Global profiling of lysine 2-hydroxyisobutyrylome in *Toxoplasma gondii* using affinity purification mass spectrometry

Lan-Bi Nie[1,2] · Qin-Li Liang[2] · Hany M. Elsheikha[3] · Rui Du[1] · Xing-Quan Zhu[2,4] · Fa-Cai Li[2,5]

(✉) Rui Du
durui197107@126.com
(✉) Xing-Quan Zhu
xingquanzhu1@hotmail.com
(✉) Fa-Cai Li
li78561270@163.com

[1] College of Animal Science and Technology, Jilin Agricultural University, Changchun, Jilin Province 130118, People's Republic of China
[2] State Key Laboratory of Veterinary Etiological Biology, Key Laboratory of Veterinary Parasitology of Gansu Province, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, Gansu Province 730046, People's Republic of China
[3] Faculty of Medicine and Health Sciences, School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus, Loughborough, LE12 5RD, UK
[4] College of Veterinary Medicine, Shanxi Agricultural University, Taigu, Shanxi Province 030801, People's Republic of China
[5] College of Veterinary Medicine, Southwest University, Chongqing 400715, People's Republic of China

Contact information for the corresponding author:
Professor Xing-Quan Zhu
Email: xingquanzhu1@hotmail.com
Fax: +86-931-8340977; Tel: +86-931-8342837

Abstract

Lysine 2-hydroxyisobutyrylation (K_{hib}) is a recently discovered and evolutionarily conserved form of protein post-translational modification (PTM) found in mammalian and yeast cells. Previous studies have shown that K_{hib} play roles in the activity of gene transcription and K_{hib}-containing proteins are closely related to the cellular metabolism. In this study, a global K_{hib}-containing analysis using the latest databases (ToxoDB 46, 8322 sequences, downloaded on April 16, 2020) and sensitive immune-affinity enrichment coupled with liquid chromatography with tandem mass spectrometry was performed. 1,078 K_{hib} modification sites across 400 K_{hib}-containing proteins were identified in tachyzoites of *T. gondii* RH strain. Bioinformatics and functional enrichment analysis showed that K_{hib}-modified proteins were associated with various biological processes, such as ribosome, glycolysis/gluconeogenesis, and central carbon metabolism. Interestingly, many proteins of the secretory organelles (e.g. microneme, rhoptry and dense granule) that play roles in the infection cycle of *T. gondii* were found to be K_{hib}-modified, suggesting the involvement of K_{hib} in key biological process during *T. gondii* infection. We also found that histone proteins, key enzymes related to cellular metabolism, and several glideosome components had K_{hib} sites. These results expanded our understanding of the roles of K_{hib} in *T. gondii* and should promote further investigations of how K_{hib} regulates gene expression and key biological functions in *T. gondii*.

Keywords *Toxoplasma gondii* · posttranslational modification (PTM) · lysine 2-hydroxyisobutyrylation · tachyzoite

Introduction

Toxoplasma gondii is an obligate intracellular apicomplexan protozoan which has a worldwide distribution in humans and animals (Montoya and Liesenfeld 2004). Infection by this parasite can cause encephalitis and retinitis, and even death particularly in immunocompromised individuals (Elsheikha 2008). *T. gondii* exists in three main developmental forms, namely tachyzoites, bradyzoites and sporozoites (Lindsay et al. 1991). Tachyzoites are responsible for the lytic cycle of invasion, replication and egress of the host cells, leading to acute toxoplasmosis (Dubey et al. 2009). *T. gondii* has received global attention because of some of its unusual biological, epidemiological and clinical features, including worldwide distribution, long-term persistent infection in the brain of the affected people (Rougier et al., 2017), a remarkable ability to cross biological barriers (Elsheikha and Khan, 2010), including the blood-brain-barrier, blood-retinal-barrier, blood-placental-barrier), infecting the developing fetus to cause miscarriage and congenital malformations (Elsheikha, 2008), and its association with neurophysiological disorders in adults (Elsheikha et al., 2016; Elsheikha and Zhu, 2016).

These facts motivated the global scientific community to have a better understanding of the biology and pathogenesis of toxoplasmosis, and to identify factors essential for the growth and development of *T. gondii*. One of the areas that has witnessed an intensive effort in the last few years is the protein post-translational modifications (PTMs) because they play essential roles in multiple cellular processes and can greatly expand the proteome diversification and complexity. PTMs are dynamic processes that involve changing of protein properties, such as physicochemical characteristics, space conformation and stability, by proteolytic cleavage or

addition of a modifying group to an amino acid (Walsh et al. 2005). A number of PTMs have been identified, and several of which, such as acetylation (Xue et al. 2013; Cobbold et al. 2016), glycosylation (Fauquenoy et al. 2008; Wang et al. 2016), palmitoylation (Foe et al. 2015; Caballero et al. 2016), phosphorylation (Treeck et al. 2011), succinylation (Li et al. 2014), and ubiquitination (Silmon de Monerri et al. 2015), have been shown to function as key regulators of diverse biological processes and functions in the Apicomplexa parasites (Yakubu et al. 2018).

