

Magn. Reson. Discuss., referee comment RC2  
<https://doi.org/10.5194/mr-2022-5-RC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on mr-2022-5

Anonymous Referee #2

---

Referee comment on "An improved, time-efficient approach to extract accurate distance restraints for  $NMR^2$  structure calculation" by Aditya Pokharna et al., Magn. Reson. Discuss., <https://doi.org/10.5194/mr-2022-5-RC2>, 2022

---

This manuscript describes an extension of recent work by the same group on the method they call NMR molecular replacement, or  $NMR^2$ . As they have described in previous papers cited in this manuscript, in the  $NMR^2$  approach the structure of a protein-ligand complex in solution is determined using a combination of intermolecular NOE data from the complex and a previously determined structure (often a crystal structure) of the free protein. Much of the interest in this method derives from its potential for use in drug discovery; particularly since the method does not involve making independent assignments for the protein NMR resonances it can in principle provide information on a timescale that is useful in the context of a ligand screening program.

However, in the context of a fragment-based drug development strategy the ligand fragments tested will often be very small molecules, with weak binding affinities and few NMR resonances, implying that the NOE data are very sparse and possibly difficult to measure. A related issue is that distance calibration of a limited set of intermolecular NOEs poses significant challenges, as there are no fixed, known intermolecular distances that can be used as standards (as would be done for intramolecular NOE calibration) and the methods employed instead, such as the ENORA approach previously published by members of the same group, are less direct and may give inaccurate distance restraints. In this manuscript the authors describe an alternative approach for calibrating intermolecular distance restraints, based on comparison of cross-peak intensities with those of diagonal peaks in the same spectrum. This results in shorter distance restraints than those obtained using ENORA, which, at least in the test system used in this manuscript, results in the ligand being pulled more deeply into the binding pocket on the protein, although still not as deeply as is seen in the related crystal structure of a closely-related complex in this case. This may mean the new solution structure resembles the true structure more closely than did the previous solution structure, though one should probably not assume that the true solution structure is identical to the crystal structure for such a weak complex.

I think the manuscript potentially represents a worthwhile step forward and should be published, subject to the following points being satisfactorily addressed:

1) The manuscript describes results for one ligand, called Compound 1, but there seems to be some confusion in Figure 2 since this shows a different ligand, at least in the case of the crystal structure. The earlier paper from the same group by Torres et al. (2020) employing the ENORA approach describes results for three closely related ligands that differ only in the substituent at position 3 of the benzene ring: in Compound 1 it is CH<sub>3</sub>, in Compound 2 it is Cl and in Compound 3 it is CF<sub>3</sub>. The crystal structure PDB 2XP6 on which Figure 2 is based definitely contains Compound 2, yet the caption to Figure 2 in the present manuscript refers only to Compound 1. Moreover, all three copies of the ligand shown in the figure have a green atom at position 3 of the benzene ring that the caption states represents chlorine, but this cannot be correct for Compound 1, which was the subject of all the NMR calculations in the present manuscript. These discrepancies need to be resolved. Maybe this is just a colouring mistake when preparing the figure, but even if this is corrected the wording in main text and caption needs to be adjusted to clarify that the crystal structure contains Compound 2, not 1, and consequently the comparison shown is not quite as direct as it currently seems. For instance, the affinity of Compound 2 is weaker than that of Compound 1, which could perhaps influence how deeply into the binding site it binds.

2) Torres et al. (2020) states that for NMR data collection for the complex of PIN1 with Compound 1, the protein concentration was 1.3 mM, the ligand concentration was 2.5 mM and the K<sub>d</sub> was 260 μM. Using the quadratic binding equation for an A + B  $\rightleftharpoons$  AB equilibrium, it is straightforward to calculate that under these conditions the protein is 84% in the bound state and the ligand is 44% in the bound state. Figures S2 and S3 of Torres et al. make clear that, as would be expected, the system is in fast exchange on the chemical shift timescale, so the NMR signals used in the NOE experiments are all averages that contain substantial contributions from both the free and bound states, in line with the populations given above. However, eqs 1 to 5 make no reference to this, and as far as I can see from the cited papers (mainly Vögeli 2014) were originally derived for intramolecular systems where the complication of a free-bound equilibrium does not exist. Is it justified to use those equations without modification when dealing with an exchanging system where a substantial proportion of the observed signals is contributed by the free states? Even if it is, I think this point needs to be carefully explained (e.g. what assumptions/definitions does it rely on?) and justified with appropriate citations in the paper. (Also, though this is not an issue with the present manuscript, it is not fully clear to me either how the presence of exchanging bound and free states was accounted for in the ENORA approach in the Torres et al. paper - unless I have misunderstood it was done by calculating an "effective correlation time" for the complex that is a population-weighted average, but that does not quite make physical sense to me, nor is the calculation fully described in that paper). The step of converting the cross-relaxation rates into distance restraints is also not described in the present paper; I feel that at the least there should be a citation given that makes clear how this was done.

3) Although the authors did also test their new method with a second example (binding of caylin-1 to HDM2), as they point out themselves that was a less demanding case where the new approach made little difference to the outcome. However, in the Torres et al. 2020 paper the authors applied the original ENORA approach to the PIN1 complexes of the two more weakly binding ligands, Compounds 2 and 3, in addition to that of Compound 1. In those cases they found the ENORA method did not converge unless they assumed the assignments of the protein methyl signals receiving intermolecular NOEs were the same as those for the (better behaved) complex of Compound 1. An obvious question would therefore be, does the new method based on diagonal-normalised restraints improve the situation sufficiently that the complexes of Compounds 2 and 3 can be solved *without*

having to use the methyl assignments from the Compound 1 complex? Did the authors try this, and if so what did they find?

4) Looking at Figure 1, it is striking that the intermolecular NOE restraints involving the methyl of methionine 130 are significantly *shorter* than the corresponding distances in the crystal structure, whereas all the other such restraints are longer than the corresponding crystal distances. The authors mention this briefly, noting that the methionine is in a solution-exposed "floppy" region, but I did not find this a very plausible explanation; in my own experience floppy regions generally give weaker NOEs, not stronger ones. Is there perhaps a real difference between the crystal and solution structures? For instance, might it be that the methionine methyl has a hydrophobic interaction involving the methyl group of Compound 1 (the ligand used in the NMR experiments) that is missing in the complex of Compound 2 (the ligand present in the crystal structure) where the corresponding substituent is a chlorine? Looking at PDB 2XP6, it seems entirely plausible that the methionine sidechain could move to a position closer to the methyl group and that this could indeed shorten the two distances involved in the restraints in question. This could also perhaps be at least part of the reason why Compound 1 binds more tightly than Compound 2.

5) It would be useful to show a line drawing of the structure of Compound 1, and perhaps also of Compound 2, to help clarify the explanation of the system (not all readers will be able immediately to convert the IUPAC name given in the introduction to a structure). Also, it would be very useful to show a second orientation of the structural representation in Figure 2, as in the current view it is almost impossible to see how far into the binding pocket the two sets of NMR-based calculations have actually pulled the ligand. A rotation of  $+60^\circ$  about x would probably achieve a clearer view of this.

6) While using diagonal peak intensities may bring advantages in the analysis, it requires that the diagonal peaks are separately resolved, thus losing the resolution advantage of separating the peaks in multiple dimensions. Is this likely to be a difficulty in more general applications of the new approach, or does the expected sparsity of ligand signals mean there is unlikely to be a problem?