OPEN ACCESS Freely available online

PLOS ONE

Evaluation and Use of *In-Silico* Structure-Based Epitope Prediction with Foot-and-Mouth Disease Virus

Daryl W. Borley[1,2]*, Mana Mahapatra[1], David J. Paton[1], Robert M. Esnouf[2], David I. Stuart[2,3], Elizabeth E. Fry[2]*

1 The Pirbright Institute, Pirbright, United Kingdom, **2** Division of Structural Biology, University of Oxford, The Henry Wellcome Building for Genomic Medicine, Headington, Oxford, United Kingdom, **3** Diamond Light Source, Harwell Science and Innovation Campus, Didcot, United Kingdom

Abstract

Understanding virus antigenicity is of fundamental importance for the development of better, more cross-reactive vaccines. However, as far as we are aware, no systematic work has yet been conducted using the 3D structure of a virus to identify novel epitopes. Therefore we have extended several existing structural prediction algorithms to build a method for identifying epitopes on the appropriate outer surface of intact virus capsids (which are structurally different from globular proteins in both shape and arrangement of multiple repeated elements) and applied it here as a proof of principle concept to the capsid of foot-and-mouth disease virus (FMDV). We have analysed how reliably several freely available structure-based B cell epitope prediction programs can identify already known viral epitopes of FMDV in the context of the viral capsid. To do this we constructed a simple objective metric to measure the sensitivity and discrimination of such algorithms. After optimising the parameters for five methods using an independent training set we used this measure to evaluate the methods. Individually any one algorithm performed rather poorly (three performing better than the other two) suggesting that there may be value in developing virus-specific software. Taking a very conservative approach requiring a consensus between all three top methods predicts a number of previously described antigenic residues as potential epitopes on more than one serotype of FMDV, consistent with experimental results. The consensus results identified novel residues as potential epitopes on more than one serotype. These include residues 190–192 of VP2 (not previously determined to be antigenic), residues 69–71 and 193–197 of VP3 spanning the pentamer-pentamer interface, and another region incorporating residues 83, 84 and 169–174 of VP1 (all only previously experimentally defined on serotype A). The computer programs needed to create a semi-automated procedure for carrying out this epitope prediction method are presented.

Citation: Borley DW, Mahapatra M, Paton DJ, Esnouf RM, Stuart DI, et al. (2013) Evaluation and Use of *In-Silico* Structure-Based Epitope Prediction with Foot-and-Mouth Disease Virus. PLoS ONE 8(5): e61122. doi:10.1371/journal.pone.0061122

Editor: Gajendra P. S. Raghava, CSIR-Institute of Microbial Technology, India

Received August 22, 2012; **Accepted** March 6, 2013; **Published** May 7, 2013

Funding: The work performed at the Institute for Animal Health is supported by the Biotechnology and Biological Sciences Research Council UK (Grant number BB/G529532/1). The work performed at Oxford is supported by the Medical Research Council and by Wellcome Trust Core Award Grant Number 090532/Z/09/Z. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: daryl.borley@pirbright.ac.uk (DB); liz@strubi.ox.ac.uk (EF)

Introduction

Foot-and-mouth disease (FMD) is an economically important disease that predominantly affects cloven-hoofed mammals, with the primary hosts being cattle, sheep, pigs and goats. The disease is currently endemic in Africa, South America and Asia. Recent outbreaks of the disease in the Far East [1] and Eastern Europe [2] demonstrate the ability of FMD to disseminate into areas previously free from disease with major economic impact. Although FMD does not have a high mortality rate in adult animals, it reduces the productivity of infected herds. It also seriously damages the economies of enzootic countries by impeding exports of livestock and livestock products.

Foot-and-mouth disease viruses (FMDV) are single stranded, positive sense RNA viruses belonging to the family *Picornaviridae* and are currently classified into 7 serotypes: A, O, C, SAT (South African territories) 1–3 and Asia-1. These serotypes share an approximate 86% amino acid identity to each other [3], however some of the capsid proteins exhibit more variation, notably VP1 which varies by 30–50% between serotypes [4]. This variation has impeded the development of vaccines that can provide cross protection both inter and intra-serotypically [5]. A vaccine providing broader antigenic coverage would be a valuable tool for the control of FMDV.

FMDV is a small icosahedral virus 30 nm in diameter, comprising 60 copies of each of the 4 structural proteins VP1-4 (see Figure 1(a)). VP1, 2 and 3 constitute the surface of the virus and are composed of 8 anti-parallel β strands linked by loops to form a β barrel. VP4 is much smaller, internal and has little secondary structure [6]. The highly mobile VP1 G-H loop protrudes from the surface of the virus and contains the arginine-glycine-aspartic acid (RGD) motif responsible for attachment to host integrins [7]. This loop is also antigenic on all serotypes of FMDV. Due to its disorganised (flexible) nature this loop is absent from many of the crystallographic structures reported, although it has been visualised when stabilised by chemical reduction [8] or bound to monoclonal antibody [9].

The role of antibody as the principal component of the immune response to FMDV is well established [10] and several immunodominant neutralising antigenic sites have been described on the surface exposed proteins, VP1-3. However, it has been shown that immunoglobulin specific protection can be achieved in

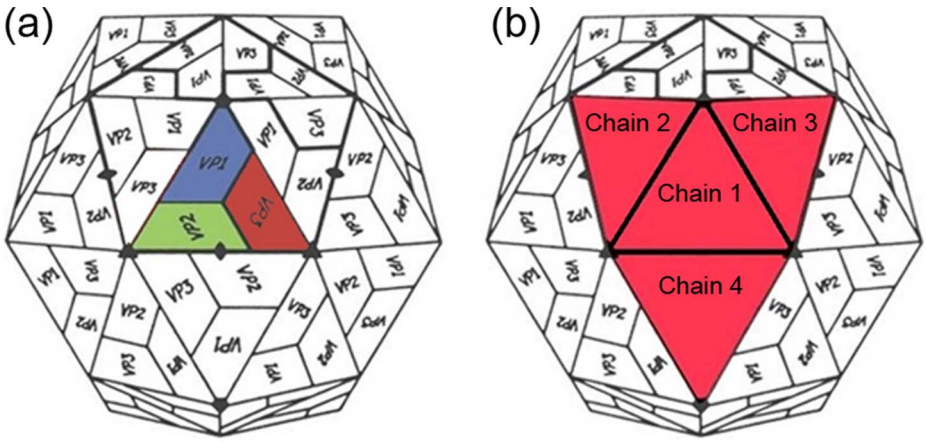

Figure 1. Building the multimer. (a) Schematic depiction of the foot and mouth disease icosahedral capsid with the individual virion peptides labelled. A viral subunit is coloured as per the standard colouring sytem; VP1 is blue, VP2 is in green and VP3 is red. (b) Schematic depiction of the of the same structure as (a) with the four subunits selected to make the multimer labelled as chains 1–4 (highlighted in red). These four protomers form the multimeric structure that was used in the analysis.
doi:10.1371/journal.pone.0061122.g001

the absence of a response to these epitopes [11], demonstrating the presence of other as yet unidentified epitopes. Traditional methods for identifying epitopes have relied on the generation of neutralising monoclonal antibody (mAb) escape mutants or peptide scanning techniques. These methods are time consuming, expensive and have generally utilised murine mAb's. Novel methods have recently been developed to correlate amino acid changes on the capsid and serological changes for multiple virus pairs [12,13]. Again, these methods require a large dataset and rely on amino acid variation within the capsid sequences to identify epitopes, thereby eliminating any possibility of identifying epitopes that are completely conserved. Such epitopes are particularly interesting as a potential route to more broadly reactive vaccines [14].

Many computational algorithms that try to predict B cell epitopes have been based on the analysis of sequence data alone. The first method described [15] used hydrophillicity scales of the amino acids averaged across the sequence tested. Subsequently, other methods have been described that utilize various sequence scoring algorithms [16,17]. An evaluation of the predictive power of these programs suggested that none could reliably determine epitopes [18], which is perhaps not a surprise as these programs would have no way of determining solvent exposure. However, it has been suggested that using a consensus of the results from B cell epitope prediction programs may increase confidence in the predictions made [19]. Additionally, these programs are limited to predicting continuous epitopes, which are believed to comprise only 10% of the total number of available epitopes [20].

To identify discontinuous epitopes the three dimensional structure must be included in the analysis [20] since this determines which areas of the sequence come into close enough proximity to form a discontinuous epitope. Recently, several web based servers (see methods section for selected list) have become available that utilize both structural and sequence information to identify conformational epitopes on the surface of a protein. As these servers use a single set of structural and sequence data they can, in principle, detect epitopes that are completely conserved in sequence both within, and between serotypes. The algorithms assign each residue a score based on the likelihood of it forming part of a conformational epitope. The performance of these algorithms has been tested using databases of known epitopes from available structural data (*i.e.* the Epitome database- http://www.rostlab.org/services/epitome/). For additional information on the individual programs see references [21–25]. Although all of these programs have reported higher levels of success in prediction of epitopes than those that relied on sequence information alone, there has been no work conducted looking specifically at the performance of these programs using 3D structure information that is available for virus capsids. In order to fully understand virus antigenicity it must be examined at the level of the capsid so that interactions between the repeated protomeric subunits are accounted for.

Table 1. Sensitivity and specificity results for polio and rhinovirus at the optimum threshold value for scoring a residue an epitope (the default given by the developers is also shown for comparison).

