

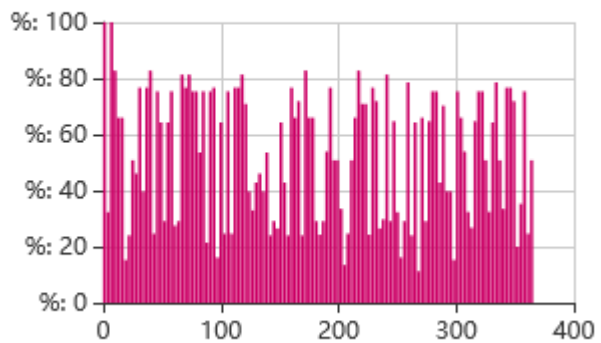
Final sequence

MEAAAKMAENSNIDDIKAPLLAALGAADLALATVNELITNLRERAEEETRTDTRS RVEESRARLTKLQEDLPEQLTELREK
 TAEELRKA AEGYLEAATSRYNELVERGEAALERLSQQSFEEVSARAEGYVDQAVELTQEALGTVASQTRAVGERAAK
 VGIELP**EAAAK**NDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNE**GPGPG**IEYNDMFNNFTVSFWLRVPKVSASHL
 EQYGT**AAYIPAF**AAAYFAVKIQAAFREQPDNIPAF**HEYGAEALERAG**RSHRGVCKVISPDGECSSISIDDL SFYV**HEY**
GAEALERAGQVSTPRFGKSGSPSIKALFQLGPPSPV**HEYGAEALERAG**

↓ H-b H ↓ TTFrC ↓ PTTG1 ↓ SP17 ↓ AKAP4 ↓ TTFrC

Figure S1: Final designed sequence

Relative Adaptiveness (optimized)



GC content (after optimization)

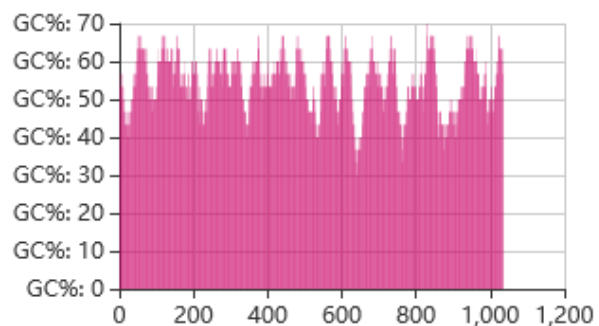


Figure S2: Graphical view of codon usage in optimized chimeric gene. CAI before optimization: N.A CAI after optimization: 0.81. GC content before optimization: N.A. GC content after optimization: 54.76%.

Table S1: TTFrC MHC-II binding peptides determined by RANKPEP.

Epitope sequence	Start position	End position
NDIISDISGFNSSVITYPDAQLVPGINGKAIHLVNNE	40	66
IEYNDMFNNFTVSFVLRVPKVSASHLEQYGT	78	108

Table S2: Conformational B-cell epitopes from full length proteins using DiscoTope server

No.	Residues	Number of residues	Score
1	_:S54, _:V56, _:E57, _:E58, _:S59, _:R60, _:A61, _:R62, _:L63, _:T64, _:K65, _:L66, _:Q67, _:E68, _:D69, _:L70, _:P71, _:E72, _:Q73, _:L74, _:T75, _:E76, _:L77, _:R78, _:E79, _:K80, _:F81, _:T82, _:A83, _:E84, _:E85, _:L86, _:A89, _:A90, _:E91, _:G92, _:Y93, _:L94, _:E95, _:A96, _:A97, _:T98, _:S99, _:R100, _:Y101, _:N102, _:E103, _:L104, _:V105, _:E106, _:R107, _:G108, _:E109, _:A110, _:A111, _:L112, _:E113, _:R114, _:L115, _:R116, _:S117, _:Q118, _:Q119, _:S120, _:F121, _:E122, _:E123	67	0.798
2	_:E2, _:A3, _:A4, _:A5, _:K6, _:M7, _:A8, _:E9, _:N10, _:S11, _:N12, _:I13, _:D14	13	0.795
3	_:A324, _:G325, _:Q326, _:V327, _:S328, _:T329, _:P330, _:R331, _:F332, _:G333, _:K334, _:G335, _:S336, _:G337, _:P338, _:K341	16	0.675
4	_:G213, _:I214, _:E215, _:Y216, _:N217, _:D218, _:M219, _:F220, _:N221, _:N222, _:F223, _:T224, _:V225, _:S226, _:F227	15	0.667
5	_:K17, _:A18, _:P166, _:E167, _:A168, _:A169, _:A170, _:K171, _:N172, _:D173, _:I174, _:I175, _:S176, _:D177, _:I178, _:S179, _:G180, _:F181, _:N182, _:S183, _:S184, _:V185, _:I186, _:T187, _:L193, _:V194, _:P195, _:G196, _:I197, _:N198, _:G199, _:K200, _:A201, _:W228	34	0.661
6	_:K233, _:V234, _:S235, _:A236, _:S237, _:H238, _:L239, _:E240, _:Q241, _:Y242, _:G243, _:T244, _:A245, _:A254, _:Y255, _:V258, _:K259, _:I260, _:Q261, _:A262, _:A263, _:F264, _:R265, _:E266, _:Q267, _:P268, _:D269, _:N270, _:I271, _:P272, _:E276	31	0.639
7	_:Y316, _:G317, _:A318, _:A320, _:L321, _:E322, _:R323, _:F344, _:Q345, _:L346, _:G347, _:P348, _:P349	13	0.608
8	_:R230, _:V231, _:P232	3	0.547

Table S3: AlgPred server reports. Prediction of allergenic proteins and mapping of IgE epitopes

Index	Allergenic status	Score	Threshold	Positive Predictive Value	Predictive Value
Prediction by SVM method based on amino acid composition	NON ALLERGEN	0.41484908	0.4	18.21% Negative	71.24%
Prediction based on SVM method based on dipeptide composition	ALLERGEN	0.027923492	0.2	63.1% Negative	85.56%
Prediction by mapping of IgE epitope	The protein sequence does not contain experimentally proven IgE epitope				
Blast RESULT	NON ALLERGEN	BLAST Results of ARPS: No Hits found			
Prediction by Hybrid Approach	NON ALLERGEN				