

3D resolved mapping of optical aberrations in thick tissues

Jun Zeng,¹ Pierre Mahou,¹ Marie-Claire Schanne-Klein,¹
Emmanuel Beaurepaire^{1,3}, and Delphine Débarre^{1,2,*}

¹Laboratory for Optics and Biosciences, Ecole Polytechnique, CNRS, INSERM,
91128 Palaiseau, France

²Laboratory of Interdisciplinary Physics, UJF-CNRS UMR 5588,
38402 St Martin d'Hères, France

³emmanuel.beaurepaire@polytechnique.edu

*delphine.debarre@polytechnique.edu

Abstract: We demonstrate a simple method for mapping optical aberrations with 3D resolution within thick samples. The method relies on the local measurement of the variation in image quality with externally applied aberrations. We discuss the accuracy of the method as a function of the signal strength and of the aberration amplitude and we derive the achievable resolution for the resulting measurements. We then report on measured 3D aberration maps in human skin biopsies and mouse brain slices. From these data, we analyse the consequences of tissue structure and refractive index distribution on aberrations and imaging depth in normal and cleared tissue samples. The aberration maps allow the estimation of the typical aplanatism region size over which aberrations can be uniformly corrected. This method and data pave the way towards efficient correction strategies for tissue imaging applications.

© 2012 Optical Society of America

OCIS codes: (180.4315) Nonlinear microscopy; (110.1080) Active or adaptive optics; (170.3880) Medical and biological imaging; (180.6900) Three-dimensional microscopy.

References and links

1. M. J. Booth, M. A. A. Neil, R. Juškaitis, and T. Wilson, "Adaptive aberration correction in a confocal microscope," *Proc. Nat. Acad. Sci.* **99**, 5788–5792 (2002).
2. P. N. Marsh, D. Burns, and J. M. Girkin, "Practical implementation of adaptive optics in multiphoton microscopy," *Opt. Express* **11**, 1123–1130 (2003).
3. M. Rueckel, J. Mack-Bucher, and W. Denk, "Adaptive wavefront correction in two-photon microscopy using coherence-gated wavefront sensing," *Proc. Nat. Acad. Sci.* **103**, 17137–17142 (2006).
4. D. Débarre, E. J. Botcherby, T. Watanabe, S. Srinivas, M. J. Booth, and T. Wilson, "Image-based adaptive optics for two-photon microscopy," *Opt. Lett.* **34**, 2495–2497 (2009).
5. N. Ji, D. E. Milkie, and E. Betzig, "Adaptive optics via pupil segmentation for high-resolution imaging in biological tissues," *Nat. Methods* **7**, 141–147 (2009).
6. R. Aviles-Espinosa, J. Andilla, R. Porcar-Guezenc, O. E. Olarte, M. Nieto, X. Levecq, D. Artigas, and P. Loza-Alvarez, "Measurement and correction of in vivo sample aberrations employing a nonlinear guide-star in two-photon excited fluorescence microscopy," *Biomed. Opt. Express* **2**, 3135–3149 (2011).
7. A. Facomprez, E. Beaurepaire, and D. Débarre, "Accuracy of correction in modal sensorless adaptive optics," *Opt. Express* **20**, 2598–2612 (2012).
8. N. Olivier, D. Débarre, and E. Beaurepaire, "Dynamic aberration correction for multiharmonic microscopy," *Opt. Lett.* **34**, 3145–3147 (2009).
9. D. Débarre, A. Facomprez, and E. Beaurepaire, "Assessing correction accuracy in image-based adaptive optics," *Proc. SPIE* **8253**, 82530F (2012).

10. H. Ait El Madani, E. Tancrède-Bohin, A. Bensussan, A. Colonna, A. Dupuy, M. Bagot, and A. -M. Pena, "In vivo multiphoton imaging of human skin: assessment of topical corticosteroid-induced epidermis atrophy and depigmentation," *J. Biomed. Opt.* **17**, 026009 (2012).
11. K. König, and I. Riemann, "High-resolution multiphoton tomography of human skin with subcellular spatial resolution and picosecond time resolution," *J. Biomed. Opt.* **8**, 432-439 (2003).
12. D. Débarre, E. J. Botcherby, M. J. Booth, and T. Wilson, "Adaptive optics for structured illumination microscopy," *Opt. Express* **16**, 9290-9305 (2008).
13. M. Kaatz, A. Sturm, P. Elsner, K. König, R. Buckle, and M. J. Koehler, "Depth-resolved measurement of the dermal matrix composition by multiphoton laser tomography," *Skin Res. Technol.* **16**, 131-136 (2010).
14. J. Livet, T. A. Weissman, H. Kang, R. W. Draft, J. Lu, R. A. Bennis, J. R. Sanes, J. W. Lichtman, "Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system," *Nature* **450**, 56-62 (2007).
15. J. Binding, J. Ben Arous, J. -F. Léger, S. Gigan, C. Boccara, and L. Bourdieu, "Brain refractive index measured in vivo with high-NA defocus-corrected full-field OCT and consequences for two-photon microscopy," *Opt. Express* **19**, 4833-4847 (2011).
16. J. Sun, S. J. Lee, L. Wu, M. Sarntinoranont, and H. Xie, "Refractive index measurement of acute rat brain tissue slices using optical coherence tomography," *Opt. Express* **20**, 1084-1095 (2012).
17. Z. Kam, B. Hanser, M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, "Computational adaptive optics for live three-dimensional biological imaging," *Proc. Nat. Acad. Sci.* **98**, 3790-3795 (2001).
18. N. Ji, T. R. Sato, and E. Betzig, "Characterization and adaptive optical correction of aberrations during in vivo imaging in the mouse cortex," *Proc. Nat. Acad. Sci.* **109**, 22-27 (2012).
19. H. Hama, H. Kurokawa, H. Kawano, R. Ando, T. Shimogori, H. Noda, K. Fukami, A. Sakaue-Sawano, and A. Miyawaki, "Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain," *Nat. Neurosci.* **14**, 1481-1488 (2011).
20. M. Hart, "Recent advances in astronomical adaptive optics," *Appl. Opt.* **49**, D17-D29 (2010).
21. N. Olivier, M. Luengo-Oroz, L. Duloquin, E. Faure, T. Savy, I. Villeux, X. Solinas, D. Débarre, P. Bourguin, A. Santos, N. Peyriéras, and E. Beaupaire, "Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy," *Science* **339**, 967-971 (2010).

1. Introduction

Over the past decade, adaptive optics (AO) has received increasing attention as a means to restore the quality of microscopy images inside thick, aberrating samples. This method, based on the use of an active element such as a deformable mirror or a spatial light modulator, consists of applying a deformation to the excitation and/or the detected wavefront that compensates the deformation induced by the sample and the optical setup. AO has proven effective to improve the signal and the resolution of images in a number of imaging modalities and most notably in nonlinear microscopy, where several methods have been developed for measuring aberrations [1, 2, 3, 4, 5, 6].

Although improvement in image quality has been reported in various contexts, less attention has been given so far to the spatial variations of aberrations which limit the quality of correction in extended samples. However, knowledge about the spatial variations of aberrations across the imaged volume is crucial for designing efficient acquisition and correction strategies. In this article, we demonstrate a simple method for mapping aberrations with 3D resolution. The method, based on the principle of image-based modal aberration correction [7], does not rely on a direct wavefront sensor and only requires that an active element such as used for aberration correction be placed on the excitation light path.

