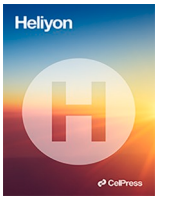
Heliyon 10 (2024) e39127

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Cofactor F_{420} tail length distribution in different environmental samples

Mathias Wunderer *, Rudolf Markt, Eva Maria Prem, Nico Peer, Andja Mullaymeri, Andreas O. Wagner

Universität Innsbruck, Department of Microbiology, Technikerstrasse 25d, 6020, Innsbruck, Austria

ARTICLE INFO

Keywords:
Environmental habitat
Microbial community
F_{420} producing microorganism

ABSTRACT

Cofactor F_{420} is an electron carrier playing a crucial role in a variety of microorganisms during redox reactions of the primary and secondary metabolism due to its low redox potential and thus arouses increasing interest. In this study, cofactor F_{420} glutamyl tail length spectra in various habitats like manure, compost, soil, and digester sludge samples and their respective microbial communities were investigated using high performance liquid chromatography and an amplicon sequencing approach A previous *in-silico* study was used to identify F_{420} producing microorganisms. The highest concentration of cofactor F_{420} could be achieved in the horse manure, digester sludge, and mixed manure samples, which was approximately 100-fold higher than in all the other samples. The high content of the cofactor in the samples with high O_2 availability pointed to the important role of the cofactor not only in redox reactions of anaerobic but also for aerobic microorganisms and indicated its ubiquitous character. The most abundant derivate was F_{420}-3 comprising the largest part of the cofactor derivates in seven out of ten samples. The high abundance of F_{420}-3 in samples with distinct properties (e.g. O_2 and H_2O availability) showed its important role in redox reactions of the primary and secondary metabolism among prokaryotes.

1. Introduction

Flavin-dependent enzymes mediate a wide range of redox reactions in the metabolism across all domains of life. Some organisms also synthesize the 5-deazaflavin structure $F_{420,}$ in which the N-5 atom is substituted through a C atom in the isoalloxazine ring system [1]. The cofactor F_{420} is structurally very similar to flavins but has distinctly different physiochemical properties. It is an obligate two-electron hydride carrier and has a very low standard redox potential and a blue-green intrinsic fluorescence with an absorption maximum at 420 nm [1]. Its properties are very similar to that of nicotinamides and are thus sometimes described as nicotinamides in flavin clothing [2]. With −340 mV, the cofactor has one of the lowest redox potentials of all known redox factors [1]. The redox potential can be reduced even further under certain physiological conditions: In CO_2-reducing methanogens, for example, the redox potential can decrease to – 380 mV when these organisms exhibit a 1:10 ratio of oxidized to reduced cofactor F_{420} [3].

Apart from its important role in redox-reactions of the primary metabolism, biochemical studies have revealed that the cofactor F_{420} can also be part of biosynthesis of various classes of secondary metabolites [4,5]. $F_{420}H_2$ dependent enzymes are involved in the reduction of tetracycline precursors [4,6], are further involved in the production of thiopeptides, lanthipeptides, 4-alkyl-L-proline

* Corresponding author. Department of Microbiology, Universität Innsbruck, Technikerstrasse 25d, 6020, Innsbruck, Austria.
E-mail address: Mathias.Wunderer@uibk.ac.at (M. Wunderer).

https://doi.org/10.1016/j.heliyon.2024.e39127
Received 3 April 2024; Received in revised form 7 October 2024; Accepted 8 October 2024
Available online 9 October 2024

derivatives, alkaloid guanipiperazine, alchivemycin A, and sulfonamide metabolites [4,5]. Besides, $F_{420}H_2$ depend enzymes are putatively involved in the production of 3-amino-5-hydroxybenzoic acid (AHBA)-containing metabolites, kasugamycin, coronafacoyls and phenazines [4]. The potential of $F_{420}H_2$ dependent enzymes is already used for some medical and industrial biotechnological applications, e. g. the biodegradation of environmental contaminants such as picrate [7]) or nitroaromatic explosives [8]. The antimicrobial prodrugs delamanid and pretomanid used to medicate tuberculosis patients is activated by an $F_{420}H_2$ dependent reductase of *Mycobacterium tuberculosis* directly within a patient's body [9,10]. However, F_{420} dependent enzymes have a considerable potential for bio catalysis, but according to Shah et al. F_{420} dependent enzymes are under explored and need to be better characterized to use their potential as a source for biocatalysts [11]. For a long time, the cofactor F_{420} could only be chemically identified in the phyla *Euryarchaeota* and *Actinobacteria* [12]. In 2017, Ney et al. could show that the five enzymes known for the biosynthesis of the cofactor are incorporated in the genome of at least 653 bacterial and 173 archaeal species and demonstrated that the cofactor is widely synthesized in soil ecosystems and is, thus far, more important in aerobic bacterial metabolism than previously thought [12].

The first discovery of the cofactor occurred in methanogenic *Euryarcheota* in the year 1972 [13] and after the structure and biosynthesis pathway was established it was assumed that it is universal in all F_{420} producing organisms. Novel investigations of F_{420} biosynthesis proclaim that its biosynthesis pathway differs in various organisms and is not universal [14]. The variations derive from the various substrate utilized to link Fo with its poly-glutamate tail to form the cofactor F_{420}. According to the current state of knowledge three different F_{420} biosynthesis pathway are known which could be found in *Actinobacteria*, *Euryarcheota*, and *Proteobacteria* [14]. All three pathways have in common that the Fo headgroup is formed from 5-Amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione and L-tyrosine. In methanogens this reaction is mediated from the enzymes CofG and CofH, while in mycobacteria the enzyme FbiC is responsible for this reaction [14].

The number of glutamates in the glutamyl chain differs among numerous organisms that express this cofactor. In mycobacteria, the cofactors mostly contain five to seven glutamate residues [15]. Cofactors with two or three glutamate residues seem to prevail in methanogens without cytochromes, whereas four to five glutamate residues are more common in methanogens with cytochromes [16]. Peck [17] and Wunderer et al. [18] also showed that the proportion of various F_{420} cofactor analogs did not remain stable in some methanogenic *Archaea* during batch cultivation and that even the prevailing carbon source can have an effect on the proportion of cofactor analogs. In Wunderer et al. [18], the proportions of various glutamate tail lengths in *Methanosarcina thermophila*, a cytochrome-containing methanogen, changed with increasing incubation time, whereas tail length-composition remained the same in *Methanoculleus thermophilus*, a methanogen without cytochromes. The length of the poly-glutamate tail modulates the binding affinity of F_{420} cofactors with oxidoreductases. Long-chain (five to eight glutamates) F_{420} cofactors bind with a six-to ten-fold higher affinity to these enzymes than short-chain (two glutamates) ones [19]. Additionally, the side chain significantly affects the kinetics of the enzyme, whereby long-chain F_{420} increases the substrate affinity (lower K_m) and reduces the turnover rate (lower k_{cat}) of these enzymes [19].

Besides the knowledge of the glutamyl tail length spectrum of certain species, orders, or even classes little is known about the glutamyl tail length spectra of complex microbial communities in various habitats. Despite its promising future applicability in industrial contexts, it is important to take one step back and evaluate the *in-situ* expression of glutamyl tail length spectra in complex, microbial habitats from a basic-scientific point of view. To our best knowledge, this is the first study dealing with glutamyl tail length distribution in natural communities. Consequently, the aim of the present study was i) to perform an inventory of F_{420} glutamyl tail length in various habitats, ii) to investigate the glutamyl tail length spectra with regard to an overall oxygen content estimate in these habitats, iii) to link physico-chemical parameters with F_{420} glutamyl tail length and iv) to compare these with microbiome data to obtain a first insight into F_{420} production by various microorganisms in various habitats.

