The manuscript presents a microstructural investigation of scaly clay from Opalinus Clay of the main fault in the Mont Terri underground laboratory.

Overall, the investigations and results are described clearly and in sufficient detail. Particularly the description and interpretation based on more qualitative methods (i.e. interpretation of images obtained by various methods) appears profound and convincing. The derived model of the evolution of scaly clay is comprehensible and quite differentiated.

Regarding the use of quantifying methods (cumulative grain size distribution obtained by sieving; frequency distribution of microlithon cross-sectional area obtained by DIA), the manuscript does not yet get as much as possible out of the data. Some shortcomings in this part of the data analysis require minor revision (but can be corrected easily).

There is also a number of technical issues, particularly concerning references to figures as well as some details within the figures.

Following is a detailed list of comments indicating required or desirable changes in the manuscript. Besides, there are some suggestions that are beyond any necessity but may further improve this fine piece of scientific work:

p.2, line 30ff: The sequence/numbering of sample types given here is switched compared to the numbering used in Figure 3. Harmonizing the numbering would improve clarity.

p.3, line 9 + 11 + 16:
The term “representative” might deserve some explanation. “Representative” with respect to which aspect? How did the authors ensure the representativeness?

p.3, line 9f: *“representative microlithons were selected for morphological description and microstructural analysis”* – does this selection take place before or after the sieving process? If it happens before sieving, it is recommended to check whether this could have caused any relevant impact on the mass balances of the sieve fractions. Usually, a selection of “representative microlithons” is not representative in all aspects, but results in some overrepresentation of larger microlithons within the selection and a corresponding underrepresentation within the remaining sample material.

p.3, line 9ff: The moisture state significantly affects the properties and behavior of a clay stone, for instance the results of the disintegration and sieving procedure. In case that the moisture state of the samples has not been equilibrated to the relative humidity of the air in the lab (i.e. dried to room conditions), evaporation during the sieving process could disturb the mass balance along the several states of the sieving process. It can also disturb the mass balance between the different sieving fractions, because during sieving smaller microlithons will lose their moisture much faster than the larger ones.
Therefore, some information should be given concerning the pretreatment of the material with respect to its moisture content (and some considerations concerning possible implications for the data interpretation, in case that the sieving has started far from moisture equilibrium).

p.3, line 27: *“unloading fractures”* – regarding my own experience with Opalinus Clay from Mont Terri it is likely, that many or even the majority of these fractures are not unloading fractures (in termes of “a result of removing the in situ confining stress”) but result from shrinkage due to sample desiccation.
In the framework of poroelastic theory (e.g. according to Biot or Bishop) shrinkage cracks

could also be termed "unloading fractures", where the unloading in effective stress results from increasing suction (whereas total stress remains unchanged). But in the context of this paper, where hydraulic-mechanical coupling is addressed only in a short section (chap. 4.2.2), subsuming shrinkage cracks under the term "unloading fractures" might be misleading.

p.4, line 8: Giving information concerning the weight of these samples is recommended (cf. comment to "p.4, line 19 and Figure 5b" below).

Figure 5a: Obviously, the assignment of data points (cumulative fraction weight F) to abscissa values (sieve size w_{sieve}) has been done according to the rule "all material that passed this sieve size or was left in this sieve". This assignment is not reasonable because a cumulative fraction weight according to this rule does not depend on the chosen abscissa value – in fact, it only depends on the next larger sieve size ("all material that passed the next larger sieve size"). E.g. the value assigned to $w_{sieve} = 0$ in the figure actually represents all material that passed the 1 mm sieve ($w_{sieve} = 1\,mm$).
Therefore, the plot has to be changed to the correct assignment of abscissa values, i.e. to the standard form of a cumulative grain size distribution curve $F(w_{sieve})$. Applying the commonly used logarithmic scaling of the abscissa is recommended (although not obligatory). As there was no sieve size that has been passed by all material, the abscissa value corresponding to the cumulative fraction weight $F = 1$ cannot be taken from a used sieve size; instead, the size of the intermediate axis of the largest microlithon can be used as an appropriate abscissa value. But caution is required, because the size of the largest microlithon is not constant but might have decreased considerably during the sieving process by disintegration of large microlithons, particularly when increasing the vibration intensity.
According to p.3 line 10, the smallest sieve size was 0.063 mm. The figure does not show data below the 1 mm sieve size. Representing the full data set is recommended.

p.4, line 19 and Figure 5b:

- What was the sieving procedure (sieving time and vibration intensity) applied to obtain the data presented in Figure 5b?
- Usually, the calculation of the quotient N_i/b_i would not be called a "normalization". It is just the calculation of the differential distribution function $f_{raw}(A) = \frac{dN_{cum}}{dA}$ (where A is the measured area of the microlithon, and $N_{cum}(A)$ is the cumulative number of microlithons with area $\leq A$); strictly speaking, it is the approximation of $f_{raw}(A)$ by a quotient of finite differences $f_{raw}(A) \approx \frac{\Delta N_{cum}}{\Delta A} = \frac{N_i}{b_i}$.

