communications earth & environment

Article

A Nature Portfolio journal

https://doi.org/10.1038/s43247-025-02450-7

# Temperature-dependent spatial and temporal trends in archaeal lipid distributions

Check for updates

Siqi Zhao[1,2], Rui Bao[1,2] ✉, Long Zhou[1], Mingzhi Liu[1,2], Xiang Sun[3], Zhen Gao[3], Yang Zhou[4], Nan Wang[5], Yipeng Wang[1,2], Junhui Chen[6], Lei Xing[1] & Chuanlun Zhang[7]

Isoprenoid glycerol dialkyl glycerol tetraethers are key archaeal lipids widely applied in biogeochemistry. Compounds with zero cyclopentane rings and those with four cyclopentane and one cyclohexane ring dominate in marine sediments and are crucial for investigating archaeal evolution and adaptation. However, the mechanisms regulating their distribution patterns remain unclear. Here we use molecular dynamics simulations to examine their structural properties under increasing temperatures and find that temperature-dependent behaviors govern their global spatial distributions. Consistent distribution patterns across different geological epochs, revealed by thousands of geochemical records, suggest that Earth's temperature may have constrained long-term temporal trends of archaeal lipids. Reinforcement learning results further show that average atmospheric carbon dioxide levels since the Late Jurassic are closely associated with these two archaeal lipid distribution patterns, enabling predictions for more ancient sedimentary records, which can benefit for understanding the long-term adaptation strategy for the evolutionary and ecological significance of archaea.

Microbial lipid biomarkers have become increasingly important in enhancing our understanding of biophysical mechanisms and biogeochemical characteristics due to their sensitivities to environmental fluctuations[1–6] and abilities to encode information about the Earth's temperature history. The distribution of microbial lipids reveals microbial life evolutions and provides insights into their biosynthetic pathways[7,8]. Archaea are one of the most important microorganisms in the three domains of life. Archaeal membrane core lipids, known as isoprenoid glycerol dialkyl glycerol tetraethers (GDGTs), are widespread in marine environment[9–14]. Marine archaea, such as ammonia-oxidizing archaea, typically synthesize GDGTs with zero to four cyclopentane rings (GDGT-0 to GDGT-4) and including crenarchaeol and its isomer, serve as important lipid biomarkers in marine sediments[13–22].

The distributions of GDGTs (such as GDGT-1 to GDGT-3, crenarchaeol isomer) in marine systems were observed by Schouten et al.[13], who introduced the $TEX_{86}$ index to correlate increasing temperature with the structural cyclization numbers of GDGTs. So far, the growing amount of spatial GDGT fractional abundance data have been published and revealed the ambiguity in the confounding environmental factors and diverse archaea communities that co-influenced the distribution abundances of GDGTs[3,13–19]. For instance, GDGT spatial distribution abundances in equatorial and polar regions, appear to be governed by distinct mechanisms, as evidenced by the atypical $TEX_{86}$ signatures observed in these areas[16,17]. Thus, the spatial distribution patterns of GDGTs, particularly on a global scale, may reflect the intertwined impacts of archaea controlling GDGT production in response to environmental changes.

Furthermore, temporal distribution patterns of GDGTs, such as $TEX_{86}$, have previously been used to reconstruct the oceanic temperature, proposing a possibly regional forcing of $CO_2$ to surface water cooling in tropical area during Eocene[21]. The subsequent improved $TEX_{86}$ palaeothermometry based on temporal GDGT distribution has precisely recorded the sea-surface temperature evolution during the Cretaceous period[22].

[1]Frontiers Science Center for Deep Ocean Multispheres and Earth System, and Key Laboratory of Marine Chemistry Theory and Technology, Ministry of Education/Institute for Advanced Ocean Studies, Ocean University of China, Qingdao, China. [2]Laboratory for Marine Geology, Qingdao Marine Science and Technology Center, Qingdao, China. [3]School of Mathematical Sciences, Ocean University of China, Qingdao, China. [4]Guangzhou Marine Geological Survey, China Geological Survey, Ministry of Land and Resources, Guangzhou, China. [5]Frontiers Science Center for Deep Ocean Multispheres and Earth System, Key Lab of Submarine, Ocean University of China, Qingdao, China. [6]Marine Bioresource and Environment Research Center, Key Laboratory of Marine Eco-Environmental Science and Technology, First Institute of Oceanography, Ministry of Natural Resources, Qingdao, China. [7]Shenzhen Key Laboratory of Marine Archaea Geo-Omics, Southern University of Science and Technology, Shenzhen, China. ✉e-mail: baorui@ouc.edu.cn

Previous studies elucidated the temperature-driven evolution and ecology of marine ammonia-oxidizing archaea by examining the relationship between the distribution patterns of GDGT-2/GDGT-3 and $TEX_{86}$ in paleo-marine environments[12]. In addition, the Earth's temperature change may cause archaeal evolution. For example, previous studies observed the consistence between *Thaumarchaeota* in marine environment, the deepest lineages of ammonia-oxidizing archaea, and a thermophilic ancestor for ammonia-oxidizing archaea[9]. These observations strengthen the importance of temporal GDGT distribution trends due to their inseparable ties with long-term temperature change and archaea evolution. Thus, understanding the mechanisms behind GDGT temporal distribution patterns can provide valuable insights into the adaptive strategies of archaea under long-term and gradually changing temperature throughout the Earth's history.

The focus of research on GDGT distribution patterns has predominantly centered on the minor abundant GDGTs such as GDGT-1 to GDGT-3 and crenarchaeol isomers in the $TEX_{86}$ index, rather than the major abundant GDGT-0 and crenarchaeol[9,11,14–22]. The exclusion of GDGT-0 and crenarchaeol from $TEX_{86}$ calculations is due to a combination of factors, including their low environmental sensitivity, contributions from multiple biological sources, and the influence of various environmental parameters such as temperature, pH, and oxygen availability[14–22]. However, GDGT-0 and crenarchaeol are closely linked to archaeal evolution, playing an important role in reconstructing evolutionary adaptations over geological timescales. For example, the observed increase in GDGT-0 production in *Methanocaldococcus jannaschii* when was cultured between 47–75 °C, suggesting a potential link between GDGT-0 and archaeal thermal adaptation[10,11]. The presence of crenarchaeol as a distinct lipid biomarker of ammonia-oxidizing archaea, highlights its potential role in tracing the evolutionary history of this group[9]. Moreover, complex depositional settings, such as sediment maturity and redox conditions, can influence the degradation rates of GDGTs in sediments. The distribution patterns of minor GDGTs with relatively low concentrations of are easily influenced, especially in diverse sedimentary archives spanning geological time periods[23–26]. Nevertheless, GDGT-0 and crenarchaeol, which are prevalent in modern and ancient archives, are valuable for elucidating the mechanisms that govern the global spatial and temporal trends in GDGT distribution.

Research on GDGT distribution patterns has primarily focused on the structural characteristics of GDGTs, as changes in the configuration of GDGT molecules reflect the physiological regulatory mechanisms of membrane lipids in archaea in response to environmental factors[9–22,27–29]. While previous studies attempted to address this issue through mesocosm experiments and environmental surveys[9–30], these approaches have technical constraints and the absence of ancient samples[11,14]. Relying solely on laboratory or environmental experiments for understanding the controlling mechanism of archaeal lipid distributions may be a challenge to visually decipher the dynamic variations in structural characteristics.

Molecular dynamics (MD) simulations are valuable for understanding the dynamic changes in structural features, such as stability and interaction ability with the environmental change[31–34]. For instance, MD simulations have elucidated the role of cyclopentane rings in the structures of GDGTs, showing their contribution to membrane stabilization at a molecular level[35,36]. These simulations have also been instrumental in creating membrane models and investigating the physiological function of membrane-spanning lipids[37–41]. In this study, MD simulations were conducted to investigate the structural properties of GDGT-0 and crenarchaeol under computer-controlled varying temperatures. Relative abundance data of GDGT-0 and crenarchaeol were collected from surface sediments in the globally modern marine environment ($n = 723$). Furthermore, the temporal distribution trends of GDGT-0 and crenarchaeol were analyzed for each epoch from the Late Jurassic period to the present day (~200 Ma, $n = 3132$). The goal of this study is to elucidate the underlying mechanisms driving the spatial and temporal distribution trends of GDGT-0 and crenarchaeol, which would be used for predicting the distribution trends of archaeal lipid archaea from geological sedimentary archives.

