Human Immunology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

journal homepage: www.elsevier.com/locate/humimm

Usefulness of the ElliPro epitope predictor program in defining the repertoire of HLA-ABC eplets

Rene J. Duquesnoy *, Marilyn Marrari

Thomas E. Starzl Transplantation Institute, University of Pittsburgh Medical Center, USA

ARTICLE INFO

Article history:
Received 1 February 2017
Revised 8 March 2017
Accepted 11 March 2017
Available online xxxx

Keywords:
HLA epitope
Antibody-verified epitope
Eplet
Epitope registry

ABSTRACT

HLA matching at the epitope level offers new opportunities to identify suitable donors for transplant patients. The International HLA Epitope Registry (www.Epregistry.com.br) describes for the various HLA loci, repertoires of eplets including those that correspond to epitopes experimentally verified with specific antibodies. There are also many eplets which have remained as theoretical entities because no informative antibodies have been found. Which of them have immunogenic potential or conversely, might be considered as non-epitopes that cannot elicit specific antibody responses? This question is important for the application of epitope-based HLA matching in clinical transplantation.

Correct predictions of B-cell epitopes on antigenic proteins are essential to the effective design of microbial vaccines and the development of specific antibodies used in immunotherapy and immunodiagnostics but prediction programs based on structural and physiochemical properties of amino acid residues are generally ineffective. Recent prediction programs based on three-dimensional structures of antigen-antibody complexes are more promising. One such program is called ElliPro developed by Ponomarenko.

This report describes studies demonstrating that ElliPro can predict alloantibody responses to HLA-ABC eplets. Antibody-verified eplets have amino acid residues with much higher ElliPro scores than eplets for which no specific antibodies have been found. The latter group includes residues with very low ElliPro scores; they appear to represent eplets that might be classified as non-epitopes.

In conclusion, ElliPro offers a new approach to characterize epitope repertoires that are clinically relevant in HLA matching.

1. Introduction

HLA antibodies are primary causes of allograft rejection and transplant failure and it is now widely accepted that they recognize specific epitopes on HLA antigens. Accordingly, HLA matching at the epitope level provides new opportunities to identify suitable donors for transplant patients. HLA epitopes can be structurally defined by eplets which are small configurations of polymorphic amino acids in antibody-accessible sequence locations of HLA molecules [1]. Eplets were originally conceived as theoretical entities with dominant roles in the specificities of HLA epitopes reactive with antibodies. The International HLA Epitope Registry (www.Epregistry.com.br) has eplet repertoires for HLA-A,B,C, HLA-DR,DQ,DP and MICA loci and they include eplets that correspond to epitopes experimentally verified with specific antibodies [2]. Many epitopes correspond to single eplets whereas others are defined by eplets paired with nearby amino acid configurations; such epitopes are referred to as eplet pairs [3,4]. The HLA Epitope Registry website includes a downloadable PDF file "EpiPedia of HLA" that summarizes the experimental evidence accumulated during the past 30 years about antibody-verified HLA epitopes.

The Registry has also many eplets which have remained as theoretical entities because no informative antibodies have been identified. The question must be raised which of them have immunogenic potential or conversely, might be considered as non-epitopes that cannot elicit specific antibody responses and therefore would play no role in HLA matching. Are there any rules that determine whether eplets are immunogenic or non-epitopes?

Abbreviations: CDR, Complementarity Determining Region; ElliPro, derived from Ellipsoid and Protrusion; HLA, Human Leukocyte Antigen; PDB, Protein Database.

* Corresponding author at: University of Pittsburgh Medical Center, Thomas E Starzl Biomedical Science Tower, Room W1552, Pittsburgh, PA 15213, USA.
E-mail address: Duquesnoyr@upmc.edu (R.J. Duquesnoy).

http://dx.doi.org/10.1016/j.humimm.2017.03.005
0198-8859/

A distinction between epitopes and non-epitopes on antigenic proteins is important for the effective design of antimicrobial vaccines and the generation of specific antibodies used in immunotherapy and immunodiagnostics [5,6]. Amino acid residues that are not a part of an antibody-antigen interface of an immune complex are referred to as non-epitope residues [7]. There are no intrinsic properties of amino acid residues that differentiate between epitopic and non-epitopic residues [8,9]. Original epitope prediction programs based on different physiochemical properties of amino acids including hydrophilicity and hydrophobicity, accessible surface area, solvent accessibility, secondary structure and chain flexibility have shown rather low accuracy rates [5,10]

Recent B-cell epitope prediction models consider also three-dimensional structures of complexes of protein antigens with antibodies determined by X-ray crystallography as summarized in various reviews [5,6,11,12]. This information together with intrinsic properties of amino acids in the protein sequences have led to new computational methods to predict epitopes.

In 2007, Ponomarenko et al. [13] analyzed a dataset of 62-antigen-antibody complexes selected from the Protein Database (PDB) of the website of the National Center for Biotechnology (http://www.ncbi.nlm.nih.gov/Structure). Their findings led to a structurally based epitope predictor tool [14]. ElliPro (derived from Ellipsoid and Protrusion) implements a modified version of Thornton's method [15] together with a residue clustering algorithm called MODELLER [16]. ElliPro approximates protein surface patches with ellipsoid structures [17] and incorporates Thornton's concept about identifying continuous epitopes in protein regions protruding from their globular surfaces [15]. Regions with high protrusion index values were shown to correspond to the experimentally determined continuous epitopes in various proteins used to generate antibodies [14]. ElliPro assigns a protrusion score to each residue on the antigenic protein and identifies epitope locations with clusters of residues with high protrusion scores. ElliPro was developed with a benchmark dataset of epitopes inferred from 3D structures of antibody-protein complexes and it showed good performance as an epitope predictor for a second dataset of 39 epitopes on protein structures [14]. ElliPro has been considered as one of the more complete tools for epitope prediction [11] and can give better predictions of non-epitopes on tertiary structures than three other recent programs [18].