	Default server	Optimal	Poliovirus		Rhinovirus	
Program	**setting**	**Threshold value**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**
Discotope	−7.7	−10.7	0.868	0.746	0.591	0.686
Seppa	1.8	1.75	0.66	0.585	0.545	0.654
Epitopia	Not given	0.174	0.528	0.788	0.727	0.766
BEPro	1.3	0.8	0.717	0.767	0.682	0.612
Ellipro	0.5	0.3	0.811	0.648	0.864	0.553

doi:10.1371/journal.pone.0061122.t001

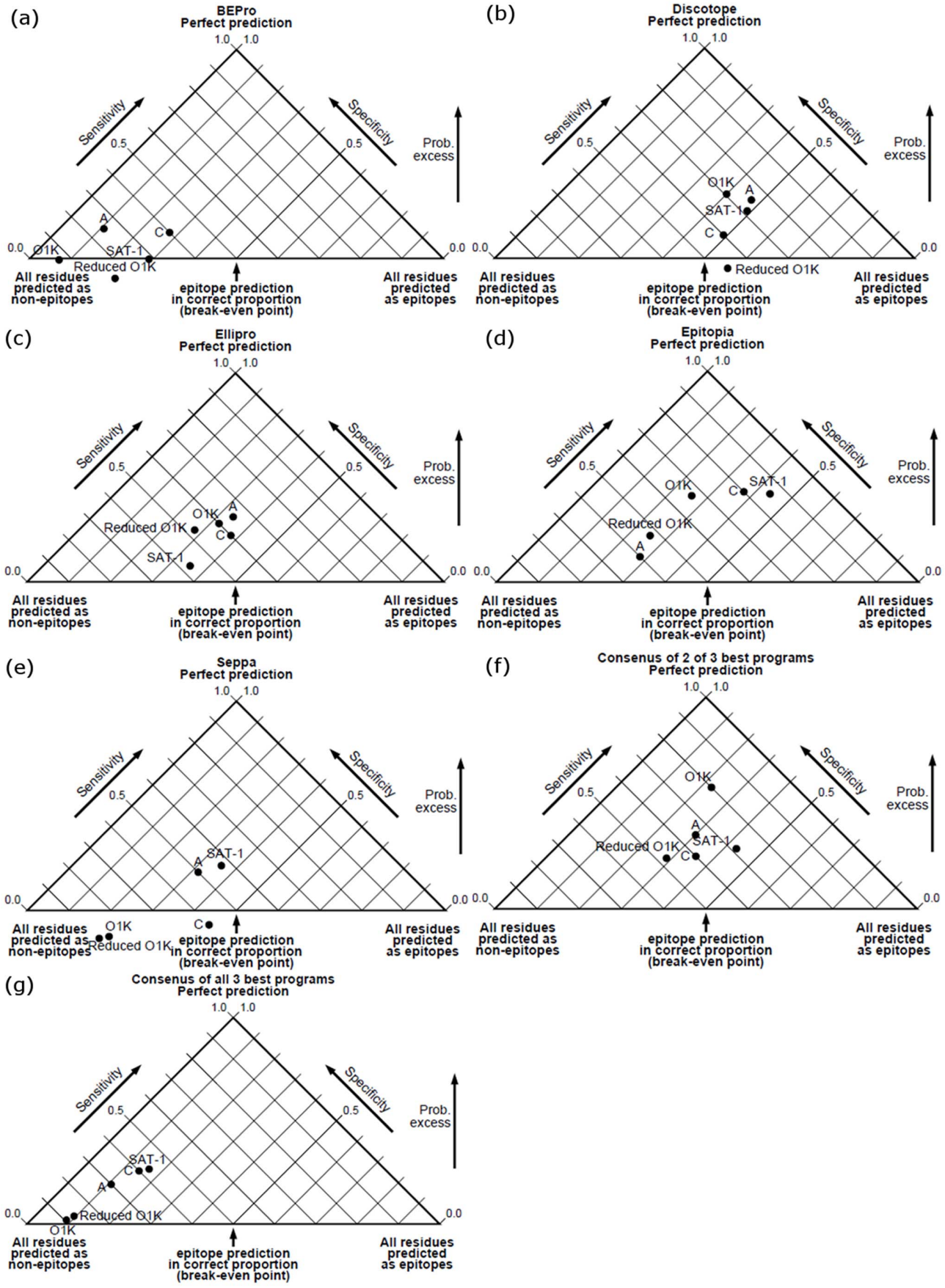
(a)
BEPro
Perfect prediction
1.0 1.0
Sensitivity
Specificity
0.5
0.5
Prob. excess
A
C
O1K
SAT-1
Reduced O1K
0.0
0.0
All residues predicted as non-epitopes
epitope prediction in correct proportion (break-even point)
All residues predicted as epitopes
(b)
Discotope
Perfect prediction
1.0 1.0
Sensitivity
Specificity
0.5
0.5
Prob. excess
O1K
A
SAT-1
C
Reduced O1K
0.0
0.0
All residues predicted as non-epitopes
epitope prediction in correct proportion (break-even point)
All residues predicted as epitopes
(c)
Ellipro
Perfect prediction
1.0 1.0
Sensitivity
Specificity
0.5
0.5
Prob. excess
Reduced O1K
O1K
A
C
SAT-1
0.0
0.0
All residues predicted as non-epitopes
epitope prediction in correct proportion (break-even point)
All residues predicted as epitopes
(d)
Epitopia
Perfect prediction
1.0 1.0
Sensitivity
Specificity
0.5
0.5
Prob. excess
O1K
C
SAT-1
Reduced O1K
A
0.0
0.0
All residues predicted as non-epitopes
epitope prediction in correct proportion (break-even point)
All residues predicted as epitopes
(e)
Seppa
Perfect prediction
1.0 1.0
Sensitivity
Specificity
0.5
0.5
Prob. excess
SAT-1
A
0.0
0.0
All residues predicted as non-epitopes
O1K
Reduced O1K
C
epitope prediction in correct proportion (break-even point)
All residues predicted as epitopes
(f)
Consenus of 2 of 3 best programs
Perfect prediction
1.0 1.0
Sensitivity
Specificity
0.5
0.5
Prob. excess
O1K
A
SAT-1
Reduced O1K
C
0.0
0.0
All residues predicted as non-epitopes
epitope prediction in correct proportion (break-even point)
All residues predicted as epitopes
(g)
Consenus of all 3 best programs
Perfect prediction
1.0 1.0
Sensitivity
Specificity
0.5
0.5
Prob. excess
SAT-1
C
A
Reduced O1K
O1K
0.0
0.0
All residues predicted as non-epitopes
epitope prediction in correct proportion (break-even point)
All residues predicted as epitopes

Figure 2. Plots of the sensitivity and specificity of each program (a–e) and a consensus of any two or all three of the top three methods at identifying the known epitopes for each FMDV structure selected represented as black dots on the pyramid plot. Pyramid probability represents scales of sensitivity and specificity, with the probability excess increasing with height and a perfect prediction (where both the sensitivity and specificity score a value of 1) coming at the tip of the pyramid. The bottom line of the plot represents the point at which results would be expected to be generated by chance.
doi:10.1371/journal.pone.0061122.g002

The aim of our work is to examine the potential power of current structure-based approaches to predict unidentified Picornavirus epitopes using an appropriate subset of the icosahedral capsid which contains a central protomeric unit augmented by neighbouring structures sufficient to ensure that epitopes spanning symmetry related protomers can be identified (Figure 1). There is potential utility in such an approach, since it could provide a much more rapid, cheap and easy means of predicting viral epitopes conserved across several serotypes compared to the expensive and time consuming serological methods. We firstly define objective measures for sensitivity and specificity and then evaluate the best parameters for a selection of programs using an independent training set of known epitopes/structures from related picornaviruses (Poliovirus and Rhinovirus) before evaluating their performance in terms of their ability to identify already known FMDV epitopes for all serotypes for which a structure is available. On the basis of these results we use a consensus of the outputs of the more reliable programs to predict further antigenic regions on the capsid of FMDV. The results, which are the first to look at virus structures specifically in this context, demonstrate that there are severe limitations with the predictive power of each of the programs evaluated; however the predictions made using the consensus data are of sufficient interest to justify further experimental investigations.

Methods

Program Selection

Five freely accessible web-based B cell epitope prediction servers were selected for evaluation. These were BEPro [22] Discotope [23], Ellipro [25], Epitopia [21] and Seppa [24]. These are believed to encompass the most recent programs available (at the time of the evaluation), with the exclusion of the EPCES and EPSVR servers [26], for which the size of the Protein data bank (PDB) files for the multimeric structures being used was prohibitive.

Generally, for each program, a single value is assigned to each residue predicting its likelihood of being an epitope. However, for Epitopia two values are included in the output for each residue, one being an immunogenicity score and the second being a probability score. For the purposes of this analysis the probability score was used to determine the optimal threshold for identification of Picornavirus epitopes. Since the two values are strongly correlated this should not impact on the results.

Structure Selection

Two Picornavirus structures were selected for program optimisation; Poliovirus P1 Mahoney strain (PDB code 1HXS- [27]) and Rhinovirus 14 (4RHV- [28]). These were chosen as they are the most well characterised Picornaviruses in terms of identified epitope residues. For FMDV, a structure was selected for each serotype (where available) for program evaluation. The list of structures used was as follows: Serotype C-S8c1 (1FMD- [29]), A1061 (1ZBE: [30]), O (O_1 Kaufbeuren- [31]), SAT-1 (2WZR- [13]). An additional structure for serotype O (O_1 Kaufbeuren- data not published) with the G-H loop present in a reduced conformation was also included in order to evaluate the impact of the presence of the G-H loop.