## Results and discussion

### Temperature control of spatial distribution trends of GDGT-0 and crenarchaeol

The relative abundances of GDGT-0 and crenarchaeol are analyzed in globally distributed surface sediments ($n = 723$, see spatial GDGT distribution data) from marine environments, as depicted in Fig. 1. Our findings show that the distribution trend of GDGT-0 gradually decreases with increasing temperature (gray dots and green line in Fig. 1a), while the trend for crenarchaeol continuously increases (gray dots and green line in Fig. 1b). To investigate the mechanisms influencing the distribution trends of GDGT-0 and crenarchaeol, we assess the structural characteristics of these compounds through MD simulations at different temperatures. Our simulation results demonstrate the structural variation trends of GDGT-0 and crenarchaeol based on the structural total energy (STE), a parameter in MD simulations that reflects molecular structural stability[31,32]. The STE values of GDGT-0 increase with temperature, whereas those of crenarchaeol decrease (Fig. 1, Supplementary Table S1). The consistent patterns between the global distribution trends and our simulated structural energies of GDGT-0 and crenarchaeol under increasing temperature, respectively, are observed (Supplementary Fig. S1).

Our study then compares the similarity of distribution trends and our simulated results of GDGT-0 and crenarchaeol based on Euclidean Distance calculation, which has been applied for comparing the similarity of two sets of data (detailed in Method)[42,43]. The Euclidean Distance results show the goodness of the shape similarity of the distribution trends and STE variation patterns of GDGT-0 and crenarchaeol, 0.95 for GDGT-0 and 0.92 for crenarchaeol (thresholds from 0 to 1, with the maximum similarity of 1). This high similarity suggests that the globally spatial distribution trends of GDGT-0 and crenarchaeol in the modern marine sediments under the increasing temperature is largely controlled by their structural properties driven by temperature. We thus suggest that the structural properties of GDGT-0 and crenarchaeol may be crucial for unraveling the overarching mechanisms controlling the temporal distribution trends of GDGT-0 and crenarchaeol in geological records.

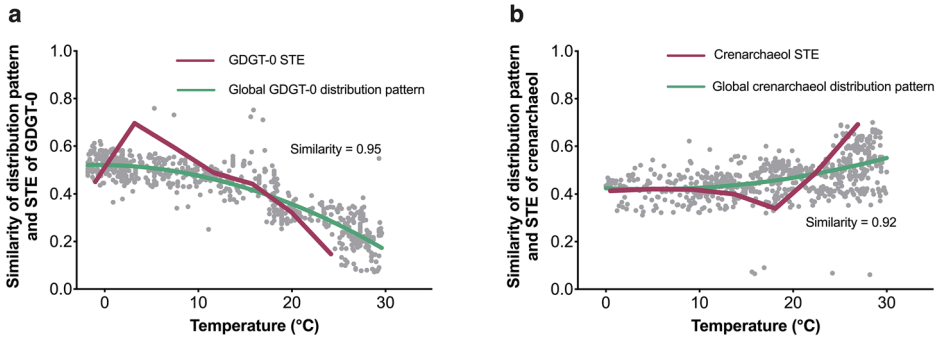



**Fig. 1 | Similarity between distribution trends and structural total energy values of GDGTs.** Distributions of GDGT-0 (**a**) and crenarchaeol (**b**) are shown as gray dots ($n = 723$), representing global relative abundances of GDGTs in marine surface sediments. Polynomial fittings of the GDGT distribution trends are shown as green solid lines, with coefficients of determination ($R^2$) of 0.78 for GDGT-0 and 0.23 for crenarchaeol. Structural total energy (STE, in kcal per mol) under increasing temperatures, calculated through molecular dynamics simulations, is shown as red solid lines for each compound.

### Temperature control of degradation ability of GDGT-0 and crenarchaeol

Temperature can influence the post-burial modification of archaeal lipids in the environment, leading to changes in the distribution of these lipids[23,24]. This can be attributed to the degradation of GDGTs in response to temperature fluctuations. Due to structural differences between GDGT-0 and crenarchaeol, the extent of degradation of these compounds may vary, impacting their distribution patterns, particularly over long geological time scales[23]. This degradation variability can affect the use of GDGT-0 and crenarchaeol distribution trends as environmental indicators over geological periods. Understanding the degradation capabilities of GDGT-0 and crenarchaeol is thus crucial for elucidating the mechanisms controlling their distribution trends in geological records.

We propose the molecule-specific surface area (MSSA, Fig. 2) of the equilibrium configurations of these archaeal lipids to evaluate variations in their degradation abilities of GDGT structures under a range of temperatures. The MSSA is a parameter that is calculated by MD simulations, which refers to the capacity of a compound communicating with ambient environments (details see Methods)[44]. The degrees of convolution in the equilibrium configurations of GDGTs represent the extent of the opportunities for communicating with the environment (Fig. 2, Supplementary Table S2). The stretched configuration with a larger MSSA value indicates a larger contact area with the external environment, and such a molecule structure change is more degradable in the environment and vice versa. We thus argue that the MSSA values may represent the degradation abilities of GDGTs.

In Fig. 3a, the MSSA values of GDGT-0 are higher than those of crenarchaeol across all temperatures within the 0–30 °C interval, indicating that the degradation ability of GDGT-0 is superior to that of crenarchaeol. The differences in their degradation ability can be attributed to their structural properties. In the core GDGT-0 structures, the only chemical groups present on the isoprenoid chains are the methyl groups, which are strong electron-donating groups[13,37,39]. In contrast, crenarchaeol contains four cyclopentane rings and one cyclohexane ring, with a weaker electron-donating ability than the methyl groups. Thus, GDGT-0 may be more active than crenarchaeol, and tend to degradation in the natural environments. Furthermore, the bond cleavage in the cyclopentane and cyclohexane structures requires a high covalent force due to the strong binding energy of the C-C bond in cycloalkyl groups[39]. These findings support our simulated results that the degradation ability of GDGT-0 is higher than that of crenarchaeol and reinforce the hypothesis that the MSSA values reflect the preferential degradation of GDGTs. Schouten et al. (2004) demonstrated that the degradation of GDGTs occurs substantially by altering the relative amounts of intact tetraether lipids under the influence of sediment thermal maturity[23], which is aligned with our simulated results.

Importantly, the MSSA values for both GDGT-0 and crenarchaeol exhibit distinct temperature-dependent trends, segregating into two temperature intervals (0–10 °C and 15–30 °C). There is a decrease in the MSSA values with increasing temperature within these intervals, with a notable inflection point at 15 °C (Fig. 3a). The similar patterns observed in the MSSA values of GDGT-0 and crenarchaeol suggest a consistent degradation behavior following their burial in sediments. This consistent degradation behavior implies that temperature-induced changes in the distributions of GDGT-0 and crenarchaeol are likely occurring proportionally[23], either positively or negatively.

This phenomenon could be tentatively explained by previous research indicating that redox conditions influence the distribution of GDGTs in oxygen-rich environments, with minimal impact on regional temperature reconstructions over millennia[21,26]. Despite variations in the degradation rates of GDGT-0 and crenarchaeol, the consistent changes in their distributions suggest little divergence in their distribution patterns when compared together. Our findings thus highlight the potential of GDGT-0 and crenarchaeol as reliable molecular markers for tracking environmental variations across extended geological time scales.

Previous research has primarily focused on using less abundant GDGTs (such as GDGT-1-3 and crenarchaeol isomer) as paleotemperature indicators through methods like $TEX_{86}$[11,14–22]. Our simulated results show that the STE values of less abundant GDGTs are relatively higher than GDGT-0 and crenarchaeol, suggesting a less stable characteristics in the structures of the less abundant GDGTs (Supplementary Fig. S2). Changes in GDGT distribution trends resulting from varying degradation abilities may be more pronounced for less abundant GDGTs compared with GDGT-0 and crenarchaeol, as these compounds are more sensitive to environmental fluctuations[14]. If the degradation abilities of less abundant GDGTs impact markedly on their distribution patterns, the suitability of their utilization as paleoenvironmental indicators, such as $TEX_{86}$, should be re-evaluated, especially in ancient sedimentary records. Thus, we also investigate the MD-based structures of GDGT-1–3 and crenarchaeol isomer to evaluate variations in their degradation abilities in response to temperature changes.