As regards acetylation, 2,876 lysine acetylation sites across 1,146 proteins have been identified in *Plasmodium falciparum* (Cobbold et al. 2016), and 411 lysine acetylation sites distributed in 274 proteins have been reported in *T. gondii* (Jeffers and Sullivan 2012). A proteomic analysis of *T. gondii* confirmed that numerous N- and O-linked glycosylated sites were found in the micronemes, rhoptries, dense granules and the components of glideosome, which are involved in motility, invasion and intracellular survival (Fauquenoy et al. 2008; Wang et al. 2016). More than 30% of the predicted proteome have been shown to be phosphorylated in *P. falciparum* and *T. gondii* (Treeck et al. 2011; Alam et al. 2015), which play crucial regulatory roles in parasite motility, energy metabolism and host-parasite interaction. In *T. gondii*, phosphorylation of a motor protein myosin A (MyoA) at two serine sites by calcium dependent kinase 3 (CDPK3) can facilitate the initiation of parasite motility and egress (Gaji et al. 2015). A phosphorylation null mutant of glycogen phosphorylase (GP^{S25A}) in *T. gondii* PRU strain resulted in amylopectin accumulation, showing that GP phosphorylation is a regulatory factor for amylopectin storage and digestion (Sugi et al. 2017). Additionally, *T. gondii* rhoptry protein 16 (ROP16) can directly phosphorylate host signal transducer and activator of transcription (STAT)-1, STAT-3, STAT-5, STAT-6 (Yamamoto et al. 2009; Ong et al. 2010;

Butcher et al. 2011; Rosowski and Saeij 2012; Jensen et al. 2013), which are critical for host defense against *T. gondii*.

Lysine 2-hydroxyisobutyrylation (K_{hib}) is an evolutionarily conserved and abundant histone mark that has been detected in eukaryotic cells (Dai et al. 2014). H4K8 K_{hib} has been shown to be involved in transcriptional activity in meiotic and post-meiotic cells (Dai et al. 2014) and glucose homeostasis in *Saccharomyces cerevisiae* (Huang et al. 2017). Also, histone K_{hib} has been detected in *Trypanosoma cruzi* (Picchi et al. 2017). An earlier study also detected K_{hib} along with crotonylation (K_{cr}) proteins in *T. gondii* (Yin et al. 2019). In the present study, using the latest databases ToxoDB 46, we identified some different K_{hib} proteins that play important roles in *T. gondii* pathobiology. The K_{hib} proteome of *T. gondii* RH tachyzoites was analyzed using liquid chromatography with tandem mass spectrometry (LC-MS-MS) coupled with highly affinity purification. More than 1,000 K_{hib} sites across 400 K_{hib} proteins were identified, and these K_{hib} proteins were mainly located in the cytoplasm, nucleus, extracellular and mitochondria, and were primarily related to ribosome, glycolysis/gluconeogenesis and central carbon metabolism in cancer.

Materials and methods

Parasite and cell culture maintenance

Toxoplasma gondii RH strain was used in this study. Tachyzoites of *T. gondii* RH strain was originally stored and provided by the Department of Parasitology, Zhongshan School of

Medicine, Sun Yat-Sen University, Guangzhou, Guangdong Province, China. This RH strain belonged to Type I (ToxoDB #10) based on genotyping using Mn-PCR-RFLP (Liu et al. 2016). Tachyzoites of *T. gondii* RH strain were maintained in human foreskin fibroblast (HFF) cells (ATCC, Manassas, VA, USA) that were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (Gibco, USA) and antibiotics (100 U/ml penicillin and 100 μg/ml streptomycin; Gibco) at 37 °C with 5% CO_2. When the infected cells were lysed (~ within 3-4 days), the parasites and cells were harvested and passed through 25-gauge syringe needles. Tachyzoites were purified from host cell debris through 3-μm membrane filters (Millipore). The purified tachyzoites were washed with phosphate buffered saline (PBS) to remove any remaining host cell debris, and the purified parasite pellets were stored at –80 °C prior to protein extraction.