We first describe the principle of the image-based modal aberration sensing method and its adaptation for 3D-resolved aberration measurements. In sections 3 and 4, we discuss the accuracy of the method as a function of the signal strength and the aberration amplitude, and derive the achievable resolution for the resulting aberration map. We then report on measured 3D aberration maps in various types of fixed and fresh tissues. We use these data to discuss the consequences of the observed heterogeneities on the quality of correction and the possible correction strategies.

2. Principle of model-based modal aberration measurement

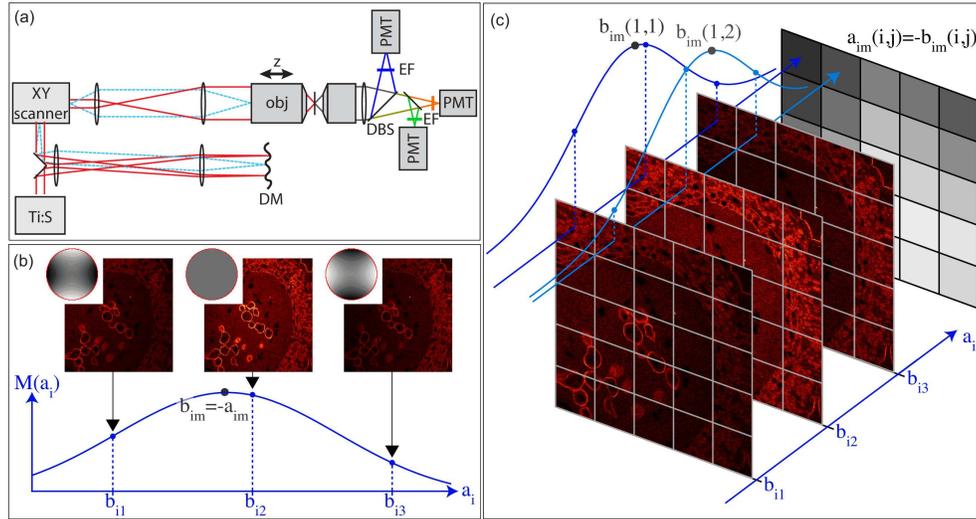


Fig. 1. Principle of spatially resolved modal aberration correction. (a), experimental setup. A titanium-sapphire laser (Ti:S) (Chameleon, Coherent Inc.) is used for excitation. The beam is reflected on a deformable mirror (DM) (Mirao 52e, Imagine Optic), scanned with galvanometric mirrors (General Scanning) and focussed using a 25x, 1.05NA, water immersion, coverslip-corrected objective (obj) (Olympus). 2PEF and second-harmonic signals are collected on photomultiplier tubes (PMT) (Photon Line) using dichroic beamsplitters (DBS) (Chroma) and short-pass emission filters (EF) (Semrock) to eliminate excitation light. (b), principle of aberration measurement. For each aberration mode i , P images are acquired with an amount of aberration b_{ij} ($j=1..P$) applied by the DM. Here, the example of Zernike astigmatism is shown. Metric M (here average image intensity) is subsequently calculated for each value of b_{ij} , and the initial amount of aberration, a_{im} , is estimated as the opposite of the calculated maximum of the curve of M . (c), extension to spatially-resolved aberration measurement: each of the P images are divided in $L \times L$ pixel subregions, and M is calculated for each sub-region of coordinates (k,l) as a function of b_{ij} . The map of aberrations in mode i is obtained as the matrix $\{a_{im}(k,l) = -b_{im}(k,l)\}_{kl}$ where $b_{im}(k,l)$ is the position of the maximum of M .

In modal adaptive optics, the aberrated wavefront at the back aperture of the objective is decomposed onto a truncated series of N aberration modes such as Zernike modes (see appendix A). The aberration is thus described as a vector of coefficients $\{a_i\}_{i=1..N}$, each coordinate a_i being the amount of aberration in mode i . The modes are chosen orthogonal to tip, tilt and defocus for the considered imaging process, i.e. addition of aberration in any mode does not shift the resulting image in the lateral plane or axial direction. For Zernike modes, this is ensured by experimentally measuring the shift of the images and subtracting the required amount of tip, tilt and defocus to compensate for this shift [7]. Coordinates $\{a_i\}_{i=1..N}$ are the sum of the aberrations introduced by the sample and the optical setup, $\{a_{im}\}_{i=1..N}$, and of the aberrations, or biases, applied using the adaptive element (in our case a deformable mirror (DM)), $\{b_i\}_{i=1..N}$. In order to correct for wavefront distortions introduced by the sample and the optical setup, we thus measure one by one the aberration coefficients $\{a_{im}\}_{i=1..N}$.

This is achieved using the following procedure (Fig. 1): for each mode i , the aberration over a given region is measured by acquiring P images of this region with known amounts of

aberration b_{ij} ($j = 1..P$) introduced by the DM. For each of these P images, the image quality is assessed using a metric M calculated from the image and that should exhibit for each mode i a single maximum when $a_i = 0$. The optimal choice for M depends on the imaging method: in this paper we focus on two-photon excited fluorescence (2PEF), for which image average intensity has been shown to be a suitable metric [4]. Using the measured values $\{M(b_{ij})\}_{j=1..P}$, the location b_{im} of the maximum for M is calculated using a model for the curve $M(a_i)$ (for example a Lorentzian or quadratic function, see [7]). The initial aberration a_{im} is then inferred as $-b_{im}$. This process is iterated for the N aberration modes, assuming that they have independent influence on M , i.e. that the measured value a_{im} for mode i does not depend on the amount of aberrations in other modes. This assumption is discussed in more details in the next section.

In the case of spatially-resolved aberration mapping, the $P \times N$ acquired images are split into sub-images of size $L \times L$ pixels and the same process for aberration retrieval is subsequently used on each sub-image to obtain, for each mode i , a map $\{a_{im}(k,l)\}_{kl}$ that has a lateral resolution of $L \times p$ where p is the size of a pixel in the original image (Fig. 1).

In principle, perfect aberration maps could be measured by this method; in practice however, the measurement accuracy depends on several parameters such as the properties of the images, the amplitude of aberrations and the chosen set of modes. Error sources and limitations are described in the next section. In this article, we have focussed on the use of a 2PEF signal to measure aberrations, and of a subset of Zernike modes. It should be noted however that the aberration maps do not depend on the signal that is used for the measurement, but only on the excitation wavelength that is used to create the signal [8]. Furthermore, the same measurement procedure can equivalently be used with different signals and/or set of aberration modes, and the analysis of the sources of error would be similar.

3. Accuracy of the aberration maps

The first parameter that limits the accuracy of the measured aberration maps is the finite number of modes considered: the wavefront would be described perfectly using an infinite set of orthogonal modes such as Zernike modes, but the series used to model it must be truncated in order to limit the number of measurements required to obtain the aberration map. Ultimately, the number of modes that can be measured is limited experimentally by the number of independent modes that the DM can produce, which is equal to the number of mirror actuators that have a non-zero influence on the wavefront over the region of interest (i.e. the back aperture of the focussing objective). In our setup, this would set an upper limit of $N=52$ different modes. In most cases however, the number of modes that are effectively used is smaller; first, because if the chosen functions are not eigenmodes of the wavefront shaper, the number of these modes that can be accurately produced is smaller than the number of eigenmodes. With our mirror, only about 15 Zernike modes (including tip, tilt, defocus and piston) could be accurately produced with significant amplitude even when the full aperture of the mirror was used. Secondly, the number of required measurements scales as $P \times N$, so that acceptable irradiation of the sample limits the number of modes that can be measured before significant bleaching and/or phototoxicity occurs. As a result, only a fraction of the aberrated wavefront can be measured which corresponds to its low spatial frequency content. However, since the same limitation applies to the correction of the aberrations, the measured part corresponds to the part that can be corrected: consequently, although the uncorrected residual wavefront cannot be measured, the gain that can be expected from aberration correction can be accurately determined.

