Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 297 (2023) 122713

Contents lists available at ScienceDirect

Spectrochimica Acta Part A:
Molecular and Biomolecular Spectroscopy

journal homepage: www.journals.elsevier.com/spectrochimica-acta-part-a-molecular-and-biomolecular-spectroscopy

Monitoring oocyte-based human pluripotency acquisition using synchrotron-based FTIR microspectroscopy reveals specific biomolecular trajectories

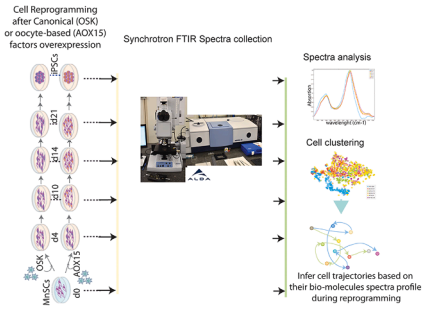
Tanja Dučić [d,*], Alicia Sanchez-Mata [a,b,1], Jesus Castillo-Sanchez [a,b,1], Manuel Algarra [c], Elena Gonzalez-Munoz [a,b,*]

[a] Biomedical Research Institute and Nanomedicine Platform (IBIMA-BIONAND), C/ Severo Ochoa, 35, Malaga, Spain
[b] Department of Cell Biology, Genetics and Physiology, University of Málaga, 29071 Málaga, Spain
[c] INAMAT² Institute for Advanced Materials and Mathematics, Department of Sciences, Public University of Navarre, Campus de Arrosadia, 31006 Pamplona, Spain
[d] ALBA Synchrotron Light Source, Carrer de la Llum 2-26, 08290 Cerdanyola del Vallès, Barcelona, Spain

HIGHLIGHTS

- FITR microspectroscopy allows identification of reprogrammed pluripotent cells.
- Pluripotent iPSCs FITR profile depends on the reprogramming combination used.
- Oocyte-based reprogramming results in specific macromolecular changes.
- iPSC reprogramming cell trajectories detection based on their-specific FITR signature.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Induced pluripotent stem cells (iPSCs)
Oocyte
Reprogramming
FTIR
Synchrotron spectroscopy

ABSTRACT

The reprogramming of human somatic cells to induced pluripotent cells (iPSCs) has become a milestone and a paradigm shift in the field of regenerative medicine and human disease modeling including drug testing and genome editing. However, the molecular processes occurring during reprogramming and affecting the pluripotent state acquired remain largely unknown. Of interest, different pluripotent states have been described depending on the reprogramming factors used and the oocyte has emerged as a valuable source of information for candidate factors. The present study investigates the molecular changes occurring in somatic cells during reprogramming with either canonical (OSK) or oocyte-based (AOX15) combinations using synchrotron-radiation Fourier transform infrared (SR FTIR) spectroscopy. The data acquired by SR FTIR indicates different representation and conformation of biological relevant macromolecules (lipids, nucleic acids, carbohydrates and proteins) depending on the reprogramming combination used and at different stages during the reprogramming process. Association analysis based on cells spectra suggest that pluripotency acquisition trajectories converge at

Abbreviations: AOX15, ASF1A, OCT4 , SOX15 combination; ESC, Embryonic Stem Cells; iPSC, induced Pluripotent Stem Cells; MnSC / MenSC, Menstrual-blood derived Stromal Cells; OSK, OCT4, SOX2, KLF4 combination; SR FTIR, Synchrotron-Radiation Fourier Transform Infrared.

* Corresponding authors at: Biomedical Research Institute and Nanomedicine Platform (IBIMA-BIONAND), C/ Severo Ochoa, 35, Malaga, Spain (E.G.-M.).
E-mail addresses: tducic@cells.es (T. Dučić), egonmu@uma.es (E. Gonzalez-Munoz).

[1] These authors contributed equally.

https://doi.org/10.1016/j.saa.2023.122713
Received 22 November 2022; Received in revised form 9 March 2023; Accepted 3 April 2023
Available online 6 April 2023

late intermediate stages while they diverge at early stages. Our results suggest that OSK and AOX15 reprogramming operates through differential mechanisms affecting nucleic acids reorganization and day 10 comes out as a candidate hinge point to further study the molecular pathways involved in the reprogramming process. This study indicates that SR FTIR approach contribute unpaired information to distinguish pluripotent states and to decipher pluripotency acquisition roadmaps and landmarks that will enable advanced biomedical applications of iPSCs.

1. Introduction

Induced pluripotent stem cells (iPSCs) are derived from somatic cells after over expression of specific "reprogramming factors", generating cells that are genetically reprogrammed into an embryonic-like pluripotent state capable of differentiating into all three germ layers [1].

The reprogramming of somatic cells to iPSCs has opened a new venue to study human development and sequences of pluripotent stages along it, and also has enabled the direct modelling of human disease and basically promises to revolutionize regenerative medicine. IPS cells offer a valuable source of patient-specific pluripotent stem cells for disease model generation, drug testing, toxicology testing and, remarkably, for regenerative medicine, and worth mentioning, they are already being used for the understanding of human diseases [2,3].

Although iPSCs can be reproducibly generated by expressing the four "Yamanaka Factors" OCT4, SOX2, KLF4, and c-MYC (OSKM) or this combination without the oncogene c-Myc (OSK), human iPSC derivation remains inefficient and variable [4]. Reprogrammed cells differ from embryonic stem cells in gene expression patterns and epigenetic profiles, as well as in their differentiation potential, and present difficulties for their cell culture. A thorough understanding of the mechanisms of reprogramming is required to overcome these barriers that hold back the clinical application of iPSCs. A variable of great importance in this regard is the proper reprogramming factors combination [5].

Human oocyte as been shown to be a source of information to understand and improve reprogramming due to the impaired reprogramming capacity of this cell type [6,7]. We have previously identified crucial reprogramming factors involved in the acquisition of pluripotency and we have developed an oocyte-based reprogramming combination (ASF1A, OCT4 and SOX15, called AOX15) that allows to reach distinguishable pluripotent states from that of the OSK combination, with different transcriptional and methylation profile and enhanced differentiation potential [8,9]. However, a full understanding of molecular changes in cellular components of the somatic cells during the reprogramming process - and the dependence on the reprogramming combination used - remains to be accomplished and a systematic analysis of biochemistry changes during the process will shed light on it.

Fourier transform infrared (FTIR) spectroscopy is a powerful analytical technique that allows the analysis of several cell components such as polysaccharides, nucleic acids, protein and lipid contents. They i generate spectra based on the intrinsic property of molecular systems to vibrate in resonance with different frequencies of infrared light. Synchrotron radiation (SR)-FTIR microspectroscopy provides excellent spectral quality at single cell level and thus offers a powerful technique to dissect cellular diversity within a complex system.

The FTIR microspectroscopy, *i.e.* spectroscopy with the micrometer resolution, monitors the overall biochemical composition in the probed bulk sample (*e.g.* tens of cubic microns) the spectroscopical signatures demonstrate slight changes in the biological macromolecular composition. Therefore there have been numerous applications of synchrotron infrared microscopy in different biological fields for probing the macromolecular chemical and conformational changes [11]. On the other hand, the FTIR imaging *i.e.* microscopy, after integration of specific molecular bands, is widely used for tissues, single cells, and subcellular imaging [12]. In addition, the synchrotron infrared beam is used for live cells studies, as there are no beam damage of the cellular structure [13,14].

High spectral resolution of synchrotron-based FTIR (SR-FTIR), combined with fast, stabile measurements makes the method suitable for the high resolution organic compounds resolving in comparison to internal "globar" source. Here, applied SR-FTIR spectroscopy has a ~ 1000-fold higher brightness than conventional sources, and improved the S/N ratio allowing stability for a rapid collection of thousands of infrared spectra over relative large areas of a sample as well as a more precise peak detection (reviewed in [15]). Furthermore, the SR-FTIR combines high spectral resolution of vibrational spectra with high spatial resolution that reaches the diffraction limit [16]. This is a established technique which enables association of vibrational peaks of the IR absorption spectra with specific chemical groups, making it appropriate for the investigation of the chemical properties of biological molecules and molecular complexes [17,18].

The mid IR spectroscopy has proven to be a sensitive analytical tool to examine the structure and conformation of lipids, proteins and nucleic acids and also of complex biological materials such as body fluids and cell cultures in tracking down of biomarkers, reviewed in [10,19,20]. On the other side, near- and far – infrared spectroscopy on the biological samples is also possible. NIR spectroscopy is still a developing discipline and the area of bioanalytical research and medical diagnosis is still in strong competition with mid-IR and Raman spectroscopy (reviewed in [21]). However, current progress includes improvements in the interpretability of NIR spectra, which are valued for their superior chemical specificity, and the progress on the instrumentation that enable environmental monitoring will further extend its bioapplications [21]. The far-infrared spectroscopy is also important for the biological applications, since it is suitable for *in-vivo* measurements without causing damage, and provide information as hydrogen bonds or DNA epigenetic changes. However, the sample preparation, the low-power sources are one of the current challenges, besides limits of the penetration depth by water absorption especially at higher frequencies of infrared [22].

Infrared spectroscopy has been identified as one of the key approaches for this field of "functional biology" as reviewed in [10]. There are a lot of information contained in FTIR spectra such as protein secondary structure – Amide I (1600–1700 cm^{-1}) and Amide II (1500–1560 cm^{-1}), features that arise primarily from the C—O and C—N stretching vibrations of the peptide backbone, respectively. The infrared signature of Amide I and II of the protein has been shown to be above all sensitive to protein secondary structure based on the vibrational frequency of the Amide I and Amide II bands, which are affected by different hydrogen-bonding environments for α-helix, β-sheet, turn and loops, and unordered structures (for a review, see [23]).

In addition to protein's conformational change, FTIR simultaneously provides information about sample biochemistry, with predominant absorption features of the lipids (region 2800–3000 cm^{-1}), carbonyl C═O groups (in the lipids and proteins), and nucleic acids, mostly absorbing in region between 1000 and 1500 cm^{-1} that contains contributions from PO_2 stretching vibrations [10].