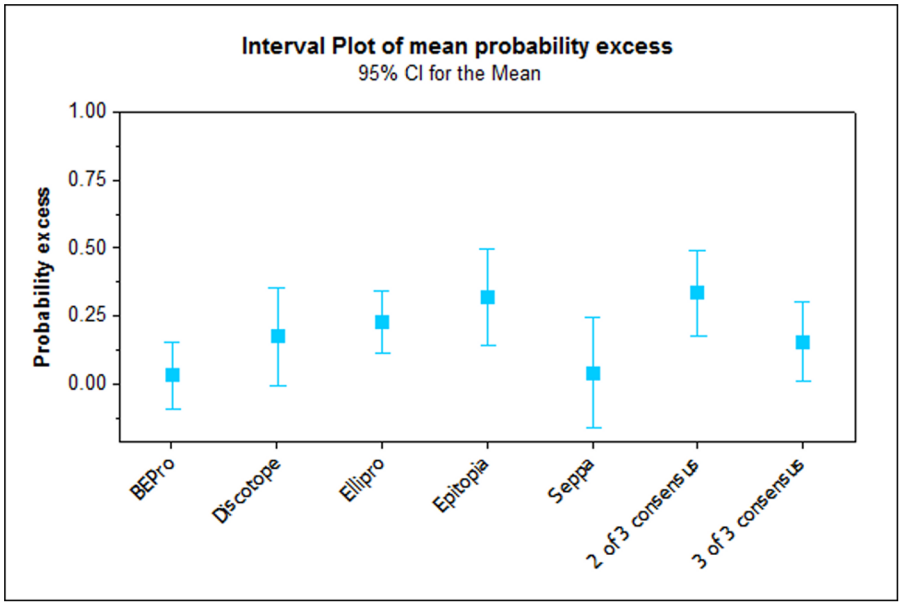

Figure 3. Interval plot showing the mean probability excess (and confidence intervals) of each program and a consensus of any two or all three of the top three methods.
doi:10.1371/journal.pone.0061122.g003

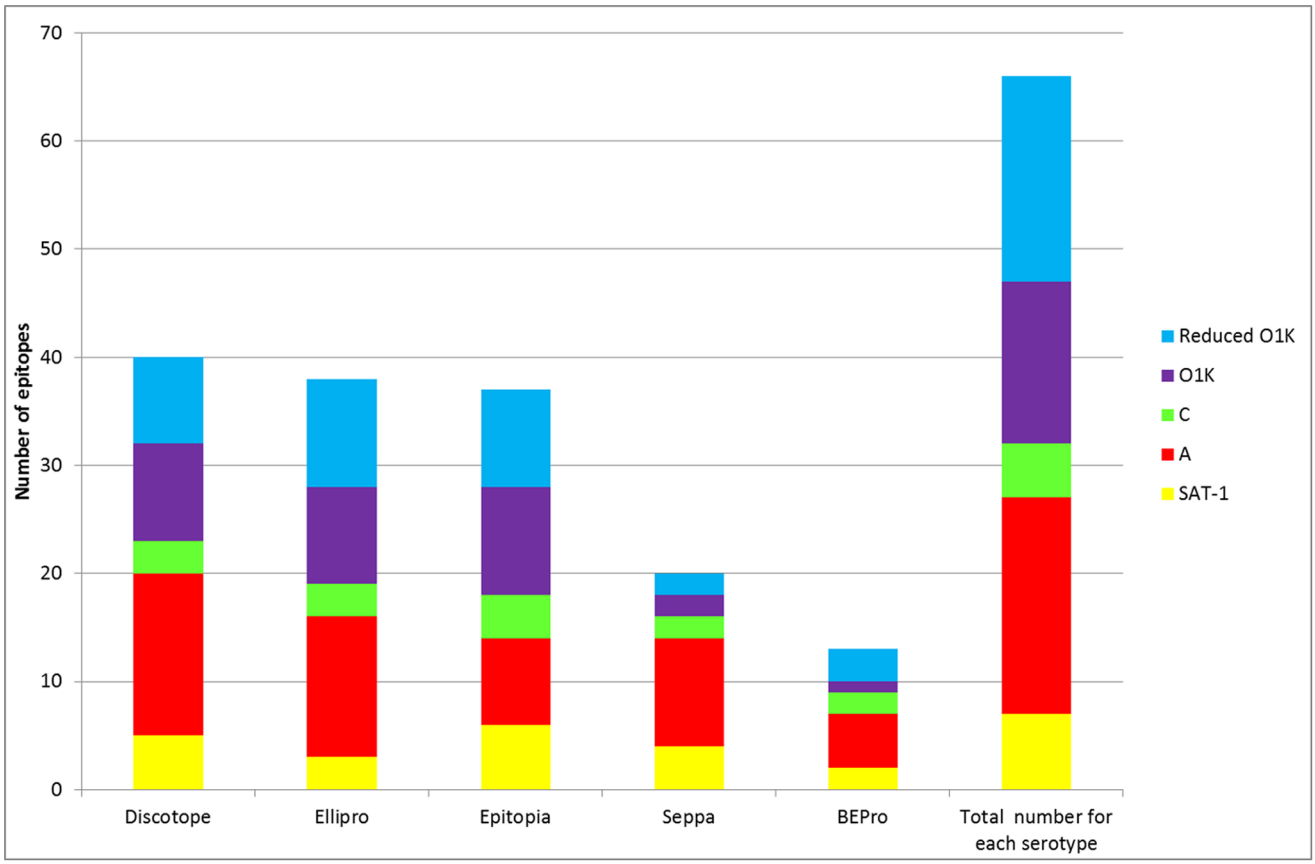

Figure 4. Graph showing the number of known epitopes for each serotype identified by each program compared to the total number of known epitopes for each serotype.
doi:10.1371/journal.pone.0061122.g004

Sequence and Conformational Comparison of Structures

Poliovirus and Rhinovirus are both members of the enterovirus genera of the Picornaviruses. The sequences of the two virus structures examined here share a 50% identity to each other while the structures have an average root-mean-square distance (RMSD) between 765 structurally equivalent C alpha atoms of 0.96 Å (calculated with SHP [32]). The sequences of these two enteroviruses have 22–24% identity with the sequences of the four FMD viruses we use in the subsequent evaluation (calculated in MEGA 5 [32] using a p-dist matrix and multiplying by 100 - used for all Identity calculations). The RMSD between Rhinovirus and serotype O FMDV C_α (excluding the G-H Loop) has previously been reported as 2.3 Å for VP1 and 1.8 Å for both VP2 and VP3 [33].

Serotypes O, A and C of FMDV share an approximate 80–82% sequence identity across the capsid proteins while the SAT-1 virus is less closely related with a sequence identity of ~65% with the other serotypes. VP1 exhibits the most variation (~30% for O, A and C and 50% for SAT-1) and VP2 exhibits the least variation of the surface exposed proteins (~17% between O, A and C and 30% for SAT-1). On average the RMSD (calculated using the structural homology program [34]) of each of the surface exposed virus proteins was 0.85 Å for VP1 (excluding the G-H loop), 0.68 Å for VP2 and 0.66 Å for VP3.

Structure Preparation

In the virus, the protein subunits are arranged with icosahedral symmetry, so the first stage in preparing the coordinates was to add the appropriate context to the selected icosahedral protomeric unit. To do this, a complete icosahedral capsid for each structure was first generated from a protomeric subunit using the 60 non-crystallographic symmetry matrices included in each coordinate file and the General Averaging Program (GAP, Stuart D.I and Grimes J.S. unpublished). From this complete structure, an icosahedral (triangular) protomeric subunit comprising VP1-4 was selected together with the protomers bounding each of the sides of the first, forming a tetrameric multimer see Figure 1(b).

The individual polypeptide chains (VP1, VP2, VP3 and VP4) that make up each protomer were relabeled [35] so each protomer was represented as a single chain, with the central protomer labeled as chain 1 and the outer protomers labeled as chains 2, 3 and 4 respectively (Figure 1(b)). This enabled the central protomer to be selected as a single chain for analysis, with all interfacing protomers being taken into account as non-selected chains where possible. Where this was not possible the entire multimer was uploaded into the epitope program and only the data for the central protomer taken forward for further analysis.

For the development of the method the surface exposure of each residue on the central protomer was determined, initially suitable criteria were established by inspection but the procedure now forms part of the software pipeline we have developed (see Files S1 and S2). In brief we use a probe the approximate size of a water molecule (1.6 Å), and any residue on the surface with greater than 1 $Å^2$ exposed was included in the analysis. The interactions of residues at interfaces between protomers were determined using