We utilize the published global $TEX_{86}$ values ($n = 723$, see spatial GDGT distribution data) to represent the distribution trends of GDGT-1–3 and crenarchaeol isomer, and compare them with the simulated MSSA values of GDGTs under the temperature range of 0–30 °C (Fig. 3b, c). The $TEX_{86}$ values are well-correlated with sea surface temperature at mid-latitudes (Fig. 3b). However, the scatter with a divergently positive bias of $TEX_{86}$ signatures at high- and low-latitude (temperature) internal are observed[13,14], indicating the unreliable application of $TEX_{86}$ at certain temperature. Our simulated results show that the discrepancies of MSSA values among GDGT-1–GDGT-3 and crenarchaeol isomer are relatively smaller at 5 and 15 °C but relatively larger at the simulated 0, 10, 20, 25, 30 °C (Fig. 3c), suggesting the different degradation abilities of GDGTs under temperature change. The similar trends between the global $TEX_{86}$ and our simulated MSSA values under a range of temperatures suggest that the unreliable application of $TEX_{86}$ during certain periods (Supplementary Discussion, Fig. S3, S4), especially in high- and low-latitude (temperature) environmental settings, may be related to the discrepancy in archaeal lipid degradation abilities.

Previous researches have noted variations in the relationship between GDGT distribution trends and temperature across different ocean regions and environments[3,11,14–19]. These differences were attributed to ecological variations in spatial distribution[19], environmental factors influencing on GDGT preservation[1,28] or the introduction of GDGTs from external sources[19,20], which may challenge traditional $TEX_{86}$-temperature calibrations. Our simulations suggest that anomalies in the global $TEX_{86}$ signatures are more likely influenced by the variable degradation ability of GDGT-1–GDGT-3, and crenarchaeol isomers driven by temperature. Thus, for long-term archaeal lipid evolution, especially in million-year timescale geological records, we suggest that compared to less abundance GDGTs, the distribution trends of GDGT-0 and crenarchaeol may be more suitable for overarching mechanisms governing archaeal lipid evolution particularly for geological sedimentary archives subject to extensive degradation, due to their consistent degradation abilities.

### Temperature-dependent distribution trends of GDGT-0 and crenarchaeol in geological records

Based on 3132 published abundance data of GDGT-0 and crenarchaeol in the marine environment since the late Jurassic (~200 Ma), we find a similar scissor-shaped distribution pattern of GDGT-0 and crenarchaeol at each geological episode (Fig. 4, Supplementary Fig. S5, see temporal GDGT distribution data). The prevailing concept posits multiple factors affecting GDGT distribution (e.g., temperature, pH, salinity)[11,14,18–22], but we observe the similar distribution pattern of GDGT-0 and crenarchaeol across the ~200 Ma, suggesting that their temporal distribution trends may be under a main controlling mechanism rather than dependent on the comprehensive regulation of these environmental factors. Due to the widespreadness of archaea under diverse and changing environments on Earth[7,8,45,46], the globally archaeal lipid distributions under the gradual-changing environments can be controlled by a consistent determining factor, further resulting in this similar temporal relative distribution trends of GDGT-0 and crenarchaeol. In contrast, the variable distribution patterns of the less abundant GDGTs may imply a confounding influencing factor (Supplementary Fig. S5). Since the archaea can adapt to the changing environmental stress

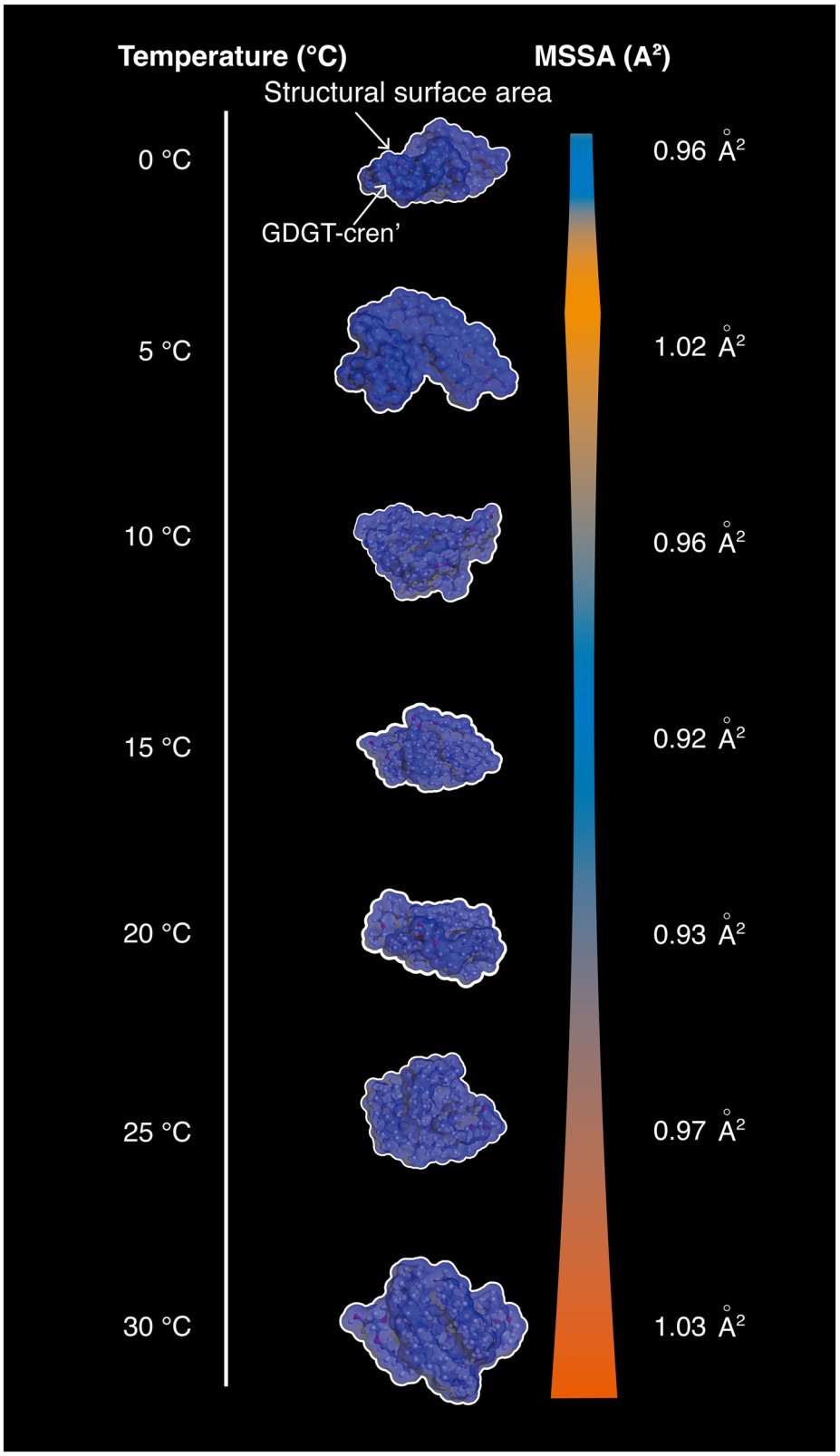



**Fig. 2 | Temperature-dependent changes in molecule-specific surface area of GDGT.** The simulated molecule-specific surface area (MSSA) values of the equilibrium configurations under a temperature range of 0–30 °C, with crenarchaeol isomer taken as an example. The MSSA of the equilibrium configurations changes from 0.96 to 1.03 Å under a temperature range of 0–30 °C.

through regulating membrane permeability, the similar distribution trends of GDGT-0 and crenarchaeol through geological time may be attributed to the adaption ability of archaea through producing archaeal lipids under the Earth's environmental change[45–47]. Thus, the distribution patterns of GDGT-0 and crenarchaeol may reveal the adaption ability of archaea during the geological time.