In addition to the biological research and diagnostic, there are numerous applications of FTIR spectroscopy. The high spatial resolution and higher spectral quality have a lot of usage in material science, extreme conditions, archeology, earth and planetary science (see review [10]). FTIR is also a key tool in criminal investigation, as the biochemical composition of fingermarks is as a source of vital forensic evidence [24].

FTIR microspectroscopy as a nondestructive technique, sensitive to conformational changes of the bio-macromolecules, has been used to study the pluripotent state of human embryonic stem cells (hESCs) with focus into the differences with their differentiated cell derivates [25–27] and there is still limited literature providing human iPSCs information, with comparative studies indicating that iPSCs show a macromolecular chemistry more similar to hESCs than to either original somatic amniotic fluid cells [28] or differentiated progeny [27] and that reduction of the reprogramming cocktail to only two factors (OCT4 and SOX2) has minimal effect on the molecular characteristics of iPSCs [29].

Our SR-FTIR data show that the representation of biological relevant macromolecules and their conformations differ when either the canonical OSK or the oocyte-based AOX15 combination are used. Additionally, we also evaluated the bio-macromolecular FTIR signatures during the reprogramming process using both combination AOX15 and OSK at different days (day 4, 10, 14, 21 and fully reprogrammed iPSCs). We clearly identified differences in the spectral macromolecular signatures during the reprogramming process. These results confirm previous transcription data of our group suggesting differential pluripotency state acquired depending on the reprogramming factors used and, interestingly suggest differential molecular pathways and molecular states to achieve pluripotency [6,8,9]. Euclidean hierarchical clustering of the

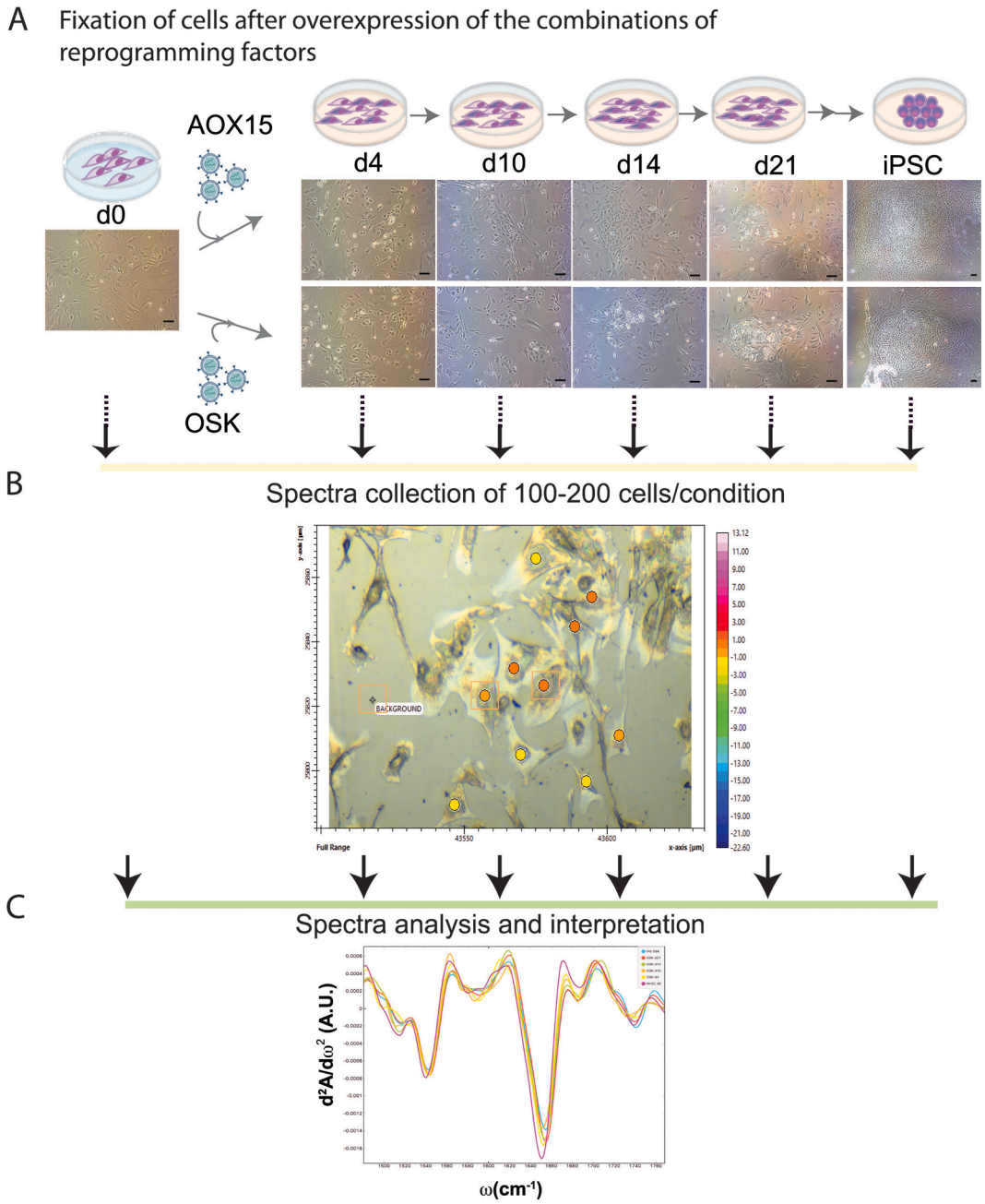


Fig. 1. Experimental workflow A. Primary somatic stromal cell lines (MnSCs) were used to perform retroviral driven reprogramming initiation by over expression of either canonical OSK (OCT4, SOX2 and KLF4) or oocyte-based AOX15 (ASF1A, OCT4 and SOX15) combination of factors. Cells were fixed at day 0, 4, 10, 14 and 21 after factor overexpression and after complete reprogramming into iPSCs forming rounded colonies. Representative bright field images of fixed cells at each stage are shown (scale bar = 20 μm) B. Between 100 and 200 cell spectra for each sample were collected in the 4000–900 cm^{-1} mid-infrared range. C. Spectral analysis for every single cell and interpretation for three different areas: 3020–2800 cm^{-1} lipids, 1800–1480 cm^{-1} proteins and esters, and 1480–900 cm^{-1} nucleic acids and carbohydrate region was applied for further statistical PCA and *t*-SNE and clustering analysis.

samples showed that AOX15 and OSK reprogrammed cells cluster together at iPSC stage and at reprogramming day 21 at all spectral regions analyzed (lipid, protein and nucleic acid) while they show different clustering at earlier reprogramming stages. Our association analysis suggests that there are significant differences in bio-macromolecular composition during reprogramming depending on the reprogramming combination used, and these differences are more evident at early stages.

2. Experimental section

2.1. Cell culture

Primary somatic menstrual-blood derived stromal cell lines (called MnsCs or MenSCs) were culture in DMEM-F12 containing 10 %FBS, 1xNEAA, 1xL-Glutamine, penicillin and streptomycin as previously described [8,9].

We used previously generated OSK and AOX15 iPS cell lines (3 cell lines from each combination) [8] that were cultured in standard human ES cell culture medium (DMEM/F12 containing 20 %KSR, 10 ng/ml of human recombinant basic fibroblast growth factor (bFGF), 1xNEAA, 1xL-Glutamine, 5.5 mM 2-ME, penicillin and streptomycin. iPS cells were cultured on top of mitomycin-C mouse fibroblasts and picked mechanically as previously described [6,8].

All cell lines were regularly tested for mycoplasma using PCR validation (Venor *GeM* Classic, Minerva Biolabs) and found to be negative. At passage 20, iPS cells were seeded on CaF_2 10 × 0.5 mm slides (transparent for infrared and VIS light) and fixed using 4 % paraformaldehyde incubation for 20 min at 25 °C to be later analyzed in ALBA by SR-FTIR beamline MIRAS.

2.2. Reprogramming assay

Primary somatic stromal cell lines were used to perform retroviral driven reprogramming initiation using both canonical OSK (OCT4, SOX2 and KLF4) and AOX15 oocyte-based combination (ASF1A, OCT4 and SOX15) as previously described using pMXs vectors [6,8]. Cells were fixed at day 0, 4, 10, 14 and 21 after factor overexpression and after complete reprogramming into iPSCs using 4 % paraformaldehyde incubation for 20 min at 25 °C to elucidate the effect of each combination reprogramming initiation on the biomacromolecules. Cells were seeded on CaF_2 10 × 0.5 mm slides (transparent for infrared and VIS light). Fixed cells were analyzed in ALBA by SR-FTIR beamline MIRAS (Fig. 1).

2.3. SR-FTIR measurements and analysis

Samples were analyzed at the beamline MIRAS BL01 (ALBA Synchrotron, Spain), by using the 3000 Hyperion microscope coupled to a Vertex 70v spectrometer and a liquid nitrogen-cooled mercury cadmium telluride (MCT) detector. The spectroscopic data were collected in transmission mode using the 36x Schwarzschild objective/condenser and an aperture size of 10 μm × 10 μm. Between 60 and 120 cell spectra for each sample were collected in the 4000–900 cm^{-1} mid-infrared range at a spectral resolution of 4 cm^{-1} with 128 co-added scans per spectrum. Spectra were collected by using the OPUS software (version 8.2, Bruker Company).

Spectral analysis for every single cell and second derivative (Savitzky-Golay filter, 17 smoothing points, third polynomial order, and vector normalization), was implemented for three different areas: 3020–2800 cm^{-1} lipid area, 1800–1480 cm^{-1} proteins and esters, and 1480–900 cm^{-1} nucleic acids and carbohydrate region. Vector normalization was applied after differentiation, as this normalization technique does not require a reference peak.

