Discovery of potent inhibitors of rumen methane metabolism via an AI-assisted workflow

Xiong Xia[1, #], Liangzhen Zheng[2, #], Jinjie Zhou[3, #], Haoran Ni[1], Guoliang Zhu[2], Hongbin Liu[1], Ruiyue Chen[1], Haimei Shi[1], Huizeng Sun[4], Meng Li[3, *], Sheng Wang[2, *], Lei Dai[1, *]

[1] State Key Laboratory of Quantitative Synthetic Biology, Shenzhen Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

[2] Shanghai Zelixir Biotech Co. Ltd, Shanghai 201210, Shanghai, China

[3]Archaeal Biology Centre, Synthetic Biology Research Center, Shenzhen Key Laboratory of Marine Microbiome Engineering, Key Laboratory of Marine Microbiome Engineering of Guangdong Higher Education Institutes, Institute for Advanced Study, Shenzhen University, Shenzhen 518060, China

[4] Institute of Dairy Science, College of Animal Sciences, Zhejiang University, Hangzhou 310058, China

[#]These authors contributed equally

[*]Correspondence:

lei.dai@siat.ac.cn (L.D.); wangsheng@zelixir.com (S.W.); limeng848@szu.edu.cn (M.L.)

Abstract

Ruminant methane emission has attracted increasing attention due to its significance in global climate change and sustainable livestock production. Methyl-coenzyme M reductase (MCR), the key enzyme involved in the terminal step of methanogenesis in the rumen microbiome, is responsible for nearly all biologically generated methane released into the atmosphere. In this study, we developed an AI-assisted workflow for discovery of novel MCR inhibitors. Ultra-fast virtual screening of ~23 million molecules yielded 26 candidate molecules for experimental validation, among them 4 molecules showed significant inhibition of methane metabolism. Subsequent structure-guided optimization led to the discovery of a sub-nanomolar inhibitor Pyrazol-5(4H)-one,3-(4-nitrophenyl) (PZON), which effectively suppressed methane production in rumen microbiome fermentation and exhibited minimal cytotoxicity. Overall, our study provides promising candidates for methane-reducing feed additives, and demonstrates the power of AI-assisted discovery of small molecules for targeted modulation of gastrointestinal microbiomes.

Introduction

Over the past few decades, remarkable progress has been made in elucidating the composition, diversity, and functional potential of the trillions of microorganisms residing in the host gastrointestinal tract [1]. Increasing evidence highlights the pivotal role of microbial metabolic activities in shaping host physiology, health, and the surrounding environment [2, 3]. These discoveries have spurred growing interest in harnessing the gastrointestinal microbiome as a promising frontier for precision modulation and therapeutic development [4]. However, despite substantial advances in microbiome research, our understanding of the precise molecular mechanisms by which microbial communities influence host biology and disease pathogenesis remains limited[5]. Consequently, there is an urgent need for innovative strategies that enable the targeted manipulation and functional dissection of individual microbial metabolic and biochemical processes within complex ecosystems [6-8].

Ruminant livestock play a crucial role in global food production but simultaneously emit substantial amounts of methane, a potent greenhouse gas and a major contributor to energy loss in the host [9]. Within the rumen, diverse microbial consortia—including bacteria, archaea, fungi, and protozoa—synergistically degrade plant polysaccharides and ferment substrates, ultimately driving methanogenesis [10-13]. Methanogenesis primarily proceeds through hydrogenotrophic, methylotrophic, and acetoclastic pathways (**Fig 1A**), with methyl-coenzyme M reductase (hereafter referred to as MCR) catalyzing the final and rate-limiting step of methane biosynthesis [14, 15]. Targeted engineering of ruminal methane metabolism is therefore essential for improving feed efficiency and mitigating environmental impacts [16-18]. Small-molecule inhibitors that selectively interfere with key microbial enzymes offer a promising avenue for precise microbial regulation without broad community disruption [6]. Current MCR inhibitors exhibit similar modes of action and do not guarantee sustained methane mitigation[16, 19, 20], calling for the development of novel inhibitors. However, conventional screening strategies are constrained by high cost, limited chemical space exploration and low hit efficiency, underscoring the need for efficient large-scale screening approaches to accelerate the discovery of structurally diverse and functionally potent inhibitors[21-23].

The need for efficient large-scale screening has prompted the development of numerous artificial intelligence (AI)-based methods, and these approaches are marked by accelerated processing, significantly lower costs, and high reliability in virtual docking [24-26]. Advance of modern structure prediction methods [27-29] demonstrated considerable accuracy in predicting protein-ligand complexes, with some models even capable of estimating binding free energies [28]. However, the predictive performance of these models is frequently unreliable for unseen ligands that are structurally distinct from training data (mainly the ligands in RCSB PDB structures) [30], posing a major challenge since such molecules comprise the vast majority of

virtual screening libraries. Meanwhile, these approaches remain prohibitively time-consuming, rendering them impractical for large to ultra-large scale molecular screening. Conventional virtual screening methods (such as AutoDock Vina) need to balance precision and speed for real-world drug discovery applications [31]. Integration of traditional docking tools [32] with deep learning-based scoring estimators provides a pragmatic approach [24, 25, 33].

In this study, we established a two-stage deep learning-based ultra-fast virtual screening workflow to identify potential inhibitors of MCR, the key enzyme involved in methanogenesis of the rumen microbiome. Our AI-assisted screening approach integrates the deep learning-based molecular representation [34] and the machine learning-based scoring function for efficient filtering and accurate predictions[35]. Using fermentation experiments of the rumen microbiome, we successfully validated multiple candidate compounds. Furthermore, structure-guided optimization led to the discovery of a sub-nanomolar inhibitor, Pyrazol-5(4H)-one,3-(4-nitrophenyl) (hereafter abbreviated as PZON). Finally, we found that PZON remarkably reduced methane production in rumen microbial fermentation *in vitro* with minimal disturbance to the overall microbiome composition. Overall, our study demonstrates the power of AI in accelerating the discovery of structurally novel and highly potent chemicals for targeted modulation of microbial metabolism in complex microbiomes.

Results

AI-assisted discovery of small molecule inhibitors for rumen methane metabolism

There are three primary methanogenesis pathways in the rumen microbiome (**Fig. 1A**). Terminal methane-yielding reaction, reduction of methyl-coenzyme M, is exclusively mediated by methyl-coenzyme M reductase (MCR), an evolutionarily conserved gatekeeper enzyme that governs the rate-limiting step shared by all anaerobic archaeal methanogenesis pathways. To identify MCR targeting compounds which would inhibit methanogenesis in the rumen microbiome, we implemented an iterative deep learning based ultra-fast virtual screening pipeline combining hybrid molecular fingerprint embeddings and multilayer perceptron regression to accelerate large-scale virtual screening against MCR (**Fig. 1B-C, Supplementary Fig. 1A**). Firstly, we randomly selected molecules (docking training dataset) from the combined library and preformed a traditional virtual screening to calculate all the binding energies (in kcal/mol) between each molecule with the receptor using AutoDock Vina (**Fig. 1B**). Next, a molecule-activity multiple layer perceptron (MLP) regression model was trained to learn the Vina docking score between a molecule (from the docking training dataset) with the MCR receptor (**Fig. 1B-C, Supplementary Fig. 1A**), whereas the molecules are represented by classical fingerprints and AI-based embeddings (see Methods).

The MLP models are trained in an iterative mode such that a best model could be harvest with limited docking efforts. After 12 iterations, we observed a best performed MLP model for predicting the AutoDock Vina scores (**Supplementary Fig. 1 B-D**). The optimum MLP model achieved excellent agreement (R^2 = 0.78, MSE = 2.67) between predicted and true Vina scores (**Fig. 1D**), demonstrating capability to replace resource-intensive docking for ultra-fast virtual screening. From the benchmarked computational time calculation, the two-stage workflow developed in this study demonstrates a 28- and 171-times speed-up comparing to the two pipelines, the brute force per-molecule Vina (iDock) virtual screening and OnionNet+SFCT virtual screening[35] (**Fig. 1E**).

We applied our AI-assisted pipeline to screen a library of ~23 million molecules to identify novel small molecule inhibitor against MCR for inhibition of rumen methane metabolism. The compound library integrated multiple sources to ensure both chemical diversity and practical applicability in the context of further development for ruminant methane mitigation. The ZINC20 Now database (21.45 million compounds) provides maximal coverage of the drug-like chemical space, ensuring broad structural exploration while retaining favorable physicochemical properties (e.g., MW < 600, logP < 5) compliant with feed additive safety and bioavailability requirements[36]. The ChemBridge lead-like subset (1.3 million compounds) offers optimized pharmacophore features and moderate molecular complexity, supporting good developability and oral bioavailability under practical dosing scenarios[37]. The COCONUT library (500,000 compounds), representing natural products and derivatives, introduces biogenic and eco-compatible structures that align with regulatory preferences for "naturally derived" or "generally recognized as safe-like" ingredients in the livestock industry[38]. The Enamine REAL Space collection (26,000 compounds) includes chemicals with focused diversity and feasibility of large-scale synthesis[39]. Collectively, this integrated 23-million-compound library balances chemical diversity, biological relevance, synthesis feasibility, and regulatory compliance—addressing the key challenges of identifying cost-effective, safe, and mechanistically precise small-molecule inhibitors for enteric methane reduction in ruminants.