In addition, we find that there is a 'backward' trend in the distribution trends of GDGT-0 and crenarchaeol with temperature at each geological episode since the late Jurassic (Fig. 4, arrow). Such 'backward' trend of the GDGT-0 and crenarchaeol distributions suggests that the maximum temperature at which archaea can survive and adapt is gradient decreasing, which is corresponding to the gradually temperature cooling from the Late

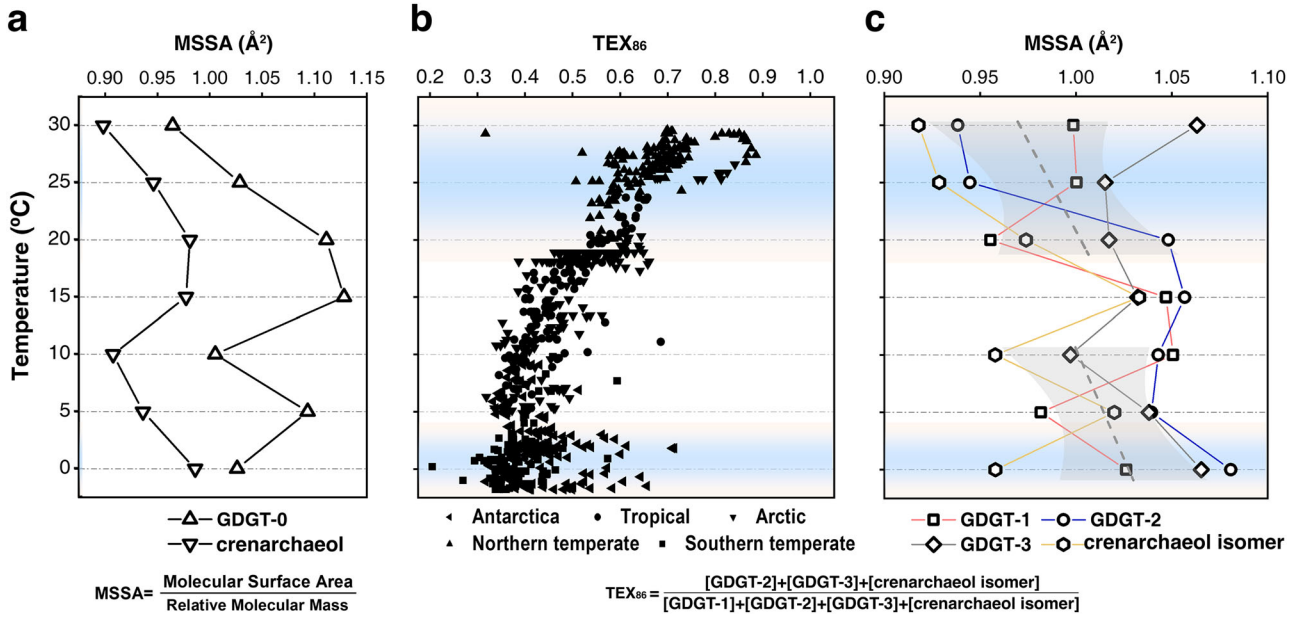



**Fig. 3 | Variations in MSSA values of GDGTs and global $TEX_{86}$ distribution patterns. a** The MSSA values of the equilibrium configurations of GDGT-0 (open upright triangles) and crenarchaeol (open inverted triangles) under a temperature range of 0–30 °C. **b** $TEX_{86}$ data of GDGTs from global marine surface sediments ($n = 723$): Antarctica (solid left-pointing triangles); Tropical (solid circles); Arctic (solid inverted triangles); Northern temperate (solid upright triangles); Southern temperate (solid squares). **c** The MSSA values of the equilibrium configurations of minor abundant GDGTs (GDGT-1 (open square), GDGT-2 (open circles), GDGT-3 (open diamonds) and crenarchaeol isomer (open hexagons)) under a temperature range of 0–30 °C. The shaded regions indicate the 95% confidence intervals of MSSA trends within two temperature intervals (0–10 °C and 15–30 °C).

Jurassic period to the present based on the planktonic foraminiferal $\delta^{18}O$ levels (Supplementary Fig. S6)[48,49]. This suggests that the Earth's temperature may be the dominant parameter influencing the temporal distribution of GDGT-0 and crenarchaeol, which is attributed to the adaption strategy of archaea in regulating membrane permeability under the Earth's temperature change.

Given the adaption strategy of archaea to the Earth's temperature can be revealed, we can thus predict the distribution trends of GDGT-0 and crenarchaeol from earlier sedimentary archives. Adaptive behaviors enable organisms to thrive in favorable conditions and avoid unfavorable environments[50]. This principle applies to archaea as well, as their evolutionary success is closely linked to their ability to adjust membrane lipid composition, metabolic pathways, and ecological strategies in response to shifting environmental conditions. Specifically, avoiding detrimental conditions is a crucial adaption strategy for the long-term evolution of archaea, allowing them to quickly escape harmful conditions and maintain the continuity of their populations[50]. Consequently, when the Earth's temperature is not conducive to archaea, archaeal cells may prioritize the basic survival resources for their essential metabolic activities, minimizing the allocation of these substance towards the synthesis of membrane lipids[45,47]. Under this harmful temperature condition, archaea may thus opt to synthesize GDGT structures that require less energy[45], such as the most dominated components of GDGT-0 and crenarchaeol, altering their distributions. It is thus possible that the modified distribution of GDGT-0 and crenarchaeol under harmful temperature conditions for archaea is recorded in the trends of their distributions for long-term geological time. By demonstrating the harmful temperature condition for archaea, the adaption strategy of them under the Earth's temperature change can be largely revealed.

In addition, because archaea inhabit a wide range of environments, it is reasonable that such a harmful temperature condition is not a specific temperature value, but rather a temperature range. This is supported by the observation from Pearson et al. (2014) that rare distributions of GDGT-0 and crenarchaeol within a specific temperature range across various archaeal survival habitats[14]. We thus suggest that there may be a harmful temperature range for archaea as part of their adaptation strategy for survival. More importantly, this harmful temperature range is likely not static, it may gradually evolve over time due to the fluctuations in the Earth's temperature. This makes it difficult to clarify the accurate harmful temperature range for archaea survival, especially during long-term geological history. Nevertheless, we still observe this "backward" trend in the GDGT-0 and crenarchaeol distributions over geological timescale (Fig. 4), suggesting the potential existence of a common factor (common harmful temperature range) that lays foundations between adaptation strategy of archaea and the Earth's temperature change, which could be reflected by the distribution trends of GDGT-0 and crenarchaeol. Thus, it is necessary to evaluate the relationship between the distribution trends of GDGT-0 and crenarchaeol under the common harmful temperature range for archaea and the Earth's temperature change.

In this study, we define the common harmful temperature range as $T_h$. The distribution patterns of GDGT-0 and crenarchaeol under $T_h$ during each epoch are here represented by $F_h$ (Fig. 5). We utilized the globally average $CO_2$ concentration spanning the last ~200 million years ($CO_2$-average, Fig. 5)[51] to represent the Earth's temperature, as the gradual changes in Earth's temperature over geological time are primarily influenced by atmospheric $CO_2$ concentrations[50,51]. To construct the relationship between $F_h$ and average $CO_2$ concentration, we perform Q-learning, a form of reinforcement learning (RL, detailed in Methods)[52,53]. Q-learning is well-suitable for our study since it is capable of improvement through interaction in highly complex and uncertain environments, reducing reliance on predefined rules[53]. Thus, by selecting the optimal solving among the vast relationships between $F_h$ and average $CO_2$ concentration (the highest $R^2$ value, Fig. 5a), we can acquire the greatest common harmful temperature range for archaea survival. Our modeling results show that the $T_h$ falls within the temperature range of 26–36 °C, with the strongest correlation coefficient between $F_h$ and $CO_2$-average across the epochs ($R^2 = 0.90$, $P < 0.0005$), which is the optimal solution in over 30,000-repeated selecting results. This indicates that temperatures between 26–36 °C may represent the harmful temperature condition for archaea, posing challenges for them in synthesizing or maintaining GDGTs. Interestingly, this finding is independently supported by Pearson et al. (2014) that have shown limited