The principal component analysis (PCA) for each data set was performed, *i.e.*, after the second derivative calculation, and normalization of spectra. The PCA analysis was performed using the Orange software (Bioinformatics Laboratory of the University of Ljubljana [30], Version 3.20.1), with the spectroscopy package [31] (Version 1.2.0) on baseline corrected and unit vector normalized spectra, for spectra without second derivative transformation. The same software was used to perform PCA analysis on the different wavenumber regions, also to calculate integral area mean values of second derivative average spectra at specific wavelengths ranges using integral from 0 method and setting the lower and upper limits of integration at desired wavelenghts and to calculate Euclidean distances among groups, computing a matrix of pairwise distances between average spectra at selected regions where data are normalized to ensure equal treatment of individual features, and also to construct *t*-SNE plots of the data with a *t*-distributed stochastic neighbor embedding method. *t*-SNE is a dimensionality reduction technique, similar to multidimensional scaling, where points are mapped to 2-D space by their probability distribution. *t*-SNE was run with a perplexity value of 30 on the principal components of the input data thus controlling the number of principal components to use when calculating distances between data points. Standardization was applied before running PCA to normalize each data set by subtracting the data set mean and dividing by the standard deviation.

3. Results

3.1. Identification of reprogrammed pluripotent state using SR FTIR spectroscopy

Original somatic menstrual-blood derived stromal cells (MnsCs) and iPSCs reprogrammed using either OSK or AOX15 combination previously generated [8] (Fig. 1) were analyzed using SR FTIR microspectroscopy.

Fig. 2 presents the bio-macromolecular FTIR signatures of somatic cells and derived pluripotent iPS cells using either OSK or AOX15 reprogramming combination. Different conditions are compared and second derivative average spectra are presented in Fig. 2 for three different areas: lipids (A), proteins and carboxyl group (B), and nucleic acids (C). As expected, each cell line has its spectral signature, and somatic cells clearly differ from iPSCs. In the lipid region, both reprogrammed cells differ from the somatic cells, and the differences concern the bands at ~2850, ~2874 and ~2925 cm^{-1} assigned to ν_sCH_2, ν_sCH_3 and $\nu_{as}CH_2$, respectively [32], which were lower in somatic cells while ~2958 cm^{-1} assigned to $\nu_{as}CH_3$ was more pronounced in the somatic cells (Fig. 2A)(SI Table S1). Consistently, the PCA analysis showed a clear contribution of these CH_3 bands in PC1 and PC2, as well as bands assigned to CH_2 groups (Fig. 2D,E) (SI Table S1). The somatic cells in the PCA plot showed clear separation regarding both PCs (mostly PC2) in comparison to reprogrammed cells (Fig. 2D) and loading plots pointed out a maximum peak at ~2870 cm^{-1} and ~2922 cm^{-1} and minimum peak at ~2907 cm^{-1}, ~2850 cm^{-1} and ~2950 cm^{-1} for PC2 (Fig. 2E). Regarding the protein/ester area, differences were observed in the band position of Amide I and Amide II, (Fig. 2B) (SI Table S1). The PCA analysis pointed out the differences in the secondary protein structure, and the PCA score showed segregation in both PC1 and PC3. The somatic cells spectral profile differed from those of both iPSCs for the signals at 1655 and 1670 cm^{-1} (α-helix structure and turn and loops structure, respectively) and the carboxyl group at 1740 cm^{-1} [32] (Fig. 2B,F,G). Differences were confirmed by PCA analysis and the loading plots pointed out a maximum peak for PC3 at ~1655 cm^{-1} and at ~1670 cm^{-1} for PC1. Regarding Amide II, the PC loadings present spectral differences for somatic cells versus pluripotent iPSCs at ~1545 to 1555 cm^{-1} assigned to the α-helix structure and the random coil protein structure, as well as at 1560 cm^{-1}, which is attributed to the β-sheet structure [13,19] (SI Table S1).

In the fingerprint area the peak at ~1070 cm^{-1} corresponding to − CO − O — C stretching related to cholesterol esters and phospholipids [32] was higher in pluripotent cells (Fig. 2C) (SI Table S1).

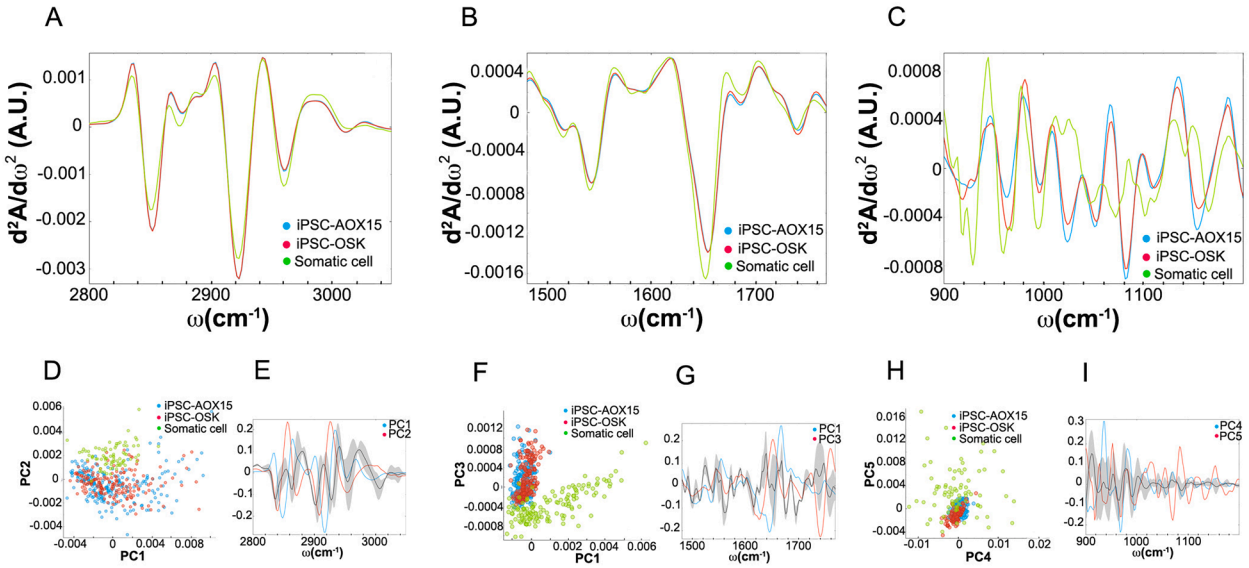


Fig. 2. Second derivative of the FTIR averaged spectra of fixed original somatic cells (green), iPSC reprogrammed with OSK combination (red), and iPSC reprogrammed with AOX15 combination (blue) (N = 120, 179, 185, respectively) in the (A) lipids' spectral region of 3020–2800 cm^{-1}, (B) proteins and carbonyl area, 1800–1480 cm^{-1}, and (C) the nucleic acids region, 1200–900 cm^{-1}. Graphs in (D,F,H) represent the PCA analysis and values of the PC scores for each region assigned above; graphs (E,G,I) show the contribution of individual absorbance to the PCAs (loading values) of the specified principal components in blue and red. Black line and grey shadow represents the average and SD contribution of all principal components explaining 80 % of the variance.

In the nucleic acid area (1200–900 cm^{-1}), several bands were considerably different between somatic cells and pluripotent iPSC which is consistent with major DNA reorganization during reprogramming of somatic cells towards iPSCs [33,34] (Fig. 2C) (SI Table S1). Analysis of the contribution of individual absorbances to PCA showed relevant vibrational bands differences at 925, 965, 1025, 1085 cm^{-1}, corresponding to Z-form DNA, DNA, strongly enhanced in Z-form DNA and B-form DNA respectively [35] (Fig. 2H,I)(SI Table S1). The band at ~1103 cm^{-1} pointed out differences in the process of DNA methylation [36] emphasizing the epigenetic changes occurring during reprogramming. The band at ~1125 cm^{-1} related to RNA [37], was higher in pluripotent cells thus suggesting high transcriptional turnover after reprogramming.

3.2. *Pluripotent iPSCs FITR profile depends on the reprogramming combination used*

When we specifically compared bio-macromolecular FTIR signatures of pluripotent iPSCs depending on the reprogramming combination used (OSK or AOX15) (Fig. 3 and SI Table S2) we confirm the biggest differences in the spectral range of nucleic acids, specifically on the DNA contribution at 960 and 980 cm^{-1}, and interestingly, on the Z-form DNA conformation at 930, 1020 and 1060 cm^{-1}, B-form DNA conformation at 1089 cm^{-1} [35] and DNA methylation at ~1103 cm^{-1} [36] thus suggesting epigenetic differences in DNA organization and structure between OSK and AOX15 iPSCs. These differences were supported by the analysis of the integral area of the specific peaks (shown in SI Fig. S1A

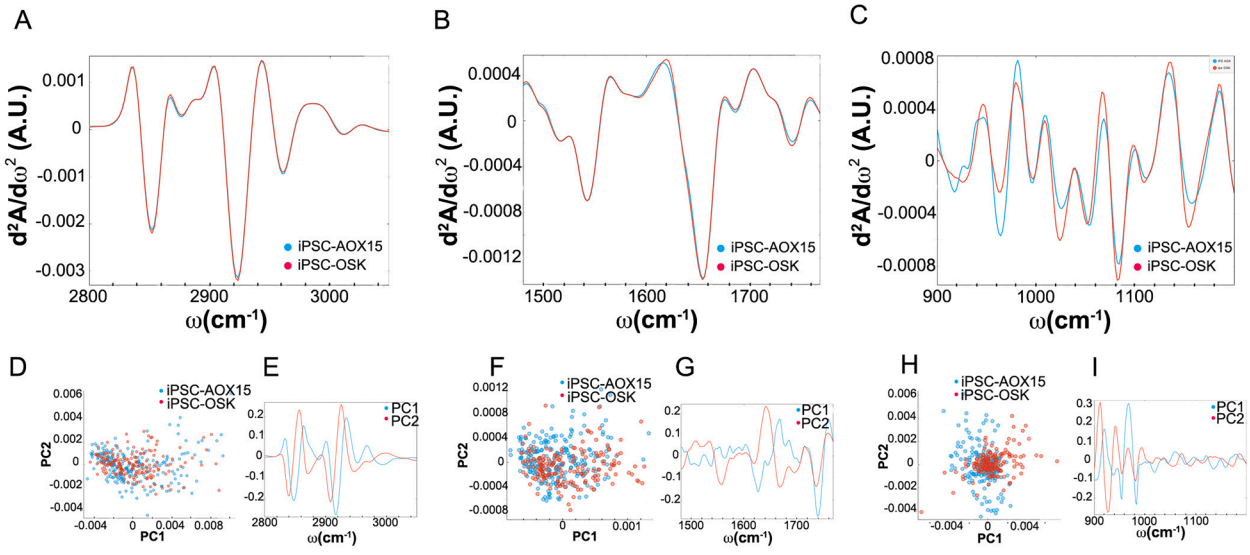


Fig. 3. Second derivative of the FTIR averaged spectra of fixed iPSCs reprogrammed with OSK combination (red), and iPSCs reprogrammed with AOX15 combination (blue) (N = 179, 185, respectively) in the (A) lipids' spectral region of 3020–2800 cm^{-1}, (B) proteins and carbonyl area, 1800–1480 cm^{-1}, and (C) the nucleic acids region, 1200–900 cm^{-1}. Graphs in (D–F) represent the PCA analysis and values of the PC1 and PC2 scores for each region assigned above; graphs (G–I) show the contribution of individual absorbance to the PCAs (loading values) of the first two principal components, PC1 (blue) and PC2 (red) explaining 80 % of the variance.

