Electronic Supplementary Material (ESI) for Chemical Science

Crosslinker Nanocarriers-based Intratumoral Delivery for Protein Complex Mapping in Mitochondria of Live Tumor-Bearing Mice

Yuwan Chen, †[ab] Wenxin Fu, †[ad] Baofu Ma [ab], Xinwei Li,[ac] Wen Zhou,[ab] Zhen Liang, [a] Kaiguang Yang,*[a] Lihua Zhang,*[a] and Yukui Zhang[a]

a. State Key Laboratory of Medical Proteomics, National Chromatographic R. & A. Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China.

b. University of Chinese Academy of Sciences, Beijing 100049, China.

c. School of Chemistry, Dalian University of Technology, Dalian 116024, China.

d. Research Center for Analytical Sciences, Northeastern University, Shenyang 110819, China.

This PDF file includes:

Supplementary Text

Figs. S1 to S11

Supplementary text

Alignment process of the cryo-EM structure of the protein complex with the AlphaFold model structure of some subunits

The structure of complex IV (5Z62) was aligned based on the supercomplex $ClCIII_2CIV$ structure ($SCI_1III_2IV_1$; PDB:5XTH) (RMSD = 0.863, 61 to 61 atoms). The AlphaFold model of protein NDUFS7 was aligned based on the $SCI_1III_2IV_1$ structure (PDB code: 5XTH) (RMSD = 0.543, 143 to 143 atoms). The AlphaFold model of protein NDUFB8 was aligned based on the $SCI_1III_2IV_1$ structure (PDB code: 5XTH) (RMSD = 0.752, 131 to 131 atoms). The AlphaFold model of NDUFV3 was aligned based on the $SCI_1III_2IV_1$ structure (PDB code: 5XTH) (RMSD = 0.773, 25 to 25 atoms). The AlphaFold model of protein NDUFA13 was aligned based on the $SCI_1III_2IV_1$ structure (PDB code: 5XTH) (RMSD = 0.996, 112 to 112 atoms). The AlphaFold model of protein NDUFA2 was aligned based on the $SCI_1III_2IV_1$ structure (PDB code: 5XTH) (RMSD = 0.756, 76 to 76 atoms). The AlphaFold model of protein UQCRC2 was aligned based on the AW chain of the $SCI_1III_2IV_1$ structure (PDB code: 5XTH) (RMSD = 0.769, 392 to 392 atoms), and on the AK chain of the $SCI_1III_2IV_1$ structure (PDB code: 5XTH) (RMSD = 0.758, 382 to 382 atoms). The AlphaFold model of protein UQCRB was aligned based on the AF chain of the $SCI_1III_2IV_1$ structure (PDB code: 5XTH) (RMSD = 0.402, 97 to 97 atoms), and on the AS chain of the $SCI_1III_2IV_1$ structure (PDB code: 5XTH) (RMSD = 0.268, 97 to 97 atoms). The AlphaFold model of protein SDHB was aligned based on the complex II structure (PDB code: 8GS8) (RMSD = 0.329, 223 to 223 atoms). The 15 crosslinked peptide identified were located in the unresolved region of the reported protein structure, including COX6C (Lys-1 to Lys-47), NDUFS7 (Lys-66 to Lys-55), NDUFA2 (Lys-13 to Lys-98, Lys-13 to Lys-64, and Lys-13 to Lys-45), NDUFB8 (Lys-32 to Lys-34), NDUFV3 (Lys-51 to Lys-56), NDUFA13 (Lys-5 to Lys-7), UQCRC2 (Lys-21 to Lys-23), UQCRB (Lys-4 to Lys-19, Lys-4 to Lys-12) and SDHB (Lys-267 to Lys-274). The inter crosslinked peptides of UQCRC2(Lys-92) and UQCRB (Lys-4) were matched to the AlphaFold prediction model.

