Peer Review File

Article Information: https://dx.doi.org/10.21037/atm-22-2293

First Round of Peer Review

Reviewer A

Comment 1:
This review article covers the optimisation of EAU induction methods in C57BL/6 mice. It highlights the importance of IRBP-1-20 and PTX doses in the development of chronic EAU in mice along with the benefit of emulsion sonication to elicit a stable and consistent antigen response, which is an important finding.
"The article is generally well written but requires some English language proofing (mainly syntax and grammar) to improve readability, for example 'human uveitis which is usually prolonged and slowly progressed' (line 76) and 'mice showed mild to moderate destruction on visual function' (line 79).
Reply 1:
We thank the reviewer's advice. We have checked text and made the revision carefully.
Changes in the text:
We have modified our text as advised (see Page 5, line 73; and Page 6, line 75)

Comment 2:
"Moreover, there are inconsistencies in the presentation of the results for each optimization step, for example, Fig. 2 and 3 do not include clinical scoring or flat mounts, which makes it difficult to draw a real conclusion. "
Reply 2:
Agreed. The Fig. 2A and 3A with clinical scoring and Fig. 2D and 3D with flat mounts were added. The relevant text was also added accordingly in the Results Section.
Changes in the text:
We added Fig. 2A, 2D and Fig.3A, 3D in the Figure Section (see Page 26, line 471; Page 27, line 486).
The relevant text was added in the Results Section (see Page 14, line 209, 217; Page 15, line 227, 232).

Comment 3:
"Animal numbers for each optimization step and assessment technique are also not very clear and an overview table should be provided in the methods section."
Reply 3:
Agreed. Table 1 containing the overview information of the mice used in our work was added in the Method Section, including the animal numbers and the results obtained via different assessment approaches. The relevant text was added as well.

Changes in the text:
We added Table 1 in the Method Section (see Page 7, line 114).
The relevant text was added (see Page 7, line 112).

Comment 4:
"Finally, a head-to-head comparison (using all assessment methods described) between the previously used (200ug IRBP and 500ug PTX, Fig. 1) and supposedly optimized (500ug IRBP and 1000ug PTX) dose for EAU induction would have been useful."
Reply 4:
We agree to the reviewer's valuable advice. A head-to-head comparison can present our results more clearly. In the Results Section, Fig 5 A-D was added with data of clinical scores, incidence of EAU, pathological scores and ratios of $CD4^{+}IL\text{-}17A^{+}$ cell in dLN. The relevant text was also added accordingly in the Results Section.
Changes in the text:
We added Fig 5 A-D in the Figure Section(see Page 30, line 518).
The relevant text was added in the Results Section (see Page 16, line 257).

Comment 5:
Below are several suggestions to further improve the article:
1) Section 3.1 and Figure 1:
"The clinical and histological scoring for C57BL/6j mice described Xu et al. ranges from 1 to 4. Yet here the clinical scores do not exceed 0.6. Did authors use a different scoring scale? If so, further details on how clinical and histological scoring was performed is required."
Reply 5:
In this work, we used the clinical scoring scale by Xu et al. (2008) and a histological scoring scale by Shao et al. (2006). According to the EAU scoring results of the 12 mice at day 18 (Fig. 1), as the reviewer pointed out, seven mice did not exhibit EAU signs and the scores of remaining 5 mice with EAU were all integer. Therefore, the overall scores of the 12 mice included with EAU and without EAU (score 0), so that the final scoring result was significantly lowered to 0.6. The relevant text was also added accordingly in the Figure Legend.

Changes in the text:
The relevant text was added in the Figure Legend (see Page 26, line 461; Page 27, line 477; Page 28, line 494; Page 29, line 509; Page 30, line 524).

Comment 6:
1) Section 3.1 and Figure 1:
"Xu et al. found that EAU peaked around day 25, while authors see a peak at day 18. What could be the reason for this?"
Reply 6:
We observed the peak EAU inflammation level at day 18, based on the data of day 14, 18, 22. The discrepancy might be ascribed to the difference in experimental conditions, such as the

immunization protocol (500 ug IRBP and 150 ug PTX in Xu et al.'s study vs 200 ug IRBP and 500 ug PTX in our study), mice raising condition and reagents used, etc.. Moreover, some previous literatures found the different EAU peak times as well, including day 16-21 (N. Chen et al. 2021; Li et al. 2020) which is consistent with our results.

Changes in the text:

We added the relevant text and the literature of Chen et al. 2021 and Li et al. 2020 in the Results Section (see Page 13, line 190; Page 23, line 430).

Comment 7:

1) Section 3.1 and Figure 1:

" Also, authors claim that 200 ug IRBP and 500 ng PTX is the most frequently used dose, yet it is not used in that combination in any of the studies listed in Table 1."