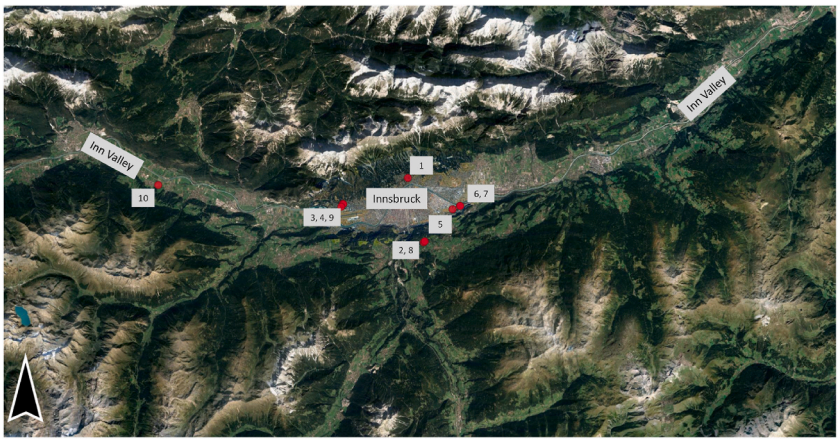


Fig. 1. Map of the Inn Valley near Innsbruck locating sampling spots (scale 1:200 000, created with QGIS 3.34.7). 1 = Mühlau (mixed forest), 2 = Lake Lans (coniferous forest), 3 = Innsbruck (meadow), 4 = Innsbruck (arable soil), 5 = Innsbruck (anaerobic digester), 6 and 7 = Innsbruck (compost), 8 = Lake Lans (swamp), 9 = Innsbruck (mixed manure), and 10 = Polling (horse manure).

2. Material and methods

2.1. Sampling and experimental design

Various environmental samples, including soil, sludge (anaerobic digestion plant), compost (composting plant), and manure were taken and analyzed in triplicates each in and around Innsbruck (Tyrol, Austria, Fig. 1, Table 1) to represent habitats with various oxygen content. The temperature was measured on-site before samples were taken. For the soil samples, a 20 × 20 cm square was cut out and the top layer (horizon A) was removed before sampling. Samples were immediately brought to the laboratory and stored at 4 °C. Soil samples were sieved to 4 mm. Due to the low water content manure samples had to be suspended in water (water: sample ratio 1: 3.5), mixed with an overhead shaker for 30 min, and vented every 10 min. Physico-chemical parameters (pH, dry weight, C:N ratio) were assessed and microbial community (16S rRNA amplicon sequencing) as well as F_{420} tail length analyses were conducted.

2.2. Physico-chemical analyses

For the pH measurement, the various samples were mixed with 0.01 M $CaCl_2$ (ratio 1: 2.5) and incubated for 3 h. Afterward, the pH was measured using an electrode (Metrohm, Swiss) according to previous protocols [20]. 125 mg heat dried (105 °C, overnight) sample material was used for the evaluation of the C: N ratio [20]. The analysis was done in triplicates on a CN Analyzer CN828-MC (LECO, USA) according to the manufacturer's protocols. Physio-chemical parameters are given in Table 2.

2.3. Cell disruption, solid-phase extraction, and F_{420} tail length variants analyses via high performance liquid chromatography (HPLC)

For cofactor F_{420} analyses, cells were disrupted through heat treatment (autoclaving) and the cofactor purified by solid-phase extraction: 5 g of the various sample materials were filled up to 10 mL with distilled water and 10 mL of cell disruption buffer was added (200 mM potassium dihydrogen phosphate, 50 mM ethylenediaminetetraacetic acid, and 1% (v/v) polysorbate 80, pH 7) and subjected to a temperature-pressure treatment applying 121 °C and 1.2 bar pressure for 30 min. Subsequently, the cell lysate was centrifuged for 5 min at 10 000×*g* and 8 mL of the supernatant was used for solid-phase extraction according to Markt et al. [21] The various tail length variants of the cofactor F_{420} were separated via reversed-phase ion-pair high performance liquid chromatography. The separation occurred on a prominence HPLC system (Shimadzu, Japan) equipped with a Gemini-NX C18 5μ 110 A 150 × 3 mm column (Phenomenex, Germany), an UV/VIS (@210 nm) and a fluorescence detector (absorption 420 nm, emission 475 nm). The column was heated at 40 °C and a flow rate of 1.0 mL/min was applied. The mobile phase was a gradient of solvent A (aqueous solution of 10 mM tetrabutylammonium hydroxide, 20 mM di-ammonium hydrogen phosphate, pH 7 adjusted with 85% phosphoric acid) and solvent B (100% acetonitrile) and varied as described in [18].

2.4. Amplicon sequencing

DNA was extracted in triplicates for each sampling location with the NucleoSpin® Soil Kit (MCHEREY&NAGEL, Germany) according to manufactures recommendations using 250–1000 mg fresh sample. Buffer SL1 was used as lysis buffer and elution occurred in 50 μL. The DNA was quantified fluorometrically using Quant-iT™ PicoGreen™ dsDNA Assay Kit (Invitrogen, USA) and an Anthos-Zenyth Multimode Detector [22]. Obtained DNA concentrations can be found in Table S2.

The small subunit (SSU) rRNA gene primers 515f and 806r [23] were used to target the V4 region, according to the Earth Microbiome Project [24]. The NGS library was prepared in-house according to well established protocols [25] with following modifications: The PCR mix contained 5 μL Q5® reaction buffer 5x (NEB, Germany), 0.25 μL Q5® Hot Start High-Fidelity DNA Polymerase (NEB, Germany), 10 mM dNTPs, 500 nM forward and reverse primer, 20% enhancer (NEB, Germany), 5 ng template and PCR-grade water to fill up to 25 μL for each reaction. Thermocycling conditions: 25 cycles of denaturation for 10 s, annealing at 57 °C (second PCR step: 56 °C) for 20 s, and elongation for 20 s. PCR products were purified using the Monarch PCR & DNA Cleanup Kit (NEB, Germany). The DNA was checked for contaminants via the NanoDrop 2000c™ system and the DNA quantity was measured via QuantiFluor® dsDNA Dye (Promega, USA). A final ready-to-load sample pool of 13 ng/μL, with a 260/280 absorbance ratio of 1.86 [25,26], was sent

Table 1
List of the various sample types with coordinates of the sample location.

Location	Sample	Longitude	Latitude	Sea level [m]
Mühlau, mixed forest (1)	soil	11°24'9.6"E	47°16'56.2"N	657
Lake Lans, coniferous forest (2)	soil	11°25'5.5"E	47°14'31.6"N	885
Innsbruck, meadow (3)	soil	11°20'20.9"E	47°16'3.6"N	613
Innsbruck, arable soil (4)	soil	11°20'21.2"E	47°16'2.3"N	613
Innsbruck, anaerobic digester (5)	sludge	11°26'47.6"E	47°15'50.7"N	653
Innsbruck, compost (6)	10-day compost	11°27'13.8"E	47°15'58.5"N	582
Innsbruck, compost (7)	6-month compost	11°27'14.1"E	47°15'58.0"N	582
Lake Lans, swamp (8)	soil	11°25'9.1"E	47°14'33.5"N	861
Innsbruck, mixed manure (9)	manure	11°20'14.5"E	47°15'56.7"N	579
Polling, horse manure (10)	manure	11°9'28.7"E	47°16'48.2N	613

Table 2
Physico-chemical properties of the various sample locations.

Sample location	Temperature [°C]	pH	Ratio [C:N]	Dry weight [g/g]
Arable Soil	20.80 ± 0.29	7.06 ± 0.11	45.70 ± 8.26	0.87 ± 0.00
Digester sludge	38.07 ± 0.13	7.40 ± 0.01	5.44 ± 0.04	0.04 ± 0.00
Swamp soil	16.27 ± 0.19	6.26 ± 0.17	14.32 ± 0.63	0.40 ± 0.03
10-days compost	53.47 ± 0.27	7.97 ± 0.02	17.84 ± 1.67	0.60 ± 0.01
6-months compost	30.07 ± 0.61	8.04 ± 0.06	14.75 ± 1.22	0.47 ± 0.02
Mixed manure	50.13 ± 0.82	8.22 ± 0.03	14.16 ± 0.35	0.27 ± 0.01
Mixed forest soil	16.43 ± 0.82	6.50 ± 0.08	16.81 ± 1.92	0.92 ± 0.02
Coniferous forest soil	14.77 ± 0.17	3.56 ± 0.19	22.71 ± 1.39	0.76 ± 0.04
Horse manure	30.43 ± 0.53	7.05 ± 0.01	12.31 ± 0.13	0.34 ± 0.03
Meadow soil	19.63 ± 0.21	7.15 ± 0.04	14.32 ± 0.27	0.75 ± 0.01

to Microsynth AG (Switzerland) for sequencing on a MiSeq™ System (Illumina®, USA) according to the company's protocols.

