Line	Statement
37	I think that with your project you need to think not about whether it is possible, because youve proved that it is, but can you compete against what is used nowadays.
38	You need to make it so that the biologist or bioinformatician is keener on using this approach than on the approach that is being used already.
51	Is [the speed of the sonification] something that youre looking to have control of?
54	You might not be interested in the whole protein anyway. Most people are not.
114	I thought that reducing the alphabet was a good idea, as it would be less confusing than hearing a lot of notes.
120	I appreciate that I couldnt get all seven of them, but I think there is still something good abut grouping the notes into groups. Unless you are specifically looking for that, you arent bothered about the different hydrophobicity or hydrophilicity of all the 20 amino acids.
125	I think that the grouping would still be a good idea, depending on the analysis that you were trying to get from it.
172	...but I dont see how that would be better than just looking at the alignment as that is very easy to see white and not white. It is much faster. I dont see what would be the added value of listening to the alignment.
178	That is if you can see. If you were a visually impaired scientist, I think this would be a really good thing to have.
184	Its a trade-off between including everything and having more information, or making it more understandable but losing information.
218	[in response to learning that they'd only heard half the alignment] God!
233	I think its useful but only to give a rough approximation to what the alignment is like compared to other alignments.
266	You might not be able to reproduce things between different people as easily if you get that difference in the way people hear different pitches. [Other participant:]Maybe you could customise it. [First Participant:] Yeah, I was going to say!
274	Theres nothing about the sound which is biological, it would just be a preference.
278	People are interested in different aspects of the same data.
279	I could see people choosing to using this technique to if they have a really long multiple sequence alignment or a lot of multiple sequence alignments and they just want to know if theres a conserved domain: just as a way of filtering through the data that theyve already got and they dont want to look up every single one. Theyre just going to use this as an initial filtering approach to look for conserved domains, so the flexibility to customise it would be good.
288	So, if you start really broad it helps you get to conserved regions, then as you get more people using it who have different ideas or different needs then you can start to get more specific, if possible.
330	But I think implementing different instruments might be a good shout, but you might not want to overcomplicate it.
334	It would take a bit of getting used to: knowing what instruments corresponds to which level of hydrophobicity. That would be quite difficult to get a handle on. If its just pitch then its easier to get a handle on at some level, but with different instruments corresponding you would have to practice and train yourself. Im not sure how easy that would be, and whether people would bother.
352	I think it comes down to trying to attune yourself to an instrument representing a level of hydrophobicity.

Table 1: *Project Judgements*: representative statements from focus group transcript concerning feedback on the project as a whole, and on the new mode of data representation for protein sequence data. Horizontal lines indicate when each sonification was played during focus group, details of these are in Table B.2.

