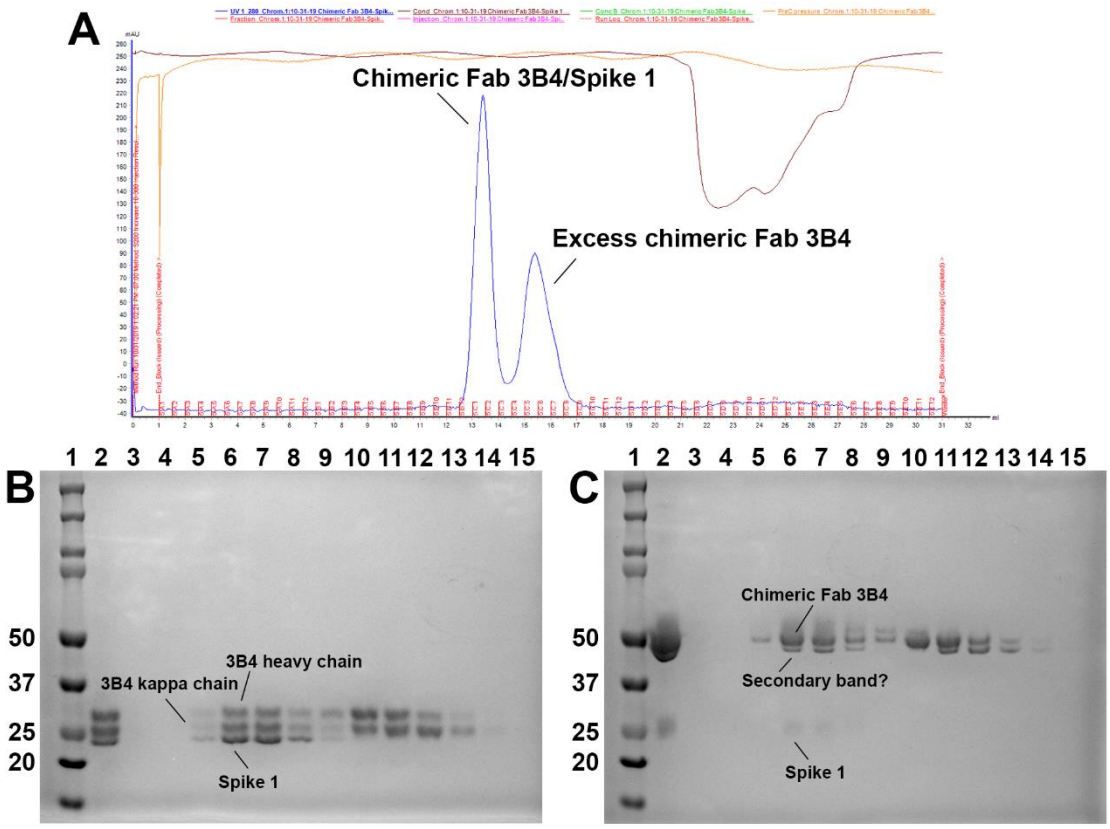

Figure A1.4: Chimeric Fab 3B4/Spike 1 Prep 1 complex preparation from 10/30/19 – 11/1/19.

(A) Size exclusion chromatography of chimeric Fab 3B4/Spike 1 Prep 1 performed 10/31/19. **(B)** Reducing and **(C)** non-reducing SDS-PAGE gels run 11/1/19 on the size exclusion chromatography of chimeric Fab 3B4/Spike 1. Order from left to right: (1) Molecular weight ladder (kD), (2) chimeric Fab 3B4/Spike 1 pre-sizing, (3) chimeric Fab 3B4/Spike 1 10 kD concentration flow-through, fractions (4) B11, (5) B12, (6) C1, (7) C2, (8) C3, (9) C4, (10) C5, (11) C6, (12) C7, (13) C8, (14) C9, (15) C10.

I set up MCSG1 – MCSG4, Grid Screen Salt HT, and Index HT screens with the purified complex (using Prep 1 at 6.68 mg/mL for MCSG1 and Prep 2 at 7.72 mg/mL for the rest), but no real crystals formed. We sent a few blobs to the synchrotron anyway, but they did not diffract (12/3/19).

A1.4 Crystallization attempts of chimeric Fab 3B4 Strep/Spike 1

Because of the inefficiency of chimeric Fab 3B4 preparation, I moved on to expressing Fabs directly in CHO-S cells. First, I tried a chimeric Fab 3B4 construct with a non-cleavable His tag. It expressed poorly and could not be purified because I had truncated the C-terminus of the heavy chain construct before the final fifth cysteine, which is needed to hold the heavy and kappa chains together.