The AlphaFold model of protein ATP5ME was aligned based on the human ATP synthase structure (PDB code: 8H9U) (RMSD = 0.510, 42 to 42 atoms). The AlphaFold model of protein ATP5F1E was aligned based on the human ATP synthase structure (PDB code: 8H9U) (RMSD = 0.669, 41 to 41 atoms). The AlphaFold model of protein ATP5MF was aligned based on the human ATP synthase structure (PDB code: 8H9U) (RMSD = 2.676, 61 to 61 atoms). The AlphaFold model of protein ATP5MG was aligned based on the human ATP synthase structure (PDB code: 8H9U) (RMSD = 0.676, 70 to 70 atoms). The AlphaFold model of protein ATP5IF1 was aligned based on the human ATP synthase structure (PDB code: 8H9U) (RMSD = 0.395, 32 to 32 atoms). The AlphaFold model of protein ATP5MPL was aligned based on the human ATP synthase structure (PDB code: 8H9U) (RMSD = 7.350, 40 to 40 atoms). The AlphaFold model of protein MT-ATP8 was aligned based on the human ATP synthase structure (PDB code: 8H9U) (RMSD = 7.164, 51 to 51 atoms). The AlphaFold model of protein ATP5PD was aligned based on the human ATP synthase structure (PDB code: 8H9U) (RMSD = 5.576, 149 to 149 atoms). The AlphaFold model of protein ATP5PB was aligned based on the human ATP synthase structure (PDB code: 8H9U) (RMSD = 4.041, 188 to 188 atoms). The 23 crosslinked peptide amino acids identified were located in the unresolved region of human ATP synthase, including ATP5ME (Lys-1 to Lys-34, Lys-69 to Lys-55, Lys-69 to Lys-34, Lys-50 to Lys-48, and Lys-49 to Lys-50), ATP5F1E (Lys-47 to Lys-49), ATP5MF (Lys-14 to Lys-17 and Lys-16 to Lys-17), ATP5MG (Lys-24 to Lys-54 and Lys-24 to Lys-11), ATP5IF1 (Lys-82 to Lys-90, Lys-71 to Lys-72, Lys-82 to Lys-83, Lys-96 to Lys-98 and Lys-100 to Lys-1033), ATP5MPL(Lys-46 to Lys-49), MT-ATP8(Lys-49 to Lys-54), ATP5PD(Lys-1 to Lys-5) and ATP5PB (Lys-233 to Lys-244, Lys-233 to Lys-249 and Lys-221 to Lys-233). The above crosslinked peptides were matched to the AlphaFold prediction model.

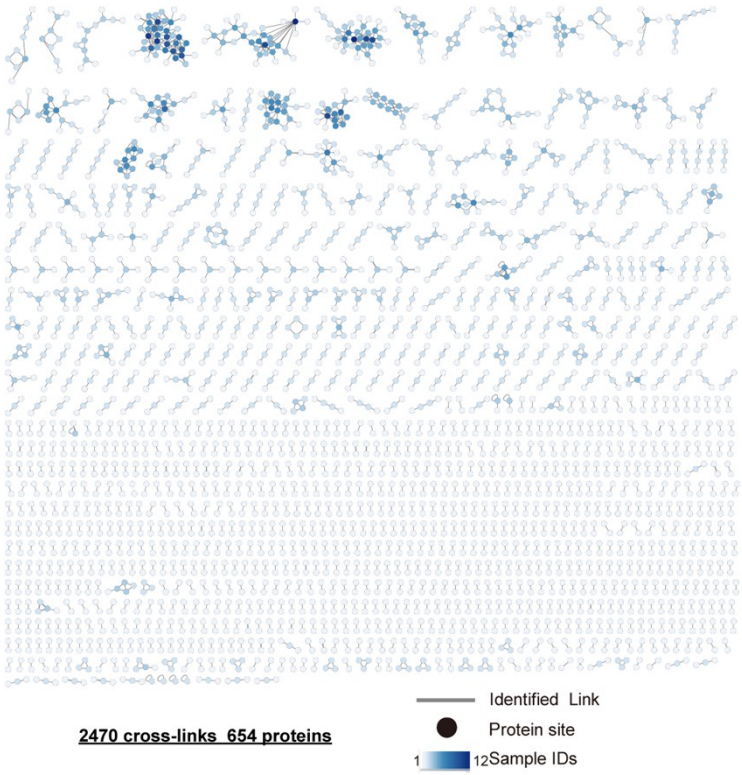
The AlphaFold model of protein MRPL12 was aligned with the t chain of MRPL12 in the ribosomal protein complex structure (PDB code:7QI4) (RMSD = 1.859, 36 to 36 atoms); aligned with the u chain of MRPL12 (PDB code:7QI4) (RMSD = 1.105, 32 to 32 atoms); aligned with the MRPL12 chain v (PDB code:7QI4) (RMSD = 1.027, 28 to 28 atoms); aligned with the MRPL12 chain w (PDB code:7QI4) (RMSD = 0.752, 26 to 26 atoms); aligned with the MRPL12 chain x (PDB code:7QI4) (RMSD = 1.199, 28 to 28 atoms) and aligned with the MRPL12 chain y (PDB code:7QI4) (RMSD = 1.199, 28 to 28 atoms. The AlphaFold model of protein MRPS28 was aligned based on the ribosomal protein complex structure (PDB code: PDB code:7QI4) (RMSD = 0.364, 86 to 86 atoms). The AlphaFold model of protein MRPL52 was aligned based on the ribosomal protein complex structure (PDB code:7QI4) (RMSD = 0.597, 78 to 78 atoms). The AlphaFold model of protein MRPL31 was aligned based on the ribosomal protein complex structure (PDB code:7QI4) (RMSD = 0.909, 101 to 101 atoms). The AlphaFold model of protein MRPL48 was aligned based on the ribosomal protein complex structure (PDB code:7QI4) (RMSD = 0.785, 148 to 148 atoms). The AlphaFold model of protein MRPL50 was aligned based on the ribosomal protein complex structure (PDB code: 7QI4) (RMSD = 0.406, 89 to 89 atoms). The 15 crosslinked peptide amino acids identified were located in the unresolved

