


A chronotherapeutics-applicable multi-target therapeutics based on AI: Example of therapeutic hypothermia

Fei Liu, Xiangkang Jiang, Jingyuan Yang, Jiawei Tao and Mao Zhang 

Corresponding author: Mao Zhang. Tel.: +86-13757119125; E-mail: z2jzk@zju.edu.cn

All authors are first authors. The author Fei Liu is superior to other authors.

Abstract

Nowadays, the complexity of disease mechanisms and the inadequacy of single-target therapies in restoring the biological system have inevitably instigated the strategy of multi-target therapeutics with the analysis of each target individually. However, it is not suitable for dealing with the conflicts between targets or between drugs. With the release of high-precision protein structure prediction artificial intelligence, large-scale high-precision protein structure prediction and docking have become possible. In this article, we propose a multi-target drug discovery method by the example of therapeutic hypothermia (TH). First, we performed protein structure prediction for all protein targets of each group by AlphaFold2 and RoseTTAFold. Then, QuickVina 2 is used for molecular docking between the proteins and drugs. After docking, we use PageRank to rank single drugs and drug combinations of each group. The ePharmaLib was used for predicting the side effect targets. Given the differences in the weights of different targets, the method can effectively avoid inhibiting beneficial proteins while inhibiting harmful proteins. So it could minimize the conflicts between different doses and be friendly to chronotherapeutics. Besides, this method also has potential in precision medicine for its high compatibility with bioinformatics and promotes the development of pharmacogenomics and bioinfo-pharmacology.

Keywords: therapeutic hypothermia, AI (artificial intelligence), drug prospecting, multi-target therapeutics, chronotherapeutics

Introduction

At present, the common strategy of drug development is derived from the paradigm 'One Drug, One Disease'. Therefore, highly selective and potent molecules are being developed for a distinct target [1]. However, the application of this strategy, such as the hybrid drugs or chimeric drugs, analyzes each target individually and could not solve the problem of complex interactions between drugs or targets [2]. Besides, this approach also has disadvantages when dealing with time-specific proteins. These defects limit the application of this strategy.

Take therapeutic hypothermia (TH) as an example. TH can limit the degree of some kinds of injuries in randomized trials [3] and animal experiments [4] and is even the only effective method for some diseases, especially hypoxic-ischemic encephalopathy (HIE). HIE often causes severe neurological sequelae, which is the main reason for the poor prognosis of patients with stroke, shock, carbon monoxide poisoning, cerebral hemorrhage and cardiac arrest [5–7].

In the research based on TH, cold shock proteins, especially cold-induced ribonucleic acid (RNA)-binding protein (CIRP), show a high expression [8] and rapid response [9]. CIRP has been shown to promote the translation of genes involved in deoxyribonucleic acid repair [10, 11], telomerase maintenance [12] and genes associated with the translational machinery [13].

However, if CIRP leaks to the intercellular substance with cell swelling and rupture, it will become a proinflammatory mediator. Extracellular CIRP showed a strong pro-inflammatory effect by binding to the Toll-like receptor 4 (TLR4)-MD2 complex on the surface of macrophage, which leads to the heavier injury of ischemic tissue [14, 15]. So, agonists applied with CIRP can effectively promote cell protection before CIRP leaks out of the cell. After the leakage of CIRP out of the cell, the application of the CIRP antagonist can effectively promote cell protection.

Thus, sophisticated drug discovery approaches, such as pharmacogenomics [16], are more appropriate for this problem. With the development of high-precision protein prediction technologies, especially AlphaFold2 [17] and RoseTTAFold [18], we can obtain the structure of proteins quickly and accurately, and widespread docking becomes possible. However, there is no report analyzing the results of multi-target data. In this article, we provide a multi-targeted drug discovery method and predict the side effects by ePharmaLib.

Materials and Methods

Experiment design

As shown in Figure 1, the representative experiment is divided into five processes: (1) protein targets were calculated by

Fei Liu is a Ph.D of Zhejiang University. His research interests focus on AI, drug discovery and therapeutic hypothermia.

Xiangkang Jiang is a Ph.D candidate in Emergency Medicine at Zhejiang University. His research interests include genomics and epidemiology.

Jingyuan Yang is a Ph.D of Zhejiang University, focusing on the application of transcriptome, metabolomics and bioinspired materials.

Jiawei Tao is a master student at Zhejiang University, mainly dedicated to research on the application of machine learning and therapeutic hypothermia.

Mao Zhang is the vice president of Second Affiliated Hospital of Zhejiang University School of Medicine, is the director of the Institute of Emergency Medicine, Zhejiang University and is an expert in cardiac arrest, ultrasound, AI, 5G and drone applications in medicine.

Received: May 25, 2022. Revised: July 15, 2022. Accepted: August 3, 2022

© The Author(s) 2022. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

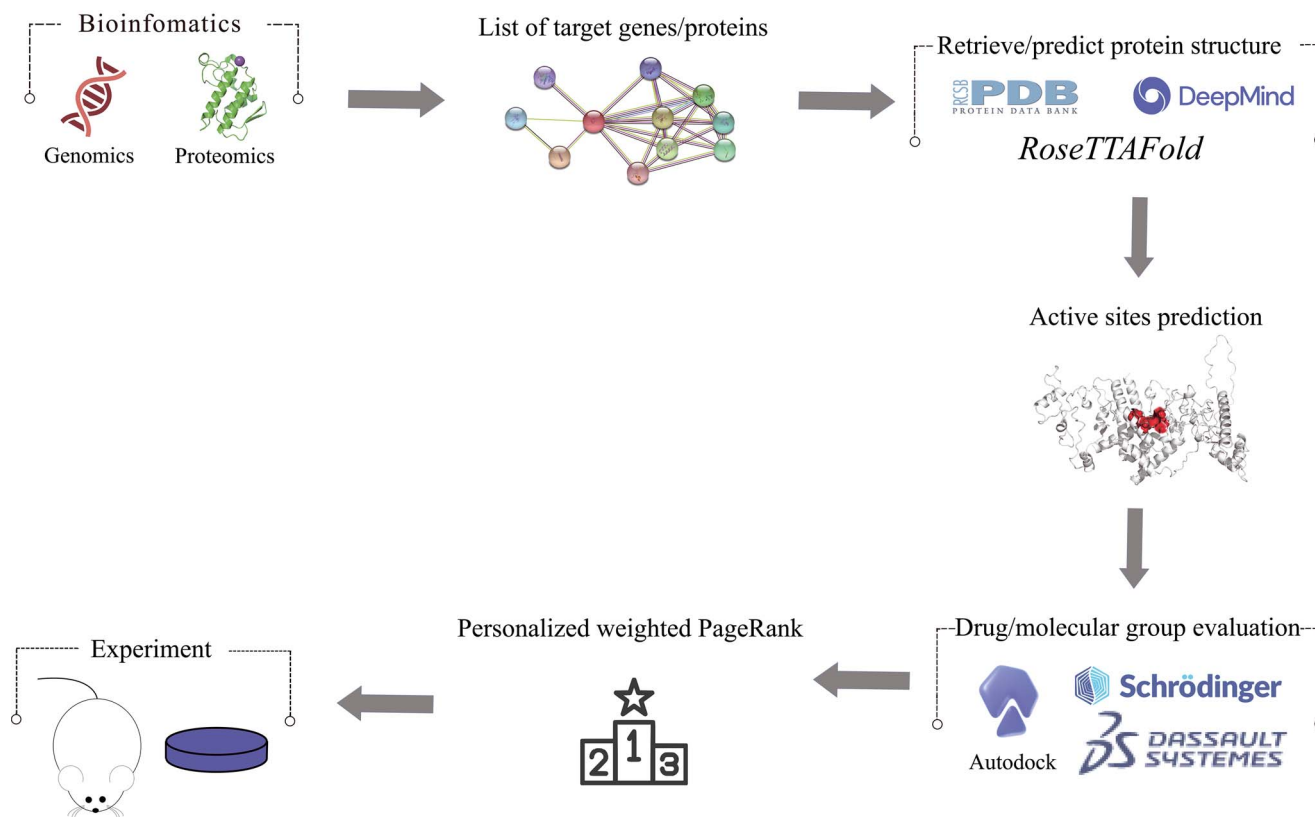


Figure 1. Representative workflow for bioinfo-pharmacology drug discovery.

bioinformatics analysis, (2) proteins and drugs 3D structure acquisition and prediction, (3) proteins' active sites prediction, (4) drug/molecular group evaluation with target proteins and (5) PageRank of docking results, protein logFC and messenger RNA (mRNA) expression. The experiment of animals or cells is referred to by authors but is not forced.