The raw reads were processed with mothur version 1.45.2 [27]. A contig file was created with the paired-end reads resulting in 6 396 503 sequences across all 32 samples (three parallels of the ten sample locations and two MOCK samples). Two different defined MOCK communities were used for the NGS validation. The ZymoBIOMICS™ Microbial Community standard (containing eight bacterial and two yeast microorganisms, further referred as Mock1) and the archaeon *Methanosarcina thermophila* DSM 1825 (DSMZ, German Collection of Microorganisms and Cell Cultures, further referred to as Mock2). Unique sequences were aligned with the SILVA V138.1 database [28]. Chimeric amplicons were removed with the VSEARCH algorithm and the sequences were classified with the k-nearest neighbor algorithm. The minimal sampling depth was 42 778 sequences per sample. All microorganisms of the MOCK community could be found after the amplicon sequencing.

2.5. Statistical and graphical analyses

All spreadsheets and the transformation (log 10) of the F_{420} data were done with Excel® (Microsoft®). The map of the sample locations was created with QGIS 3.34.7 Prizren (QGIS Development Team, Swiss). Graphs and analyses of variance (ANOVA) for F_{420} data were created with Statistica™ 13 (TIBCO® Software Inc., USA), whereas RStudio 2023.09.1 + 494 with R version R-4.4.1 (Posit PBC, USA) was used for further analyses and graphical illustrations: The heatmap, showing the most abundant F_{420} producing taxa (≥50reads) over all samples, was done with the packages *readxl* [29], *pheatmap* [4,30] and *extrafont* [31]. Hierarchical clustering was used for defining clusters. Stacked bar plots of metagenomic data were done with the packages as described earlier as well as the packages microbiome [32], *randomcoloR* [33] and *dplyr* [34], *tidyverse* [35], and *ggeasy* [36]. Nonmetric multidimensional scaling was done with the packages *vegan* version 2.6–6.1 [37], *readxl* and *extrafont*. F_{420} tail length data were *Hellinger* transformed prior ordination using *Bray* distance. Oxygen availability was dummy coded (1: low, 2: medium, 3: high) and environmental data were overlaid as vectors using the *vegan* command *envfit*. Stress value was 0.8 and the non-metric fit showed an R^2 of 0.99.

$$F_{420}\ average\ length = \frac{\sum_{k=F_{420}-2}^{F_{420}-10} F_{420}k*k}{\sum_{k=F_{420}-2}^{F_{420}-10} F_{420}k} \quad (1)$$

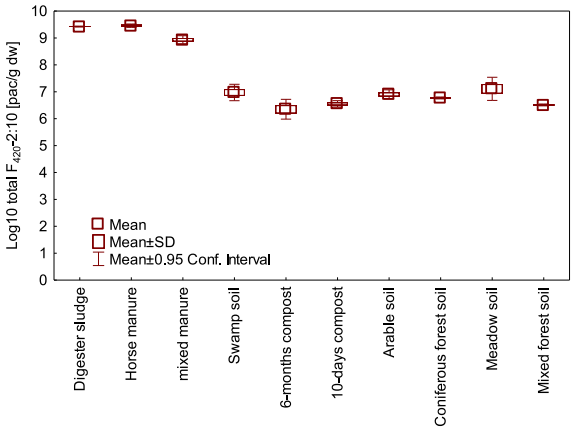


Fig. 2. Total concentration of F_{420} derivates (F_{420}-2 - F_{420}-10; F_{420}-2:10) for various sample locations, shown as log 10 peak area count/g dry weight [pac/g dw].

2.6. Identification of potentially F_{420} producing microorganisms

The results of the *in silico* analysis part from Ney et al. [12] were used to identify microorganisms potentially able to produce cofactor F_{420}. Ney et al. provided in their study a list of the 653 bacterial and 173 archaeal species (whose genomes were available on NCBI in 2017) which incorporated all five known enzymes for F_{420} biosynthesis in their genome. This list was compared with the results of the amplicon sequencing of the presented study to identify potential F_{420} producing microorganisms at genus level and to compare them with the glutamyl tail length spectra of the respective sample. The genome of many potentially F_{420} producing microorganisms is still unknown and could therefore not be included in this study. For this reason, the identified F_{420} producing genera likely represent an underestimation of the actual F_{420} producing microbes.

3. Results and discussion

3.1. Distribution of $F_{420\text{-}n}$ in different habitats

The total concentration of the F_{420} cofactor derivates of the various sample locations were logarithmized (log 10) and are shown in Fig. 2. Horse manure samples showed the highest amount of F_{420} cofactor derivates, followed by digester sludge samples and mixed manure samples. In all other samples, cofactor concentrations were rather similar (Fig. 2) and approximately two orders of magnitude lower.

The high concentration of F_{420} in samples with limited or very low O_2 availability (digester sludge, mixed manure, and horse manure) indicates the important role of the redox cofactor F_{420} during the absence of O_2 as a terminal electron acceptor. However, the swamp soil and the two compost types, also characterized by limited O_2 availability at least in specific niches, showed significantly lower concentrations of cofactor F_{420}. However, considerable amounts of the cofactor could also be isolated from samples with a better O_2 availability (arable soil, mixed forest soil, coniferous forest soil, and meadow soil). This confirms the findings from Ney et al. [12] that the cofactor F_{420} is far more important in aerobic habitats than previously thought.

The proportions of the various cofactor F_{420} derivates (F_{420}-2: F_{420}-10) of the total F_{420} derivates in the respective sample location are shown in Fig. 3 and the standard deviation of each cofactor derivate in Table S1. In the arable soil samples, all F_{420} cofactor derivates could be proven except the cofactor with two glutamate residues (F_{420}-3 to F_{420}-10). The cofactor with six glutamate residues accounted for the largest share reflecting 28.9% of total F_{420}, followed by F_{420}-5 at 25.5%, F_{420}-7 at 16.9%, F_{420}-8 at 10.2%, and F_{420}-4 at 9.6% (Fig. 3). The cofactors with three, nine, and ten glutamate residues contributed less than 5% of total F_{420} derivates. The glutamyl tail length spectrum in the meadow soil samples was very similar to the glutamyl tail length spectrum of the arable soil. The most abundant cofactor was F_{420}-5 at 26.7%, followed by F_{420}-6 at 25.3%, F_{420}-7 at 15.6%, F_{420}-4 at 11.3%, F_{420}-8 at 9.2%, and F_{420}-3 at 8.1%; F_{420}-2 was also not detected (Fig. 3). The cofactors with nine and ten glutamate residues again were least abundant and contributed less than 5%.

In digester sludge, all of the F_{420} cofactors (F_{420}-2: F_{420}-10) were found, whereby cofactors with two and three glutamate residues constituted the majority: Cofactor F_{420}-2 contributed 50.1% and F_{420}-3 44.9% to total F_{420} derivate pool. The cofactor with four glutamate residues reached 2.6% and the cofactors with five to ten glutamate residues stayed under 1% of total F_{420} (Fig. 3). In the swamp soil, all F_{420} cofactors could be evidenced except for F_{420}-7. The most abundant cofactor was F_{420}-3 at 32%, followed by F_{420}-5 at 24%, F_{420}-9 at 18.9%, F_{420}-4 at 13.3%, and F_{420}-6 at 5.5%. The cofactors with two, eight, and ten glutamate residues were the least abundant with less than 5%.

In the 10-days compost, all cofactors were found except the cofactor with two glutamate residues; predominant was cofactor F_{420}-3 reflecting more than half of total F_{420} (53.3%). The second most abundant cofactor was F_{420}-8 at 11.1%, followed by F_{420}-5 at 8.3%, F_{420}-7 at 8.2%, F_{420}-4 at 8%, and F_{420}-6 at 6.3% (Fig. 3). The least abundant cofactors were those with nine and ten glutamate residues

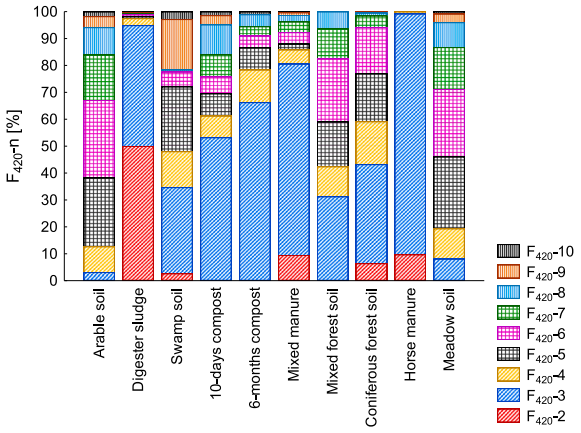


Fig. 3. Proportions [%] of the various F_{420} derivates of total F_{420} in the respective sample.