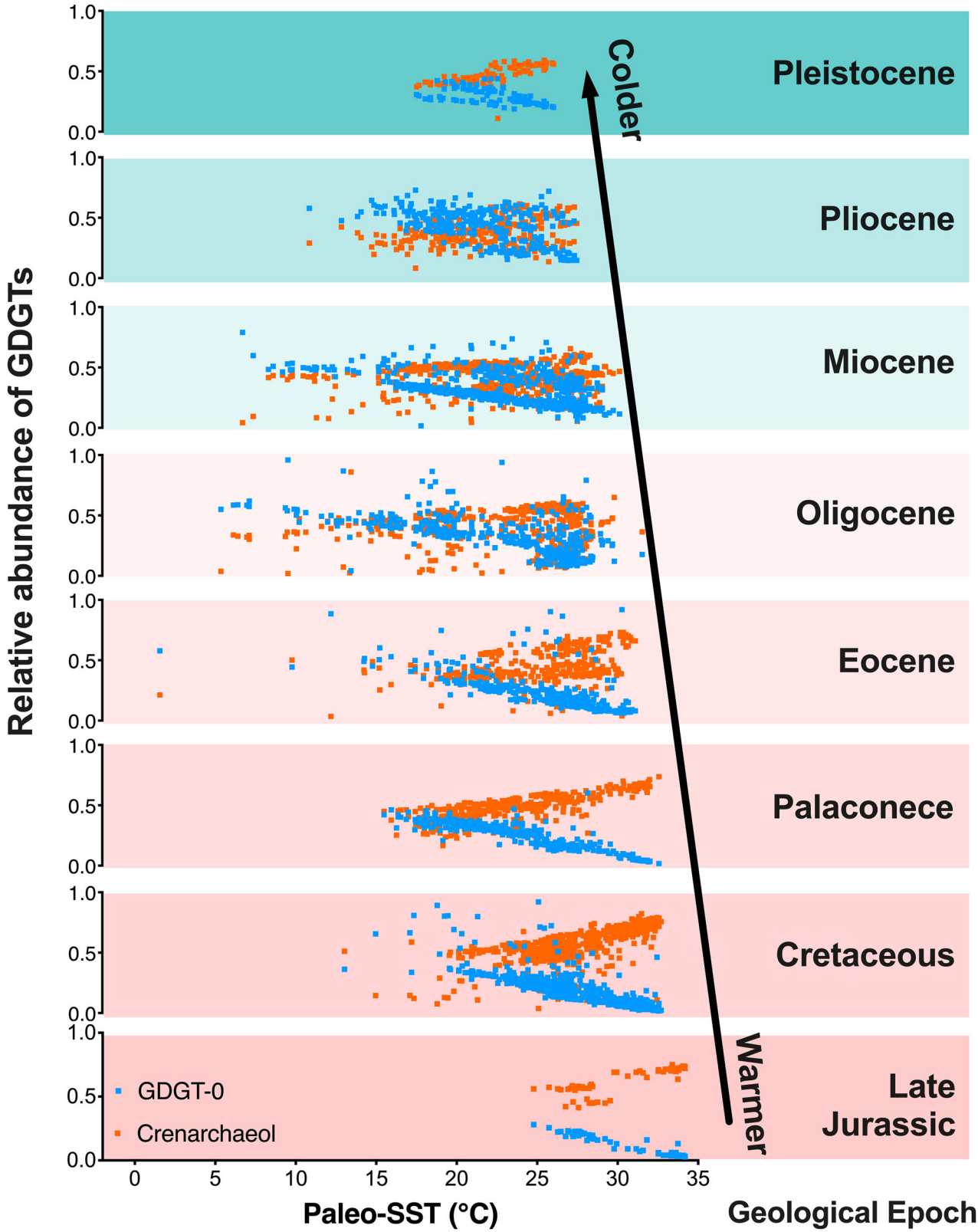



**Fig. 4 | Temporal distribution trends of GDGT-0 and crenarchaeol from the Late Jurassic to the Pleistocene.** Distribution trends of GDGT-0 (blue dots) and crenarchaeol (orange dots) across geological time from the Late Jurassic to the Pleistocene, based on global sedimentary records ($n = 3132$).

detection of GDGT-0 and crenarchaeol at temperatures ranging from 26–36 °C based on both experimental surveys and culture experiments[14], consistent with our modeling results.

Hence, we define this harmful temperature range from 26 to 36 °C as the optimal $F_h$. In Fig. 5b, we compare the variations between the optimal $F_h$ and the average $CO_2$ concentrations. We find that there are consistent variation trends between the optimal $F_h$ and the average $CO_2$ concentrations from each epoch, implying that the optimal $F_h$ may be influenced by the global atmosphere $CO_2$ variation. Generally, the elevated atmosphere $CO_2$ may lead to the increase in the Earth's temperature, which forces archaea

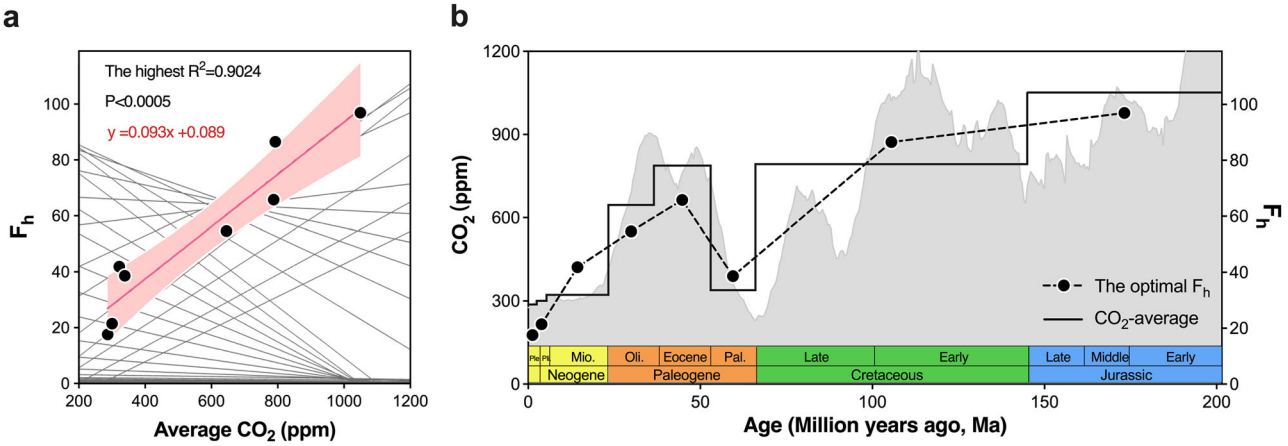



**Fig. 5 | Relationship between GDGT distributions and atmospheric $CO_2$ concentrations across geological epochs. a** Relationship between the distribution proportions of GDGT-0 and crenarchaeol of each epoch under harmful temperature range calculated by reinforcement learning ($F_h$) and average $CO_2$ concentration. Red line represents the highest correlation coefficient result between $F_h$ and $CO_2$-average (the optimal $F_h$ $R^2$ = 0.9024) with shaded 95% confidence among the fitness lines of all RL results (gray lines, $n$ = 30000). **b** Comparation between GDGT-0 and crenarchaeol proportions of each epoch under the optimal $F_h$ (highest correlation coefficient result) calculated by reinforcement learning (black circle blocks) and average $CO_2$ concentration ($CO_2$-average, solid black line).

species to adapt[53–55]. This may lead to an expansion in the distributions of archaea species that need to adapt to elevated temperatures, causing a corresponding increase in GDGT production by this group of archaea, thereby leading to their increasing distributions in the optimal $F_h$. Thus, the optimal $F_h$ of GDGT-0 and crenarchaeol can thus represent the highest potential factor in the relationship between $F_h$ and $CO_2$-average, based on current data. By utilizing this relationship, we propose that we can predict the optimal $F_h$ of GDGT-0 and crenarchaeol under the harmful temperature range further back in time than what is currently available.

The phylogenetic analyses have estimated the evolutionary timelines of *Thaumarchaeota*, a predominant archaeal group in marine environments, to have originated approximately 1017 Ma ago in shallow ocean habitats and 643 Ma ago in deep-sea habitats[9,56]. However, such emergence event of archaea, especially in their early evolutionary stages cannot be directly revealed using the existing GDGT database due to the situation that GDGTs are only preserved in immature marine sediments dating back to the past 192 Ma[11]. Consequently, inferring the evolution of archaea in response to environmental changes solely based on the existing GDGT distribution data is challenging. However, our study highlights the importance of predicting trends in the distribution of GDGTs reaching back to earlier periods, which may provide valuable information about the evolutionary stages of archaea, such as their emergence or mass reproduction in natural environments.

Throughout Earth's history, the pronounced increases in the global $CO_2$ concentration have been associated with the major biological extinction events[44]. However, archaea date back to the origin of life on Earth at least 3.8 billion years ago, and they exist well beyond any extinction events[45], highlighting their quintessential role in shaping the evolutionary stages of life on Earth[45]. Thus, by predicting the distribution trends of archaeal lipids in million to billion-year timescale geological records can provide a comprehensive and coherent chain of evidence of evolution stages of archaea, or/and even the whole life on Earth. The high optimal $F_h$ of GDGT-0 and crenarchaeol in Jurassic period (~150–200 Ma). This finding strongly suggests that archaea were confronted with substantial environmental changes. In response, numerous archaea underwent adaptive processes (Fig. 5b). This can be supported by previous studies, showing the evolution timelines of ammonia-oxidizing archaea in the marine environment were before 140 Ma[22,57,58]. The consistent timeline of ammonia-oxidizing archaea emergence in the marine environment reinforces the practicality and reliability of our prediction method in providing valuable information of how archaea have historically adapted to the varying environmental conditions.

