We thank the reviewers for their constructive and insightful comments and have revised our manuscript to address each of these as follows.

Reviewer #1

The very well written and well-illustrated manuscript by Robertson et al. from the Bax group entitled "Four-dimensional NOE-NOE spectroscopy of SARS-CoV-2 Main Protease to facilitate resonance assignment and structural analysis" describes two 4D [1H,1H]-NOESY NMR experiments and their application to a large protein system of 2×306 residues. The detailed precise description allows not only to reproduce the pulse sequences straight-forwardly, but also highlights the importance that are in the details such as the absence of phase distortion of the spectrum relevant for the application of NUS obtained by the introduction of two composite 1H 180° pulses, and the optimization of the phase cycle to 4 and the excitation of resonances only to the "amide side" of the water frequency to enable high resolution acquisition. Of high interest is the NOESY-NOESY-TROSY experiment reminiscent of the first 3D experiments (by the Kaptein group) because of its information content for assignment and distance restraint collection along with its inherent high signal to noise because it comprises a relaxation-based magnetisation transfer. This is well illustrated for the large system under study.

In short, the presented manuscript is recommended for publication. The only small suggestions perhaps are

(i) highlight the two composite pulses in the pulse sequence

Response: We agree that the drawing of the composite 180° pulses was confusing. They are now depicted the same as regular 180° pulses but colored green.

(ii) indicate, that the 4D NOESY-NOESY-TROSY does not comprise NOEs from amide protons that are upfield of the water resonance.

Response: We added "As a result, no NOE peaks from amide protons resonating near water or upfield from water were observed" (line 326-327).

Reviewer #2

The manuscript presents the setup and combined use of two 4D NOESY experiments for resonance assignment and and structural studies of larger proteins by NMR. The combination of 4D TROSY-NOESY-TROSY and selective 4D NOESY-NOESY-TROSY turns out to be essential to simplify the complex 1H-1H dipolar relaxation network of the 712 res. homodimeric main protease of SARS-CoV-2. The authors indicate that the presented experiments were invaluable to validate the resonance assignments and subsequent structural analysis of this medically highly relevant protein, an exciting result, the detailed results of which will be presented elsewhere. This manuscript focuses in particular on the recording of the two types of 4D NOE-based spectra and presents several examples.

To make the 4D recording feasible, the spectra were recorded using NUS sampling, with a sampling sparsity of 0.54% for the 4D TROSY-NOESY-TROSY and 1.69% for the 4D NOESY-NOESY-TROSY. The resulting spectral quality is excellent. I have to compliment the authors as well that they show and describe the pulse sequences in recommendable detail and provide the scripts and details for processing.

The description of the NUS usage and detailed description of its results makes this paper very valuable and opens practical implementation and realistic acquisition times of these 4D acquisition methods. However, I suggest to provide more detail for the NUS sampling in the supplement, to be able to reproduce the NUS time series, also for different or future instruments.

Response: We thank the reviewer for the kind comments. We have added the acquisition parameters as an SI table, and we have uploaded the time domain data, pulse programs and Bruker acquisition parameter files to https://zenodo.org/record/4625615

The 4D NOESY-NOESY-TROSY experiment, a concatenation of 3D NOESY-NOESY and 2D TROSY (or 2D NOESY and 3D NOESY-TROSY…) has not been demonstrated before, is new, as well as the combination of both 4D's for detailed analysis, which presents a major step forward for studying large proteins by NMR.

However, the 4D TROSY-NOESY-TROSY is not really new. It is rather an improved version of similar experiments, not only the 4D HMQC-NOESY-TROSY by the authors (Barnes et al, 2019) and 4D HMQC-NOESY-HMQC (Grzesiek et al, 1995), both referred to. There is also similarity to the 4D TROSY-NOESY-TROSY by Xia et al (2000, JBNMR 18, 261), which could be cited. Would there also have been an advantage of recording a diagonal-free version of this (Diercks et al 2010, JACS 132, 2138)?

Response: The TROSY-NOESY-TROSY by Xia et al is now cited, with a brief description of the minor difference between their experiment and ours, mainly our implementation of gradient encoding/decoding for both TROSY detections before and after the NOE mixing, thereby allowing us to reduce the phase cycling steps from 8 to 4 to maximize sampling of the time domain data. The clever Diercks diagonal suppression scheme was not implemented because, at the long mixing

times used, these diagonals provide convenient reference anchors during analysis and are sufficiently attenuated such as not to obscure nearby peaks or make the NUS reconstruction more difficult. The required increased length of the ST2-PT vs the INEPT transfer step was another minor consideration against using diagonal suppression. We now refer to the Diercks paper in lines 68-73.

