SUPPORTING INFORMATION

pH and temperature dependent peptide binding to the *Lactococcus lactis* oligopeptide-binding protein A (OppA) measured with a fluorescence anisotropy assay

Stevie Norcross$^{1,2,3}$, Ashwin Sunderraj$^1$, and Mathew Tantama$^{1,2,3,*}$

$^1$Department of Chemistry,
$^2$Institute for Integrative Neuroscience,
$^3$Institute for Inflammation, Immunology, and Infectious Disease,
Purdue University,
560 Oval Drive Box 68,
West Lafayette, IN 47907, USA

* Corresponding author
E-mail: mtantama@purdue.edu
Phone: 765-494-5312
Determination of $K_D$ of the Unlabeled Competitor: Fitting to a Competitive Binding Model

We directly fitted competitor dose response data to a two ligand complete competition binding model because the Cheng-Prusoff equation is not appropriate. The Cheng-Prusoff equation ($IC_{50} = K_i (1 + [S]/K_M)$) is used in ideal cases in which it is known a priori that the competitor is not high affinity. In our competition binding assays, we make no assumption about the binding affinity of the unlabeled competitor. Furthermore, we designed our assay conditions such that the concentration of the OppA protein is in excess of the labeled bradykinin-SR101 (BK-SR101) peptide in order to achieve approximately 50% occupancy for good sensitivity and a large dynamic range for fluorescence anisotropy changes. Therefore, we fitted our competition dose-response data directly to the complete competitive binding model described by Wang (FEBS Lett. 1995, 360:111) and Roehrl et al. (Biochemistry 2004, 43:16056).

![Diagram of competitive binding model]

We define in molar units: $P$, the concentration of unbound free OppA protein; $L$, the concentration of unbound free labeled BK-SR101 peptide; $U$, the concentration of unbound free unlabeled competitor; $PL$, the concentration of the labeled OppA-(BK-SR101) complex; $PU$, the concentration of the unlabeled OppA-competitor complex; $PT = P + PU + PL$, the total OppA protein concentration; $LT = L + PL$, the total labeled BK-SR101 peptide concentration; $UT = U + PU$, the total unlabeled competitor concentration; $KL$, the dissociation constant for the labeled OppA-(BK-SR101) complex; $KU$, the dissociation constant for the unlabeled OppA-competitor complex. As previously described by Wang (FEBS Lett. 1995, 360:111) and Roehrl et al. (Biochemistry 2004, 43:16056), an analytical solution can be written for the fraction of labeled BK-SR101 peptide that is bound to the OppA protein ($F_{PL}$):

$$F_{PL} = \frac{PL}{PT} = \frac{(2\sqrt{(d^2 - 3e)} \cdot \cos(\theta/3) - d)}{3KL + (2\sqrt{(d^2 - 3e)} \cdot \cos(\theta/3) - d)}$$

$$d = KL + KU + LT + UT - PT$$

$$e = (UT - PT) \cdot KL + (LT - PT) \cdot KU + KL \cdot KU$$

$$f = -KL \cdot KU \cdot PT$$
\[ \theta = \cos^{-1} \left( \frac{-2d^3 + 9de - 27f}{2 \sqrt{(d^2 - 3e)^3}} \right) \]

For our assay conditions at each pH and temperature, we experimentally determined the minimum anisotropy, \( r_{\text{min}} \), of the labeled BK-SR101 alone. We also determined the maximum anisotropy, \( r_{\text{max}} \), of the labeled OppA-(BK-SR101) complex without any competitor present. Therefore, we can write a standard saturation binding equation for the total anisotropy (\( r \)):

\[ r = r_{\text{min}} + (r_{\text{max}} - r_{\text{min}}) \cdot F_{\text{PL}} \]

By rearranging this equation, anisotropy from the competitor dose-response experiments was converted to the fraction bound labeled BK-SR101:

\[ F_{\text{PL}} = \frac{r - r_{\text{min}}}{r_{\text{max}} - r_{\text{min}}} \]

