

Reducing the immunogenicity of pulchellin A-chain, ribosome-inactivating protein type 2, by computational protein engineering for potential new immunotoxins

Reza Maleki¹, Libing Fu², Ricardo Sobhie Diaz³, Francisco Eduardo Gontijo Guimarães⁴, Otávio Cabral-Marques^{5,6,7,8}, Gustavo Cabral-Miranda^{*5,9}, Mohammad Sadraeian^{*2}

¹ Adelaide Medical School, University of Adelaide, Adelaide, SA, Australia.

² Institute for Biomedical Materials and Devices (IBMD), Faculty of Science, University of Technology Sydney, Sydney, NSW 2007, Australia

³ Laboratório de Retrovirologia, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo 04039-032, SP, Brazil

⁴ Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, São Carlos, SP, CEP 13560-970, Brazil

⁵ Department of Immunology, Institute of Biomedical Sciences, University of São Paulo (ICB/USP), São Paulo 05508000, SP, Brazil

⁶ Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, SP, Brazil.

⁷ Laboratory of Medical Investigation in Pathogenesis and Directed Therapy in Onco-Immuno-Hematology (LIM-31), Department of Hematology and Cell Therapy, Hospital das Clínicas HCFMUSP, University of São Paulo, São Paulo, Brazil.

⁸ Interunit Post Graduate Program on Bioinformatics, Institute of Mathematics and Statistics (IME), University of São Paulo (USP), São Paulo, SP, Brazil.

⁹ Institute of Research and Education in Child Health (PENSI), São Paulo 01228200, SP, Brazil

*** Corresponding authors:**

Gustavo Cabral-Miranda, email: gcabral.miranda@usp.br, +55(11)3091-7206

Mohammad Sadraeian, email: mohammad.sadraeian@uts.edu.au, +61-431963783

*** Current addresses:**

Department of Immunology, Institute of Biomedical Sciences, University of São Paulo, São Paulo 05508000, SP, Brazil

Institute for Biomedical Materials and Devices (IBMD), Faculty of Science, University of Technology Sydney, Sydney, NSW 2007, Australia

Supplementary Materials

Modeling servers	Verify3D (%) of the residues had an averaged 3D1D score ≥ 0.2	ERRAT Overall Quality Factor	ProSA-WEB ZScore	Ramachandran Plot (%) Favored(F) additionally allowed (AA) generously allowed (GA) disallowed(D)	
Galaxy (100% Amino acid Coverage)					
Model 1	98.01 %	96.69	-7.59	F=93.9 GA=0.0	AA=5.7 D=0.4
Model 2	97.61 %	95.06	-7.55	F= 94.7 GA=0.0	AA=4.8 D=0.4
Model 3	96.70 %	96.70	-7.63	F= 93.4 GA= 0.4	AA=5.7 D=0.4
Model 4*	98.01 %	95.88	-7.61	F=94.7 GA= 0.0	AA=4.8 D=0.4
Model 5	98.41 %	95.88	-7.57	F= 93 GA=0.0	AA=6.6 D=0.4
I-TASSERS					
Model C-score (1.53)	94.02 %	97.52	-7.78	F=82.5 GA= 2.6	AA=14.9 D=0.0
RaptorX					
Model	94.42 %	81.07	-7.73	F= 92.1 GA=0.0	AA=7.5 D=0.4
Phyre2					
Model	97.61 %	81.89	-7.56	F= 85.5 GA= 0.4	AA=13.6 D=0.4
SWISS-MODEL					
Model	98.01 %	91.73	-7.83	F= 88.6 GA= 0.0	AA=11 D=0.4

Table S1: Evaluation of 3D structure models of PAC by Verify 3D, ERRAT, ProSAWEB, and Ramachandran Plot. Galaxy model 4 was chosen for refinement process.

