Supporting Information

Single-Molecule Counting Coupled to Rapid Amplification Enables Detection of α-Synuclein Aggregates in Cerebrospinal Fluid of Parkinson’s Disease Patients

Akshay Bhumkar, Chloe Magnan, Derrick Lau, Eugene Soh Wei Jun, Nicolas Dzamko, Yann Gambin,* and Emma Sierecki*

anie_202014898_sm_miscellaneous_information.pdf
SUPPLEMENTARY INFORMATION

CSF samples used in this article were obtained from the MJFF LRRK2 Cohort Consortium (LCC), which is coordinated and funded by the Michael J Fox Foundation for Parkinson’s disease research. For up to date information on the LCC study, visit https://www.michaeljfox.org. The investigators within the LCC contributed to the design and implementation of the LCC and/or provided data and collected biospecimens, but did not participate in the analysis or writing of this report.

LCC investigators include Susan Bressman (Beth Israel Medical Center, NY), Nir Giladi (Tel-Aviv Sourasky Medical Centre, Israel), Karen Marder (Columbia University Medical Centre, NY), Jose Marti Masso (Hospital Donostia, Spain), Eduardo Tolosa (Hospital Clinic de Barcelona, Spain), Jan Aasly (University Hospital, Norwegian University of Science and Technology, Norway), Daniela Berg (University of Tuebingen, Germany), Thomas Gasser (University of Tuebingen, Germany), Alexis Brice (Hospital de la Pitie-Salpetriere, France), Jean-Christopher Corval (Hospital de la Pitie-Salpetriere, France), Piu Chan (Xuanwu Hospital of Capital Medical University, China), Emily Drabant (23andMe, CA), Tatiana Foroud (Indiana University School of Medicine, IL), Faycal Hentati (National Institute of Neurology, Tunisia), Matthew Farrer (National Institute of Neurology, Tunisia), Connie Maras (Toronto Western Hospital, Canada), Anthony Lang Maras (Toronto Western Hospital, Canada) and Birgitt Schuele (The Parkinson’s Institute, CA).

METHODS

Expression and purification of recombinant α-synuclein wild-type and K23Q
Wild-type αSyn (WT αSyn) monomer was recombinantly produced from E.coli BL21 DE3 as previously described in Subramaniam et. al44. Purified WT αSyn monomers were flash frozen and stored at -80 °C as 300 µM aliquots until further use. The K23Q point mutation was introduced to the WT αSyn expression vector via site-directed mutagenesis using specific forward (GAAAACCAGCAGGGTGGCAGAAGCAGC) and reverse primer with mutation site highlighted in red (CCCTGCTGGGTTTTCTCAGCAGCC). The resultant K23Q αSyn T7-7 vector was used to produce monomeric αSyn K23Q (K23Q) using similar protocol as for the WT αSyn. The purified K23Q was flash frozen and stored at -80 °C as 300 µM aliquots until further use.

In vitro generation of recombinant α-synuclein fibrils
WT αSyn fibrils were obtained as described before22 by incubating 200µM of monomeric WT αSyn in PBS (pH 7.4) at 45 °C with continuous stirring by a Teflon bar at 500 rpm. At 24 h intervals, the solution was sonicated using a ultrasonic water bath sonicator for 15min. After 72 h, the concentration of the fibrils was calculated (as monomer equivalent) and aliquots of 20µL at 50µM concentration
were stored flash frozen at -80°C until further use. Prior to freezing, 0.1% sodium azide was added to the samples to avoid bacterial growth during the storage.

**Single-molecule instrumentation**

The 3d-printed housing described by Brown et al.\(^{22}\) was fitted with a collimated laser diode, at a wavelength of 450 nm (CPS-450, ThorLabs), a neutral density filter (1.0 density), a dichroic beamsplitter (Di02-R488-25 × 36, Semrock) and a 525/50 nm band-pass filter (FF03-525/50-25, Semrock). A × 40/1.2NA water immersion objective (C-Apochromat, Zeiss) was used to focus the illumination beam and collect the emitted fluorescence which was measured on an avalanche photodiode (APD) detector (50 μm diameter active area, PD-050-CTC, Micro Photon Devices). Data acquisition software were described previously.

**Fluorescence measurements and single molecule counting**

20μL of sample was placed in a custom-built multi-well silicon plate with a 70x80 mm microscope glass coverslip (ProSciTech). Fluorescence of ThT, measured as the number of photons collected by 10 ms time bins, was recorded over the desired amount of time. Data were analysed using a custom Matlab routine, to determine the total fluorescence intensity (integral of all signal), average intensity and the brightness parameter calculated as: \( B = \frac{\sigma^2}{\mu} \), where \( \sigma \) is the standard deviation and \( \mu \), the mean of the data.