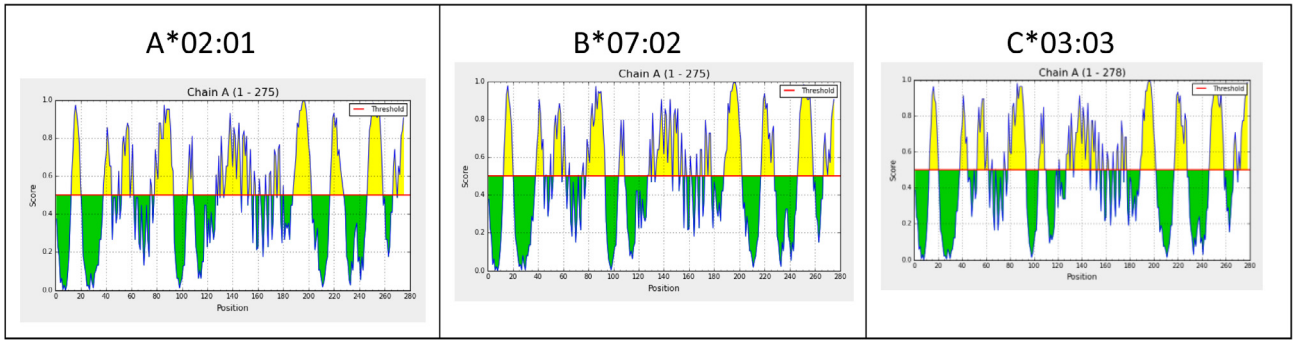

Fig. 1. ElliPro score charts for three HLA molecules: A*02:01, B*07:02 and C*03:03.

Table 1
ElliPro scores of polymorphic residues on A*02:01 that are monomorphic on other loci or which correspond to eplets that have or have not been antibody-verified.

Polymorphic Residue	ElliPro Score	Monomorphic at	Polymorphic Residue	ElliPro Score	Antibody-verified Eplet	Polymorphic Residue	ElliPro Score	Antibody-verified Eplet?
17R	0.940	BC	62G	0.459	62GE	9F	0.035	Not found
56G	0.836	BC	62G	0.459	62GK	43Q	0.658	Not found
82R	0.832	C	79G	0.703	79GT	44R	0.661	Not found
83G	0.888	C	80T	0.634	79GT	63E	0.290	Not found
90A	0.944	B	107W	0.591	107W	65R	0.566	Not found
109F	0.487	BC	127K	0.593	127K	66K	0.265	Not found
149A	0.852	BC	144K	0.736	144K	67V	0.190	Not found
150A	0.700	BC	142T	0.797	144TKH	70H	0.136	Not found
158A	0.544	C	145H	0.888	144TKH	73T	0.273	Not found
161E	0.588	BC	145H	0.888	145KHA	74H	0.186	Not found
166E	0.607	C	151H	0.809	150AAH	76V	0.527	Not found
167W	0.404	C	253Q	0.941	253Q	77D	0.345	Not found
186K	0.405	BC				81L	0.576	Not found
245A	0.369	C				95V	0.161	Not found
246A	0.540	BC				97R	0.055	Not found
276P	0.935	BC				99Y	0.034	Not found
						102D	0.143	Not found
						105S	0.669	Not found
						114H	0.120	Not found
						116Y	0.143	Not found
						152V	0.419	Not found
						156L	0.257	Not found
						163T	0.363	Not found
						184A	0.312	Not found
						193A	0.946	Not found
						194V	0.950	Not found
						207S	0.247	Not found

Please cite this article in press as: R.J. Duquesnoy, M. Marrari, Usefulness of the ElliPro epitope predictor program in defining the repertoire of HLA-ABC eplets, Hum. Immunol. (2017), http://dx.doi.org/10.1016/j.humimm.2017.03.005

Table 2
ElliPro scores for five class I alleles with antibody-verified eplets and polymorphic residues on eplets for which no specific antibodies have been found.