Additionally, the accuracy of the measured aberration values for the N chosen modes relies on the fact that these N modes have independent influence on metric M . This is however rarely the case in practice; as previously investigated [7], linear and nonlinear crosstalks induce shift in the maximum value determined for one mode depending on the amount of aberration

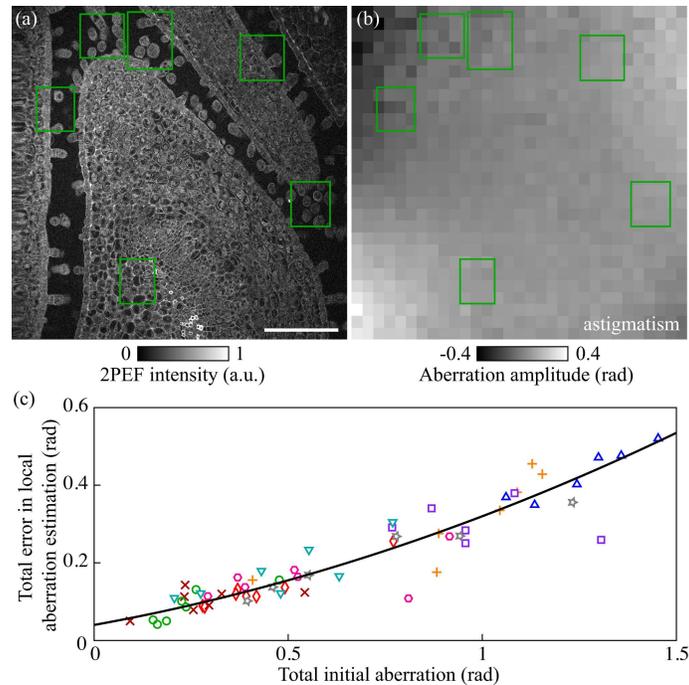


Fig. 2. Calibration of the accuracy of the spatially resolved aberration measurement. (a), example of test sample used for calibration (fixed, stained pine root slice). The green rectangles indicate the locations of the smaller regions over which aberrations were measured iteratively. Scale bar, $100\mu\text{m}$. (b), aberration map obtained on the sample in (a) for astigmatism ($z=5$, see Appendix A). The green rectangles mark the same location as in (a), and delimitate the region over which the aberration map was averaged to compare the spatially-resolved measurement with the iterative result. (c), Error in the spatially resolved measurements as a function of the total amount of aberrations in the 11 Zernike modes under investigation ($z=5$ to 15 , see Appendix A), estimated as the difference between the spatially-resolved value at a given location and the iterative (accurate) local measurement at the same location (green rectangles in (a) and (b)). Both axes are expressed in terms of rms value of the phase profile, or equivalently of the norm of the Zernike coefficient vector. Results for each sample are plotted with a different symbol and colour. The black line is a quadratic fit to the data [9].

in other modes. The amplitude of this effect increases with the initial amount of aberrations and depends on the amplitude of crosstalk between the modes. The resulting error can be reduced by iterating the correction process in the case of correction of a single region, but not in the case of simultaneous measurement on several sub-regions with different aberrations unless aberrations are measured and corrected independently for each subregion. This solution was not investigated here since the required time and exposure would not be compatible with imaging of biological samples.

In order to assess the accuracy of our spatially resolved measurements, we have compared the aberrations measured when correcting a single small region iteratively in conditions that have been determined to yield accurate results [7] with the corresponding values on the aberration map of a larger region, as a function of the amount of aberrations. Since the small regions were larger than the size of the sub-regions, the aberration values on the map were averaged over the zone corresponding to the small region. Several samples with different spatial features were

used. All were fixed, so that variations between the two sets of measurements could be avoided, and robustly stained so as to minimise the effect of photobleaching between the two sets of measurements. An example of such a sample (fixed pine root slice) along with the accuracy results are presented on Fig. 2. In this experiment and in all subsequent experiments, we used 11 low order Zernike modes ($z=5$ to 15, see Appendix A) to model aberrations. The number of measurements per aberration mode P was chosen equal to 5 since this has proven to be an optimal value [7]. Variations in the initial amount of aberrations were obtained by considering different regions in the samples.

As expected, we found that the total error in the spatially resolved measurement increases with the amount of aberrations, and that its amplitude matches previous measurements of the error for a single iteration measurement of aberrations [9]. Over the whole range that was investigated here, the error is roughly of the order of 30% of the aberrations that are measured. This value is therefore a good estimate of the error bars that should be added to the subsequently determined aberration maps for small to moderate amounts of aberrations (up to 2 rad rms). It can be expected, however, that the measurement error might become comparable to the measured amplitudes when considering very large aberrations (greater than 3-4 rad rms). In order to minimise this error, a correction of the average aberration over the entire image can be performed before spatially varying aberrations are measured. In the following experiments, the average system aberration (as measured on a thin test sample) was corrected before aberration maps measurements.

4. Influence of the sub-region size

The measurement error curve shown on Fig. 2 indicates the presence of a non-zero error in the absence of aberrations : this is due to the statistical error arising from the shot noise on the measurements of M . The amplitude dc of this error, in each mode, depends on the total number of photons per sub-region detected during the measurement process for this mode, N_{tot} , as

$$dc = \sqrt{\frac{N_{tot} + P B}{F}}, \quad (1)$$

where B is the dark noise in each sub-image, P is the number of measurements per mode as previously discussed, and F is an amplitude factor related to the modulation depth of M over the range encompassed by the aberration biases (see [7] for more details). In the case of an average correction of aberration, this error can easily be reduced to a negligible value as the total number of photons required to achieve a Strehl ratio greater than 0.9 for 11 modes (or equivalently total residual aberration smaller than 0.325 rad rms; diffraction-limited performance is achieved for a Strehl ratio above 0.8) is only of the order of $10^4 - 10^5$ photons. In contrast for spatially resolved measurements, the same number of photons should be detected for each subregion of much smaller size. When the photon flux is limited (weak signal and/or light sensitive sample), decreasing the size of each sub-regions might lead to excessive error.

Another effect might influence the accuracy of the results: theory indeed predicts that the average image intensity has a single maximum in the absence of aberrations independently of the sample structure under the assumption that the spatial variations of this structure are averaged during the calculation of M [4]. If instead M is calculated in one point only, M might reach a maximum for a non-zero amount of aberrations that increases the size of the focal spot in the case where a region generating a large signal lies near the considered point. As a result, the size of the sub-regions should be chosen large enough so that the increase in size of the focal spot due to aberrations over the range encompassed by the measurement bias is small in comparison. In practice, this limits the sub-region size to a minimum of a few microns.

An additional difficulty in the case where such small sub-regions are used arises in the case of sparsely labelled samples for which some sub-regions might correspond to no-signal areas in the sample: in this case, no measurement can be performed in the corresponding sub-region.

To take into account these various constraints, the measurements and subsequent analysis of aberrations were performed as follows: first, the $P \times N$ images with different biases applied were acquired using the maximum sampling and acquisition time compatible with photobleaching, phototoxicity and movement of the sample over time. The collected signal for each image was thus maximised. Even so, in the various samples that were analysed, the number of photons was usually the limiting factor when reducing the size L of the analysis sub-regions. After acquisition of the images, we thus estimated the sub-region size that would correspond to a given average number of photons, and thus a given statistical error [7]. The number of photons for the aberration maps presented in this article was on average $\approx 10^5$ photons per aberration mode and per subregion, resulting in a sub-region size L of about $10\text{-}20\mu\text{m}$. The aberration maps were then calculated using this grid and finally, if required, a mask was applied to set to zero the aberration values for the $L \times L$ pixels corresponding to dark regions in the original images (see for example Fig. 3(a)). As demonstrated in Fig. 2(c), this procedure permits limiting the shot noise-related error and the influence of the sample structure on the measurements.

