Supporting Information: Systematic Mutagenesis of Oncocin Reveals Enhanced Activity and Insights into the Mechanisms of Antimicrobial Activity

Pin-Kuang Lai, Kathryn Geldart, Seth Ritter ,Yiannis N. Kaznessis, and Benjamin J. Hackel*

Table S1 Purity data from the SPOT synthesis of monosubstituted peptides. These peptides were selected randomly from the library for quality control. MW_{calc} is theoretical molecular weight and MW_{anal} is the molecular weight from analysis

Sequence (first eleven residues)	MW_{calc}	MW_{anal}	Purity
VDKPPYLPRPR	2446.88	2446.6	69.1
HDKPPYLPRPR	2484.89	2484.8	34.8
VTKPPYLPRPR	2432.9	2431.8	64.5
VDKQPYLPRPR	2477.89	2477.3	62.2
VDKPPNLPRPR	2397.81	2396.4	72.5
VDKPPYAPRPR	2404.80	2404.2	66.3
VDKPPYCPRPR	2436.86	2436.4	52.5
VDKPPYMPRPR	2464.92	2464.0	56.0
VDKPPYLYRPR	2512.93	2512.4	65.5
VDKPPYLPRHR	2486.90	2486.3	62.5
VDKQPYLPRPR	2477.89	2476.6	54.6
VDKPPYLPRPR	2446.88	2445.6	64.4

Table S2 Purity data from the SPOT synthesis of multisubstituted peptides. These peptides were selected randomly from the library for quality control. MW_{calc} is theoretical molecular weight and MW_{anal} is the molecular weight from analysis

Sequence (first eleven residues)	MW_{calc}	MW_{anal}	Purity
VDKPPYLPRPR	2446.88	2446.1	60.9
VDKKPYKPRKR	2523.99	2524.0	61.8
VDKKHYKPRRR	2592.03	2592.1	67.5
VDKRPRLPRRR	2558.04	2557.3	61.9
VDKRHRRPRKR	2613.07	2610.8	67.2