I suggest to additionally deposit the pulse sequences, parameters and processing scripts of the supplement electronically. This would also allow to add the file bits.jfy.

Response: We now have uploaded the bits.jfy file, pulse programs, acquisition parameter files and processing scripts, together with the raw data sets on the zenodo.org webserver.

Some detailed remarks:

p.4 l.103/104 Writing it this way "data matrix consists of 1536 x 90* x 91* x 90* complex points" can be confusing. It reads as if 90*90*91*8 fids were recorded, which with ns=4 would be a 1.5 yr experiment. In fact, the experiment recorded only 31896 fids and thus the full 'data' matrix (if created at all) is largely empty (How does Bruker store the sparse NUS data? Is the large sparse data matrix indeed filled with zeros? Even for archiving storage?).*

Response: Thanks. We have added SI Table S1 that lists the number of sampled data points and other pertinent information explicitly. For all NUS experiments, only the data points selected in a NUS list were recorded. Bruker only stores the recorded data in the order they were acquired, it does not store zeros for those points that are not selected. During the processing within NMRPipe, however, the recorded FIDs are first sorted according to the sampling schedule and the missing data points are filled with zeros, which are subsequently replaced with the reconstructed values by SMILE.

p.4 l.104 91 (1H) t2 increments is bit counterintuitive number, which may occur when t2 was the last incremented time and the measurement halted, though not unusual when planning subsequent measurements. Since it could also be a stalling measurement related to an instrument hardware/software issue, into which others may also run when using this sequence and parameter setup. Was it coincidence, planned, or is a warning footnote in the supplement at place?*

Response: The number of maximum t_2 ^{1}H increments was set to 91 in order to reach an acquisition time of 20 msec. We indicated there were 91* complex data points, i.e., 91 real and 91 imaginary. It does not require the experiment to be intentionally or accidentally halted to reach an odd increment number in any indirect dimension. This is particularly so for a NUS experiment in which the data was recorded according to a randomized, pre-determined, sampling list consisting of 3987 (t_3, t_2, t_1) vectors (8 FIDs per vector for the quadrature detection of the 3 indirect dimensions). As a result, the last increment in any indirect dimension may appear in the beginning of the list and therefore be recorded at the very early stage of the experiment.

p.4 l.109 Recording time 88 h? (2 + 0.2 + 0.12 + 0.03) x4 x31896 = 83 h. Where did the extra 5h for the two 4D's go into? Or did I miss a delay? There appears a hidden delay. Is this internal handling of NUS or I/O by Bruker? If there is a hidden delay, is it constant or could it be

interfering with steady-state? Or was there a typing error?

Response: The reviewer is correct that inadvertently we had simply listed the Bruker d1 delay as the recycling delay, but an additional 70ms delay is set aside for Topspin to write the data to the hard drive and also for the so-called MC work used to handle the loops for each indirect dimension. This has now been corrected. In addition, the time for the sum of incremented t_1, t_2, and t_3, 128 dummy scans, some water flipback pulses, and some gradient pulses and their recovery delays were not included in the reviewer's calculation. Each of these delays is now accessible in the newly uploaded parameter files (https://zenodo.org/record/4625615), so we prefer to simply summarize the total data collection time which is most relevant to the user.

p.6 l.171 idem. Recording time 110 h, (1.7 + 0.3 + 0.05) x4 x43856 = 100 h. Where is the extra 10 h?

Response: As per above response, the 70 ms delay was inadvertently not included in the recycling delay (now corrected) and the reviewer neglects the durations of the evolution/detection periods.

The spectra Fig.2b and Fig.3 (diagonal signals) show double Lorentzian lineshapes. This may be cause of some overlap in Fig.2b and obscure signals close to the diagonal in Fig.3. The related 4D HMQC-NOESY-TROSY in Barnes et al. (2019) in fact showed more Gaussian lineshapes. Or, would the current data processing be optimal for S/N for this sample?