and SI Table S3). The peak area at ~1125 cm^{-1}, associated with RNA, was also found to be statistically significantly different between iPSCs groups (SI Fig. S1F). Also, the peak area at ~1070 cm^{-1} corresponding to —CO—O—C stretching vibrations in cholesterol esters and phospholipids was higher in OSK than in AOX15 iPSCs (SI Fig. S1G). In the protein region, Amide I band shows fewer differences between OSK and AOX15 iPSCs (Fig. 3B and SI Fig. S1H) that might indicate changes in the amount, structure or function of the proteins (SI Table S2) [38,39]. The ratio of band intensity at 1740 cm^{-1} (originated from the stretching vibrations of νC═O of the ester carbonyl groups) with the sum of CH_2 and CH_3 band intensities (both asymmetric and symmetric) has been described to be correlated to lipid peroxidation or reorganisation [12] and we identify differences between AOX15 and OSK iPSCs in this ratio (SI Fig. S1I). Also vibration of CH_2 at higher frequencies (2925 and 2852 cm^{-1}) that are associated with longer acyl chains and increased fluidity of the membranes [29] were found significant different between iPSCs groups (SI Fig. S1J). The significant shift to higher wavenumbers is associated with a lower conformational order in the plasma membrane was observed earlier [40], their orientation and conformation in cells, showing an increase in lipid peroxidation, but also in membrane instability was reported in [41].

3.3. Macromolecular changes can be detected during the reprogramming process

The bio-macromolecular FTIR signatures were evaluated during the reprogramming process using either combination AOX15 (Fig. 4A–C) or OSK (Fig. 4D-F) on different days (days 4, 10, 14, 21 and fully reprogrammed the iPSCs). FTIR simultaneously provides information about sample biochemistry, with predominant absorption features of the proteins (Amide I and II: 1480–1700), lipid parts (region 2800–3000 cm^{-1}), carbonyl C═O groups, and nucleic acids and sugars that mostly absorb in region between 1000 and 1500 cm^{-1} [10]. Spectra are compared and average spectra are presented in Fig. 4 (and SI Fig. S2) for

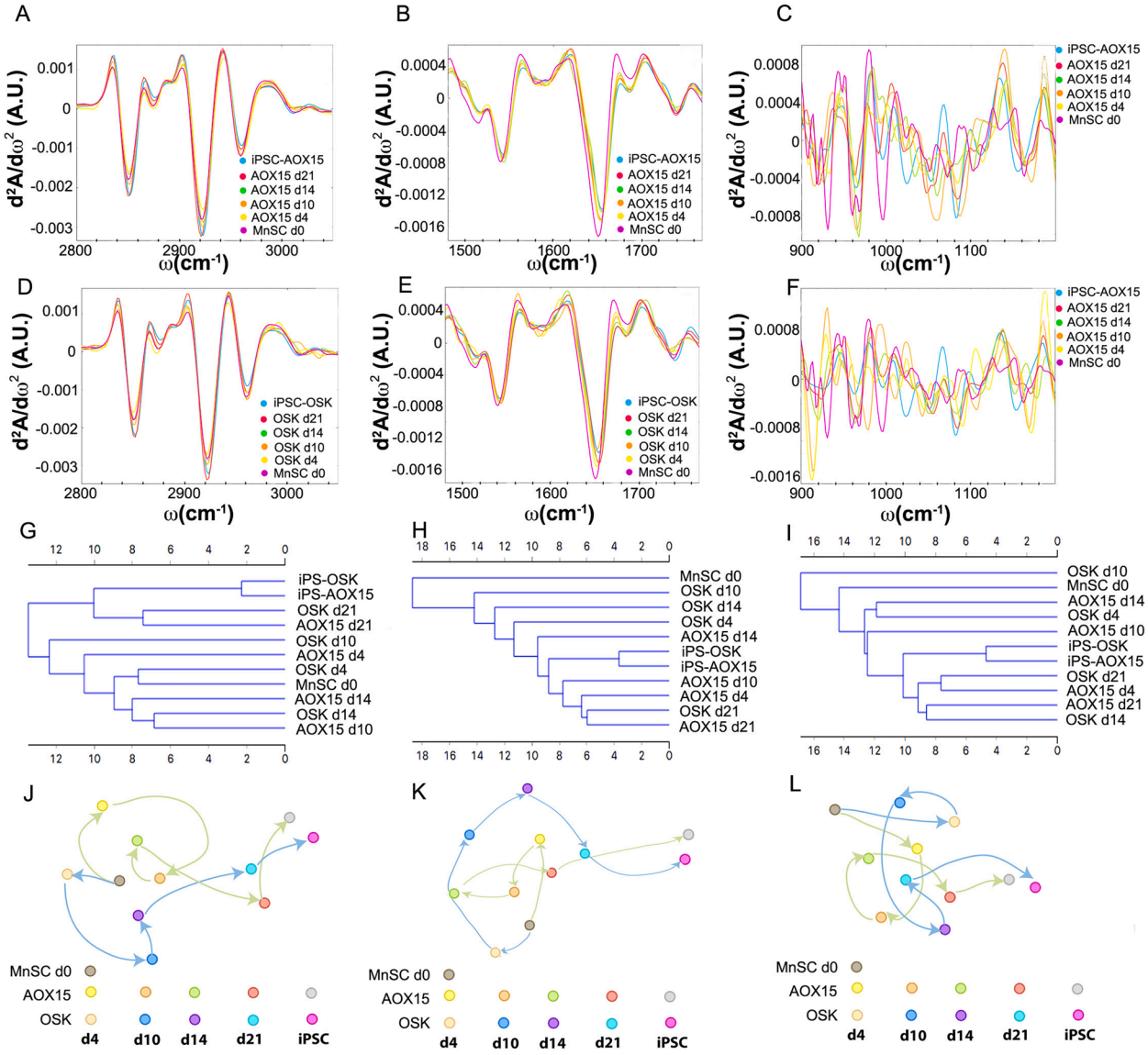


Fig. 4. Second derivative of the FTIR averaged spectra of fixed original somatic cells (MnsC d0) and of AOX15 (A-C) or OSK (D–F) reprogramming intermediates at different days (d4, d10, d14, d21 and fully reprogrammed iPSCs) in the (A,D) lipids' spectral region of 3020–2800 cm^{-1}, (B,E) proteins and carbonyl area, 1800–1480 cm^{-1}, and (C,F) the nucleic acids region, 1200–900 cm^{-1}. (G–I) Hierarchical clustering of AOX15 and OSK intermediates using Euclidean distances based on their average spectra in the lipid (G), proteins and carbonyl groups (H) and nucleic acid (I) regions. (J–L) *t*-SNE representation of second derivative average spectra of OSK and AOX15 reprogramming intermediates groups at different days (d4, d10, d14, d21 and fully reprogrammed iPSCs) at three different areas: lipids (F), proteins and carbonyl group (K), and the nucleic acids (L).

three different areas: lipids (A, D), Amides I and II and carboxyl group (B, E), and the nucleic acids (C, F). We clearly identified differences in the spectral signatures at all regions during the reprogramming process using both combinations AOX15 and OSK. In the lipid region, were observed differences concern the bands at ~2850, ~2874 and ~2925 cm^{-1} assigned to ν_sCH_2, ν_sCH_3 and $\nu_{as}CH_2$, respectively [32] (Fig. 4A,D) (SI Table S4). Consistently, the PCA analysis showed the contribution of these bands (SI Fig. S3 and SI Table S4). Regarding the protein/ester area, differences were mainly observed in the band position of Amide I and Amide II, (Fig. 4B,E) (SI Table S4). The PCA analysis pointed out the differences in the secondary protein structure, including the signals at ~1620, ~1650 and ~1670 cm^{-1} (intramolecular β-sheet structure, α-helix structure and turn and loops structure, respectively) and the carboxyl group at 1740 cm^{-1} [32] (SI Fig. S3B,E and SI Table S4). Regarding Amide II, the PC loadings present spectral differences at ~1545–1555 cm^{-1} assigned to the α-helix structure and the random coil protein structure, [13,19] (SI Fig. S3B,E and SI Table S4). In the nucleic acid area (1200–900 cm^{-1}), several bands were considerably different (Fig. 4C,F, SI Fig. S3C,F and SI Table S1). PCA analysis showed differences of individual vibrational bands contribution at 925, 965, 1025, 1085 cm^{-1}, corresponding to Z-form DNA, DNA, strongly enhanced in Z-form DNA and B-form DNA respectively [35] as well as ~1103 cm^{-1} and ~1125 cm^{-1} bands related to DNA methylation [36] and RNA [37], respectively, emphasizing major DNA reorganization and epigenetic changes during reprogramming of somatic cells towards iPSCs [33,34].

Additionally, the Euclidean distances were calculated based on the average of each spectral region and performed hierarchical clustering of the samples (Fig. 4G–I). AOX15 and OSK reprogrammed cells cluster together at iPSC stage at all spectral regions (lipid, protein and nucleic acid) and at reprogramming day 21 at lipid and protein spectral regions while they segregate separately at earlier reprogramming stages.