  There is no normalization (in the narrower sense) of the differential distribution function applied in figure 5b. Because $N_{cum}(A)$ will increase proportionally with the size of the investigated sample, a normalization with respect to the mass of the investigated sample is required to ensure comparability between curves resulting from different samples. (Obviously, the masses of sample BPS12-SC1 and BFS1-SC1 do not differ by more than one order of magnitude, otherwise, the presented curves would display a larger difference). Therefore, I recommend to replace the presented $f_{raw}(A)$-curves by the normalized $f(A)$-curves, where $f = \frac{f_{raw}}{M_{sample}} \approx \frac{\Delta N_{cum}}{M_{sample}\,\Delta A} = \frac{N_i}{M_{sample}\,b_i}$.

- In any case, the ordinate displays a dimensionful quantity. Therefore, the dimension must be stated in the figure, i.e. "[1/µm²]" in case of f_{raw}, respectively "[1/(kg µm²)]" for the normalized quantity f.

p.4, line 21ff and Figure 5b:

Although values below 0.64 mm² were excluded in the calculation of the regression lines, the corresponding data points should be displayed in Figure 5b (just for completeness of the presentation, and to illustrate why the chosen restriction of the data base for the calculation of the regression lines is reasonable or even necessary).

Obviously, the power-law regression has been performed as a least square regression in logarithmized data ($\ln(f)$ resp. $\ln(f_{raw})$) instead of linear data (f resp. f_{raw}). This approach is required in a data set that covers several orders of magnitude. In Figure 5b the result of the regression is not displayed in the logarithmized form $\ln\left(f(A)\right) = \ln(a_0) - a_1 \ln(A)$, but in the linear form $f(A) = a_0\, A^{-a_1}$ to highlight the power-law. Therefore, indicating that the regression has been performed in logarithmic regime could prevent misconception.

For mathematical reasons the power-law regression for $f(A)$ can only be valid in a limited range of microlithon area A. A particle size distribution $f(A) = a_0\, A^{-a_1}$ valid for $0 \le A < \infty$ would result in a cumulative particle volume diverging toward infinity. Therefore, indicating an observation-based estimate for the validity range is desirable.

In the following, I will briefly show that the available data from sieving and DIA together with an assumption of largely size-independent average shape of microlithons (i.e. the average axes-ratios do not vary substantially with microlithon size) are sufficient to derive an estimate for the validity range:
The available data comprise

- $\{M_k \mid k = 1 \ldots n\}$ Masses of sieve fractions, where $\{M_k \mid k = 1 \ldots j\}$ with $j \le n$ are those fractions that have also been analyzed by DIA.
- $\{A_{k,i} \mid i = 1, \ldots, n_k\}$ All microlithon areas registered by DIA in sieve fraction k.

According to the similarity assumption the following relationship for average microlithon volume $V(A)$ and average microlithon mass $m(A)$ applies:

$$m(A) \sim V(A) \sim A\sqrt{A} \qquad \Rightarrow \qquad m(A) = c\, A^{\frac{3}{2}}$$

The value of c can be derived from the sieving and DIA data, because the sum of masses of all registered microlithons must be equal to the cumulated mass of the investigated sieving fractions:

$$\sum_{k=1}^{j} M_k = \sum_{k=1}^{j}\sum_{i=1}^{n_k} m\left(A_{k,i}\right) = \sum_{k=1}^{j}\sum_{i=1}^{n_k} c\, A_{k,i}^{\frac{3}{2}} \qquad \Rightarrow \qquad c = \frac{\sum_{k=1}^{j} M_k}{\sum_{k=1}^{j}\sum_{i=1}^{n_k} A_{k,i}^{\frac{3}{2}}}$$

Regarding the power-law microlithon size distribution $f(A) = a_0\, A^{-a_1}$, the cumulative microlithon mass fraction distribution $F(A)$ results:

$$F(A) = \int_0^A f(\tilde{A})\, m(\tilde{A})\, d\tilde{A} = \int_0^A a_0\, \tilde{A}^{-a_1}\, c\, \tilde{A}^{\frac{3}{2}}\, d\tilde{A} = \frac{a_0\, c}{\frac{5}{2} - a_1}\, A^{\frac{5}{2} - a_1}$$

This not only highlights that a power-law microlithon size distribution causes a divergent cumulative mass distribution, it also allows to determine an upper limit A_{max} of the validity range for the power-law by

$$F(A_{max}) = 1 \qquad \Rightarrow \qquad A_{max} = \left(\frac{\frac{5}{2} - a_1}{a_0\, c} \right)^{\frac{1}{\frac{5}{2} - a_1}}$$

One cannot derive a lower limit A_{min} of the validity range for the power-law with similar mathematical arguments. But the exclusion of all data with $A < 0.64\ \mathrm{mm}^2$ from the power-law regression limits the range where the validity of the power-law is proven.

Figure 7 & 8: Obviously, the captions given to Figure 7 and 8 have been confused:
The caption given to Figure 8 corresponds to Figure 7.
The caption given to Figure 7 corresponds to Figure 8.
In the following, I indicate several wrong references to these figures. These comments are based on the assumption, that you will keep the sequence of the figures and exchange the captions. If you keep the sequence of captions instead and exchange the figures, most of these comments are obsolete but other references to these figures in the manuscript will require correction in this case. In any case, check the references to these figures carefully.