Reply 7:

We thank the reviewer's valuable advice. The different protocols of EAU inducement used in the multiple studies were listed in Table 2, including the discrepancy in combination of IRBP and PTX. Based on these data, we adopted the moderate dose of IRBP and PTX in order to obtain the high EAU incidence. In addition, the protocol with equal IRBP and PTX doses (Ke et al. 2007) was added in Table 2.

Changes in the text:

We have modified our text as advised (see Page 12, line 182).

We added the literature of Ke et al. 2007 and Fang et al. 2010 in the Table 2 (see Page 12, line 186; Page 22, line 388; Page 23, line 417).

Comment 8:

1) Section 3.1 and Figure 1:

"Finally, the number of animals used per assessment technique is not entirely clear. For clinical scoring the n=5 (one or both eyes?). For pathological (later called histological) scoring the incidence seen above the bar graph is 2/10 on day 10, for example. Does this indicate that EAU was induced in 10 mice (or 10 eyes?) but only 2 out of 10 showed signs of EAU on day 10?"

Reply 8:

In the Method Section, the supplementary Table 1 gives the mice number of different mice groups used in our study. For clinical scoring, the n=5 denotes the 5 mice. The both two eyes of each mice was evaluated and the average scores of the two eyes were used. For pathological/histological scoring, 2/10 denotes that 2 out of 10 mice developed EAU. The changes were made in the revised manuscript accordingly.

Changes in the text:

We added Table 1 in the Method Section (see Page 7, line 114).

We have modified the text as advised in Methods Section (see Page 9, line 132).

Comment 9:

2) Figure 2:"Higher magnification HE images are required to highlight the pathologies used for histological scoring. Since there was no significant difference in the incidence or IL-17A %, was the optimal dose selection solely based the histological scoring? Why was

clinical scoring not included here?"

Reply 9:

We replaced the Fig 2D with higher magnification HE images, in which the black arrow denotes retinal folds and white star denotes edema. Since no significant difference in the incidence of EAU or IL-17A % in dLN was observed with treatment of different IRBP doses, we hence added the data of clinical scoring and flat mounts (Fig. 2A and 2E). Based on the overall results of EAU level assessment, the 500 ng IRBP seems to be the optimal condition for EAU induction. As for Th17% level, the relatively high ratio of Th17 in all groups at day 18 (about 4%) might be the reason for the insignificant difference. The number of mice used for Th17 testing was relatively small (n=5), which may also result in the statistical deviation.

Changes in the text:

We replaced the Fig. 2Din the Figure Section (see Page 26, line 471).

We added Fig. 2A and 2E in the Figure Section (see Page 26, line 471).

Comment 10:

3) Figure 3:"There is no significant difference in incidence %, histological scoring and IL-17A % between 1000 ng and 500 ng PTX. As such, the 'optimal' dose selection was based on a trend rather than significance. Again, why was no clinical scoring included here? This should be further clarified in the discussion."

Reply 10:

Agreed. The Fig. 3A and 3D containing the data of clinical scoring and flat mounts were added in the revised manuscript. Although no statistically significant difference in incidence %, histological scoring and IL-17A % between 1000 ng and 500 ng PTX was observed, 1000 ng PTX ought to be more effective than 500ng PTX in EAU induction according to the overall EAU assessment indices. Besides, the relatively small sample size used also can be the reason. The relevant text is revised accordingly as well.

Changes in the text:

We added Fig. 3A and 3D in the Figure Section(see Page 27, line 486).

We have modified our text as advised in Discussion Section (see Page 18, line 305).

Comment 11:

5) The authors claim that histological scoring is the 'gold standard', but little can be seen on the histological images provided. Moreover, histological scoring is exceedingly difficult as one needs to obtain serial eye sections, scan for pathologies and distinguish between real pathologies and artifacts created by the staining procedure. In comparison, clinical scoring (fundus and OCT imaging) can be used to track the progression of the disease in the entire eye (rather than a section) in real time without sacrificing the animal. Please comment.

Reply 11:

We appreciate the reviewer's valuable advice. Yes, the histological scoring has the weakness in terms of sample preparation and image evaluation, although it was taken as the most commonly used approach in some literature(Agarwal et al 2012; Klímová et al. 2016). In our study, actually both the historical staining image and clinical scoring were applied for the EAU assessment. Efforts were made to archive the reliable and efficient results of sample evaluation as far as possible in our work. Firstly, all the historical samples were processed by

same one person following the same staining protocol. At scoring stage, two independent observers completed the assessment based on the low and high magnification images as described in the Methods Section. However, we really found the discrepancy of scoring results in different histological sections. Therefore, we applied the retinal flat mounts as well to obtain more complete and overall features of the retinal inflammation. Secondly, the ocular fundus imaging was adopted to score the relevant lesions, which can provide the continuous information of EAU progress. Together, we agree to the reviewer's comments and also believe that the multiple approaches to assessing the EAU lesions should be taken into account in order to obtain more reliable and efficient scoring results, as each approach has its merits and shortcomings.

