THE UNIVERSITY OF RHODE ISLAND



Immunogenicity Analysis of Chinese Hamster Ovary (CHO) **Host Cell Protein Contaminants in Therapeutic Protein Formulations**



Andres H. Gutiérrez¹, Lenny Moise^{1,2}, Frances Terry², Chris Bailey-Kellogg³, William D. Martin², Anne S. De Groot^{1,2}

¹Institute for Immunology and Informatics, University of Rhode Island, Providence, RI; ²EpiVax, Inc., Providence, RI; ³Dartmouth College, Hanover, NH

Abstract

Concerns over immune responses to contaminating host cell proteins (HCPs) in therapeutics have recently emerged. For example, the presence of Chinese hamster ovary (CHO)-derived HCPs contributed to the cancellation of two phase III clinical trials in 2012 for Inspiration's IB1001, a recombinant factor IX produced in CHO cells. These trials were cancelled due to the development of anti-CHO antibodies at higher levels than expected in patients treated with a protein drug. Publication of the CHO-K1 genome and transcriptome offers an opportunity to analyze and understand the immunogenic potential of CHO HCPs. We analyzed the CHO-K1 genome, the CHO-K1 transcriptome, published CHO proteins identified in biotherapeutics, and mouse secreted proteins downloaded from the LOCATE and UniProtKB/Swiss-Prot databases. These datasets were screened for potential CD4⁺ T cell immunogenicity using the epitope-mapping algorithm EpiMatrix and overall protein immunogenicity scores calculated. Human homologs discovered by BLAST were similarly screened for T-cell epitope content; predicted immunogenicity was then adjusted for human homology. Further study is required to assess the potential for autoimmune responses, as several widely publicized cases of severe adverse immune responses to autologous therapeutic proteins are published and the FDA and EMA take the potential for such effects very seriously. Finally, we built a website (CHOPPI) to give protein manufacturers free access to information on potential immunogenicity of CHO-K1 sequences. We look forward to its further development in collaboration with protein therapeutics industry, as this and additional means of performing immunogenicity analysis continue to be refined.

Immunogenicity Analysis

Immunogenic potential of validated HCP contaminants was assessed by epitope density and ranked on an immunogenicity scale, which was developed to evaluate therapeutic proteins of interest against known immunogenic and non-immunogenic proteins.



CHOPPI Web Tool

CHOPPI (CHO Protein Predicted Immunogenicity) is a web tool developed to address two questions:

(1) What are the possible contaminants? CHOPPI integrates the genome of the CHO-K1 cell line, a CHO transcriptome, a mouse secretome, and SignalP-based signal peptide prediction (indicating likely secretion).

(2) Which are likely to be immunogenic? CHOPPI assesses predicted T cell epitope content by EpiMatrix, as well as compares the proteins to their human homologs and their predicted epitopes to predicted epitopes within the human genome.

CHO Cell Protein

- CHO-derived Host Cell Protein (HCP) contaminants in biotherapeutics can potentially generate immune responses.
- Of significant concern are immune responses that may be crossreactive with human T cell epitopes.



Figure 3. EpiMatrix Protein Immunogenicity Scale. Validated CHO HCP contaminants with scores ≥ 20 represent approximately the top 20% of the distribution. High-scoring proteins are considered potentially immunogenic. Conversely, proteins scoring below -20 are categorized as possessing "low" immunogenic potential. All scores normalized for protein length

Immunogenicity comparison			
minunogeneity companyon	Immunodeni	city com	narison
	manogem	Ulty Com	panson

