Combined Bioinformatic and Rational Design Approach to Develop Antimicrobial Peptides against *Mycobacterium tuberculosis*

C. Seth Pearson[1], Zachary Kloos[2+], Brian Murray[1+], Ebot Tabe[2+], Monica Gupta[2], Jun Ha Kwak[1], Pankaj Karande[1], Kathleen McDonough[2,3,*] and Georges Belfort[1,*]

[1]Howard P Isermann Department of Chemical and Biological Engineering and Center for Biotechnology and Interdisciplinary Studies, RPI, Troy, NY 12180, [2]Division of Infectious Diseases, Wadsworth Center, NY State Dept. of Health, Albany NY, 12208, and [3]Department of Biomedical Sciences, University at Albany, SUNY

[+]These authors contributed equally to the work

* **Corresponding authors**: Email belfog@rpi.edu and kathleen.mcdonough@health.ny.gov

AMP Purification

Peptide purification was necessary due to inefficient cleavage of Pbf (pentamethyldihydro-benzofuransulfonyl) side-chain protecting groups. Purification was run on a C18 column with isocratic elution at 30% acetonitrile (ACN) at 50°C. Purity of product was confirmed using MALDI-TOF MS (**Fig. S1**)

In vivo *Activity and Toxicity against M. tuberculosis*

The disk diffusion assay (**Fig. S2**) showed all three designed AMPs, B1, B2 and B3, were toxic to *M. tuberculosis in vivo*. Using 100 μg, all three AMPs were potent enough to produce a dark ring of dead cells around the inserted disk. B3 appeared the most potent followed by B1 then B2. 10 μg of kanamycin and water were used as positive and negative controls, respectively. For quantitative MIC values of the AMPs, the Alamar blue assay was performed (**Table 1**).

Relevant Tables of AMPs

The full list of database peptides used in the bioinformatic analysis is given in **Table S1**. For comparison purposes, the peptide sequences of the four designed peptides are shown above four natural antimicrobial peptides with similar properties **Table S2**. Finally, in **Table S3** we show the results of the database filtration using the complete database and a modified version of the database with only four peptides from Ramon-Garcia *et al.* (Ref. 9).