The acceptable excitation intensity and imaging time might vary significantly with the type of sample, the imaging depth, or the excitation wavelength. Furthermore for a given illumination intensity the number of detected photons depends on the scattering properties of the sample, its staining, the detection efficiency of the optical setup, etc. Values for the excitation intensities and pixel dwell times required for aberration map measurements, or alternatively achievable resolution of the aberration maps, should thus be determined on a case-by-case basis. The resolution that can be achieved on a given sample with a given setup can for example be estimated by measuring the number of photons collected on a typical image of such sample.

5. Spatially-resolved aberration measurements on human skin biopsies

We first studied the spatial distribution of aberrations in fresh unstained human skin biopsies. The results presented here were obtained on abdominal skin of a 31-year-old caucasian female. The sample was kept frozen until a few minutes before the experiment. After thawing, the epithelium side was covered with Phosphate Buffer Saline (PBS) and a microscope coverslip for imaging. Endogenous fluorescence and second-harmonic generation (SHG) excited at 780nm were used as sources of contrast for imaging the structure of the tissue (Fig. 3(a)-(d) and [Media 1](#)). In the superficial layer (stratum corneum, $0\text{-}20\mu\text{m}$) and in the skin surface ridges (Fig. 3(a) and red arrow in (b)), keratin was the main source of fluorescence. At intermediate depth (stratum granulosum, stratum spinosum and stratum basale, $20\text{-}60\mu\text{m}$) shown on Fig. 3(b), bright melanin fluorescence (yellow arrow) and fluorescence from the cell cytoplasm (blue arrow) were visible, whereas fluorescence at larger depth ($\geq 60\text{-}70\mu\text{m}$) mainly arose from elastin in the dermal papillae (purple arrow), where a SHG signal from fibrillar collagen (shown in green) was also detected [11, 10].

For aberration measurement, blue (410-490nm) fluorescence from the different above cited endogenous fluorophores was used (Fig. 3(e), top). This signal conveniently had a sufficient intensity in all probed parts of the sample to allow measurement of aberrations. Prior to spatially resolved measurements, system aberrations at the centre of the field of view were corrected on a thin sample such as that from Fig. 2. The resulting aberration maps are presented in Fig. 3(e) for spherical aberration and the total aberration amplitude. Before measurement, the correction collar of the focussing objective was adjusted manually to optimise image quality around a depth of $100\mu\text{m}$, so that the average spherical aberration value at lesser depths was negative. Measurement was performed starting at a depth of $40\mu\text{m}$ to avoid imaging the superficial cor-

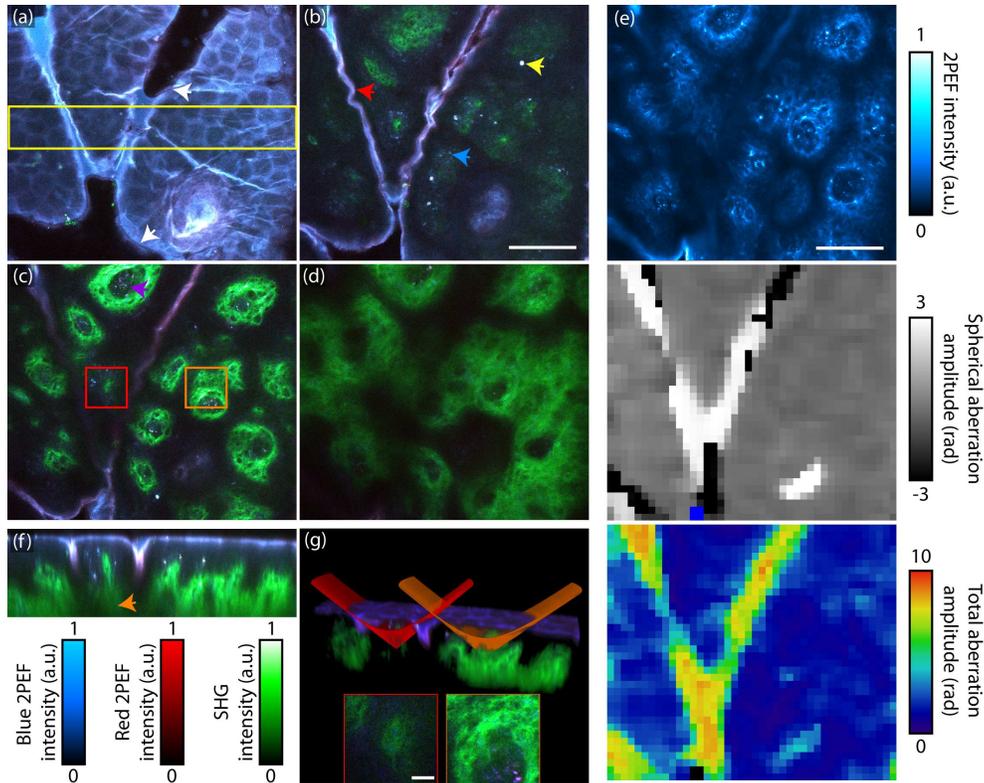


Fig. 3. Aberration measurements on human skin biopsies. (a)-(d), blue (410-490nm) and red (500-600nm) endogenous 2PEF signals and SHG signal (390nm, shown in green) recorded at $10\mu\text{m}$ (a), $35\mu\text{m}$ (b), $60\mu\text{m}$ (c) and $105\mu\text{m}$ (d) below skin surface. The strong aberrations introduced by the stratum corneum are evidenced by the darker pattern in deeper images (white ellipses). [Media 1](#) shows the structure of the sample as a function of depth. Scale bar, $100\mu\text{m}$. (e), endogenous fluorescence signal at a depth of $80\mu\text{m}$ (top), and the corresponding recorded spherical aberration amplitude (middle) and total aberration amplitude (bottom). Aberrations increase as a function of depth ([Media 2](#)) and reflect the structure of the superficial layer. The blue areas in the middle images and the black areas in the bottom images correspond to an absence of measurement due to a lack of signal from the sample. Scale bar, $100\mu\text{m}$. (f), XZ reslice of the data along the middle of the yellow box in (a). The 2PEF and SHG signals are strongly attenuated under the stratum corneum folds (orange arrow) compared with the surrounding regions. (a)-(d) and (f) share the same colour scale shown below (f). (g) and [Media 3](#), 3D reconstruction of the boxed area in (a). The excitation cones corresponding to the effective NA used for imaging (0.95) are plotted for the two positions corresponding to the red and orange boxed area in (a). The corresponding images are shown as inset at $80\mu\text{m}$ deep, demonstrating the correlation between the distortion of the excitation wavefront by the skin ridges and the loss of image resolution. Scale bar, $15\mu\text{m}$.

neocyte layer: indeed this layer corresponds to a region where the signal varies with depth on a very short scale, which may result in erroneous aberration measurement [12, 4]. Aberration maps at depths 0 and $20\mu\text{m}$ in [Media 2](#) are therefore arbitrarily set to zero.