We used the dimensionality reduction technique *t*-distributed stochastic neighbour embedding (*t*-SNE) [42] on average spectral regions (lipid, protein and nucleic acids) for each cell group to visualize cell trajectories during reprogramming. The *t*-SNE plots show that cells separated into either OSK or AOX15 reprogramming trajectories and that they get closer at the latest stages (d21 and iPSC) (Fig. 4J–L). These analyses suggest that there are significant differences in biomacromolecular composition during the reprogramming process depending on the reprogramming combination used, and these differences are more evident during the early stages (day 4–day14).

3.4. *Cellular trajectories can be detected during the reprogramming process based on their intermediate cell-specific FITR signature*

In addition, each reprogramming method was analyzed separately, OSK (Fig. 5A–F) and AOX15 (Fig. 5G–L). We performed *t*-SNE and hierarchical clustering based on Euclidean distances of single-cell FTIR signatures based on their second derivative average spectra at specific wavelength ranges corresponding to lipids, proteins and nucleic acids

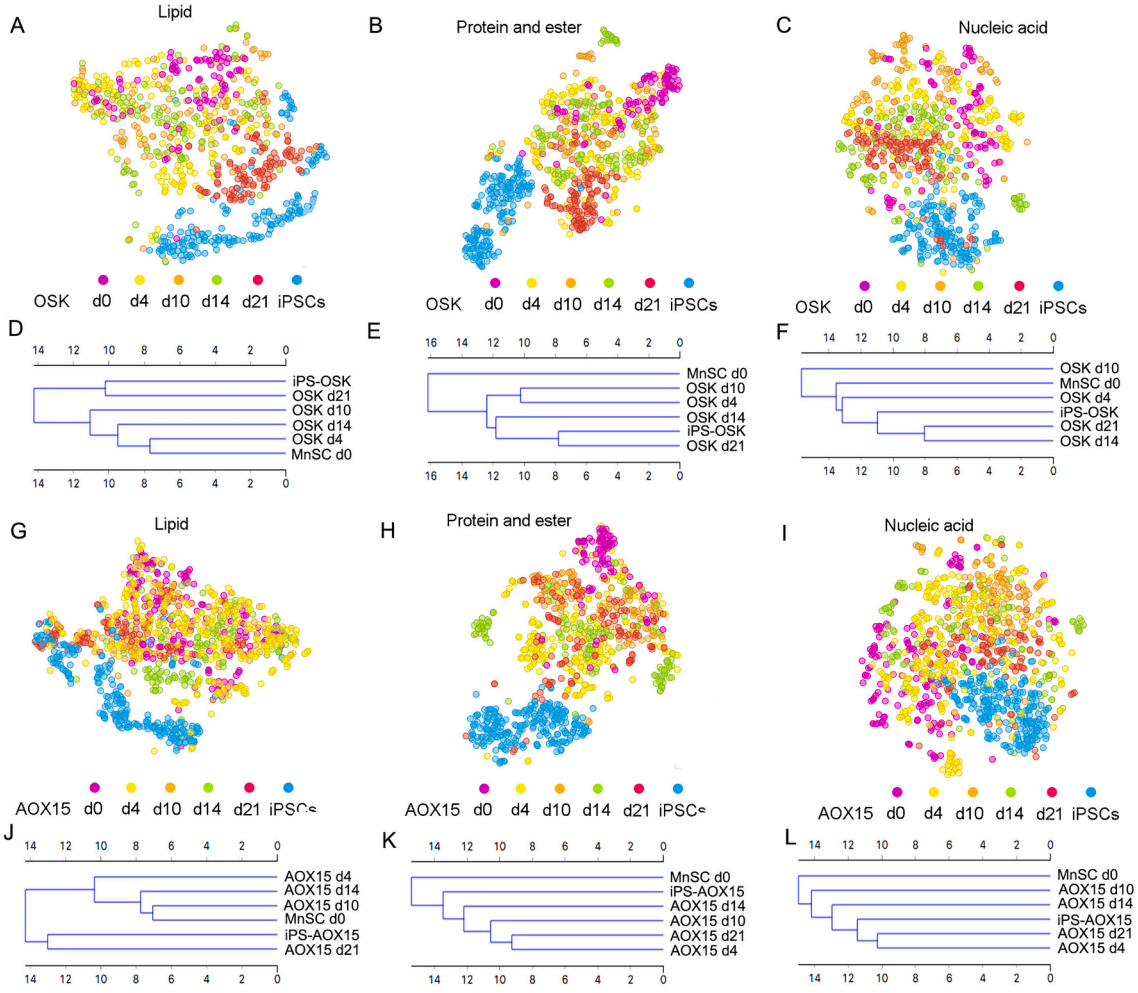


Fig. 5. *t*-SNE representation of second derivative averaged spectra of OSK (A-C) and AOX15 (G-I) reprogramming intermediate cells at different days (d4, d10, d14, d21 and fully reprogrammed iPSCs) at three different areas: lipids (A,G), proteins and carboxyl group (B,H), and the nucleic acids (C,I). Hierarchical clustering of OSK (D–F) and AOX15 (J–L) intermediates using Euclidean distances based on their average spectra in the lipid (D,J), proteins and carbonyl groups (E,K) and nucleic acid (F,L) regions.

spectra. While both OSK and AOX15 reprogramming induced changes in lipid composition that segregate late (d21 and iPSC stage) from early stages (d4–d14) (Fig. 5 A,D,G,J), protein and nucleic acid signature trajectories behave differently depending on the reprogramming method. OSK reprogramming induces progressive changes in protein and nucleic acid composition that allow the sequential cell clustering over time (Fig. 5B,C,E,F). However, AOX15 overexpression promotes protein and nucleic acid sudden changes at the earliest stage (d4) that make these cells cluster with cells at the latest stages (d21) (Fig. 5H,I,K,L). In the nucleic acid spectral range, several bands were considerably different at d4 after AOX15 reprogramming (SI Fig. S4A). Analysis of the contribution of individual absorbance to PCA showed relevant vibrational bands significant differences corresponding to strongly enhanced Z-form DNA, B-form DNA, DNA methylation and RNA [35–37] (SI Fig. S4B,C). These differences were supported by the analysis of the integral area of the specific peaks (SI Fig. S4D–H and SI Table S5). Also, the peak area at ~1070 cm^{-1} corresponding to cholesterol esters and phospholipids was lower four days after AOX15 induction (SI Fig. S4I and SI Table S5). In the protein region, analysis of second derivative spectra and PCA also showed differences in Amide I bands four days after AOX15 reprogramming (SI Fig. S4J-L) and we found significant differences in integral areas of the regions between 1649 and 1660 cm^{-1}, 1665 and 1680 cm^{-1}, and bands centered between 1680 and 1699 cm^{-1} that is attributed to α-helix, turn and loops structures and β-turn structures, respectively [43–47] (SI Fig. S2M–O and SI Table S5).

We focused on nucleic acid spectra changes along reprogramming using both reprogramming factors combinations (Fig. 5C,F,I,L), as remodelling of the epigenome is a crucial barrier that must be overcome for efficient somatic cell reprogramming [34]. We found that day 10 represents a segregating point for cells during both OSK (Fig. 5C,F) and AOX15 (Fig. 5I,L) reprogramming methods.

Ten days after AOX15 overexpression, several bands were different to somatic cells in the nucleic acid area, (Fig. 6A) and the PCA analysis showed segregation of these cell groups mainly in PC6 (Fig. 6B). Analysis of contribution of individual absorbances to PCA (Fig. 6C) and the analysis of the integral area of the specific peaks (SI Fig. S5 and Table 1) showed relevant vibrational bands differences at ~925, ~965, ~1025, ~1085 cm^{-1}, ~1103 cm^{-1} and ~1125 cm^{-1} among others, corresponding to Z-form DNA, DNA, strongly enhanced in Z-form DNA, B-form DNA, DNA methylation and RNA, respectively [35].

OSK reprogramming also caused differences in the nucleic acid spectra after 10 days (Fig. 6D) that enabled cell population identification through PCA analysis (Fig. 6E). Although the analysis of the contribution of individual absorbances to PCA (Fig. 6F) confirmed differences in the DNA spectra, analysis of the integral area of mentioned specific peaks related to known functional DNA conformations (SI Fig. S6A and Table 1) shows significant differences focused solely into ~930 cm^{-1} Z-form of DNA.

The comparison of nucleic acid FTIR profile ten days after reprogramming initiation using either AOX15 or OSK combinations, even if as