A*02:01 Residue	A*02:01 ElliPro	A*02:01 Eplet	A*24:02 Residue	A*24:02 ElliPro	A*24:02 Eplet	B*07:02 Residue	B*07:02 ElliPro	B*07:02 Eplet	B*44:02 Residue	B*44:02 ElliPro	B*44:02 Eplet	C*03:03 Residue	C*03:03 ElliPro	C*03:03 Eplet
62G	0.459	62GE	62E	0.598	62EE	65Q	0.540	65QIA	41T	0.922	41T	21H	0.404	21H
62G	0.459	62GK	65G	0.416	65GK	69A	0.299	69AA	69T	0.351	69TNT	76V	0.527	76VRN
79G	0.703	79GT	79R	0.726	79Rpr	69A	0.299	70IAQ	71T	0.344	69TNT	80N	0.650	80N
80T	0.634	79GT	80I	0.635	80I	71A	0.312	69AA	79R	0.726	82LR	163L	0.397	163LW
107W	0.591	107W	81A	0.647	81ALR	79R	0.726	80N	80T	0.634	80T	173K	0.780	173K
127K	0.593	127K	82L	0.830	82LR	80N	0.650	80N	81A	0.647	81ALR	194V	0.950	193PV
144K	0.736	144K	83R	0.878	82LR	81L	0.576	80N	82L	0.830	82LR	219W	0.905	219W
142T	0.797	144TKH	127K	0.593	127K	163E	0.365	163EW	83R	0.878	82LR			
145H	0.888	144TKH	142I	0.808	138MI	177D	0.716	180E	131S	0.747	131S	9Y	0.039	No Ab
145H	0.888	145KHA	144K	0.736	144K	178K	0.380	177DK	156D	0.259	156DA	11A	0.217	No Ab
151H	0.809	150AAH	144K	0.736	144KR	180E	0.455	180E	163L	0.397	163LS/G	66K	0.265	No Ab
253Q	0.941	253Q	151H	0.809	150AAH				167S	0.333	166ES	73T	0.273	No Ab
			166D	0.594	166DG	9Y	0.039	No Ab	199V	0.879	199V	77S	0.373	No Ab
9F	0.035	No Ab	167G	0.343	166DG	24S	0.069	No Ab				79R*	0.726	No Ab
43Q	0.658	No Ab				45E	0.315	No Ab	9Y	0.039	No Ab	91R	0.856	No Ab
44R*	0.661	No Ab	9S	0.043	No Ab	46E	0.469	No Ab	24T	0.075	No Ab	94I	0.260	No Ab
63E	0.290	No Ab	43Q	0.658	No Ab	62R*	0.498	No Ab	32L	0.084	No Ab	95I	0.165	No Ab
65R	0.566	No Ab	44R	0.661	No Ab	63N	0.283	No Ab	45K	0.275	No Ab	97R	0.055	No Ab
66K	0.265	No Ab	63E	0.290	No Ab	66I	0.248	No Ab	46E	0.469	No Ab	99Y	0.034	No Ab
67V	0.190	No Ab	66K	0.265	No Ab	67Y	0.178	No Ab	62R*	0.498	No Ab	103V	0.336	No Ab
70H	0.136	No Ab	67V	0.190	No Ab	70Q	0.142	No Ab	63E	0.290	No Ab	113Y	0.077	No Ab
73T	0.273	No Ab	70H	0.136	No Ab	74D	0.202	No Ab	65Q*	0.540	No Ab	114D	0.119	No Ab
74H	0.186	No Ab	73T	0.273	No Ab	76E*	0.557	No Ab	66I	0.248	No Ab	116Y	0.143	No Ab
76V*	0.527	No Ab	74D	0.202	No Ab	77S	0.373	No Ab	67S	0.222	No Ab	152E	0.421	No Ab
77D*	0.345	No Ab	76E*	0.557	No Ab	94T	0.266	No Ab	70N	0.171	No Ab	156L	0.257	No Ab
81L*	0.576	No Ab	77N	0.362	No Ab	95L	0.160	No Ab	74Y	0.166	No Ab	184H	0.397	No Ab
95V	0.161	No Ab	95L	0.160	No Ab	97S	0.035	No Ab	76E*	0.557	No Ab			
97R	0.055	No Ab	97M	0.032	No Ab	99Y	0.034	No Ab	77N	0.362	No Ab			
99Y	0.034	No Ab	99F	0.037	No Ab	103V	0.336	No Ab	94I	0.260	No Ab			
102D	0.143	No Ab	102D	0.143	No Ab	113H	0.090	No Ab	95I	0.165	No Ab			
105S	0.669	No Ab	105S	0.669	No Ab	114D	0.119	No Ab	97R	0.055	No Ab			
114H	0.120	No Ab	114H	0.120	No Ab	116Y	0.143	No Ab	99Y	0.034	No Ab			
116Y	0.143	No Ab	116Y	0.143	No Ab	11S	0.213	No Ab	103V	0.336	No Ab			
152V	0.419	No Ab	152V	0.419	No Ab	152E	0.421	No Ab	113Y	0.077	No Ab			
156L	0.257	No Ab	156Q	0.232	No Ab	156R	0.223	No Ab	114D	0.119	No Ab			
163T	0.363	No Ab	163T	0.363	No Ab	194I	0.943	No Ab	116D	0.119	No Ab			
184A	0.312	No Ab	194I	0.943	No Ab				11A	0.217	No Ab			
193A	0.946	No Ab							12M	0.381	No Ab			
194V*	0.950	No Ab							152V	0.419	No Ab			
207S	0.247	No Ab							194I	0.943	No Ab			
	p = 0.00039			p = 0.00026			p = 0.0011			p = 0.00046			p = 0.0021	

Alleles have the following PDB numbers: A*02:01 (1JF1), A*24:02 (2BCK), B*07:02 (5EO1), B*44:02 (1M6O) and C*03:03 (1EFX).
No Ab means that no specific antibodies have (yet) been identified to corresponding eplets.

Table 3
ElliPro scores of polymorphic residues that are monomorphic for other class I loci.

Polymorphic Residue	Monomorphic at	ElliPro Score	Polymorphic Residue	Monomorphic at	ElliPro Score	Polymorphic Residue	Monomorphic at	ElliPro Score	Polymorphic Residue	Monomorphic at	ElliPro Score
1G	AB	0.474	52I	AC	0.457	144Q	B	0.727	183E	C	0.280
6R	ABC	0.005	56G	BC	0.836	145R	B	0.986	184P	B	0.341
12V	AC	0.387	59Y	AC	0.487	147W	A	0.426	186K	BC	0.405
14R	AB	0.760	69R	C	0.380	149A	BC	0.852	189M	A	0.608
16G	AB	0.963	82R	C	0.832	150A	BC	0.700	189V	BC	0.608
17R	BC	0.940	83G	C	0.888	151R	BC	0.826	193P	BC	0.941
19E	BC	0.723	90A	B	0.944	158A	C	0.544	199A	AC	0.879
21R	AB	0.404	91G	AB	0.856	161E	BC	0.588	207G	BC	0.247
24A	A	0.078	103V	AC	0.336	166E	C	0.607	211A	AB	0.017
30D	AC	0.010	105P	BC	0.645	167W	C	0.404	219R	AB	0.905
32Q	AC	0.084	107G	BC	0.632	171Y	AC	0.301	245A	C	0.369
35R	AB	0.283	109L	BC	0.487	173E	AB	0.780	246A	BC	0.540
41A	AC	0.886	109F	BC	0.487	177E	A	0.737	248V	AB	0.846
43P	BC	0.658	127N	BC	0.537	178T	AC	0.380	253E	BC	0.946
45G	C	0.422	131R	AC	0.819	180Q	AC	0.475	261V	AB	0.261
46E	AC	0.469	138M	A	0.921	182T	A	0.331	273R	AB	0.813
49A	AB	0.381	138T	B	0.921	182A	BC	0.331	275E	AB	0.773
52V	C	0.457	143T	C	0.580	183D	AB	0.280	276P	BC	0.935

This report describes a study how the application of ElliPro can distinguish between HLA-ABC eplets that are immunogenic and eplets that might be considered non-epitopes.