Protein extraction

The frozen tachyzoite pellets were resuspended and mixed with lysis buffer (8 M urea, 2 mM ethylenediaminetetraacetic acid (EDTA), 5 mM dithiothreitol (DTT), 3 μM trichostatin A (TSA), 50 mM nicotinamide (NAM) and 1% protease inhibitor cocktail) and then sonicated on ice. The cell debris was removed by centrifugation for 10 min at 4 ℃ and 20,000 *g*. The proteins were precipitated with 20% TCA for 2 h at 4 ℃. The supernatant was discarded by centrifugation at 12,000 *g* for 3 min at 4 ℃. The remaining precipitate was desalted with cold acetone three times. The protein was dissolved in urea buffer and the protein concentration was determined using a Bradford protein assay kit and bovine serum albumin as a standard. Protein

was digested with trypsin twice at trypsin to protein ratios of 1:50 and 1: 100 overnight.

Western blotting

The parasite lysates were separated by SDS-PAGE, and transferred to polyvinylidene difluoride membranes (PVDF, Millipore). The Khib proteins were detected by incubation of the membrane with primary pan anti-Khib antibody (PTM Biolabs) and followed by incubation with secondary antibodies coupled with horseradish peroxidase (Thermo-Fisher Scientific, Waltham, MA). The signals of horseradish peroxidase (HRP) were detected by an enhanced chemiluminescence kit (Pierce).

Enrichment of 2-hydroxyisobutyrylated peptides

To enrich the K_{hib} peptides, the tryptic peptides dissolved in NETN buffer (100 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl, 0.5% NP-40, pH 8.0) were incubated with prewashed anti-K_{hib} agarose-conjugated beads (PTM Biolabs, Hangzhou, China) with gentle shaking at 4 ℃ overnight. The beads were washed four times with NETN buffer, and three times with ddH_2O (pH 8.0). The bound peptides were eluted with 0.1% trifluoroacetic acid (TFA) and dried by a vacuum. The resulting peptides were cleaned by C18 ZipTips (Millipore Corp., Bedford, MA) according to the manufacturer's instructions, prior to LC-MS/MS analysis.

LC-MS/MS analysis

The enriched K_{hib} peptides were reconstituted in solvent A (0.1% formic acid in water) and loaded onto a C18 reverse-phase pre-column (Thermo-Fisher Scientific, Waltham, MA) to separate peptides. The gradient used was programed as: 6–23% solvent B (0.1% formic acid in 98% acetonitrile) for 26 min, 23–35% for 8 min and climbing to 80% in 3 min, then holding at 80% for the last 3 min. The eluted peptides were subjected to a NanoSpray Ionization source followed by MS/MS in Q Exactive (Thermo-Fisher Scientific) coupled online to the UPLC. Intact peptides were detected at a resolution of 70,000 in the Orbitrap. Peptides were selected for MS/MS analysis using NCE setting as 30; ion fragments were detected at a resolution of 17,500 in the Orbitrap. For MS scans, the m/z scan range was 350-1,800.

Database search

Maxquant search engine (v.1.5.2.8) was used to process the MS/MS data. The mass spectra data were quried in UniProt *T. gondii* database against the ToxoDB 46, 8322 sequences, downloaded on April 16, 2020, and concatenated with reverse decoy database. Trypsin/P was allowed up to four missing cleavages, specified as the cleavage enzyme. Mass tolerances for precursor ions was set to 10 ppm. K_{hib} on lysine K_{hib} was specified as a variable modification, while cysteine carbamidomethylation of cysteine was set as a fixed modification parameter. False discovery rate (FDR) thresholds for peptides was set to 1%. All the other parameters in MaxQuant analysis were set to default values. The Maxquant label free quantification (LFQ) algorithm (Cox et al., 2014) was used to perform the label-free quantification. The site of lysine K_{hib} site probability

localization was set as >0.75.

Bioinformatic analysis

Gene Ontology (GO) annotation of proteins was performed to identify the enriched functional categories using UniProt-GOA (http://www.ebi.ac.uk/GOA/) and ToxoDB 46 database. When an identified protein was not annotated by UniProt-GOA and ToxoDB database, the InterProScan was used to annotate protein's GO function by the alignment of protein sequence. The lysine 2-hydroxyisobutyrylated proteins were classified into three categories based on GO annotation: biological process, cellular component, and molecular function. Domains of 2-hydroxyisobutyrylated proteins were annotated by InterProScan, using the InterPro domain database, based on the protein sequence alignment. Kyoto Encyclopedia of Genes and Genomics (KEGG) was applied to identify the protein pathway. Protein subcellular location was predicted by Wolfpsort (https://wolfpsort.hgc.jp/). The sequence model contained amino acids in specific position of modified-21-mers (10 amino acids upstream and downstream of the Khib site) was analyzed by MoMo (http://meme-suite.org/tools/momo). The *T. gondii* proteome database was used as a background parameter, and other parameters were set as default. The GO, KEGG, and domain enrichment analysis of 2-hydroxyisobutyrylated proteins were performed using a two-tailed Fisher's exact test. The P-value < 0.05 was considered to be significant. Differentially 2-hydroxyisobutyrylated proteins were searched against the search tool for retrieval of interacting genes/proteins (STRING) database (http://string-db.org/) to obtain the protein-protein interaction (PPI) network. All parameters were set as default except

the interaction score that was set ≥ 0.7. Cytoscape (version 3.5.0) software was used to visualize the PPI network.