In addition, the adaption strategy of archaea is expressed as the evolving GDGT synthesis pathways though long-term geological time, including methanogenesis, sulfide oxidation and ammonia oxidation, and are linked to the timing of key nodes of geologic Earth's events[46]. Specifically, large-scale volcanic eruptions (e.g., the Siberian Traps at 252 Ma and the Deccan Traps at 66 Ma) and the formation of hydrothermal vents on the seafloor can release vast amounts of $CO_2$ into the atmosphere[46]. In contrast, orogeny, or mountain-building events, expose new rock surfaces, increasing weathering processes that absorb $CO_2$ from the atmosphere[47]. Such geologic events may lead to the alteration in production GDGTs by archaea, raising the importance of predicting the distribution trends of GDGT-0 and crenarchaeol during these geologic events[9,10]. More importantly, archaea inhabit nearly all environments on Earth, including those occupied by macroscopic organisms, as well as extreme and deep biosphere environments[1,14]. The strategies for adaptation to these extreme habitats may differ considerably from those employed in typical marine environments, as demonstrated in our study. This notion is supported by previous cultivation experiments indicating that methanogenic archaea, such as *Methanocaldococcus jannaschii*, exhibit an optimal growth temperature range of 47–75 °C, which likely reflects the conditions found in hydrothermal environments[10,11]. This underscores the necessity for further exploration. Future investigations may aim to predict the distribution trends of archaeal lipids across various environmental contexts while also assessing the adaptive differences among distinct archaeal communities as data become available. In summary, our study uses the varying environmental conditions to predict the distribution trend of GDGTs, which is an important indicator advancing our understanding of the evolutionary adaption strategy of archaea, bridging the gap between phylogenetic information and sedimentary records.

The spatiotemporal distribution patterns of GDGT-0 and crenarchaeol contain valuable evolutionary information about archaea, yet have been previously overlooked. Through MD simulations and large GDGT datasets investigation, this research enhances our understanding the significance of GDGT-0 and crenarchaeol in elucidating the adaptive strategies of archaea in response to Earth's temperature changes. While previous research has primarily focused on the optimal conditions for archaeal survival[51], it is often the case that the adaptive strategies for archaea evolution over time is substantially influenced by unfavorable conditions. Building upon these findings, we have first utilized reinforcement learning to link GDGT distribution trends under unfavorable conditions with global $CO_2$ concentrations, identifying the most influential common factor of this relationship across geological time. Our study lays the groundwork for future detailed investigations to better comprehend Earth's molecular fossil record and the nature of ancient microbial life.

## Methods

### MD simulations

In this study, MD simulations were performed using the Material Studios 8.0 software package (Accelerys Software Inc.). The COMPASSII force field was adopted for its broad coverage of covalent molecules, including most common organics, small inorganic molecules, and polymers, and was appropriate for our situation with GDGTs[59,60]. GDGT structures (GDGT-0 ~ 3, crenarchaeol and crenarchaeol isomer) were constructed as parallel configurations[38]. Specifically, crenarchaeol and its isomer (crenarchaeol isomer) were constructed as trans of crenarchaeol and cis of crenarchaeol isomer[37]. A periodic box was constructed to represent a simulated marine system. To simulate the seawater environment, GDGT molecules, water molecules, and major elements of seawater were added to the system[48,49]. The MD simulation model (a cubic box with a 34.56 Å edge under periodic boundary conditions) contained one optimized GDGT molecule, 1000 water molecules, and randomly dispersed ions[44,61]. To simulate a practical seawater environment, specific ions ($Na^+$, $K^+$, $Cl^-$, and $Mg^{2+}$) that represent major elements in marine systems were added in proportion to the concentration of each ion in the seawater while maintaining a charge balance[61–63]. The density of the system was set at an average seawater density of 1.03 g/cm$^3$. All simulations were performed under the same conditions. The stimulated temperatures utilized in this study ranged from 0 to 30 °C, with intervals of 5 °C, as depicted in Figs. 1, 2, and 3. This temperature range corresponds to the known habitats of marine archaea, such as ammonia-oxidizing archaea in modern marine environments[14,64]. However, while some marine archaea may inhabit environments beyond this range, our selected temperature intervals primarily reflect surface ocean conditions in the modern marine setting. Our MD system was simulated for 600 ps with an interval of 1 fs, which was sufficiently enough for the system to reach equilibrium[62].

### Parameters provided by the MD simulations

Parameters obtained from MD simulations are widely used to investigate the stability and physicochemical properties of various compounds, including metallic elements[65], crystalline compounds[66], and macromolecular compounds, such as proteins[67]. In this study, we utilized the potential and kinetic energies of the system, known as the structural total energy (STE), to explore the biophysical regulation of individual GDGT configurations under the influence of temperature[31,68,69]. The STE provides valuable information about the structural stability and physicochemical properties of molecules in response to environmental changes[69]. By comparing the STE values of different GDGT configurations at varying temperatures, we were able to gain valuable insights into the physical, chemical, and biological properties of these molecules.

Moreover, the molecule surface area of GDGT molecules provides insights into their interaction ability with the surrounding environment, including intermolecular forces and hydrogen bonds. The curly configurations of GDGT molecules indicate low exposure to the external environment and limited accessibility to interactions, suggesting their stability and persistence in environments[70]. The molecule surface area is commonly used to identify the stability of organic compounds[44], and protein macromolecule persistence[70] in the environment, making it a valuable tool for investigating the stability and persistence of GDGT molecules. Due to the differences in relative molecular masses between GDGT configurations, the bias may exist in comparing the molecule surface areas of different GDGT structures (e.g., GDGT-1–3 and crenarchaeol isomer), and especially at the molecular level, such a discrepancy may be magnified. We thus divided the molecule surface area values by the relative molecular masses of the corresponding GDGTs to eliminate this influence, expressed as the molecule-specific surface area (MSSA, Figs. 2, 3b). We proposed estimating the MSSA of GDGT molecules to better understand their susceptibility to degradation during sediment burial from an MD perspective.

### Shape similarity

The shape similarity of the distribution and simulated result trends of GDGT-0 and crenarchaeol was performed utilizing Matrix Laboratory (MATLAB). We first fit the curve of GDGT distribution trends with a second-order polynomial to enable comparisons with STE results. For the purpose of eliminating the magnitude discrepancy and comparing the shapes of the distribution and STE trends of GDGTs, we normalized the data in all curves. Then we utilized Procrustes Analysis to obtain the matching corresponding coordinates of STE data as close as distribution data of GDGTs. Procrustes Analysis is a statistical method for analyzing the shape distributions of different curves based on matching corresponding points (coordinates) in two datasets by translating, rotating, and scaling the coordinates of points, which is widely used in bioinformatics, microbiology and so on[42,43]. In order to quantify the similarity, we calculated the Euclidean Distance of the new datasets of STE and the distribution trends of GDGT-0 and crenarchaeol, which is widely used to quantize the similarity of two different curves[43]. The similarity between two curves is defined as $e^{-(\text{Euclidean distance})}$, and its thresholds value ranges is 0-1. The closer to 1, the more similar the two curves are.

### Reinforcement learning

To reveal the influence of the Earth's temperature on archaeal lipid distribution trends, we constructed the relationship between average $CO_2$ concentration and the distribution trends of GDGT-0 and crenarchaeol according to geological period (Fig. 4). We performed the RL by MATLAB to find the harmful temperature range of archaea ($T_h$). $T_h$ was randomly set, and for each epoch, the proportions of GDGT-0 and crenarchaeol under the random $T_h$ compared with their distributions under the whole epoch were expressed as $F_h$, which were variable according to $T_h$. The decision-making rule of RL in our study follows iteration the highest correlation coefficient between the average $CO_2$ and $F_h$ (the highest $R^2$). The Q-learning pseudocode was listed as followed:

Initialize $Q(s,a)$ arbitrarily
Repeat (for each episode):
  Initialize $s$
  Repeat (for each step of episode):
    Choose $a$ from $s$ using policy derived from $Q$ (e.g., $\varepsilon$-greedy)
    Take action $a$, observe $r$, $s'$
    $Q(s,a) \leftarrow Q(s,a) + [r + \max_a' Q(s',a') - Q(s,a)]s \leftarrow s'$;
  until $s$ is terminal

### Data collection and statistical analyses

In order to investigate the distribution mechanism of GDGTs in marine environments, we conducted a comprehensive analysis of publicly available data from previous studies. The data was categorized as follows:

1) To assess the total energy of GDGT-0 and crenarchaeol in response to growth temperature, we compared global distribution data for 723 GDGT-0 and crenarchaeol samples in marine environments to the results obtained from MD simulations.