Focus Group Representative Statements

Line	Statement
39	I think it is kind of fast, if you want to make things out of it. Id rather hear it with more spacing between each sound, if I was to analyse it. But then again, there is a difference between each thing. So, I know that this is one thing, and this is the next thing. So maybe more space...
43	Not sure I agree, it seemed to go on for a long time
49	Im not sure youd always want to be going slower, I think you wouldnt have time if you had several proteins to look at.
64	You could really hear it this time. The hydrophobicity at least, the part that was low pitched
66	Yeah, you could hear low-low-low then high. [second participant:]...Then little high one.
68	I would not say a conserved domain, but a repeated motif.
72	The only thing I noticed that was different was that the high-pitched ones were a bit more frequent than the first one. Probably I wasnt focusing enough in detail and that could be a flaw of the method. It depends on your purpose for listening to it...
95	Im struggling to hear repeated patterns, to be honest. There could have been something there, but it was hard to discern. You had that high note going all through it...
105	It was not easy to hear the motif change.
112	It just made it sound like there was a hydrophillic thing coming in all the time every 2 amino acids, which stopped you picking up on the longer structures.
115	I could hear subtle changes, but I think it was that high note throughout that was throwing me off. You could tell there was a pattern, but in terms of the number of times that pattern came up, or even what that pattern was, Im not sure you could hear that. [Second Participant:] Yeah.
120	I appreciate that I couldnt get all seven of them
122	If you could make the difference in pitch more pronounced. In the previous example the lower pitch felt lower.
124	I think the same, there didnt seem to be a huge difference between the highs and the lows.
125	Here, even the lowest pitch felt quite high.
146	I think you can really hear where the conserved regions are. It does sound very different, going along high and the you get the *dum-dum-dum* for a bit which gives you an idea where to look, or if there is a conserved region. If that was just your initial question. I felt that I could hear that more clearly than the first ones...
152	I think the extremes are very easy to tell, but everything that isnt really low or high is a blur. It was easier to understand that one than the previous.
171	I think as [other participant] said, its easy to discern between the highly conserved and not highly conserved...
185	I think its very easy to understand the very high pitch and the very low pitch, but everything in between is not at all.
221	But you still can pick out conserved regions. Every now and again you can hear that one note coming out strong. Especially when it happened a lot, you got it over a few amino acids. I dont think it was as clear as the previous, as all the notes playing at the same time is quite a lot to take in. I dont know if I could tell, with everything else going on, the difference between where all the amino acids were conserved, or whether there was one and gaps.
226	I think all of this gives a flavour of what the protein is like, but Id struggle to get detailed information out of it. Maybe Im just too stupid, but I think Id struggle to compare this alignment to another alignment that youd hear just afterwards. You might get a general idea of which one was more conserved...
236	This was easier to discern than the previous one because it was challenging and the other was boring. As in why am I putting attention?. As well I was trying to find a pattern.
247	And while I do agree that it is more challenging, there are positives with it being boring, as Pablo said, with regards to analysis.
251	Whereas when I hear the same sound, I can more easily find a pattern across what Im listening to, and usually patterns are what youre looking for in this kind of stuff.
258	I think that one is better as long as the sound within the spectrum are more different.
260	I do think there are positives to listening to less sounds than to more difference.
284	I was going to say about the difficulty to getting detailed information out of it, so you might want to use it as an approximation just to get an initial idea about something.
296	As an only source of information this is difficult
302	As I cant tell what Im listening to.
325	I did like the fact that there were different instruments, as I think you can tell the difference more clearly between different instruments than just different notes. The whole difference in sound? I dont know.
328	I dont know if I could take on board both things.
329	I can definitely tell the different instruments, but the different sound within the same instruments? No, not really.
333	Yeah, its a lot to take in, I think.
338	I think it was the most informative, because of the different instruments. I think it is more easily distinguishable...
348	Trying to do the reduced alphabet is to simplify things, but then using different instruments with different pitches within the reduced alphabet undoes that work of simplification. Then you get more complex again. Maybe having the full alphabet with full pitches is easier than reduced with different pitches.

Table 2: *Analytic Judgements*: representative statements from focus group transcript concerning feedback on the information carrying aspects of the sonification. Horizontal lines indicate when sonifications are played, details of these are in Table B.2.

Line	Statement
24	It is much more diverse that I expected. I was expecting something very boring, but it has a lot of diversity.
30	Its slightly chaotic, but not completely chaotic. Its like a scary movie tune, like a soundtrack
31	Its very Carpenter!
44	It goes on for a long time, but I think that that is nice. When you work with any genetic data, you can forget how big it is. When you look at the 4 lines on the page, it looks like a small piece of protein. But listening to every single one of these amino acids as a note, even going quite fast, you see that it is a lot of information. I think it is good for giving you an appreciation of how big that dataset is.
97	You had that high note going all through it, a beep, higher than the others. And there were bits where it sounded like a guitar coming in like acoustic guitar. That was a bit gentler
102	[in response to being asked if it was more musical than the previous example] Yeah, probably was to me. I found it easier to relate to: a repetitive musical structure. It was a bit easier to listen to easy listening proteins!
115	It sounded very nice
149	I felt that I could hear that more clearly than the first ones, but maybe thats because it sounded less like music and more discrete, it was just high-high-high then low-low-low.
220	I thought it was going to be chaos, especially when it started-it was just like someone bashing at notes.
236	and the [entropy sonification] was boring.
339	I think it is more easily distinguishable, but its also the least musical. Its the weirdest sound.
340	Its the most variable sound, so it doesnt conform to the normal structures of music.

Table 3: *Aesthetic Judgements*: representative statements from focus group tran-script concerning feedback concerning the beauty and feeling of the sonifications. Hor-izontal lines indicate when sonifications are played, details of these are in Table B.2.