The data source of bioinformatics analysis

We retrieved the original data of mRNA expression under hypothermia treatment from the website of the National Center for Biotechnology Information (GSE54229). The research was reported by Ilmjärvi et al. [9]. In their research, mouse embryonic fibroblasts were exposed to mild hypothermia (32°C) or normothermia (37°C) to gain the transcription response induced by hypothermia. After 0.5, 1, 2, 4, 8 and 18 h of hypothermia, cells were collected for bioinformatics analysis.

Expression profile analysis

The log₂ fold-change (log₂FC) and *P*-value of each group were calculated. Top 3 log₂FC mRNAs with *P*-values <0.05 were selected from each group to enter the next step. If there exists mRNA with protein structure prediction failing, the mRNA would be skipped.

In the step, R 3.6.1 was used to detect the differential expressed compared to matched normothermia samples. The clustering of genes was calculated by the 'dist' and 'hclust' functions of R. The visualization of gene expression and clustering was performed by the 'dendextend' package.

3D data of proteins and small molecular drugs

All proteins were first searched on PubMed to see if there was protein clipping like cleaved caspase-3 [19]. Then, the 3D

structures of proteins were first searched from Protein Data Bank (PDB), which is used for biological-related ligand-protein interaction. In this article, no protein structure is listed on the PDB website. All the protein structures were predicted by AlphaFold2 and RoseTTAFold.

AlphaFold2 was developed by Google and it is the champion of the 14th Critical Assessment of Structure Prediction. In August 2021, in collaboration with the European Molecular Biology Laboratory-European Bioinformatics Institute, AlphaFold submitted a protein structure prediction database of model organism proteomes. RoseTTAFold is based on the Rosetta software, which is designed for macromolecular modeling, docking and design [20] RoseTTAFold also has good application [21] in the research of protein structure prediction. Finally, protein structures with fewer irregular regions will be selected for the next step.

All the selected protein structures need to be pretreated before they can be docked. Water molecules, ions and solvent molecules were removed if they existed. The missing hydrogen atoms were then filled in and the Gasteiger charge was calculated for the protein. Lastly, the file format were changed to PDBQT-format.

The 3D structures of 8697 drugs (DrugBank, 5.1.8) were downloaded from DrugBank Online (<https://go.drugbank.com/>). Drugs classified by the FDA as approved, experimental, nutraceutical and investigational were included in the research. The illicit and withdrawn drugs were excluded. Specific classification rules refer to the FDA's web site (<https://www.accessdata.fda.gov/>). After downloading the original data from DrugBank, we split each drug molecule into the PDBQT-format file. We added the missing hydrogen and calculated the Gasteiger charge for each one. Lastly, all molecules were energy-minimized before docking.

Table 1. The target proteins of different groups

Group	Target	logFC	Harmful for cell	Personalization
0.5 h	CIART	0.46	×	0.42
	CHAC1	0.43	✓	1.34
	NUDT22	0.40	✓	1.32
1 h	CDSN	0.52	×	0.41
	NR1D1	0.50	×	0.41
	CHAC1	0.50	✓	1.41
2 h	CIRP	0.80	×	0.36
	ARMCX5	0.66	×	0.39
	CCDC122	0.49	✓	1.41
4 h	CIRP	1.17	×	0.31
	RAMP3	0.91	×	0.35
	CEACAM1	0.87	✓	1.82
8 h	CIRP	1.57	✓	2.97
	RAMP3	1.33	×	0.28
	NQO1	1.18	✓	2.27
18 h	CIRP	1.71	✓	3.27
	NQO1	1.55	✓	2.93
	RAMP3	1.32	×	0.29

Visualize evolutionary conservation and active site prediction

Visualize evolutionary conservation was performed by the ConSurf server [22]. In a typical ConSurf application, through BLASTed [23] against the UNIREF-90 database [24] and aligning using MAFFT [25], the evolutionarily conserved positions are analyzed by the Rate4Site algorithm.

Then, the consensus approach-D (COACH-D) [26] was used to predict the active site of target proteins. The COACH-D uses five different methods to predict the binding sites of protein ligands. Four of these methods are COFAC-TOR [27], FINDSITE [28], TM-SITE [29] and S-SITE [29]. These methods predict binding sites by matching the query structure and sequence with the ligand-binding template in BioLiP [30], which is a semi-manual functional database [31] based on the PDB.

Virtual screening of potential compounds

To evaluate the hit compounds obtained from DrugBank and to calculate their interaction and binding posture in the active site of target proteins, the molecular docking method was carried out through QuickVina 2 [32]. QuickVina 2 uses the calculation of shape and electrostatic potential similarity of binding pockets to select molecules, which may exhibit binding patterns like those of binding pockets.

3D files of target proteins were dehydrated and hydrogenated. Then, proteins were saved as PDBQT files using AutoDock. AutoDock assisted in assigning Gasteiger charges and adding polar hydrogen atoms to both the proteins and the compounds.

Molecular dynamics simulation

The molecular dynamics (MD) simulation was performed by GRO-MACS [33]. First, a protein–drug complex was prepared, including adding hydrogenation and balancing charge. Then, we added a solvent so that the target protein and small molecules are coated. The forcefield was Chemistry at HARvard Macromolecular Mechanics 36. The simulation time was set as 50 ns for the speed of calculation. The simulation temperature was 309.15 K (36°C) and the pressure was 1 atm. Root-mean-square deviation (RMSD)

and root-mean-square fluctuation (RMSF) were calculated based on the first frame.

Personalization-weight-PageRank

We use personalization-weight-PageRank to rank the cross-level data of docking scores and differences in protein expression. To facilitate understanding, we took a simplified case in the Additional file Part 1.

PageRank is a comprehensive rank algorithm designed by Google and named after Larry Page [34]. It is one of the most famous ranking algorithms of network nodes based on the Markov process. PageRank has been applied in medical domains with success [35, 36].

Personalization and weight represent two different levels of score data. The weight of PageRank allows all nodes to be initially assigned with different weights/probabilities [37]. In this article, the weights of rank were set to docking scores of proteins and drugs. The higher the docking value, the higher is the connection rate of the complex.

Personalization of PageRank reinforces the connection intensity between the nodes, which makes the result more personalized and realistic [38]. In this article, personalization is influenced by protein functions. If the protein performs a negative influence, such as promoting apoptosis, the personalization will be calculated by $2^{(\text{fold change})}$ to ensure they are >1 . Meanwhile, if the protein plays a positive role in the group, the personalization will be set as $1/(\text{fold change} + 1)$ to <1 . The personalization values of all the drugs are set to 0 to prevent iterations of the drugs themselves from going wrong.

The calculation process is like putting all proteins and all drugs in the solution, then simulating the connections between all proteins and drugs by calculation. The damping factor is set to 0.85 to simulate the metabolism of proteins and drugs.

The whole calculation is based on Python 3.8.10. The relating Python libraries include NetworkX, Pandas and NumPy. We use Pandas and NumPy to import all the docking data into a matrix for PageRank calculating. The protein expression value is then imported by the PageRank personalization parameter of NetworkX. Lastly, we can get a comprehensive ranking of drugs.