We used the Matlab non-linear least squares fitting function \texttt{lsqcurvefit} to fit our competition dose-response data and determine the unlabeled competitor dissociation constant (referred to as the "fitted \( K_U \)"). We compared the Trust Region Reflective and Levenberg-Marquardt optimization algorithms. Both algorithms produced similar results, but the Trust Region Reflective converged more robustly, and therefore we report results using the default Trust Region Reflective algorithm in Table I and Table III in the main text. Matlab code is included at the end of the Supplemental Information. Dose response data for each replicate were fitted (Figures S4 and S5), and then the fitted \( K_U \) values were averaged and reported in in Table I and Table III in the main text.

We achieve greater accuracy in determining \( K_U \) by directly fitting dose-response data to the complete competition binding model, especially when the Cheng-Prusoff equation is not valid. However, we expect that there will still be limitations to this fitting approach determined by experimental assay parameters. In particular, we expect that when the true competitor affinity is very high beyond a certain threshold, the fitted \( K_U \) values will become indistinguishable. In order to determine this threshold (effectively the lower limit of fitted \( K_U \) estimation), we simulated anisotropy dose-response data in which we varied the theoretical \( K_U \) value for the unlabeled competitor. We then fit the simulated data and determined at what simulated \( K_U \) values the fitted \( K_U \) values no longer reliably estimate the simulated \( K_U \) values. Across different pH and temperature, the OppA-(BK-SR101) binding affinity was approximately \( K_L = 10 \mu M \), which we set as an experimental constant. We therefore explored \( K_U \) values \( \pm 3 \log_{10} \) units around \( K_L \) by simulating data for \( K_U = 10^{-8} \) M to \( 10^{-2} \) M at 0.1 \( \log_{10} \) unit increments. We also set experimental assay parameters for total OppA protein concentration \( P_T = 11.25 \mu M \), total labeled BK-SR101 peptide concentration \( L_T = 0.75 \mu M \), and the approximate competitor peptide concentration range \( U_T = 10^{-8} \) M to \( 10^{-2} \) M. The standard deviation for our replicate anisotropy measurements was approximately 2% of the mean, and for a normal distribution, the full width at half maximum (FWHM) is \( 2 \cdot (2 \ln 2)^{1/2} \) times the standard deviation. Therefore to achieve an approximate 5% FWHM, we added normally distributed random numbers scaled to a maximum percent error range
of -10% to 10% to the simulated data. Using these experimental parameters, for each theoretical $K_U$ value we simulated 1000 dose-response data sets and then fitted the simulated data to determine the fitted $K_U$ value (Figure S6). We then analyzed the relationship between the fitted $K_U$ versus the simulated $K_U$ (Figure S6). Below a simulated $K_U$ value of $10^{-7}$ M, fitting became more error prone. We then estimated the probability that the fitted $K_U$ was within an N-fold range of the simulated $K_U$ (Figure S6). At a simulated $K_U$ value of $10^{-7}$ M, there was only a 50% probability that the fitted $K_U$ would be within 2-fold of the simulated $K_U$ value. Therefore, we made the conservative decision that for any experimental data that results in a fitted $K_U \leq 10^{-7}$ M, we consider it a “high affinity” competitor and we report $K_U \leq 10^{-7}$ M in Table III in the main text.
**Figure S1.** N-terminal sequence removed from the *Lactococcus lactis* OppA to prevent palmitoylation, where C is the N-palmitoyl cysteine.