Results

Proteome-wide analysis of lysine 2-hydroxyisobutyrylation sites and proteins in *T. gondii*

To reveal the 2-hydroxyisobutyrylated proteins present in *T. gondii*, western blotting analysis using pan anti-K_{hib} antibody was performed and showed a wide range of bands in the parasite tachyzoite lysate (Figure 1A). Subsequently, a proteomic analysis based on LC-MS/MS and immune affinity was used to identify the global K_{hib} proteome of *T. gondii*. To determine the quality of MS data, the mass error of identified peptides was checked. The peptide mass error was < 4 ppm, suggesting the accuracy of the MS data (Figure 1B). Most of the identified peptides fell in the range of 7 to 17 amino acids in length, which were consistent with the properties of trypsin peptides (Figure 1C).

In the present study, three parallel experiments (designated Exp 1, Exp 2 and Exp 3) were performed, 673 K_{hib} sites on 297 K_{hib}-containing proteins were identified in Exp 1, 676 K_{hib} sites across 301 K_{hib}-containing proteins were identified in Exp 2, and 659 K_{hib} sites distributed on 297 K_{hib}-containing proteins were identified in Exp 3. Of these K_{hib} sites, about 47% were identified in at least two parallel experiments, indicating a high accuracy of these sites. Among the identified proteins, over 64% K_{hib}-containing proteins consisted of 1 or 2 K_{hib} sites, about 9% K_{hib}-containing proteins contained > 5 K_{hib} sites (Figure 1D).

Functional annotation of the K_{hib}-containing proteins of *T. gondii*

To have better understanding of the putative functions of the K_{hib}-containing proteins in *T. gondii*, GO functional classification of all K_{hib}-containing proteins were determined based on their biological processes, cellular components and molecular functions (Figure 2A-C). Within the biological processes, most K_{hib}-containing proteins were involved in cellular metabolic processes, organic substance metabolic processes, and primary metabolic processes, accounting for 12% of all K_{hib}-containing proteins, respectively (Figure 2A). For the cellular components, the majority of K_{hib}-containing proteins were enriched in intracellular (23%) (Figure 2B). Molecular functions analysis showed that 16%, 12% and 12% of the K_{hib}-containing proteins were associated with protein binding, organic cyclic compound binding and heterocyclic compound binding, respectively (Figure 2C). For the subcellular localization, the K_{hib}-containing proteins were mainly distributed in the cytoplasm (29%), nucleus (19%), extracellular (18%), and mitochondria (17%) (Figure 2D).

Motifs analysis of lysine 2-hydroxyisobutyrylated peptides

To characterize the K_{hib}-containing peptides, the specific amino acid biases adjacent to K_{hib} sites in all the identified K_{hib}-containing peptides were analyzed by Motif-x algorithm. In total, seven conserved motifs were identified, namely $K_{hib}X_1I$, $K_{hib}X_5K$, KX_7K_{hib}, $K_{hib}X_4K$, KX_9K_{hib}, KX_6K_{hib} and $K_{hib}X_6K$ (motif score > 6.7, X represents an amino acid residue) (Figure 3A). The

enriched and depleted amino acid residues surrounding the K_{hib} of all motifs are shown in a heatmap (Figure 3B). Most positions of I, K, M, V and Y amino acid residues around K_{hib} site were overrepresented, whereas R, S, P, G, E amino acid residues were underrepresented in the majority of positions (Figure 3B) (Red indicates that this amino acid is significantly enriched near the modification site, and green indicates that this amino acid is significantly reduced near the modification site).

Functional enrichment analysis

To reveal the biological functions of Khib-containing proteins, an enrichment analysis of the GO, KEGG, and domain databases was performed. GO enrichment analysis showed three categories, including cellular component, molecular function and biological process, were enriched (Figure 4A). For the cellular component, the K_{hib}-containing proteins were mainly enriched in mitochondria (Figure 4A). For the molecular function, most K_{hib}-containing proteins were associated with structural constituent of carbon-oxygen lyase activity, hydro-lyase activity, and box C/D snoRNA binding (Figure 4B). For the biological processes, the majority of the K_{hib}-containing proteins were significantly related to ADP metabolic process, nucleoside diphosphate phosphorylation and nucleotide phosphorlation (Figure 4C). Protein domain enrichment analysis revealed that K_{hib}-containing proteins were enriched in thioredoxin, proteasome subunit, ribosomal protein family and AHPC/TSA family (Figure 5A). KEGG enrichment analysis indicated that most K_{hib}-containing proteins participated in ribosome, glycolysis/gluconeogenesis and central carbon metabolism (Figure 5B), suggesting K_{hib}

involvement in energy metabolism processes.