Next, I redesigned the chimeric Fab constructs to include the missing cysteine and one non-cleavable Strep tag instead of His tag in case the EDTA in the CHO-S cell media interferes with His-tag purifications. Chimeric Fab 3B4 Strep expressed well but was difficult to purify – most of the protein would flow through the column and end up in the column load flow-through instead of binding because the tag was too close to the end of the protein. I did manage to purify enough to form the complex with Spike 1 and purify by size exclusion chromatography on 7/31/20. Chimeric Fab 3B4 Strep/Spike 1 elutes ~1 mL after chimeric Fab 3E8 Strep/Spike 8 on the S200 10/300, further evidence for the 1:2 chimeric Fab 3B4 Strep:Spike 1 ratio. I set up an MCSG1 screen of chimeric Fab 3B4 Strep/Spike 1 at 3.83 mg/mL on 8/7/20 which did not result in any real hits. I moved on to a 2x Strep-tagged construct.

A1.5 Crystallization attempts of thrombin-digested chimeric Fab 3B4/Spike 1

Last, I tried a chimeric Fab 3B4 2x Strep construct, for which the 2x Strep tag is cleavable, and which expressed and purified well. I digested with thrombin, formed the complex with Spike 1 using a 2x molar excess chimeric Fab 3B4 2x Strep, and ran

size exclusion chromatography on 10/14/20 (Figure A1.5). This time, digested excess chimeric Fab 3B4 eluted as two overlapping peaks, likely corresponding to the two double bands always seen with CHO-S-expressed chimeric mAb 3B4 and chimeric Fab 3B4 constructs. I had not seen these poorly resolved peaks before, but instead a single peak; maybe I had not had enough sample before to see two peaks. Below is the chromatogram of the size exclusion chromatography run as well as the non-reducing gel of sizing fractions.

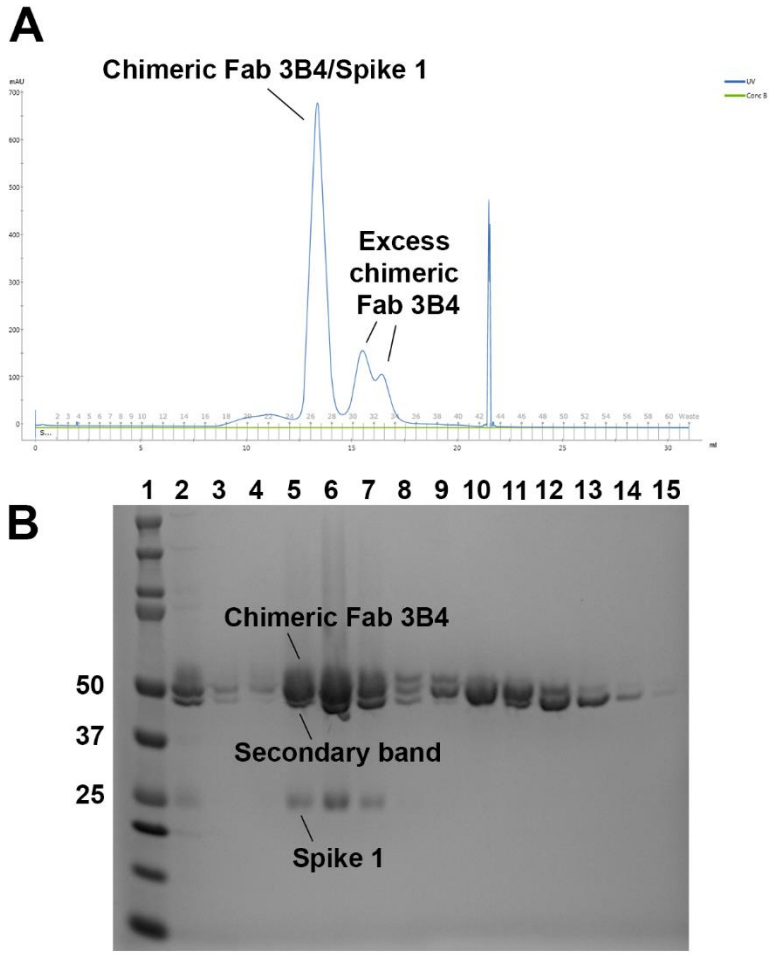

Figure A1.5: Chimeric Fab 3B4/Spike 1 Prep 1 complex preparation from 10/11/20 – 10/15/20.