Figure 7b, 8b, 11, and 14:
There are several rectangles or circles inserted in Figure 7b, 8b, 11, and 14. These indicate details that are shown scaled up in Figure 7c+d, 8c+d, 9b+c, 12, 13c+d, and 15b+c.
According to the scale bars displayed in these figures, the size of the inserted rectangle/circle often does not agree with the size of the corresponding scaled up detail. Either the reason might be an inappropriate size of the inserted rectangle/circle or incorrect scale bars – therefore, check both carefully.
Deviations of ≥15% occur in the following cases:

rectangle/circle indicating detail	size according to scale bar	corresponding detail	size according to scale bar
Figure 7b, inset e	79×74 µm	Figure 9b	62×58 µm
Figure 8b, inset c	69×51 µm	Figure 8c	50×37 µm
Figure 8b, inset d	81×59 µm	Figure 8d	23×17 µm
Figure 8b, inset e	67×51 µm	Figure 9c	102×47 µm
Figure 11, inset c	Ø 12 mm	Figure 12	Ø 19 mm
Figure 14, inset a	24×16 mm	Figure 13c+d	30×19 mm

p.5, line 21: Wrong reference to “Figure 8c”. It must be “Figure 7c” instead.

p.5, line 28: Wrong reference to “Figure 8b and d”. It must be “Figure 7b and d” instead.

p.6, line 4f: *“A simple geometric estimate based on microlithon shape and size distribution shows a total sheared volume fraction of 0.5% for sample BSF1-SC1.”* – The geometric estimate might be simple, but it definitely requires some nontrivial assumptions concerning at least two aspects:

- Thickness of shear bands: Is the estimate based on the assumption of an average thickness of shear bands, independent from the size of the microlithon (which appears reasonable, regarding the impression from Figure 7 to 9)? Indicating the assumed value of average thickness of shear bands might be useful for the reader.

- Microlithon shape: The ratio of surface area to volume $\frac{A_{surf}}{V}$ of microlithons is not self-evident, when the available information concerning microlithon shape is essentially limited to 2D (i.e. the frequency distribution of cross sections A, measured more or less in the orientation of the largest cross section of any microlithon). Which assumptions where employed to derive $\frac{A_{surf}}{V}$ from A?

p.6, line 18: Wrong reference to “Figure 8b”. It must be “Figure 7b” instead.

Figure 14: Wrong reference “a) is a detail shown in Figure 14c and d” in the caption of Figure 14. It must be “a) is a detail shown in Figure 13c and d”.

Figure 15: Wrong reference to “Figure 7c” (in the figure as well as in the caption). It must be “Figure 8c” instead.
To me the evolutionary step from stage ① to stage ② appears not very clear or intuitively comprehensible. From my point of view, it is desirable to make this a bit clearer, either by textual explanation or even better by appropriate amendments in the figure. However, I concede that this might be very (or even too) difficult. Thus, the following is just a suggestion: Indicating the folded bedding and the direction of movement of the conjugate shears (cf. p.8, line 17ff) in stage ② could possibly be helpful.

p.8, line 22: *“more curved convex shear zones”* should be replaced by “more curved shear zones”. The term “convex” appears inappropriate in this context. There are two different meanings of “convex”, but both do not apply (or cannot be demonstrated) in this context:

- “convex” can describe the surface of a body. In this case it is not a property of the surface by itself, but depends on the viewing direction to this surface (from one side, the surface appears convex; from the other side it is concave). In case of the surface of a body, “outwards” defines the preferred viewing direction, facilitating to differentiate e.g. between a convex and a concave lens. Because the shear zones usually separate two microlithons, there is no preferred viewing direction to a shear zone. Therefore, if it is termed “convex” with respect to one microlithon, it has to be termed “concave” with respect to the other one. Therefore, “curved” already provides a complete description.

- “convex” can describe a surface in a mathematical sense, denoting a positive Gaussian curvature (which is an intrinsic property of the surface, independent from the viewing direction). This can possibly apply to most of the shear zones – but for fundamental mathematical reasons this cannot be proven from a 2D-picture (like Figure 15b and Figure 12, which are referenced as a proof in this context).

p.8, line 27: Wrong reference to “Figure 7c”. It must be “Figure 8c” instead.

p.8, line 32: *“the soft material of the shear zones which becomes volumetrically more important”* – according to p.6, line 5 and p.7, line 24, the volume fraction of shear zone material is low (not more than 0.5 to 1%). Therefore, the phrase “volumetrically more important” appears somewhat inadequate. In fact, the soft material of the shear zones becomes mechanically more important, although its volumetric increase is limited to a very small amount.

p.9, line 7: Wrong reference to “Figure 7d”. It must be “Figure 8d” instead.

p.10, line 11: *"The size distribution of the microlithons follows a power law."* – indicating the validity range for the power-law distribution (cf. comment to "p.4, line 21ff and Figure 5b" above) is recommended, particularly within the conclusions.