K_{hib}-containing proteins involved in carbohydrate metabolism

To understand the roles of the K_{hib} in carbohydrate metabolism, many important proteins related to glycolysis/gluconeogenesis, ribosome, glyoxylate and dicarboxylate metabolism, and propanoate metabolism were analyzed. In glycolysis/gluconeogenesis process, 15 K_{hib}-containing proteins were identified, of which 9 proteins contained > 2 K_{hib} sites, including glyceraldehyde-3-phosphate dehydrogenase (9 K_{hib} sites), glucose-6-phosphate isomerase (4 K_{hib} sites), phosphoglycerate mutase (8 K_{hib} sites), lactate dehydrogenase (7 K_{hib} sites), pyruvate kinase (5 K_{hib} sites), enolase 2 (17 K_{hib} sites), phosphoglycerate kinase (7 K_{hib} sites), fructose-1,6-bisphosphate aldolase (9 K_{hib} sites), fructose-bisphospatase II (3 K_{hib} sites), triose-phosphate isomerase (4 K_{hib} sites). In the ribosome process, 54 proteins were found to have K_{hib} sites, which were mainly ribosomal proteins and fructose-1,6-bisphosphate aldolase. Additionally, K_{hib}-containing proteins were enriched in other carbohydrate metabolism pathways, including central carbon metabolism in cancer, aminoacyl-tRNA biosynthesis, and glyoxylate and dicarboxylate metabolism.

K_{hib}-containing proteins involved in *T. gondii* infection

Many secreted proteins from the parasite microneme (MIC), rhoptry (ROP) and dense granule (GRA) play critical roles in the invasion, virulence, replication and egress. A total of 8

295 microneme proteins were identified to be K_{hib}, including MIC1, MIC2, MIC4, MIC5, MIC8,
296 MIC11, apical membrane antigen 1 (AMA1) and subtilisin 1 (SUB1). Of these MICs, MIC1,
297 MIC2, MIC4, MIC5, MIC11 contained 5, 5, 5, 4, 5 K_{hib} sites. In the rhoptry proteins, many
298 ROPs and RONs contained K_{hib} sites, including ROP4 (3 K_{hib} sites), ROP5 (2 K_{hib} sites), ROP7
299 (4 K_{hib} sites), ROP15 (1 K_{hib} sites), ROP17 (1 K_{hib} sites), ROP35 (1 K_{hib} sites), ROP40 (1 K_{hib}
300 sites), RON2 (5 K_{hib} sites), RON5 (2 K_{hib} sites), RON8 (1 K_{hib} sites). For the dense granule
301 proteins, GRA2, GRA3, GRA7, GRA12 consisted of K_{hib} sites, accounting for 4, 2, 2, 4 K_{hib}
302 sites, respectively. Moreover, several crucial components for gliding motility were identified,
303 including GAP50 (4 K_{hib} sites), Myosin A (8 K_{hib} sites) and TgMLC1 (3 K_{hib} sites). For the
304 histone proteins, H2A, H2B, H3, H4 were consisted of many K_{hib} sites.
305
306 **PPI network of lysine K_{hib}-containing proteins in *T. gondii***
307
308 To study the cellular processes regulated by K_{hib} in *T. gondii*, the K_{hib} PPI network was
309 visualized by Cytoscape software. A total of 273 K_{hib}-containing proteins were mapped to the
310 protein interaction database (Figure 6). The K_{hib}-containing proteins were associated with
311 ribosome, glycolysis/gluconeogenesis, aminoacyl-tRNA biosynthesis and proteasome.
312
313 ## Discussion
314
315 Studies on PTMs in *T. gondii* are essential to provide valuable information on protein changes
316 and the underling processes that mediate the parasite interaction with the host cells. In recent