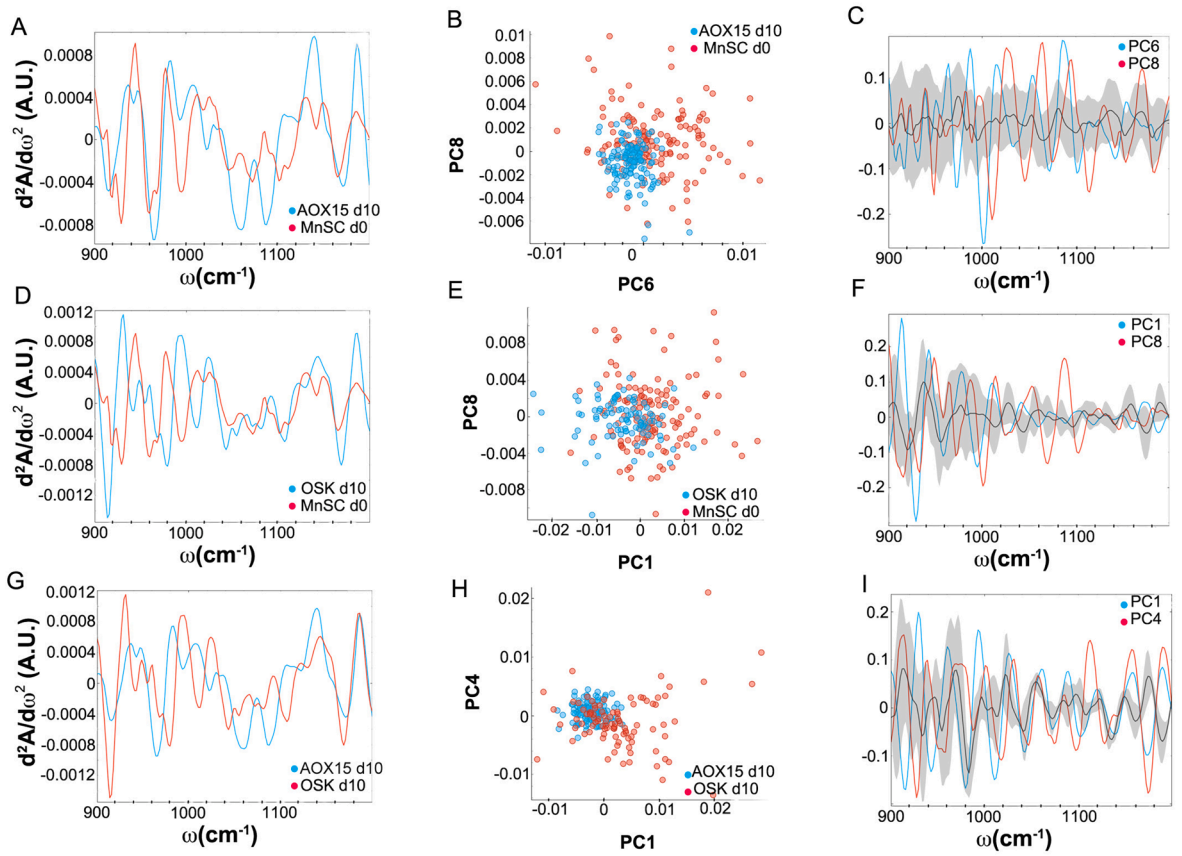


Fig. 6. Second derivative of the FTIR averaged spectra of fixed original somatic cells (MnSC d0) and ten days after AOX15 (A) or OSK (D) reprogramming initiation (AOX15 d10 or OSK d10 respectively) in the nucleic acids region, 1200–900 cm^{-1}. Graph in (B,E) represent the PCA analysis and values of the PC scores for the nucleic acid region for each reprogramming AOX15 (B) or OSK (E) method. Graphs (C,F) show the contribution of individual absorbance to the PCAs (loading values) of the specified principal components in blue and red. Black line and grey shadow represents the average and SD contribution of all principal components explaining 80 % of the variance. Graph in (G) show second derivative of the FTIR averaged spectra of fixed cells after ten days of AOX15 or OSK reprogramming initiation in the nucleic acids region, 1200–900 cm^{-1}. (H) represents the PCA analysis and values of the PC scores for the nucleic acid region and (I) show the contribution of individual absorbance to the PCAs (loading values) of the specified principal components in blue and red. Black line and grey shadow represents the average and SD contribution of all principal components explaining 80 % of the variance.

Table 1
Integral area mean values of second derivative average spectra of MnSC at day 0 (d0) and ten days after AOX15 or OSK over expression (AOX15 d10 or OSK d10) at specific wavelengths ranges related to specific molecules vibration ± standard deviation (SD). T-test score, P-value indicating significant (green) or not significant (red) differences between d0 and AOX15 or OSK d10 groups. Related to SI Figs. S5–S7.

~ Peak wavelength	Molecule/Conformation	Integral Area Mean Value ± SD MnSC d0	Integral Area Mean Value ± SD AOX15 or OSK d10	MnSC d0 vs. AOX15 d10 or OSK d10		AOX15 d10 vs OSK d10	
				T-score	P value	T-score	P value
~ 930 cm^{-1}	Z-form DNA	0.0066 ± 0.0227	AOX15 d10	3.7	p= 0.0001	4.28	0.0001
			0.0007± 0.0061				
			OSK d10	5.8	p= 0.0001		
			0.008 ± 0.0156				
~ 1020 cm^{-1}	Strongly enhanced z-form DNA	0.0039 ± 0.0080	AOX15 d10	4.8	p= 0.0001	4.8	0.0001
			0.0001 ± 0.0045				
			OSK d10	1.1	p= 0.29		
			0.0042 ± 0.0064				
~ 1059 cm^{-1}	Strongly enhanced z-form DNA	0.0025 ± 0.0098	AOX15 d10	8.5	p= 0.0001	5.67	0.0001
			0.0104 ± 0.0049				
			OSK d10	1.4	p= 0.152		
			0.0043± 0.0096				
~ 1089 cm^{-1}	B-form DNA	0.0026 ± 0.0071	AOX15 d10	7.9	p= 0.0001	5.49	0.0001
			0.0087± 0.0055				
			OSK d10	1.7	p= 0.099		
			0.0041± 0.0066				
~ 1100 cm^{-1}	DNA Methylation	0.0007 ± 0.0016	AOX15 d10	5.9	p=0.001	5.09	0.0001
			0.0002 ± 0.0010				
			OSK d10	0.02	p= 0.98		
			0.0007± 0.0016				
~ 1125 cm^{-1}	RNA	0.0039 ± 0.0079	AOX15 d10	2.6	p=0.01	2.8	0.06
			0.0021 ± 0.0025				
			OSK d10	1.1	p= 0.28		
			0.0052± 0.0095				

mentioned, they both induce major nucleic acid remodelling, their effect on nucleic acid composition and conformation is different, as shown by second derivative average spectra comparison, PCA and analysis of the contribution of individual absorbances to PCA (Fig. 6G–I). This divergence was supported by analysis of the integral area of the specific peaks that show significant differences in specific bands related to known functional DNA conformations (SI Fig. S7 and Table 1). The results suggest that OSK and AOX15 reprogramming operates through differential mechanisms affecting nucleic acids reorganization and day 10 comes out as a candidate hinge point to study the molecular pathways involved in each reprogramming process.

4. Discussion

The conversion of human adult somatic cells to pluripotent cells through the expression of defined factors (iPSCs) is a long and complex process that produces cells similar to embryonic stem cells (ESCs) that vary in their developmental potential. Understanding the molecular mechanisms of the human reprogramming process is essential to improve the efficiency and quality of the resulting induced pluripotent stem cells (iPSCs), which is crucial for potential therapeutic applications, and to address fundamental questions about the control of cell identity [48].

Mammalian metaphase II (MII) oocyte has widely been shown to have an unpaired reprogramming capacity with enough epigenetic power to reprogram a somatic cell into an embryonic state in only two cell cycles. Thus, the generation of pluripotent cells using MII oocyte through somatic cell nuclear transfer (or SCNT) in mice and humans is 10-to 100-fold more efficient, as well as less variable, faster, and it is considered less stochastic than iPSC derivation [7,49,50].

We hypothesized that oocyte reprogramming factors bear responsibility for the exceptional reprogramming capacity of SCNT. By analyzing its specific factors [51–53] we have previously identified that the oocyte-enriched factors ASF1A and SOX15, crucial for human pluripotency and we have shown that their overexpression along with OCT4 can reprogram human somatic cells. Our previous results revealed that the oocyte-based reprogramming combination (AOX15) generate iPSCs with a transcriptional and genome methylation profile distinguishable to the canonical OSK combinations. This alternative pluripotent state has relevant functional consequences, showing increased differentiation potential, including challenging PGC-like generation [7–9].

4.1. Macromolecular identification of the pluripotent acquisition

The SR FTIR data indicate that the pluripotent state is readily recognizable from that of MnsC somatic cell, based on the different profiles at major bio-macromolecular components (lipids, proteins and nucleic acids), confirming the feasibility of this approach to distinguish somatic vs pluripotent state as previously shown using amniotic fluid cells instead [28], and opening an intriguing technique on the intact cells to explore molecular components with an important role in cell identity. Our results pointed out that pluripotency acquisition affects lipid metabolism based on the increase in peaks related to $\nu_s CH_2$, $\nu_s CH_3$ and $\nu_{as} CH_2$ [32], while $\nu_{as} CH_3$ contribution was more pronounced in the somatic cells. It has been described that $\nu_{as} CH_2/\nu_{as} CH_3$ is related to fatty acid acyl chain oxidation [54] thus suggesting an increase of this pathway in pluripotent cells. The contribution of bands related to cholesterol esters and phospholipids [32] was also higher in pluripotent cells indicating that the changes also extend to the increase in lipids essential for plasma membranes. Glucose metabolism has been studied in pluripotent stem cells that suffer a shift from oxidative phosphorylation towards aerobic glycolysis [55–57]. However, there is still limited information focused on lipid metabolism and recent studies show a role of c-MYC inducing remodeling of the lipid content, as well as the saturation and length of their acyl chains in mouse cells during pluripotency acquisition [58]. Reprogramming combinations used in this work lack c-

MYC and thus our data suggest that lipid reprogramming also happens without c-MYC overexpression in human cells and invite to extend lipidomic characterization using liquid chromatography–mass spectrometry.

Peaks associated with DNA structure show that reprogramming also affects DNA structure with changes in Z-form and B-form DNA contribution that has been associated with histone acetylation [35], and the increase in DNA methylation related to the band at ~1103 cm^{-1} confirms major epigenetic differences after pluripotency acquisition that have been shown by standard molecular approaches including chromatin immunoprecipitation, DNA sequencing based on bisulfite conversion [59].

4.2. Recognition of the specific pluripotent state

Our analysis reveals differences between AOX15 and OSK iPSCs as well. The biggest dissimilarities concentrate on the nucleic acid contribution and conformation to the spectra, where differences in B-form and Z-form DNA suggest epigenetic differences [60]. Z-DNA has been analyzed for its relevance in regulating transcription. Z-DNA is thought to form in the promoter region of the open chromatin structure induced by a chromatin remodeler and to maintain the open structure stably via binding by proteins such as ADAR1 or Nrf2 in order to support the downstream transcription event [61,62]. The differential DNA-form structures observed between AOX15 and OSK iPSCs invite us to explore proteins participating in such DNA-form changes, including histone acetylases and Z-form binding proteins such us ADAR1, Nrf2 or DLM-1 [63], as reprogramming factor candidates. Of interest, ADAR1 has already been described to be involved in the regulation of reprogramming human fibroblasts to induced pluripotent stem cells through post-transcriptional RNA editing mechanism and molecular mechanisms involved and its relationship with DNA conformation still need to be addressed [64].