2. Methods and results

ElliPro is available on the website-based Immune Epitope Database (www.iedb.org) [19,20] and its analysis is based on three-dimensional protein structures. The website of the National Center for Biotechnology (http://www.ncbi.nlm.nih.gov/Structure) has many HLA molecular structures. Their Protein Data Bank (PDB) numbers can be entered as the first step of the ElliPro analysis. Class I molecules have three components: α chain, β2-microglobulin and bound peptide and the polymorphic α chain has always been selected for ElliPro analysis.

For example, one of the A*02:01 molecules has 1JF1 as its PDB. ElliPro generates lists of predicted linear epitopes on the α chain as determined from "minimum scores" and predicted discontinuous epitopes based on "maximum distance" specifications. With the default settings in the program, there are 9 predicted linear epitopes with numbers ranging from 5 to 28 residues along with 7 predicted discontinuous epitopes with numbers of residues ranging from 3 to 51 (details are on www.iedb.org). Testing other HLA alleles generated similarly complex epitope descriptions and we have concluded that these predictions cannot be applied to HLA epitopes defined by eplets and eplet pairs. Therefore, we chose another approach to analyze the ElliPro data.

ElliPro generates scores for each amino acid residue in the protein sequence; they are between 0.001 and 1.000. These scores are not the same for a given residue in the protein chain but they have wide ranges determined by their sequence location and the contribution of nearby residues in ellipsoid structures. For instance, the A*01:01 α chain has 22 glycine and 25 arginine residues; their ElliPro scores range from 0.033 to 0.974 and 0.004 to 0.949, respectively.

Fig. 1 shows the picturized ElliPro score charts for three HLA molecules: A*02:01, B*07:02 and C*03:03. The upper part shows the sequence locations of residues with high ElliPro scores which as predicted by Ponomarenko [14] would contribute to antibody-reactive epitopes. The lower part shows the residue locations with low ElliPro scores. It can be readily seen that the profiles are similar for different class I alleles thereby suggesting that epitopes reside in similar sequence locations. Fig. 1 has ElliPro scores for both monomorphic and polymorphic residues, the latter being essential components of epitopes.

Antibody responses to HLA mismatches are specific for epitopes which must have minimally one polymorphic amino acid residue present on the immunizing allele but absent on any allele of the antibody producer. Such nonself residue on a mismatched eplet can be interpreted as the "driving force" of the antibody response specific for that epitope.

The A*02:01 allele has many sequence locations with polymorphic residues and two rules have been applied to determine if these residues are nonself or self. First, a polymorphic residue on A*02:01 can be monomorphic for the other class I loci of the antibody producer, in this case HLA-B and/or HLA-C. Table 1 shows 16 such residues on A*02:01. For instance, polymorphic residue 90A is monomorphic for HLA-B whereas polymorphic 149A is monomorphic for HLA-B and HLA-C. Such residues cannot be considered as driving forces of the HLA antibody response although they might be part of an epitope specifically recognized by antibody.

Second, polymorphic residues might be shared with other alleles controlled by the same locus and/or additional loci. For instance, 127K is shared by A2, A23, A24, A68 and A69 alleles and cannot be considered a nonself driving force on A*02:01 for antibody producers who type for A23, A24, A68 and/or A69. Similarly, residue 62G shared by A2, B57 and B58 cannot be presented by A*02:01 as a driving force for antibody producers who type for B57 and/or B58. These considerations reflect the well-recognized concept that an epitope-specific antibody response is influenced by the HLA type of the antibody producer.

The HLA Epitope Registry has ten eplets on A*02:01 that have been experimentally verified with specific antibodies. They are 62GE, 62GK, 79GT, 107W, 127K, 144TKH, 144K, 145KHA, 150AAH and 253Q; each eplet has at least one nonself residue which might represent the driving force of the eplet-specific antibody response. Table 1 shows the ElliPro scores for such polymorphic residues for the antibody-verified eplets on A*02:01; their median value is 0.736, range = 0.459–0.941. Conversely, A*02:01 has a second group of polymorphic residues for which no corresponding antibody-verified epitopes have been found; their median ElliPro score is 0.273, range = 0.034–0.950.

This ElliPro analysis was done with more than thirty molecular models of class I alleles and Table 2 shows data for five representative alleles along with PDB numbers: A*02:01 (1JF1), A*24:02 (2BCK), B*07:02 (5EO1), B*44:02 (1M6O) and C*03:03 (1EFX). The table does not display the residues that are monomorphic for other class I loci. For each allele, the ElliPro scores were signifi-

Please cite this article in press as: R.J. Duquesnoy, M. Marrari, Usefulness of the ElliPro epitope predictor program in defining the repertoire of HLA-ABC eplets, Hum. Immunol. (2017), http://dx.doi.org/10.1016/j.humimm.2017.03.005

Table 4
ElliPro scores for the current repertoire of antibody-verified class I eplets.