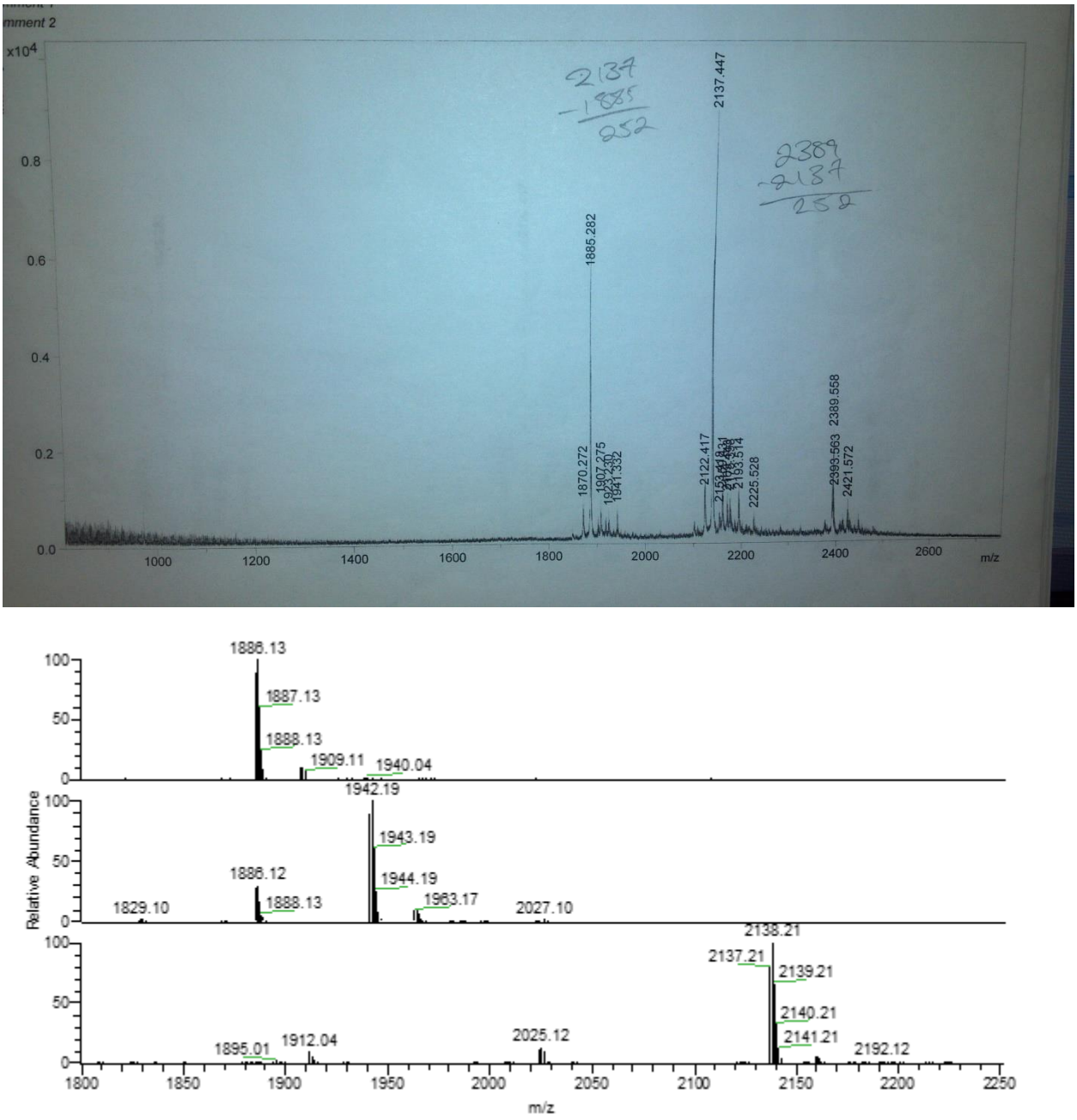

Figure S1. MALDI-TOF MS of crude synthesis product for B3 (top), isocratic elution at 30% ACN (second from top), isocratic elution at 35% ACN (third from top), and isocratic elution at 80% ACN (bottom).

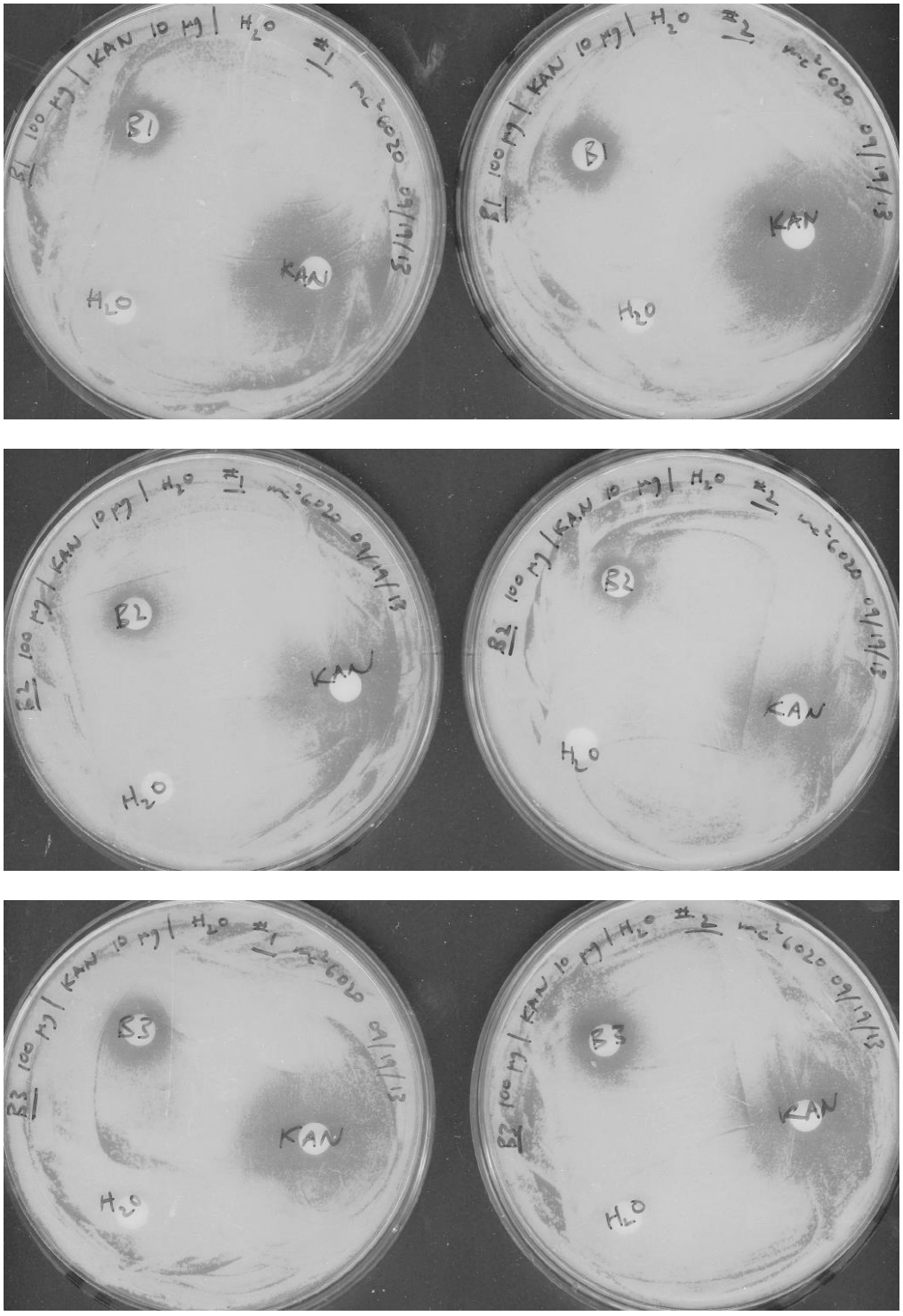

Figure S2. Disk diffusion assay for the three designed AMPs, B1 (top), B2 (middle) and B3 (bottom) were performed in duplicate. Disks with 100 μg AMP were placed onto lawns of *M. tuberculosis*. 10 μg kanamycin disks and water disks used as positive and negative controls, respectively. Dark areas indicated the presence of dead cells.