CHOPPI					e	The	11-			(- 1
CHO protoip pr	odictod immuno	dopicity		Carl	(C)	1	US	ers ma	ay re	triev	e lists
Cho protein pr	a subscription		1103	of proteins from the							
					[home] [help]		, IO aor	nomo	200	ordinc
							U	i yei		acc	orung
Welcome to CHOPPI, w proteins in CHO-based p	hich identifies potential protein production. Plea	immunogenicity risks from ho se see the help page for furth	ost contaminant ner information.				to s	specifi	ed fil	ters.	
Look up			Show 100 ÷ entries						S	earch:	
Enter an id or accessior	number (gi or gb), a pr	otein name, or a portion of ar	protein 🍦	immuno. 🔻	transcr. 🛊	secr. 🛊	SignalP 🛊	validated 🛊	human 🔶	9mers 🛊	unique 9mers
Ex: "interleukin" or "EGW	07755" or "mkflsardfhplafl	glmla"	Collagen alpha-1(III) chain	35.79	100	96	no	100	92	228	60
			Lysosomal protective protein	34.4	100	92	yes	100	87	109	28
			Glutathione S-transferase P	25.16	99	0	no	100	87	45	14
			Annexin A1	24.87	100	0	no	100	89	75	20
look up		Annexin A2	14.87	100	0	no	100	97	39	4	
		Nuclear prelamin A recognition factor	-3.77	100	0	no	100	83	54	13	
		Actin, cytoplasmic 2	-9.36	100	0	no	100	100	39	0	
Filter			Serine protease HTRA1	-13.02	99	93	no	100	91	38	6
Specify values for criter	tia: leave as "" to not f	ilter: defaults are from the pa	Peptidyl-prolyl cis-trans isomerase B	-13.65	100	0	no	100	94	31	5
specify values for criter	ia, leave as to not i	itter, defaults are from the pe	Macrophage-capping protein	-17.11	100	0	no	100	93	50	5
immunogenicity	min score +		Vinculin	-22.45	100	0	no	100	98	164	3
	max score +		Metalloproteinase inhibitor 2	-25.28	100	99	no	100	93	29	1
transcriptome	min %id 95 ‡	at %coverage 99 ÷	Procollagen C-endopeptidase enhancer 1	-28.49	92	86	yes	100	72	66	23
mouse secretome	min %id 30 ‡	at %coverage 🧐 🗧	Stathmin-4	-28.74	64	0	no	100	99	27	1
SignalP	required? 🗹		Elongation factor 2	-30.01	100	0	no	100	99	131	3
validated HCPs	min %id 🛛 💠	at %coverage 🛛 🗧 🗧	Transketolase	-33.53	100	0	no	100	95	87	11
human proteome	min %id 🛛 💠	at %coverage ÷	Actin, cytoplasmic 1	-34.19	100	0	no	100	100	49	0
	max %id ÷	at %coverage +	Dystroglycan	-36.63	100	0	yes	100	93	121	18
			Cathepsin Z	-39.59	100	90	yes	100	84	36	11
filter			Heat shock cognate 71 kDa protein	-39.63	98	0	no	100	98	29	5
			Beta-2-microglobulin	-43.93	100	70	no	100	70	4	3
			Galectin-1	-45.49	100	0	no	100	91	16	2
			78 kDa glucose-regulated protein	-48.51	100	0	yes	100	99	86	4
			Elongation factor 2	-50.97	100	0	no	100	100	8	0
			Peroviredovin-1	-51.76	100	0	no	100	95	24	5