Models	Verify3D (%) of the residues had an averaged 3D1D score ≥ 0.2	ERRAT Overall Quality Factor	ProSA-WEB ZScore	Ramachandran Plot (%) Favored(F) additionally allowed (AA) generously allowed (GA) disallowed(D)	
GALAXY Model 4 (Before refinement)	98.01%	95.88	-7.61	F=94.7 GA= 0.0	AA=4.8 D=0.4
GALAXY Model 4 (After refinement)					
Model 1	96.81%	95.43	-7.81	F= 93.4 GA= 0.0	AA=6.6 D=0.0
Model 2*	98.01%	97.51	-7.74	F=93.4 GA=0.0	AA=6.1 D=0.4
Model 3	97.61%	93.41	-7.67	F=92.5 GA= 0.0	AA=7 D=0.4
Model 4	97.61%	94.60	-7.68	F= 92.1 GA= 0.0	AA=7.5 D=0.4
Model 5	94.82%	95.85	-7.53	F= 92.5 GA= 0.0	AA=7.5 D=0.0
Model 6	96.81%	94.21	-7.71	F=93 GA= 0.0	AA=7 D=0.0
Model 7	96.41%	96.26	-7.45	F= 92.5 GA= 0.0	AA=7.5 D=0.0
Model 8	96.81%	95.83	-7.6	F=92.4 GA=0.0	AA=7.9 D=0.0
Model 9	96.41%	96.26	-7.74	F=93 GA=0.0	AA=7 D=0.0
Model 10	96.81%	93.77	-7.47	F=92.1 GA=0.0	AA=7.9 D=0.0

Table S2: Validation scores of Galaxy Model after refinement. Selected model 2 of refinement obtained the highest validation score.

Residue suggestion	Discoptope score Threshold: -3.7	EPSVR SCORE	SEPPA Score Threshold:0.089	Ellipro Conformational Score Threshold:0.5	DEPTH of residue (Å)	OSP of residue	Pseudo $\Delta\Delta G$ of protein	Stability
Wild: T82	-3.58	94	0.190	0.734	3.2	0.15	-	-
T82A*	-4.964	76	0.169	0.733	3.1	0.22	0.12	↑
T82V	-4.781	81.00	0.277	0.734	3.1	0.17	0.07	↑
T82L	-5.246	91.00	0.278	0.732	3.3	0.13	0.5	↑
Wild: T100	-3.376	90	0.223	0.734	3.4	0.19	-	-
T100A*	-4.4	86.00	0.337	0.732	3.1	0.2	0.35	↑
T100V	-4.707	84.00	0.471	0.732	3.4	0.16	0.42	↑
T100L	-4.793	85.00	0.467	0.734	3.3	0.18	0.85	↑
Wild: D101	-2.414	90	0.247	0.734	3.3	0.11	-	-
D101A	-4.331	89.00	0.311	0.735	3.2	0.08	-0.21	↓
D101V*	-4.331	82.00	0.260	0.735	3.3	0.1	-0.26	↓
D101L	-4.769	92.00	0.322	0.735	3.2	0.06	0.2	↑
Wild: Q121	-1.651	87	0.088	0.572	3.4	0.12	-	-
Q121A*	-3.963	67.00	0.046	0.567	3	0.13	0.24	↑
Q121V	-3.921	64.00	0.080	0.568	3.2	0.15	0.22	↑
Q121L	-4.24	69.00	0.080	0.572	3.2	0.1	0.56	↑
Wild: N146	-2.539	96	0.234	0.713	3.2	0.15	-	-
N146A	-4.97	78.00	0.199	0.723	3	0.2	0.03	↑
N146V*	-5.225	75.00	0.200	0.723	3.2	0.17	0.02	↑
N146L	-5.226	90.00	0.203	0.724	3.2	0.14	0.25	↑
Wild: D147	-3.658	98	0.075	0.758	3.5	0.37	-	-
D147A*	-5.852	96.00	0.063	0.722	3.5	0.34	0.15	↑
D147V	-5.648	96.00	0.055	0.724	3.6	0.31	0.02	↑
D147L	-6.012	98.00	0.055	0.723	3.5	0.37	0.12	↑
Wild: R149	-3.654	94	0.073	0.758	3.6	0.28	-	-
R149A*	-6.155	88.00	0.051	0.654	3.1	0.41	0.41	↑
R149V	-6.376	92.00	0.050	0.654	3.4	0.41	-0.33	↓
R149L	-6.394	91.00	0.050	0.62	3.7	0.34	0.39	↑

Table S3: Conformational B cell epitope scores and SDM2 results of all potential mutants.