Table S1. List of *M. tuberculosis* database peptide sequences with references.

Peptide Sequences	Reference
W K W L K K W I K	Ramón-García et al. AAC, 2013, **57**(5); 2295-2303
R W R R K W W W W	
W R K F W K Y L K	
K R W W K W W R R	
R R W W K W W W R	
K I W W W W R K R	
R L W W W W R R K	
K W K W W W R K I	
R I R R W K F R W	
R L K R W W K F L	
R W W R W R K W W	
K R W W W W R F R	
K R W W R K W W R	
R R W W R W V V W	
W F K M R W W G R	
K F K W W R M L I	
R I K R W W W W R	
R W R W W W R V Y	
L K R R W K W W I	
R R R I K I R W Y	
R L W W K I W L K	
K R R W R I W L V	
F F I Y V W R R R	
I R M R I R V L L	
R W W R K I W K W	
L R F I L W W K R	
R W W I R I R W H	
R R R W W K L M M	
L R R W I R I R W	
R K F R W W V I R	
H Q F R F R F R V R R K	
I L R W K W R W W R W R R	
I L P W K W R W W K W R R	
F I K W K F R W W K W R K	
R G G R L C Y C R R R F C V C V G R	Zhao et al. FEBS Lett., 1994, **346**(2-3); 285-258
K W K S F L K T F K S A K K T V L H T A L K A I S S	Jiang et al. Protein Pept Lett., 2011, **18**(3); 241-252
K W K S F L K T F K S A K K T V L H T L L K A I S S	
K W K S F L K T F K S L K K T V L H T L L K A I S S	
K W K S F L K T F K S L K K T V L H T L L K L I S S	
K W K S F L K T F K S L K K T K L H T L L K L I S S	
A C Y C R I P A C I A G E R R Y G T C I Y Q G R L W A F C C	Madison et al. Infect. Immun., 2007, **75**(10); 4780-91
V V C A C R R A L C L P R E R R A G F C R I R G R I H P L C C R R	Ganz et al. J Immunol., 1989, **143**(4); 1358-65
L L G D F F R K S K E K I G K E F K R I V Q R I K D F L R N L V P R T E S	Zelezetsky et al. J Biol. Chem., 2006, **281**(29);19861-71
R R R P R P P Y L P R P R P P P F F P P R L P P R I P P G F P P R F P P R F P	Gudmundsson et al. PNAS, 1995, **92**(15);7085-9

Table S2. Linear sequences of designed AMPs with similar literature peptides for comparison

Peptide Names	Peptide Sequences	Reference
B1	I L S L R W R W K W W K K	
B2	I L S L R W W R K W W K K	
B3	I L S L R W R W W K W K K	
B4	I R K L K S W K W L R W L	
Indolicidin*	I L A W K W A W W A W R R	Selsted et al. J Biol. Chem., 1992, **267**(7); 4292-5
Tritrpticin	V R R F P W W W P F L R R	Lawyer et al. FEBS Lett., 1996, **390**(1); 95-8
Lactoferricin b 15	P M F K C W R W Q W R W K K L G A M	Tian et al. Appl Microbiol Biotechnol., 2007, **75**(1); 117-24
Lactoferricin-cecropin A hybrid protein	K W K L F K K F K C W R W Q W R W K K L G A	Feng et al. J Ind Microbiol Biotechnol., 2014, **41**(3); 527-34

--- Peptides under the dashed lines are peptides of similar amino acid composition from the literature

*The C-terminal of indolicidin is modified with $–NH_2$

Table S3. Comparison of database filtration including and excluding the majority of peptides from Ramon-Garcia et al. (Ref. 9)

	With 34 Ramon-Garcia, et al. (Ref. 9) Peptides			With 4 Ramon-Garcia, et al. (Ref. 9) Peptides		
	Number	**Most Common**	**Prediction**	**Number**	**Most Common**	**Prediction**
Hydrophobic Residues	6.51 (7)	W (36%) L (18%) I (13%)	W,W,W,W,L, L,I	5.85 (6)	L (24%) W (13%) I (11%)	W,W,L,L,L,I
Charged Residues	5.17 (5)	K (56%) R (40%)	K,K,K,R,R	4.39 (4-5)	K (51%) R (49%)	K,K,(K),R,R
Polar Residues	1	S (51%)	S	2-3	S (51%)	S,S,(S)