Changes in the text:

We addedthe literature of Agarwal et al. 2012 and Klímová et al. 2016inDiscussion Section (see Page 20, line 334; Page 22, line 399; Page 23, line 451).

We have modified our text as advised in Discussion Section (see Page 19, line324;Page 20, line 328, 334).

Comment 12:

6) The discussion requires more work as to justify the experimental choices made by the authors in terms of the sequential order of the experiments and the success of the optimised dose. A head-to-head comparison between the 'previously used' and 'optimized' dose would have been useful to be able to draw a real conclusion.

Reply 12:

Agreed. We added the content of a head-to-head comparison between the previously used and optimized dose in the Results Section 5 (Fig. 5A-D). The results showed significantly statistical difference in EAU incidence, histological scores and IL-17A% in dLN, suggesting that the better efficiency of the optimized protocol than the previously used protocol. The changes were made accordingly in the revised manuscript.

Changes in the text:

We added Fig. 5 in the Figure Section(see Page 30, line 518).

The relevant text was added in Results Section (see Page 16, line 257).

Reviewer B

The authors were trying to address an interesting and important issue in the EAU field. However, the authors failed to demonstrate a clear procedure with significant improvement in the efficiency of EAU induction. With the current standard protocol, which was originally developed in Dr. Rachel Caspi's group in NIH, one can already achieve the efficiency of EAU induction that was reported in the current manuscript. The current manuscript can be considered for publication, if authors could further optimize their protocol, and address the following questions:

Comment 1:

1. Present a table to clearly compare the EAU induction efficiency produced by their method and by the current standard protocol, along with proper statistical test. The comparison should include the data from day 7 to day 14 post IRBP immunization.

Reply 1:

Agreed. The original protocol applied in our study was similar to the one reported by Dr. Rachel Caspi group. However, the efficiency of EAU induction was not satisfactory and hence the protocol was optimized in this work. The comparison between the original and optimized protocols was made with recruited data of day 7, 14, 18, 22. The specific data, including clinical scores, incidence of EAU, pathological scores and ratios of $CD4^{+}IL\text{-}17A^{+}$ cell in dLN, were shown in Fig 5 A-D "Head-to-head comparison between primary (original protocol) and optimized protocols for EAU induction" in the Results Section 3.5.. For unltrasonic emulsification protocol analysis, we added the additional data of day 7 in these three weeks based on the original day 14, 18, 22 data obtained previously.

Changes in the text:

We added Fig. 5 in the Figure Section(see Page 30, line 518).

The relevant text was added in Results Section (see Page 16, line 257).

Comment 2:

2. EAU scoring is quite sensitive to the scorer's experience and the instrument, especially at the onset stage of EAU. The score should be reported from at least two independent scorers, together with details about the scoring criterial and instruments.

Reply 2:

Agreed. EAU clinical scoring was completed by two independent scorers. The scoring results that obviously differed between the two independent scorers will be arbitrated by third experienced scorer. In this work, we used the detailed clinical scoring scale by Agarwal et al. 2012 and Xu et al. 2008, and the histological scoring scale by Agarwal et al. 2012 and Shao et al. 2006. All the historical samples were processed by same one person following the same staining protocol. At scoring stage, two experienced independent observers (together with third experienced scorer if necessary) completed the assessment based on the low and high magnification images as described in the Methods Section. Moreover, both the clinical scoring using ophthalmoscope and histological scoring based on HE staining section and retinal flat mounts were applied in order to obtain more complete and reliable results in our work.

Changes in the text:

We have modified our text as advised in Methods Section (see Page 9, line 132; Page 10, line 139).

Second Round of Peer Review

Review Comments

While the authors have made great efforts to address our comments, some issues remain:

Comment 1: The manuscript contains many typographical, syntax, grammar and formatting errors, for example, the full form of CFA in the abstract is incorrect. A syntax example includes line 133 'The both two eyes of each mice were evaluated and the average score of the two eyes were used.' The manuscript still requires significant English proofing as the authors only corrected the two examples we provided but did not improve the entire manuscript. This makes it extremely difficult for the reader to follow.

Reply 1:

We are so sorry about the negligence in the text writing of the paper. We went through the entire paper and corrected the typo, formatting and grammar errors in the revised manuscript. We are also glad to receive the language editing with the help of the editorial office to further improve the manuscript.