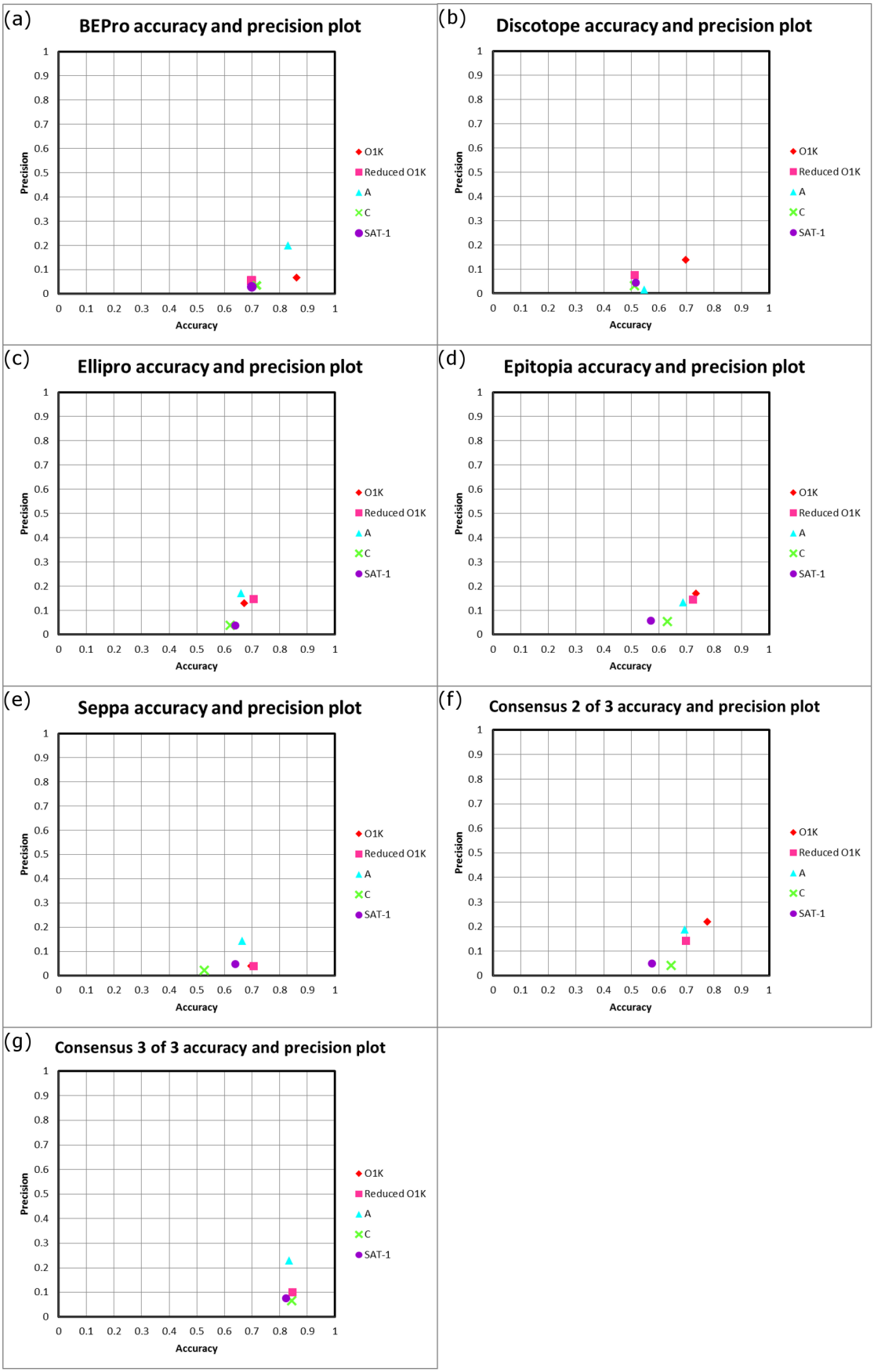
(a)
BEPro accuracy and precision plot
(b)
Discotope accuracy and precision plot
(c)
Ellipro accuracy and precision plot
(d)
Epitopia accuracy and precision plot
(e)
Seppa accuracy and precision plot
(f)
Consensus 2 of 3 accuracy and precision plot
(g)
Consensus 3 of 3 accuracy and precision plot
Precision
Accuracy
O1K
Reduced O1K
A
C
SAT-1

Figure 5. Plots of the the precision and accuracy of each program (a–e) and a consensus of any two or all three of the top three methods at identifying the known epitopes for each serotype with a structure available.
doi:10.1371/journal.pone.0061122.g005

the PISA program [36]. Only surface exposed residues were included in the analysis.

Identification of Picornavirus Epitopes

A literature search was conducted to collate all identified epitopes for each of the structures selected. A total of 53 and 22 neutralising epitopes were identified for Poliovirus serotype 1 [37–43] and Rhinovirus serotype 14 [44,45] respectively. For each serotype of FMDV the following numbers of epitopes were identified; 25 for serotype A [30,46–50], 20 for serotype O [51–56], 16 for serotype SAT-1 [57] and 10 for C [29,58–60]. Asia-1 [61] and SAT-2 epitopes [62] were also identified, although they were not included in the analysis as there was no available capsid structure for either serotype. All of the epitopes described here, with the exception of residue 188 of VP2 which was identified using bovine mAbs, [54] are derived from murine mAbs. Only the exact residues identified were examined.

For all of the FMD structures, with the exception of the reduced O_1 Kaufbeuren (O_1K) structure, the immunogenic and mobile G-H loop was not represented and therefore any antigenic residues previously described on this loop were not included in the program evaluation. Additionally, as the analysis utilises structural data, only epitopes identified on the native capsid were included (i.e. monoclonal antibody escape mutants).

Comparisons of Genetic Diversity

The genetic diversity of FMDV viruses both within serotype O and between serotypes was analysed using the Espript program [63] and a Risler similarity scoring matrix [64]. Firstly the amino acid sequences of capsids from either serotype O (N = 105) or all seven serotypes (N = 255- 105 for serotype O, 50 for serotype A, 28 for serotype Asia-1, 28 for SAT-1, 25 for SAT-2, 7 for SAT-3 and 12 for serotype C) were aligned to the amino acid sequence of the reduced serotype O1K structure file using the Clustal X program [65]. This alignment was subsequently divided into the individual proteins VP1-VP3 and any amino acid residues not found within the pdb file were removed. This meant that residues 212 and 213 of VP1 and residues 1–4 of VP2 were not included in the analysis. Each sequence file and the corresponding PDB file for each capsid protein, VP1-3, were uploaded into the program Espript.

The similarity scores generated for each residue were output in the B factor column of a new PDB file, with scores given from 100 (identical) to 0 (the most diverse). These scores were reversed so 100 represented the most diverse residues to create a 'percentage diversity' score.

Determination of Program Performance

In order to evaluate the performance of each program the following parameters were determined (where 'truth' is defined as agreement with the experimental data described above):

Sensitivity = number of true positives/(number of true positives+number of false negatives)

Specificity = number of true negatives/(number of true negatives+number of false positives)

Accuracy = (true positives+ true negatives)/(true positives+number of false positives+ false negatives+true negatives)

Precision = true positives/(true positives+false positives)

Probability excess = (sensitivity+specificity) –1.

Statistical Analysis

To determine the significance of each of the results an analysis of variance was performed using a general linear model with the virus name and the program as fixed effects. All comparisons were performed using the Tukey method for multiple comparisons, with confidence limits set at 95%. All statistical functions performed were completed using the Minitab 16 statistical software package (Minitab Inc, USA).

Partial Automation of the Procedure for Identifying Virus Epitopes

The above procedure has been semi-automated for the use of other workers to predict picornavirus epitopes using two programs to handle the preparation (EPIPREP- file S1) and analysis of the results (EPI_PRESENT- file S2) respectively. EPIPREP takes a standard PDB file, converts the multiple chains of an icosahedral asymmetric unit to a single super-chain (providing the re-mapping information to the user), generates the full icosahedral structure from the symmetry information contained in the PDB file and selects the neighbours making the greatest contacts with the reference unit, The contacting units are numbered uniquely and an output file produced which can then be directly loaded into the epitope detection servers.

For data analysis the text files from the servers are input to EPI-PRESENT and formatted as a single merged table for the users, with residues on the inner surface of the capsid removed. Input files and a command line are also generated to facilitate graphical analysis with RIVEM [66].

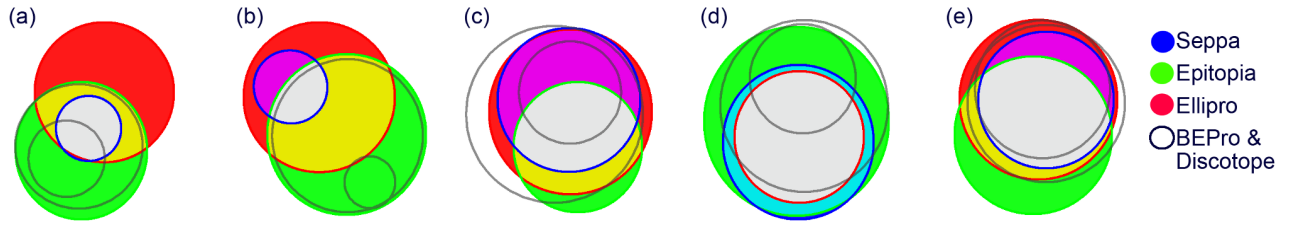

Figure 6. Venn diagrams showing the regognition of residues by each program for FMDV serotype O1K-Reduced (a), O1K (b), A1061 (c), SAT-1 (d) and Cs8 c1 (e). The key on the far right hand side indicates the colour of each program as represented on the Venn Diagram. For clarity the two worst perfoming algorithms are not coloured. In regions of overalap the colour is represented as the sum of the RGB colour channels of the overlapping mathods Diagrams made using the Venn master program (Kestler et al., 2008). Note that formally there need be no perfect projection of the multi-dimensional overlap information into the Venn diagram, so these represent best approximations.
doi:10.1371/journal.pone.0061122.g006