(A) Size exclusion chromatography of chimeric Fab 3B4/Spike 1 performed 10/14/20. **(B)** Non-reducing SDS-PAGE gel run 10/15/20 on the size exclusion chromatography of chimeric Fab 3B4/Spike 1. Order from left to right: (1) Molecular weight ladder (kD), (2) chimeric Fab 3B4/Spike 1 pre-sizing, (3) chimeric Fab 3B4/Spike 1 10 kD concentration flow-through, fractions (4) 24, (5) 25, (6) 26, (7) 27, (8) 28, (9) 29, (10) 30, (11) 31, (12) 32, (13) 33, (14) 34, (15) 35

I concentrated the purified complex to 6.10 mg/mL and set up MCSG1 – MCSG4 screens. No real crystal hits were identified and I had already been working

on this project for almost two years, so I decided to focus on obtaining a structure for 3E8/Spike 8 instead.

Further attempts to crystallize 3B4/Spike 1 could involve a massive complex of 3B4/3H4/Spike 1, since 3H4 is predicted to bind a separate epitope from 3B4 based on the Arias lab's escape mutation studies. Additionally, it could be confirmed whether chimeric mAb 3B4 and/or chimeric Fab 3B4 2x Strep expressed in HEK 293 cells do not have the cleaved secondary band, which would result in a more homogeneous sample and maybe lead to crystal formation in complex with Spike 1.

Appendix 2: Variable Region Sequencing of Anti-VA1 Antibodies 2A2 and 7C8

A2.1 Introduction

The Arias lab sent us RNA for two new antibodies, 2A2 and 7C8, which target human astrovirus VA1. In their email, they told us they would send the extracted RNA from the two hybridomas precipitated in isopropanol and an aliquot of 300 μL (diluted 1:2 in glycerol) of each antibody. However, when I looked in the bag of samples, I could not tell from the labeling which tubes contained RNA and which contained purified antibody. I spun all four tubes down and none resulted in a pellet. I then arbitrarily labeled them 2A2 Tube 1 or Tube 2 and 7C8 Tube 1 or Tube 2 to keep track of which I used in RT-PCR, since I would expect amplification off the sample in the tube containing RNA but not the tube containing purified antibody.

A2.2 RT-PCR on 2A2 and 7C8

I attempted amplification of kappa, lambda, and heavy chains using template RNA from 2A2 Tube 1 and 7C8 Tube 1. 2A2 kappa, 7C8 kappa, and 7C8 heavy amplified, but not 2A2 heavy (Figure A2.1).

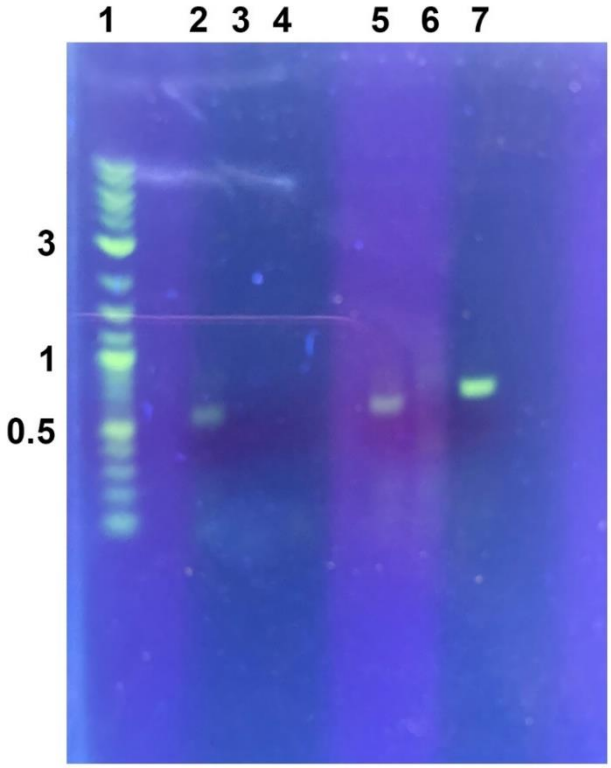

Figure A2.1: RT-PCR of kappa, lambda, and heavy chains using 2A2 Tube 1 and 7C8 Tube 1.

Order from left to right: (1) kb ladder, (2) 2A2 kappa, (3) 2A2 lambda, (4) 2A2 heavy, (5) 7C8 kappa, (6) 7C8 lambda, (7) 7C8 heavy.