317 years, proteomic identification of K_{hib} on histone and non-histone proteins have been reported
318 in many species. A total of 6548 K_{hib} sites distributed on 1,725 proteins were discovered in
319 mammalian cells (Huang et al. 2018). In plants, 9,916 K_{hib} sites across 2,512 proteins were
320 identified in developing rice seeds (Meng et al. 2017), and 11,976 K_{hib} sites in 3,001 proteins
321 were found in *Physcomitrella patens* (Yu et al. 2017). In *S. cerevisiae*, 1,458 K_{hib} sites on 369
322 proteins were identified, many of which were enriched in the ribosome and
323 glycolysis/glycogenesis pathways (Huang et al. 2017). In *Proteus mirabilis*, 4,735 K_{hib} sites on
324 1,051 proteins were identified, and many K_{hib}-containing proteins were associated with
325 metabolic pathways, such as glycolysis/glycogenesis (Dong et al. 2018). In *T. gondii*, 9,502 K_{hib}
326 sites on 1,950 proteins were identified in the tachyzoites of *T. gondii* RH strain purified from
327 peritoneal fluid of mice (Yin et al. 2019).

328 In the present study, we determined the K_{hib} profile of *T. gondii* RH tachyzoites purified
329 from HFF monolayers and explored the potential involvement of the identified K_{hib}-containing
330 proteins in the infection processes by analyzing the K_{hib} proteome using a high-resolution LC-
331 MS/MS coupled with immune purification. We searched the latest version of the ToxoDB 46
332 database against ME49 strain (8322 sequences, accessed on April 16, 2020) and identified 1,078
333 K_{hib} sites across 400 K_{hib}-containing proteins. For protein extraction, we used lysis buffer to
334 lyse tachyzoites, it is inevitable that some insoluble cell membrane proteins may not be
335 dissolved completely and removed with cell debris, but this will not have much impact on the
336 experimental results in general as some previous results showed that PTMs have few or even
337 no modification sites on the cell membrane (Meng et al. 2017; Sun et al. 2017; Wu et al. 2018;
338 Nie et al. 2020). In a recent study, 2-hydroxyisobutyrylated proteins were mostly related to fatty

339 acid degradation (Yin et al. 2019); however, in our study K_{hib}-containing proteins were
340 primarily involved in ribosome, glycolysis/gluconeogenesis and central carbon metabolism in
341 cancer. In a previous study, the proteins were mainly distributed in the nucleus (Yin et al. 2019);
342 whereas in the present study, proteins were mostly abundant in the cytoplasm. These differences
343 may be caused by the updated databases and the different growth conditions of *T. gondii* RH
344 strain used in both studies. Moreover, PPI analysis suggested that abundant interactions
345 involved in important cellular processes were regulated by K_{hib} modification.

346 Comparative analysis between human cells (Huang et al. 2018), *Oryza sativa* (Meng et al.
347 2017), *P. patens* (Yu et al. 2017) and *T. gondii* showed that K_{hib} motif patterns are different from
348 each other. However, the K, V, Y residues were overrepresented in most positions around the
349 K_{hib} sites between *T. gondii* and *P. patens*, and I and V residues were overrepresented in the
350 majority of positions in *T. gondii* and *O. sativa*, but P and S residues were underrepresented in
351 *T. gondii* and *O. sativa*. The sequence logos showed a strong bias for isoleucine (I) downstream
352 of the K_{hib} sites, which was similar to the K_{mal} bias for cysteine (C) detected in *T. gondii* (Nie
353 et al. 2020), but was different from a recent study that reported that leucine (L), lysine (K),
354 tyrosine (Y) and valine (V) occurred upstream of the K_{hib} sites (Yin et al. 2019). This difference
355 may be due to the different versions of the database used between the two studies.

356 The carbohydrate metabolism, including glycolysis/gluconeogenesis, citrate cycle,
357 glyoxylate and dicarboxylate metabolism, starch and sucrose metabolism, pyruvate metabolism,
358 and fructose and mannose metabolism, participates in the lytic cycle of *T. gondii*. Our analysis
359 of the K_{hib} proteomic database indicated that some of the modified proteins participated in
360 metabolism processes, which is consistent with that of a previous study (Yin et al. 2019). For

example, enolase 2 (ENO2) is an essential factor for the growth of *T. gondii* (Mouveaux et al. 2014). Fructose-1, 6-bisphosphate aldolase (ALD) is required for energy metabolism rather than host-cell invasion in *T. gondii* (Shen and Sibley 2014). In the glycolysis/gluconeogenesis and citrate cycle processes, there were many K_{hib}-modified enzymes which are important for the energy supply of *T. gondii*, especially in the tachyzoites (fast replicating stage) and bradyzoites (slow replicating stage) (Nitzsche et al. 2017; Shukla et al. 2018). Additionally, the alpha-1,4 glucan phosphorylase containing the K_{hib} sites is involved in amylopectin digestion, which is crucial for the development of *T. gondii* bradyzoite and latent infection (Sugi et al. 2017).