Over the imaged area ($430\times 380\mu\text{m}$), aberrations are strongly correlated to the topology of the overlying stratum corneum with large values detected underneath the folds of the corneocyte

layer. A similar result was obtained for the 10 other probed modes (not shown), as evidenced on the total aberration amplitude map. Aberrations as large as 10 rad rms were measured at moderate depth ($120\mu\text{m}$) and correspond to highly degraded regions in the corresponding images (Fig. 3(d)). It should be noted that the measurement error on such amounts of aberrations is too large to be analysed quantitatively. In particular, in strongly aberrated area, artefacts arise due to the rapid decrease of signal intensity with depth, as visible in the spherical aberration amplitude map: at intermediate depth ($40\text{--}60\mu\text{m}$), the folds introduce negative spherical aberration; however at greater depth ($\geq 80\mu\text{m}$) the sign of the measured spherical aberration changes in regions with large aberrations. We attributed this change to the fact that negative spherical aberration elongates the 2PEF point spread function towards the upper regions of the tissue where chromophore density is higher, resulting in more signal despite the decrease in excitation peak intensity. Aberration measurement based on local intensity variations with the amount of introduced aberrations therefore indicates an opposite (thus positive) coefficient for spherical aberration.

Because large aberrations are mainly created at the superficial folds, the degradation of image quality varies significantly depending on the imaged region, as evidenced on Fig. 3(f) and (g), showing a XZ reslice of the SHG/2PEF images of Fig. 3(a)-(d) and a 3D reconstruction of the same data, respectively. Figure 3(g) and [Media 3](#) emphasize the link between the degradation of the image quality and the ridges of the corneocyte layer encompassed by the excitation light cone. The combination of the measured aberration maps and imaging experiments demonstrate the predominant role of the topology of the stratum corneum in aberration creation at moderate depths: this is explained by the structure of the outermost stratum corneum layers, which consist of a dense assembly of keratin-filled flattened corneocytes surrounded by ordered lipid phases. The refractive index of this layer is thus higher than the surrounding immersion medium and skin layers, thereby introducing wavefront distortions particularly pronounced in fold regions.

Most studies using *in vivo* skin multiphoton imaging have reported high-resolution images down to depths not exceeding $150\mu\text{m}$ [13, 10]. Our results explain the difficulty of maintaining satisfying image quality beyond the epidermis.

6. Aberration measurements and modelling in fixed brain tissues

We also investigated the spatial variations of aberrations in brain tissue samples. The data presented in this article were obtained on $\approx 400\mu\text{m}$ -thick slices of fixed adult mouse hippocampus. Wild type and Thy1-Brainbow-1.0 (L) adult mice [14] were anesthetized and perfused with 4% Paraformaldehyde. Dissected brains were post-fixed overnight in the same fixative prior to sectioning with a vibrating microtome. Sections were mounted in a sealed chamber in PBS or glycerol-based mounting medium (Vectashield, Vector Labs) prior to imaging.

We first investigated the case of a wild-type mouse brain slice mounted in PBS. The results are presented on Fig. 4. The average intensity of the blue (410-490nm) endogenous 2PEF excited at 820nm (Fig. 4(a)) was used as a metric for measuring the effect of aberrations on image quality. Fluorescence is mainly found in the neuropil regions with blood vessel and neuron bodies visible as dark regions. Similarly to what we found on skin biopsies, we measured large amplitudes of aberrations at moderate depths (up to 9 rad rms total aberration at $80\mu\text{m}$, Fig. 4(b)). Again, system aberrations at the centre of the field of view were corrected on a test sample before measurements. Additionally, residual aberrations introduced by the optical setup were removed from the maps presented on Fig. 4 by subtracting the measured aberration maps at the surface of the tissue. These aberrations reached a maximum amplitude around the edges of the field of view of about 0.7 rad rms, and had therefore a limited impact on the images compared to the sample-induced aberrations.

The effect of the aberrations introduced by the sample could directly be observed on the

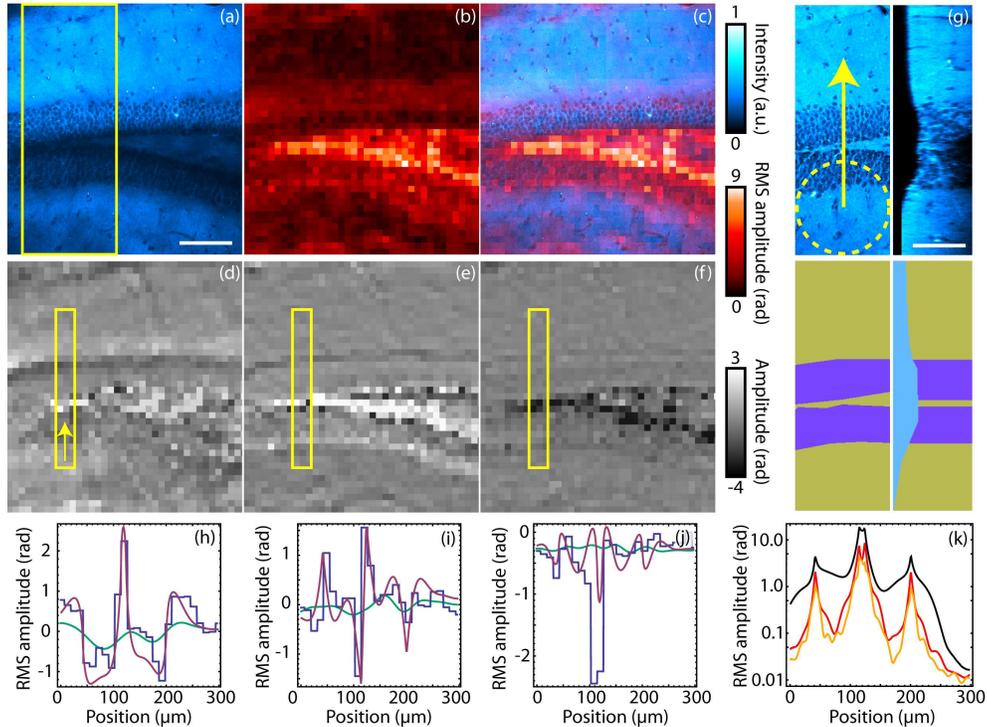


Fig. 4. Aberration measurement on fixed mouse hippocampus mounted in PBS. (a) and [Media 4](#), endogenous 2PEF image at a depth of $80\mu\text{m}$. (b), corresponding total aberration map, and (c) superposition of (a) and (b). (d)-(f), aberration coefficient maps for (d) astigmatism ($z=5$), (e) coma ($z=8$) and (f) spherical aberration ($z=11$). The 3 maps share the same colour scale. (g) and [Media 5](#), region of the tissue modelled for aberration simulation (top), corresponding to the yellow bow in (a). Dotted circle, limit of the excitation cone at the surface of the tissue. Arrow, line along which aberrations were calculated. bottom, model used for the 3D refractive index map. Left, xy image $10\mu\text{m}$ below the surface of the tissue; right, xz reslice. Scale bars, $100\mu\text{m}$. (h)-(j), experimental (blue) and simulated aberration profile with (purple) and without (green) index variation within the tissue at a depth of $80\mu\text{m}$. (h), astigmatism; (i), coma; (j), spherical aberration. (k), residual aberration profile at the same depth after correction of 0 (black), 11 (red) and 32 (orange) Zernike modes.

2PEF images, where regions with the largest aberrations exhibited a reduced signal level, as evidenced by the superposition of the aberration map and the 2PEF image (Fig. 4(c)). This combined image also illustrates the role of the sample geometry in the spatial variations of aberrations, with the largest aberrations located around heterogeneities such as the boundaries between the region containing neuron bodies and the dense neuropil regions. Similar results are found when considering various aberration modes such as astigmatism ($z=5$, Fig. 4(d)), coma ($z=8$, Fig. 4(e)) and spherical aberration ($z=11$, Fig. 4(f)). These three modes accounted here for most of the measured aberration.