In addition, differences in specific bands of lipids and proteins secondary structures also contribute to their distinguishable molecular profile. Of interest, the transition between pluripotent states, such as the ESCs conversion from naive to the primed state, changes their lipid metabolism dramatically, increasing fatty acid synthesis, accumulating lipids, and impeding the use of palmitate as an energy source [65] and it has been shown that changes in the metabolome regulate the epigenetic landscape between pluripotent states. Specifically, nicotinamide *N*-methyltransferase (NNMT) a crucial enzyme involved in metabolic regulation, is required to establish and maintain histone H3 repressive mark (H3K27me3) in the naive state [66]. NNMT is especially abundant in adipose tissue and the liver and its expression affects lipid accumulation during cell differentiation processes [67,68] showing the interconnection between cell identity and biomolecules presence.

4.3. Macromolecular monitoring during pluripotency acquisition

FITR microspectroscopy analysis provides worthy information for the study of reprogramming progression and proposes interesting intermediate cell stage points to study their transcriptional and epigenetic signature to bring the molecular understanding of pluripotency acquisition.

There are few recent studies that reveal the details of reprogramming towards human pluripotency and a comprehensive roadmap of the OSK-driven process is still under construction using complex single-cell transcription analysis, high-throughput DNA methylation profiling and sequencing of accessible chromatin (ATAC-seq) [69–72]. This work contributes not only molecular information, promising candidate time points and intermediate states to delve into molecular factors and pathways involved in reprogramming but also to distinguish among pluripotent states acquired using different triggering combinations.

We have monitored the dynamics of bio-macromolecules chemistry during either oocyte-based AOX15 or canonical OSK pluripotency induction process. Hierarchical clustering based on Euclidean distances and *t*-SNE dimensionality reduction technique of specific spectra ranges reflects biological cell trajectories. Our data suggest that somatic cells undergo differential macromolecular pathways during reprogramming depending on the reprogramming combination used. Interestingly, pluripotency acquisition trajectories seem to converge at late intermediate stages while they diverge at early stages. Association analysis based on nucleic acid and protein spectra suggests that AOX15 reprogramming causes more abrupt changes starting at day 4 than OSK reprogramming, which follows a more gradual pattern of change over time. Nucleic acid spectra profile dynamics suggest day 10 after factor overexpression as a divergent intermediate state using both reprogramming combinations. At this stage, AOX15 reprogramming causes significant changes in peaks related to DNA conformation and methylation and to RNA content as well, while OSK reprogramming effect was milder and significantly different to AOX15's within these areas.

In this study, we employed SR-FTIR microspectroscopy analysis to study macromolecular characteristics of pluripotency acquisition and thus conclusions must be cautious to the limitation of using solely this experimental technique that even if it provides excellent spectral quality at single cell level, the biological significance should be further proven using cellular and functional assays in forward studies. In addition, despite relatively easy classical sample preparation for FTIR, including the fixation with the PFA, the cells are slightly chemical changed in comparison to their physiological conditions. The future studies will include the live cells imaging and spectroscopy in order to obtain information in close to physiological conditions of cells.

Overall, these data support the hypothesis that the pluripotent state of human cells is not a single state. There are already ground-breaking publications in this field demonstrating the existence of the naïve and basic states of human iPSCs and ESCs, and yet, the difficulty and complexity in defining these states are being confirmed [73,74]. Further progress is needed to study the existence of additional pluripotency states or, probably more accurately, sequences of pluripotency states, which may resemble the natural progression after fertilization. For this, the oocyte constitutes a precious source of information as a reprogramming cell par excellence that allows revealing factors, pathways and crucial molecular changes in the process.

Our data suggest that cell identity must be understood as a bio-macromolecular interconnected network where FTIR spectral analysis can contribute valuable information based on cell's biomolecule profiles highlighting unexplored connections among their abundance and conformation.

5. Conclusions

This study explores the bio-macromolecular signature of cells during and after pluripotency acquisition using either canonical OSK reprogramming combination or oocyte-based AOX15 cocktail. Primary, we show that pluripotent cells generated using either of the mentioned combinations are readily distinguished from somatic MnsC with respect to all bio-macromolecules analyzed including lipid, protein and nucleic acid composition and conformation. Second, we demonstrate that the FITR spectral profile of iPSCs depends on the reprogramming combination used, with differences in all macromolecules studied and especially in the nucleic acid spectra, where changes in the symmetric phosphate band vibration, related to different structures of DNA, points toward possible epigenetic differences. Third, we observed macromolecular changes during the reprogramming process. Using hierarchical clustering and *t*-SNE analysis based on FITR spectra we constructed cell trajectories during reprogramming, which show that cells separated into either OSK or AOX15 reprogramming paths and that they get closer at the latest stages while they differ at earlier phases. Last, we focused on nucleic acid spectra changes along reprogramming using both factors combinations and found that day 10 after initiation of reprogramming, represents a segregating point for cell trajectories using both OSK and

AOX15 combinations. In addition, at this stage, AOX15 and OSK cell intermediates show different nucleic acid composition and conformational changes suggesting that OSK and AOX15 reprogramming operate through differential mechanisms affecting nucleic acids reorganization.

CRediT authorship contribution statement

Tanja Dučić: Investigation, Writing – review & editing. **Alicia Sanchez-Mata:** Investigation. **Jesus Castillo-Sanchez:** Investigation. **Manuel Algarra:** Resources. **Elena Gonzalez-Munoz:** Conceptualization, Funding acquisition, Methodology, Investigation, Resources, Formal analysis, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors thank ALBA Synchrotron facility for beamtime allocation and financial support from the Proposal No. 2021085254 and excellent working conditions. E.G-M acknowledge financial support from Ministerio de Ciencia e Innovación del Gobierno de España (grant number PID2021-124033OB-I00) and from Consejería Economía y Conocimiento Junta de Andalucía-FEDER (grant number UMA18-FEDERJA-107). Funding for open access charge was provided by Universidad de Málaga / CBUA.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.saa.2023.122713.