For the top-ranked 0.1 million molecules after the first-stage screening, the second-stage screening based on OnionNetSFCT+Vina was applied to identify the most promising candidates (**Supplementary Fig. 1A**). Top candidates were then subjected to human evaluation of docking pose and ligand-protein interactions. Finally, 26 candidate compounds were selected for experimental validation (**Supplementary Table S1**). The SFCT+Vina binding scores for the top-ranked molecules ranged from 1.055 to 2.366, with a median value of 1.758. Specifically, the candidate compounds exhibited scores between 1.644 and 2.366 (**Fig. 1F**), with lower scores indicating stronger binding affinity. For comparison, previously reported MCR inhibitors[40], including bromoform, nitro-alcohol/esters, Coenzyme-B analogs, statins and pterins, showed a broader score distribution, ranging from 1.432 to 2.897 and a median of 2.204.

The top 2774 ranked molecules exhibited varying levels of structure similarity (Tanimoto coefficient ranging from 0 to 0.592, with a median value of 0.048) to methyl-CoM, the natural substrate of MCR. Out of the 26 candidate compounds selected, 24 compounds showed low structural similarity (Tanimoto coefficient < 0.3) to methyl-CoM (**Fig. 1G**). Structural diversity analysis (based on t-distributed stochastic neighbor embedding, t-SNE) indicated that the candidate compounds possess high structural diversity relative to both the top-ranked molecules identified through SFCT+Vina docking and known MCR inhibitors (**Fig. 1H**). Notably, these candidate compounds have high binding affinity to MCR and low similarity to methyl-CoM, highlighting their potential as structurally novel MCR inhibitors.

Validation and optimization of candidate MCR inhibitors

To validate the inhibitory potential of the candidate compounds, we established an *in vitro* batch fermentation system using cryopreserved rumen inoculum and methane quantification assay (**Fig. 2A, Supplementary Fig. 2**). The system reliably reproduced the methane-suppressing effect of the known MCR inhibitor BES at concentration of 10 μM (**Fig. 2B**), confirming its suitability for compound screening. We screened the 26 candidate compounds at 98.31 μM and found 4 hit compounds that significantly reduced methane production after 72 h fermentation of ruminal microbiota (collected from Xiangxi cattle, see Methods) by 57% to 76% (**Fig. 2C**). These 4 hit compounds, with chemical structures clearly distinguished from known MCR inhibitor [40], possess aromatic ring scaffolds and are enriched with highly polar heteroatoms (e.g., N, O, Cl) (**Fig. 2D**). Visual inspections revealed the structural basis of potential interactions between the hit compounds and the enzyme active sites (e.g. Arg120 on chain C, Tyr367 on chain B, Tyr333 on chain A and Ni^{2+} ion of cofactor F430) (**Fig. 2E**). The Arg120 residue on chain C is capable of forming a hydrogen-bonding network, which stabilizes the substrate-binding conformation. Meanwhile, the catalytic Ni^{2+} ion can engage in metal coordination chemistry by accepting lone-pair electrons from the substrate's nitrogen or oxygen atoms, thereby further enhancing the stability of the enzyme–substrate complex. In addition, the phenolic hydroxyl oxygen of Tyr367 residue on chain B and the Tyr333 residue on chain A can participate in pi-pi interactions with the delocalized pi-electron system of the substrate's aromatic ring. Assisted by the hydrophobic microenvironment of the active pocket, this interaction promotes the preferential orientation of the substrate's aromatic side chain toward the sulfur atom of coenzyme B, thereby increasing the binding affinity and competitively blocking the entry of endogenous substrates. Collectively, the cooperative effects of hydrogen bonding, metal coordination, and pi-electron interactions synergistically stabilize substrate binding and achieve competitive inhibition of the native substrate.

We further optimized the hit compounds to improve their inhibitory activity. The oblong shape of the pocket intuitively guides ligand design where the ligand should adopt a conformation

that positions a polar 'anchor' for charge stabilization at one end (for Arg120 on chain C), while displaying electronegative pharmacophores to maintain coordination bonds with the metal ion at the other[41]. Following this rationale, 11 structurally optimized derivatives were obtained (**Fig. 3A, Supplementary Fig. 3A-C, Supplementary Table S1**). Among these, Pyrazol-5(4H)-one,3-(4-nitrophenyl) (PZON) showed markedly enhanced activity, suppressing methane production by >97% at 98.31 μM (**Fig. 3B**, **Supplementary Fig. 3D-F**). PZON showed dose-dependent inhibitory activity of methane with an IC_{50} of 0.24 μM, while the IC_{50} before optimization was 1.56 μM (**Fig. 3C**). In the binding mode of PZON, the two oxygen atoms of the nitro group attached to the phenyl side chain forms a stable hydrogen bond with Arg120 on chain C, while the delocalized pi-electron system of the azole five-membered ring engages in a pi-pi interaction with Tyr333 (chain A) and Tyr367 (chain B) (**Fig. 3D**). In addition, the nitrogen atom of the substrate coordinates with the catalytic Ni^{2+} ion, thereby inducing a conformational adjustment of the side chain that positions the carbonyl oxygen toward the sulfur atom of coenzyme B to facilitate further interaction (**Fig. 3D**). These synergistic interactions collectively stabilize the substrate binding within the active site.

Our results demonstrate that PZON exerts measurable anti-methanogenic activity within complex rumen microbiomes, indicating its potential as an effective agent for mitigating methane emissions in ruminants. Furthermore, we assessed the anti-methanogenic effects of PZON in pure cultures of *Methanobrevibacter ruminantium* M1 (DSM 1093), a representative ruminal methanogen, at concentrations ranging from 3.33 μM to 30.0 μM (**Fig. 3E**). PZON completely inhibited the growth of *Methanobrevibacter ruminantium* and its methanogenesis throughout the 7-day incubation period. We next assessed the cytotoxicity of PZON against multiple cell lines at concentrations ranging from 1.53 μM to 100 μM. We found that PZON exhibited no cytotoxicity at 100 μM (**Fig. 3F**), which is >400-fold greater than the IC_{50}. The potent anti-methanogenic activity of PZON, together with its minimal cytotoxicity, underscores its potential as livestock feed additives aimed at reducing methane emission.

Evaluation of PZON using *in vitro* fermentation of the rumen microbiome

We further evaluated the anti-methanogenic activity of PZON across different host-derived rumen microbiomes, and to assess its potential influence on microbial metabolism and community structure using an *in vitro* batch fermentation model. We conducted experiments using rumen microbial communities collected from Xiangxi beef cattle, Holstein dairy cow, Jinjiang beef cattle, and Buffalo which were incubated with 2.95 μM PZON for 72 h (**Fig. 4A**). Comparing to DMSO control, methane production was markedly reduced after 24 h of incubation (**Fig. 4B**). After 72 h, methane suppression in the rumen microbial communities of Xiangxi beef cattle, Holstein dairy cow, Jinjiang beef cattle and buffalo reached 95.36%, 73.15%, 86.02%, and 94.38%, respectively (**Fig. 4B**).

To assess potential metabolic shifts induced by PZON, we profiled volatile fatty acids (VFAs) in the fermentation samples. VFAs profiles remained largely unchanged across different rumen microbiomes (**Supplementary Fig. 4A**), except that PZON treatment reduced the production of acetate in the Xiangxi cattle microbiota (**Fig. 4C, Supplementary Fig. 4B-C**). Together, these results indicate that PZON acts as a potent methane inhibitor in multiple ruminant species without influencing the fermentative function.

To further investigate the response of the microbial community to PZON, we performed quantitative PCR (qPCR) and metagenomic sequencing. PZON treatment led to a reduction in methanogen abundance as inferred from 16S rRNA gene copy numbers (**Fig. 4D**), while the total bacterial load remained largely unchanged (**Supplementary Fig. 5**). Notably, the most significant reduction in methanogens occurred in rumen microbiomes from Holstein cow and Jinjiang cattle (**Fig. 4D**). Metagenomic analysis showed that neither alpha diversity nor principal coordinate analysis (PCoA) indicated substantial alterations in the overall microbial community structure following PZON treatment (**Fig. 4E-F**), although the rumen microbiome of Jinjiang cattle showed greater susceptibility to PZON-induced structural perturbations (**Fig. 4E-F**). To gain deeper insight into compositional changes, we conducted differential abundance analysis using Microbiome Multivariable Associations with Linear Models (MaAsLin2) [42]. Rumen methanogenesis is predominantly mediated by archaea belonging to the phyla *Euryarchaeota* — particularly *Methanobacteriota*, largely represented by the genus *Methanobrevibacter* — and *Thermoplasmatota*, the latter comprising methylotrophic methanogens of the order *Methanomassiliicoccales*[43]. As expected, at the phylum level, only *Euryarchaeota* and *Thermoplasmatota* that include methanogenic archaea were significantly reduced following PZON treatment (**Fig. 4G**). Collectively, these findings demonstrate that PZON consistently suppresses methanogenesis across rumen microbiomes from diverse ruminant host, primarily through the reduction of methanogenic archaea, while preserving overall microbial diversity and fermentative functionality.