Most of K_{hib}-containing proteins in carbohydrate metabolism are also enriched in other species. K_{hib}-containing proteins were strongly enriched in the *S. cerevisiae* glycolysis/gluconeogenesis pathway (Huang et al. 2017). In mammalian cells, several important enzymes were heavily modified, which is required for the glycolysis pathway, such as alpha-enolase (ENO1) and fructose-bisphosphate aldolase (ALD) (Huang et al. 2018). In *O. sativa* seeds, most K_{hib}-containing proteins were enriched in glycolysis/gluconeogenesis, citrate cycle, and starch and sucrose metabolism (Meng et al. 2017). These facts showed that the K_{hib} modification could play key roles in glucose metabolism.

The lytic cycle of *T. gondii*, including invasion, replication and egress, is largely regulated by three secretory organelles, including microneme, rhoptry and dense granule. Many of the proteins secreted by these organelles were identified to be K_{hib}. For example, the largest identified K_{hib} protein was a chaperonin protein BiP with 26 sites, which is different from a previous report which showed the rate-limiting enzyme phosphofructokinase PFKII as the most

significantly modified protein (Yin et al. 2019). The second was the heat shock protein HSP70 with 25 sites, which has also been found malonylated with five K_{mal} sites in *T. gondii* (Nie et al. 2020). AMA1 and MIC2 were identified as K_{hib}-containing proteins, which are involved in the attachment of extracellular parasites to the host membrane (Carruthers and Sibley 1997). In rhoptry proteins, RON2 containing 5 K_{hib} sites can interact with AMA1 to maintain the moving junction (MJ) integrity and is essential for the internalization of *T. gondii* (Lamarque et al. 2014). The rhoptry kinase ROP17 (1 K_{hib} sites) can manipulate monocyte migration to facilitate *T. gondii* dissemination (Drewry et al. 2019). GRA7 (2 K_{hib}) facilitates the virulence in mice (Alaganan et al. 2014), and GRA 12 (4 K_{hib}) plays an important role in mediating parasites resistance to host gamma interferon (Fox et al. 2019). These results indicated that K_{hib} can play key roles in the lytic cycle of *T. gondii*.

In *T. gondii* invasion and egress, the glideosome provides the power in gliding motivity (Frénal et al. 2017). Several components of the glideosome were identified as K_{hib}-containing proteins, including GAP45 (4 K_{hib} sites), GAP50 (4 K_{hib} sites), Myosin A (8 K_{hib} sites) and TgMLC1 (3 K_{hib} sites). Interestingly, changing the PTM sites of the glideosome components usually impair the invasion, egress and motility of *T. gondii*. The phosphorylation of Myosin A by CDPK3 contributes to the initiation of motility in *T. gondii* egress (Gaji et al. 2015). Mutations in acylation sites of GAP45 impair pellicle integrity in *T. gondii* invasion (Frénal et al. 2010). Thus, it will be interesting to study the K_{hib} sites of some important proteins in the lytic cycle of *T. gondii* in the future.

In conclusion, this study provided a new proteome dataset of K_{hib}, and identified 1,078 K_{hib} modification sites across 400 K_{hib}-containing proteins in *T. gondii*. These K_{hib}-containing

proteins participate in various cellular processes, such as ribosome, glycolysis/gluconeogenesis and central carbon metabolism. These data expanded our understanding of K_{hib} and provided new resources for further investigation of the roles of the lysine 2-hydroxyisobutyrylation in regulating different biological processes of *T. gondii*.

Data and materials availability The mass spectrometry data have been submitted to the ProteomeXchange Consortium with the identifier PXD019326.

Acknowledgements We would like to thank PTM Biolabs (Hangzhou, China) for assistance in LC-MS/MS analysis.

Funding information Project support was provided by the National Natural Science Foundation of China (Grant No. 31802182), the Fundamental Research Funds for the Central Universities (Grant No. SWU119030), the National Key Research and Development Program of China (Grant No. 2017YFD0500403) and the International Science and Technology Cooperation Project of Gansu Provincial Key Research and Development Program (Grant No. 17JR7WA031).

Competing interests The authors declare that they have no competing interests.