Polymorphic Residue	Antibody Verified Eplet	ElliPro Score	Polymorphic Residue	Antibody Verified Eplet	ElliPro Score	Polymorphic Residue	Antibody Verified Eplet	ElliPro Score	Polymorphic Residue	Antibody Verified Eplet	ElliPro Score
1C	1C	0.474	71T	69TNT	0.344	107W	107W	0.591	163E	163EW	0.365
21H	21H	0.404	71S	71SA	0.276	127K	127K	0.593	163L	163LS/G	0.397
41T	41T	0.922	71T	73TTS	0.344	131S	131S	0.747	163L	163LW	0.397
44K	44KM	0.642	73A	73AN	0.273	138K	138K	0.921	163R	163R	0.489
45T	45RT	0.372	76A	76ANT	0.558	142I	138MI	0.808	163R	163RG	0.489
46A	45RMA	0.378	76E	76ESI	0.557	142T	144TKH	0.797	163R	163RW	0.489
56R	56R	0.836	76V	76VRN	0.527	144K	144K	0.736	166D	166DG	0.594
62E	62EE	0.598	77D	71ATD	0.345	144K	144KR	0.736	167G	166DG	0.343
62G	62GE	0.459	79G	79GT	0.703	145L	144QL	0.886	167S	166ES	0.333
62G	62GK	0.459	79R	138Mpr	0.726	145H	144TKH	0.888	173K	173K	0.780
62G	62GRN	0.459	80T	79RT	0.634	145H	145KHA	0.888	177K	177KT	0.748
62L	62LQ	0.486	80I	80I	0.635	149T	145RT	0.852	177D	180E	0.716
62Q	62QEpr	0.473	80K	80K	0.648	149T	149TAH	0.852	178K	177DK	0.380
62R	62RR	0.498	80N	80N	0.650	150V	150VHA	0.679	180E	180E	0.455
65G	65GK	0.416	80T	80TLR	0.634	151H	150AAH	0.809	194L	193PL	0.936
65Q	65QIA	0.540	81A	81ALR	0.647	152A	151AHA	0.455	194V	193PV	0.950
69A	69AA	0.299	81L	81ALR	0.576	156D	156DA	0.259	199V	199V	0.879
69A	70IAQ	0.299	82L	82LR	0.830	158T	158T	0.544	219W	219W	0.905
69T	69TNT	0.351	83R	82LR	0.878	158V	158V	0.588	253Q	253Q	0.941
71A	69AA	0.312	90D	90D	0.957	161D	161D	0.547			

Table 5
Sequence locations and <0.300 ElliPro scores of 81 polymorphic residues for which no eplet-specific antibodies have been found.

Polymorphic Residue	<0.300 ElliPro Score	Antibody Verifiable Eplet	Polymorphic Residue	<0.300 ElliPro Score	Antibody Verifiable Eplet	Polymorphic Residue	<0.300 ElliPro Score	Antibody Verifiable Eplet	Polymorphic Residue	<0.300 ElliPro Score	Antibody Verifiable Eplet
4F	0.223	No?	66N	0.244	No?	95F	0.170	No?	114N	0.120	No?
6K	0.005	No?	67S	0.222	No?	95I	0.165	No?	114D	0.119	No?
9S	0.043	No?	67C	0.208	No?	95V	0.161	No?	114E	0.118	No?
9Y	0.039	No?	67V	0.190	No?	95L	0.160	No?	114Q	0.118	No?
9H	0.037	No?	67F	0.187	No?	97R	0.055	No?	114R	0.099	No?
9T	0.037	No?	67Y	0.178	No?	97W	0.047	No?	116H	0.146	No?
9D	0.036	No?	67M	0.176	No?	97I	0.046	No?	116L	0.146	No?
9F	0.035	No?	70S	0.178	No?	97T	0.040	No?	116Y	0.143	No?
11A	0.217	No?	70N	0.171	No?	97V	0.036	No?	116S	0.133	No?
11S	0.213	No?	70Q	0.142	No?	97S	0.035	No?	116F	0.127	No?
24T	0.075	No?	70K	0.141	No?	97N	0.033	No?	116D	0.119	No?
24S	0.069	No?	70H	0.136	No?	97M	0.032	No?	156L	0.257	No?
30G	0.010	No?	73I	0.273	No?	99C	0.038	No?	156Q	0.232	No?
31S	0.078	No?	73T	0.273	No?	99S	0.038	No?	156R	0.223	No?
32L	0.084	No?	74D	0.202	No?	99F	0.037	No?	156W	0.176	No?
35Q	0.197	No?	74N	0.200	No?	99Y	0.034	No?	207S	0.247	No?
45K	0.275	No?	74H	0.186	No?	102D	0.143	No?	211T	0.017	No?
63E	0.290	No?	74Y	0.166	No?	102H	0.143	No?	261M	0.261	No?
63N	0.283	No?	94T	0.266	No?	113H	0.090	No?			
66K	0.265	No?	94I	0.260	No?	113Y	0.077	No?			
66I	0.248	No?	95W	0.214	No?	114H	0.120	No?			

cantly higher for the antibody-verified eplets than for the polymorphic residues with corresponding eplets for which as indicated no specific antibodies have been found. Mann-Whitney U statistical analysis showed that all p values were 0.002 or lower.

For each group without antibody-verified epitopes, the ElliPro scores were sorted from highest to lowest values. Several residues have ElliPro scores that are comparable to those with antibody-verified epitopes. Could such residues define other epitopes that have not yet been identified with informative specific antibodies? Some residues such as 194V and 44R on A*02:01 are marked with asterisks because they appear to be driving forces for different antibody-verified epitopes presented by other alleles. Conversely, Table 2 shows many other residues have very low ElliPro scores; do they reflect non-epitopes?

These questions have been addressed by determining the ElliPro scores for the complete sequence of polymorphic residues. This analysis was possible because data were available for more than thirty class I alleles with molecular structures determined by X-ray crystallography. We noted that the ElliPro scores of polymorphic residues had generally no more than 15% variations between the different alleles. Besides the thirty alleles with molecular structure information, the kits for antibody-binding assays have more than sixty additional alleles. Most of them have the same polymorphic residues and assignments of ElliPro scores were based on comparisons with alleles with defined molecular structures.

The ElliPro scores have been arranged into four groups of polymorphic residues. Table 3 shows polymorphic residues that are monomorphic for one or two other class I loci. Although many of them have high ElliPro scores, they cannot be considered as driving forces in epitope-specific antibody responses because these residues are self on one or more loci of the antibody producer. These residues can nevertheless participate in epitopes defined by eplets and eplet pairs.