Discussion

In summary, we identified a novel MCR inhibitor with sub-nanomolar potency through an AI-assisted virtual screening pipeline followed by structure-based rational optimization. The compound effectively suppressed methanogenesis without impairing fermentative activity or causing major shifts in the overall microbial community structure. Its strong anti-methanogenic efficacy, combined with minimal cytotoxicity, highlights the potential of this inhibitor as a targeted methane mitigation agent for application in the livestock industry.

The identification of lead compounds for a specific molecular target is a critical step in the drug discovery process and very often the rate-limiting step [44]. Our study addresses this challenge by employing an AI-based ultra-fast virtual screening strategy, which represents a significant improvement in efficiency and coverage for compound screening. Our approach leverages a deep learning framework that integrates hybrid molecular fingerprint embeddings, combining graph neural network-based representation (NYAN) with Morgan fingerprint, to train a multiple layer perceptron (MLP) regression model for predicting molecular docking scores [34, 35]. This methodology enabled the comprehensive screening of a large library containing over 23 million molecules within a drastically shortened timeframe (less than 3 days), substantially reducing computational demands and time costs compared to conventional docking methods [25]. More importantly, the workflow demonstrated a robust capability to explore under-explored regions of the chemical space [24], successfully identifying non-classical and structurally novel compounds that diverge significantly from known MCR inhibitor scaffolds [40]. Therefore, the development of AI-assisted molecular representation learning and ultra-fast virtual screening frameworks is transforming the discovery of small-molecule inhibitors, enabling efficient, scalable, and mechanism-guided exploration of vast chemical spaces far beyond the reach of conventional methodologies.

Recent advances in gut microbiome research have increasingly focused on elucidating underlying mechanisms and leveraging microbial modulation to improve host phenotypes [2, 3, 18, 45]. Consequently, the ability to precisely regulate specific metabolic functions within complex microbial communities has become particularly critical[6]. These developments collectively highlight the growing demand for efficient, rapid, and accurate strategies to design and implement targeted microbiome interventions [6, 7, 46]. By successfully applying AI-assisted ultra-fast screening platform to address the need of ruminant methane mitigation, we identified a sub-nanomolar inhibitor of methanogenic metabolism with potency comparable to known MCR inhibitors such as BES and halogenated compounds[20]. PZON represents an attractive feed-additive candidate because it is composed solely of C, H, O, and N, in contrast to bromine-containing inhibitors[20, 47]. Such a halogen-free composition would minimize toxicity and residue concerns in livestock production[16, 20]. Moreover, the framework presented in our study provides a generalizable approach for rapidly identifying small-molecule inhibitors capable of modulating metabolic pathways in complex microbial ecosystems. This paradigm holds potential for broader applications in microbiome-directed interventions, such as human gut microbiome modulation[45].

This study has several limitations. While our AI-assisted workflow facilitated hit discovery and lead optimization, it does not take into account properties such as stability, solubility and the difficulty of synthesis. Also, fermentation experiments of the rumen microbiome were performed *in vitro*, which do not fully reproduce the physiological complexity of live ruminants.

Finally, for future applications as ruminant feed additives, the environmental persistence, safety profile and ecological impacts of PZON and its derivatives need to be fully evaluated. Future investigations *in vivo* are warranted to establish bioavailability, efficacy, safety and dosing regimens.

Methods

Protein structure and small-molecule libraries used in virtual screening

The MCR crystal structure (RCSB PDB: 1HBN) from *Methanothermobacter thermautotrophicus* was used as the receptor in all docking and screening. In this structure, the two cofactor molecules F430 (nickel porphinoid F430) and TP7 (coenzyme B) were kept for a more accurate inhibitor-receptor modelling and removing all crystal water molecules, ions and other solvent molecules. To keep the integrity of the binding pocket, the 6 heteromers were kept as a whole. The protein molecules in this receptor structure were prepared by prepare_receptor4.py in MGLTools and the cofactor molecules were prepared by prepare_ligand4.py and then concatenated with the protein molecules for the following AutoDock Vina based docking and screening.

In this study a combined library with around 23 million molecules was aggregated from commercially accessible molecules libraries ZINC20 Now (maximal coverage, 21.45M)[36], ChemBridge (lead-like, 1.3M, https://chembridge.com/), COCONUT (natural products, 500k)[38], and Enamine REAL Space (focused diversity, 26000, https://enamine.net/). The molecules in the screening library were prepared by generating energy optimized 3-dimensional conformations using Rdkit toolkit and hydrogen atoms were also added at pH=7.4. For molecule docking, the molecules were all processed by prepare_ligand4.py script in MGLTools.

Deep learning accelerated virtual screening with iterative model training

We implemented an iterative deep learning based ultra-fast virtual screening pipeline combining hybrid molecular fingerprint embeddings and multilayer perceptron regression to accelerate large-scale virtual screening against MCR (**Fig. 1B-C**). Firstly, we randomly selected molecules (docking training dataset) from the combined library and preformed a traditional virtual screening to calculate all the binding energies (in unite kcal/mol) between each molecule with the receptor using iDock (a fork version of AutoDock Vina), with docking box size = 1.2 nm and exhustiveness=32 (**Fig. 1B**). Next, a molecule-activity MLP regression model was trained to learn the Vina docking score (generated by iDock software) (**Fig. 1C**). In this stage, for each molecule, a NYAN encoder-decoder molecule fingerprint embedding (with 64 dimensions) and the morgen fingerprint (512 dimensions) were combined to train a MLP model using scikit-learn toolkit with Adam optimizer and an initial learning rate 0.001 [24]. For this MLP model, the ReLU activation was applied for all hidden layers (1024, 512, 256, 128 and 64 neurons, respectively) (**Fig. 1C**). A test set composed by 50k molecules were randomly picked and docked to obtain their binding energies (AutoDock Vina docking scores). During the iterative training process (with maximum 50 iterations), in each iteration, 10k molecules were randomly selected and docked, and then trained (together with the molecules selected in

previous iterations) by the MLP model (**Fig. 1B**). The performance of the model is evaluated by the independent test set using different metrics such as RMSE, the coefficient of determination, Pearson's correlation and Spearman's correlation. With this trained best MLP model selected by optimal coefficient of determination, all the binding energies of the molecules in the combined virtual screening library were predicted (**Fig. 1B, Supplementary Fig. 1A**). For the top-ranked 0.1 million molecules (**Supplementary Fig. 1A**), a OnionSFCT+Vina based virtual screening protocol[35] was adopted to find the most possible inhibitors with box-size=1.2 nm and other default parameters. The top-ranked molecules with lowest SFCT+Vina (where the weight factor for SFCT is 0.8 as indicated by the original research) scores were visually inspected to evaluate the existence of the potential important binding patterns, and to purchase molecules for experiment validation (**Supplementary Fig. 1A, Supplementary Table 1**).

Structural similarity analysis

Structural relationship between the molecules to natural substrate methyl-CoM were analyzed by a three-dimensional molecular similarity analysis. All molecular structures were first converted into RDKit molecular objects, and circular fingerprints were generated using the Morgan algorithm. For each molecule, we generated a binary fingerprint vector via RDKit. Pairwise structural similarities between molecules were then quantified using the Tanimoto coefficient, which measures the degree of overlap between two molecular fingerprints and yields a similarity score. The similarity score ranges from 0 to 1, with higher values indicating a closer structural resemblance to methyl-CoM.

Structural diversity visualization

Structural diversity was assessed using t-distributed Stochastic Neighbor Embedding (t-SNE), a dimensionality-reduction method that preserves local similarity relationships when projecting high-dimensional molecular representations into a two-dimensional space[48].