(less than 5%). The glutamyl tail length spectrum in the 6-months compost samples was very similar to the glutamyl tail length spectrum in the 10-days compost samples. In the 6-months compost, cofactors with three and four glutamate residues became more abundant, while at the same time, a decrease of cofactors with six to ten glutamate residues was observed. It could be shown that during the maturation of the compost (10-days to 6-months), the glutamyl tail length spectrum is slightly shifting towards F_{420} cofactors with shorter glutamyl tails. However, in the 6-months compost samples, the cofactors with two and nine glutamate residues could not be found, whereas cofactor with three glutamate residues was the most abundant at 66.3%, followed by F_{420}-4 at 12.1%, and F_{420}-5 at 8.3% (Fig. 3). Cofactors with six, seven, eight, and ten glutamate residues were reflecting less than 5% of total F_{420} derivates.

Besides digester sludge, mixed manure was the only sample where all cofactors (F_{420}-2: F_{420}-10) could be verified (Fig. 3). F_{420}-3 was predominant at 71%, followed by F_{420}-2 at 9.2%, and F_{420}-4 at 5.3%. The cofactors with five to ten glutamate residues were below 5% of total F_{420}. Horse manure had a similar glutamyl tail length spectrum when compared to mixed manure regarding short chained F_{420} cofactors. The most abundant cofactor was F_{420}-3 at 89.5%, F_{420}-2 s at 9.7% and F_{420}-4 still accounted for 0.8%. Mid- and long-chained F_{420} cofactors (five to ten glutamate residues) were completely missing in horse manure samples.

In the mixed forest soil, the cofactor with three glutamate residues was the most abundant at 31.3% of total F_{420}, followed by F_{420}-6 at 23.5%, and F_{420}-5 at 16.6%. F_{420}-4 and F_{420}-7 accounted for 11.1% and F_{420}-8 for 6.3% (Fig. 3). The cofactors with two, nine, and ten glutamate residues could not be detected in the mixed forest soil. The glutamyl tail length spectrum of cofactor F_{420} in the coniferous forest soil was very similar to that in the mixed forest soil, except for some differences regarding cofactors F_{420}-3 and F_{420}-8. In the coniferous forest soil, the predominant cofactor was F_{420}-3 at 36.8%, followed by F_{420}-5 at 17.8%, F_{420}-6 at 17.2%, F_{420}-4 at 15.9%, and F_{420}-2 at 6.5%. The cofactors with seven, eight, and nine glutamate residues contributed with less than 5% to the total F_{420} derivate pool, and the cofactor with ten glutamate residues could not be proven in the coniferous forest soil.

The average glutamyl tail length of the cofactor F_{420} was calculated for each sample and is shown in Fig. 4. The shortest average glutamyl tail length could be proven in the digester sludge samples with 2.62 ± 0.005 glutamate residues while the second shortest tail length could be verified in horse manure with 2.91 ± 0.008, followed by mixed manure with 3.49 ± 0.009, 6-months compost with 3.85 ± 0.471, coniferous forest soil with 4.22 ± 0.701, 10-days compost with 4.62 ± 0.391, mixed forest with 4.91 ± 0.293, swamp soil with 5.12 ± 0.134, meadow soil with 5.73 ± 0.104, and arable soil with 6.02 ± 0.056. However, according to the KRUSKAL-WALLIS ANOVA, the average F_{420} glutamyl tail length in the digester sludge samples was significantly shorter than in the arable soil and meadow soil samples. The average glutamyl tail length of the cofactor in the horse manure was also significantly shorter than in the arable soil, whereas no significant differences could be found between all other sample locations. Additionally, the average F_{420} glutamyl tail length for samples with various oxygen availability (low, medium, and high) was calculated (Fig. S1). In samples with low oxygen availability the shortest average glutamyl tail length could be proven with 3.01 ± 0.362 glutamate residues, which was significantly shorter than in the samples with medium and high oxygen availability. Samples with medium oxygen availability exhibited a mean glutamyl tail length of 4.53 ± 0.634 whereas samples with higher oxygen availability showed a mean glutamyl tail length of 5.22 ± 0.806.

The cofactor F_{420} has one of the lowest redox potentials of all currently known redox factors [1] and is therefore perfectly suited to take up electrons which accrue during the oxidation of various compounds (also various environmental pollutants). Ney et al. [19] could demonstrate that the length of the glutamyl tail is decisive for the binding affinity and turnover rate of the cofactor with oxidoreductases. Long-chain (five to eight glutamates) F_{420} binds with a six-to ten-fold higher affinity to oxidoreductases than short-chain F_{420} (two glutamates) [19]. Conversely, this suggests that a shorter glutamyl chain length of the cofactor F_{420} leads to a lower binding affinity and higher turnover rates and indicates the high potential for bioremediation of environments with shorter F_{420} glutamyl tail length as observed in the present study for e. g. digester sludge.

NMDS analysis (Fig. 5) demonstrates F_{420} tail length in the relationship to environmental data. While habitats with low oxygen concentrations (digester sludge, mixed manure, horse manure) tended to result in shorter tail length spectra, longer ones were found with higher oxygen availability, thus, confirming results from analysing mean F_{420} tail length (Fig. 4).

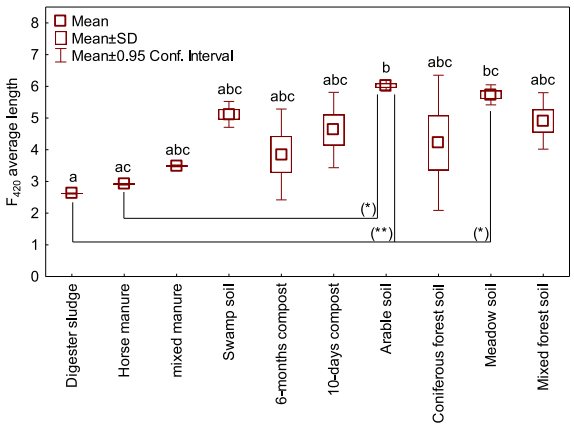


Fig. 4. Mean F_{420} glutamyl tail length of samples from various locations. Significant differences ($p < 0.05$) are indicated by various characters.

3.2. Microbial community composition

The relative abundance of the 50 most abundant genera across all samples is depicted in Fig. 6 to microbiologically characterize the samples briefly. Additionally, a relative abundance plot of 15 most abundant genera was made and is shown in Fig. S2. The abundance of microorganisms in the various soils was very similar except for the coniferous forest soil. In the arable soil, *Nitrososphaeraceae* unspecified genus was the most abundant taxa, followed by *Vicinamibacterales* genus as well as another unspecified genus of *Vicinamibacteraceae* (Fig. 6). In the mixed forest soil, *Vicinamibacterales* genus was the most abundant microorganism, followed by not specified genera of *Rokubacteriales* and *Vicinamibacteraceae* (Fig. 6), whereas meadow soil contained the same genera as the arable soil, with *Vicinamibacterales* genus being the most abundant one. Also, in the swamp soil the *Vicinamibacterales* genus was the most abundant one, followed by the genera *Rokubacteriales* and *Subgroup_17* which belongs to the class *Vicinamibacteria* (Fig. 6). In the coniferous forest soil, which showed the greatest differences in microbial abundance among the studied soils, the genus *Acidothermus,* which belongs to the class *Actinobacteria,* was most abundant. The second and third most abundant genera were the *Isosphaeraceae* and the *Gemmataceae* with no cultivated representatives (Fig. 6).

In digester sludge samples, a genus from the *Anaerolineaceae* was the most abundant one and accounted for almost 50% of the sequences, followed by *Methanosaeta* spp. and *Bacteriodetes_vadinHA17* (Fig. 6). In the 10-days compost samples, the genus *SBR 1031,* which belongs to the class *Anaerolineae,* was the predominant genus, followed by *Anaerolineaceae* genus which was the most abundant genus in the digester sludge samples. The third most abundant genus in the 10-days compost was *OM190* which belongs to the phylum *Planctomycetota* (Fig. 6). The 6-months compost samples were very similar to the 10-days compost samples. The predominant genus was again *SBR1031*, followed by *Anaerolineaceae genus* and *Vicinamibacterales* genus.