In order to clarify the link between the tissue structure and the aberrations, we modelled the wavefront deformation caused by focussing a laser beam within the sample under investigation. We used the images to obtain a simple model of the 3D distribution of refractive index in the brain tissue. In turn, we used this map to estimate the phase accumulated along a given direction from the objective to the focal point. By deducing the local phase for all points within the

excitation cone, the distorted wavefront could be retrieved and fitted onto Zernike modes. The process was iterated for various positions of the focal spot to obtain the spatial variation of the phase and of the fitting Zernike coefficients.

The part of the sample that we modelled is shown in Fig. 4(a) as a yellow box. The corresponding 2PEF image and simplified geometry are shown in Fig. 4(g) and [Media 5](#): the immersion medium (blue) is water, with refractive index of 1.33 at 820nm. The neuropil regions (yellow) contains densely packed axons (in the central area surrounding by the grey matter) or dendrites (outside the grey matter region), and was therefore attributed a higher refractive index n_2 , whereas the cell layer (purple) was considered having an intermediate homogeneous refractive index n_3 (with $n_2 \geq n_3 \geq n_1$). Recent studies have measured *in vivo* refractive indices for brain in the near infrared to be around 1.355-1.37 (grey matter) and 1.4 (white matter) [15, 16]. Due to the dense structure of the neuropil we set $n_2 = 1.39$, whereas the value of n_3 was approximately adjusted to fit the experimental aberration data, resulting in $n_3 = 1.355$. This value is consistent with the previously cited *in vivo* measurements, although in our case the tissue was fixed and the cell bodies mostly filled with PBS.

We chose to model aberrations at the depth of $80\mu\text{m}$, which was the maximum depth investigated here, and where the aberrations were the largest. The width of the modelled part was chosen so as to fit the width of excitation cone when the focal spot was located at this depth. The wavefront was therefore calculated along a line in the centre of the chosen area, pictured as a yellow arrow in (g). The results for modes 5, 8 and 11 are plotted in Fig. 4(h)-(j). For comparison, aberrations measured around the same location (yellow box in (d)-(f)) were also plotted on the same graphs. Measured values were averaged over a width of 2-3 pixels in the aberration maps (which corresponds to $\approx 50\mu\text{m}$) in order to limit fluctuations due to the local structure of the sample that are not taken into account in our simplified model. Finally, the offset value for the simulated spherical aberration ($z=11$, Fig. 4(j)) was adjusted to account for the unknown amplitude of aberration in this mode introduced by the objective correction collar.

Despite the simplicity of this modelling, the simulated curves fit the experimental data remarkably well. Characteristic features of the astigmatism and coma curves (Fig. 4(h) and (i)) are accurately reproduced, and the relative amplitudes of the curves are correct although a single free parameter (n_3) was used. A discrepancy is visible in the case of spherical aberration, for which the dip at position $\approx 100\mu\text{m}$ does not appear in the simulations. This might be due to imperfect modelling of the corresponding region of the tissue, or to measurement uncertainty due to large aberration values as discussed in section 3.

These simulations provide a useful tool to investigate the origin of the measured aberrations. Contrary to the case of human skin, aberrations arise here mostly from in-depth heterogeneity in the tissue rather than from the surface of the sample. This is evidenced in Fig. 4 (h)-(j), where we have plotted in green the simulated aberration in the case of a homogeneous sample ($n_2 = n_3 = 1.38$): the calculated aberration coefficients are much smaller in this case and do not fit the experimental measurements. Although a fraction of the measured astigmatism indeed stems from the index contrast at the non-flat surface of the tissue, most of the image degradation can instead be attributed to the index difference between the cell bodies and the neuropil. This results in aberrations rapidly increasing with depth around the heterogenous regions.

Interestingly, the simulations also provide insight into the total amount of aberrations present in the sample, including higher order modes that have not been measured directly: having validated our model by comparing to experimental measurements for low order Zernike modes, we can use the calculated phase to obtain a rough estimate of the remaining phase aberration after correction of a given number of modes. The results are plotted on Fig. 4(k) for 0 (black), 11 ($z=5$ to 15, red) and 32 ($z=5$ to 36, orange) corrected Zernike modes. As expected, correction of low order Zernike modes significantly improves wavefront quality. Large distortions

of up to 5 rad rms nevertheless remain near the boundaries of the cell layer, where the index contrast is the largest. However, these residual distortions are not significantly reduced when correcting for a larger number of modes, since in our model they stem from abrupt phase steps. Such large phase gradients cannot be compensated due to the finite numerical aperture of the focussing objective, and therefore limit the efficiency of aberration correction. Although this extension of our simple model should be taken with care, this finding indicates that restoring diffraction-limited focussing in such area of the sample might not be achievable.

7. Effect of index matching

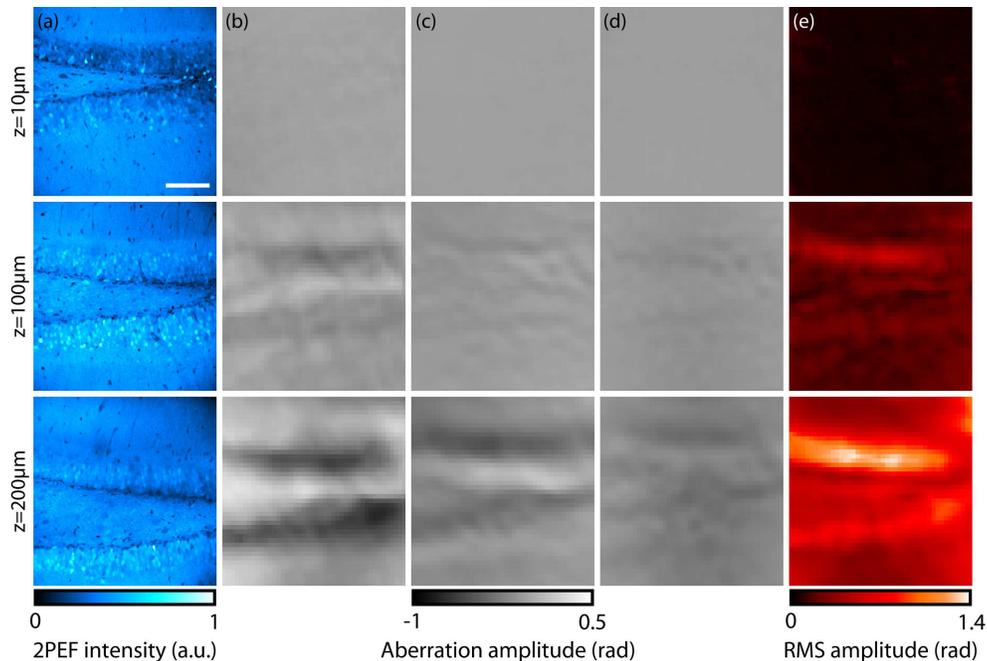


Fig. 5. Aberration measurement on fixed Brainbow mouse hippocampus mounted in Vectashield. (a) and [Media 6](#), CFP and endogenous 2PEF image at depths of 10, 100 and 200 μm . (b)-(d), aberration coefficient maps for (b), astigmatism ($z=5$); (c), coma ($z=8$); (d), spherical aberration ($z=11$). The three maps share the same colour scale. (e), maps of the total aberration amplitude. Scale bar, 100 μm . At each depth, measurement of the 11 aberration maps took about 6 minutes and induced $\approx 7\%$ decrease of the fluorescent signal due to photobleaching.

