

# **Significant protein sequence alignment between peanut allergen epitopes and vaccine antigens**

Vinu Arumugham

Nov 2016

[vinucubeacc@gmail.com](mailto:vinucubeacc@gmail.com)

## Significant protein sequence alignment between peanut allergen epitopes and vaccine antigens

### Background

Ahmed et al.<sup>1</sup> demonstrated that Pandemrix induced narcolepsy was caused by molecular mimicry between influenza nucleoproteins (NP) and the human hypocretin receptor (HCRT).

So, antibodies created against influenza NP (X-179A), cross reacted against HCRT, resulting in hypocretin dysregulation and narcolepsy.

The specific influenza NP sequence they identified was **YDKEEIRRIWR**, that mimicked HCRT sequence **YDDEEFLRYLWR**.

This raises the possibility of other such cross reactions due to molecular mimicry.

Stanley et al.<sup>2</sup> identified the following IgE epitopes of the Ara h 2 peanut allergen.

HASARQQWEL  
QWELQGDR  
DRRCQSQLER  
LRPCEQHLMQ  
KIQRDEDS  
RDPYSP  
SQDPYSPS  
LQGRQQ  
KRELRLNL  
QRCDLDVE

I used the BLAST (Basic Local Alignment Search Tool)<sup>3</sup> of the National Center for Biotechnology Information (NCBI) to search for protein sequence matches to the above epitopes.

Specifically, Protein BLAST or the blastp tool.

<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>

I selected the following organisms in the BLAST Non-redundant protein sequences (nr) database for the search:

Homo sapiens (to determine HCRT baseline protein match score and expect values)

Clostridium tetani (vaccine antigen)

Corynebacterium diphtheriae (vaccine antigen)

Bordetella pertussis (vaccine antigen)

Haemophilus influenzae (vaccine antigen)

Streptococcus pneumoniae (vaccine antigen)

human poliovirus (vaccine antigen)

Saccharomyces cerevisiae (vaccine contaminant)

### **Expect value**

The "expect" value or e-value depends on the selected database size. The selected database size was therefore kept constant for all searches. This allowed comparison of e-values across search results.

The e-value and score obtained by the influenza NP (X-179A) to HCRT sequence match was used as the baseline. We know that this level of sequence match can cause cross reacting antibodies.

[https://blast.ncbi.nlm.nih.gov/blast/Blast.cgi?CMD=Web&PAGE\\_TYPE=BlastDocs&DOC\\_TYPE=FAQ#expect](https://blast.ncbi.nlm.nih.gov/blast/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=FAQ#expect)

### **Baseline**

blastp search result for **YDKEEIRRIWR**. HCRT is also known as an orexin receptor.

>ABR23648.1 orexin receptor 2 variant 1A, partial [Homo sapiens]

ABR23649.1 orexin receptor 2 variant 1B, partial [Homo sapiens]

ABR23650.1 orexin receptor 2 variant 1C short, partial [Homo sapiens]

ABR23651.1 orexin receptor 2 variant 1C long, partial [Homo sapiens]

Length=74

Score = 19.3 bits (38), Expect = 1436

Identities = 7/12 (58%), Positives = 7/12 (58%), Gaps = 1/12 (8%)

Query 1 YDKEEIRR-IWR 11

YD EE R WR

Sbjct 34 YDDEEFLRYLWR 45

Higher scores (and lower e-value) mean better match and higher likelihood of cross reaction.

Results

Peanut Ara h 2 epitope	DRRCQSQLER		HASARQQWEL		KIQRDEDS		KRELRNL		LQGRQQ		LRPCEQHLMQ		QRCDLDVE		QWELQGDR		RDPYSP		SQDPYSPS	
	Score	Expect	Score	Expect	Score	Expect	Score	Expect	Score	Expect	Score	Expect	Score	Expect	Score	Expect	Score	Expect	Score	Expect
Saccharomyces cerevisiae	22.3	99	24	24	22.3	63			18.5	907	22.3	98	20.6	258	22.7	44	20.2	223	20.2	367
Clostridium tetani	18.9	1679	18.9	1678	19.7	530	22.7	3.4	15.9	7618			20.2	372	19.7	526	17.6	1844	18.9	1068
Streptococcus pneumoniae	22.3	98	23.5	34	23.5	22	22.7	3.4	22.3	39	21.8	140	22.7	45	21.8	89	21	111	21	181
Corynebacterium diphtheriae	21.4	200	23.1	48	19.3	748	22.7	3.4			24.4	17	20.2	369	21.8	90	20.2	225	20.2	369
Bordetella pertussis	20.6	404	24	24	20.2	370	19.7	40	19.3	450	21.4	199	21.4	131	21	182	20.2	224	20.2	368
Haemophilus influenzae	19.7	824	19.7	825	21	183	19.7	40	19.3	450	24.4	17	20.2	369			20.2	224	20.2	332
Human poliovirus 1	17.2	7042															13.8	48175		
Baseline																				
HCRT	19.3	1436																		
Low or no cross reaction																				

A majority of hits yield higher scores than baseline, indicating high cross reactivity.

Sample sequence matches

>WP\_046656073.1 transcriptional regulator [Clostridium tetani]

Length=576

Score = 22.7 bits (46), Expect = 3.4

Identities = 6/6 (100%), Positives = 6/6 (100%), Gaps = 0/6 (0%)

Query 2 RELRNL 7

RELRNL

Sbjct 492 RELRNL 497

>WP\_010934900.1 magnesium transporter CorA [Corynebacterium diphtheriae]

CAE49742.1 Putative magnesium and cobalt transport protein [Corynebacterium diphtheriae]

KKA80801.1 magnesium transporter CorA [Corynebacterium diphtheriae]

OFI51823.1 magnesium transporter CorA [Corynebacterium diphtheriae]

OFI61790.1 magnesium transporter CorA [Corynebacterium diphtheriae]

Length=366

Score = 24.4 bits (50), Expect = 17

Identities = 6/8 (75%), Positives = 7/8 (88%), Gaps = 0/8 (0%)

Query 3 PCEQHLMQ 10

P EQH+MQ

Sbjct 78 PSEQHMMQ 85

## Conclusion

While conformation and surface exposure must be factored in, the sequence matches seem significant enough to warrant further research. Also, Pandemrix was a single dose. Childhood vaccines containing the above antigens are administered numerous times, thus increasing the probability of synthesizing antibodies to these protein sequences.

The Prevnar 13 vaccine contains *Streptococcus pneumoniae*, *Corynebacterium diphtheriae*, *Saccharomyces cerevisiae* proteins and is administered 4 times, making it especially interesting.

These results suggest that much more care needs to be exercised during vaccine design. Careful selection of vaccine targets, excipients, growth media, etc. is essential to avoid off-target immune responses that can result in serious adverse events such as life-threatening food allergies.<sup>4</sup>

## References

1. Ahmed SS, Volkmuth W, Duca J, Corti L, Pallaoro M, Pezzicoli A, et al. Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2 (ABSTRACT ONLY). *Sci Transl Med*. 2015;7(294):294ra105–294ra105.
2. Stanley JS, King N, Burks AW, Huang SK, Sampson H, Cockrell G, et al. Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2. *Arch Biochem Biophys*. 1997;342(2):244–53.
3. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389–402.
4. Arumugham V. Evidence that Food Proteins in Vaccines Cause the Development of Food Allergies and Its Implications for Vaccine Policy. *J Dev Drugs*. 2015;4(137):2.