region of the ribosomal protein complex structure, including MRPL12 (Lys-150 to Lys-183, Lys-150 to Lys-138, Lys-150 to Lys-142, Lys-185 to Lys-178, Lys-163 to Lys-183, Lys-173 to Lys-162, Lys-162 to Lys-147 and Lys-142 to Lys-144), MRPS28 (Lys-71 to Lys-75), MRPL52 (Lys-117 to Lys-110), MRPL31 (Lys-107 to Lys-97, Lys-195 to Lys-198 and Lys-145 to Lys-146), MRPL48 (Lys-1 to Lys-7) and MRPL50 (Lys-32 to Lys-34), the above crosslinked peptides were matched to the AlphaFold prediction model.

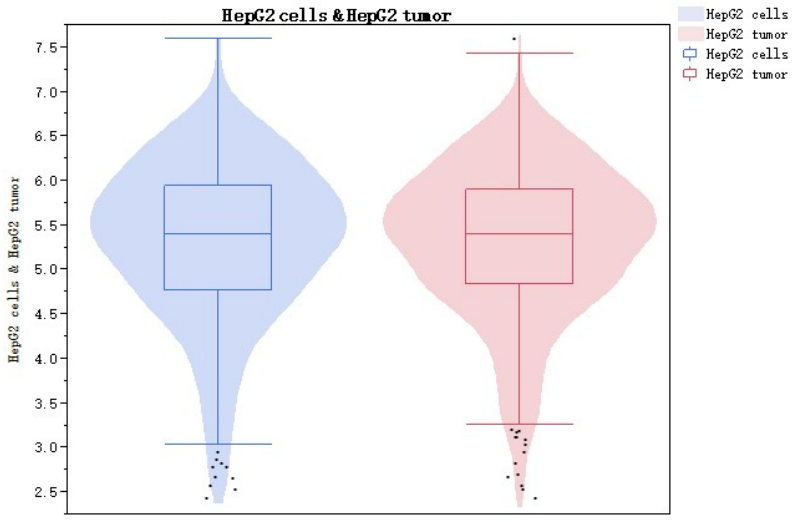
Reference database

The copy number of proteins was obtained from Ref 1. The Human.MitoCarta3.0 database was obtained from the MitoCarta3.0 website (https://www.broadinstitute.org/mitocarta)[2]. The experimental structures of the proteins were obtained from the PDB database. The detected residues located in the unreported region of the reported structure or the region lacking the experiment structure of the predicated structure of the proteins were obtained using the Uniprot plug in the AlphaFold model (https://alphafold.ebi.ac.uk/entry/)[3].

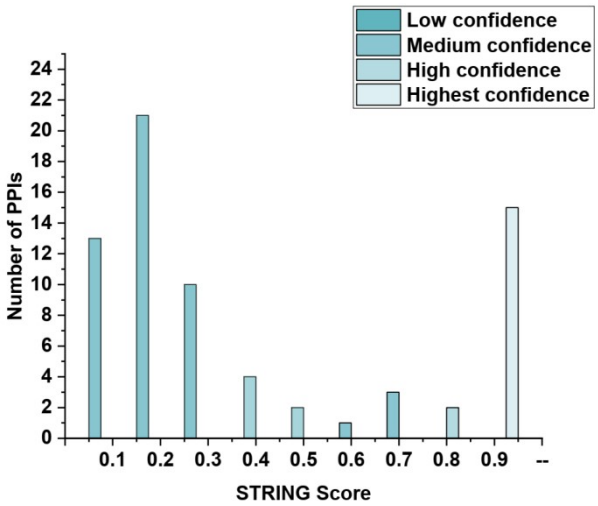
Supplementary Figures



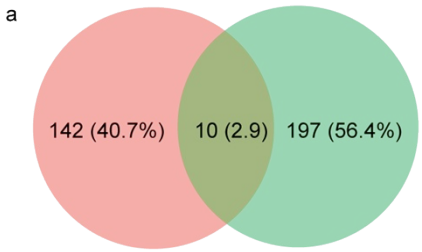
Supplementary Figure 1. Map of interaction sites of crosslinked peptides identified in mitochondria of tumor cells of tumor-bearing mice. Analysis of the crosslinked peptide results identified in all fractionations in a global view. Protein-Protein interactions (PPIs) were determined through the targeted delivery of crosslinkers to the mitochondria in the HepG2 cells via NPs for crosslinking. The nodes represented individual protein sites, and the lines represented all crosslinks identified between two proteins. Each node was colored based on the number of samples in which each protein interaction was observed. An interactive network depicting site-to-site interactions is available in Source Data file.