Alternatively, aberrations can be reduced by use of an appropriate mounting medium that produces index matching between different regions of the sample. We investigated the 3D aberration maps obtained for a slice of mouse brain hippocampus embedded in Vectashield, a mixture of glycerol and water with refractive index of ≈ 1.44 in the visible and near infrared range. System aberrations were first corrected on a test sample and the remaining aberration maps measured on this test sample were subtracted from the maps measured on the brain slice. The results are presented on Fig. 5: again, aberrations were mostly composed of astigmatism ($z=5$, Fig. 5(b)), coma ($z=8$, Fig. 5(c)) and spherical aberration ($z=11$, Fig. 5(d)). However, the aberration maps for these three modes exhibited much lower amplitudes than in the case of the PBS mounted brain slice: at comparable depth (80 μm below the surface), the total aberration

amplitude in the most aberrated areas was reduced by a factor of ten. As a result, neuronal cell bodies were still clearly visible at a depth of $400\mu\text{m}$ (Media 6, Media 7, and Fig. 6).

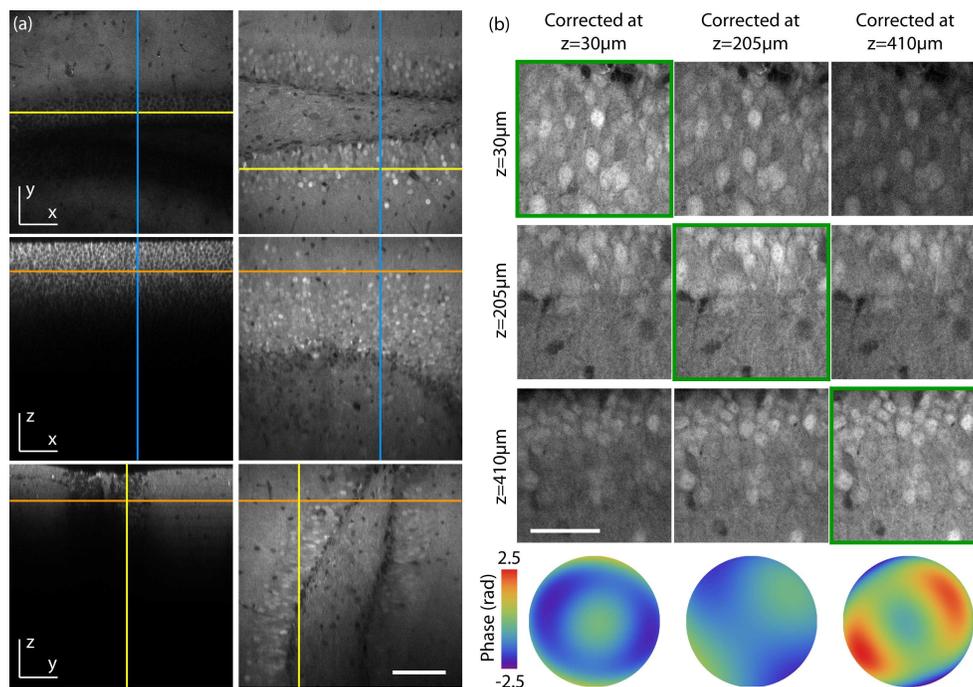


Fig. 6. Deep imaging within fixed mouse hippocampus. (a) and Media 7, comparison of 2PEF imaging over a $400 \times 400 \times 420\mu\text{m}$ volume of a PBS-mounted (left) and a Vectashield-mounted (right) sample. The lines on the images indicate the location of the three visualised planes. Top, xy view (cut along the orange lines); middle, xz view (cut along the yellow lines); bottom, yz view (cut along the blue lines). Scale bar, $100\mu\text{m}$. (b), residual aberration correction on a Vectashield-mounted sample. Within a $100 \times 100\mu\text{m}$ aplanatism region, the quality of images can be maintained throughout the tissue by adjusting the correction phase. Scale bar, $50\mu\text{m}$.

Index-matching using Vectashield did not however cancel all the aberrations introduced by the sample, and the aberration maps of the Vectashield-mounted hippocampus are similar to the one measured in the PBS-mounted sample, indicating that the residual aberrations stem from similar index variations. A possible explanation for this result is that although the refractive index of Vectashield is higher than that of the neuropil, the resulting index of the mounted grey matter has a lower value resulting from residual PBS in the slice and from the remaining structures of cells. Since the amplitude of aberration is nevertheless reduced compared to the case of tissues mounted in PBS, the decrease of the 2PEF signal with depth and of the resolution of the image is also significantly reduced, as shown on Fig. 6: the penetration depth is here increased by a factor of ≈ 5 .

Furthermore, the aberration maps of Fig. 5 show that aplanatism regions, where aberrations can be considered as spatially constant, are typically $50\text{-}100\mu\text{m}$ in size at depth of several hundreds of microns, thereby enabling aberration correction on subregions of the sample. An example of such correction is shown on Fig. 6(b) for different imaging depths. The correction phase consists mostly of astigmatism and spherical aberration in the chosen region. As previously, the correction collar of the focussing objective was set so as to optimise image quality at a

depth around $200\mu\text{m}$ within the sample, which corresponds to the middle of the slice thickness. These results indicate that by combining index matching with residual aberration correction, image quality can be preserved throughout the total thickness ($410\mu\text{m}$) of the sample.

8. Discussion

In this paper, we have presented a new method for measuring the spatial variations of aberrations with 3D resolution within biological samples. We have characterised the achievable spatial resolution, the maximum measurable aberration amplitude and the error on the measurements. Finally, we have measured aberration variations in thick biological tissues, namely human skin biopsies and mouse brain slices. On such samples, we have demonstrated aberration measurements on a $400 \times 400 \times 420\mu\text{m}$ volume (for Vectashield-embedded mouse hippocampus) with a spatial resolution of $10 \times 10 \times 10$ microns.

Previous studies have already reported on the possibility of measuring spatially resolved optical aberrations [17, 18]. However, no aberration map was presented, and the resolution of these methods was not investigated. In the case of Kam et al. [17], although the method for estimating the focal spot distortion is elegant, it relies on the combination of Nomarski and fluorescence contrasts, which may be complicated to set up in an existing microscope, and impossible to use on thick, scattering samples for which no image is obtained in transmission. On the contrary, our approach relies only on the use of the wavefront-shaping device that is also used for aberration correction, and on the signal that is also used to image the sample (2PEF in the examples presented here).

Our data permits estimating the degradation of image quality with depth as well as the possible improvement that can be expected from aberration correction. Two main aspects can thereby be investigated.

First, our data gives indications on the distribution of refractive indices within the sample, and hence on the phase gradients that are introduced during propagation of the excitation beam. In turn, this permits estimating not only the gain in signal and resolution that can be expected from aberration correction, but also the distortions that cannot be corrected using adaptive optics. Indeed, adaptive optics rely on the principle that optical heterogeneities are of small enough amplitude that their effect on the propagation of the excitation beam can be described in terms of wavefront aberration that affect focussing. This does not stand, however, in the case of large index variations. First, because rays experiencing such large variations might experience significant reflected fractions; for example when focussing near the cell region vertical boundary in the hippocampus as modelled in Fig. 4(g), about 20% of the excitation light suffers total internal reflection before reaching the focal spot. Secondly, because the resulting large phase gradients correspond to large changes in the local wave vectors that might fall outside of the range encompassed by the focussing numerical aperture; in this situation, it is impossible to compensate for the distortion as the corrected ray would fall outside of the microscope NA. As demonstrated on the example of fixed mouse hippocampus slices, this effect can be mitigated when possible by the use of an appropriate immersion medium, and more generally by the use of optical clearing agents that limit refractive index variations in fixed samples [19]. The remaining aberrations that stem from smaller index contrasts are then corrected more efficiently using adaptive optics (see Fig. 6).