The rank of drug combinations

In addition to the comprehensive ranking of drugs, we also try to generate the rank of drug combinations. Similarly, the calculation places all drugs of combination and target proteins in a solution to bind free.

First, all drugs will be grouped according to the docking results of drugs in each combination. In this article, to reduce the amount of calculation, we selected the top 20 drugs of each protein to include in the drug combination pool. Then, all the combinations were performed personalization-weight-PageRank against all protein targets. The sum of each score of all drugs in the combination is the final score of the combination. Lastly, we get the rank of combinations.

To make the distribution of combinations clear, we propose the drug-protein expression fit score (DPEFS) to show the data distribution pattern. The calculation is as follows: the PageRank values of all proteins were summed by multiplying logFC, then divided by the total PageRank values of drugs and were finally divided by the PageRank values of specific proteins for standardized calculation. It is used for standardized calculation for comparing different combinations.

DPEFS evaluates the combination by referring to the protein expression trend. The higher the DPEFS, the better is the fitness. In actual drug design, DPEFS is relatively high and PageRank score is relatively low, indicating that drugs of combination are relatively moderate, which suggests a negative outcome. All code

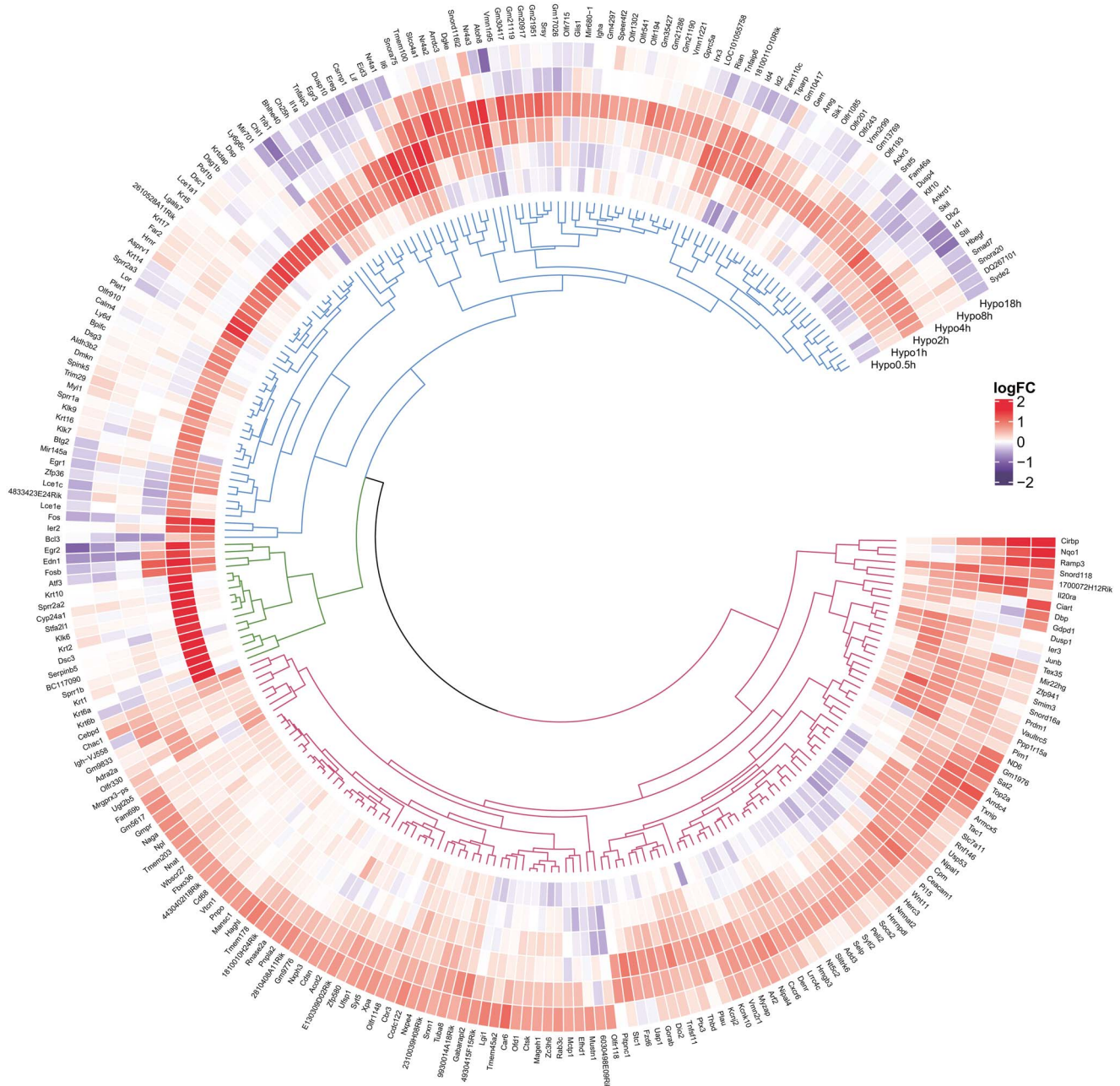


Figure 2. Circular visualization of expression patterns and clustering of hypothermia treatment.

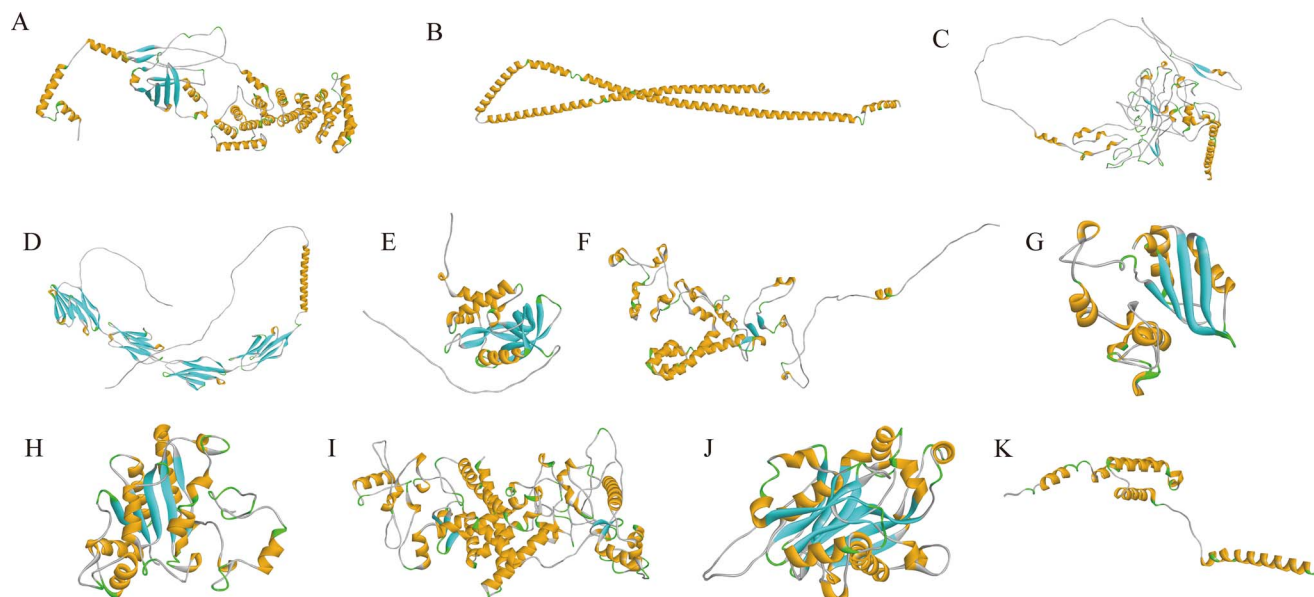


Figure 3. The 3D structures of target proteins. (A) ARM CX5, (B) CCDC, (C) CDSN, (D) CEACAM1, (E) CHAC1, (F) CIART, (G) CIRP, (H) NQO1, (I) NR1D1, (J) NUDT22 and (K) RAMP3.

can be found on GitHub (<https://github.com/FeiLiuEM/PageRank-weight-drug>).