Please cite this article in press as: R.J. Duquesnoy, M. Marrari, Usefulness of the ElliPro epitope predictor program in defining the repertoire of HLA-ABC eplets, Hum. Immunol. (2017), http://dx.doi.org/10.1016/j.humimm.2017.03.005

Table 6
Sequence locations, >0.300 ElliPro scores and mismatch probabilities of 49 polymorphic residues that have corresponding eplets for which specific antibodies have not been found.

Polymorphic Residue	>0.300 ElliPro Score	Probability Residue Mismatch	Residue Frequency	Antibody Verifiable Eplet	Polymorphic Residue	>0.300 ElliPro Score	Probability Residue Mismatch	Residue Frequency	Antibody Verifiable Eplet
16S	0.963	3.2%	3.3%	Unlikely	12M	0.381	22.9%	64.5%	Possible
17S	0.940	1.9%	1.9%	Unlikely	14W	0.760	9.9%	11.1%	Possible
19K	0.723	0.3%	0.3%	Unlikely	44R	0.661	19.6%	26.9%	Possible
43R	0.658	2.5%	2.6%	Unlikely	45E	0.315	24.1%	40.5%	Possible
56E	0.836	0.1%	0.1%	Unlikely	46E	0.469	11.9%	86.2%	Possible
59H	0.487	0.0%	0.0%	Unlikely	49E	0.381	9.9%	11.1%	Possible
77G	0.371	0.1%	0.1%	Unlikely	65R	0.566	9.3%	89.6%	Possible
103M	0.333	0.1%	0.1%	Unlikely	77D	0.345	23.1%	36.3%	Possible
152R	0.448	0.7%	0.7%	Unlikely	77N	0.362	24.4%	57.9%	Possible
152T	0.420	0.4%	0.4%	Unlikely	77S	0.373	23.2%	63.2%	Possible
152W	0.448	1.3%	1.3%	Unlikely	91R	0.856	6.7%	7.2%	Possible
162D	0.503	0.0%	0.0%	Unlikely	103L	0.322	11.8%	86.3%	Possible
163T	0.363	4.1%	95.7%	Unlikely	105S	0.669	22.0%	67.3%	Possible
170G	0.628	1.5%	1.5%	Unlikely	143S	0.580	6.2%	6.6%	Possible
184R	0.321	0.8%	0.8%	Unlikely	147L	0.426	22.7%	35.0%	Possible
186R	0.405	0.7%	0.8%	Unlikely	152A	0.455	24.9%	47.5%	Possible
193L	0.932	2.4%	2.5%	Unlikely	152E	0.421	13.7%	83.6%	Possible
245T	0.369	0.1%	0.1%	Unlikely	152V	0.419	23.4%	62.8%	Possible
245V	0.369	3.3%	3.4%	Unlikely	171H	0.301	8.7%	9.6%	Possible
248M	0.846	2.8%	2.8%	Unlikely	184A	0.312	25.0%	48.2%	Possible
253K	0.933	0.1%	0.1%	Unlikely	184H	0.397	21.8%	68.0%	Possible
266C	0.489	1.5%	1.6%	Unlikely	193A	0.946	25.0%	52.0%	Possible
					194I	0.943	6.5%	93.0%	Possible
					246S	0.540	14.4%	17.4%	Possible
					275G	0.773	11.9%	13.7%	Possible
					275K	0.773	11.2%	12.9%	Possible
					276L	0.935	23.3%	36.9%	Possible

Table 7
ElliPro scores of critical contact residues in antibody-verified eplet pairs.

Eplet Pair	Residue	ElliPro
44RT + 69TNT	69T	0.351
62QIA + 76ESN	76E	0.557
69AA + 65QI	65Q	0.540
69AA + 76E	76E	0.557
69TNT + 80N	80N	0.650
79GT + 90D	90D	0.957
80I + 69TNT	69T	0.351
80I + 90A	90A	0.944
80K + 14R	14R	0.760
82LR + 138M	138M	0.921
82LR + 138T	138T	0.921
82LR + 144QR	144Q	0.727
82LR + 145R	145R	0.986
82LR + 145RA	145R	0.986
82LR + 90A	90A	0.944
131S + 163T	163T	0.385
138MI + 79GT	79G	0.703
144KR + 127K	127K	0.593
144KR + 151H	151H	0.809
163EW + 66I	65Q	0.540
163EW + 73TE	76E	0.557
163LW + 65QIT	65Q	0.540
	Mean	0.695 ± 0.213

Table 4 shows for 79 antibody-verified eplets polymorphic residues considered as driving forces for antibody responses. Their ElliPro scores are much higher than the remaining 130 residues listed in Tables 5 and 6 and for which no corresponding antibody-verified epitopes have been found. Their median scores were 0.576 and 0.223 respectively, and the Mann-Whitney *U* test yielded an extremely significant p value of 0.011×10^{-12}.

Table 4 has 75/79 antibody-verified epitopes for which their polymorphic residues have ElliPro scores above an arbitrarily chosen cut-off value of 0.300. Four eplets 69AA, 70IAQ, 71S, 156DA had lower values which were still above 0.250. Nevertheless, the 0.300 cut-off value seems a reasonable estimate although 0.250–0.300 range represents an indeterminate range.

Table 5 shows 81 polymorphic residues with lower scores; they belong to eplets which might be considered non-epitopic because they cannot function as driving forces of epitope-specific antibody responses. Eplets fully defined by non-epitopic residues appear clinically irrelevant as mismatches.

Table 6 shows 49 residues with >0.300 ElliPro scores; they might represent eplets that can might be antibody-verifiable if such experimental evidence becomes available. From residue frequencies estimated in the general population one can calculate the probability of their chances of being a mismatch. The left part of Table 6 shows residues with <5% probabilities; it seems highly unlikely that they will be identified as driving forces of eplet-specific antibody responses. The right part has the remaining residues that determine eplets as candidates for antibody-verified epitopes.