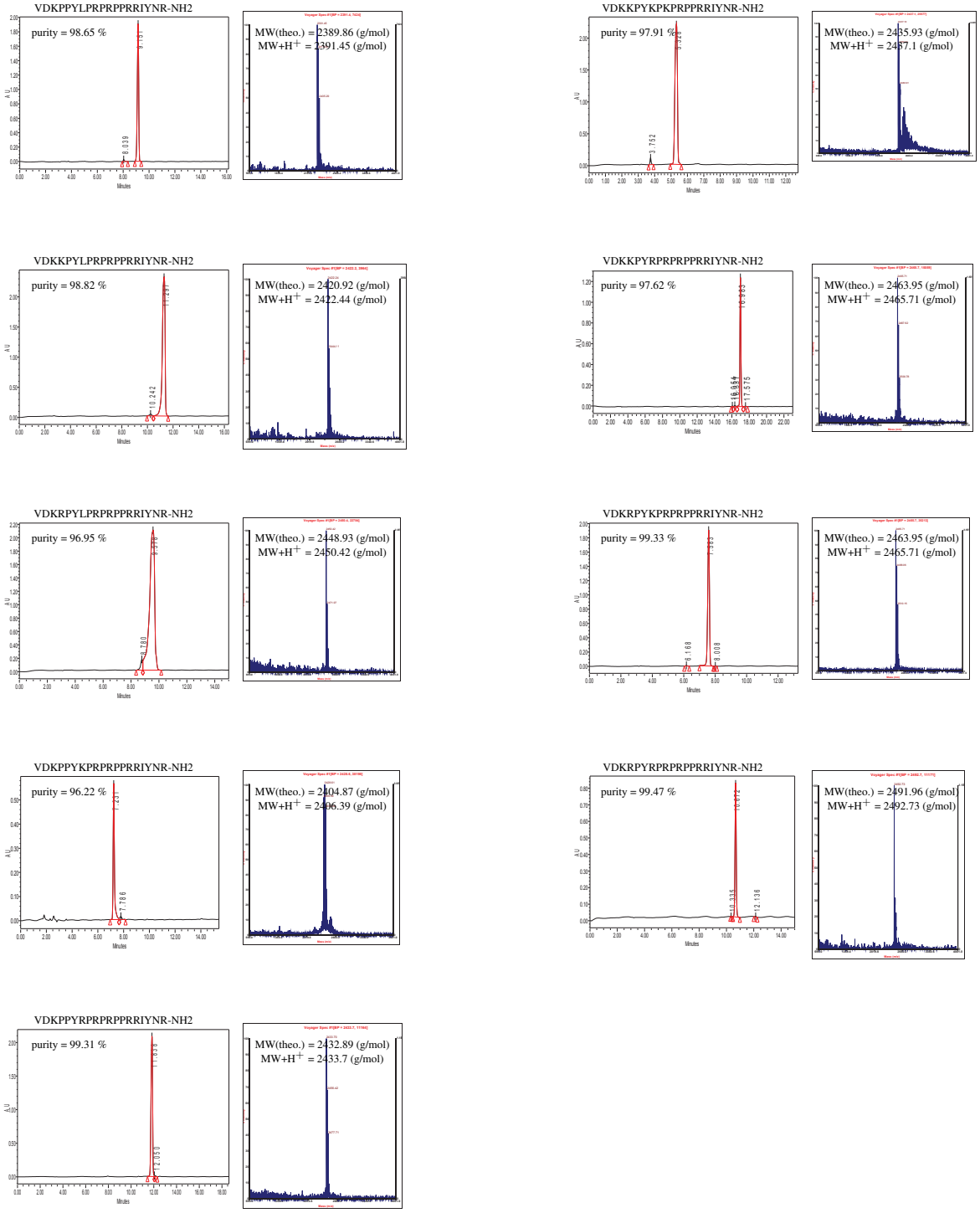

Figure S1 Characterization, HPLC and MS data of the purified peptides from united biosystems. Purity is measured by HPLC, 220 nm, C18, and linear gradient. MW(theo.) is the theoretical molecular weight.

```
clc; clear; close all;
rng default % For reproducibility

% Eq. (6) in the paper
pmu_sig_x = @(mu, sigma, X)prod(0.5*(erf((X-mu)/sqrt(2)/sigma) - erf((X./2-mu)/sqrt(2)/sigma)));

% MIC data from triplicates
obs = [0.2,0.2,0.2; 0.4,0.2,0.2; 0.4,0.4,0.4; 0.8,0.4,0.4; 0.8,0.8,0.4; 0.8,0.8,0.8; 1.6,0.8,0.8;
       1.6,1.6,0.8; 1.6,1.6,1.6; 3.2,1.6,0.8; 3.2,1.6,1.6; 3.2,3.2,1.6; 3.2,3.2,3.2; 6.4,3.2,3.2;
       6.4,6.4,3.2; 6.4,6.4,6.4; 12.8,6.4,6.4; 12.8,12.8,6.4; 12.8,12.8,12.8];

exp = length(obs(:,1)); n = 2000; mus = linspace(0, 12.8, n); ys = zeros(exp, n);

% Numerical integration of sigma
for k = 1:exp
    for i = 1:n
        ys(k,i) = integral(@(sigma)arrayfun(@(sigma)pmu_sig_x(mus(i),sigma, obs(k,:)), sigma), ...
            0.001, 40);
    end
end
ys = ys./sum(ys,2); % normalize mu distribution

subplot(2, 1, 1)
plot(mus, ys);

subplot(2, 1, 2)
sumys=cumsum(ys,2); plot(mus, sumys);

% find mode of each mu distribution
p_max = zeros(exp,1); idx = zeros(exp,1);

for i = 1:exp
    [p_max(i),idx(i)]=max(ys(i,:));
end

mu = mus(idx)'; % store MIC values in variable mu

% find standard error
sig=zeros(exp,1);

for i = 1:exp
    temp = find(sumys(i,:)<0.841 & sumys(i,:)>0.159);
    sig(i) = 0.5*(mus(temp(end))-mus(temp(1)));
end

% Hypothesis test of the MIC values
sample=5000; trial=3; mu_stat = zeros(exp,trial); mean_stat = zeros(exp,sample);

for s = 1:sample
    rng_i = rand(exp,trial);
    for i = 1:exp
        for j = 1:trial
            if(rng_i(i,j)<sumys(i,1)), rng_i(i,j)=sumys(i,1); end
            mu_stat(i,j)=mus(find(sumys(i,:)≤rng_i(i,j), 1, 'last'));
        end
    end
    mean_stat(:,s) = mean(mu_stat,2)';
end

% calculate p-values, change index number to compare different sets of data
p = sum(mean_stat(1,:) < mean_stat(3,:))/sample;
```

