

A graduate student in your research lab is studying the importance of a three-way junction RNA secondary structure in terms of the ability of this RNA to be bound by protein G. She brings you the following data regarding the significance of base pair $G_{11} \bullet C_{22}$.

nucleotides at positions 11-22	K_D (nM)*
$G_{11} C_{22}$	2
$G_{11} G_{22}$	2000

*These dissociation constants were obtained by fitting the fraction of RNA bound at various protein concentrations (0 to 10 μ M) with a standard binding isotherm, assuming one RNA binding site per protein G.

What would you conclude from these data?

The experimental results from the experiment in which the $G_{11} C_{22}$ base pair was destroyed by changing the 22nd nucleotide to G *suggests* to me that the base pair is essential for recognition by protein G. With the base pair destroyed, the K_D increases by 1000-fold (which is weaker binding). However, the fact remains that it could be the actual identity of the 22nd nucleotide matters more than the potential for a base pair at this position. I would therefore do the next experiments as listed below.

What experiments would you suggest that your graduate student perform next and why?

First, I would make the $C_{11} C_{22}$ variant to see the actual sequence at the 11th position matters in addition to the fact that a base pair exists at this location. Then I would make the $C_{11} G_{22}$ variant to see if restoring the base pair restores activity.

AFTER these experiments I could try these:

- Maybe try an AU base substitution to see if this relatively weaker base pair (2 rather than 3 H-bonds holding it together) retains affinity for protein G binding.
- Maybe look at mutating the bases above or below the G_{11} - C_{22} base pair to see if base stacking is important or to see if an entire stem is important here (rather than just an isolated base pair).

What would you say to your graduate student if these data had been obtained using a maximum protein concentration of 2.5 μ M?

I would suggest that the K_D values for the G_{11} - G_{22} and C_{11} - C_{22} variants are not valid. They certainly still suggest much weaker binding for these two variants so the overall conclusions are fine, but they can't be accurate numbers.

I say this because if you think of the hyperbolic binding curve that we studied in class, you need numbers well above and below the K_D value in order to have a complete, fit-able curve. If your protein concentrations only go roughly as high as the K_D you can really only say that this is an estimate, or perhaps a lower limit, for the true value of K_D .

So this question was asking you to think about what data you need to have to determine the K_D .