Supplementary Figure 2. Violin plot and box plot comparison of the copy-number distribution of proteins covered by crosslinked peptides identified in tumor-bearing mice and cultured HepG2 cell[4]. Violin plot of the distribution of the copy number of proteins detected in all crosslinked peptides spanned a dynamic range of five orders of magnitude, which correlated with the protein expression abundance. For boxplots, the center line, boxes and whiskers represent the median, inner quartiles, and rest of the data distribution Data was represented as mean values ± SD, n = 3 biological triplicates.



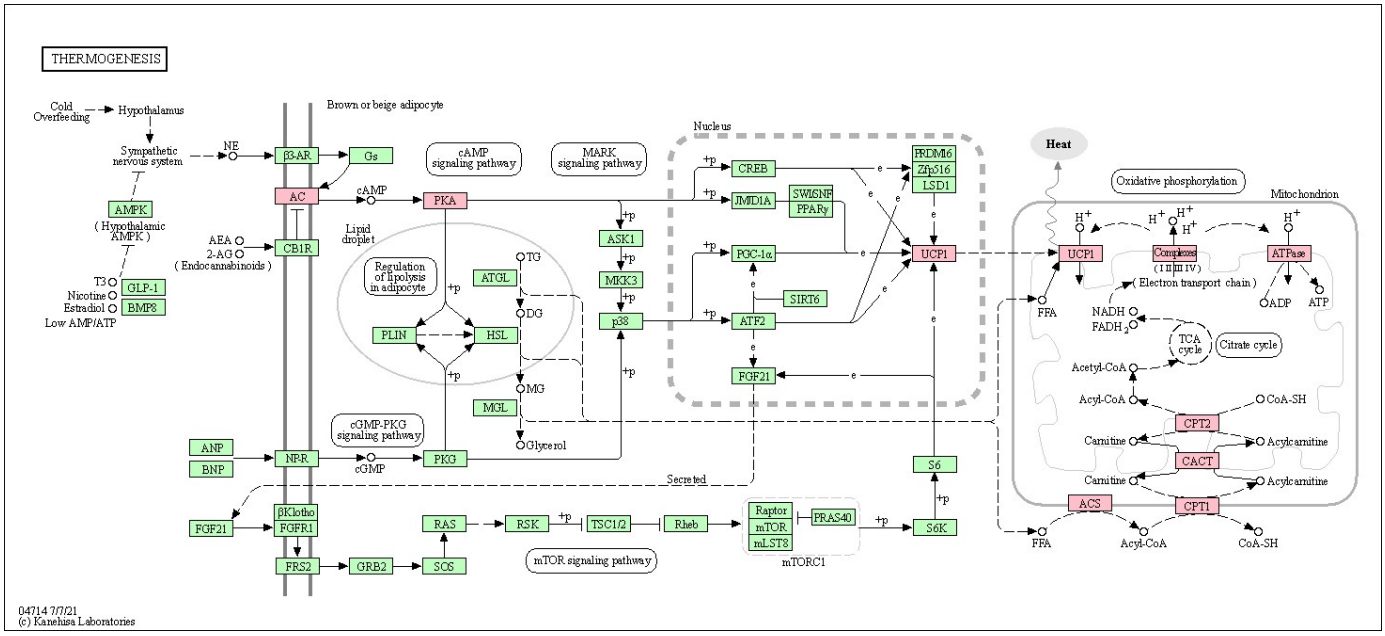
Supplementary Figure 3. Inter-type identified PPIs and scored in the string database. Low confidence = 0.15-0.4; medium confidence = 0.4-0.7; high confidence = 0.7-0.9; highest confidence = 0.9-1.



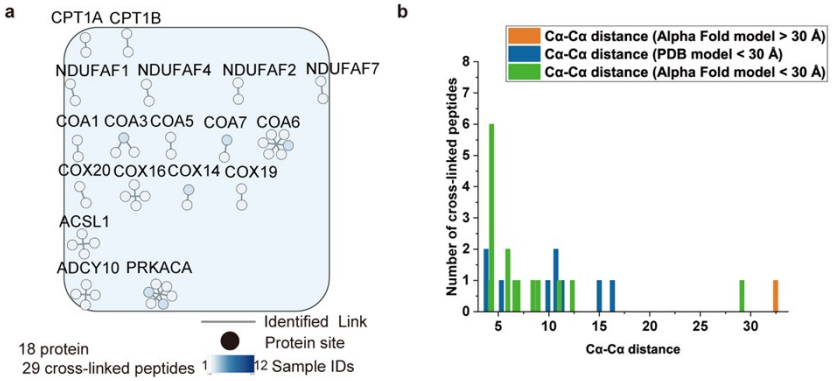
b **Overlap_10 PPIs**

ATP5PF-ATP5PD
HSPD1-SND1
ATP5F1D-ATP5F1E
PHB2-PHB
UQCRC2-UQCRB
ATP5F1A-ATP5F1B
HSPA9-TRAP1
ATP5PO-ATP5PB
HSPD1-MDH2
ATP5PB-ATP5PD