Chain ID	O1K	Reduced O1K	A	C	SAT-1	Known Antigenic site?	
VP1	VP1:ASN 46 VP1:GLN 47	VP1:ASN 46			VP1:ASN 46 VP1:ASP 47	45 &48 Antigenic site on type O, 48 on SAT-1	
				VP1:THR 83		Antigenic Site on type A	A
	VP1:GLY 84		VP1:GLY 84				
			VP1:GLN 169			Antigenic Site on type A	B
	VP1:THR 171	VP1:THR 171		VP1:GLU 171			
	VP1:ARG 172	VP1:ARG 172		VP1:THR 172		173 Antigenic site on type A	
	VP1:THR 174	VP1:THR 174					
					VP1:TYR 195		
					VP1:ASP 196		
					VP1:HIS 197		
					VP1:ASN 198		
			VP1:ILE 192		VP1:GLY 199		
			VP1:LYS 193		VP1:ARG 200	Antigenic Site on type C (193)	
		VP1:ILE 194	VP1:VAL 194	VP1:ILE 194	VP1:ASP 201		C
	VP1:HIS 195	VP1:HIS 195	VP1:THR 195	VP1:GLN 195	VP1:ARG 202		
	VP1:PRO 196	VP1:PRO 196	VP1:SER 196	VP1:PRO 196	VP1:TYR 203		
	VP1:THR 197	VP1:THR 197	VP1:GLN 197	VP1:THR 197	VP1:LYS 204		
	VP1:GLU 198	VP1:GLU 198	VP1:ASP 198	VP1:GLY 198		Antigenic site on type O and A	
	VP1:ALA 199	VP1:ALA 199	VP1:ARG 199	VP1:ASP 199			
				VP1:PRO 204		Antigenic site on type A	
			VP1:ILE 205				
VP2			VP2:ASP 68	VP2:ASP 68		67 Antigenic site on type Asia-1	
				VP2:VAL 70		Antigenic site on type O	
				VP2:PRO 71		Antigenic site on type O	
	VP2:SER 72	VP2:SER 72	VP2:ASP 72	VP2:SER 72	VP2:ASP 72	Antigenic on all tested serotypes	D
				VP2:GLN 73			
				VP2:ASN 74		Antigenic site on type C and Asia-1	
				VP2:ASP 134	VP2:ASP 134	Potential Antigenic site on type O	
				VP2:PRO 186			
				VP2:THR 188		Antigenic site on type O	
				VP2:THR 189			
	VP2:ASN 190	VP2:ASN 190	VP2:SER 190	VP2:ASN 190	VP2:ASP 190		E
	VP2:THR 191	VP2:THR 191	VP2:ASN 191	VP2:THR 191	VP2:GLN 191		
	VP2:GLU 192	VP2:GLU 192		VP2:ALA 192	VP2:THR 192		
				VP2:GLY 193	VP2:ILE 193		
					VP2:GLY 194		
				VP2:GLN 195			
				VP2:GLN 196		Antigenic site on type A	
VP3		VP3:GLU 58			VP3:ASN 58	Antigenic site on type O, A, C and Asia-1	
		VP3:GLY 59			VP3:GLY 59	Antigenic site on type A and Asia-1	
		VP3:GLY 60			VP3:VAL 60		
		VP3:VAL 61				Antigenic site on type A	
					VP3:GLN 66		
					VP3:SER 67		
			VP3:ALA 68		VP3:ASN 68		
	VP3:ASP 69	VP3:ASP 69	VP3:ASP 69		VP3:SER 69	Antigenic site on type A	F
			VP3:ASP 70		VP3:GLY 70		
	VP3:ASP 71	VP3:ASP 71	VP3:THR 71		VP3:SER 71		
					VP3:THR 112		
					VP3:GLY 113		
			VP3:SER 114				
			VP3:ASP 116				
					VP3:GLU 135	Antigenic site on SAT-1	
					VP3:THR 191		
			VP3:HIS 192				
			VP3:GLY 193		VP3:THR 193		G
			VP3:LYS 194		VP3:HIS 194		
				VP3:ASP 195		Antigenic site on type A	
			VP3:GLU 196		VP3:LYS 196		
			VP3:ASN 197		VP3:ASP 197		
					VP3:ALA 199		
				VP3:GLN 219		218 antigenic site on Asia-1	
				VP3:GLN 220	VP3:ARG 220		
					VP3:GLN 221		

Figure 7. List of residues selected by a consensus of the three best performing programs (Discotope, Ellipro and Epitopia) for each selected FMDV structure compared to locations of known antigenic sites of all serotypes. Those residues coloured red are an already known epitope on at least one serotype of FMDV, those in blue are adjacent to a known epitope of FMDV. Regions A–G are predicted to be antigenic on the majority of the serotypes tested and are coloured the same on Figures 8 and 9. The remaining residues are coloured grey, as Figures 8 and 9. Note that the SAT-1 virus VP1 has a incorporated several additional residues into the G–H loop and the A serotype also aligns slightly differently to the O and C structures, therefore the numbering is different in region C as they are aligned according to position on structure. All other residues are in approximately the same location relative to each other.
doi:10.1371/journal.pone.0061122.g007

Executables of EPIPREP and EPI_PRESENT are available from the authors on request and source code is provided as Supplementary Information (files S1 and S2).

Results

Program Optimisation

For each program in the trial, the program authors had previously established an optimal value for the threshold at which a residue is predicted to be part of an epitope by optimisation against a broad set of epitopes in a database of test structures. The metric used to set the threshold was the so-called receiver operating characteristic area under the curve (ROC AUC) value: (this is simply the area under the curve obtained by plotting true positives against false positives, both as fractions) a value greater than 0.5 represents a robust predictor performing better than chance with a value of 1 indicative of a perfect predictor. However, as this threshold is set against a broad range of epitopes it may not be appropriate for a particular system, such as virus capsids, where neotopes are likely to arise due to the juxtaposition of repeated protomer subunits.

To evaluate the utility of the suggested threshold values the performance of each of the programs was tested at 6 different threshold values against both Poliovirus and Rhinovirus. The results demonstrate that for each program a single threshold value yielded the optimal performance against both of the viruses; however, the performance between Polio and Rhinovirus is in several cases markedly different at this value (Table 1). For example, the Discotope server performed better at identifying Polio epitopes than at identifying Rhinovirus epitopes. The threshold values determined by our enterovirus based analyses were in many cases different to the default threshold suggested by the program developers, for example the Discotope server suggests a default threshold of -7.7, whereas our evaluation determined the optimal threshold to be -10.7. Additionally, the results demonstrated the generally poor sensitivity and specificity of each program tested. The picornavirus optimised threshold values were then used in the analysis of the FMDV outputs.

Performance of each Program at Identifying Known FMDV Epitopes

The best performing programs in terms of sensitivity, specificity and in the average probability excess were Epitopia and Ellipro, followed by Discotope (see Figures 2 and 3). These programs also performed well in terms of overall number of FMDV epitope residues identified (Figure 4), with Discotope correctly identifying 61% of the total of the previously described epitopes for all serotypes of FMDV, followed by Ellipro (58%) and Epitopia (56%). Seppa and BEPro performed poorly in terms of sensitivity and specificity - in some cases the programs performed worse than if the residues had been selected by chance, and average probability excess, with Epitopia performing significantly ($P<0.05$) better than both of these programs (Figures 2 and 3), in addition BEPro and Seppa also identified the least number of epitopes (Figure 4), with Seppa identifying 34% and BEPro 22% of the total number respectively.

All of the programs performed poorly in terms of precision (Figure 5), with most results below 0.2. However all performed better in terms of accuracy of the results obtained, with BEPro performing the best (average 0.761), followed by Ellipro, Epitopia, Seppa and Discotope (Figure 5).

Although the programs utilise different algorithms there was considerable overlap in the residues identified, and there were epitopes on each serotype that were not identified by any of the programs (Figure 4). This suggests that although the programs utilise different algorithms for identifying epitopes they are not fully independent in terms of the outputs (this is seen graphically in Figure 6).

Using a Consensus Approach to Identify Novel Epitopes

To investigate the potential of using a consensus approach to predict novel epitopes the results for the three best performing programs in terms of average probability excess were selected (Epitopia, Ellipro and Discotope - see section above). The residues that were predicted to be antigenic by a consensus of all three programs on the selected structures were tabulated and compared to the known antigenic residues described previously (see Figure 7). Figure 2(g) indicates, as expected, that the sensitivity of such an approach is limited, however the significantly ($P<0.01$) increased accuracy obtained over any individual program (except the poorly performing program, BEPro $P = 0.12$) or using a consensus of 2 of the 3 programs (Figure 5(g)) validates the use of this approach. Additionally for the purpose of predicting novel epitopes using such a conservative approach is appropriate as the number of false positives should be as small as possible. It is also apparent that selecting the consensus results from any two of the best three methods gives an overall improvement in consistency, with all the sensitivity and specificity results in the right proportions (figure 2). Significant ($P<0.01$) improvements in the average probability excess were obtained compared to both Seppa and BEPro, and marginal, but not statistically significant, average probability excess 0.338 compared to 0.319 for the best performing program, Epitopia (Figure 2(f) and Figure 3). Whilst there was no improvement in precision or accuracy over any single program, it is possible that this approach may be useful for some applications.

The majority of residues identified using this triple-consensus approach are located at regions previously described as antigenic on at least one serotype. Additionally, several of these regions were also selected for the majority (or all) of the serotypes tested (see Figure 7). This suggests that some of these residues may comprise presently unidentified epitopes on certain serotypes. The regions predicted as antigenic on the majority of serotypes are mapped onto the surface of a serotype O structure in Figures 8 and 9.