The most abundant genus in horse manure samples was *Lysinibacillus* belonging to the order *Bacillales* and second most abundant was the genus *WCHB1-41* which belongs to the class *Kiritimatiellae*, followed by *Akkermansia* spp. belonging to the order *Verrucomicrobiales* (Fig. 6). By contrast, in the mixed manure the most abundant genus was the *Hydrogenispora* which accounted for approximately 50% of the sequences. The second most abundant genus was *Truepera,* followed from the genus *Lysinibacillus* (Fig. 6).

This short description of the dominant microbial genera in the various samples serves as an overview and general characterization of the samples but cannot inform on microbial F_{420} producing microorganisms. In a next step, microbiome data were used to compare microorganisms found in this study with known F_{420} producers that were shown to be capable of synthesizing F_{420}-n genetically [12].

3.3. F_{420} producing community

The results of the *in silico* analysis of Ney et al. [12] were used to identify microorganisms capable of producing cofactor F_{420} in various environmental samples. Fig. 7 is linking the abundance of F_{420} producing microorganisms (showing an abundance of more than 50) with the origin of sample. Regarding the F_{420} producing microbial community soil samples clearly separated from those of manure, compost and sludge habitats (Fig. 7). While in soil abundant bacterial genera were mainly responsible for the composition of the F_{420}-n pool, in manure and sludge, but also in compost, methanogens accounted for F_{420} production. In the following, the most important producers of F_{420} in the various habitats are pinpointed.

The bacterial genus *Acidothermus* was the most abundant F_{420} producer in the coniferous forest soil, followed by the genera *Bradyrhizobium*, *Conexibacter*, and *Mycobacterium* (Fig. 7). Nothing is known about the glutamyl tail length spectra of the genus *Acidothermus,* but it belongs to the class of *Actinobacteria* and according to Daniels et al. [38] they should produce only small amount of the cofactor F_{420} compared to methanogenic Archaea. *Bradyrhizobium* belongs to the class of *Alphaproteobacteria* and the genus

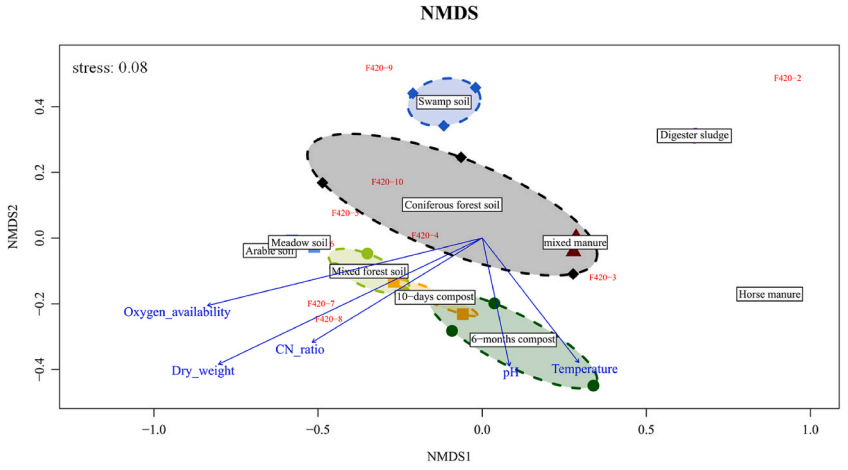


Fig. 5. Nonmetric multidimensional scaling (NMDS) ordination of F_{420} tail length data of all investigated habitats. Blue vectors show environmental data and F_{420} cofactor derivates are plotted in red. Each point represents a sample and dashed circles enclose all points of the respective habitat. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

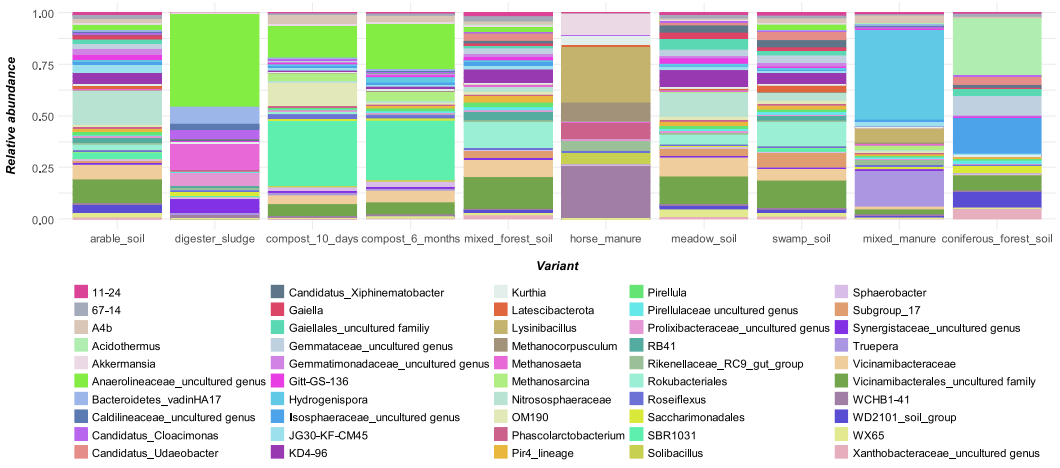


Fig. 6. Relative abundance plot of the 50 most anundant genera across all samples.

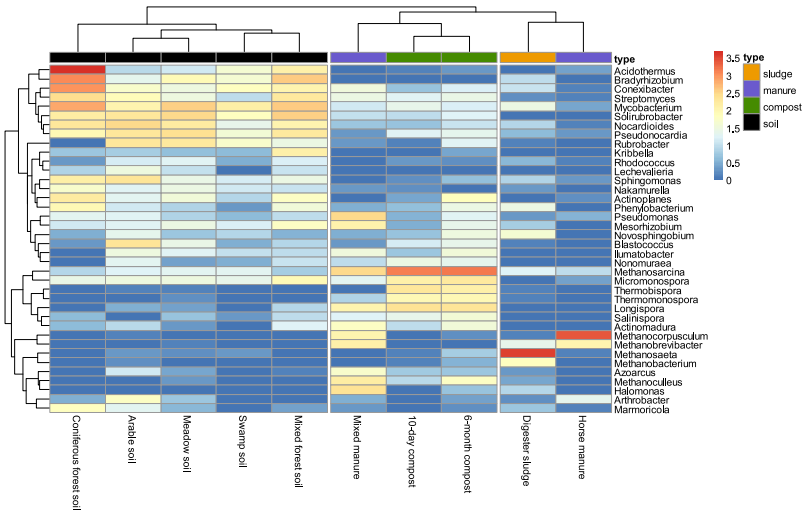


Fig. 7. Heatmap (log 10) of F_{420} producing genera according to in-silico analysis of Ney et al. [12]. Genera with an abundance of n ≥ 50 across all samples were included.

Conexibacter to *Thermoleophilia*; however, details on their glutamyl tail length spectra are missing so far. Members of the genus *Mycobacterium* should tend to attach mainly five to seven glutamate residues to the cofactor F_{420} [15]. Nevertheless, the cofactor F_{420}-3 was the most abundant cofactor derivate in the coniferous forest soil, followed from the cofactors with five and six glutamate residues (Table 3). In the arable soil, *Nocardioides* was most abundant genus, followed by the bacterial genera *Solirubrobacter, Pseudonocardia, Blastococcus, Sphingomonas*, and *Rubrobacter* (Fig. 7). The genera *Nocardioides, Pseudonocardia, Blastococcus* are aerobic members of the class Actinobacteria and should only produce small amounts of the cofactor F_{420} [38]. Moreover, the genus *Solirubrobacter* belongs to the class *Thermoleophilia*, *Sphingomonas* to *Alphaproteobacteria*, and *Rubrobacter* to *Rubrobacteria*. So far, their glutamyl tail length spectra have not been investigated. In the arable soil, F_{420} with six and five glutamate residues were predominant in the present study (Table 3). In the meadow soil, the genera *Mycobacterium*, *Solirubrobacter, Pseudonocardia, Rubrobacter, and Nocardioides* were predominant (Fig. 7). The genus *Mycobacterium*, a member of the class *Actinobacteria*, is a known F_{420} producer and was described to attach mainly five to seven glutamates to the cofactor according to Bair et al. [15]. *Solirubrobacter* is a genus of the class *Thermoleophilia, Pseudonocardia and Nocardioides* belong to *Actinobacteria* and the genus *Rubrobacter* to the class *Rubrobacteria*. Their F_{420} glutamyl tail length spectra have not been investigated so far. In meadow soil samples, F_{420} with five and six glutamate residues were dominating which is consistent with the high abundance of the genus *Mycobacterium* in these samples (Table 3). In the mixed forest, the genera *Mycobacterium, Solirubrobacter, and Bradyrhizobium* were predominant (Fig. 7). The genus *Mycobacterium* belongs to the class Actinobacteria and was described to mainly attach five to seven glutamate residues to the cofactor F_{420} [15]. The genus *Solirubrobacter* belongs to the class *Thermoleophilia* and *Bradyrhizobium* is a member of the order *Rhizobiales* which belongs to the class *Alphaproteobacteria*. So far, nothing is known about the glutamyl tail length of these two genera. However, in the present study the cofactors

Table 3
Summary of the dominant F_{420} producing genera and F_{420} cofactors of the respective sampling area.