All molecular structures were first parsed using RDKit. Molecules obtained from structural databases were imported from PDB files, and compounds in the candidate library were loaded from SDF files. MACCS fingerprints were then generated for each molecule to represent structural features in a high-dimensional vector space. The resulting fingerprint matrix was embedded into two dimensions using t-SNE to visualize both global and local structural relationships among molecules. To further delineate structural organization, K-means clustering was applied to the t-SNE coordinates to identify groups of structurally related compounds. Molecules analyzed in this study and previously reported MCR inhibitors were highlighted on the t-SNE map to facilitate direct comparison of their distribution within the structural space, thus providing insights for compound prioritization and structure – activity relationship

377 analysis.
378
379 ## Lead compound optimization
380 To expand upon the initial hits, a virtual library of structurally related compounds (Tanimoto
381 similarity $\geqslant$ 0.5) was retrieved from the Molport webserver. After computational docking, the
382 resulting binding poses were carefully examined. Candidate selection for experimental testing
383 was guided by visual inspection, prioritizing molecules that preserved the key structural
384 features of the lead compound. These features included electronegative pharmacophores
385 capable of coordinating metal ions, as well as substituents positioned to engage in critical
386 interactions with residues such as Arg120 on chain C, consistent with our prior design strategy.
387 The prioritized compounds were subsequently purchased for experimental validation
388 (**Supplementary Table 2**).
389
390 ## Rumen fluid collection and processing
391 Ruminal fluid used in this study were collected from Xiangxi beef cattle, Holstein dairy cow,
392 Jinjiang cattle and Buffalo. All procedures involving animals were approved by the Animal Use
393 and Health Committee of Zhejiang University (Hangzhou, China, No. 12410). Rumen content
394 of Holstein dairy cow, Jinjiang cattle and buffalo were collected using a stomach tube before
395 morning feeding, with the first 150 mL of rumen contents discarded to avoid saliva
396 contamination. Rumen content of Xiangxi beef cattle with permanent ruminal fistulas were also
397 collected before morning feeding. Immediately after collecting from rumen, the contents were
398 transfer to a 50 mL Corning tube containing 8 mL 80% glycerol (v/v, in anaerobic and sterile
399 phosphate-buffered saline, with 0.5% L-cysteine hydrochloride) making total volume up to 40
400 mL, homogenized by shaking and transported to the laboratory in dry-ice. When arrived
401 laboratory, samples were rapidly thawed at 37 ℃ water bath and immediately transferred to an
402 anaerobic workstation (85% N_2, 10% H_2, and 5% CO_2, COY). 40 mL of each rumen sample
403 was suspended into 40 mL 50% glycerol (v/v, in sterile phosphate-buffered saline, with 0.1%
404 L-cysteine hydrochloride), homogenized by vertexing, and then filtered with sterile nylon mesh
405 to remove large particles in rumen content. Aliquots of the suspension were stored in sterile
406 cryogenic vials and frozen at -80 °C for long-term storage until processing for DNA extraction
407 and culturing so that the rumen-derived community could be revived (thawed) for repeatable
408 experiments.
409
410 ## *In vitro* fermentation of ruminal microbiome
411 The *in vitro* fermentation was cultured using Balch tube (18 x 150 mm) with butyl stoppers and
412 aluminum seals. The anaerobic medium was prepared from 400 mL of distilled water, 0.1 mL
413 of liquid A, 200 mL of liquid B and 200 mL of liquid C [49], and a modification with 100 mL

clarified rumen fluid (ELITE-MEDIA). Liquid A consisted of 13.2 g $CaCl_2 \cdot 2H_2O$, 10.0 g $MnCl_2 \cdot 4H_2O$, 1.0 g $CoCl_2 \cdot 6H_2O$, 8.0 g $FeCl_2 \cdot 6H_2O$ dissolved in 100 mL distilled water containing 0.1% L-cysteine hydrochloride. Liquid B contained 5.7 g Na_2HPO_4, 6.0 g KH_2PO_4, 0.6 g $MgSO_4 \cdot 7H_2O$ in 1 L distilled water supplemented with 0.1% L-cysteine hydrochloride. Liquid C contained 4.0 g NH_4HCO_3, 35 g $NaHCO_3$ in 1 L distilled water with 0.1% L-cysteine hydrochloride. Aliquots of rumen microbiome (50 µL) were inoculated into 9.25 mL anaerobic medium containing 10 µL compound stock or DMSO (vehicle control) for biological triplicate in the anaerobic workstation. Cultures were incubated at 37 °C for 72 h with shaking at 150 rpm. Candidate compound and sodium 2-bromoethanesulfonate (BES, Aladdin, Cat#B105365) were solvent in DMSO for 1000× stock.

After incubation, 50 µL gas sample was collected from the headspace of Balch tube for CH_4 determine, using a 500 µL sample-Lock Hamilton syringe inserting the butyl stoppers with guided of 18 G syringe needle [50]. Liquid sample was collected from each tube and frozen at -80 °C for further analysis. The relative methane production was quantified as the percentage of methane in the treatment group relative to the DMSO control. The inhibition ratio was calculated as the percentage reduction in methane production in treated cultures compared with the DMSO control.

Quantification of methane and volatile fatty acids by gas chromatography

The methane concentrations in headspace gas samples were quantified using gas chromatography (GC, Agilent 8890, Santa Clara, CA, USA) equipped with HP-PLOT Q column (30 m × 0.53 mm × 40 µm film thickness) and a flame ionization detector (FID), with nitrogen as the carrier gas, following a protocol adapted from reference [50]. The oven temperature was maintained at 80 °C for 30 min, and the FID was operated at 200 °C. Calibration curves were generated using methane standard gases across a broad concentration range, including 50 µL injections of 98.8 ppm, 2.1% and 100% methane, as well as 10 µL injections of 2.1% and 100% methane.

Volatile fatty acids (VFAs) in the post-fermentation liquid samples were analyzed using gas chromatography as previously described[51]. For the sample preparation, fermentation broth was centrifugated at 15,000 × g for 10 min at 4 °C, and 300 µL of the resulting supernatant was acidified with 60 µL 25% (w/v) metaphosphoric acid. The mixture was vortexed for 1 min until homogeneous and stored at -20 °C for 8 h to precipitate proteins. Samples were subsequently thawed on ice and centrifuged again under the same conditions. The clarified supernatant was transferred to a clean tube, mixed with an equal volume of ethyl acetate, and vortexed vigorously for 1 min to extract VFAs. The organic phase was collected and transferred into a glass insert within a GC vial and capped tightly. Quantification of acetic, propionic, iso-butyric,

butyric, iso-valeric and valeric acids was performed by Agilent 8890 GC equipped with DB-Wax capillary column (30 m × 0.53 mm × 1 μm film thickness) and FID. Total VFAs concentrations were calculated as the sum of analyzed acetic, propionic, iso-butyric, butyric, iso-valeric and valeric acids.

Culture of *Methanobrevibacter ruminantium*

Methanobrevibacter ruminantium M1 (DSM 1093) was cultivated in 10 mL of DSM 1523 medium in Balch tubes (18 x 150 mm) under strictly anaerobic conditions. Tubes were sealed with butyl rubber septa secured by aluminum crimps, and the headspace was filled with an H_2/CO_2 mixture (80:20, v/v). After inoculation with 0.5 mL of an actively growing preculture, cultures were incubated horizontally at 37 °C without shaking. Methanogenesis and growth kinetics were monitored daily for 7 days measuring CH_4 production and optical density at 600 nm (OD_{600}). Experimental treatments were conducted in biological triplicate and included:(i) Control, Medium + 1% (v/v) DMSO (solvent control); (ii) Positive control; 1 mM BES; and (iii) PZON at final concentrations of 0.33, 3.33, or 30.0 μM. BES and PZON were dissolved in DMSO to prepare 100× stock solutions and were added to the cultures at the time of inoculation.

In vitro toxicity evaluation of PZON

Human intestinal epithelial cells (Caco-2, ATCC HTB-37) and bovine renal epithelial cells (MDBK, ATCC CCL-22) were cultured in 96-well plates using Dulbecco's Modified Eagle Medium basic (DMEM, Gibco, Cat#11965092) supplemented with 10%(v/v) Fetal Bovine Serum (FBS, Gibco, Cat#10437028) and 1% (v/v) Penicillin-Streptomycin (Solarbio, Cat#P1010). After seeding, cells were allowed to attach for 12 h at 37 °C in a humidified atmosphere containing 5% CO_2. Following attachment, the culture medium was replaced with fresh DMEM containing 10% FBS and serial dilutions of PZON (final concentrations: 100, 50, 25, 12.5, 6.25, 3.125, 1.563 μM). Vehicle control wells received 0.1% DMSO, and SDS (2%) was used as a positive control for cytotoxicity. Cells were incubated with compounds for 24 h under the same conditions. Cell viability was assessed using the MTT assay according to the manufacturer's instructions.