References

[1] K. Takahashi, et al., Induction of pluripotent stem cells from adult human fibroblasts by defined factors, Cell 131 (5) (2007) 861–872.
[2] K. Takahashi, S. Yamanaka, A decade of transcription factor-mediated reprogramming to pluripotency, Nat. Rev. Mol. Cell Biol. 17 (3) (2016) 183–193.
[3] A. Sharma, et al., Multi-lineage human iPSC-derived platforms for disease modeling and drug discovery, Cell Stem Cell 26 (3) (2020) 309–329.
[4] K.K. Haridhasapavalan, et al., An insight into reprogramming barriers to iPSC generation, Stem Cell Rev. Rep. 16 (1) (2020) 56–81.
[5] Y. Shi, et al., Induced pluripotent stem cell technology: a decade of progress, Nat. Rev. Drug Discov. 16 (2) (2017) 115–130.
[6] E. Gonzalez-Munoz, et al., Cell reprogramming. Histone chaperone ASF1A is required for maintenance of pluripotency and cellular reprogramming, Science 345 (6198) (2014) 822–825.
[7] E. Gonzalez-Munoz, J.B. Cibelli, Somatic cell reprogramming informed by the oocyte, Stem Cells Dev. (2018).
[8] L. Lopez-Caraballo, et al., Analysis of menstrual blood stromal cells reveals SOX15 triggers oocyte-based human cell reprogramming, iScience 23 (8) (2020), 101376.
[9] A. Sanzhez-Mata, A. Ferez-Gomez, E. Gonzalez-Munoz, Protocol to reprogram human menstrual blood-derived stromal cells to generate AOX15-iPSCs, STAR Protoc 1 (3) (2020), 100183.
[10] P. Dumas, M.M.C.C.L., IR Spectroscopy and spectromicroscopy with synchrotron radiation, in: Synchrotron Light Sources and Free-Electron Lasers: Accelerator Physics, Instrumentation and Science Applications, 2020, pp. 2059–2113.
[11] L.M. Miller, P. Dumas, From structure to cellular mechanism with infrared microspectroscopy, Curr. Opin. Struct. Biol. 20 (5) (2010) 649–656.
[12] T. Ducic, et al., Multimodal synchrotron radiation microscopy of intact astrocytes from the hSOD1 G93A rat model of amyotrophic lateral sclerosis, Anal. Chem. 91 (2) (2019) 1460–1471.
[13] T. Ducic, et al., Live-cell synchrotron-based FTIR evaluation of metabolic compounds in brain glioblastoma cell lines after riluzole treatment, Anal. Chem. 94 (4) (2022) 1932–1940.
[14] E. Mitri, et al., SU-8 bonding protocol for the fabrication of microfluidic devices dedicated to FTIR microspectroscopy of live cells, Lab Chip 14 (1) (2014) 210–218.
[15] M.J. Baker, et al., Using Fourier transform IR spectroscopy to analyze biological materials, Nat. Protoc. 9 (8) (2014) 1771–1791.
[16] M.R. Max Diem, C. Matthäus, M. Miljkovic, L. Miller, P. Lasch, Comparison of Fourier transform infrared (FTIR) spectra of individual cells acquired using synchrotron and conventional source, Infrared Phys. Technol. 45 (5–6) (2004) 331–338.
[17] K. Wehbe, et al., Investigation of blood vessels in glioblastoma at a micrometric scale: a comparative study by synchrotron and conventional micro-FTIR, Anal. Methods 5 (24) (2013) 6925–6932.
[18] C. Sandt, et al., Use of infrared microspectroscopy to elucidate a specific chemical signature associated with hypoxia levels found in glioblastoma, Analyst 141 (3) (2016) 870–883.
[19] K.W. Malek, K.R. Bambery, FTIR imaging of tissues: techniques and methods of analysis, in: Optical Spectroscopy and Computational Methods in Biology and Medicine, Springer Netherlands, 2014, pp. 419–473.
[20] D. Finlayson, C. Rinaldi, M.J. Baker, Is infrared spectroscopy ready for the clinic? Anal. Chem. 91 (19) (2019) 12117–12128.
[21] K.B. Beć, J. Grabska, C.W. Huck, Near-infrared spectroscopy in bio-applications, Molecules 25 (2020) 12.
[22] A. Heidari, Terahertz technology in the future of health and medical applications, in: S. Adibi (Ed.), Mobile HeAlth: A Technology RoAd MAp, Springer International Publishing, Cham, 2015, pp. 663–670.
[23] L.M. Miller, P. Dumas, Chemical imaging of biological tissue with synchrotron infrared light, Biochimica et Biophysica Acta (BBA) - Biomembranes 1758 (7) (2006) 846–857.
[24] M.O. Amin, E. Al-Hetlani, I.K. Lednev, Trends in vibrational spectroscopy of fingermarks for forensic purposes, TrAC Trends Anal. Chem. 143 (2021), 116341.
[25] D. Ami, et al., Embryonic stem cell differentiation studied by FT-IR spectroscopy, Biochim. Biophys. Acta 1783 (1) (2008) 98–106.
[26] P. Heraud, et al., Fourier transform infrared microspectroscopy identifies early lineage commitment in differentiating human embryonic stem cells, Stem Cell Res. 4 (2) (2010) 140–147.
[27] J. Cao, et al., Fourier transform infrared microspectroscopy reveals unique phenotypes for human embryonic and induced pluripotent stem cell lines and their progeny, J. Biophotonics 7 (10) (2014) 767–781.
[28] C. Sandt, et al., Identification of spectral modifications occurring during reprogramming of somatic cells, PLoS ONE 7 (4) (2012) e30743.
[29] C. Sandt, J. Frederick, P. Dumas, Profiling pluripotent stem cells and organelles using synchrotron radiation infrared microspectroscopy, J. Biophotonics 6 (1) (2013) 60–72.
[30] J. Demšar, A. Erjavec, Č. Gorup, T. Hocevar, M. Milutinovič, M. Martin Možina, M. Polajnar, M. Toplak, A. Starič, M. Štajdohar, L. Umek, L. Žagar, J. Žbontar, M. Žitnik, B. Zupan, Orange: data mining toolbox in Python, J. Mach. Learn. Res. 14 (2013) 2349–2353.
[31] M. Toplak, S. Read, C. Sandt, S.M. Rosendahl, L. Vaccari, J. Demšar, F. Borondics, Infrared orange: connecting hyperspectral data with machine learning, Synchrotron Radiat. News 30 (2017) 40–45.
[32] K. Loutherback, et al., Microfluidic approaches to synchrotron radiation-based Fourier transform infrared (SR-FTIR) spectral microscopy of living biosystems, Protein Pept. Lett. 23 (3) (2016) 273–282.
[33] A.S. Knaupp, et al., Transient and permanent reconfiguration of chromatin and transcription factor occupancy drive reprogramming, Cell Stem Cell 21 (6) (2017) 834–845 e6.
[34] D. Li, et al., Chromatin accessibility dynamics during cell fate reprogramming, EMBO Rep. 22 (2) (2021), e51644.
[35] F. Zhang, et al., Histone acetylation induced transformation of B-DNA to Z-DNA in cells probed through FT-IR spectroscopy, Anal. Chem. 88 (8) (2016) 4179–4182.
[36] L. Li, A. Puretzky, R. Riehn, H.D. Hallen, DNA methylation detection using resonance and nanobowtie-antenna-enhanced Raman spectroscopy, Biophys. J. 114 (11) (2018) 2498.
[37] G. Gioacchini, et al., A new approach to evaluate aging effects on human oocytes: Fourier transform infrared imaging spectroscopy study, Fertil. Steril. 101 (1) (2014) 120–127.
[38] K.Z. Liu, et al., Quantitative determination of apoptosis on leukemia cells by infrared spectroscopy, Apoptosis 6 (4) (2001) 269–278.
[39] F. Karimi, et al., Infrared microspectroscopy studies on the protective effect of curcumin coated gold nanoparticles against H2O2-induced oxidative stress in human neuroblastoma SK-N-SH cells, Analyst 146 (22) (2021) 6902–6916.
[40] R. Mendelsohn, C.R. Flach, D.J. Moore, Determination of molecular conformation and permeation in skin via IR spectroscopy, microscopy, and imaging, Biochim. Biophys. Acta 1758 (7) (2006) 923–933.
[41] M. Kreuzer, et al., Lipids status and copper in a single astrocyte of the rat model for amyotrophic lateral sclerosis: correlative synchrotron-based X-ray and infrared imaging, J. Biophotonics 13 (10) (2020), e202000069.
[42] L.H. van der Maaten, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (2008) 2579–2605.
[43] M. Pezolet, et al., Conformation of wheat gluten proteins. Comparison between functional and solution states as determined by infrared spectroscopy, FEBS Lett. 299 (3) (1992) 247–250.
[44] N. Wellner, P.S. Belton, A.S. Tatham, Fourier transform IR spectroscopic study of hydration-induced structure changes in the solid state of omega-gliadins, Biochem. J 319 (Pt 3) (1996) 741–747.

[45] R. Sarroukh, et al., Transformation of amyloid beta(1–40) oligomers into fibrils is characterized by a major change in secondary structure, Cell. Mol. Life Sci. 68 (8) (2011) 1429–1438.
[46] N.K. Bhatia, et al., Curcumin binds to the pre-fibrillar aggregates of Cu/Zn superoxide dismutase (SOD1) and alters its amyloidogenic pathway resulting in reduced cytotoxicity, Biochim. Biophys. Acta 1854 (5) (2015) 426–436.
[47] I.J., S.A.J., Peak fitting applied to Fourier transform infrared and Raman spectroscopic analysis of proteins, Appl. Sci. 10(17) (2020) 5918.
[48] E. Apostolou, M. Stadtfeld, Cellular trajectories and molecular mechanisms of iPSC reprogramming, Curr. Opin. Genet. Dev. 52 (2018) 77–85.
[49] D.P. Wolf, et al., Concise review: embryonic stem cells derived by somatic cell nuclear transfer: a horse in the race? Stem Cells 35 (1) (2017) 26–34.
[50] C. Gouveia, et al., Lessons learned from somatic cell nuclear transfer, Int. J. Mol. Sci. 21 (7) (2020).
[51] A.M. Kocabas, et al., The transcriptome of human oocytes, PNAS 103 (38) (2006) 14027–14032.
[52] S. Assou, et al., The human cumulus–oocyte complex gene-expression profile, Hum. Reprod. 21 (7) (2006) 1705–1719.
[53] S. Assou, et al., A gene expression signature shared by human mature oocytes and embryonic stem cells, BMC Genomics 10 (2009) 10.
[54] E. Giorgini, et al., Vibrational characterization of female gametes: a comparative study, Analyst 139 (20) (2014) 5049–5060.
[55] C.D. Folmes, et al., Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming, Cell Metab. 14 (2) (2011) 264–271.
[56] K. Ito, T. Suda, Metabolic requirements for the maintenance of self-renewing stem cells, Nat. Rev. Mol. Cell Biol. 15 (4) (2014) 243–256.
[57] J. Prieto, et al., MYC induces a hybrid energetics program early in cell reprogramming, Stem Cell Rep. 11 (6) (2018) 1479–1492.
[58] J. Prieto, et al., c-MYC triggers lipid remodelling during early somatic cell reprogramming to pluripotency, Stem Cell Rev. Rep. 17 (6) (2021) 2245–2261.
[59] J. Chen, Perspectives on somatic reprogramming: spotlighting epigenetic regulation and cellular heterogeneity, Curr. Opin. Genet. Dev. 64 (2020) 21–25.
[60] N. Beknazarov, S. Jin, M. Poptsova, Deep learning approach for predicting functional Z-DNA regions using omics data, Sci. Rep. 10 (1) (2020) 19134.
[61] R. Liu, et al., Regulation of CSF1 promoter by the SWI/SNF-like BAF complex, Cell 106 (3) (2001) 309–318.
[62] A. Maruyama, et al., Nrf2 activation is associated with Z-DNA formation in the human HO-1 promoter, Nucleic Acids Res. 41 (10) (2013) 5223–5234.
[63] S.H. Kim, et al., Unveiling the pathway to Z-DNA in the protein-induced B-Z transition, Nucleic Acids Res. 46 (8) (2018) 4129–4137.
[64] I. Germanguz, et al., ADAR1 is involved in the regulation of reprogramming human fibroblasts to induced pluripotent stem cells, Stem Cells Dev. 23 (5) (2014) 443–456.
[65] D.D. Ehnes, et al., Combinatorial metabolism drives the naive to primed pluripotent chromatin landscape, Exp. Cell Res. 389 (2) (2020), 111913.
[66] H. Sperber, et al., The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition, Nat. Cell Biol. 17 (12) (2015) 1523–1535.
[67] M. Komatsu, et al., NNMT activation can contribute to the development of fatty liver disease by modulating the NAD (+) metabolism, Sci. Rep. 8 (1) (2018) 8637.
[68] W. Xu, et al., Effect of nicotinamide N-methyltransferase on lipid accumulation in 3T3-L1 adipocytes, Bioengineered 13 (5) (2022) 12421–12434.
[69] K. Takahashi, et al., Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm, Nat. Commun. 5 (2014) 3678.
[70] D. Cacchiarelli, et al., Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency, Cell 162 (2) (2015) 412–424.
[71] Y. Wang, et al., Unique molecular events during reprogramming of human somatic cells to induced pluripotent stem cells (iPSCs) at naive state, Elife 7 (2018).
[72] X. Liu, et al., Reprogramming roadmap reveals route to human induced trophoblast stem cells, Nature 586 (7827) (2020) 101–107.
[73] T.W. Theunissen, R. Jaenisch, Mechanisms of gene regulation in human embryos and pluripotent stem cells, Development 144 (24) (2017) 4496–4509.
[74] S.N. Hassani, et al., Transition of inner cell mass to embryonic stem cells: mechanisms, facts, and hypotheses, Cell. Mol. Life Sci. 76 (5) (2019) 873–892.