Models	Verify3D (%) of the residues had an averaged 3D1D score ≥ 0.2	ERRAT Overall Quality Factor	ProSA-WEB ZScore	Ramachandran Plot (%) Favored(F) additionally allowed (AA) generously allowed (GA) disallowed(D)
PAC wild type	98.01%	97.51	-7.74	F=93.4 AA=6.1 GA=0.0 D=0.4
PAC Mutants				
T82A	98.01%	97.51	-7.57	F=93.4 AA=6.1 GA=0.0 D=0.4
T100A	98.01%	97.51	-7.63	F=93.4 AA=6.1 GA=0.0 D=0.4
D101V	98.01%	95.85	-7.54	F=93.4 AA=6.6 GA=0.0 D=0.4
Q121A	98.01%	97.51	-7.71	F=93.4 AA=6.1 GA=0.0 D=0.4
N146V	98.01%	97.51	-7.68	F=93.4 AA=6.1 GA=0.0 D=0.4
D147A	98.01%	97.51	-7.69	F=93.4 AA=6.1 GA=0.0 D=0.4
R149A	98.01%	97.51	-7.73	F=93.4 AA=6.1 GA=0.0 D=0.4

Table S4: Validation scores of wild type and mutants form of PAC. Mutant models were obtained from SDM2 server.

Galaxy server	Verify3D (%) of the residues had an averaged 3D1D score ≥ 0.2	ERRAT Overall Quality Factor	ProSA-WEB ZScore	Ramachandran Plot (%) Favored(F) additionally allowed (AA) generously allowed (GA) disallowed(D)	
PAM					
Model 1*	98.01 %	97.10	-6.95	F= 93.4 GA= 0.0	AA= 6.1 D= 0.4
Model 2	97.61 %	94.65	-7.01	F= 93.4 GA= 0.0	AA= 6.1 D= 0.4
Model 3	94.02 %	97.52	-6.93	F= 93 GA= 0.0	AA=6.6 D= 0.4
Model 4	98.41 %	96.70	-7.02	F= 92.5 GA= 0.0	AA= 7 D= 0.4
Model 5	97.61 %	97.11	-6.91	F= 92.5 GA= 0.0	AA= 7 D= 0.4

Table S5: Evaluation of 3D structure of a pulchellin containing all mutations model with Verify 3D, ERRAT, ProSAWEB, and Ramachandran Plot.

Galaxy server	Verify3D (%) of the residues had an averaged 3D1D score ≥ 0.2	ERRAT Overall Quality Factor	ProSA-WEB ZScore	Ramachandran Plot (%) Favored(F) additionally allowed (AA) generously allowed (GA) disallowed(D)	
PAM model 1 (Before refinement)	98.01 %	97.10	-6.95	F= 93.4 GA= 0.0	AA= 6.1 D= 0.4
PAM model 1 (After refinement)					
Model 1	94.42%	98.75	-7.15	F= 94.3 GA= 0.0	AA=5.7 D=0.0
Model 2	97.61%	98.34	-7.17	F= 94.7 GA= 0.0	AA=5.3 D=0.0
Model 3*	98.41	97.09	-7.15	F=94.7 GA= 0.0	AA=5.3 D=0.0
Model 4	96.02%	98.34	-7.09	F= 93.9 GA= 0.0	AA=6.1 D=0.0
Model 5	97.61%	96.69	-6.98	F= 93.9 GA= 0.0	AA=6.1 D=0.0
Model 6	94.82%	97.92	-7.03	F=94.3 GA=0.0	AA=5.7 D=0.0
Model 7	96.81%	97.92	-6.93	F= 93.9 GA= 0.0	AA= 6.1 D=0.0
Model 8	98.01%	97.91	-7.09	F= 94.3 GA=0.0	AA=5.7 D=0.0
Model 9	94.82%	94.19	-7.15	F= 94.7 GA= 0.0	AA= 5.3 D=0.0
Model 10	94.42%	98.75	-7.15	F=93.0 GA= 0.0	AA=7.0 D=0.0