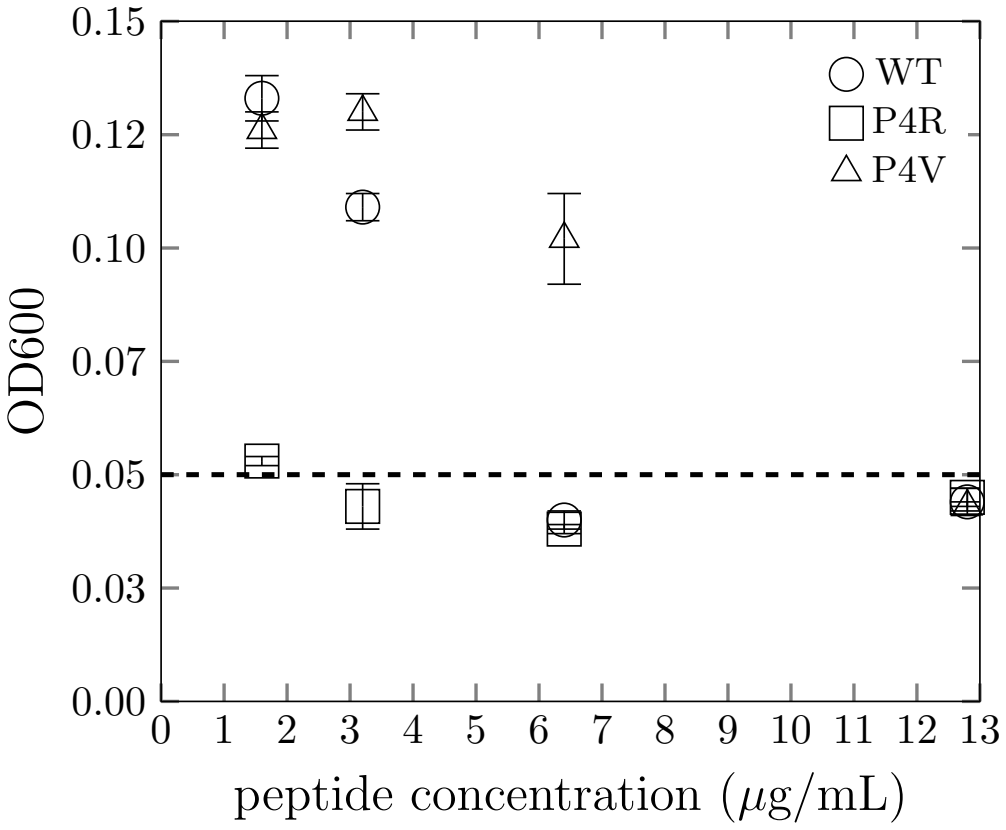

Figure S2 OD600 values of three oncocin variants including the wild-type (WT), an improved mutant P4R, and a hindered mutant P4V at different concentrations. The values were averaged from a triplicate. The dashed line at 0.05 is the cutoff to determine if there is bacteria growth.