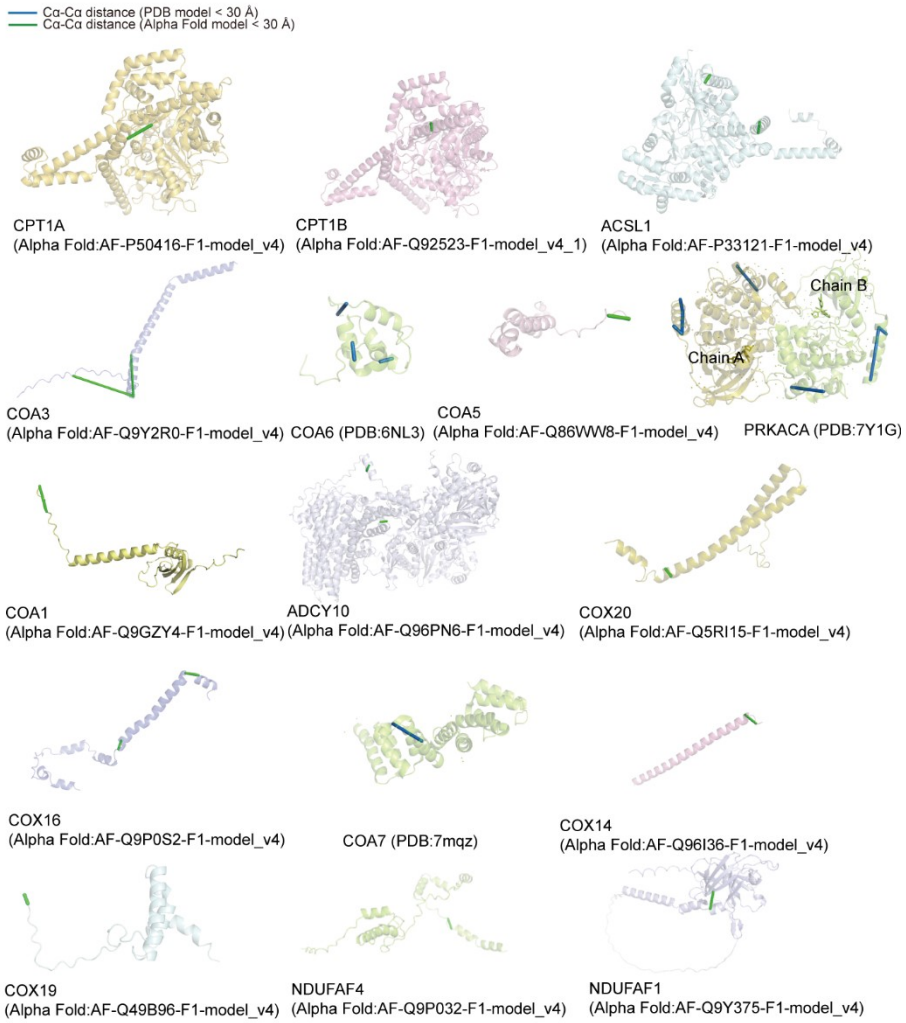
Supplementary Figure 4. PPIs comparison of crosslinked peptides identified in tumor-bearing mice and cultured HepG2 cells[4]. a Venn diagram of PPIs identified in HepG2 cells from tumor-bearing mice versus cultured HepG2 cells. b Ten overlapping PPIs detected between the two groups.