Next, I amplified kappa and heavy chains using template RNA from 2A2 Tube 2 and also used 2A2 Tube 1 as a control. 2A2 kappa amplified from the RNA from both tubes, but 2A2 heavy only amplified from 2A2 Tube 2 (Figure A2.2).

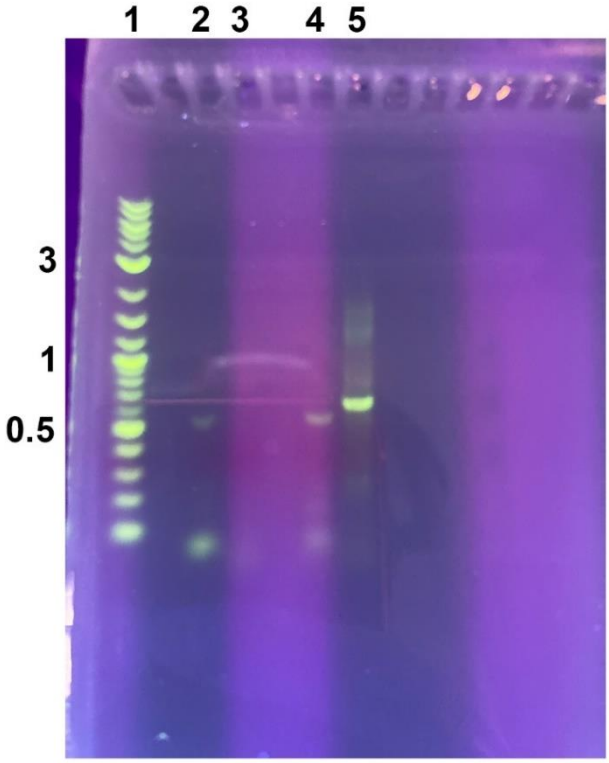

Figure A2.2: RT-PCR of kappa and heavy chains using 2A2 Tube 1 and Tube 2.

Order from left to right: (1) kb ladder, (2) 2A2 kappa Tube 1, (3) 2A2 heavy Tube 1, (4) 2A2 kappa Tube 2, (5) 2A2 heavy Tube 2.

A2.3 Sequencing results of 2A2 and 7C8

I blunt-end cloned and sequenced colonies from 2A2 kappa Tube 1, 2A2 kappa Tube 2, 2A2 heavy Tube 2, 7C8 kappa Tube 1, and 7C8 kappa Tube 2. Sequence alignments for miniprepped colonies are followed by the consensus sequences for each antibody chain.

```
lm0685-M13_fwd_A10.ab1_frame_-2_translation      VPTQVLGLLLLWLTDARCDIQMTQSPASLSVSVGETVTITCRASENIYSNLAWYQQKRGK   201
lm0686-M13_fwd_B10.ab1_frame_-1_translation      VPTQVLGLLLLWLTDARCDIQMTQSPASLSVSVGETVTITCRASENIYSNLAWYQQKRGK   207
lm0689-M13_fwd_E10.ab1_frame_-3_translation      VPTQVLGLLLLWLTDARCDIQMTQSPASLSVSVGETVTITCRASENIYSNLAWYQQKRGK   208
lm0721-M13_fwd_C12.ab1_frame_1_translation       VPTQVLGLLLLWLTDARCDIQMTQSPASLSVSVGETVTITCRASENIYSNLAWYQQKRGK   102
                                                 ************************************************************

lm0685-M13_fwd_A10.ab1_frame_-2_translation      SPQLLVHAARDLATGVPSRFSGSGSGTQYSLKINSLQSEDFGTYYCQHFWETPWTFGGGT   261
lm0686-M13_fwd_B10.ab1_frame_-1_translation      SPQLLVHAARDLATGVPSRFSGSGSGTQYSLKINSLQSEDFGTYYCQHFWETPWTFGGGT   267
lm0689-M13_fwd_E10.ab1_frame_-3_translation      SPQLLVHAARDLATGVPSRFSGSGSGTQYSLKINSLQSEDFGTYYCQHFWETPWTFGGGT   268
lm0721-M13_fwd_C12.ab1_frame_1_translation       SPQLLVHAARDLATGVPSRFSGSGSGTQYSLKINSLQSEDFGTYYCQHFWETPWTFGGGT   162
                                                 ************************************************************

lm0685-M13_fwd_A10.ab1_frame_-2_translation      KLEIKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVRANSADI-HHTGGRSS   320
lm0686-M13_fwd_B10.ab1_frame_-1_translation      KLEIKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVRANSADI-HHTGGRSS   326
lm0689-M13_fwd_E10.ab1_frame_-3_translation      KLEIKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINEGRILQISITLAAARAC   328
lm0721-M13_fwd_C12.ab1_frame_1_translation       KLEIKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVRANSSTL-AAVTSGSE   221
                                                 *******************************************          . . :
```