One region that is predicted for all serotypes but not previously identified as an epitope in the literature is residues 190 to 192 of VP2. This region is around the 3-fold axis and is 17 Å (from the C_α of the central amino acid VP2 191 to the C_α of VP2 72) away from residue 72 of VP2, which has previously been described as an

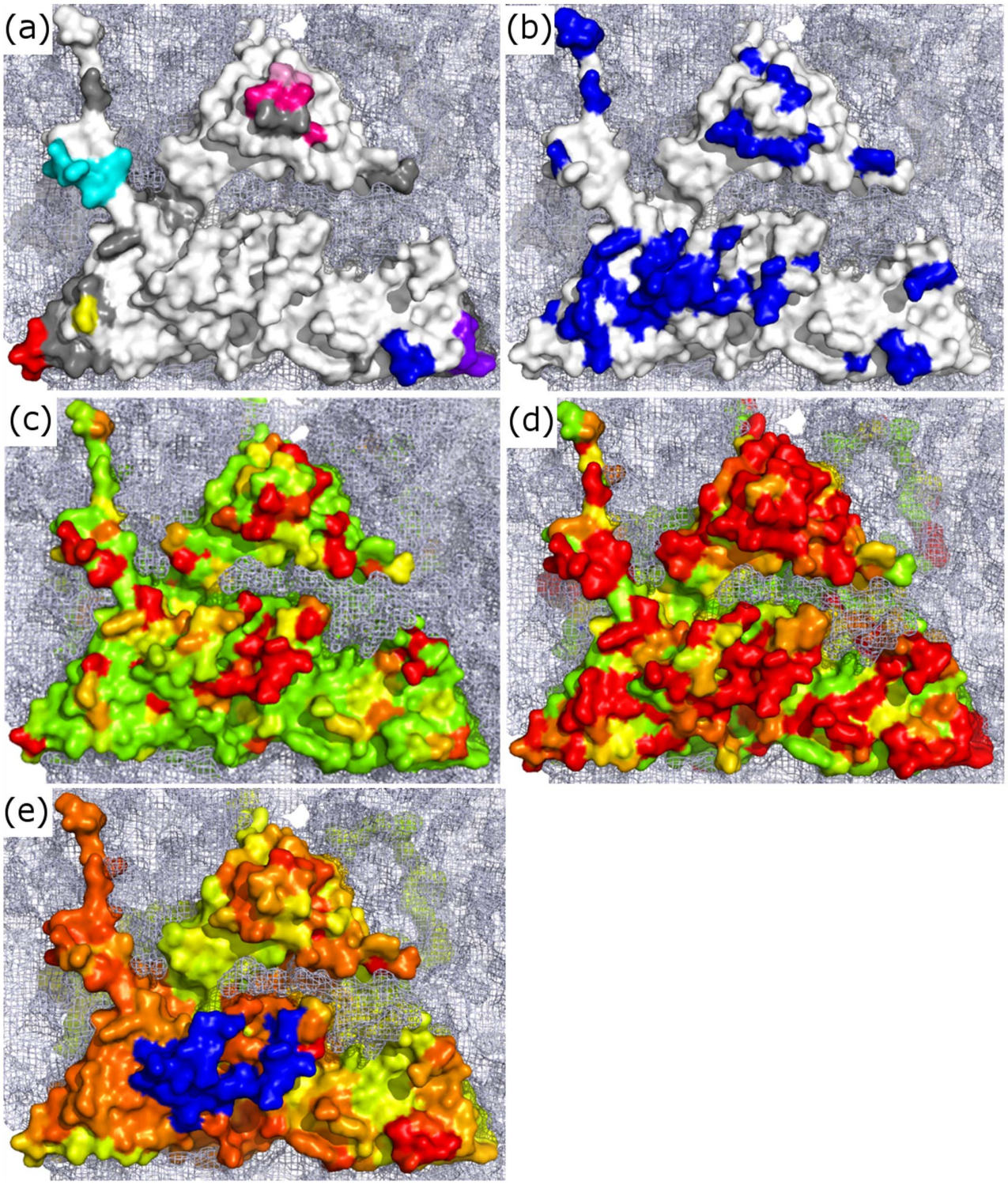

Figure 8. A side-by-side comparison of the molecular surface of five reduced O1 Kaufbeuren protomers made using Pymol, with the molecular surface of neighbouring protomers at each two fold axis of symmetry shown as a light blue mesh (Schrödinger LLC). Protomer (a) shows the residues in the regions A–G from Figure 7 coloured accordingly mapped onto the molecular surface (white). Protomer (b) shows the location of all described epitopes in blue for each serotype used in the analyses. Protomer (c) is a pictorial representation of the sequence variability of residues between 105 serotype O virus capsid sequences and Protomer (d) is a representation of the sequence variability of residues between 255 capsid sequences from all 7 serotypes. Both (c) and (d) were constructed using the ESPRIPT program [63] as described in the methods, the coloring goes from green (most conserved) to red (least conserved). Those sites predicted by a consensus of all programs on three or more structures tend to be located at areas of variability both inter and intra-serotypically. Region VP3 69–71 (highlighted in red on (a) and (c)) is completely conserved within serotype O, however it appears to be highly divergent between serotypes. (e) is the average RMSD across the serotypes coloured from green (smallest) to red (largest). Blue indicates the GH loop, missing on all but one structure.
doi:10.1371/journal.pone.0061122.g008

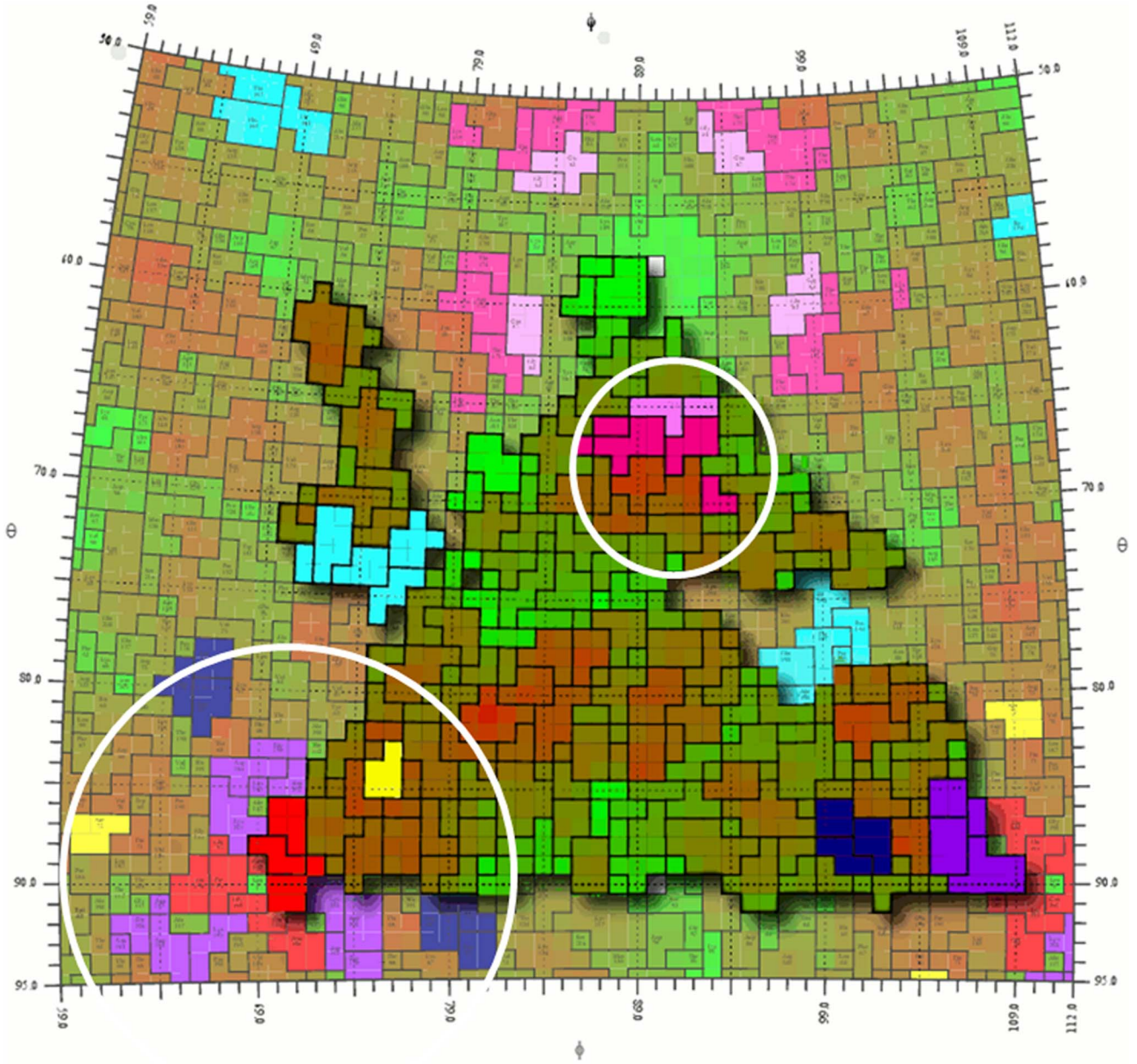

Figure 9. Roadmap illustration of a zoomed in area of the reduced O1K capsid generated using the RIVEM program (Xiao & Rossmann, 2007). The location of a single protomer is outlined. The occupancy of each residue on the surface is shown as an area corresponding to the surface accessibility with a black outline, with the protomer coloured according to the radius of each residue from the centre of the icosahedral capsid, coloured from dark green (least exposed) to brown (most exposed). The residues predicted by the best performing programs on the majority of the structures evaluated are coloured as in Figure 6. When the individual areas are examined it can be seen that several regions are in very close proximity across the 3-fold axis, potentially forming a conformational epitope. Additionally two of the regions are adjacent to each other near the 5-fold axis, again potentially forming a conformational epitope. These two regions are circled in white.
doi:10.1371/journal.pone.0061122.g009

epitope on all serotypes tested and is also predicted across all serotypes tested. This region also lies close to two other regions predicted on multiple serotypes, VP3 69–71 and 193–197 (~25 Å and ~14 Å from the C_α of the central amino acid VP2 191 to the C_α of VP3 70 and 195 respectively), potentially forming an antigenic site spanning the 3-fold axis across pentamers. A further area of interest lies towards the 5-fold axis of symmetry, with two regions (VP1 residues 83–84 and 169–174) identified as potentially forming a conformational epitope, again on multiple serotypes. The location of these regions is shown in Figure 9 (coloured as in Figure 7).