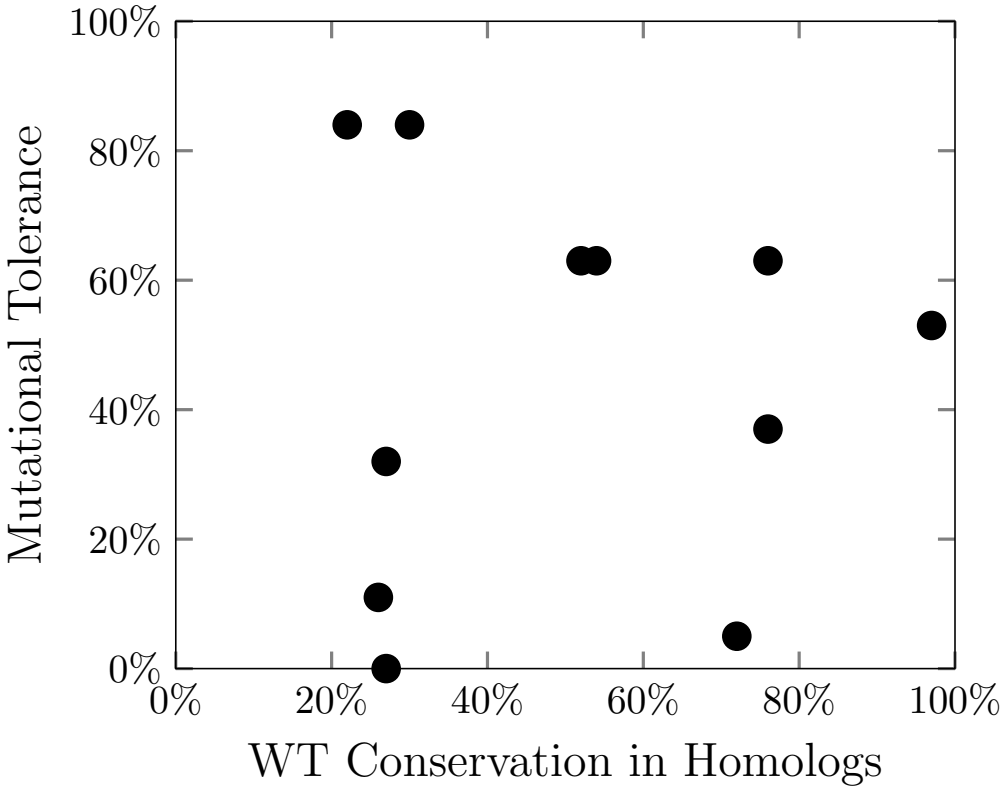

Figure S3 Mutational tolerance does not strongly negatively correlate with wild-type conservation in natural homologs. Mutational tolerance is computed as the frequency of single mutations at a site that exhibit MIC values that are not twofold weaker than wild-type (via data in Figure 2). Wild-type conservation in homologs is computed as the frequency of that site's wild-type amino acid within 35 homologs identified via BLAST of the oncocin sequence.