Response: The reviewer is correct. Whenever signal decay is exponential, either due to the transverse relaxation or apodization, this results in Lorentzian line shape. When decay is dominated by apodization with Gaussian-like windows (such as cosine-bell), the line shape takes on a more Gaussian character, but the extent of Gaussian character depends on whether transverse relaxation or apodization dominates the decay of the processed time domain data. With the acquisition times comparable to the T_2 values in the indirect dimensions, and much longer than T_2 in the direct dimension, we applied a composite window function (a cosine function with 15 Hz additional exponential line broadening), as now clarified in lines 146-149.

p.4 l.105 Since Bruker and TopSpin 3.1 is not the only and/or forever configuration for NMR acquisition, the NUS release version and the method to generate the random numbers, distribution function, parameters and seeds could be provided (or the generated array of delays), and certainly when deposited.

Response: We agree with the reviewer and have uploaded the sampling schedule with the uploaded data sets at zenodo.org. In section 2.2, we now also include additional details on how this schedule was generated under Topspin (line 126-131).

p.5. l.117 Caption Fig. 1. I suggest to write 'selective EBURP2', to be consistent with description on p.10 l.277. Only on p.10 it became clear that the f1,f2 time-evolution periods in the 4D cover only the spectral region downfield from the water line. Maybe this can already be mentioned on p.6 in the Methods, also to fully appreciate the f1,f2 sampling in that section. To be complete, it could be indicated in the description of the experiment of Fig.1b, that it is a selective 4D NOESY-NOESY-TROSY, and that it thus focusses on the amide proton NOE networks which fully makes

sense for a perdeuterated protein.

Response: Thanks! In Fig. 1 caption, we now write 'selective EBURP2', and also point out that the selective pulses were used to excite the amide protons downfield from the water signal without perturbing the residual aliphatic protons that may result from imperfect deuteration (line 322-325). We have also inserted a sentence at the top of p.5, pointing out that for perdeuterated proteins we only sample the region downfield from the water resonance (line 117-120).

p.6 l.161 Add the ref. (Ying et al) for SMILE already here, since it appears now only later on p.10

Response: Thanks. Fixed.

p.8 l.208 Labelling error: …are highlighted in the regular 2D 1H-15N TROSY-HSQC spectrum of Figure 2a. This is 2b in the Figure and caption. Idem,… a-helix (Fig. 2b,c), b-sheet (Fig. 2d,e), and a loop region (Fig. 2f,g). Description in text and Figure do not correspond.

Response: Thanks for catching these errors. They are now fixed.

Fig. 2b. (and discussion p.8. l.224 etc) Diagonal intensities can vary, due to differences in longitudinal relaxation. However, the 'diagonal' cross-peak of L232 in Fig. 2b appears almost the same intensity as the sequential cross-peak to V233. Idem M235 and A234. Is this true, illusion, or due to contour levels chosen? If true, that diagonal and cross-peak have equal intensity, the NOE mixing time would be far beyond an optimum NOE cross-peak intensity, or not? (at least in 2-spin approximation, cf. Fig.3, Jeener et al 1979, J Chem Phys 71, 4546). Would 50 ms as in the f1,f2 planes of Fig.3 for this large 712 res. protein have been more appropriate, in hindsight?

Response: The reviewer's observation is correct. A long mixing time is purposely used to link to a maximum number of neighbors. The 50-ms mixing time of the 4D NOESY-NOESY-TROSY experiment is then used to get a semi-quantitative measure for the actual distance. We have included a sentence to this effect in the Concluding remarks section (line 415-421).

p.12 l.338 Such weak direct NOEs are well possible in perdeuterated proteins (Koharudin et al. 2003 JMR 163, 228), but of course here they could also correspond to a relay via a -OH.

Response: We fully agree, these indirect NOEs are very helpful during the assignment process. We have added an explicit sentence and a reference to the Koharudin paper in the Concluding remarks section (line 415-418).

Depending on the deuteration level the 4D NOESY-NOESY-TROSY may also show NOE cross-peaks to aromatic protons. How complete was the deuteration of the SARS-CoV-2 main protease? From the description of the protein production on p.4, the description of the isotope labelling for this protein is missing. When available, the isotope levels could then be indicated in Section 2.1 as well.