Secondly, measuring aberration maps within thick samples permits estimating the variation scale of aberrations within the sample. This scale determines the characteristic size of the aplanatism regions, over which a homogeneous correction can be applied to restore a nearly diffraction-limited focus. The size of such regions is critical when estimating the benefits of aberration correction for microscopy: indeed, all the schemes demonstrated so far have used a spatially constant correction in the imaging plane. The quality of correction is therefore di-

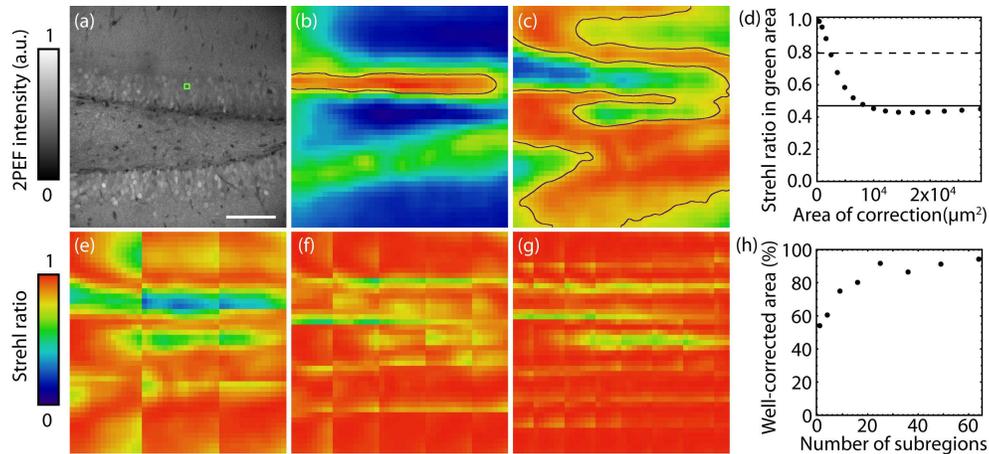


Fig. 7. Aberration correction over a large area. (a), 2PEF image of a Vectashield-embedded brain slide $150\mu\text{m}$ below the surface. Scale bar, $100\mu\text{m}$. (b) (resp. (c)), map of the Strehl ratio after spatially unresolved aberration correction over the green boxed region in (a) (resp. the entire area in (a)). (d), Strehl ratio in the green boxed region after correction over a region centred on this boxed region and of varying size. The dotted line is the lower limit for diffraction-limited focussing and the solid line is the value achieved after correction over the entire image. (e)-(g), Strehl ratio map after spatially resolved correction over (e), 9; (f), 25; and (g), 64 subregions. (h), percentage of the image area for which the Strehl ratio is above 0.8 as a function of the number of subregions used for spatially resolved correction. (b)-(g) maps share the same colour scale and pixel size ($10\mu\text{m}$).

rectly related to the size of the imaged area compared to the size of the aplanatism region. When applying a spatially invariant correction, two approaches are possible: the correction can be optimised on a small region of interest, and the same correction is used for the whole field of view (Fig. 7(b)). In this case, the quality of correction might degrade quickly with the distance to the region of interest, resulting in a very accurate correction over a fairly small area. An alternative approach is to use an averaged correction over a larger region (Fig. 7(c)). The quality of correction for each point however degrades when the size of the chosen region becomes larger than the characteristic scale of aberration variation (Fig. 7(d)).

Spatially varying correction is therefore required (Fig. 7(e)-(h)) when imaging over a region larger than the variation scale of aberrations. This could be achieved by multiconjugate adaptive optics [20], where two or more correction devices with a fixed correction applied are optically conjugated with different planes to create an appropriate aberration combination at different points in the focal plane. Alternatively, for scanning imaging techniques, the scanning pattern can be adapted to the geometry of the sample [21] so that regions of constant aberrations are scanned in turn and aberration correction is updated for each region. In both cases, one difficulty is that if the imaged area is divided into B subregions corresponding to different corrections, the number of measurements required to achieve correction is multiplied by the same factor B . This limits the number of subregions that can be independently corrected, depending on the robustness of the sample and on the time that can be allowed for correction.

On the other hand, the correction procedure would be greatly simplified in situations where the optical properties of the sample are known and predictive correction can be used: in this case, correction can be applied without measuring aberrations. As shown in this study, mapping of aberrations permits gaining such information, that could possibly be used on similar samples.

9. Conclusion

In conclusion, we have presented a technique for mapping optical aberrations within thick samples with 3D resolution in an adaptive microscope, without modification of the optical setup. This method allows characterisation of the aberration properties of various tissues with a lateral resolution of $\approx 10\mu\text{m}$, paving the way towards efficient correction strategies.

Appendix A: Zernike modes and Strehl ratio definition

Table 1 shows the 15 Zernike modes that were used in this paper. If the aberrated wavefront is described by the vector $\{a_i\}_{i=1..N}$ of aberration amplitude in Zernike modes 1 to N , the total root mean square (rms) aberration a (in rad) is obtained as the norm of $\{a_i\}$.

The Strehl ratio is defined as the ratio of the peak intensity of the focussed (aberrated) beam at the focal point to the peak intensity in the absence of aberration. It can be expressed as a function of the amount of aberrations in different Zernike modes as

$$S(a) = \exp[-a^2] = \exp\left[-\sum_{i=1}^N a_i^2\right] \quad (\text{A.1})$$

Table 1. Zernike modes 1 to 15 and numbering scheme. The modes are expressed over the unit disk as functions of r and θ with $0 < r < 1$ and $0 < \theta < 2\pi$.

Index, z	Zernike mode	Name
1	1	Piston
2	$2r \cos \theta$	Tip
3	$2r \sin \theta$	Tilt
4	$\sqrt{3}(2r^2 - 1)$	Defocus
5	$\sqrt{6}r^2 \cos 2\theta$	Astigmatism, 1st order
6	$\sqrt{6}r^2 \sin 2\theta$	Astigmatism, 1st order
7	$2\sqrt{2}(3r^3 - 2r) \cos \theta$	Coma
8	$2\sqrt{2}(3r^3 - 2r) \sin \theta$	Coma
9	$2\sqrt{2}r^3 \cos 3\theta$	Trefoil
10	$2\sqrt{2}r^3 \sin 3\theta$	Trefoil
11	$\sqrt{5}(6r^4 - 6r^2 + 1)$	Spherical aberration
12	$\sqrt{10}(4r^4 - 3r^2) \cos 2\theta$	Astigmatism, 2nd order
13	$\sqrt{10}(4r^4 - 3r^2) \sin 2\theta$	Astigmatism, 2nd order
14	$\sqrt{10}r^4 \cos 4\theta$	Tetrafoil
15	$\sqrt{10}r^4 \sin 4\theta$	Tetrafoil

Acknowledgements

This work was supported by the Agence Nationale de la Recherche (RIB2007 and PIRIBio 2009 programs). The authors thank Maxwell Zimmerley, Jean Livet, Katherine Matho, Willy Supatto and Ana-Maria Pena for discussions and comments. The skin samples were provided by Ana-Maria Pena (L'Oréal Research and Innovation, Aulnay sous Bois, France). The brain tissue samples were prepared and provided by Jean Livet and Katherine Matho (Institut de la Vision, Paris).