The HLA Epitope Registry lists 22 class I eplet pairs that are recorded as antibody-verified epitopes (Table 7). Each epitope is determined by an eplet reacting with one CDR of antibody together with a second critical residue configuration within a 15-Å radius that contacts another CDR of antibody. Their mean ElliPro score is 0.695 ± 0.213 which is comparable to the 0.597 ± 0.205 value for antibody-verified eplets. Eplet pairs represent examples of discontinuous epitopes and these findings indicate that critical contact residues paired to eplets have relatively high ElliPro scores.

3. Discussion

A very recent invited overview in Transplantation addressed the question: Are we ready for epitope-based HLA matching in clinical organ transplantation? [21]. HLA epitopes are defined by antibody reactivity patterns with HLA alleles representing strings of eplets determined with the HLAMatchmaker algorithm [1]. The HLA Epi-

Please cite this article in press as: R.J. Duquesnoy, M. Marrari, Usefulness of the ElliPro epitope predictor program in defining the repertoire of HLA-ABC eplets, Hum. Immunol. (2017), http://dx.doi.org/10.1016/j.humimm.2017.03.005

tope Registry has extensive repertoires of antibody-verified eplets and they constitute the principal basis for epitope-based matching. However, what do we do with many other eplets for which no specific antibodies have been found and which ones reflect non-epitopes incapable of eliciting antibody responses? Should they be excluded from epitope-based matching strategies but how can we distinguish between the epitope and non-epitope nature of eplets? Continuously looking for specific antibodies but never finding them would be an ineffective method.

Theoretical concepts about structures and intrinsic properties of amino acids in protein epitopes have led to numerous epitope predictor algorithms aimed to distinguish between epitopes and non-epitopes on antigenic proteins. Unfortunately, none of these methods are very reliable regarding the development of microbial vaccines and specific antibodies for immunotherapy and immunodiagnostics [5,10]. We have tested six linear and two discontinuous B-cell epitope prediction programs available on the website-based Immune Epitope Database (www.iedb.org) but none of them (data not shown) except ElliPro, yielded statistically significant data that distinguish antibody-verified eplets from eplets for which no specific antibodies have been identified. Furthermore, the validity of ElliPro could only be demonstrated by focusing on polymorphic amino acid residues in eplets.

Epitope prediction programs appear to work best within the same protein family with common physiochemical and structural properties [22]. The class I system of HLA-ABC is an ideal model for evaluating epitope prediction programs because of its high degree of structural similarity. Moreover, there is considerable information about HLA antibodies analyzed in sensitive antibody assays with large single allele panels and immunizer-specific epitopes can be identified from the HLA type of the antibody producer. The HLA Epitope Registry has recorded many eplets and eplet pairs as antibody-verified epitopes and with ElliPro we have addressed the question why many other eplets have remained at the theoretical stage because no specific antibodies have been identified.

Our analysis focused on the ElliPro scores of polymorphic residues in eplets that were considered the driving forces for specific antibody responses. In other words, ElliPro scores appear to reflect the ability of eplets to induce specific antibody responses. The data show that the ElliPro scores for antibody-verified eplets were much higher than polymorphic residues for which no eplet-specific antibodies have been found and the difference was statistically highly significant. These findings suggest that for an eplet to induce a specific antibody response it must have at least one non-self residue with a high ElliPro score.

The HLA Epitope Registry has currently more than twenty antibody-verified epitopes that are defined by eplets paired with other configurations. In our experience these configurations generally consist of shared self-residues present on one or more alleles of the antibody producer [23,24]. Their ElliPro scores are comparable to those for the antibody-verified eplets. These configurations serve as critical contact sites for CDRs that are different from CDRs reacting with eplets. The increased binding of such CDRs appears to be a result of the well-recognized affinity maturation process of the antibody response. Our findings suggest that residues with high ElliPro scores might play an important role in the affinity maturation of such CDRs. Our structural epitope model for HLA has always been based on the participation of any residues within a 15 Å radius of the immunizing eplet [25] and accordingly, structural epitopes must have residues with relatively high ElliPro scores.

Many residues without corresponding eplet-specific antibodies give ElliPro scores below the 0.300 cut-off. We postulate that such eplets might be considered as non-epitopes incapable of inducing specific antibodies. We have identified 59 such eplets in the HLA Epitope Registry and their ElliPro scores will be recorded. Eventually, these non-epitopic eplets will be removed from the Registry.

The remaining residues had ElliPro scores higher than 0.300 cut-offs. Such residues will offer some guidance in identifying eplets as possible candidates for experimental verification with specific antibodies. It should be noted that such antibodies seem unlikely for eplets with low probabilities of being mismatched. Moreover, successful detection of informative antibodies in sera might also be masked by other HLA antibodies induced by immunizing alleles.

This study suggests that the ElliPro algorithm is a useful tool to distinguish between class I eplets as antibody-verified epitopes or as non-epitopes. ElliPro scores however, cannot be used as estimates of eplet immunogenicity. Other epitope prediction programs should be considered to further assess the immunogenicity of HLA eplets and the clinical relevance of eplet-based HLA matching in transplantation. A similar ElliPro analysis is being performed for class II epitopes encoded by HLA-DR, DQ and DP loci and the results will be presented soon.