A custom-built Python/R data analysis (see SI) was used to automatically detect the presence of fluorescent peaks, and record, for each individual event, the maximal intensity, width (number of time bins) and total intensity (area under the curve). The detection threshold was set at mean + 3 standard deviations.

**Amplification of in vitro generated α-synuclein fibrils**

50 µM WT fibrils aliquots were thawed, and the fibrils were diluted to 2 µM in PBS (pH 7.4) before being sonicated using an ultrasonic water bath sonicator for 10 min. 300µM aliquots of monomeric αSyn K23Q were thawed and the solution was centrifuged through 100 kDa cut-off filters (Amicon Ultra, Millipore) immediately prior to use. The concentration of the flow-through was measured via A280 before being diluted to 30 µM for the amplification assay. The amplification assay was set-up in flat capped PCR strip tubes with final volume of each assay of 20-30 µL. Filter sterilized 100 µM ThT was added to each tube to a final concentration of 10 µM. Depending on the protocol, the PCR tubes were either placed in a thermal mixer/shaker (Eppendorf ThermoMixer-C) set at 37°C with 500 rpm shaking or in a PCR thermal mixer (BioRad C1000 Touch Thermal Mixer) set at the required temperature. For time course measurements, a master mix was prepared and either separated into individual 20µL aliquots or aliquots were taken and analysed over time.

**Direct detection and amplification of aggregates in CSF samples**

2µL of freshly thawed CSF samples were added directly to a well of a custom-built multi-well silicone plate with a microscope glass coverslip containing a solution of 20µM monomeric K23Q αSyn and 2µM ThT, as above. Multiple 100 seconds fluorescence time traces were recorded with a time binning of 10 ms. The mixture was then transferred to a PCR tube, sonicated using an ultrasonic water bath sonicator for 10 min and incubated at 55°C for 5 hours in a PCR thermal mixer (BioRad C1000 Touch Thermal Mixer). After amplification, the mixture was transferred in a clean well of the
silicone plate and fluorescence was measured (100s time traces, 10 ms binning time). Data were analysed as described above.

Statistics
Data are indicated as mean ± s.e.m. if not otherwise specified. Values were collected in Microsoft Excel and compared using unpaired two-tailed Student’s t-tests.

Software for fingerprinting of ThT-positive species

An algorithm was written in Python to analyse the raw traces of AttoBright and deduce the number of ThT-positive species and fingerprint their amplitude and duration. The script is available in the Supplementary Information section.

The raw photon counts (number of photons collected per 10 ms, I(t)) were first filtered with a low-pass filter with a frequency of 50 Hz. The peaks were identified on the resulting filtered trace I_F(t)), to eliminate the probability of very short fluctuations of intensity. To identify a peak, we calculated the standard deviation (SD) of the raw ThT trace in segments of 500 data points (5s). The segments with the lowest standard deviations were considered as baseline ThT signal and did not contain ThT-positive events.

ThT-positive events are considered at bursts in fluorescence where the filtered signal I_F(t) is at least 3xSD above the average signal <I(t)>. The amplitude of the peaks was calculated from the raw I(t) signal to avoid dampening by the temporal filter. The duration of the peak was calculated as the number of time points where I(t) is above the average background <I(t)> signal.
Supplementary Figure 1

Figure S1: Effect of temperature and amplification time on αSyn WT monomers.

20 μM αSyn WT (PBS, pH 7.4) was incubated with 10 μM ThT, at different temperatures (50-70°C) without shaking, immediately after filtration using a filter with 100kDa cut-off. T0-T6 correspond to t=0, 1h, 2h, 3h, 4h,5h and 6h. The dash line represents the average number of aggregates detected for the human αSyn WT at time 0 (negative control). The line was included to facilitate comparison of higher number of aggregates produced during the time course temperature dependent incubation of human αSyn WT.
Supplementary Figure 2

Figure S2: Effect of temperature and amplification time on αSyn K23Q monomers

20 μM αSyn K23Q mutant (PBS, pH 7.4) was incubated with 10 μM ThT, at different temperatures (50-70°C) without shaking, immediately after filtration using a filter with 100kDa cut-off.
Supplementary Figure 3