Sample location	Dominant F_{420} producing genera	Dominant F_{420}-n [%]
Coniferous forest soil	*Acidothermus, Bradyrhizobium, Conexibacter, Mycobacterium*	F_{420}-3 [36.8] F_{420}-5 [17.8] F_{420}-6 [17.2]
Arable soil	*Nocardioides, Solirubrobacter, Pseudonocardia, Blastococcus, Sphingomonas, Rubrobacter*	F_{420}-6 [28.9] F_{420}-5 [25.5] F_{420}-7 [16.9]
Meadow soil	*Mycobacterium, Solirubrobacter, Pseudonocardia*	F_{420}-5 [26.7] F_{420}-6 [25.3] F_{420}-7 [15.6]
Swamp soil	*Mycobacterium, Conexibacter, Solirubrobacter, Rubrobacter*	F_{420}-3 [32.0] F_{420}-5 [24.0] F_{420}-9 [18.9]
Mixed forest soil	*Mycobacterium, Solirubrobacter, Bradyrhizobium*	F_{420}-3 [31.3] F_{420}-6 [23.5] F_{420}-5 [16.6]
Mixed manure	*Pseudomonas, Methanosarcina, Halomonas, Methanoculleus*	F_{420}-3 [71.0] F_{420}-2 [9.2] F_{420}-4 [5.3]
10 days-compost	*Methanosarcina, Longispora, Thermobispora*	F_{420}-3 [53.3] F_{420}-8 [11.1] F_{420}-5 [8.3]
6-months compost	*Methanosarcina, Longispora, Micromonospora*	F_{420}-3 [66.3] F_{420}-4 [12.1] F_{420}-5 [8.3]
Digester sludge	*Methanosaeta, Methanobacterium, Novosphingobium*	F_{420}-2 [50.1] F_{420}-3 [44.9] F_{420}-4 [2.6]
Horse manure	*Methanocorpusculum, Methanobrevibacter, Arthrobacter*	F_{420}-3 [89.5] F_{420}-2 [9.7] F_{420}-4 [0.8]

with three, six, and five glutamate residues were the most abundant in the mixed forest soil which is consistent with the high abundance of members of the genus *Mycobacterium* (Table 3). In the swamp soil no clear dominant genus could be found (Fig. 7). The most abundant genera in these samples were *Mycobacterium, Conexibacter, Solirubrobacter,* and *Rubrobacter* (Fig. 7). As already mentioned in the genus *Mycobacterium* the cofactors with five to seven glutamate residues should be dominant [15]. Whereas, the genera *Conexibacter and Solirubrobacter* are members of the class *Thermoleophilia* and the genus *Rubrobacter* belongs to the *Rubrobacteria*. These three genera have in common that the glutamyl tail spectra of their cofactor F_{420} derivates have not been investigated so far. The dominant cofactor derivates in the swamp soil were the F_{420}-3, F_{420}-5, and F_{420}-9. The high abundance of the F_{420}-5 in these samples concurs with the high abundance of the genus *Mycobacterium* known for predominant F_{420}-5 to F_{420}-7 production (Table 3).

In the various soil samples, similar F_{420} genera popped up e.g. like *Mycobacterium*, *Solirubrobacter*, *Pseudonocardia,* or *Rubrobacter*. The appearance of mainly aerobic mircoorganisms in coniferous and mixed coniferous forest-, arable-, meadow-, and swamp soil samples being responsible for F_{420} production proofs the *in-silico* analysis of Ney et al. [12] stating that the ability of F_{420} production is more wide-spread in bacterial genomes also in aerobic habitats than previously assumed for a long time. Regarding F_{420} production anaerobic microorganisms colonizing anaerobic soil niche in the commonly as aerobic considered habitat did not seem to contribute significantly to the F_{420} pool when looking at 16S rRNA gene-based abundance data.

Another picture was obtained for habitats that are considered as (mainly) anaerobic. In the mixed manure samples the genera *Pseudomonas, Methanosarcina, Halomonas,* and *Methanoculleus* were the predominant F_{420} producing microorganisms (Fig. 7). *Pseudomonas* ssp. *and Halomonas* ssp. belong to the class of *Gammaproteobacteria* and to our best knowledge nothing is known about the glutamyl tail length of the cofactor F_{420} in members of these genera. In contrast, according to Gorris and van der Drift [16] and Wunderer et al. [18] members of the archaeal genus *Methanosarcina* should tend to attach mainly four to five glutamate residues to the cofactor, whereas in *Methanoculleus* ssp. cofactors with three and four glutamate residues should be predominant [18]. The cofactor with three glutamate residues was clearly the most abundant cofactor in these samples, reflecting 71% of all F_{420} cofactor derivates (Table 3). This could indicate a major influence of hydrogenotrophic methanogens (*Methanoculleus* sp.) towards acetoclastic methanogens (*Mehtanosarcina* sp.) on the glutamyl tail length spectra in these samples. Wunderer et al. [18] could already show that the cofactor F_{420} content in *Methanoculleus thermophilus* was 100-fold higher than in *Methanocarcina thermophila*. Regarding F_{420} producing microorganisms 10-days and 6-months compost samples were very similar. The archaeal genus *Methanosarcina* was the predominant genus in both composts, followed by the bacterial genus *Longispora*. The third most abundant F_{420} producer in the 10-days and 6-months compost was *Thermobispora* spp. and *Micromonospora* spp., respectively (Fig. 7). The genus *Methanosarcina* can use a broader spectrum of carbon sources and can perform acetoclastic, methylotrophic, or even hydrogenotrophic methanogenesis [39]. However, members of the genus *Methanosarcina* are specialized in methanogenesis at high acetate concentrations [40], and are therefore indicators for high acetate concentrations. According to Gorris and van der Drift [16] and Wunderer et al. [18], F_{420} with four and five glutamate residues are predominant in *Methanosarcina* spp. Regarding the genera *Longispora, Thermibospora,* and *Micromonospora*, to

our best knowledge, nothing is known about their glutamyl tail length spectra, however, as all three belong to the class *Actinobacteria* these organisms should only have contributed small amounts of the cofactor F_{420} to the overall F_{420} pool [38]. Nevertheless, in both compost types F_{420}-3 was the dominant F_{420} cofactor, reflecting more than 50% of total F_{420} (Table 3). In digester sludge, the archaeal genus *Methanosaeta*, a strictly acetoclastic methanogenic group [41], was clearly the most abundant F_{420} producing microorganism based on 16S rRNA gene analysis, followed by archaeal *Methanobacterium* and bacterial *Novosphingobium* (Fig. 7). According to Gorris and van der Drift [16] and Wunderer et al. [18] methanogens with cytochromes mainly tend to attach four to five glutamates to the cofactor F_{420}. Wunderer et al. [18] could also show that in *Methanosarcina thermophila*, a facultative acetoclastic methanogen, the cofactors with four and five glutamate residues dominated regardless of the used substrate. The genus *Methanobacterium* is described as hydrogenotrophic and can use CO_2/H_2 or formate as substrate [42], and according Wunderer et al. [18], hydrogenotrophic methanogens mainly produce F_{420} with two to four glutamate residues. The cofactor with two and three glutamate residues constituted the largest fraction of the F_{420} cofactor spectrum (95%) in digester sludge samples (Table 3), indicating that hydrogenotrophic methanogens like *Methanobacterium* sp. and *Methanobrevibacter* spp. (Fig. 7) had the greatest impact on the glutamyl tail length spectrum in the digester sludge even though the genus *Methanosaeta*, an obligate acetoclastic methanogen turned out to be more abundant in these samples. However, Wunderer et al. [18] showed that the cofactor F_{420} concentration during acetoclastic methanogenesis was 100-fold lower than during hydrogenotrophic methanogenesis possibly explain the high abundance of *Methanosaeta* spp. and concurrent high concentrations of F_{420}-2:3. During hydrogenotrophic methanogenesis, cofactor F_{420} is directly involved as a redox cofactor for the methylene-H_4MPT-dehydrogenase and reductase reaction in the methanogenesis process, while it is not directly involved in acetoclastic methanogenesis and therein only required for other metabolic purposes, like the reduction of NADP or the detoxification of O_2 [1]. These preliminary results show that the DNA-based, relative abundance of microbial taxa do not necessarily reflect physiological properties and activities [43]. This seems especially true when comparing DNA approaches with F_{420} glutamyl tail length profiles which have not been considered so far in natural habitats.