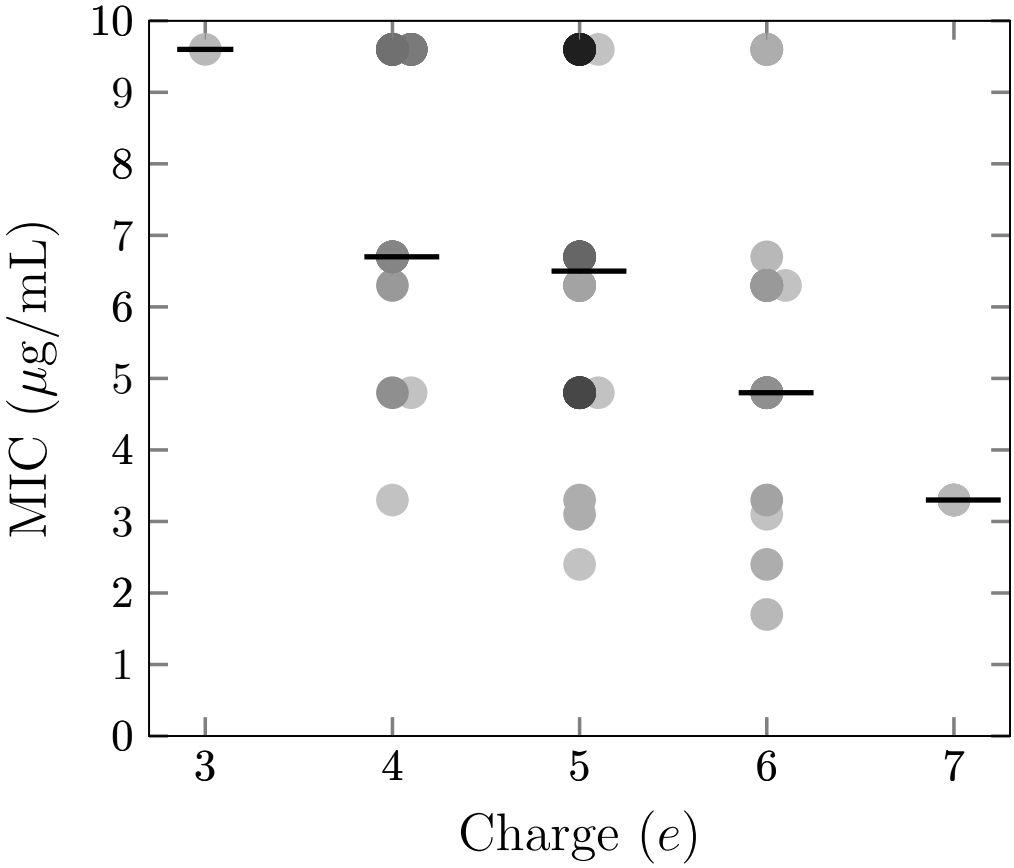

Figure S4 Peptide activity varies with net charge. MIC on *E. coli* JW0013 is plotted against net charge for 140 monosubstituted variants of Onc18 produced via SPOT synthesis. The darkness of each data point is proportional to the number of overlapping peptides at that MIC/charge combination. Lines indicate median values.

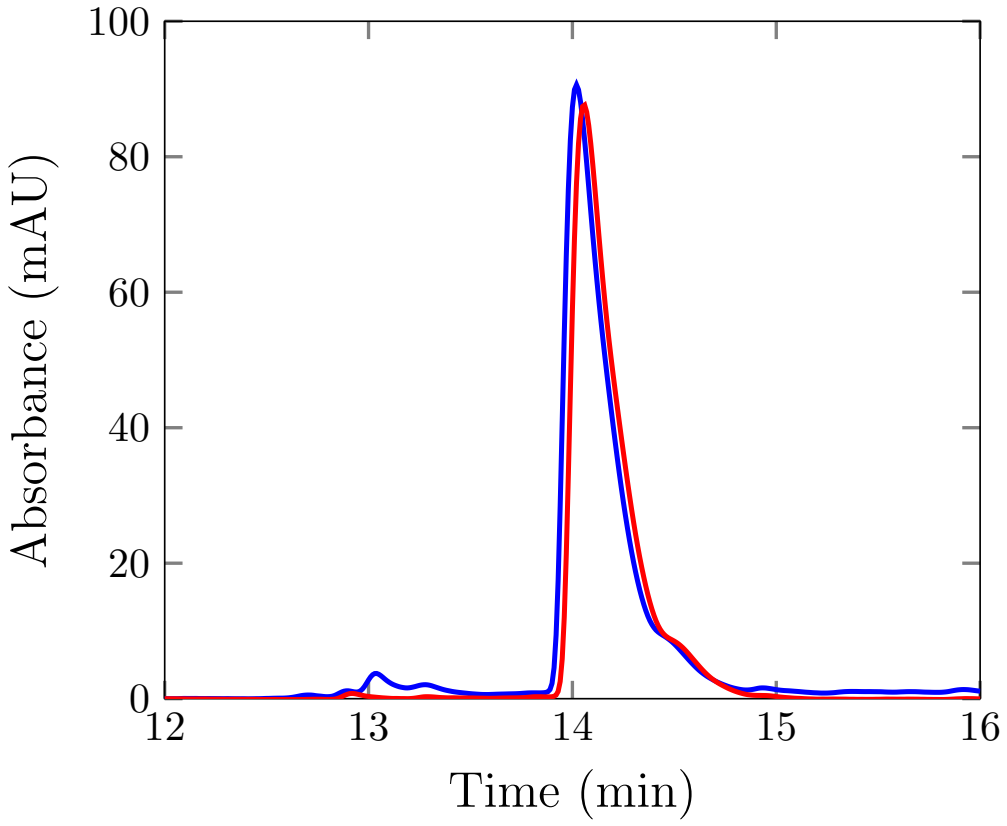

Figure S5 The blue curve represents the sample that had P4K/L7K incubated with cell-free supernatant for 2 hrs. The cell-free supernatant was prepared using a centrifuge from a cell culture grown in 20% M9 broth for 2 hours. The red curve is a control with P4K/L7K in the sterile water for 2 hrs. The peak area of the blue curve was 22.5 (mAU $\times$ min), and the peak area of the red curve was 21.9 (mAU $\times$ min).