Response: Efforts were made to achieve >98% perdeuteration of the protein by preparing all reagents for the large- (and final starter-) culture in 99.8% $^{2}H_2O$ (filter sterilized) and using >97%

deuterated carbon sources (Isotec, Sigma). We now have added a sentence in Section 2.1 (line 105-108), and a more detailed description in the Supporting Information. No NOE crosspeaks were observed to any non-exchangeable sidechain protons in these experiments.

Reviewer #3

The reviewed manuscript submitted by Ad Bax and collaborators describes a four dimensional NOE experiments strategy applied for amide signal assignment of large dimeric SARS-1 CoV-2 protease.

The paper is well and clearly written and the detailed description of experiments is provided in Supporting Information. The presented experimental results are of high quality and are reported convincingly. The presented approach is interesting and scientifically sound, and there is a novelty in this work. Additionally, the very important nowadays subject of study and presented methods make this article certainly interesting for Journal readership. Therefore, undoubtedly I am recommending the paper for publication.

Minor points:

- would be very informative to present a full 2D Trosy-hsqc spectrum of studied protein. The spectrum presented on Fig. 2b seems to be a fragment.

Response: The reviewer is correct that the ^{1}H-^{15}N TROSY spectrum presented in Fig 2b is not the full spectrum. Consistent ^{1}H-^{15}N spectral limits were chosen for the illustrations in order to improve the clarity of the 4D cross section panels in Fig. 2 and 3. The full spectrum is now included in the Supporting Information, as indicated in the updated figure 2 caption.

- how many amide peaks could be observed on 2D correlation spectrum?

Response: 261 non-proline backbone amide peaks were observed in the ^{1}H-^{15}N TROSY spectrum, with an additional *ca*. 40 peaks from sidechain amides and the minor protein conformer. We now include the following sentence in the revised manuscript (lines 235-239) and refer to the full TROSY spectrum in Fig S1.

"A 2D TROSY spectrum (Fig. S1) of this sample allows identification of 261 backbone amide peaks out of 293 non-proline residues, suggesting the feasibility of implementing the TROSY version of the 4D NOESY experiment."

- Authors informed that assignment will be published elsewhere, however would be good to provide information regarding how many not-proline amide diagonal peaks were detected in 4D experiments out of almost 300 possible? How, mentioned at p.3 isomerization of P184 influenced the assignment process? Is it possible to assign both isomers?

Response: While 261 non-proline backbone amide peaks were observed, we now have added on

line 250-252: “A total of 231 peaks, out of the 261 peaks in the 2D TROSY spectrum, are detected as (semi-) resolvable diagonal peaks in the projected ^{15}N-^{1}H (F_3,F_4) plane (data not shown). These numbers do not include the doubling of resonances associated with isomerization of P184.”

And on p.13 (line 348-351: “Compared to the 4D TROSY-NOESY-TROSY pulse scheme, the 4D NOESY-NOESY-TROSY experiment avoids the lossy magnetization transfer step from ^{1}H to ^{15}N and back (leading to a slightly larger number of 241 diagonal peaks on the ^{15}N-^{1}H (F_3,F_4) projected plane, compared to 231 for TROSY-NOESY-TROSY).”

Reviewer #4

General comments

This article by Bax and colleagues describes new 4D-TROSY-NOESY-TROSY and 4D-NOESY-NOESY-TROSY experiments applied to the perdeuterated 68-kDa dimer of the SARS-CoV-2 M^{pro} protease expressed in a D_2O E. coli culture. The 4D-TROSY-NOESY-TROSY is an improved version of earlier 4D ^{15}N/^{15}N-edited NOESYs, whereas the 4D-NOESY-NOESY-TROSY is a new combination of the 3D NOESY/NOESY pioneered by the Kaptein group with a TROSY detection. The 4Ds were acquired with extended acquisition times in all indirect dimensions using NU sampling within about 4 days experimental time. The spectral resolution and sensitivity are excellent and a number of very clear and extensive $^{1}H^{N}$ NOE networks in alpha helices and beta sheets are observed. The 4D-NOESY-NOESY-TROSY in particular is well suited for larger proteins since magnetization losses are minimal during the two NOE transfer steps. As the authors indicate, the NOESYs will be very helpful for obtaining assignments of such larger proteins by combining their information with that of the less-sensitive conventional triple-resonance assignment experiments and/or that of existing structures. The manuscript is very well written. The experiments are well documented by the printout of the pulse sequences and processing scripts in the supplementary information, which should make it easy for others to use the sequences.