Amplification of αSyn fibrils as a function of the temperature of incubation. αSyn (K23Q) monomer (20 μM) was incubated in the presence of 100 pM seeds and 10 μM ThT at different temperatures, without shaking, in a PCR machine. Fluorescence time traces were obtained before, and after 3h and 5h amplification. Aggregation was assessed by measuring the B parameter. As reported before, B, calculated as the square of the standard deviation normalized by the mean fluorescence, reports on the inhomogeneity of the traces that arise from the presence of fluorescent bursts when a ThT-positive fibril enters the detection volume. This parameter is concentration-independent, making B a useful screening tool. Note that in the absence of seeds, at t=5h and t=6h, the traces for monomeric αSyn K23Q do not show any detectable spontaneous aggregation at 55℃ (Figure S2). Therefore, this optimal temperature of 55℃ was chosen for the experiments.
Aggregation of αSyn as a function of time in the absence (red squares) or presence (black circles) of 100 nM seeds. Amplification of the seeds was carried out at 55°C without shaking and aggregation was measured at regular intervals. In seeded reactions, the B parameter increases as a function of time to plateau after 4h. Rare events can be detected in unseeded controls at short times (typically between 30min and 2h), however these events disappear in all cases after 4h of incubation at 55°C (also shown in Figure S2). We did not attempt to characterise these aggregates further as they do not proceed to forming ThT positive species that could perturb our sensitivity; these aggregates either dissolve spontaneously or will bind to the plastic container after 4h of incubation.
Figure S5: Estimation of the number and size of the fluorescence bursts showed slight differences with 55°C amplification leading in the formation of more ThT-positive events, especially at the latter time points. Comparison of the fluorescence time traces (Figure 3D), after 5h, shows similar profiles, with little difference in maximal peak size.

Comparison of amplification protocols. Amplification of 100 nM αSyn fibrils was realized in presence of 20 µM K23Q monomer and 10 µM ThT at either 37 °C with orbital shaking at 500 rpm (dark circles) or 55 °C without shaking (dark squares). Fluorescence time traces were obtained at regular intervals. Data were analysed and compared using A) the B parameter as before; B) the number of ThT-positive events per minute or C) the average peak size. D) Typical fluorescence time traces. (Left) 55 °C non-shaking treatment traces obtained at 0h (dark green) and 5h (dark red) and (right) 37 °C with 500 rpm orbital shaking before (light green) and after 5h amplification (light red). All data were acquired on AttoBright using a 405 nM excitation source and a 40x water-immersion objective. Error bars are mean ± s.d of 3 × 300 s measurements.
Figure S6: Examples of traces (sample7, direct detection and amplification);

(left): traces analysed with the single molecule counting script, the * shows the peaks that are detected above background. (right): raw traces represented with the same y-axis scale, to show the absence of peaks of significant amplitude before amplification.
Supplementary Figure 7

Figure S7: Reproducibility of amplification on PD patient (S4) and analysis of intensity and diffusivity of Synuclein aggregates after amplification. (Left): traces obtained in two independent amplification (5h, 55°C) for sample S4. (Right, bottom): typical trace recorded for S4 before amplification. (Right, top): fingerprinting of individual peaks measured in the two independent amplifications of S4.
Figure S8: Number of peaks detected in PD patients (S series) compared to control patients (N series). The unpaired t-test with Welch’s correction (which does not assume that data have the same standard deviation) indicates a significant difference between the group of PD patients and the group of controls, with P value <0.0001.
Figure S9: influence of the acquisition time on AttoBright.

Our experimental plan was to use as little CSF as possible per experiment (10%) and conduct as many separate rounds of experiments as possible to compare PD patients and controls. To avoid possible bias, all 16 samples were amplified and measured at roughly the same time and compared with non-amplified samples. We chose a measurement time of 100s per sample. For diagnostic purposes, it would be more efficient to run a single long read of 10min, before and after amplification for 5h. For short times, we only sample a small volume of the sample and we expect the number of events to scale approximately linearly with the duration of the read. With an acquisition of 10 min (600s) we expect to observe a 6-fold increase in the number of events detected compared to our current 100s protocol.
Supplementary Figure 10

Figure S10: influence of CSF % in the assay on the expected number of events.

With duration of 10min per run, we can predict the number of events detected as a function of CSF concentration: These experiments would enable to check (1) that the number of peaks detected increases linearly with the duration of the run, this would be a first quality check to show that all events are randomly distributed and not grouped in time, which could indicate a source of variability and artefacts. (2) that the number of events increases linearly with the amount of sample tested in the mix, as expected. With larger numbers of events detected (up to 40 aggregates are expected at 40% CSF), statistics of number and size of the individual aggregates would be sufficient to gain new information between patients, and not only between PD positive and negative controls.
Table S1

demographic data associated with the CSF samples used in this study.