Collagen alpha-1(III) chain (CH244242922) (PHECK990961	The tool presents protein sets in
immunogenicity score 35.79	HTML tables sortable by column and
9mers 228 total: 168 predicted cross-reactive to human + 60 ur	nique to CHO allows export to CSV-format files
transcriptome 100% id at 24% coverage	
SignalP no	
validated 100% id at 100% coverage human proteome 92% id at 24% coverage	eroxiredoxin-1 [GI:344238489; GB:EGV94592]
im	munogenicity score -51.76
Transcriptome	9mers 24 total: 19 predicted cross-reactive to human + 5 unique to CHO
	transcriptome 100% id at 100% coverage
contig % id % coverage	mouse secretome no matches
JP038052 100 24 JI888352 100 24	SignalP no
JP038052 100 7	validated 100% id at 100% coverage
JI888352 100 7	numan proteome 95% to at 100% coverage
JI894745 56 23	
JP061650 56 23	anscriptome
JI884887 4Z Z3 ID041582 42 23	
COI	ntig % id % coverage
JI8	86573 100 100
Mouse secretome JPC	77204 75 100
protein % id % coverage	1/296 75 100 144317 75 100
procollagen, type III, alpha 1 96 24	75251 69 96
unnamed protein product 58 23 Jpc	056430 69 96
procollagen, type I, alpha 2 58 23	76437 63 93
pro-alpha-2(I) collagen 58 23	051534 63 93
mouse a1(XI) collagen chain 41 23	75976 33 97
JPC	044567 33 97
Validated HCPs	
Ma	ouse secretome
protein % id % coverage	matchas
Cottagen alpha-1(iii) chain 100 100 no	matches
Human proteome Va	lidated HCPs
protein % id % coverage	otein % id % coverage
Collagen alpha-1(III) chain 92 24 Per	roxiredoxin-1 100 100
Collagen alpha-1(II) chain 69 22	
Collagen alpha-1(I) chain 65 22	
Collagen alpha-2(V) chain 57 23	iman proteome
Pulmonary surfactant-associated protein A2 45 8	Kid K coverage
Pulmonary surfactant-associated protein A1 45 8	rovirodovip_1 75 100
Collagen alpha-1(III) chain 43 11	roxiredoxin-2 75 100
Collagen alpha-1(XI) chain 42 23 Per	roxiredoxin-4 68 99
Collagen alpha-2(XI) chain 40 22 Th	ioredoxin-dependent peroxide reductase, mitochondrial 63 93
Collagen alpha-1(XXVII) chain 33 22 Per	roxiredoxin-6 31 97
Collagen alpha-1(XXIV) chain 31 23	



Figure 1. Immune response to mammalian proteins.

FOREIGN

- The availability of the CHO genome and transcriptome has made it possible to apply immunoinformatics tools to HCP analysis.
- This significantly accelerates research on HCP immunogenicity and cross-reactivity.



- Immunogenicity scores calculated for proteins sourced from three different datasets.
- (above 20) are considered potentially High-scoring proteins immunogenic.
- Proteins scoring between -20 and 20 are less likely to generate a significant immune reaction.
- Major validated HCP contaminants peak co-locates with secreted proteins peak, suggesting lower potential immunogenicity for this set,
- A minor peak higher on the scale suggests that some HCP do carry a risk of immunogenicity.



Figure 5. CHOPPI web tool (http://clover.cs.dartmouth.edu/~cbk/choppi/). Each protein is characterized in terms of homology criteria and signal peptide prediction, along with assessments of epitope count and density, and count and density of epitopes unique to CHO according to a human cross-reactivity analysis.

Conclusions

• The presence of HCP in therapeutic protein products may contribute to adverse clinical side effects and may impact drug efficacy.

Figure 2. Databases used for immunogenicity and cross-reactivity analysis. Illustration of the sources used to analyzed potential CHO HCP immunogenicity Sources include CHO genome, CHO transcriptome, published CHO proteins identified in biotherapeutic products (Experimental HCP contaminants), and mouse secreted proteins downloaded from LOCATE and UniProtKB/Swiss-Prot.



• The impact of cross-conserved epitopes on human immune responses may include induction of effector T cells, induction of Tregs, or anergy. • Our ability to score HCP for immunogenicity opens the door for future analysis such as the screening of whole proteins from the CHO

transcriptome.

- Future HCP immunogenicity scores will be adjusted based on the prevalence of the human homolog in human extracellular fluids (such as serum).
- In vitro studies will be required to determine the impact of our findings and their implications.

Discussions leading to the development of CHOPPI were supported by NIH U19 grant AI082642 (to ADG). We thank our collaborators at iCubed, EpiVax, and Dartmouth for their contributions to this work.