As such this article should be very interesting and helpful for the readership of ‘Magnetic Resonance’.

Specific Comments

1. *Figure 2b shows a large region of a ^{1}H-^{15}N correlation spectrum. The figure legend is missing a description of subpanel b. Is this a 2D TROSY? A proper description in the legend as well as experimental details should be provided.*

   Response: Indeed, this spectrum showed only a region of the 2D TROSY spectrum. A brief description has been added in the Fig. 2 legend and more experimental detail is provided with the full spectrum in the Supporting Information (Fig. S1). See also our response to Reviewer 3.

2. *The number of resonances of the $^{1}H^{15}N$ 2D in Figure 2b is rather small for a 306-residue protein.*

I counted about one hundred resonances. Is this due to missing ^{1}H back exchange? The authors should discuss this point and indicate which percentage of the amides did exchange.

Response: As described above in our response to Referee 3, Figure 2b is a subsection of the $^{1}H^{15}N$-TROSY spectrum; a larger region is now included in the Supporting Information. As also added in the main text, we were able to identify 261 peaks out of 293 non-proline residues in the full 2D TROSY spectrum. It was established in subsequent measurements that about 20 resonances were missing due to incomplete back exchange. Others are presumed to be missing due to exchange broadening, as most of the remaining missing peaks correspond to residues in the immediate vicinity of resonances that exhibit field-dependent line broadening. We now note on p.9 (line 238-240): "Conformational exchange on a time scale that results in extensive line broadening and incomplete back exchange of amides when the protein was purified in $^{1}H_2O$ (2 days at 25 ˚C) are the primary causes for the absence of the *ca.* 30 amide signals." As the current manuscript is aimed at highlighting the merits of the 4D NOE experiments, not the structure or dynamics of the protease, we believe a more detailed analysis falls outside of the scope of the present study.

3. *In this respect what would be the prospects for using similar sequences for protonated large proteins?*

Response: The principal limitation in these NOESY-type experiments when applied to protonated proteins is spin diffusion to sidechain protons. As we point out in the Introduction "This decrease in sensitivity was remedied by generating the protein in a highly perdeuterated state, while keeping the solvent-exchangeable backbone amide protons protonated (Torchia et al., 1988;Lemaster and Richards, 1988)" In addition to spin diffusion, the TROSY effect is also reduced when the experiment is carried out for a large protonated protein such as M^{pro}. We added two sentences (lines 443-448) at the end of the Concluding remarks that for these reasons the experiment is best suited for perdeuterated proteins, but that analogous experiments on methyl-protonated but otherwise perdeuterated proteins should be readily feasible.

4. *The authors use a rather high concentration of 0.9 mM dimer/1.8 mM monomer. Do they observe any aggregation? What are the prospects for applying the sequences to lower concentrations?*

Response: No aggregation was observed over the 2 months that this sample was being used, but we note that this work is carried out on an active site mutant of the protease which enhances its long-term stability. As with all NMR experiments, S/N is approximately proportional to sample concentration and working at the high concentrations benefits S/N of these experiments that involve multiple magnetization transfer steps, in particular NOE experiments where magnetization from a single nucleus is distributed over many neighbors. We have added a brief paragraph (line 409-414), pointing to the caveats of required long term stability and high sample concentrations.

5. *Could the authors please also indicate $^{1}H^{N}$ and possibly ^{15}N T_2 relaxation times?*

Response: We indicated that the rotational correlation time of the protein is ~27 ns (line 229), a result based on analysis of the ^{15}N relaxation data. Details of this analysis, including observation of a small subset of more flexible residues and the presence of exchange-broadened resonances will be reported elsewhere and are beyond the scope of this study. We also note that for the data

reported, it is the ^{1}H and ^{15}N TROSY-R_2 values that are most relevant. These are in line with theoretical predictions based on a τ_c=27 ns value.

6. *What is the reason for having 0.3 mM DSS in the buffer?*

   Response: In Section 2.2, we have inserted that DSS serves as an internal chemical shift reference (line 113). Please note that DSS is the IUPAC ^{1}H chemical shift reference.

7. *It would be helpful to not only provide the pulse programs and processing scripts as PDFs but also as plain text files.*

   Response: We agree and have included the pulse programs and scripts as text files with the uploaded raw data sets onto the zenodo.org web server.