Pharmacophore model analysis

All the selected drugs were analyzed by ePharmaLib. ePharmaLib [39] is an open source library of 15 148 e-pharmacophores modeled from solved structures of pharmaceutically relevant protein–ligand complexes of the screening PDB. It can be used for target fishing of phenotypic hits, side effect predictions, drug repurposing and scaffold hopping. In this article, after ranking, we predicted drug side effects targets by comparing top-ranking drugs with 15 148 e-pharmacophores. We then searched UniProt and Genecard websites to obtain the potential effect of the target, which is also equivalent to the side effects of drugs.

Results

Protein expression analysis and clustering after hypothermia expose

Figure 2 shows the expressions of different mRNAs of different groups after hypothermia. From the inside to the outside, the rings were divided into hypothermia 0.5 h group, hypothermia 1 h group, hypothermia 2 h group, hypothermia 4 h group, hypothermia 8 h group and hypothermia 18 h group.

As shown in Table 1, in each group, we selected the top three expression protein targets. In the hypothermia 0.5 h group, the target proteins are circadian-associated transcriptional repressor (CIART), glutathione-specific gamma-glutamylcyclotransferase 1 (CHAC1) and uridine diphosphate glucose pyrophosphatase nudix hydrolase 22 (NUDT22). The target proteins of the hypothermia 1 h group are CHAC1, corneodesmosin (CDSN) and nuclear receptor subfamily 1 group D member 1 (NR1D1). The target proteins of the hypothermia 2 h group are cold-induced RNA-binding protein (CIRP), armadillo repeat-containing X-linked protein 5 (ARM CX5) and coiled-coil domain-containing protein 122 (CCDC122). The target proteins of the hypothermia 4 h group are CIRP, receptor activity-modifying protein 3 (RAMP3) and carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1). The

target proteins of the hypothermia 8h and 18h groups are the same: CIRP, RAMP3 and NAD(P)H dehydrogenase [quinone] 1 (NQO1).

Within the targets, CHAC1 could enhance apoptosis [40]. NUDT22 is an Mg^{2+} -dependent UDP-glucose and UDP-galactose hydrolase [41], while high glucose shows a negative effect in HIE, like stroke [42]. CCDC122 is potentially pro-inflammatory [43]. CIRP can effectively reduce cell death in the early stage of hypothermia therapy. However, it has a strong pro-inflammatory effect outside the cell, leading to cell killing. There is no definitive research on the timing of this shift. Referring to the previous article [44], we conservatively believed that CIRP could be identified as a negative protein from the 8H group. CEACAM1 [45] and NQO1 [46] promote apoptosis. All the other targets are shown protective effects or do not have enough data. The personalization values are shown in Table 1. All the structures of target proteins were predicted by AlphaFold2 and RoseTTAFold in the Additional file (Supplementary Figure 1 are available online at <https://academic.oup.com/bib>). Eligible structures are shown in Figure 3 and as per the rules given in the Materials and Methods section.

Visualize evolutionary conservation and structure–function relationship-based binding site prediction

The conservation analysis of all the target proteins is listed in Figure 4A–K. The redder the amino acid, the higher is the possibility of the amino acid sequence with function. Then, we identified its structure–function relationship by the COACH-D server. The results showed a familiar result of conservation analysis, which is listed in Figure 4L–V. As shown in Table 2, the range around 3–5 Å of the active site was used for the setting of the receptor pocket of the target proteins that were used for virtual screening.

Virtual screening of target proteins' antagonists

We utilized the virtual screening technique to identify potential antagonists exhibiting an adequate binding affinity. We started with a chemical database consisting of 8697 drug molecules

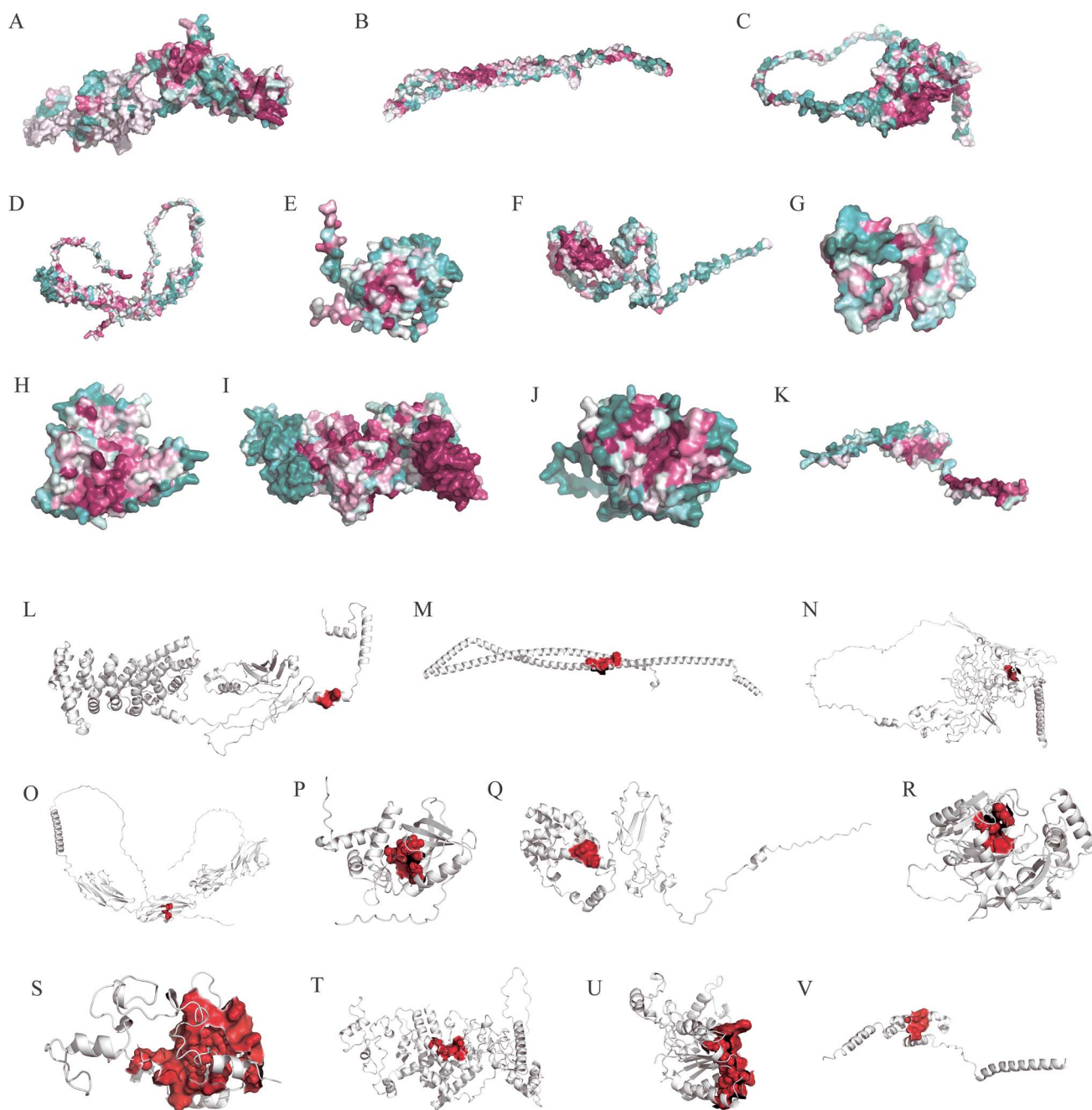


Figure 4. The ConSurf analysis and predicted active sites of target proteins. The upper 11 pictures are ConSurf analysis results. The darker the color, the more conserved the amino acids and the higher the probability of active sites. The last 11 pictures are predicted active sites. The colored areas indicate potential binding pockets of targets. (A and L) ARMCX5, (B and M) CCDC, (C and N) CDSN, (D and O) CEACAM1, (E and P) CHAC1, (F and Q) CIART, (G & R) CIRP, (H and S) NQO1, (I and T) NR1D1, (J and U) NUDT22 and (K and V) RAMP3.

and isolated a set of compounds satisfying the threshold of a high docking score. The results of the best match complexes are shown in Figure 5 and all the results are listed in the Additional file (Supplementary Table 1 are available online at <https://academic.oup.com/bib>).