2) To investigate the effect of the Earth's temperature on the archaeal lipids during geological time, we collected 3132 GDGT-0 and crenarchaeol samples from the Late Jurassic to Pleistocene.

3) To evaluate the impact of temperature on the structures of GDGT-1–3 and crenarchaeol isomer, we compared 723 $TEX_{86}$ values and associated mean annual sea surface temperature to the MSSA of GDGT-1–3 and crenarchaeol isomer. Mean annual sea surface temperature were collected through satellite database from Tierney and Tingley[71]. Additionally, we included data that contained both global distribution and $TEX_{86}$ values of individual GDGTs.

The above data were collected from Tierney and Tingley[71], Kim et al.[72], Ho et al.[19], Leider et al.[73], Wei et al.[74], Hernandez-Sanchez et al.[75], Trommer et al.[76], Jia et al.[77], Rattanasriampaipong et al.[12].

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Data of spatial and temporal GDGT distribution, molecular dynamics simulations, sensitivity analysis and similarity analysis results are available in figshare at [https://doi.org/10.6084/m9.figshare.22698850].

## Code availability

Codes of similarity analysis and reinforcement learning are available in figshare at [https://doi.org/10.6084/m9.figshare.22698850].

Received: 8 March 2023; Accepted: 5 June 2025;
Published online: 04 August 2025

## References

1. Pearson, A. et al. Factors controlling the distribution of archaeal tetraethers in terrestrial hot springs. *Appl. Environ. Microbiol.* **74**, 3523–3532 (2008).
2. Zhang, C. et al. Lipid biomarkers and carbon isotope signatures of a microbial (Beggiatoa) mat associated with gas hydrates in the Gulf of Mexico. *Appl. Environ. Microbiol.* **71**, 2106–2112 (2005).
3. Volkman, J. K. Marine Organic Matter: Biomarkers, Isotopes and DNA. In *The Handbook of Environmental Chemistry* (ed. Volkman, J. K.) Vol. 2N (Springer, 2006).
4. Sinninghe et al. Origin of organic sulphur compounds and sulphur-containing high molecular weight substances in sediments and immature crude oils. *Org. Geochem.* **13**, 593–606 (1988).
5. Ingalls, A. E. et al. Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. *Proc. Natl Acad. Sci. USA.* **103**, 6442–6447 (2006).
6. Summons, R. E. et al. Lipid biomarkers: molecular tools for illuminating the history of microbial life. *Nat. Rev. Microbiol.* **20**, 174–185 (2022).
7. Brocks, J. J. & Banfield, J. Unravelling ancient microbial history with community proteogenomics and lipid geochemistry. *Nat. Rev. Microbiol.***7**, 602–609 (2009).
8. Yasuhara, M. et al. Time machine biology cross-timescale integration of ecology, evolution, and oceanography. *Oceanography* **33**, 16–28 (2020).
9. Abby, S. S. et al. Ancestral reconstructions decipher major adaptations of ammonia-oxidizing archaea upon radiation into moderate terrestrial and marine environments. *mBio* **11**, https://doi.org/10.1128/mbio.02371-20 (2023).
10. Taubner, R. et al. Lipidomics and comparative metabolite excretion analysis of methanogenic srchaea reveal organism-specific adaptations to varying temperatures and substrate concentrations. *mSystems* **8**, e01159-22 (2023).
11. Sprott, G. D. et al. Proportions of diether, macrocyclic diether, and tetraether lipids in Methanococcus jannaschii grown at different temperatures. *J. Bacteriol.* **173**, 3907–3910 (1991).
12. Rattanasriampaipong, R. et al. Archaeal lipids trace ecology and evolution of marine ammonia-oxidizing archaea. *Proc. Natl Acad. Sci. USA*. **119**, e2123193119 (2022).
13. Schouten, S. et al. Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures. *Earth Planet. Sci. Lett.* **204**, 265–274 (2002).
14. Pearson, A. & Ingalls, A. E. Assessing the use of archaeal lipids as marine environmental proxies. *Annu. Rev. Earth Pl. Sc.* **41**, 359–384 (2013).
15. Sinninghe Damsté, J. S. et al. Widespread occurrence of structurally diverse tetraether membrane lipids: evidence for the ubiquitous presence of low-temperature relatives of hyperthermophiles. *Proc. Natl Acad. Sci. USA.* **97**, 14421–14426 (2000).
16. Schouten, S. et al. The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: a review. *Org. Geochem.* **54**, 19–61 (2013).
17. Jones, T. D. et al. OPTiMAL: a new machine learning approach for GDGT-based palaeothermometry. *Clim. Past* **16**, 2599–2617 (2020).
18. Turich, C. et al. Lipids of marine archaea: patterns and provenance in the water-column and sediments. *Geochim. Cosmochim. Acta* **71**, 3272–3291 (2007).
19. Huguet, C. et al. Effects of long term oxic degradation on the $TEX_{86}$ and BIT organic proxies. *Org. Geochem.* **40**, 1188–1194 (2009).
20. Ho, S. L. et al. Appraisal of $TEX_{86}$ and $TEX_{86}^{L}$ thermometries in subpolar and polar regions. *Geochim. Cosmochim. Acta* **131**, 213–226 (2014).
21. Inglis, G. N. et al. Descent toward the Icehouse: Eocene sea surface cooling inferred from GDGT distributions. *Paleoecanogr. Paleocl.* **30**, 1000–1020 (2015).
22. O'Brien, C. L. et al. Cretaceous sea-surface temperature evolution: Constraints from $TEX_{86}$ and planktonic foraminiferal oxygen isotopes. *Earth-sci. Rev.* **172**, 224–247 (2017).
23. Schouten, S. et al. The effect of maturity and depositional redox conditions on archaeal tetraether lipid palaeothermometry. *Org. Geochem.* **35**, 567–571 (2004).
24. Kusch, S. et al. Factors influencing $^{14}C$ concentrations of algal and archaeal lipids and their associated sea surface temperature proxies in the Black Sea. *Geochim. Cosmochim. Acta* **188**, 35–57 (2018).
25. Cobban, A. et al. Multiple environmental parameters impact lipid cyclization in Sulfolobus acidocaldarius. *Environ. Microbiol.* **22**, 4046–4056 (2020).
26. Qin, W. et al. Confounding effects of oxygen and temperature on the $TEX_{86}$ signature of marine Thaumarchaeota. *Proc. Natl Acad. Sci. USA.* **112**, 10979–10984 (2015).
27. Zeng, Z. et al. GDGT cyclization proteins identify the dominant archaeal sources of tetraether lipids in the ocean. *Proc. Natl Acad. Sci. USA.* **116**, 22505–22511 (2019).
28. Elling, F. J. et al. Influence of temperature, pH, and salinity on membrane lipid composition and $TEX_{86}$ of marine planktonic thaumarchaeal isolates. *Geochim. Cosmochim. Acta* **171**, 238–255 (2015).
29. Baxter, A. J. et al. Seasonal and multi-annual variation in the abundance of isoprenoid GDGT membrane lipids and their producers in the water column of a meromictic equatorial crater lake (Lake Chala, East Africa). *Quaternary Sci. Rev.* **273**, 107263 (2021).
30. Raberg, J. H. et al. Near-universal trends in brGDGT lipid distributions in nature. *Sci. Adv.* **8**, 1–12 (2022).
31. He, X. & Zhang, J. Z. H. A new method for direct calculation of total energy of protein. *J. Phys. Chem.* **122**, 1–3 (2005).
32. Proença, B. et al. Stability and structural analysis of $A_6R$ polypeptide nanosheets: a theoretical study using the classical molecular dynamics simulation. *J. Phys. Chem. C* **122**, 24445–24453 (2018).
33. Bruce, C. D. et al. Molecular dynamics simulations of sodium dodecyl sulfate micelle in water: the behavior of water. *J. Phys. Chem. B* **106**, 10902–10907 (2002).
34. Wall, M. E. et al. Conformational dynamics of a crystalline protein from microsecond-scale molecular dynamics simulations and diffuse X-ray scattering. *Proc. Natl Acad. Sci. USA.* **111**, 17887–17892 (2014).
35. Huguet, C. et al. Molecular dynamics simulation study of the effect of glycerol dialkyl glycerol tetraether hydroxylation on membrane thermostability. *Biochim. Biophys.Acta (BBA)—Biomembr.* **1859**, 966–974 (2017).
36. Chong, P. L. Archaebacterial bipolar tetraether lipids: Physico-chemical and membrane properties. *Chem. Phys. Lipids* **163**, 253–265 (2010).
37. Sinninghe Damsté, J. S. et al. The enigmatic structure of the crenarchaeol isomer. *Org. Geochem.* **124**, 22–28 (2018).
38. Liu, X. et al. Glycerol configurations of environmental GDGTs investigated using a selective *sn2* ether cleavage protocol. *Org. Geochem.* **128**, 57–62 (2019).
39. Sinninghe et al. Crenarchaeol: the characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota. *J. Lipid Res.* **43**, 1641–1651 (2002).
40. Ventura, G. T. et al. Molecular evidence of Late Archean archaea and the presence of a subsurface hydrothermal biosphere. *Proc. Natl. Acad. Sci. USA.* **104**, 14260–14265 (2007).
41. Gabriel, J. L. & Chong, P. L. Molecular modeling of archaebacterial bipolar tetraether lipid membranes. *Chem. Phys. Lipids* **105**, 193–200 (2000).
42. Oladyshkin, S. et al. Global sensitivity analysis: A flexible and efficient framework with an example from stochastic hydrogeology. *Adv. Water Resour.* **37**, 10–22 (2012).