Table S6: Validation scores of PAM model after refinement. Model 3 of refinement was selected due to the highest validation score.

Models	Interaction residues with CGAGAG according to Auto dock Vina	Predicted pulchellin interaction residue according to homology to other RIPs
PAC wild type	I21-L24-R25-L28-L39-L56-V66-I68-N72-A73-Y74-V75-Y78-G111-S112-Y113-L116-R124-M155-I156-I158-I159-A162-S163-E164-A166-R167-N196-W198-D199-N200-L242-L244	N72-Y74-Y113-R124-Q160-E164-R167-E195-N196-W198
PAC Mutants		
T82A	N72-Y74-V75-L87-D89-G111-S112-Y113-S114-L116-Q117-S123-R124-I127-A152-M155-I156-I159-E164-E195-N196-W198-D199-N200-S202-L244	N72-Y74-Y113-R124-Q160-E164-R167-E195-N196-W198
T100A	L28-L39-P42-T71-N72-A73-Y74-V75-V76-L87-D89-G111-S112-Y113-L116-R124-I159-S163-E164-R167-Y168-Y170-I171-L194-E195-N196-W198-D199-S202-L242-L244	N72-Y74-Y113-R124-Q160-E164-R167-E195-N196-W198
D101V	N72-Y74-V75-V76-D89-D110-G111-S112-Y113-S114-L116-R124-I156-I159-E164-L192-E195-N196-N197-W198-D199-N200-I218	N72-Y74-Y113-R124-Q160-E164-R167-E195-N196-W198
Q121A	L39-V70-T71-N72-A73-Y74-V76-D89-G111-S112-Y113-S114-L116-R124-I156-I159-E195-N196-N197-W198-D199-N200-L244-F245	N72-Y74-Y113-R124-Q160-E164-R167-E195-N196-W198
N146V	Y17-F20-R25-L28-V54-L56-V66-N72-A73-Y74-V75-Y78-L87-D89-G111-S112-Y113-L116-R124-L134-M155-I156-I158-I159-Q160-M161-A162-S163-E164-A166-R167-E195-N196-W198-D199-A241-L242-L244	N72-Y74-Y113-R124-Q160-E164-R167-E195-N196-W198
D147A	L24-R25-L28-L39-N72-A73-Y74-V75-L87-D89-G111-S112-Y113-S114-R124-M155-I159-S163-E164-A166-R167-L192-E195-N196-N197-W198-200N-L242-L244	N72-Y74-Y113-R124-Q160-E164-R167-E195-N196-W198
R149A	F20-L24-L28-V54-V66-I68-N72-A73-Y74-V75-L87-D89-G111-S112-Y113-S114-I158-I159-A162-S163-E164-A166-R167-Y168-I171-E195-N196-W198-D199-N200-L244	N72-Y74-Y113-R124-Q160-E164-R167-E195-N196-W198
PAM	L24-L28-I52-I68-N72-A73-Y74-V75-L87-D89-G111-S112-Y113-S114-L116-M155-I159-S163-E164-A166-R167-Y168-L194-E195-N196-N197-W198-D199-N200-L201-S202-L240	N72-Y74-Y113-R124-Q160-E164-R167-E195-N196-W198

Table S7: Pulchellin interaction amino acids (wild type and mutants) in the docking process. Residues of active site in N146V contributed the most, while Q121A contributed the least in docking process. Highlight amino acids indicate shared active site residues between homology prediction and molecular docking.