Line	Statement
150	Maybe that was easier for my brain to associate with things going on. Though that could be very subjective.
155	The thing I pay attention to is the high notes, and I zone out for the low notes. So, if youre really keen on
	hearing the conserved bits, then you could flip it round and make those the high notes, so they stand out. If
	what youre interested in is breaks in your conservation, then it makes sense to make those as the high notes.
165	I dont know, I guess youd get used to it I suppose. I still think theres a tendency for my hyperac-
	tive mind to zone out unless theres something to hang onto, so I might still struggle to pay attention to
	the low bits. That could be good if you dont want to pay attention to the low bits, but it might be bet-
	ter to be the other way around. Even if I listened a lot, I might well still zone out on the low notes.
223	. I dont think it was as clear as the previous, as all the notes playing at the same time is quite a lot to take in.
240	that in relation to the strong sound and lower sounds; for me, not sure about other people, but for me a high lower sound is
	louder than a strong high sound.
244	A loud low pitch sound is louder than a loud high pitch sound. Hen it is a louder sound, I can tell more easily the lower
	pitched one than the higher pitched one.
247	For me it wasnt easy to tell a louder high pitch, compared to a louder low pitch.
249	My visual memory is much better than my listening memory, I can tell what Im listening to at the moment and I can
	recognise what is present in all of them or is hydrophobic, but I wont remember this two seconds later. Whereas when
	I hear the same sound, I can more easily find a pattern across what Im listening to, and usually patterns are what youre
	looking for in this kind of stuff. I do think there might be a point for it being more of the same sound than more dif-
	ferent things, as I do think there might be a threshold for how much you can tell by listening to this sort of thing.
262	I find the opposite pattern with hearing loudness. I think higher pitches are easier for me to find as louder.
264	Which could be a flaw, as some people are more biased towards the hydrophobic notes, and some people
	towards the hydrophilic ones, or some people towards the more conserved areas, and the less conserved.
274	At the end of the day it doesnt really matter what the sound is. Theres nothing about the sound which is biological
291	Maybe people who are blind would hear it more accurately. Maybe their senses are more attuned.
335	It would take a bit of getting used to: knowing what instruments corresponds to which level of hy-
	drophobicity. That would be quite difficult to get a handle on. If its just pitch then its easier to get
	a handle on at some level, but with different instruments corresponding you would have to prac-
	tice and train yourself. Im not sure how easy that would be, and whether people would bother.
354	It might be possible but might take a lot of training. It might be something a computer would be better at than a human like
	machine learning which might defeat the purpose slightly!

Table 4: *Psychoacoustic Judgements*: representative statements from focus group transcript concerning feedback concerning the subjective psychoacoustic responses of participants as to how they experienced the sound. Horizontal lines indicate when sonifications are played, details of these are in Table B.2.

Line	Statement
40	Id rather hear it with more spacing between each sound, if I was to analyse it.
123	If you could make the difference in pitch more pronounced. In the previous example the lower pitch felt lower.
156	So, if youre really keen on hearing the conserved bits, then you could flip it round and make those the high notes, so they stand out. If what youre interested in is breaks in your conservation, then it makes sense to make those as the high notes.
186	If you took a similar approach to the last method, and instead of calculating mathematically the pitches, you grouped them with discrete sounds, you could then say whether it is all conserved or 25% conserved etc. You would lose information as you are grouping things but... [Moderator:] But you think it might aid in understanding? [Participant:] Yeah.
192	Yeah, like a percentage cut-off for conservation, like 0% to 100%
194	And then it would be easier to find things that are very conserved, things that are maybe more or less But then it depends on what you are interested in, as you are losing information. But just listening to very different sounds might make it more difficult to make associations between things within the main alignment.
230	I think I would need to look at something and complement it with something else to get the detail to an adequate level and have something that is measurable rather than just a getting a feeling that something this is different.
253	I do think there might be a point for it being more of the same sound than more different things, as I do think there might be a threshold for how much you can tell by listening to this sort of thing.
258	I think that one is better as long as the sound within the spectrum are more different.
260	I do think there are positives to listening to less sounds than to more difference. But thats me.
269	If I want to listen to the hydrophobic as the high pitched one then I could pick that and if you wanted the opposite you could pick that instead.
276	And you could customise it so its just four sounds, or to listen to everything thats conserved, or nothing thats conserved, or quartiles in between. As its very dependent on what your listening for and for the data. People are interested in different aspects of the same data.
293	Id love to be hearing this at the same time as knowing exactly where the note was coming from
297	if you were to implement this along with a visual analysis... As [participant] said, I cant see where I am: I try and then nope. How would you make it so that Im looking at something and I know? Would you have an arrow pointing at it?
301	You could have one of those balls like at karaoke.
309	[Moderator:]You could have a bar that goes across, like in so many music applications. That could be on top one of these [MView] and go across. Also, sonically we could put in clicks. On this viewer, do you see the little 1 and 2 on top? These correspond to the 100th and 200th residue. The idea is that a little click would allow people to realign, though it might be superfluous once youve got the viewer. Also, it might be good for sectioning off so you can identify that it occurs between 1 click and 2 clicks. Then you can just go back to click-click. This might help perceive size. [participant:] Thats something to think about if you have a really long alignment.
350	Trying to do the reduced alphabet is to simplify things, but then using different instruments with different pitches within the reduced alphabet undoes that work of simplification. Then you get more complex again. Maybe having the full alphabet with full pitches is easier than reduced with different pitches.

Table 5: *Further Work*: representative statements from focus group transcript concerning suggestions for improvements or further work. Horizontal lines indicate when sonifications are played, details of these are in Table B.2.