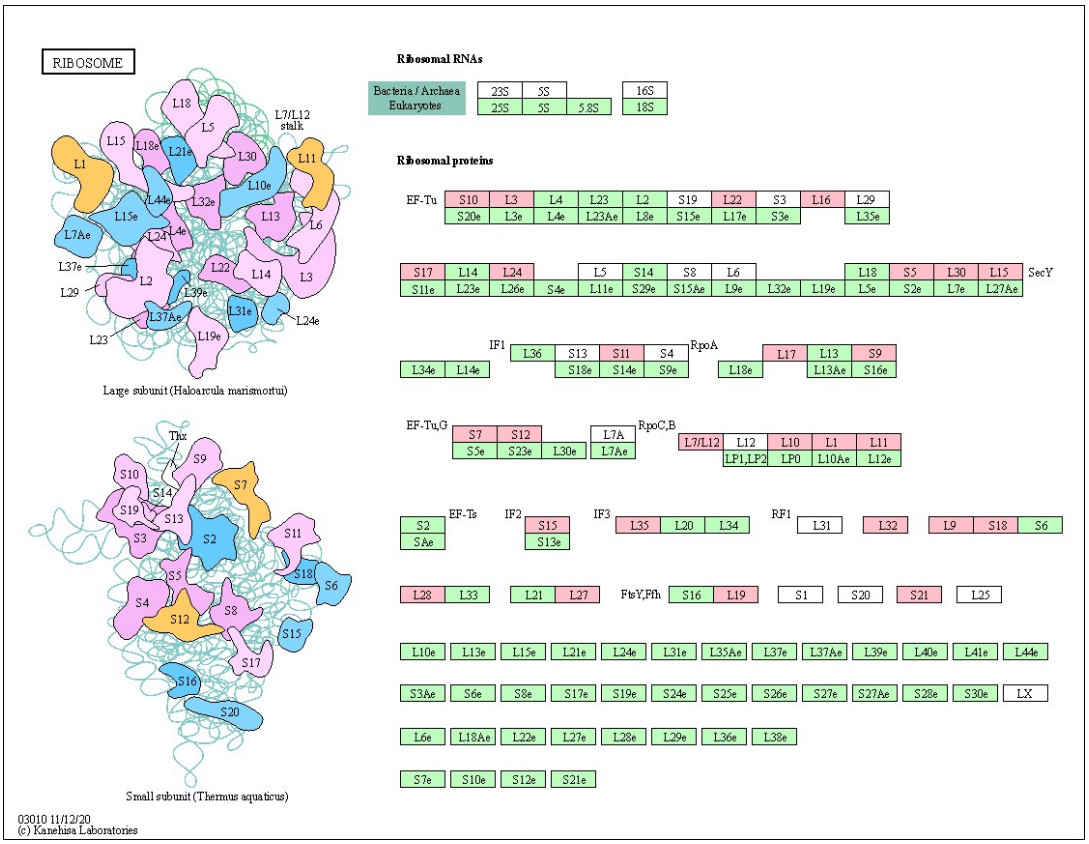
Supplementary Figure 5. Pathway analysis of the identified oxidatively phosphorylated proteins and thermogenic metabolic proteins in all fractionations using KEGG mapper. KEGG mapper was used to analyze the biological pathways of the proteins in oxidatively phosphorylated proteins and thermogenic metabolic proteins, which were identified by crosslinked peptides. The protein subunits identified by the crosslinked peptides were shown in pink, and those not identified were shown in green. The proteins from bacteria or archaea were shown in white (abbreviation: B/A).



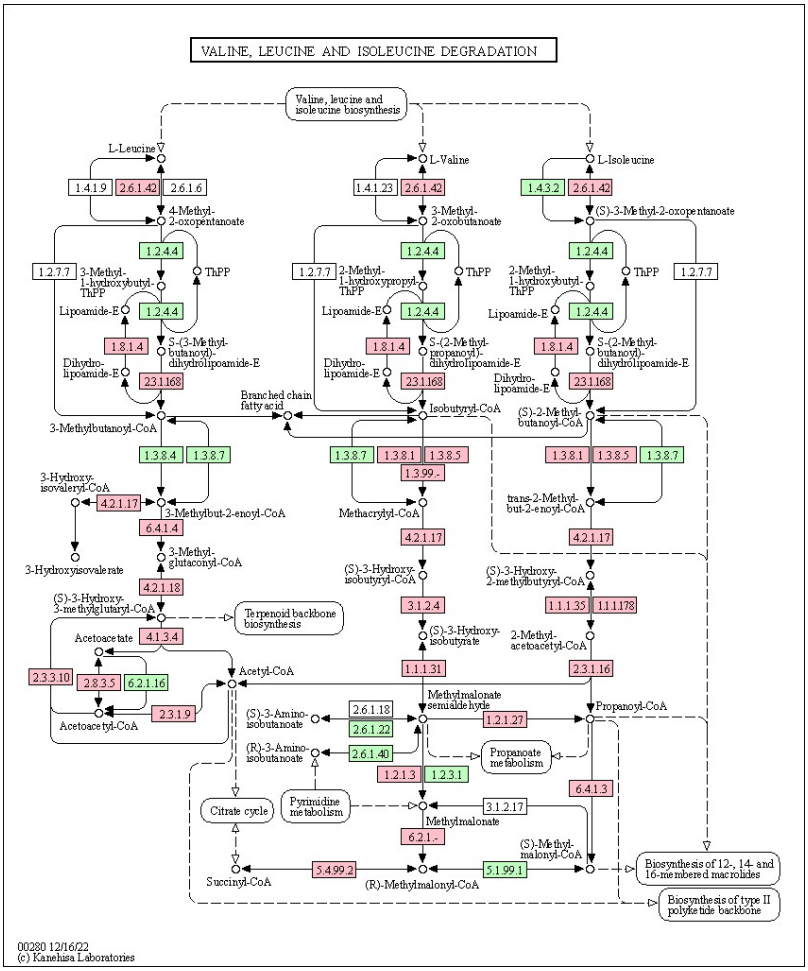
Supplementary Figure 6. Crosslinked peptides of thermogenic metabolic proteins in all fractionations. a. The crosslinked peptides of thermogenic metabolic proteins were selected from Fig. S1 and displayed individually, where the nodes represent individual proteins, and the lines represent all crosslinks identified between two proteins. Each node was colored based on the number of samples in which each protein interaction was observed. b. Histograms showing the length distribution of all distance restraints on the thermogenic metabolic proteins in Fig. S7 after docking.