When the sequences of these putative epitopes are compared between different serotype O viruses (Figure 8) it can be seen that VP2 190–192 and VP1 194–199 in particular are highly variable, and so could contribute to the antigenic variation observed within this serotype. The average amino acid diversity of all surface exposed residues across the serotypes (when aligned to the reduced O1K structure) is 53.0%, this is markedly less than both the amino acid diversity of 64.5% within the predicted sites (Figure 8), and

the diversity observed across known antigenic sites (67.8%). In contrast one region (VP3 69–71) is completely conserved in sequence within serotype O and so antibodies against this putative epitope might be cross-reactive against all serotype O viruses. Nevertheless even this region is not conserved across all serotypes and appears to be located on the most significant region for structural variation (Figure 8). It would appear from the variability observed between the serotypes that the prospect of finding an inter-serotypically conserved epitope is poor, at least on those parts of the capsid that are most exposed (Figure 8).

The average RMSD values (when aligned to the O_1 Kaufbeuren structure) for the predicted sites, at 1.05 Å and for the known antigenic sites (1.13 Å) are both slightly greater than the average of the surface exposed residues (1.00 Å). For more detail see Table S1.

The Impact of G-H Loop Conformation

The presence or absence of the G-H loop did not appear to greatly influence the results, with a similar number of epitopes identified by each program for both structures. The same regions were also predicted as being antigenic on both structures. One difference observed was the decreased sensitivity and specificity of the programs (in particular Discotope) at identifying already known epitopes on the reduced (ordered loop) form of the O_1Kaufbeuren (Figure 2). This could be due to the inclusion of four additional epitopes located on the G-H loop that are not present on the non-reduced structure. As the G-H loop is considered to be a linear epitope these programs may not be well suited to identifying these residues.

Discussion

A number of methods have been developed to identify epitopes, generally refined by training against globular protein/antibody complexes. The description of the methods they use is beyond the scope of this paper but they tend to analyse surface accessibility in conjunction with patch location or clustering and may include residue propensity information. The methods are therefore not entirely independent but the nature of their correlations is not fully characterised.

We have applied the programs to a slightly different problem – the detection of epitopes on the surface of virus capsids of known structure. We have developed a pipeline to allow for the complex symmetry of the virus (including a definition of the exterior surface) without needing to analyse the entire capsid. Our tests have focused on picornaviruses, and we found that the optimal threshold values of the programs for identifying both Polio and Rhinovirus epitopes differed from those identified as optimal by the program developers. This suggests that the best method for utilising these servers to predict viral epitopes is to identify the optimal threshold value for each program using a virus related closely to that being examined as a training set.

The evaluation performed in this paper has demonstrated, even after a recalibration of the methods using an appropriate training set, considerable limitations in the sensitivity and specificity of each program when applied to Picornavirus epitopes in the context of the intact capsid. Note that there may well be as yet undiscovered additional antigenic residues on the capsid, so that some predictions classified as false positives might be genuine. However any improvement in performance would be likely offset by the fact that some of the negatives classified currently as correct would probably be reclassified as false negatives.

We have found that the predictions are rendered marginally more reliable by using a consensus of any two of the best three performing programs (Figure 2(f) and figure 3). However when a consensus of all three best performing programs is taken, although the sensitivity of the results decreases, the specificity and accuracy increase significantly, indicating that this conservative approach results in less false positives and gives more confidence in the results. This is of particular importance when using the program outputs as predictions to base time consuming and expensive experiments on. We therefore examined the triple consensus results to see if there are residues which may represent undiscovered FMDV epitopes on each serotype evaluated. Several such putative epitopes were discovered and there are circumstantial reasons to believe that these predictions contain some truth. The majority of the residues predicted are located on or adjacent to a known epitope identified by monoclonal antibody techniques on at least one other serotype of FMDV. Potentially these residues are epitopes on all serotypes given the strong structural similarity of the FMDV capsids but due to the random nature of mAb selection and the discrepancy in the depth of investigation using these techniques between serotypes they may simply not have yet been discovered. Furthermore several regions are also predicted to be epitopes across serotypes. Finally the putative epitopes are on highly variable and mobile regions of the capsid which would be consistent with these regions coming under selective pressure, potentially from antibodies.

The consensus results predicted one completely novel epitope at residues 190–192 of VP2, residues which are highly variable in serotype O. Although this region is novel it should be noted that it is only slightly downstream of an epitope identified on serotype O [54]. Mapping this region to the capsid suggests that there might be a conformational epitope including both VP2 and VP3 residues spanning the 3-fold axis (and hence the pentamer-pentamer interface). A second conformational epitope is also suggested on VP1 towards the 5-fold axis at the centre of the pentamer. Thus our proof-of-principle application to FMDV has given some interesting results, which could be tested using a reverse genetics approach.

In conclusion, we believe the evaluation performed here, the first to specifically examine the outputs of these programs from the perspective of intact capsid structures, by highlighting the limited predictive power of an individual program at identifying established epitopes, suggests that it may be prudent to develop a program more specifically designed for the task of identifying viral epitopes, for example in using exclusively only epitope/paratope structures derived from viruses. However taking a triple consensus of the results from the best performing programs, suitably tuned could already provide a quick, cheap, and reasonably reliable method for predicting epitopes on several virus serotypes simultaneously. We have therefore developed a semi-automated pipeline, with executable programs available from the authors and source code provided in Supplementary Information.

Supporting Information

Table S1 Average diversity and RMSD of surface exposed residues, antigenic and predicted sites from all the serotypes evaluated.
(XLSX)

File S1 The EPIPREP source code.
(ZIP)

File S2 The EPI_PRESENT source code.
(F)

Acknowledgments

The authors would like to thank Dr Thomas Bowden, Mr Daniel Balls and Mrs Claire Reigate for their advice and input into the paper. David J. Paton is a Jenner Investigator.

Author Contributions

Conceived and designed the experiments: DWB EEF. Performed the experiments: DWB. Analyzed the data: DWB. Contributed reagents/materials/analysis tools: DIS EEF RME. Wrote the paper: DB EEF DIS DJP MM RME.