DNA extraction of fecal samples and metagenomic library preparation

DNA of fermentation fluid samples was extracted using the QIAamp PowerFecal DNA kit (Qiagen, Cat#51804-50) following standard manufacturer procedures. DNA samples were resuspended in Buffer C6 and quantitated using the NanoDrop (ThermoFisher Scientific). Metagenomic sequence libraries were prepared with at least 2 μg of total DNA using The Hieff NGS® OnePot Pro DNA Library Prep Kit V2® (Yeasen, Cat#12195ES96), following the

manufacturer's instructions. The resulting library DNA was cleaned up and size-selected with Hieff NGS® DNA Selection Beads (Yeasen, Cat#12601ES56), and quantified using the dsDNA High Sensitivity kit on a Qubit (Thermo Fisher). Libraries were further pooled together at equal molar ratios, and the purity and library length distribution were assessed using Bioanalyzer High Sensitivity DNA Kit (Agilent). Sequencing was performed on the DipSeq (150 bp paired-end reads; BGI Genomics Co., Ltd.) platform.

Quantitative PCR analysis of bacterial and methanogen abundance

The abundances of total bacteria and methanogens in the fermentation samples were quantified by quantitative PCR (qPCR) using specific primers (**Supplementary Table 3**) as previously describe [52]. Absolute abundances were calculated from standard curves generated for each target group.

Bioinformatics analysis

Metagenomic raw sequences were subjected to quality trimming using Fastp (v0.23.2). Reads aligning to the host genome were identified and removed using Bowtie2 (v2.3.5). Taxonomic classification was performed using Kraken2 (v2.1.2) against the GTDB-r207 database. The number of sequences reads assigned to individual species was computed using Bracken (v2.6.1). The resulting species abundance table was normalized to relative abundance for downstream diversity analyses. Alpha diversity (including Shannon and Simpson indices) and beta diversity were assessed using the vegan package (v2.6.4) in R (version 4.2.0). Beta diversity was calculated based on Bray-Curtis dissimilarity, and the resulting distance matrix was visualized using principal coordinate analysis (PCoA). Permutational multivariate analysis of variance (PERMANOVA) was applied to test for significant differences in community structure among experimental groups.

Statistical analysis

Statistical analyses were performed using R (version 4.2.0). Normality and homogeneity of variances were assessed using Shapiro–Wilk test and Levene's test, respectively. To test difference of molecule screening and optimization, t-test compared to DMSO control with Bonferroni post hoc test was used. To test differences of two factors (Treat × Group), two-way ANOVA with Bonferroni's multiple comparisons test or nonparametric Scheirer–Ray–Hare test followed by Bonferroni correction was used based on normality. To identify microbial taxa significantly associated with experimental factors, we performed multivariate statistical analysis using MaAsLin2 (Microbiome Multivariable Associations with Linear Models, v2.1.0)[42] in R. The normalized phylum-level relative abundance matrix was used as the response variable, while experimental metadata such as treatment group and individual host factors were included as fixed effects as appropriate. Significance levels were denoted as

follows: *$p < 0.05$, **$p < 0.01$, and #$p < 0.1$ (marginally significant). All experiments were performed in biological triplicates (n = 3), as indicated in the figure legends.

Data availability

All sequencing data generated in this study have been submitted to China National Center for Bioinformation database (accession number PRJCA049369).

Code availability

The code for AI-based screening workflow can be found on GitHub (https://github.com/zelixirSH/zFastVS.git).

Acknowledgments

The authors express their sincere gratitude to Prof. Weijun Shen (College of Animal Science and Technology, Hunan Agricultural University), Assoc. Prof. Yongqing Guo (College of Animal Science, South China Agricultural University), Dr. Chao Chen (College of Animal Science and Technology, Huazhong Agricultural University) and Dr. Chuanhui Xu (Institute of Animal Husbandry and Veterinary, Jiangxi Academy of Agricultural Science) for their generous support and for providing the essential materials used in the *in vitro* fermentation of the ruminal microbiome. This work was funded by National Natural Science Foundation of China (32401321, 32225003), China Postdoctoral Science Foundation (2023M743686) and Guangdong Major Project of Basic and Applied Basic Research (2023B0303000017).

Contributions

L.D., X.X., S.W., and L.Z. conceived the project. X.X. and J.Z. conducted the experiments. L.Z. and G.Z. developed the AI-assisted virtual screening and conducted hit optimization. X.X. analyzed the experimental data. H.L. and H.N. contributed to the analysis of metagenomic data. H.L., R.C. and H. S provided critical feedback on data analysis. H. S participated in discussions and provided constructive advice on experimental design and data interpretation. X.X., L.Z. and G.Z. wrote the initial draft. L.D., S.W. and M.L. supervised the project and reviewed the manuscripts. All other authors discussed the results and approved the manuscript.

Competing interests

L.D. is the cofounder of SynBiome Biotech, Ltd. S.W. and L. Z. are the cofounders of Zelixir Biotech Co., Ltd. The authors have filed patent applications on small molecule inhibitors reported in this paper.

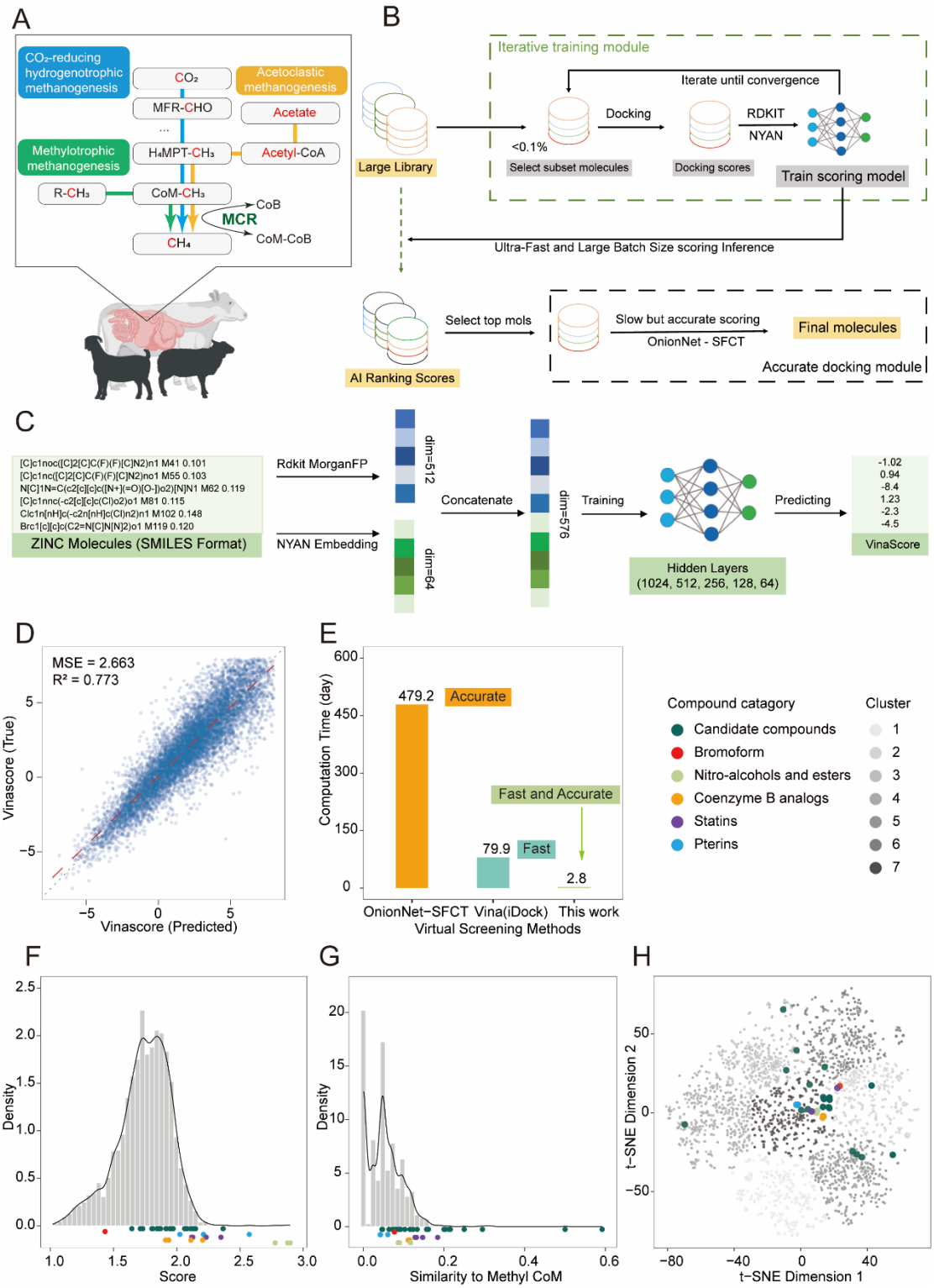

Figure 1. AI-assisted virtual screening of small molecule inhibitors for rumen methane metabolism.

A) Primary methanogenesis pathways in ruminal gastrointestinal microbiome. Major metabolites are boxed, and carbon groups involved in methanogenesis reaction are indicated in red.