MD simulations and binding free energy analysis

We performed the MD simulation of 11 complexes to measure the stability of the protein–ligand complex. RMSD profiles of the protein are shown in Figure 6A, which indicates that all systems were relatively stable during the entire simulation run. Moreover,

the RMSF profiles of protein are measured to evaluate the moving of each amino acid. All proteins are available for further analysis (Figure 6B).

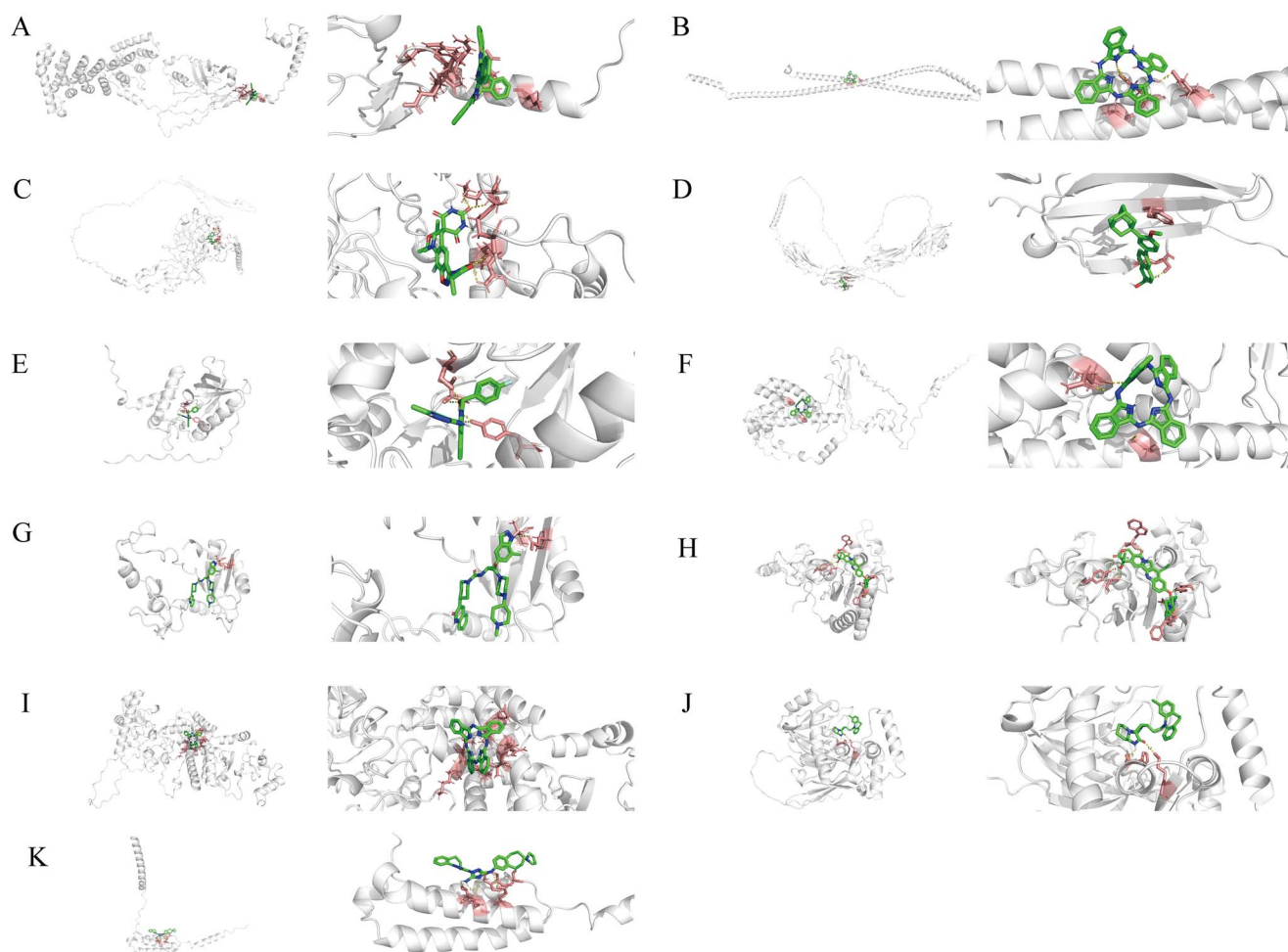
The RMSD of drug atoms was also conducted to predict the stability of the atoms in docked complexes (Figure 6C). Most compounds exhibited a consistently low RMSD, suggesting that these compounds formed stable complexes.

Drug rank of TH in different groups

We ranked all drugs using PageRank. First, we used PageRank for all the drugs and the results obtained are shown in Table 3.

Table 2. The docking parameters of target proteins

Protein name	Protein source	X	Y	Z	LEN-X	LEN-Y	LEN-Z
ARMCX5	RoseTTAFold	-7.06	31.20	-43.05	29.25	28.50	24.75
CDSN	RoseTTAFold	51.80	53.51	31.17	39.75	20.25	21.00
CEACAM1	AlphaFold2	-41.13	3.42	5.67	39.75	22.50	22.50
CHAC1	AlphaFold2	-3.04	-0.53	0.22	23.25	25.50	35.25
CIART	RoseTTAFold	36.79	-16.59	-36.23	28.50	38.25	22.50
CIRP	RoseTTAFold	14.48	2.86	-9.56	17.25	15.00	19.50
NQO1	AlphaFold2	-1.86	-16.32	-9.20	29.25	29.25	28.50
NR1D1	RoseTTAFold	0.34	35.70	10.28	39.75	25.50	24.75
NUDT22	AlphaFold2	10.08	-12.04	14.09	17.25	38.25	22.50
RAMP3	AlphaFold2	0.00	-10.93	-5.29	24.00	24.00	21.00
CCDC122	RoseTTAFold	119.27	23.34	8.09	36.00	47.25	47.25

**Figure 5.** The best docking molecular for each protein. (A) ARMCX5, (B) CCDC, (C) CDSN, (D) CEACAM1, (E) CHAC1, (F) CIART, (G) CIRP, (H) NQO1, (I) NR1D1, (J) NUDT22 and (K) RAMP3.

The higher the rank, the more protective the effect is in TH. The lower the rank is, the more damaging it is in TH. From Table 3, the recommended drugs in both H8 h and H18 h groups may damage the neurological function. The impact should be considered for different diseases. The prediction side effect targets of each drug were calculated by ePharmaLib. Side effects were referenced to the UniProt database. The original data are given in the Additional file (Supplementary Tables 2–15 are available online at <https://academic.oup.com/bib>).

The two-drug combinations are ranked in Table 4 and three-drug combinations are ranked in the Additional file

(Supplementary Table 16 are available online at <https://academic.oup.com/bib>). For comprehensive rank, the results of PageRank were listed. For drug combination ranks, the percentages of each drug's value in the combination were calculated. And, DPEFS was calculated for analyzing the distribution differences of drug combinations.