References

1. Paton DJ, King DP, Knowles NJ, Hammond J (2010) Recent spread of foot-and-mouth disease in the Far East. Vet Rec 166: 569–570.
2. Valdazo-Gonzalez B, Knowles NJ, Wadsworth J, King DP, Hammond JM, et al. (2011) Foot-and-mouth disease in Bulgaria. Vet Rec 168: 247.
3. Yang M, Clavijo A, Suarez-Banmann R, Avalo R (2007) Production and characterization of two serotype independent monoclonal antibodies against foot-and-mouth disease virus. Veterinary Immunology and Immunopathology 115: 126–134.
4. Knowles NJ, Samuel AR (2003) Molecular epidemiology of foot-and-mouth disease virus. Virus Research 91: 65–80.
5. Brehm KE, Kumar N, Thulke HH, Haas B (2008) High potency vaccines induce protection against heterologous challenge with foot-and-mouth disease virus. Vaccine 26: 1681–1687.
6. Fry E, Logan D, Archarya R, Fox G, Rowlands D, et al. (1990) Architecture and topography of an aphthovirus. Seminars in Virology 1: 439–451.
7. Fox G, Parry NR, Barnett PV, McGinn B, Rowlands DJ, et al. (1989) The cell attachment site on foot-and-mouth disease virus includes the amino acid sequence RGD (arginine-glycine-aspartic acid). J Gen Virol 70 (Pt 3): 625–637.
8. Logan D, Abu-Ghazaleh R, Blakemore W, Curry S, Jackson T, et al. (1993) Structure of a major immunogenic site on foot-and-mouth disease virus. Nature 362: 566–568.
9. Hewat EA, Verdaguer N, Fita I, Blakemore W, Brookes S, et al. (1997) Structure of the complex of an Fab fragment of a neutralizing antibody with foot-and-mouth disease virus: positioning of a highly mobile antigenic loop. EMBO J 16: 1492–1500.
10. Juleff N, Windsor M, Lefevre EA, Gubbins S, Hamblin P, et al. (2009) Foot-and-mouth disease virus can induce a specific and rapid CD4+ T-cell-independent neutralizing and isotype class-switched antibody response in naive cattle. J Virol 83: 3626–3636.
11. Dunn CS, Samuel AR, Pullen LA, Anderson J (1998) The biological relevance of virus neutralisation sites for virulence and vaccine protection in the guinea pig model of foot-and-mouth disease. Virology 247: 51–61.
12. Maree FF, Blignaut B, Esterhuysen JJ, de Beer TA, Theron J, et al. (2011) Predicting antigenic sites on the foot-and-mouth disease virus capsid of the South African Territories types using virus neutralization data. J Gen Virol 92: 2297–2309.
13. Reeve R, Blignaut B, Esterhuysen JJ, Opperman P, Matthews L, et al. (2010) Sequence-based prediction for vaccine strain selection and identification of antigenic variability in foot-and-mouth disease virus. PLoS Comput Biol 6: e1001027.
14. Corti D, Voss J, Gamblin SJ, Codoni G, Macagno A, et al. (2011) A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. Science 333: 850–856.
15. Hopp TP, Woods KR (1981) Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci U S A 78: 3824–3828.
16. Jameson BA, Wolf H (1988) The antigenic index: a novel algorithm for predicting antigenic determinants. Comput Appl Biosci 4: 181–186.
17. Odorico M, Pellequer JL (2003) BEPITOPE: predicting the location of continuous epitopes and patterns in proteins. J Mol Recognit 16: 20–22.
18. Blythe MJ, Flower DR (2005) Benchmarking B cell epitope prediction: underperformance of existing methods. Protein Sci 14: 246–248.
19. Yang X, Yu X (2009) An introduction to epitope prediction methods and software. Rev Med Virol 19: 77–96.
20. Van Regenmortel MHV (1996) Mapping Epitope Structure and Activity: From One-Dimensional Prediction to Four-Dimensional Description of Antigenic Specificity. Methods 9: 465–472.
21. Rubinstein ND, Mayrose I, Martz E, Pupko T (2009) Epitopia: a web-server for predicting B-cell epitopes. BMC Bioinformatics 10: 287.
22. Sweredoski MJ, Baldi P (2008) PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure. Bioinformatics 24: 1459–1460.
23. Haste Andersen P, Nielsen M, Lund O (2006) Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Sci 15: 2558–2567.
24. Sun J, Wu D, Xu T, Wang X, Xu X, et al. (2009) SEPPA: a computational server for spatial epitope prediction of protein antigens. Nucleic Acids Res 37: W612–616.
25. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, et al. (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics 9: 514.
26. Liang S, Zheng D, Standley DM, Yao B, Zacharias M, et al. (2010) EPSVR and EPMeta: prediction of antigenic epitopes using support vector regression and multiple server results. BMC Bioinformatics 11: 381.
27. Miller ST, Hogle JM, Filman DJ (2001) Ab initio phasing of high-symmetry macromolecular complexes: successful phasing of authentic poliovirus data to 3.0 A resolution. J Mol Biol 307: 499–512.
28. Arnold E, Rossmann MG (1988) The use of molecular-replacement phases for the refinement of the human rhinovirus 14 structure. Acta Crystallogr A 44 (Pt 3): 270–282.
29. Lea S, Hernandez J, Blakemore W, Brocchi E, Curry S, et al. (1994) The structure and antigenicity of a type C foot-and-mouth disease virus. Structure 2: 123–139.
30. Fry EE, Newman JW, Curry S, Najjam S, Jackson T, et al. (2005) Structure of Foot-and-mouth disease virus serotype A10 61 alone and complexed with oligosaccharide receptor: receptor conservation in the face of antigenic variation. J Gen Virol 86: 1909–1920.
31. Lea S, Abu-Ghazaleh R, Blakemore W, Curry S, Fry E, et al. (1995) Structural comparison of two strains of foot-and-mouth disease virus subtype O1 and a laboratory antigenic variant, G67. Structure 3: 571–580.
32. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.
33. Acharya R, Fry E, Stuart D, Fox G, Rowlands D, et al. (1989) The 3-Dimensional Structure Of Foot-And-Mouth-Disease Virus At 2.9-A Resolution. Nature 337: 709–716.
34. Stuart DI, Levine M, Muirhead H, Stammers DK (1979) Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 A. J Mol Biol 134: 109–142.
35. Kleywegt GJ, Zou JY, Kjeldgaard M, Jones TA (2001) Crystallography of Biological Macromolecules. In: Rossmann MGA, E., editor. International Tables for Crystallography. 353–356, 366–367.
36. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372: 774–797.
37. Page GS, Mosser AG, Hogle JM, Filman DJ, Rueckert RR, et al. (1988) Three-dimensional structure of poliovirus serotype 1 neutralizing determinants. J Virol 62: 1781–1794.
38. Minor PD (1986) Antigenic structure of poliovirus. Microbiol Sci 3: 141–144.
39. Minor PD (1990) Antigenic structure of picornaviruses. Curr Top Microbiol Immunol 161: 121–154.
40. Minor PD, Ferguson M, Evans DM, Almond JW, Icenogle JP (1986) Antigenic structure of polioviruses of serotypes 1, 2 and 3. J Gen Virol 67 (Pt 7): 1283–1291.
41. Minor PD, Schild GC, Bootman J, Evans DM, Ferguson M, et al. (1983) Location and primary structure of a major antigenic site for poliovirus neutralization. Nature 301: 674–679.
42. Diamond DC, Jameson BA, Bonin J, Kohara M, Abe S, et al. (1985) Antigenic variation and resistance to neutralization in poliovirus type 1. Science 229: 1090–1093.
43. Wiegers K, Uhlig H, Dernick R (1989) N-AgIB of poliovirus type 1: a discontinuous epitope formed by two loops of VP1 comprising residues 96–104 and 141–152. Virology 170: 583–586.
44. Sherry B, Mosser AG, Colonno RJ, Rueckert RR (1986) Use of monoclonal antibodies to identify four neutralization immunogens on a common cold picornavirus, human rhinovirus 14. J Virol 57: 246–257.
45. Sherry B, Rueckert R (1985) Evidence for at least two dominant neutralization antigens on human rhinovirus 14. J Virol 53: 137–143.
46. Mahapatra M, Seki C, Upadhyaya S, Barnett PV, La Torre J, et al. (2011) Characterisation and epitope mapping of neutralising monoclonal antibodies to A24 Cruzeiro strain of FMDV. Vet Microbiol 149: 242–247.
47. Saiz JC, Gonzalez MJ, Borca MV, Sobrino F, Moore DM (1991) Identification of neutralizing antigenic sites on VP1 and VP2 of type A5 foot-and-mouth disease virus, defined by neutralization-resistant variants. J Virol 65: 2518–2524.
48. Saiz JC, Gonzalez MJ, Morgan DO, Card JL, Sobrino F, et al. (1989) Antigenic comparison of different foot-and-mouth disease virus types using monoclonal antibodies defining multiple neutralizing epitopes on FMDV A5 subtypes. Virus Res 13: 45–60.
49. Thomas AA, Woortmeijer RJ, Puijk W, Barteling SJ (1988) Antigenic sites on foot-and-mouth disease virus type A10. J Virol 62: 2782–2789.
50. Baxt B, Vakharia V, Moore DM, Franke AJ, Morgan DO (1989) Analysis of neutralizing antigenic sites on the surface of type A12 foot-and-mouth disease virus. J Virol 63: 2143–2151.
51. Aktas S, Samuel AR (2000) Identification of antigenic epitopes on the foot and mouth disease virus isolate O1/Manisa/Turkey/69 using monoclonal antibodies. Rev Sci Tech 19: 744–753.
52. Kitson JD, McCahon D, Belsham GJ (1990) Sequence analysis of monoclonal antibody resistant mutants of type O foot and mouth disease virus: evidence for the involvement of the three surface exposed capsid proteins in four antigenic sites. Virology 179: 26–34.

53. Crowther JR, Farias S, Carpenter WC, Samuel AR (1993) Identification of a 5th neutralizable site on type-o foot-and-mouth-disease virus following characterization of single and quintuple monoclonal-antibody escape mutants. Journal of General Virology 74: 1547–1553.
54. Barnett PV, Samuel AR, Pullen L, Ansell D, Butcher RN, et al. (1998) Monoclonal antibodies, against O1 serotype foot-and-mouth disease virus, from a natural bovine host, recognize similar antigenic features to those defined by the mouse. J Gen Virol 79 (Pt 7): 1687–1697.
55. Mahapatra M, Aggarwal N, Cox S, Statham RJ, Knowles NJ, et al. (2008) Evaluation of a monoclonal antibody-based approach for the selection of foot-and-mouth disease (FMD) vaccine strains. Vet Microbiol 126: 40–50.
56. Crowther JR, Farias S, Carpenter WC, Samuel AR (1993) Identification of a fifth neutralizable site on type O foot-and-mouth disease virus following characterization of single and quintuple monoclonal antibody escape mutants. J Gen Virol 74 (Pt 8): 1547–1553.
57. Grazioli S, Moretti S, Barbieri I, Crosatti M, Brocchi E (2006) Use of monoclonal antibodies to identify and map new antigenic determinants involved in neutralisation on FMD viruses type SAT 1 and SAT 2. Report of the Session of the Research Foot and Mouth Disease,Group of the Standing Committee of the European commision of Foot and Mouth disease, Paphos, Cyprus: FAO. 287–297.
58. Mateu MG, Hernandez J, Martinez MA, Feigelstock D, Lea S, et al. (1994) Antigenic heterogeneity of a foot-and-mouth disease virus serotype in the field is mediated by very limited sequence variation at several antigenic sites. J Virol 68: 1407–1417.
59. Mateu MG, Martinez MA, Capucci L, Andreu D, Giralt E, et al. (1990) A single amino acid substitution affects multiple overlapping epitopes in the major antigenic site of foot-and-mouth disease virus of serotype C. J Gen Virol 71 (Pt 3): 629–637.
60. Diez J, Davila M, Escarmis C, Mateu MG, Dominguez J, et al. (1990) Unique amino acid substitutions in the capsid proteins of foot-and-mouth disease virus from a persistent infection in cell culture. J Virol 64: 5519–5528.
61. Grazioli S, Fallacara F, Brocchi E (2004) Mapping of neutralising sites on FMD virus type Asia 1 and relationships with sites described in other serotypes. Report of the Session of the Research Foot and Mouth Disease,Group of the Standing Committee of the European commision of Foot and Mouth disease, Greece: FAO. 277–287.
62. Crowther JR, Rowe CA, Butcher R (1993) Characterization of monoclonal antibodies against a type SAT 2 foot-and-mouth disease virus. Epidemiol Infect 111: 391–406.
63. Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15: 305–308.
64. Risler JL, Delorme MO, Delacroix H, Henaut A (1988) Amino acid substitutions in structurally related proteins. A pattern recognition approach. Determination of a new and efficient scoring matrix. J Mol Biol 204: 1019–1029.
65. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876–4882.
66. Xiao C, Rossmann MG (2007) Interpretation of electron density with stereographic roadmap projections. J Struct Biol 158: 182–187.