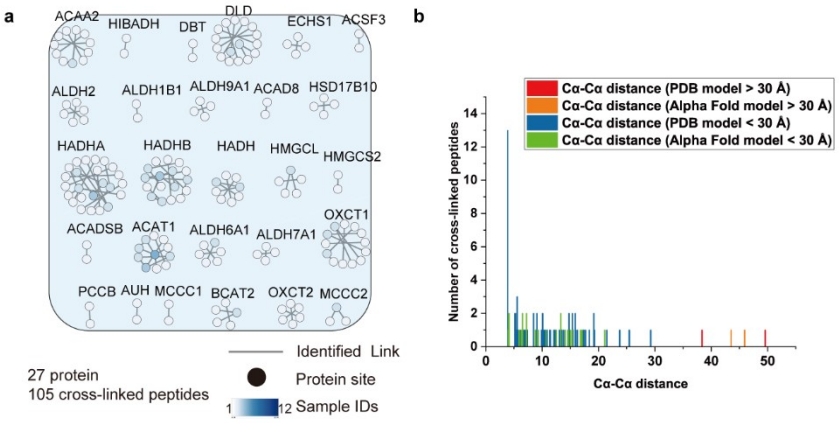
Supplementary Figure 7. Crosslinked sites mapped on thermogenic metabolic proteins in all fractionations. The crosslinked sites were mapped onto the corresponding protein structures. Crosslinks within 30 Å are shown in blue (PDB model) or green (AlphaFold model).



Supplementary Figure 8. Pathway analysis of the identified mitochondrial ribosomal proteins in all fractionations using KEGG mapper. KEGG mapper was used to analyze the biological pathways of the proteins in mitochondrial ribosomal proteins, which were identified by crosslinked peptides. The protein subunits identified by the crosslinked peptides were shown in pink, and those not identified were shown in green. The proteins from bacteria or archaea were shown in white (abbreviation: B/A).



Supplementary Figure 9. Pathway analysis of the identified branched-chain amino acids degradation proteins in all fractionations using KEGG mapper. KEGG mapper was used to analyze the biological pathways of the proteins in mitochondrial valine, leucine and isoleucine degradation, which were identified by crosslinked peptides. The protein subunits identified by the crosslinked peptides were shown in pink, and those not identified were shown in green. The proteins from bacteria or archaea were shown in white (abbreviation: B/A).



Supplementary Figure 10. Crosslinked peptides of branched-chain amino acids degradation proteins in all fractionations. a Crosslinked peptides involved in mitochondrial valine, leucine and isoleucine degradation were selected from Fig. S1 and displayed individually. In the network, nodes represent proteins and lines represent identified crosslinks between them. Node colors indicate the number of samples in which each protein interaction was observed. b Histograms show the length distribution of all distance restraints on the protein structures (from Fig. S11) after docking.