Table S3 Antimicrobial activity of the multisubstituted oncocin derivatives from SPOT synthesis against *E. coli* strain JW0013. The substituted residues are underlined. The net charge of histidine is assigned 0.1

Sequence (First 11 amino acids)	Net Charge	MIC [μg/mL]
VDKKPYLPRRR	+7	1.7 ± 0.8
VDKKPYKPRKR	+8	2.4 ± 0.6
VDKKPYRPRKR	+8	3.1 ± 1.3
VDKKPYRPRRR	+8	1.6 ± 0.7
VDKKPRLPRPR	+7	1.6 ± 0.7
VDKKPRLPRKR	+8	1.7 ± 0.8
VDKKPRLPRRR	+8	1.7 ± 0.8
VDKKPRKPRPR	+8	3.3 ± 1.5
VDKKPRKPRKR	+9	4.8 ± 1.1
VDKKPRKPRRR	+9	3.1 ± 1.3
VDKKPRRPRPR	+8	3.3 ± 1.5
VDKKPRRPRKR	+9	4.8 ± 1.1
VDKKPRRPRRR	+9	4.8 ± 1.1
VDKKHYLPRPR	+6.1	2.4 ± 0.6
VDKKHYLPRKR	+7.1	1.7 ± 0.8
VDKKHYLPRRR	+7.1	2.4 ± 0.6
VDKKHYKPRKR	+8.1	3.1 ± 1.3
VDKKHYKPRRR	+8.1	4.8 ± 1.1
VDKKHYRPRPR	+7.1	1.6 ± 0.7
VDKKHYRPRKR	+8.1	3.1 ± 1.3
VDKKHYRPRRR	+8.1	2.4 ± 0.6
VDKKHRLPRPR	+7.1	2.4 ± 0.6
VDKKHRLPRKR	+8.1	2.4 ± 0.6
VDKKHRLPRRR	+8.1	1.7 ± 0.8
VDKKHRKPRPR	+8.1	3.3 ± 1.5
VDKKHRKPRKR	+9.1	3.1 ± 1.3
VDKKHRKPRRR	+9.1	3.1 ± 1.3
VDKKHRRPRPR	+8.1	3.1 ± 1.3
VDKKHRRPRKR	+9.1	3.1 ± 1.3
VDKKHRRPRRR	+9.1	2.4 ± 0.6

(Table S3 continued)

Sequence (First 11 amino acids)	Net Charge	MIC [μg/mL]
VDKRPYLPRKR	+7	1.7 ± 0.8
VDKRPYLPRRR	+7	1.6 ± 0.7
VDKRPYKPRKR	+8	3.1 ± 1.3
VDKRPYKPRRR	+8	1.7 ± 0.8
VDKRPYRPRKR	+8	2.4 ± 0.6
VDKRPYRPRRR	+8	2.4 ± 0.6
VDKRPRLPRPR	+7	1.3 ± 1.0
VDKRPRLPRKR	+8	1.6 ± 0.7
VDKRPRLPRRR	+8	2.4 ± 0.6
VDKRPRKPRPR	+8	3.3 ± 1.5
VDKRPRKPRKR	+9	3.3 ± 1.5
VDKRPRKPRRR	+9	4.8 ± 1.1
VDKRPRRPRPR	+8	2.4 ± 0.6
VDKRPRRPRKR	+9	3.3 ± 1.5
VDKRPRRPRRR	+9	3.3 ± 1.5
VDKRHYLPRPR	+6.1	1.7 ± 0.8
VDKRHYLPRKR	+7.1	1.6 ± 0.7
VDKRHYLPRRR	+7.1	1.6 ± 0.7
VDKRHYKPRPR	+7.1	1.3 ± 1.0
VDKRHYKPRKR	+8.1	1.7 ± 0.8
VDKRHYKPRRR	+8.1	1.7 ± 0.8
VDKRHYRPRKR	+8.1	1.6 ± 0.7
VDKRHYRPRRR	+8.1	1.3 ± 1.0
VDKRHRLPRPR	+7.1	1.3 ± 1.0
VDKRHRLPRKR	+8.1	1.6 ± 0.7
VDKRHRLPRRR	+8.1	1.6 ± 0.7
VDKRHRKPRPR	+8.1	1.7 ± 0.8
VDKRHRKPRKR	+9.1	1.7 ± 0.8
VDKRHRKPRRR	+9.1	3.1 ± 1.3
VDKRHRRPRPR	+8.1	3.3 ± 1.5
VDKRHRRPRKR	+9.1	2.4 ± 0.6
VDKRHRRPRRR	+9.1	2.4 ± 0.6