43. Oladyshkin, S. & Nowak, W. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliab. Eng. Syst. Safe.* **106**, 179–190 (2012).
44. Devarajan, D. et al. Molecular dynamics simulation of the structures, dynamics, and aggregation of dissolved organic matter. *EST.* **54**, 13527–13537 (2020).
45. Valentine, D. L. Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat. Rev. Microbiol.* **5**, 316–323 (2007).
46. Cavicchioli, R. Archaea — timeline of the third domain. *Nat. Rev. Microbiol.* **9**, 51–61 (2011).
47. Baker, B. J. et al. Diversity, Ecology and evolution of Archaea. *Nat. Microbiol.* **5**, 887–900 (2020).
48. Westerhold, T. et al. An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science* **369**, 1383–1387 (2020).
49. Hansen, J. et al. Climate sensitivity, sea level and atmospheric carbon dioxide. *Phil. Trans. R. Soc. A.* **371**, 20120294 (2013).
50. Hutchins, D. A. & Fu, F. Microorganisms and ocean global change. *Nat. Rev. Microbiol.* **2**, 17508 (2017).
51. Foster, G. L. et al. Future climate forcing potentially without precedent in the last 420 million years. *Nat. Commun.* **8**, 14845 (2017).
52. van Otterlo, M. & Wiering, M. In *Reinforcement Learning. Adaptation, Learning, and Optimization* Vol 12 (eds Wiering, M. & van Otterlo, M.) (Springer, 2012).
53. Luksys, G. et al. Stress, genotype and norepinephrine in the prediction of mouse behavior using reinforcement learning. *Nat. Neurosci.* **12**, 1180–1186 (2009).
54. Cavicchioli, R. et al. Scientists' warning to humanity: microorganisms and climate change. *Nat. Rev. Microbiol.***17**, 569–586 (2019).
55. Hoffmann, A. A. & Sgrò, C. M. Climate change and evolutionary adaptation. *Nature* **470**, 479–485 (2011).
56. Follows, M. J. et al. Emergent biogeography of microbial communities in a model ocean. *Science* **315**, 1843–1846 (2007).
57. Yang, Y. et al. The evolution pathway of ammonia-oxidizing Archaea shaped by major geological events. *Mol. Biol. Evol.* **38**, 3637–3648 (2021).
58. Ren, M. et al. Phylogenomics suggests oxygen availability as a driving force in Thaumarchaeota evolution. *ISME J.* **13**, 2015–2161 (2019).
59. Bunte, S. W. & Sun, H. Molecular modeling of energetic materials: the parameterization and validation of nitrate esters in the COMPASS Force Field. *J. Phys. Chem. B* **104**, 2477–2489 (2000).
60. Sun, H. et al. COMPASS II: extended coverage for polymer and drug-like molecule databases. *J. Mol. Model.* **22**, 1–10 (2016).
61. Vollmayr-Lee, K. Introduction to molecular dynamics simulations. *Am. J. Phys.* **88**, 401–422 (2020).
62. Zubeltzu, J. et al. Lignin solvation by ionic liquids: the role of cation. *J. Mol. Liq.* **303**, 112588 (2020).
63. Zhao, T. et al. Molecular dynamics study of alkyl benzene sulfonate at air/water interface: effect of inorganic salts. *J. Phys. Chem. B* **114**, 5025–5033 (2010).
64. Millero, F. J. *Chemical Oceanography* (Taylor & Francis, 2006).
65. Materer, N. et al. Molecular surface structure of ice (0001): dynamical low-energy electron diffraction, total-energy calculations and molecular dynamics simulations. *Surf. Sci.* **381**, 190–210 (1997).
66. Liu, N. et al. Crystal morphology of 3,4-bis(3-nitrofurazan-4-yl) furoxan (DNTF) in a solvent system: molecular dynamics simulation and sensitivity study. *Crystengcomm* **18**, 2843–2851 (2016).
67. Glaser, F. et al. A method for localizing ligand binding pockets in protein structures. *Proteins Struct Funct. Bioinform.***62**, 479–488 (2006).
68. Ensing, B. et al. Energy conservation in adaptive hybrid atomistic/coarse-grain molecular dynamics. *J. Chem. Theory Comput.* **3**, 1100–1105 (2007).
69. He, X. & Zhang, J. Z. H. The generalized molecular fractionation with conjugate caps/molecular mechanics method for direct calculation of protein energy. *J. Phys. Chem.* **124**, 184703 (2006).
70. Huang, B. & Schroeder, M. LIGSITEcsc: predicting ligand binding sites using the Connolly surface and degree of conservation. *BMC Struct. Biol.* **6**, 1–11 (2006).
71. Tierney, J. E. & Tingley, M. P. A $TEX_{86}$ surface sediment database and extended Bayesian calibration. *Sci Data* **2**, 1–10 (2015).
72. Kim, J. et al. New indices and calibrations derived from the distribution of crenarchaeal isoprenoid tetraether lipids: Implications for past sea surface temperature reconstructions. *Geochim. Cosmochim. Acta* **74**, 4639–4654 (2010).
73. Leider, A. et al. Core-top calibration of the lipid-based $UK_{37}$' and $TEX_{86}$ temperature proxies on the southern Italian shelf (SW Adriatic Sea, Gulf of Taranto). *Earth Planet. Sci. Lett.* **300**, 112–124 (2010).
74. Wei, Y. et al. Spatial variations in archaeal lipids of surface water and core-top sediments in the South China Sea and their implications for paleoclimate studies. *Appl. Environ. Microbiol.* **77**, 7479–7489 (2011).
75. Hernández-Sánchez, M. T. et al. Variations in GDGT distributions through the water column in the South East Atlantic Ocean. *Geochim. Cosmochim. Acta* **132**, 337–348 (2014).
76. Trommer, G. et al. Distribution of Crenarchaeota tetraether membrane lipids in surface sediments from the Red Sea. *Org. Geochem.* **40**, 724–731 (2009).
77. Jia, G. et al. Archaeal tetraether lipids record subsurface water temperature in the South China Sea. *Org. Geochem.* **50**, 68–77 (2012).


## Acknowledgements

This work is funded by the National Natural Science Foundation of China (42525604), the National Natural Science Foundation of China (42450135), the Fundamental Research Funds for the Central Universities (202042010), the Taishan Young Scholars (tsqn202103030), and Shandong Province Natural Science Foundation (ZR2021JQ12). We would also like to thank all reviewers for their constructive comments, which helped improve the manuscript. No specific permissions were required for the sampling conducted in this study.

## Author contributions

S.Z. and R.B. designed and performed research; S.Z. and L.Z. performed the simulation work; R.B. supervised the research; S.Z., R.B., M.L., X.S., Z.G., Y.Z., N.W., and Y.W. carried out the calculation and analysis; S.Z., R.B., M.L., X.S., Z.G., Y.Z., N.W., Y.W., J.C., L.X., and C.Z. edited and reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s43247-025-02450-7.

**Correspondence** and requests for materials should be addressed to Rui Bao.

**Peer review information** *Communications Earth & Environment* thanks César Ramírez and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Mojtaba Fakhraee, Somaparna Ghosh. A peer review file is available.

**Reprints and permissions information** is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025