Discussion

In this report, a chronotherapeutics-friendly multi-target drug discovery method is proposed by taking TH as an example. This

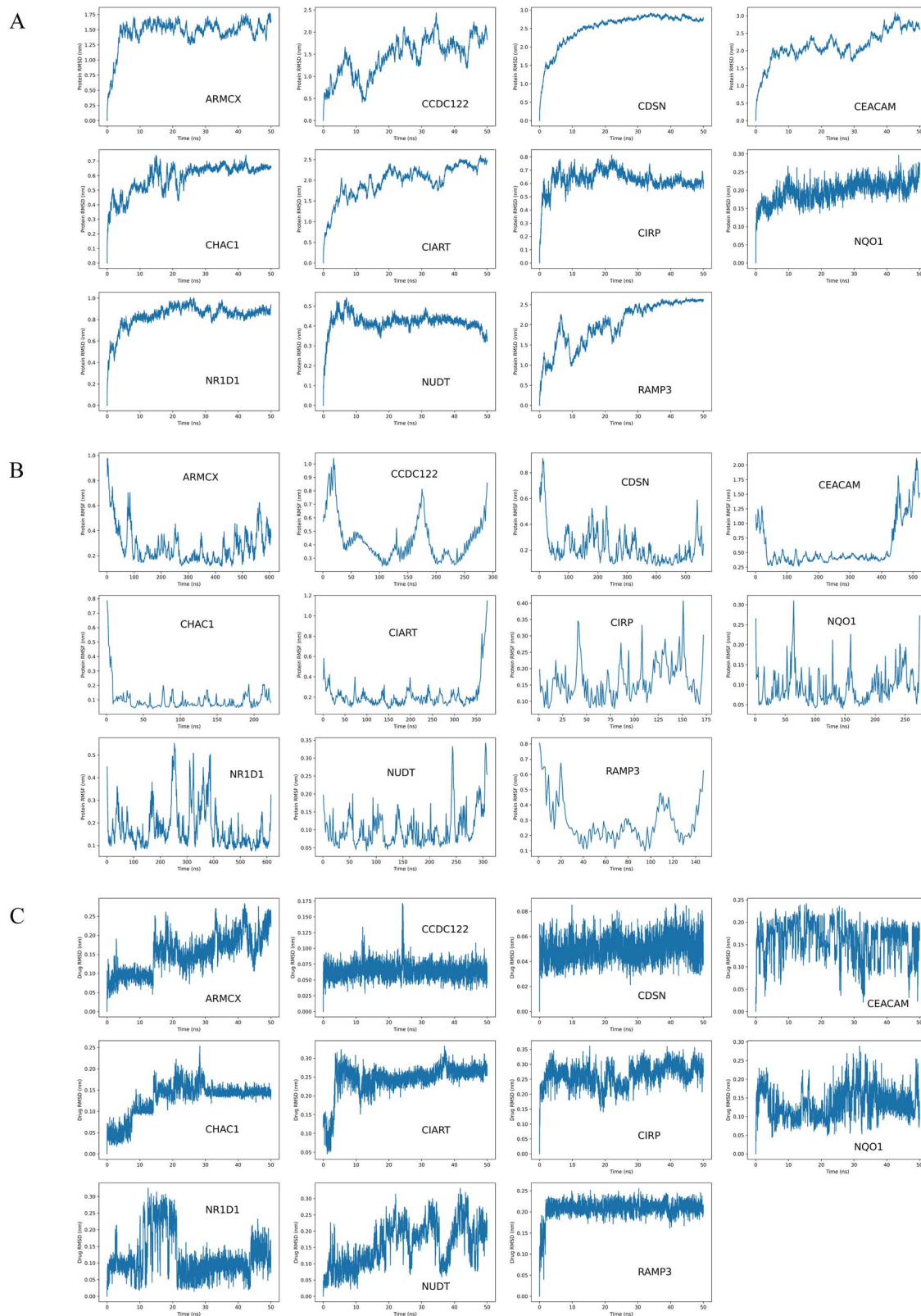


Figure 6. The RMSD and RMSF of MD simulation. (A) The RMSD of proteins. (B) The RMSF of proteins. (C) The RMSD of each molecular of proteins.

is the first application of PageRank, the first attempt of artificial intelligence (AI) protein prediction in the field of multi-target drug discovery and the first application of ePharmaLib.

In this paper, we chose the mRNA expression differences. Due to the high compatibility of PageRank, except for basic mRNA

expression differences, the analysis of Weighted Gene Coexpression Network [47] or Gene Regulatory Networks [48] will be better because they could provide numerical results of all proteins. For the same reason, PageRank has a good compatibility with existing pharmacological techniques such as pharmacophore models [49]

Table 3. The comprehensive rank of drugs and their predicted targets by pharmacophore models at different groups.

Group	Drug	Weighted_ personalized_ PageRank (*10 ⁻⁴)	Predicted targets by pharmacophore models	Predicted side effects
h0.5h	Fluzoparib	0.81	TYRO3, ACES, NOS2	Neuron protection from excitotoxic injury, platelet aggregation, cytoskeleton reorganization, innate immune response, inflammation
	Lorecivivint	0.80	CAH4, NOS2	pH homeostasis, inflammation
	Vactosertib	0.80	CAH5A, NOS2, CAH4	Mitochondrial function, pH homeostasis, inflammation
h1 h	Lorecivivint	0.81	CAH4, NOS2	pH homeostasis, inflammation
	AC-430	0.81	IF4E3, NOS2, CAH4	Protein synthesis, pH homeostasis, inflammation
	Raltegravir	0.80	NOS2, CAH4, CAH5A	Mitochondrial function, pH homeostasis, inflammation
h2 h	Phthalocyanine	0.90	ACES, NOS2	Neurological function, inflammation
	Vazegepant	0.87	NNRE, LGUL, NOS2	Cholesterol metabolism, osteoclastogenesis, oxidation, inflammation
	Bemcentinib	0.86	CRYM, LGUL, NOS2	Thyroid function, osteoclastogenesis, oxidation, inflammation
h4 h	Phthalocyanine	0.85	ACES, NOS2	Neurological function, inflammation
	3-(2-AMINOQUINAZOLIN-6-YL)-4-METHYL-N-[3-(TRIFLUOROMETHYL)PHENYL]BENZAMIDE	0.84	ACES, CN37, CAH5A	Neurological function, mitochondrial function
	Bemcentinib	0.84	CRYM, LGUL, NOS2	Thyroid function, osteoclastogenesis, oxidation, inflammation
h8 h	Phthalocyanine	0.87	ACES, NOS2	Neurological function, inflammation
	MK-3207	0.86	NOS2, CN37, ACES	Neurological function, inflammation
	Lifirafenib	0.85	EPHB2, ACES, NOS2	Development, neurological function, inflammation
h18 h	Phthalocyanine	0.87	ACES, NOS2	Neurological function, inflammation
	MK-3207	0.86	NOS2, CN37, ACES	Neurological function, inflammation
	Lifirafenib	0.85	EPHB2, ACES, NOS2	Development, neurological function, inflammation