**Figure S2.** Protein sequence for the OppA construct. Color: 6x His tag and T7 leader from pRSETB vector, OppA.

**Figure S3.** N-terminally labeled bradykinin peptide (SR101-bradykinin) structure, where the bradykinin peptide sequence is shown in bold (RPPGFSPFR).
Figure S4. Fluorescence anisotropy competition dose-response data was fitted using a complete competition model to determine the $K_D$ values of the unlabeled competitor peptides (referred to as fitted $K_U$ values in the Supplemental Discussion). Data from Figure 3 in the main text is replotted with the average fitted $K_U$ curves.
Figure S5. Fluorescence anisotropy competition dose-response data was fitted using a complete competition model to determine the $K_D$ values for unlabeled bradykinin at different pH and temperature (referred to as fitted $K_U$ values in the Supplemental Discussion). Data from Figure 5 in the main text is replotted with the average fitted $K_U$ curves.
Figure S6. Determining the limit of $K_U$ estimation under experimental conditions. (A) An example of 1 simulated dose-response data set: simulated $K_U = 10.0 \, \mu M$; fitted $K_U = 9.6 \pm 3.2 \, \mu M$ (mean ± 95% confidence interval). (B) The fitted $K_U$ values from 1000 simulated data sets is plotted against the simulated $K_U$ value. Below a simulated $K_U$ value of 0.1 µM, fitting and estimation cannot distinguish between $K_U$ values for high affinity competitors. (C) The 1000 simulated data sets were analyzed to estimate the probability that the fitted $K_U$ valued would fall within N-fold of the true simulated $K_U$ value.
Figure S7. OppA protein aggregation after ramping to 55°C for a bradykinin-SR101 direct dose response assay at pH 6. Wells shown in triplicate, vertically. Column 7 = 8 μM, column 8 = 15 μM, column 9 = 25 μM, column 10 = 45 μM, column 11 = 100 μM OppA.
Figure S8. Chromatograph for size exclusion chromatography (SEC) purification of OppA protein. The bottom trace is a magnified view of the peak.
Figure S9. Time course showing fluorescence anisotropy reaches steady-state over the equilibration time for a direct binding assay with bradykinin-SR101 at pH 6, 30°C. Lines shown from bottom to top indicate OppA concentrations of 0.1, 0.5, 2, 5, 8, 15, 25, 45 and 100 μM.

Figure S10. Electrostatic surface models for OppA in the closed state without any ligand present calculated at different pH. Surface charge was calculated and represented using PROPKA and APBS for the L. lactis OppA (PDB 3DRG) closed conformation with the bradykinin ligand removed from the structure prior to calculation.
Figure S11. (A) N-terminal sequence removed from the *Bacillus subtilis* AppA to prevent palmitoylation, where C is the N-palmitoyl cysteine. (B) Protein sequence for the AppA construct. Color: 6x His tag and T7 leader from pRSETB vector, AppA. (C) Direct binding of labeled bradykinin to *B. subtilis* AppA at pH 7, 37°C measured by fluorescence anisotropy. Bradykinin labeled with sulforhodamine 101 on the N-terminus (SR101-RPPGFSPFR) exhibited higher affinity for AppA (K_D = 50 μM) compared to bradykinin labeled on the C-terminus (RPPGFSPFRK-SR101) (K_D = 71 μM). Lines show the fit for n = 1. (D) Temperature dependence of AppA peptide binding affinity. Direct binding of SR101-bradykinin was measured by fluorescence anisotropy at pH 7 and varying temperatures. Lines show the average fit for n = 3. Error bars are stdev. (E) Bradykinin competition binding to AppA measured by fluorescence anisotropy at pH 7, 37°C. Bradykinin is able to completely displace SR101-bradykinin at an IC50 of 4.4 ± 0.2 μM. Lines show the average fit for n = 3. Error bars are stdev.

Table S1. Characterization of AppA direct binding of SR101-bradykinin at pH 7, varying temperature (n = 3, mean ± stdev). * Not determined because of protein aggregation.

| Temperature (°C) | K_D (μM) |
|-----------------|----------|
| 24              | 67 ± 5   |
| 30              | 78 ± 2   |
| 37              | 75 ± 4   |
| 45              | 230 ± 60 |
| 55              | - *      |
SUPPORTING METHODS

AppA Plasmid Construct The *Bacillus subtilis* oligopeptide-binding protein (AppA) amino acid sequence is from subspecies subtilis strain 168 (GenBank accession AAA62358.1). Similar to OppA, the AppA N-terminal signal peptide for palmitoylation and surface tethering was removed (Figure S11) and the nucleotide sequence was synthesized as a gBlock. The AppA gBlock was cloned into a pRSETB vector by Gibson Assembly for bacterial expression and purification (Figure S11).