The examples listed in the comments were also corrected. The full form of CFA in the abstract has been corrected as "complete Freund' s adjuvant" (See Page 3, Line 40). The asyntactic sentence in line 133 was modified as "The two eyes of each mouse were evaluated and the average score of the two eyes was used" (See Page 10, Line 144).

Comment 3: The term 'sick animals' is confusing. Do authors mean animals with induced EAU? Sick animals imply other issues and thus exclusion from the study.

Reply 3:

Yes, the term "sick animal" means the mice with induced EAU in our manuscript. According to the name used in the previous literature (*Nature Medicine,* 2014; 20(6):633), the term "sick animal" was substituted with "EAU mice". (See Table 1, Page 8, Line 119).

Comment 7: Ke et al. used 200 ug IRBP and 500 ng PTX for adoptive transfer of EAU induction and not for active immunisation. This should be mentioned.

Reply 7:

Agreed. The relevant content of "An adoptive transfer model of EAU" was added in Table 2.

(See Page 13, Line 201).

Comment 11: The authors acknowledge a discrepancy in histological scoring and therefore used retinal flat mounts to assess overall inflammation. Has this technique been used in literature before? If so, authors should give reference to the relevant study including some details on the grading scale used. Also, Kilmova et al. did not perform histological scoring at different time points. They only did an assessment at day 35. This should be corrected.

Reply 11:

Yes, the retinal flat mounts have been used in the literature before as listed in the following text (Ref. 1-6) (See in the Discussion Section, Page 20, Line 335). We added the three major references regarding retinal flat mounts in the revised manuscript. All these studies did not define the grading scale of retina inflammation based on the retinal flat mounts technique.

We agree with the reviewer's comment that Kilmova et al. (Ref. 34) only reported the assessment of histological score on day 35. We replaced it with other three references which presented the successive data of histological scores throughout the EAU induction process (Ref.39-41 in the revised manuscript) (See Page 20, Line 330).

The references about the retinal flat mounts:

1. Dick AD, Forrester JV, Liversidge J, Cope AP. The role of tumour necrosis factor (TNF-alpha) in experimental autoimmune uveoretinitis (EAU). Prog Retin Eye Res. 2004 Nov;23(6):617-37.

2. Chu CJ, Herrmann P, Carvalho LS, Liyanage SE, Bainbridge JW, Ali RR, Dick AD, Luhmann UF. Assessment and in vivo scoring of murine experimental autoimmune uveoretinitis using optical coherence tomography. PLoS One. 2013 May 14;8(5):e63002.

3. Umazume A, Kezuka T, Matsuda R, Usui Y, Takahashi H, Yamakawa N, Yashiro T, Nishiyama C, Goto H. Role of PU.1 Expression as an Inflammatory Marker in Experimental Autoimmune Uveoretinitis. Ocul Immunol Inflamm. 2018;26(6):951-963.

4. Tan J, Liu H, Huang M, Li N, Tang S, Meng J, Tang S, Zhou H, Kijlstra A, Yang P, Hou S. Small molecules targeting RORγt inhibit autoimmune disease by suppressing Th17 cell differentiation. Cell Death Dis. 2020 Aug 22;11(8):697.

5. Okunuki Y, Mukai R, Nakao T, Tabor SJ, Butovsky O, Dana R, Ksander BR, Connor KM. Retinal microglia initiate neuroinflammation in ocular autoimmunity. Proc Natl Acad Sci U S A. 2019 May 14;116(20):9989-9998.

6. Santeford A, Wiley LA, Park S, Bamba S, Nakamura R, Gdoura A, Ferguson TA, Rao PK, Guan JL, Saitoh T, Akira S, Xavier R, Virgin HW 4th, Apte RS. Impaired autophagy in macrophages promotes inflammatory eye disease. Autophagy. 2016 Oct 2;12(10):1876-1885.

Comment 12: The discussion still requires more work to justify the sequential order of experiments.

Reply 12:

Agreed. We have made the extensive revision to make the content be more logical and succinct throughout the discussion section.

Comment 13. The efficacy of EAU induction was not that much improved to previously reported methods.

Reply 13:

Agreed. As shown in the Discussion Section, EAU induction rates in C57BL/6 mice varied in the range of 30%-70% in the previous studies (Ref. 9,10,23,24,28,29) (See Page 17, Line 289). According to our data, the incidence and severity of EAU were low (incidence: 37.5%; pathological score: 0.3) in mice group immunized with the primary protocol. Comparatively, with the optimized conditions, the induction rate was increased to 82.4% and the pathological score was around 2.0. The variation of EAU induction may be ascribed to the discrepancy of the reagent used, animal and experiment conditions, etc. among different research teams. Nevertheless, we tuned down the tone for the relevant contents in the discussion section.

(See in the Discussion Section, Page 20 , line 341).