Figure A2.3: Sequence alignment of 2A2 kappa from Tube 1 (first three sequences) and Tube 2 (last sequence).

```
lm0714-M13_fwd_B06.ab1_frame_2_translation       LVLILKGVQCDVQLVESGGGLVQPGGSRKLSCAASGFTFSYFGMHWVRQAPKKGLEWVAY   133
lm0723-M13_fwd_E12.ab1_frame_2_translation       LVLILKGVQCDVQLVESGGGLVQPGGSRKLSCAASGFTFSYFGMHWVRQAPXKGLEWVAY   133
lm0738-M13_fwd_F11.ab1_frame_-3_translation      LVLILKGVQCDVQLVESGGGLVQPGGSRKLSCAASGFTFSYFGMHWVRQAPEKGLEWVAY   174
lm0739-M13_fwd_C08.ab1_frame_-3_translation      LVLILKGVQCDVQLVESGGGLVQPGGSRKLSCAASGFTFSYFGMHWVRQAPEKGLEWVAY   203
lm0743-M13_fwd_G08.ab1_frame_-1_translation      LVLILKGVQCDVQLVESGGGLVQPGGSRKLSCAASGFTFSYFGMHWVRQAPEKGLEWVAY   196
lm0744-M13_fwd_H08.ab1_frame_-2_translation      LVLILKGVQCDLQLVESGGGLVQPGGSRKLSCAASGFTFSYFGMHWVRQAPEKGLEWVAY   193
                                                 ***********:************************************************ ********

lm0714-M13_fwd_B06.ab1_frame_2_translation       ISSGSNTIYYADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCXRAYYGNHYYAMDF   193
lm0723-M13_fwd_E12.ab1_frame_2_translation       ISSGSNTIYYADTVKGRFTISRDNPKNTLFLQMTSLXSEDTAMYYCXRAYYGNHYYAMDF   193
lm0738-M13_fwd_F11.ab1_frame_-3_translation      ISSGSNTIYYADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARAYYGNHYYAMDF   234
lm0739-M13_fwd_C08.ab1_frame_-3_translation      ISSGSNTIYYADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARAYYGNHYYAMDF   263
lm0743-M13_fwd_G08.ab1_frame_-1_translation      ISSGSNTIYYADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARAYYGNHYYAMDF   256
lm0744-M13_fwd_H08.ab1_frame_-2_translation      ISSGSNTIYYADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARAYYGNHYYAMDF   253
                                                 ************************************ ********* *************

lm0714-M13_fwd_B06.ab1_frame_2_translation       WGPGTSVTVSSAKTTAPSVYPLAPVCGDTTGSSVTLGCLVKGYFPEPVTLTXNSGSQGRI   253
lm0723-M13_fwd_E12.ab1_frame_2_translation       WGPXTSVTVSSAKTTAPSVYPLXXVCXDTTGSSVTLGCLVKGYFPEPVTLTXNSGSXXRI   253
lm0738-M13_fwd_F11.ab1_frame_-3_translation      WGPGTSVTVSSAKTTAPSVYPLAPVCGDTTGSSVTLGCLVKGYFPEPVTLTWNSGSQGRI   294
lm0739-M13_fwd_C08.ab1_frame_-3_translation      WGPGTSVTVSSAKTTAPSVYPLAPVCGDTTGSSVTLGCLVKGYFPEPVTLTWNSGSQGRI   323
lm0743-M13_fwd_G08.ab1_frame_-1_translation      WGPGTSVTVSSAKTTAPSVYPLAPVCGDTTGSSVTLGCLVKGYFPEPVTLTWNSG-QGRI   315
lm0744-M13_fwd_H08.ab1_frame_-2_translation      WGPGTSVTVSSAKTTAPSVYPLAPVCGDTTGSSVTLGCLVKGYFPEPVTLTWNSGSQGRI   313
                                                 *** ******************  ** ************************ ***   **
```