B) Iterative deep docking pipeline combining hybrid molecular fingerprint embeddings and

multilayer perceptron regression. A small subset of compounds randomly sampled from a large virtual library was docked to the target protein using low-accuracy docking (iDock, a speed-up version of Vina) to generate training labels (vina scores). Molecular features were computed via ligand-based descriptors (e.g., NYAN and fingerprints), and a multiple-layer perceptron (MLP) network was trained to predict docking scores. New batches of molecules were iteratively sampled, docked, and added to the training set until the model converged. The final trained model was then used to predict docking scores for the entire library. Top-ranked compounds were selected for high-accuracy docking (OnionNet+SFCT), followed by expert curation and *in vitro* activity testing.

C) Molecular Label (Vina Score) Prediction Model. The SMILES representation of small molecules is processed through a graph neural network model (NYAN, dimension 64) and molecular fingerprints (dimension 512) for molecular feature encoding and encoding fusion. The fused features are then fed into the MLP model with hidden layers of dimensions (1024, 512, 256, 128, 64). Finally, an output layer (dimension = 1) produces a single label, which is the predicted Vina score.

D) Correlation between predicted and true Vina scores. A scatter plot of the predicted versus experimental Vina scores for the validation set, generated by the top-performing model selected by highest R^2. MSE, Mean Square Error.

E) The benchmarked computational time for three pipelines based on a supercomputing cluster composed by 50 AMD EPYC 7642 CPU workstations (with 192 cores and 377GB memory per workstation). The two-stage ultra-fast pipeline enables a 28.5 times and a 171.1 times speed-up comparing to pure iDock and OnionNet+SFCT virtual screening.

F-H) Distribution of SFCT+Vina scores, similarity to methyl-CoM and t-SNE visualization for top-ranked molecules predicted by the model, compared with known MCR inhibitors [40]. Compound classes are color-coded as follows: candidate compounds (dark cyan), bromoform (red), nitro-alcohols and esters (desaturated green), coenzyme B analogs (orange), pterins (blue), and statins (purple).

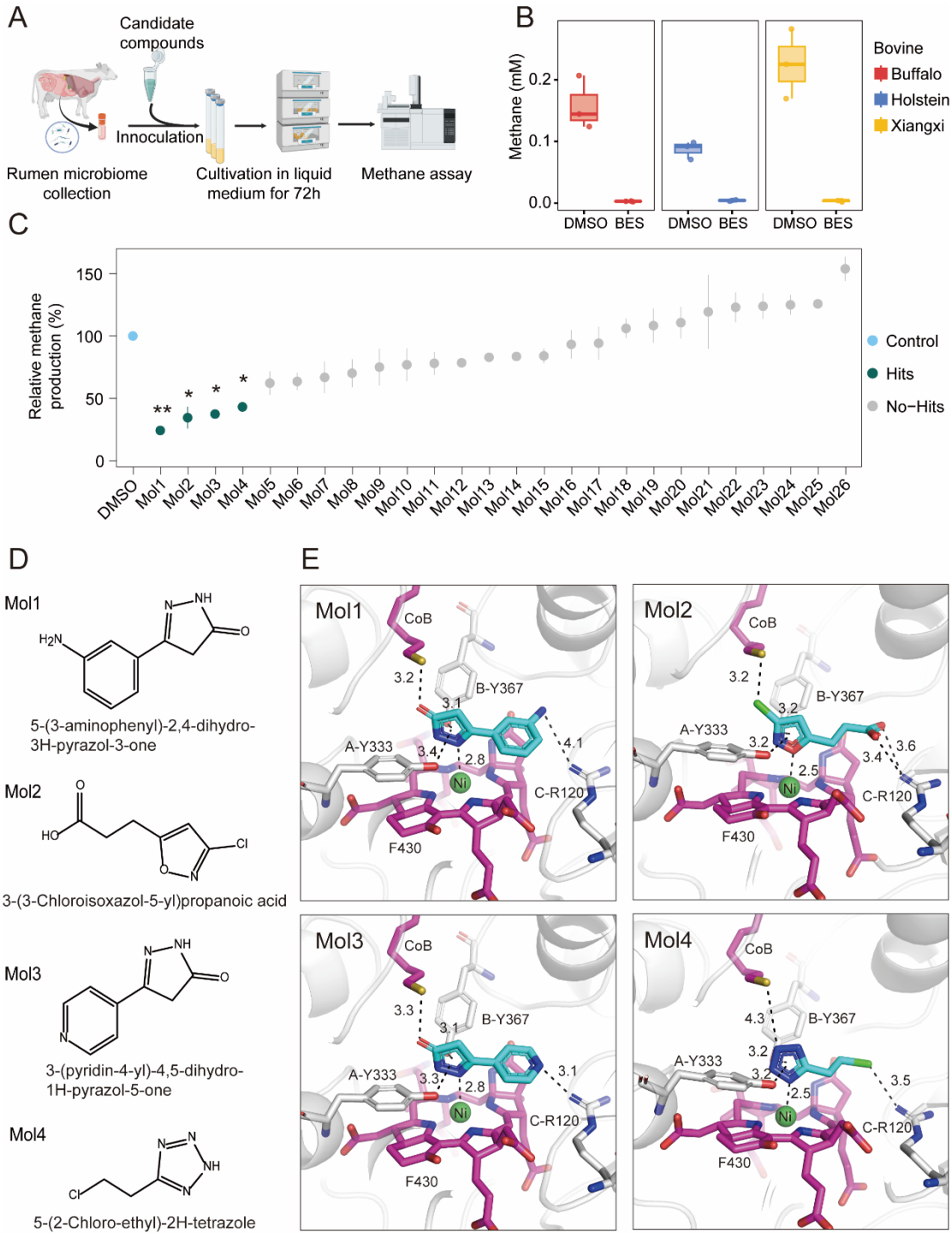

Figure 2. Experimental validation of candidate MCR inhibitors.

A) Workflow of *in vitro* rumen microbial fermentation.

B) Effect of BES on methane production at 10 μM following *in vitro* rumen microbiome fermentation. Rumen microbial inoculum was collected from Buffalo (red), Holstein dairy cattle (blue), and Xiangxi beef cattle (yellow). Box plots show the median (center line), interquartile range (IQR; box boundaries), and full data range (whiskers). Assays were performed in triplicate (n = 3). BES, sodium 2-bromoethanesulfonate.

C) Relative methane production of candidate compounds at 98.31 μM. Data are presented as

602 mean ± SEM, based on triplicate assays (n = 3). Statistical significance was determined by t-
603 test comparing to DMSO control with Bonferroni post hoc test (compared with DMSO control,
604 $*p < 0.05$, $**p < 0.01$).
605 **D)** Chemical structures of hit compounds.
606 **E)** Docking poses of hit compounds at the MCR active site. Interactions are shown as dashed
607 lines, with numerical labels indicating interatomic distances (Å). Cofactors are shown in
608 magenta (carbon atoms), the Ni^{2+} ion in green, and the hit molecules in cyan (carbon atoms).
609 Interacting residues in MCR are depicted in gray (carbon atoms), with close contacts
610 highlighted by dashed lines and corresponding interatomic distances (Å). Oxygen, nitrogen,
611 chlorine and sulfur atoms are colored red, blue, green, and yellow, respectively.
612