AppA Expression and Purification The *Bacillus subtilis* AppA was expressed and purified similar to OppA from *Lactococcus lactis*, except that large AIM cultures (250 mL) were grown at 37°C overnight (~16 hrs), which gave us the highest BsAppA expression levels. Concentration of purified AppA was determined by absorbance at 280 nm using an extinction coefficient of 97860 M⁻¹*cm⁻¹ calculated using the Northwestern University Peptide Properties Calculator.

AppA Fluorescence Anisotropy Assay Preliminary direct dose response binding for AppA was performed the same as the direct dose response binding for OppA. However, the AppA anisotropy assays were performed at 37°C in pH 7 Assay Buffer. SR101-bradykinin was selected for all following AppA anisotropy assays because of its higher affinity for AppA compared to bradykinin-SR101 (Figure S11). Follow up direct dose response assays were performed at varying temperatures by combining 100 μL of 1.5 μM SR101-RPPGFSFPR and 50 μL purified AppA to wells in a 96-well non-binding microplate (Figure S11). Final concentrations of AppA varied from 0.01 to 150 μM.

Competition dose response binding was performed similar to OppA, except a mixture of 1 μM SR101-bradykinin and 25 μM AppA was prepared, and 150 μL of the mixture was added to each well in a 96-well non-binding microplate. A higher concentration of AppA was used due to its lower affinity for the dye-labeled peptide (Table S1), however, this assay should be further optimized to improve the dynamic range (Figure S11).
MATLAB CODE FOR FITTING COMPETITION DOSE RESPONSE DATA

NorcrossMyFunFractionBound.m

```matlab
function y=NorcrossMyFunFractionBound(KU,UT,KL,PT,LT)
% Output: y=Fraction Bound of labeled ligand for a complete direct
% competitive competition fluorescence anisotropy assay.
% Inputs: KU = dissociation constant for unlabeled competitor
% UT = Total concentration of unlabeled competitor
% KL = dissociation constant for fluorescently labeled ligand
% PT = Total concentration of protein
% LT = Total concentration of labeled ligand
% M.Tantama 12/10/2018

localVarD=KL+KU+LT+UT-PT;
localVarE=(UT-PT)*KL+(LT-PT)*KU+KL*KU;
localVarF=-KL*KU*PT;
localVarG=localVarD.^2-3*localVarE;
localVarTheta=acos((-2*localVarD.^3+9*localVarD.*localVarE-
27*localVarF)./(2*sqrt(localVarG.^3)));
localVarH=2*sqrt(localVarG).*cos(localVarTheta/3)-localVarD;
y=localVarH./(3*KL+localVarH);

% END NorcrossMyFunFractionBound
```