The genus *Methanocorpusculum* was predominant in horse manure, which otherwise showed the highest similarity with digester sludge samples regarding F_{420} producing microorganisms. The second and third most abundant F_{420} producer based on the findings of Ney et al. [12] were *Methanobrevibacter* spp. and *Arthrobacter* spp., respectively (Fig. 7). The members of the genus *Methanocorpusculum* reduce CO_2 to methane with different electron donators depending on the respective species [44]. Similarly, the genus *Methanobrevibacter* uses the CO_2 reducing pathway and therefore tends to mainly attach two to four glutamate residuals [1,18] to the cofactor F_{420}. Nothing is known about the glutamyl tail length spectrum of the cofactor F_{420} in the genus *Arthrobacter*. However, this genus belongs to the class *Actinobacteria* and Daniels et al. [38] showed that several members of this class contained F_{420} although to a less extent (at maximum 1% of the total amount of the cofactor that can be synthetized by methanogens). This indicates that the influence of these organisms on the glutamyl tail length spectrum should be rather low. In the horse manure samples, only F_{420} cofactors with two, three, and four glutamates could be proven (Table 3), suggesting that mainly hydrogenotrophic methanogens were responsible for the determination of the glutamyl chain length spectrum in these samples [16,18]. This is in accordance with the obtained metagenomic data (Fig. 6).

4. Conclusion

The highest concentration of all F_{420} cofactor derivates could be isolated from digester sludge, mixed manure, and horse manure going along with the predominantly anoxic characteristics of these habitats and indicating the important role of the cofactor for redox reactions in the absence of O_2 as a terminal electron acceptor. The cofactor content in all other samples was significantly lower. The swamp soil and the two compost samples had at least niches with limited O_2 availability but showed about two orders of magnitude lower cofactor contents than the strictly anaerobic habitats. A possible reason might be the higher contribution of hydrogenotrophic methanogens in digester sludge and horse manure samples, whereas, with regard to 16S rRNA gene sequencing data, acetoclastic methanogenesis (*Methanosarcinales*) was more abundant in compost samples. Hydrogenotrophic methanogenic cells need substantially more F_{420} than those of acetoclastic methanogens as it is directly involved in hydrogenotrophic methanogenesis and thus lower F_{420} concentrations were found in those samples. By contrast, the high cofactor content in the various aerobic soils confirms that F_{420} is far more important in aerobic habitats than previously assumed and proved its ubiquitous abundance. The high concentration of the F_{420}-3 in samples with distinct properties like fluctuating O_2 and H_20 availability points out its important role as electron carrier during redox reactions of the primary and secondary metabolism among prokaryotes.

These results give a preliminary overview of F_{420} tail length spectra in various natural habitats with complex microbial interaction and highlight variations between data derived from metagenomic analyses and physiological parameters. However, further investigations are needed to enlighten the role of F_{420} producing Archaea and Bacteria in various, natural habitats.

CRediT authorship contribution statement

Mathias Wunderer: Writing – review & editing, Writing – original draft, Visualization, Formal analysis. **Rudolf Markt:** Resources, Methodology, Investigation, Data curation, Conceptualization. **Eva Maria Prem:** Writing – review & editing, Visualization. **Nico Peer:** Data curation. **Andja Mullaymeri:** Data curation. **Andreas O. Wagner:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Funding acquisition.

Data availability statement

The raw reads for the 16S amplicon sequencing are available on NCBI (National Center for Biotechnology Information) under the BioProject ID: PRJNA997793. Non-sequencing data will be made available on request. These data were not deposited into a publicly available repository.

Submission declaration

The present work has not been published previously, is not under consideration for publication elsewhere, and its publication is approved by all authors.

Ethics declarations

This publication does not deal with human embryos and fetuses, vulnerable individuals or groups, patients, invasive techniques, personal data or sensitive personal data or dual use issues. According to the Ethics Self-Assessment of the H2020 Program, there are no ethical concerns connected to this publication.

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.

Acknowledgement

The study was supported by the Tyrolean Science Fund project AP718024 and by Publikationsfonds der Universität Innsbruck.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e39127.

References

[1] C. Greening, F.H. Ahmed, A.E. Mohamed, B.M. Lee, G. Pandey, A.C. Warden, C. Scott, J.G. Oakeshott, M.C. Taylor, C.J. Jackson, Physiology, biochemistry, and applications of F420- and fo-dependent redox reactions, Microbiology and molecular biology reviews MMBR 80 (2) (2016) 451–493, https://doi.org/10.1128/MMBR.00070-15.
[2] C. Walsh, Naturally occurring 5-deazaflavin coenzymes: biological redox roles, Acc. Chem. Res. 19 (7) (1986) 216–221, https://doi.org/10.1021/ar00127a004.
[3] L.M.I. de Poorter, W.J. Geerts, J.T. Keltjens, Hydrogen concentrations in methane-forming cells probed by the ratios of reduced and oxidized coenzyme F420, Microbiology (Reading, England) 151 (Pt 5) (2005) 1697–1705, https://doi.org/10.1099/mic.0.27679-0.
[4] G. Bashiri, Cofactor F420, an emerging redox power in biosynthesis of secondary metabolites, Biochem. Soc. Trans. 50 (1) (2022) 253–267, https://doi.org/10.1042/BST20211286.
[5] G. Bashiri, F420-dependent transformations in biosynthesis of secondary metabolites, Curr. Opin. Chem. Biol. 80 (2024) 102468, https://doi.org/10.1016/j.cbpa.2024.102468.
[6] P. Wang, G. Bashiri, X. Gao, M.R. Sawaya, Y. Tang, Uncovering the enzymes that catalyze the final steps in oxytetracycline biosynthesis, J. Am. Chem. Soc. 135 (19) (2013) 7138–7141, https://doi.org/10.1021/ja403516u.
[7] T. Jirapanjawat, B. Ney, M.C. Taylor, A.C. Warden, S. Afroze, R.J. Russell, B.M. Lee, C.J. Jackson, J.G. Oakeshott, G. Pandey, C. Greening, The redox cofactor F420 protects mycobacteria from diverse antimicrobial compounds and mediates a reductive detoxification system, Appl. Environ. Microbiol. 82 (23) (2016) 6810–6818, https://doi.org/10.1128/AEM.02500-16.
[8] S. Ebert, P.-G. Rieger, P.-J. Knackmuss, Function of coenzyme F420 in aerobic catabolism of 2,4,6-trinitrophenol and 2,4-dinitrophenol by Nocardioides simplex FJ2-1A, J. Bacteriol. 181 (9) (1999) 2669–2674, https://doi.org/10.1128/jb.181.9.2669-2674.1999.
[9] S.E. Cellitti, J. Shaffer, D.H. Jones, T. Mukherjee, M. Gurumurthy, B. Bursulaya, H.I. Boshoff, I. Choi, A. Nayyar, Y.S. Lee, J. Cherian, P. Niyomrattanakit, T. Dick, U.H. Manjunatha, C.E. Barry, G. Spraggon, B.H. Geierstanger, Structure of Ddn, the deazaflavin-dependent nitroreductase from Mycobacterium tuberculosis involved in bioreductive activation of PA-824, Structure (London, England 1993) 20 (1) (2012) 101–112, https://doi.org/10.1016/j.str.2011.11.001.
[10] Chauhan Shive Murat Singh, Chemistry and biology of coenzyme F420 in tuberculosis treatment, Chemical Biology LETTERS 11 (3) (2024) 666–673, https://doi.org/10.62110/sciencein.cbl.2024.v11.666.
[11] M.V. Shah, J. Antoney, S.W. Kang, A.C. Warden, C.J. Hartley, H. Nazem-Bokaee, C.J. Jackson, C. Scott, Cofactor F420-dependent enzymes: an under-explored resource for asymmetric redox biocatalysis, Catalysts 9 (10) (2019) 868, https://doi.org/10.3390/catal9100868.
[12] B. Ney, F.H. Ahmed, C.R. Carere, A. Biswas, A.C. Warden, S.E. Morales, G. Pandey, S.J. Watt, J.G. Oakeshott, M.C. Taylor, M.B. Stott, C.J. Jackson, C. Greening, The methanogenic redox cofactor F420 is widely synthesized by aerobic soil bacteria, The ISME journal 11 (1) (2017) 125–137, https://doi.org/10.1038/ismej.2016.100.
[13] P. Cheeseman, A. Toms-Wood, R.S. Wolfe, Isolation and properties of a fluorescent compound, Factor420, from Methanobacterium strain M.o.H, J. Bacteriol. (1972) 527–531.
[14] R. Grinter, C. Greening, Cofactor F420: an expanded view of its distribution, biosynthesis and roles in bacteria and archaea, FEMS Microbiol. Rev. 45 (5) (2021), https://doi.org/10.1093/femsre/fuab021.
[15] T.B. Bair, D.W. Isabelle, L. Daniels, Structures of coenzyme F(420) in Mycobacterium species, Arch. Microbiol. 176 (1–2) (2001) 37–43.
[16] L.G. Gorris, C. van der Drift, Cofactor contents of methanogenic bacteria reviewed, Biofactors 4 (3–4) (1994) 139–145.
[17] Peck, Changes in Concentrations of Coenzyme F420 Analogs during Batch Growth of Methanosarcina Barkeri and Methanosarcina Mazei, 1989.