Conflict of interest statement

The authors have no conflicts of interest to disclose as described by Human Immunology

References

[1] R.J. Duquesnoy, A structurally based approach to determine hla compatibility at the humoral immune level, Human Immunol. 67 (2006) 847–862.
[2] R.J. Duquesnoy, M. Marrari, L.C.D. da M. Sousa, J.R.P. de M. Barroso, K.M. de SE Aita, A.S. da Silva, S.J.H. do Monte, 16th IHIW: a website for the antibody-defined HLA epitope Registry, Int. J. Immunogen. 40 (2013) 54–59.
[3] R.J. Duquesnoy, A. Mulder, M. Askar, M. Fernandez-Vina, F.H.J. Claas, HLAMatchmaker-based analysis of human monoclonal antibody reactivity demonstrates the importance of an additional contact site for specific recognition of triplet-defined epitopes, Hum. Immunol. 66 (2005) 749–761.
[4] M. Marrari, J. Mostecki, A. Mulder, I. Balazs, F. Claas, R. Duquesnoy, Human monoclonal antibody reactivity with HLA class I epitopes defined by pairs of mismatched eplets and self eplets, Transplantation 90 (2010) 1468–1472.
[5] J.A. Greenbaum, P.H. Andersen, M. Blythe, H.H. Bui, R.E. Cachau, J. Crowe, M. Davies, A.S. Kolaskar, O. Lund, S. Morrison, B. Mumey, Y. Ofran, J.L. Pellequer, C. Pinilla, J.V. Ponomarenko, G.P. Raghava, M.H. van Regenmortel, E.L. Roggen, A. Sette, A. Schlessinger, J. Sollner, M. Zand, B. Peters, Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools, J. Mol. Recogn.: JMR. 20 (2007) 75–82.
[6] Y. El-Manzalawy, V. Honavar, Recent advances in B-cell epitope prediction methods, Immun. Res. 6 (Suppl 2) (2010) S2.
[7] P. Haste Andersen, M. Nielsen, O. Lund, Prediction of residues in discontinuous B-cell epitopes using protein 3D structures, Protein Sci.: a publication of the Protein Society. 15 (2006) 2558–2567.
[8] J. Sun, T. Xu, S. Wang, G. Li, D. Wu, Z. Cao, Does difference exist between epitope and non-epitope residues? Analysis of the physicochemical and structural properties on conformational epitopes from B-cell protein antigens, Immun. Res. 7 (2011) 1–11.
[9] I. Sela-Culang, V. Kunik, Y. Ofran, The structural basis of antibody-antigen recognition, Front Immunol. 4 (2013) 302.
[10] M.J. Blythe, D.R. Flower, Benchmarking B cell epitope prediction: underperformance of existing methods, Protein Sci.: a publication of the Protein Society. 14 (2005) 246–248.
[11] R.E. Soria-Guerra, R. Nieto-Gomez, D.O. Govea-Alonso, S. Rosales-Mendoza, An overview of bioinformatics tools for epitope prediction: implications on vaccine development, J. Biomed. Inform. 53 (2015) 405–414.
[12] B. Yao, D. Zheng, S. Liang, C. Zhang, Conformational B-cell epitope prediction on antigen protein structures: a review of current algorithms and comparison with common binding site prediction methods, PLoS ONE 8 (2013) e62249.
[13] J.V. Ponomarenko, P.E. Bourne, Antibody-protein interactions: benchmark datasets and prediction tools evaluation, BMC Struct. Biol. 7 (2007) 64.
[14] J. Ponomarenko, H.H. Bui, W. Li, N. Fusseder, P.E. Bourne, A. Sette, B. Peters, ElliPro: a new structure-based tool for the prediction of antibody epitopes, BMC Bioinformatics 9 (2008) 514.
[15] J. Thornton, M. Edwards, W. Taylor, D. Barlow, Location of 'continuous' antigenic determinants in the protruding regions of proteins, EMBO J. 5 (1986) 409–413.
[16] N. Eswar, M. Marti-Renom, M. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali, Comparative protein structure modeling with modeller, Curr. Prot. Bioinf. 5 (2006) 1–30.
[17] W. Taylor, J. Thornton, W. Turnell, An ellipsoidal approximation of protein shape, J. Mol. Graph. 1 (1983) 30–38.

Please cite this article in press as: R.J. Duquesnoy, M. Marrari, Usefulness of the ElliPro epitope predictor program in defining the repertoire of HLA-ABC eplets, Hum. Immunol. (2017), http://dx.doi.org/10.1016/j.humimm.2017.03.005

[18] J. Ren, Q. Liu, J. Ellis, J. Li, Tertiary structure-based prediction of conformational B-cell epitopes through B factors, Bioinformatics (Oxford, England) 30 (2014) i264–273.
[19] Y. Kim, J. Ponomarenko, Z. Zhu, D. Tamang, P. Wang, J. Greenbaum, C. Lundegaard, A. Sette, O. Lund, P.E. Bourne, M. Nielsen, B. Peters, Immune epitope database analysis resource, Nucleic Acids Res. 40 (2012) W525–530.
[20] N. Salimi, W. Fleri, B. Peters, A. Sette, The immune epitope database: a historical retrospective of the first decade, Immunology 137 (2012) 117–123.
[21] R.J. Duquesnoy, Invited overview: are we ready for epitope-based hla matching in clinical organ transplantation?, Transplantation (2017) (In Press)
[22] E. Kozlova, B. Viart, R. de Avila, L. Felicori, C. Chavez-Olortegui, Classification epitopes in groups based on their protein family, BMC Bioinf. 16 (Suppl 19) (2015) S7.
[23] R.J. Duquesnoy, Human leukocyte antigen epitope antigenicity and immunogenicity, Curr. Opin. Organ Transpl. 19 (2014) 428–435.
[24] R.J. Duquesnoy, Reflections on HLA Epitope-Based Matching for Transplantation, Front. Immunol. 7 (2016) 469.
[25] R.J. Duquesnoy, Clinical usefulness of HLA matchmaker in HLA epitope matching for organ transplantation, Curr. Opin. Immunol. 20 (2008) 594–601.

Please cite this article in press as: R.J. Duquesnoy, M. Marrari, Usefulness of the ElliPro epitope predictor program in defining the repertoire of HLA-ABC eplets, Hum. Immunol. (2017), http://dx.doi.org/10.1016/j.humimm.2017.03.005