Table 4. Rank of two drug combinations of different groups

Group	Drug 1	Drug 2	PageRank value of drug1	Percentage	PageRank value of drug2	Percentage	Personal- ized_weight _PageRank	DPEFS
h0.5h	Lorecivivint	Fluzoparib	0.41	49.85	0.42	50.15	0.83	2.30
	Vactosertib	Fluzoparib	0.41	49.56	0.42	50.44	0.83	2.33
	2'-deoxy-N-(naphthalen-1-ylmethyl)guanosine 5'-(dihydrogen phosphate)	Fluzoparib	0.41	49.68	0.42	50.32	0.83	2.32
h1 h	Lorecivivint	AC-430	0.41	49.67	0.42	50.33	0.83	4.36
	Lorecivivint	Raltegravir	0.42	50.12	0.41	49.88	0.83	4.40
	Lorecivivint	Vactosertib	0.42	50.21	0.41	49.79	0.83	4.37
h2 h	Phthalocyanine	Vazegepant	0.42	51.42	0.40	48.58	0.82	4.65
	Phthalocyanine	Bemcentinib	0.43	51.84	0.40	48.16	0.82	4.71
	Phthalocyanine	Lifirafenib	0.43	52.28	0.39	47.72	0.82	4.64
h4 h	3-(2-AMINOQUINAZOLIN-6-YL)-4-METHYL-N-[3-(TRIFLUOROMETHYL)PHENYL]BENZAMIDE	CD564	0.41	50.25	0.41	49.75	0.82	1.60
	Adapalene	3-(2-AMINOQUINAZOLIN-6-YL)-4-METHYL-N-[3-(TRIFLUOROMETHYL)PHENYL]BENZAMIDE	0.41	49.90	0.41	50.10	0.82	1.60
	3-(2-AMINOQUINAZOLIN-6-YL)-4-METHYL-N-[3-(TRIFLUOROMETHYL)PHENYL]BENZAMIDE	Phthalocyanine	0.42	50.52	0.41	49.48	0.82	1.60
h8 h	Phthalocyanine	MK-3207	0.41	49.91	0.41	50.09	0.83	3.94
	Phthalocyanine	Lifirafenib	0.42	50.17	0.41	49.83	0.83	3.93
	Lifirafenib	MK-3207	0.41	49.75	0.42	50.25	0.83	3.94
h18 h	Phthalocyanine	MK-3207	0.41	49.91	0.41	50.09	0.83	4.95
	Lifirafenib	MK-3207	0.41	49.77	0.42	50.23	0.83	4.95
	Phthalocyanine	Lifirafenib	0.42	50.14	0.41	49.86	0.83	4.94

and Quantitative Structure–Activity Relationship [50]. The ability to apply bioinformatics directly also indicates that this method can be applied to personal precision medicine.

AlphaFold2 and RoseTTAFold were used for protein structure prediction. And the number of selected proteins predicted by AlphaFold2 in this research was close to that of RoseTTAFold. During the process of protein structure prediction, we found that for some proteins, the structures predicted by RoseTTAFold have less irregular structure than that of AlphaFold2. This may be due to the 2D distance map level being transformed and integrated by RoseTTAFold during neural network training [18], while AlphaFold2 only paired the structure database and genetic database [17]. We also find a phenomenon that the predicted protein structures were relatively unstable under MD simulation than preview reports of other protein structures detected by X-ray.

We use the PageRank algorithm to rank all the drugs and combinations. The application of PageRank is suitable. The combination of drug molecules is a memoryless stochastic process, which meets the qualifications of the Markov process. The comprehensive analysis involves the free docking of proteins with all drugs. The process is similar to putting all proteins and drugs into a solution and docking freely. Besides, the method has good compatibility with the wide compatibility of PageRank. In theory, all the technologies with numerical results can be ranked by the method.

In addition, due to the advantages of PageRank, we can adjust the weights of different proteins to suppress the negative proteins with minimal impact on the positive ones. This property is beneficial for chronotherapeutics. Chronotherapeutics aim at treating illnesses according to the endogenous biologic rhythms, which moderate xenobiotic metabolism and cellular drug response [51]. With the new method, it is difficult for one administration to affect the effect of the next one. So, it is suitable for drug discovery of chronotherapeutics. Coincidentally, the CIRP mentioned in this report is classified as rhythm protein [52]. Body temperature also changes with circadian rhythms and regulates protein function [53, 54]. This is one of the reasons why TH was used as an example.

In the article, we also use pharmacophore models (ePharmaLib) to predict the side effects of drugs. For drug discovery, researchers can get the properties of drugs more efficiently. For clinical use, we can know the side effects of drugs much earlier. More precise and rational procedures can be performed for different patients. The authors were impressed by the efficient retrieval analysis of 15 148 pharmacophore models of ePharmaLib.

The new method matches the characteristics of pharmacogenomics [55] and brings new ideas for clinical drug development. For tumors, it is possible to effectively antagonize the most abnormally expressed proteins while limiting the affection of normal proteins. It could effectively improve the development of pharmacogenomics. Given this approach's close association with bioinformatics, bioinformatics can predict the most suitable drug for every patient in most diseases. Referring to the correlation between bioinformatics and genomics, we proposed to classify the new method as bioinfo-pharmacology.

This research has some defects. (1) For the speed of calculating, we only choose the top three mRNAs for docking and the top one complex for MD simulation. (2) Theoretically, pharmacophore modeling has a better improvement under PageRank. But considering lacking related copyright and the purpose of the article, we chose to dock FBI-approved drugs with target proteins to explore the interaction between proteins and drugs.

In summary, this paper proposes a new method through protein structure predicting and PageRank. The results provide

medical clues for the treatment of TH. This method takes a new attempt at drug discovery, which might make a little bit of a difference in pharmacology.

Key Points

- A new multi-target discovery method based on high precision AI (AlphaFold2 and RoseTTAFold) for protein structure prediction was proposed for drug discovery.
- A highly compatible numerical analysis method was proposed and evaluated to measure the effective effects of drugs and minimize the harmful effects.
- It is the first application of ePharmaLib, which is the only open source pharmacophore model tool available.

Supplementary data

Supplementary data are available online at <https://academic.oup.com/bib>.

Data availability

All the data that support the findings of this research are available for email from authors. All the PageRank code and data could be found at <https://github.com/FeiLiuEM/PageRank-weight-drug>.

Authors' contributions

F.L. designed the research and wrote the article. X.J. drew the pictures. J.Y. and J.T. made spreadsheets. M.Z. modified the article.

Acknowledgements

We thank Professor David Silver and Mu Li for algorithm teaching and consulting, Muhan Yu of Universität Hamburg for ePharmaLib consulting, Prof. Tie Xu and Yeping Du for the support of medical knowledge and Zaizai Cao, Xiangjie Lin and Yuanyuan Hao for the algorithm discussion.

Funding

The authors have stated that no funding relationships exist.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors allow the publication of this article.

References

1. Rodríguez-Soacha DA, Scheiner M, Decker M. Multi-target-directed-ligands acting as enzyme inhibitors and receptor ligands. *Eur J Med Chem* 2019;**180**:690–706.
2. Bawa P, Pradeep P, Kumar P, et al. Multi-target therapeutics for neuropsychiatric and neurodegenerative disorders. *Drug Discov Today* 2016;**21**:1886–914.