[18] M. Wunderer, R. Markt, N. Lackner, A.O. Wagner, The glutamyl tail length of the cofactor F420 in the methanogenic Archaea Methanosarcina thermophila and Methanoculleus thermophilus, The Science of the total environment 809 (2022) 151112, https://doi.org/10.1016/j.scitotenv.2021.151112.
[19] B. Ney, C.R. Carere, R. Sparling, T. Jirapanjawat, M.B. Stott, C.J. Jackson, J.G. Oakeshott, A.C. Warden, C. Greening, Cofactor tail length modulates catalysis of bacterial F420-dependent oxidoreductases, Front. Microbiol. 8 (2017) 1902, https://doi.org/10.3389/fmicb.2017.01902.
[20] F. Schinner, R. Öhlinger, E. Kandeler, R. Margesin (Eds.), Methods in Soil Biology, Springer, Berlin, Heidelberg, 1996.
[21] R. Markt, M. Wunderer, E.M. Prem, M. Mutschlechner, N. Lackner, A.O. Wagner, Extraction of cofactor F420 for analysis of polyglutamate tail length from methanogenic pure Cultures and environmental samples, J. Vis. Exp. 176 (2021), https://doi.org/10.3791/62737.
[22] A.O. Wagner, N. Praeg, C. Reitschuler, P. Illmer, Effect of DNA extraction procedure, repeated extraction and ethidium monoazide (EMA)/propidium monoazide (PMA) treatment on overall DNA yield and impact on microbial fingerprints for bacteria, fungi and archaea in a reference soil, Applied soil ecology a section of Agriculture, Ecosystems & Environment 93 (2015) 56–64, https://doi.org/10.1016/j.apsoil.2015.04.005.
[23] A. Apprill, S. McNally, R. Parsons, L. Weber, Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton, Aquat. Microb. Ecol. 75 (2) (2015) 129–137, https://doi.org/10.3354/ame01753.
[24] J.A. Gilbert, J.K. Jansson, R. Knight, The Earth Microbiome project: successes and aspirations, BMC Biol. 12 (2014) 69, https://doi.org/10.1186/s12915-014-0069-1.
[25] E.M. Prem, R. Markt, N. Lackner, P. Illmer, A.O. Wagner, Microbial and phenyl acid dynamics during the start-up phase of anaerobic straw degradation in meso- and thermophilic batch reactors, Microorganisms 7 (12) (2019), https://doi.org/10.3390/microorganisms7120657.
[26] E.M. Prem, S.F. Duschl, A.O. Wagner, Effects of increasing phenyl acid concentrations on the AD process of a multiple-biogas-reactor system, Biomass Bioenergy 168 (2023) 106686, https://doi.org/10.1016/j.biombioe.2022.106686.
[27] P.D. Schloss, S.L. Westcott, T. Ryabin, J.R. Hall, M. Hartmann, E.B. Hollister, R.A. Lesniewski, B.B. Oakley, D.H. Parks, C.J. Robinson, J.W. Sahl, B. Stres, G. G. Thallinger, D.J. van Horn, C.F. Weber, Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities, Appl. Environ. Microbiol. 75 (23) (2009) 7537–7541, https://doi.org/10.1128/AEM.01541-09.
[28] C. Quast, E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, F.O. Glöckner, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools, Nucleic acids research 41 (Database issue) (2013) D590–D596, https://doi.org/10.1093/nar/gks1219.
[29] Wickham and Bryan, readxl: Read Excel Files, R package version 1.3. 1 (2019).
[30] R. Kholde, Pheatmap: Pretty Heatmaps, 2019. R package version 1.0. 12.
[31] W. Chang, Extrafont: Tools for Using Fonts, 2014. R package version 0.17, 2014.
[32] S.A. Shetty, L. Lahti, Microbiome data science, J Biosci 44 (5) (2019), https://doi.org/10.1007/s12038-019-9930-2.
[33] R. Ammar, randomcoloR: Generate Attractive Random Colors, 2019.
[34] H. Wickham, R. Francois, L. Henry, K. Müller, Dplyr: A Grammar of Data Manipulation, 2015.
[35] H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the tidyverse, JOSS 4 (43) (2019) 1686, https://doi.org/10.21105/joss.01686.
[36] J. Carroll, A. Schep, J. Sidi, Geasy: Easy Access to 'ggplot2'Commands, 2020.
[37] P. Dixon, VEGAN, a package of R functions for community ecology, J Vegetation Science 14 (6) (2003) 927–930, https://doi.org/10.1111/j.1654-1103.2003.tb02228.x.
[38] L. Daniels, N. Bakhiet, K. Harmon, Widespread distribution of a 5-deazaflavin cofactor in actinomyces and related bacteria, Syst. Appl. Microbiol. 6 (1) (1985) 12–17, https://doi.org/10.1016/S0723-2020(85)80004-7.
[39] N. Lackner, A. Hintersonnleitner, A.O. Wagner, P. Illmer, Hydrogenotrophic methanogenesis and autotrophic growth of Methanosarcina thermophila, Archaea 2018 (2018) 4712608, https://doi.org/10.1155/2018/4712608.
[40] W. Gujer, J.B. Zehnder, Conversion processes in anaerobic digestion, Wat.Sci.Tech. 15 (1983) 12 7–167. Copenhagen.
[41] Koji Mori, Iino Takao, Suzuki Ken-Ichiro, Yamaguchi Kaoru, Kamagata Yoichi, Aceticlastic and NaCl-requiring methanogen "Methanosaeta pelagica" sp. nov., isolated from marine tidal flat sediment, Appl. Environ. Microbiol. 78 (9) (2012) 3416–3423, https://doi.org/10.1128/AEM.07484-11.
[42] J. Benstead, D.B. Archer, D. Lloyd, Formate utilization by members of the genus Methanobacterium, Arch. Microbiol. 156 (1) (1991) 34–37, https://doi.org/10.1007/BF00418184.
[43] J. de Vrieze, L. Regueiro, R. Props, R. Vilchez-Vargas, R. Jáuregui, D.H. Pieper, J.M. Lema, M. Carballa, Presence does not imply activity: DNA and RNA patterns differ in response to salt perturbation in anaerobic digestion, Biotechnol. Biofuels 9 (2016) 244, https://doi.org/10.1186/s13068-016-0652-5.
[44] G. Zellner, E. Stackebrandt, P. Messner, B.J. Tindall, E. Conway de Macario, H. Kneifel, U.B. Sleytr, J. Winter, Methanocorpusculaceae fam. nov., represented by Methanocorpusculum parvum, Methanocorpusculum sinense spec. nov. and Methanocorpusculum bavaricum spec. nov, Arch. Microbiol. 151 (5) (1989) 381–390, https://doi.org/10.1007/BF00416595.