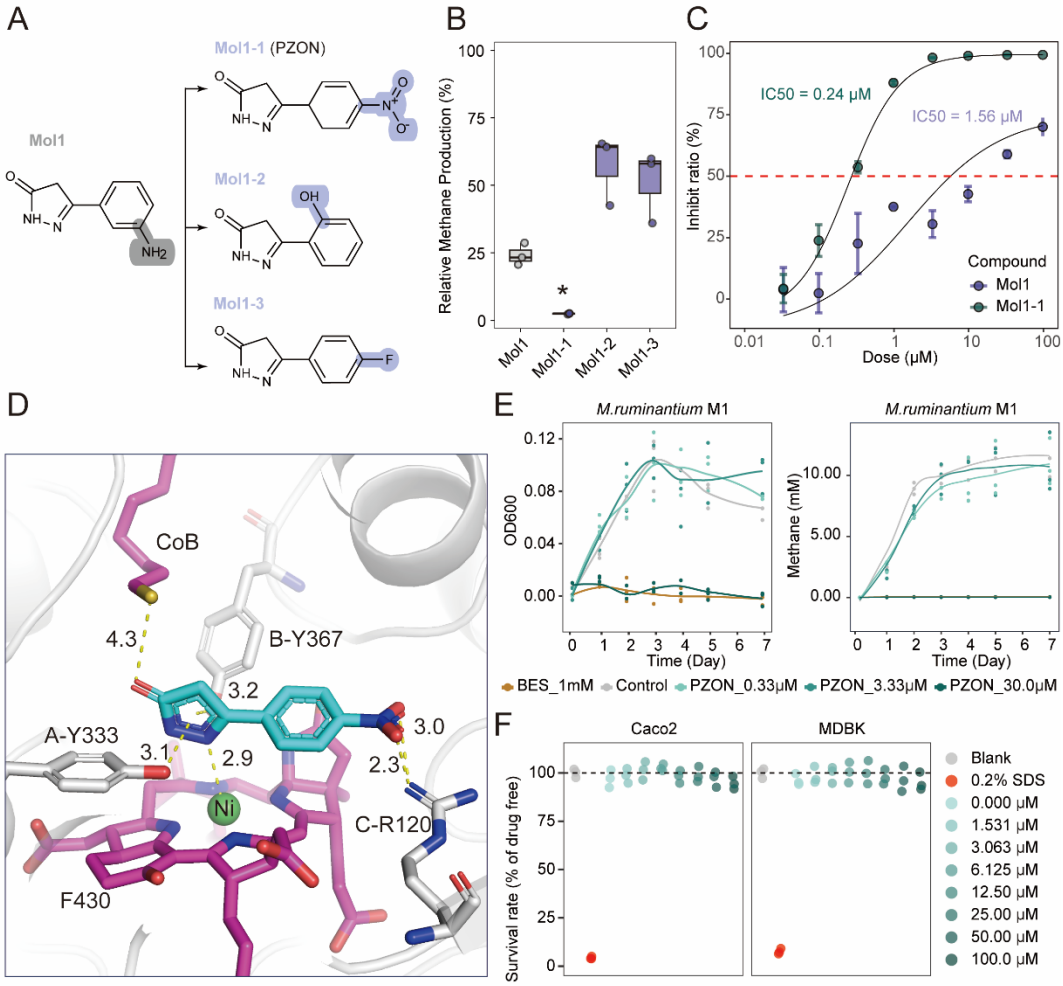

Figure 3. Identification of a potent MCR inhibitor PZON via structure optimization.

A) Rational structural optimization of Mol1. Gray represents the group planned for modification, and purple indicates the modified group.

B) Effect of optimized Mol1 derivatives on methane production at 98.31 μM after 72 h of *in vitro* rumen microbial fermentation. Statistical significance was determined by t-test comparing to Mol1 group with Bonferroni post hoc test (compared with control, *$p < 0.05$). Data were show in boxplot, and each dot represents a technical replicate, and assays were performed in triplicate (n = 3).

C) Concentration-response curves of PZON and Mol1. Inhibition data across a range of concentrations were fitted using a four-parameter log-logistic model. Data are presented as mean ± SEM, based on triplicate assays (n = 3).

D) Docking pose of compound PZON at the MCR active site. Interactions are depicted as dashed lines with numerical labels indicating interatomic distances (Å). In the enzyme–substrate complex, PZON (cyan) and the cofactor F430 (magenta) are accommodated within the active pocket. Cofactors are shown in magenta (carbon atoms), the Ni^{2+} ion in green, and the hit molecule in cyan (carbon atoms). Interacting MCR residues are shown in gray (carbon atoms), with close contacts highlighted by dashed lines and interatomic distances (Å). Oxygen,

nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively.

E) Impact of PZON on growth and methane production of *Methanobrevibacter ruminantium* M1 (DSM 1093). Lines represent group mean.

F) Cytotoxicity of PZON in cultured human colon carcinoma cell line (Caco-2) and Madin-Darby bovine kidney (MDBK) cells. Each PZON concentration showed $p > 0.05$ comparing to Blank based on t-test with Bonferroni post hoc test. Each dot denotes an individual measurement. Assay were performed in triplicates (n = 3).

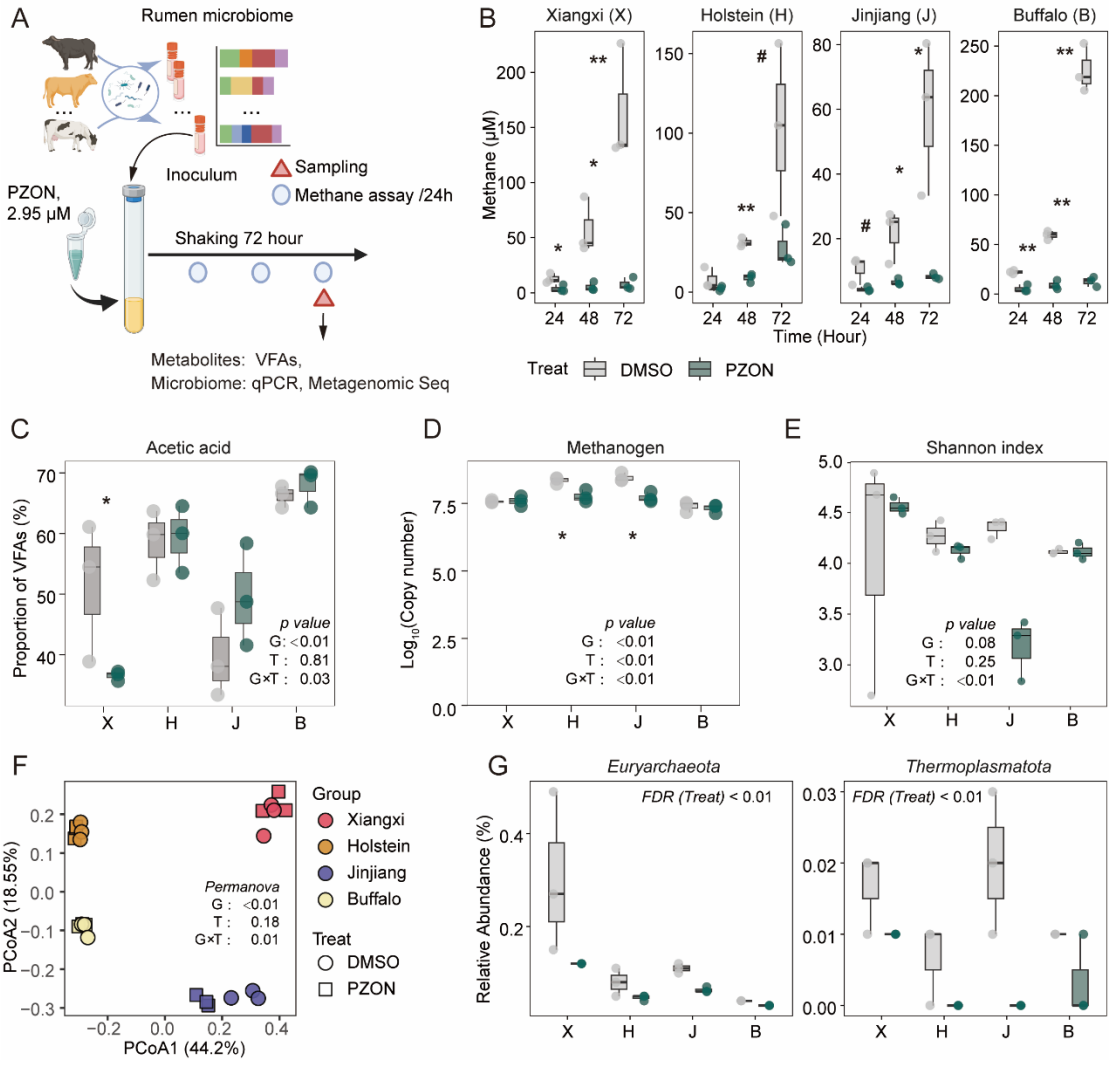

Figure 4. Evaluation of MCR inhibitor PZON using *in vitro* fermentation of rumen microbiome.

A) Workflow of rumen microbiome collection and *in vitro* rumen microbial fermentation.

B) Impact of PZON on methane production in the *in vitro* fermentation of rumen microbiome culture for 72 h.

C) Proportion of acetic acid in volatile fatty acids in cultures of rumen microbiome for 72 h.

D) Absolute abundance of methanogen in cultures of rumen microbiome for 72 h. Absolute abundance was estimated using the qPCR assay.

E-F) Alternation of rumen microbiome in response to PZON treatment *in vitro*. Shannon index and PCoA (principal coordinate analysis) based on Bray-Curtis matrices both showed no significant differences between the cultures with the DMSO treatment and the PZON treatment.

G) Relative abundance of *Euryarchaeota* and *Thermoplasmatota* in cultures of rumen microbiome for 72 h. Relative abundance was estimated using Kraken2. Effect of PZON on phylum-level taxonomy was analyzed by MaAsLin2 and FDR < 0.05 considered as significantly affected by PZON treatment.

Statistical significance of B), C) and D) were determined by two-way ANOVA with Bonferroni

post hoc test. Statistical significance of E) was determined by Scheirer–Ray–Hare test with Bonferroni post hoc test under each group. Statistical significance of F) was determined by PERMANOVA based on the Bray-Curtis dissimilarity. #p <0.1, *p < 0.05, **p < 0.01. Data were show in boxplot, and each dot represents a technical replicate, and assays were performed in triplicate (n = 3).