NorcrossFitCompetitionData.m

```matlab
% Non-linear least squares fitting of competition binding data
% M.Tantama 12/10/2018
%
% KU = dissociation constant for unlabeled competitor (Molar)
% UT = Total concentration of unlabeled competitor (Molar)
% UT can be a (1 x m) row vector
% KL = dissociation constant for fluorescently labeled ligand (Molar)
% PT = Total concentration of protein (Molar)
% LT = Total concentration of labeled ligand (Molar)

% Parameters to be passed to NorcrossMyFunFractionBound during fitting
PT = 11.25e-6; %Total concentration OppA protein in assay
LT = 0.75e-6; %Total concentration labeled BK-SR101 peptide in assay

% Initial guess for KU and bounds
x0KU = 1e-9;
lbKU = 0;
ubKU = 1;

% Read in data from Excel file
% Needs to have a sheet labeled 'data' with first column of headers and
% data arrange by rows. First row should be log10[competitor] ending with a
% final dummy 0. Subsequent rows below should have fraction bound values
% already calculated from anistropy values followed by the Kd of the
% labeled BK-SR101 for the given pH and temperature.
[numdata,txtdata]=xlsread('fpdata.xlsx','data');
xvals=numdata(1,1:size(numdata,2)-1);
```
KLs=numdata(:,size(numdata,2));             %Grabs the KL values
resultsholder=zeros(size(numdata,1)-1,3);
for dataindex=2:size(numdata,1);
    yvals=numdata(dataindex,1:size(numdata,2)-1);
    KL=KLs(dataindex);
    expfunc=@(xvar,yvar)NorcrossMyFunFractionBound(xvar,yvar,KL,PT,LT);
    [fittedKU,resnorm,residual,exitflag,output,lambda,jac]=lsqcurvefit(expfunc,x0
    KU,10.^xvals,yvals,lbKU,ubKU);
    ci=nlparci(fittedKU,residual,'jacobian',jac);   %nlparci requires stats
    resultsholder(dataindex-1,:)=[fittedKU, cierror, resnorm];
end; %dataindex
NorcrossFittingLimits.m
% Simulate and fit direct competition binding data.
% M.Tantama 12/10/2018
% % KU = dissociation constant for unlabeled competitor (Molar)
% % UT = Total concentration of unlabeled competitor (Molar)
% % KL = dissociation constant for fluorescently labeled ligand (Molar)
% % PT = Total concentration of protein (Molar)
% % LT = Total concentration of labeled ligand (Molar)

KLorder=-5;
KL = 10.^KLorder;       %Kd labeled BK-SR101
PT = 11.25e-6;          %Total concentration OppA protein in assay
LT = 0.75e-6;           %Total concentration lableled BK-SR101 in assay
results=[];
for KUindex=(KLorder+3):-0.1:(KLorder-3),
  for repindex=1:1000,
    % Generate test data with random noise
    testdata=[]; xvals=[];
    testKU=10.^KUindex; %Simulated KU
    maxLogU = -2; %Max log10[total unlabeled competitor]
    minLogU = -8; %Min log10[total unlabeled competitor]
    xLogU=[minLogU:0.5:maxLogU]; %[total unlabeled competitor]
    for index=1:5;
      fracbnd=NorcrossMyFunFractionBound(testKU,10.^xLogU,KL,PT,LT);
      xvals=[xvals,xLogU];
      testdata=[testdata,fracbnd+(-1+2*rand(size(fracbnd)))/10];
    end;
    % Initial guess for KU and bounds
    x0KU = testKU/10;
    lbKU = 0;
    ubKU = 1;
    % Create anonymous function and pass the experimental parameters
    % KL, PT, LT as constants
    expfunc=@(xvar,yvar)NorcrossMyFunFractionBound(xvar,yvar,KL,PT,LT);
    % Fit and 95% confidence intervals
    [fittedKU,resnorm,residual,exitflag,output,lambda,jac]=lsqcurvefit(expfunc,x0
    KU,10.^xvals,testdata,lbKU,ubKU);
    ci=nlparci(fittedKU,residual,'jacobian',jac); %nlparci requires
    stats toolbox, return 95% conf. int.
    cierror=ci(2)-fittedKU;
    results=[results;KL,testKU,fittedKU,cierror,resnorm];
  end;%repindex
  % You can uncomment below to plot simulated data and fitted curves
  % but reduce the number of replicate data sets repindex at line 21
  % otherwise you will generate thousands of figures.
  % %Create fitted curve
  % fitxmin=floor(min(xvals))-0.5;
  % fitxmax=ceil(max(xvals))+0.5;
  % fitxvals=[fitxmin:0.1:fitxmax];
  % fityvals=NorcrossMyFunFractionBound(fittedKU,10.^fitxvals,KL,PT,LT);
  % %Plot
  % figure;
  % plot(xvals,testdata,'ko');
  % hold on;
  % xlabel('log10[Total Unlabeled]');
  % ylabel('Fraction Labeled Bound');
plot(fitxvals,fityvals,'r--');
legend(['Test Data KU=',num2str(testKU,1)],'Fit');
title(['fittedKU,3,+/-',num2str(cierror,3),'
(',num2str(cierror/fittedKU,3),',')]);
hold off;
end; %KUindex