| sample | diagnosis | age | gender | Hoehn and Yahr |
|--------|-----------|-----|--------|----------------|
| N4     | control   | 63  | F      | 0              |
| N5     | control   | 41  | M      | 0              |
| N6     | control   | 49  | F      | 0              |
| N7     | control   | 42  | F      | 0              |
| N8     | control   | 71  | F      | 0              |
| N9     | control   | 56  | M      | 0              |
| N10    | control   | 47  | F      | 0              |
| N11    | control   | 45  | M      | 0              |
| S1     | PD        | 66  | M      | 2              |
| S2     | PD        | 63  | M      | 2.5            |
| S3     | PD        | 55  | F      | 1.5            |
| S4     | PD        | 52  | M      | 2              |
| S5     | PD        | 60  | M      | 2              |
| S6     | PD        | 56  | M      | 2              |
| S7     | PD        | 71  | M      | 2              |
| S8     | PD        | 67  | F      | 2              |
### Library
import matplotlib.pyplot as plt
import numpy as np
import matplotlib.image as mpimg
import os
from skimage import exposure
from scipy.ndimage import gaussian_filter
from skimage.morphology import reconstruction
import pandas as pd
from skimage import color
from os import chdir
import scipy.signal as signal
from scipy.signal import butter, filtfilt
from matplotlib.colors import LogNorm
import re

### file loading
```
c = os.getcwd()
repertoire = os.walk(c)
liste_elem = []
for root, dirs, files in repertoire :
    liste_elem.append([root, files])
#
### data processing
```
def butter_lowpass_filter(data, cutoff, fs, order):
    normal_cutoff = cutoff / nyq
    # Get the filter coefficients
    b, a = butter(order, normal_cutoff, btype='low', analog=False)
    y = filtfilt(b, a, data)
    return y

### peak detection
```
# Filter requirements.
T = 3000 # Sample Period
fs = 1000.0 # sample rate, Hz

cutoff = 50 # desired cutoff frequency of the filter, Hz , slightly higher than actual 1.2 Hz
nyq = 0.5 * fs # Nyquist Frequency
order = 2 # sin wave can be approx represented as quadratic
n = int(T * fs) # total number of samples
rapport = 3
taille = 500 # Calcul ecart type
```

for elem in liste_elem :
    chdir(elem[0])
    liste_fichier = elem[1]
    # print (liste_elem)
    liste_y = pd.read_table(fichier)
    label = liste.columns[0]
    liste_y = liste_y[label].tolist()
    y = butter_lowpass_filter(liste_y, cutoff, fs, order)
liste_x = np.linspace(0,len(liste_y)-1,len(liste_y))
liste_prom = []
liste_mean = []
for i in range (1,len(y),taille):
    liste_prom.append(np.std(liste_y[i:i + taille])*3)
    liste_mean.append(np.mean(liste_y[i:i + taille]))
prominence_min = min(liste_prom)
back = liste_mean[liste_prom.index(prominence_min)]
peaks, properties = signal.find_peaks(y, prominence =(prominence_min,None))
widths = signal.peak_widths(y, peaks)
liste_p = []
for p in peaks :
    liste_p.append(y[p])
liste_left = widths[2]
liste_right = widths[3]
liste_intensite = []
liste_integrale = []
for i in range(len(peaks)) :
    left = int(liste_left[i])
    right = int(liste_right[i])
    liste_integrale.append(sum(liste_y[left:right])-len(liste_y[left:right])*back)
    liste_intensite.append(max(liste_y[left:right])
fig = plt.figure(figsize=(50,30))
plt.plot(liste_x,liste_y, label = "prominence : " + str(prominence_min))
plt.plot(liste_x,y, label = "background : " + str(back))
plt.legend(fontsize = 60)
plt.plot(peaks, liste_p, 'o', markeredgecolor = "r")
plt.ylabel("Intensity", fontsize = 60)
plt.xlabel("Time (ms)", fontsize = 60)
plt.title("fichier", fontsize = 80)
plt.tick_params(labelsize = 40)
plt.savefig("fichier+.png")
plt.show()
data = [[len(peaks)],
        peaks,
        liste_intensite,
        widths[0],
        liste_integrale,
        [back]]
df = pd.DataFrame(data, index = ['Peak number',
                                'Peak position',
                                'Peak intensity',
                                'Peak width',
                                'Peak intengrale',
                                'Background'])
df.to_csv("fichier+.csv")
plt.hist(liste_integrale, 20)
plt.title("Integrale : " + fichier)
plt.savefig("fichier+_.png")
plt.show()
plt.hist(liste_intensite, 20)
plt.title("Intensitee : " + fichier)
plt.savefig("fichier+_.png")
plt.show()
# ###########################################################################
# ## data saving
# ###########################################################################
# L = blobs_dog
# dfObj = pd.DataFrame(L)
# dfObj.to_csv("fichier_ref.csv")