Reference

1. Peng, H., Ruiz-Moreno, A.J. & Fu, J. Multi-dimensional metagenomics. *Nature Reviews Bioengineering* (2025).
2. Sinha, A.K., Laursen, M.F. & Licht, T.R. Regulation of microbial gene expression: the key to understanding our gut microbiome. *Trends in microbiology* **33**, 397-407 (2025).
3. Joos, R. *et al.* Examining the healthy human microbiome concept. *Nature Reviews Microbiology* **23**, 192-205 (2025).
4. Sheth, R.U., Cabral, V., Chen, S.P. & Wang, H.H. Manipulating Bacterial Communities by in situ Microbiome Engineering. *Trends in Genetics* **32**, 189-200 (2016).
5. Noecker, C. & Turnbaugh, P.J. Emerging tools and best practices for studying gut microbial community metabolism. *Nature Metabolism* **6**, 1225-1236 (2024).
6. Woo, A.Y.M. *et al.* Targeting the human gut microbiome with small-molecule inhibitors. *Nat Rev Chem* **7**, 319-339 (2023).
7. Volpe, M.R. *et al.* A small molecule inhibitor prevents gut bacterial genotoxin production. *Nat Chem Biol* (2022).
8. Wallace, B.D. *et al.* Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. *Science* **330**, 831-835 (2010).
9. Shima, S., Huang, G., Wagner, T. & Ermler, U. Structural Basis of Hydrogenotrophic Methanogenesis. *Annu Rev Microbiol* **74**, 713-733 (2020).
10. Stewart, R.D. *et al.* Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery. *Nature biotechnology* **37**, 953-961 (2019).
11. Xie, F. *et al.* An integrated gene catalog and over 10,000 metagenome-assembled genomes from the gastrointestinal microbiome of ruminants. *Microbiome* **9** (2021).
12. Peng, X. *et al.* Genomic and functional analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiomes. *Nat Microbiol* **6**, 499-511 (2021).
13. Matthews, C. *et al.* The rumen microbiome: a crucial consideration when optimising milk and meat production and nitrogen utilisation efficiency. *Gut microbes* **10**, 115-132 (2018).
14. Garcia, P.S., Gribaldo, S. & Borrel, G. Diversity and Evolution of Methane-Related Pathways in Archaea. *Annu Rev Microbiol* **76**, 727-755 (2022).
15. Duin, E.C., Prakash, D. & Brungess, C. Methyl-coenzyme M reductase from Methanothermobacter marburgensis. *Methods in enzymology* **494**, 159-187 (2011).
16. Xie, X. *et al.* Mitigating enteric methane emissions: An overview of methanogenesis, inhibitors and future prospects. *Animal Nutrition* **21**, 84-96 (2025).
17. Wang, K., Xiong, B. & Zhao, X. Could propionate formation be used to reduce enteric methane emission in ruminants? *Science of The Total Environment* **855** (2023).
18. Mizrahi, I., Wallace, R.J. & Moraïs, S. The rumen microbiome: balancing food security and environmental impacts. *Nature Reviews Microbiology* **19**, 553-566 (2021).
19. Patra, A.K. & Puchala, R. Methane mitigation in ruminants with structural analogues and other chemical compounds targeting archaeal methanogenesis pathways. *Biotechnology advances* **69** (2023).
20. Fouts, J.Q., Honan, M.C., Roque, B.M., Tricarico, J.M. & Kebreab, E. Enteric methane

mitigation interventions. *Translational animal science* **6**, txac041 (2022).

21. Sadybekov, A.V. & Katritch, V. Computational approaches streamlining drug discovery. *Nature* **616**, 673-685 (2023).
22. Tropsha, A., Isayev, O., Varnek, A., Schneider, G. & Cherkasov, A. Integrating QSAR modelling and deep learning in drug discovery: the emergence of deep QSAR. *Nature reviews. Drug discovery* **23**, 141-155 (2024).
23. Smith, J.S., Isayev, O. & Roitberg, A.E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chemical science* **8**, 3192-3203 (2017).
24. Lyu, J. *et al.* Ultra-large library docking for discovering new chemotypes. *Nature* **566**, 224-229 (2019).
25. Gentile, F. *et al.* Artificial intelligence–enabled virtual screening of ultra-large chemical libraries with deep docking. *Nature Protocols* **17**, 672-697 (2022).
26. Luttens, A. *et al.* Rapid traversal of vast chemical space using machine learning-guided docking screens. *Nature Computational Science* **5**, 301-312 (2025).
27. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).
28. Passaro, S. *et al.* Boltz-2: Towards Accurate and Efficient Binding Affinity Prediction. *bioRxiv : the preprint server for biology* (2025).
29. Chai Discovery, t. *et al.* Chai-1: Decoding the molecular interactions of life. *bioRxiv : the preprint server for biology*, 2024.2010.2010.615955 (2024).
30. Masters, M.R., Mahmoud, A.H. & Lill, M.A. Investigating whether deep learning models for co-folding learn the physics of protein-ligand interactions. *Nature communications* **16**, 8854 (2025).
31. Trott, O. & Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry* **31**, 455-461 (2010).
32. Alhossary, A., Handoko, S.D., Mu, Y. & Kwoh, C.K. Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics (Oxford, England)* **31**, 2214-2216 (2015).
33. Gorgulla, C. *et al.* An open-source drug discovery platform enables ultra-large virtual screens. *Nature* **580**, 663-668 (2020).
34. Lam, H.Y.I. *et al.* Application of variational graph encoders as an effective generalist algorithm in computer-aided drug design. *Nature Machine Intelligence* **5**, 754-764 (2023).
35. Zheng, L. *et al.* Improving protein-ligand docking and screening accuracies by incorporating a scoring function correction term. *Briefings in bioinformatics* **23** (2022).
36. Irwin, J.J. *et al.* ZINC20—A Free Ultralarge-Scale Chemical Database for Ligand Discovery. *Journal of Chemical Information and Modeling* **60**, 6065-6073 (2020).
37. ChemBridge Lead-Like and Drug-Like Compound Database. Available online: https://chembridge.com/screening-compounds/lead-like-drug-like-compounds/.
38. Chandrasekhar, V. *et al.* COCONUT 2.0: a comprehensive overhaul and curation of the collection of open natural products database. *Nucleic acids research* **53**, D634-d643 (2025).

39. Enaine REAL Space. Available online: https://enamine.net/compound-collections/real-compounds.
40. Aryee, R. *et al.* Exploring putative enteric methanogenesis inhibitors using molecular simulations and a graph neural network. *bioRxiv : the preprint server for biology* (2024).
41. Grabarse, W. *et al.* On the mechanism of biological methane formation: structural evidence for conformational changes in methyl-coenzyme M reductase upon substrate binding. *J Mol Biol* **309**, 315-330 (2001).
42. Mallick, H. *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLoS computational biology* **17**, e1009442 (2021).
43. Tang, Y. *et al.* New insights into the enteric methane production based on the archaeal genome atlas of ruminant gastrointestinal tract. *J Adv Res* **74**, 13-24 (2025).
44. Gahbauer, S. *et al.* Docking for EP4R antagonists active against inflammatory pain. *Nature communications* **14** (2023).
45. Wang, K. *et al.* Microbial-host-isozyme analyses reveal microbial DPP4 as a potential antidiabetic target. *Science* **381**, eadd5787 (2023).
46. Challa, A.P. *et al.* Virtual Screening for the Discovery of Microbiome beta-Glucuronidase Inhibitors to Alleviate Cancer Drug Toxicity. *J Chem Inf Model* **62**, 1783-1793 (2022).
47. Ahmad, I., Rawnsley, R.P., Bowman, J.P. & Omede, A.A. Graduate Student Literature Review: Limitations in feeding red seaweed Asparagopsis species for enteric methane mitigation in ruminants. *Journal of dairy science* **108**, 13270-13288 (2025).
48. Maaten, L.v.d. & Hinton, G.E.J.J.o.M.L.R. Visualizing Data using t-SNE. *Journal of Machine Learning Research* **9**, 2579-2605 (2008).
49. Menke, K.H. *et al.* The estimation of the digestibility and metabolizable energy content of ruminant feedingstuffs from the gas production when they are incubated with rumen liquor in vitro. *The Journal of Agricultural Science* **93**, 217-222 (1979).
50. Aldridge, J., Catlett, J., Smith, M. & Buan, N. Methods for Detecting Microbial Methane Production and Consumption by Gas Chromatography. *Bio-Protocol* **6** (2016).
51. Xia, X., Wei, H., Hu, L. & Peng, J. Hydratability and improved fermentability in vitro of guar gum by combination of xanthan gum. *Carbohydr Polym* **258**, 117625 (2021).
52. Liu, H. *et al.* Ecological dynamics of the gut microbiome in response to dietary fiber. *The ISME Journal* **16**, 2040-2055 (2022).