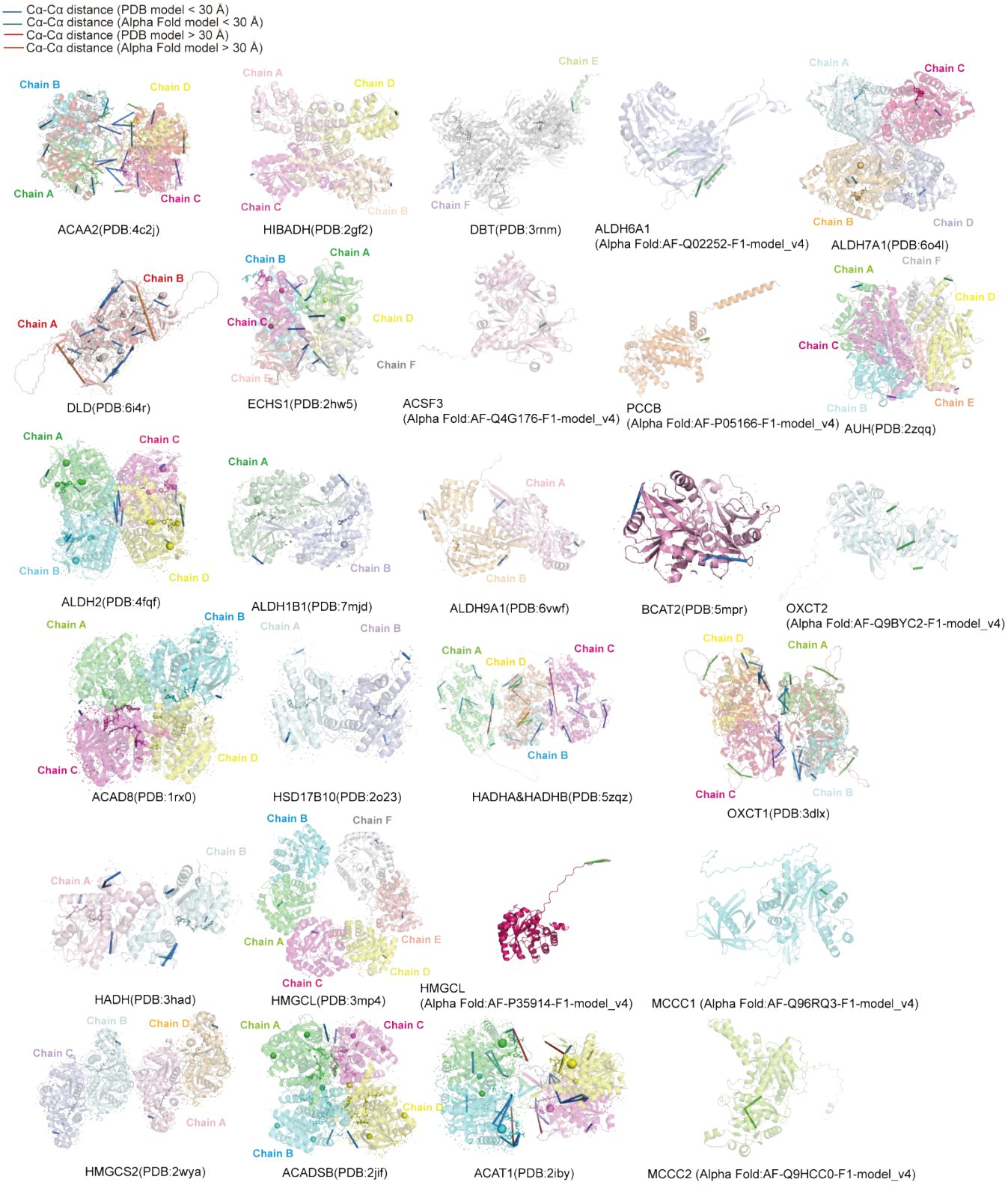


Figure 11. Structural analysis of crosslinked peptides from branched-chain amino acids degradation proteins in all fractionations. The crosslinked sites were mapped onto the corresponding protein structures. Crosslinks within 30 Å are shown in blue (PDB model) or green (AlphaFold model), while those exceeding 30 Å are shown in red (PDB) or orange (AlphaFold).

References

1. J. R. Wisniewski, A. Vildhede, A. Noren and P. Artursson, *J Proteomics*, 2016, **136**, 234-247.
2. S. Rath, R. Sharma, R. Gupta, T. Ast, C. Chan, T. J. Durham, R. P. Goodman, Z. Grabarek, M. E. Haas, W. H. W. Hung, P. R. Joshi, A. A. Jourdain, S. H. Kim, A. V. Kotrys, S. S. Lam, J. G. McCoy, J. D. Meisel, M. Miranda, A. Panda, A. Patgiri, R. Rogers, S. Sadre, H. Shah, O. S. Skinner, T. L. To, M. A. Walker, H. Wang, P. S. Ward, J. Wengrod, C. C. Yuan, S. E. Calvo and V. K. Mootha, *Nucleic Acids Res*, 2021, **49**, D1541-D1547.
3. K. Tunyasuvunakool, J. Adler, Z. Wu, T. Green, M. Zielinski, A. Zidek, A. Bridgland, A. Cowie, C. Meyer, A. Laydon, S. Velankar, G. J. Kleywegt, A. Bateman, R. Evans, A. Pritzel, M. Figurnov, O. Ronneberger, R. Bates, S. A. A. Kohl, A. Potapenko, A. J. Ballard, B. Romera-Paredes, S. Nikolov, R. Jain, E. Clancy, D. Reiman, S. Petersen, A. W. Senior, K. Kavukcuoglu, E. Birney, P. Kohli, J. Jumper and D. Hassabis, *Nature*, 2021, **596**, 590-596.
4. Y. Chen, W. Zhou, Y. Xia, W. Zhang, Q. Zhao, X. Li, H. Gao, Z. Liang, G. Ma, K. Yang, L. Zhang and Y. Zhang, *Nat Commun*, 2023, **14**, 3882.