Figure A2.4: Sequence alignment of 2A2 heavy from Tube 2.

```
lm0696-M13_fwd_D11.ab1_frame_-2_translation      LLISASVMMSRGQIVLSQSPAILSASPGEKVTMTCRASSSVSYMYWYQQKPGSSPKPWIY   207
lm0697-M13_fwd_E11.ab1_frame_3_translation       LLISASVMMSRGQIVLSQSPAILSASPGEKVTMTCRASSSVSYMYWYQQKPGSSPKPWIY   115
lm0717-M13_fwd_E06.ab1_frame_-3_translation      LLISASVMMSRGQIVLSQSPAILSASPGEKVTMTCRASSSVSYMYWYQQKPGSSPKPWIY   207
lm0718-M13_fwd_F06.ab1_frame_3_translation       LLISASVMMSRGQIVLSQSPAILSASPGEKVTMTCRASSSVSYMYWYQQKPGSSPKPWIY   121
                                                 ************************************************************

lm0696-M13_fwd_D11.ab1_frame_-2_translation      ATSNLASGVPARFSGSGSGTSYSLTISRVEAEDAATYYCQQWSSDPPMTFGGGTKLEIKR   267
lm0697-M13_fwd_E11.ab1_frame_3_translation       ATSNLAPGVPARFSGSGSGTSYSLTISRVEAEDAATYYCQQWSSDPPMTFGGGTKLEIKR   175
lm0717-M13_fwd_E06.ab1_frame_-3_translation      ATSNLASGVPARFSGSGSGTSYSLTISRVEAEDAATYYCQQWSSDPPMTFGGGTKLEIKR   267
lm0718-M13_fwd_F06.ab1_frame_3_translation       ATSNLASGVPARFSGSGSGTSYSLTISRVEAEDAATYYCQQWSSDPPMTFGGGTKLEIKR   181
                                                 ****** ********************************************************
```

Figure A2.5: Sequence alignment of 7C8 kappa from Tube 1.

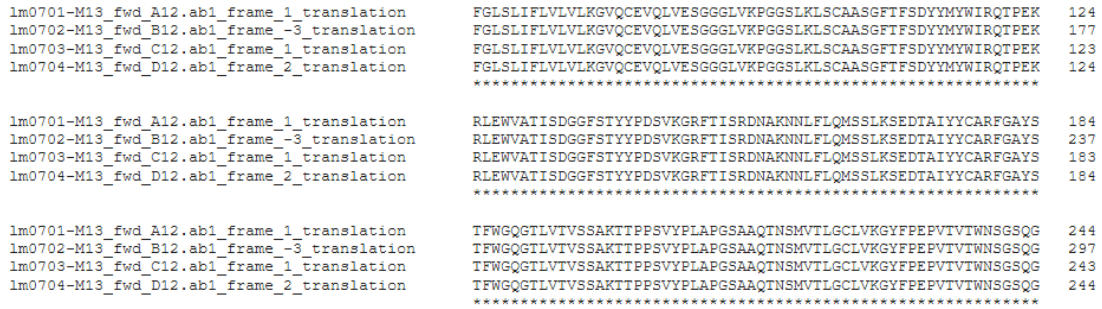

Figure A2.6: Sequence alignment of 7C8 heavy from Tube 1.

Consensus sequence for 2A2 kappa:

DIQMTQSPASLSVSVGETVTITCRASENIYSNLAWYQQKRGKSPQLLVHAAR

DLATGVPSRFSGSGSGTQYSLKINSLQSEDFGTYYCQHFWETPWTFGGGTKL

EIKR

Consensus sequence for 2A2 heavy:

DVQLVESGGGLVQPGGSRKLSCAASGFTFSYFGMHWVRQAPEKGLEWVAYI

SSGSNTIYYADTVKGRFTISRDNPKNTLFLQMTSLRSEDTAMYYCARAYYGN

HYYAMDFWGPGTSVTVSS

Consensus sequence for 7C8 kappa:

QIVLSQSPAILSASPGEKVTMTCRASSSVSYMYWYQQKPGSSPKPWIYATSNL

ASGVPARFSGSGSGTSYSLTISRVEAEDAATYYCQQWSSDPPMTFGGGTKLEI

KR

Consensus sequence for 7C8 heavy:

EVQLVESGGGLVKPGGSLKLSCAASGFTFSDYYMYWIRQTPEKRLEWVATIS

DGGFSTYYPDSVKGRFTISRDNAKNNLFLQMSSLKSEDTAIYYCARFGAYSTF

WGQGTLVTVSS