3. Lascarrou J-B, Merdji H, Le Gouge A, et al. Targeted temperature management for cardiac arrest with nonshockable rhythm. *N Engl J Med* 2019;**381**:2327–37.
4. Kim JY, Kim JH, Park J, et al. Targeted temperature management at 36 °C shows therapeutic effectiveness via alteration of microglial activation and polarization after ischemic stroke. *Transl Stroke Res* 2022;**13**:132–41.
5. Hosseini M, Wilson RH, Crouzet C, et al. Resuscitating the globally ischemic brain: TTM and beyond. *Neurother J Am Soc Exp Neurother* 2020;**17**:539–62.
6. Hazinski MF, Nolan JP, Aickin R, et al. Part 1: executive summary: 2015 international consensus on cardiopulmonary resuscitation and emergency cardiovascular care science with treatment recommendations. *Circulation* 2015;**132**:S2–39.
7. Lemiale V, Dumas F, Mongardon N, et al. Intensive care unit mortality after cardiac arrest: the relative contribution of shock and brain injury in a large cohort. *Intensive Care Med* 2013;**39**:1972–80.
8. Rosenthal L-M, Leithner C, Tong G, et al. RBM3 and CIRP expressions in targeted temperature management treated cardiac arrest patients—a prospective single center study. *PLoS One* 2019;**14**:e0226005.
9. Ilmjärv S, Hundahl CA, Reimets R, et al. Estimating differential expression from multiple indicators. *Nucleic Acids Res* 2014;**42**:e72.
10. Yang R, Zhan M, Nalabothula NR, et al. Functional significance for a heterogenous ribonucleoprotein A18 signature RNA motif in the 3'-untranslated region of ataxia telangiectasia mutated and Rad3-related (ATR) transcript. *J Biol Chem* 2010;**285**:8887–93.
11. Haley B, Paunesku T, Protić M, et al. Response of heterogeneous ribonuclear proteins (hnRNP) to ionising radiation and their involvement in DNA damage repair. *Int J Radiat Biol* 2009;**85**:643–55.
12. Zhang Y, Wu Y, Mao P, et al. Cold-inducible RNA-binding protein CIRP/hnRNP A18 regulates telomerase activity in a temperature-dependent manner. *Nucleic Acids Res* 2016;**44**:761–75.
13. Zhong P, Huang H. Recent progress in the research of cold-inducible RNA-binding protein. *Future Sci OA* 2017;**3**:FSO246.
14. Sakurai T, Kashida H, Watanabe T, et al. Stress response protein cirp links inflammation and tumorigenesis in colitis-associated cancer. *Cancer Res* 2014;**74**:6119–28.
15. Sakurai T, Kashida H, Komeda Y, et al. Stress response protein RBM3 promotes the development of colitis-associated cancer. *Inflamm Bowel Dis* 2017;**23**:57–65.
16. Wake DT, Ilbawi N, Dunnenberger HM, et al. Pharmacogenomics: prescribing precisely. *Med Clin North Am* 2019;**103**:977–90.
17. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;**596**:583–9.
18. Baek M, DiMaio F, Anishchenko I, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 2021;**373**:871–6.
19. Kothakota S, Azuma T, Reinhard C, et al. Caspase-3-generated fragment of gelsolin: effector of morphological change in apoptosis. *Science* 1997;**278**:294–8.
20. Leman JK, Weitzner BD, Lewis SM, et al. Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat Methods* 2020;**17**:665–80.
21. Humphreys IR, Pei J, Baek M, et al. Computed structures of core eukaryotic protein complexes. *Science* 2021;**374**:eabm4805.
22. Ashkenazy H, Abadi S, Martz E, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 2016;**44**:W344–50.
23. Biegert A, Söding J. Sequence context-specific profiles for homology searching. *Proc Natl Acad Sci U S A* 2009;**106**:3770–5.
24. Suzek BE, Wang Y, Huang H, et al. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinforma Oxf Engl* 2015;**31**:926–32.
25. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;**30**:772–80.
26. Wu Q, Peng Z, Zhang Y, et al. COACH-D: improved protein-ligand binding sites prediction with refined ligand-binding poses through molecular docking. *Nucleic Acids Res* 2018;**46**:W438–42.
27. Roy A, Yang J, Zhang Y. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Res* 2012;**40**:W471–7.
28. Brylinski M, Skolnick J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. *Proc Natl Acad Sci U S A* 2008;**105**:129–34.
29. Yang J, Roy A, Zhang Y. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinforma Oxf Engl* 2013;**29**:2588–95.
30. Yang J, Roy A, Zhang Y. BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. *Nucleic Acids Res* 2013;**41**:D1096–103.
31. Rose PW, Prlić A, Altunkaya A, et al. The RCSB protein data bank: integrative view of protein, gene and 3D structural information. *Nucleic Acids Res* 2017;**45**:D271–81.
32. Alhossary A, Handoko SD, Mu Y, et al. Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinforma Oxf Engl* 2015;**31**:2214–6.
33. Van Der Spoel D, Lindahl E, Hess B, et al. GROMACS: fast, flexible, and free. *J Comput Chem* 2005;**26**:1701–18.
34. Page L, Brin S, Motwani R, Winograd T. The PageRank citation ranking: bringing order to the web. 1999. <http://ilpubs.stanford.edu:8090/422/>.
35. Kalecky K, Cho Y-R. PrimAlign: PageRank-inspired Markovian alignment for large biological networks. *Bioinforma Oxf Engl* 2018;**34**:i537–46.
36. Chen L, Heikkinen L, Wang C, et al. Trends in the development of miRNA bioinformatics tools. *Brief Bioinform* 2019;**20**:1836–52.
37. Xing W, Ghorbani A. Weighted PageRank algorithm. *Proc Second Annu Conf Commun Netw Serv Res* 2004;**2004**:305–14.
38. Avrachenkov K, van der Hofstad R, Sokol M. Personalized PageRank with Node-Dependent Restart. *Algorithms Models Web Graph* 2014;**8882**:23–33.
39. Mounbock AFA, Li J, Tran HTT, et al. ePharmaLib: a versatile library of e-pharmacophores to address small-molecule (poly-)pharmacology. *J Chem Inf Model* 2021;**61**:3659–66.
40. Mungrue IN, Pagnon J, Kohannim O, et al. CHAC1/MGC4504 is a novel proapoptotic component of the unfolded protein response, downstream of the ATF4-ATF3-CHOP cascade. *J Immunol Baltim Md* 1950 2009;**182**:466–76.
41. Carter M, Jemth A-S, Carreras-Puigvert J, et al. Human NUDT22 is a UDP-glucose/galactose hydrolase exhibiting a unique structural fold. *Struct Lond Engl* 1993 2018;**26**:295–303.e6.
42. Zhang S, Zuo W, Guo X-F, et al. Cerebral glucose transporter: the possible therapeutic target for ischemic stroke. *Neurochem Int* 2014;**70**:22–9.
43. Leturiondo AL, Noronha AB, Mendonça CYR, et al. Association of NOD2 and IFNG single nucleotide polymorphisms with leprosy in the Amazon ethnic admixed population. *PLoS Negl Trop Dis* 2020;**14**:e0008247.

44. Denning N-L, Yang W-L, Hansen L, et al. C23, an oligopeptide derived from cold-inducible RNA-binding protein, suppresses inflammation and reduces lung injury in neonatal sepsis. *J Pediatr Surg* 2019;**54**:2053–60.
45. Zhao C, Dang Z, Sun J, et al. Up-regulation of microRNA-30b/30d cluster represses hepatocyte apoptosis in mice with fulminant hepatic failure by inhibiting CEACAM1. *IUBMB Life* 2020;**72**:1349–63.
46. Feng Y, Cui R, Li Z, et al. Methane alleviates acetaminophen-induced liver injury by inhibiting inflammation, oxidative stress, endoplasmic reticulum stress, and apoptosis through the Nrf2/HO-1/NQO1 signaling pathway. *Oxid Med Cell Longev* 2019;**2019**:7067619.
47. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform* 2008;**9**:559.
48. Alexiou A, Chatzichronis S, Perveen A, et al. Algorithmic and stochastic representations of gene regulatory networks and protein-protein interactions. *Curr Top Med Chem* 2019;**19**:413–25.
49. Sliwoski G, Kothiwale S, Meiler J, et al. Computational methods in drug discovery. *Pharmacol Rev* 2014;**66**:334–95.
50. Muratov EN, Bajorath J, Sheridan RP, et al. QSAR without borders. *Chem Soc Rev* 2020;**49**:3525–64.
51. Ballesta A, Innominato PF, Dallmann R, et al. Systems chronotherapeutics. *Pharmacol Rev* 2017;**69**:161–99.
52. Morf J, Rey G, Schneider K, et al. Cold-inducible RNA-binding protein modulates circadian gene expression posttranscriptionally. *Science* 2012;**338**:379–83.
53. Murayama Y, Kori H, Oshima C, et al. Low temperature nullifies the circadian clock in cyanobacteria through Hopf bifurcation. *Proc Natl Acad Sci U S A* 2017;**114**:5641–6.
54. Farsi H, Harti D, Achaâban MR, et al. Melatonin rhythm and other outputs of the master circadian clock in the desert goat (*Capra hircus*) are entrained by daily cycles of ambient temperature. *J Pineal Res* 2020;**68**:e12634.
55. Cecchin E, Stocco G. Pharmacogenomics and personalized medicine. *Gen* 2020;**11**:E679.