References

Alaganan A, Fentress SJ, Tang K, Wang Q, Sibley LD (2014) *Toxoplasma* GRA7 effector increases turnover of immunity - related GTPases and contributes to acute virulence in the

mouse. Proc Natl Acad Sci U S A 111:1126-1131

Alam MM, Solyakov L, Bottrill AR, Flueck C, Siddiqui FA, Singh S, Mistry S, Viskaduraki M, Lee K, Hopp CS, Chitnis CE, Doerig C, Moon RW, Green JL, Holder AA, Baker DA, Tobin AB (2015) Phosphoproteomics reveals malaria parasite protein kinase G as a signalling hub regulating egress and invasion. Nat Commun 6:7285

Butcher BA, Fox BA, Rommereim LM, Kim SG, Maurer KJ, Yarovinsky F, Herbert DR, Bzik DJ, Denkers EY (2011) *Toxoplasma gondii* rhoptry kinase ROP16 activates STAT3 and STAT6 resulting in cytokine inhibition and arginase-1-dependent growth control. PLoS Pathog 7:e1002236

Caballero MC, Alonso AM, Deng B, Attias M, de Souza W, Corvi MM (2016) Identification of new palmitoylated proteins in *Toxoplasma gondii*. Biochim Biophys Acta 1864:400-408

Carruthers VB, Sibley LD (1997) Sequential protein secretion from three distinct organelles of *Toxoplasma gondii* accompanies invasion of human fibroblasts. Eur J Cell Biol 3:114-123

Cobbold SA, Santos JM, Ochoa A, Perlman DH, Llinás M (2016) Proteome-wide analysis reveals widespread lysine acetylation of major protein complexes in the malaria parasite. Sci Rep 6:19722

Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics 13:2513-2526

Dai L, Peng C, Montellier E, Lu Z, Chen Y, Ishii H, Debernardi A, Buchou T, Rousseaux S, Jin F, Sabari BR, Deng Z, Allis CD, Ren B, Khochbin S, Zhao Y (2014) Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. Nat Chem Biol

10:365-70

Drewry LL, Jones NG, Wang Q, Onken MD, Miller MJ, Sibley LD (2019) The secreted kinase ROP17 promotes *Toxoplasma gondii* dissemination by hijacking monocyte tissue migration. Nat Microbiol 4:1951-1963

Dong H, Guo Z, Feng W, Zhang T, Zhai G, Palusiak A, Rozalski A, Tian S, Bai X, Shen L, Chen P, Wang Q, Fan E, Cheng Z, Zhang K (2018) Systematic identification of lysine 2-hydroxyisobutyrylated proteins in *Proteus mirabilis*. Mol Cell Proteomics 17:482-494

Dubey JP, Lindsay DS, Lappin MR (2009) Toxoplasmosis and other intestinal coccidial infections in cats and dogs. Vet Clin North Am Small Anim Pract 39:1009-1034

Elsheikha HM (2008) Congenital toxoplasmosis: priorities for further health promotion action. Public Health 122:335-353

Elsheikha HM, Büsselberg D, Zhu XQ (2016) The known and missing links between *Toxoplasma gondii* and schizophrenia. Metab Brain Dis 31:749–759

Elsheikha HM, Khan NA (2010) Protozoa traversal of the blood–brain barrier to invade the central nervous system. FEMS Microbiol Rev 34 (4):532–553

Elsheikha HM, Zhu XQ (2016) *Toxoplasma gondii* infection and schizophrenia: An inter-kingdom communication perspective. Curr Opin Infect Dis 29:311–318

Fauquenoy S, Morelle W, Hovasse A, Bednarczyk A, Slomianny C, Schaeffer C, Van Dorsselaer A, Tomavo S (2008) Proteomics and glycomics analyses of N-glycosylated structures involved in *Toxoplasma gondii*--host cell interactions. Mol Cell Proteomics 7:891-910

Foe IT, Child MA, Majmudar JD, Krishnamurthy S, van der Linden WA, Ward GE, Martin BR, Bogyo M (2015) Global analysis of palmitoylated proteins in *Toxoplasma gondii.* Cell

Host Microbe 18:501-511

Fox BA, Guevara RB, Rommereim LM, Falla A, Bellini V, Pètre G, Rak C, Cantillana V, Dubremetz JF, Cesbron-Delauw MF, Taylor GA, Mercier C, Bzik DJ (2019) *Toxoplasma gondii* Parasitophorous Vacuole membrane-associated dense granule proteins orchestrate chronic infection and GRA12 underpins resistance to host gamma interferon. mBio 10:e00589-19

Frénal K, Dubremetz JF, Lebrun M, Soldati-Favre D (2017) Gliding motility powers invasion and egress in Apicomplexa. Nat Rev Microbiol 15:645-660

Frénal K, Polonais V, Marq JB, Stratmann R, Limenitakis J, Soldati-Favre D (2010) Functional dissection of the apicomplexan glideosome molecular architecture. Cell Host Microbe 8:343-357

Gaji RY, Johnson DE, Treeck M, Wang M, Hudmon A, Arrizabalaga G (2015) Phosphorylation of a myosin motor by TgCDPK3 facilitates rapid initiation of motility during *Toxoplasma gondii* egress. PLoS Pathog 11:e1005268

Huang J, Luo Z, Ying W, Cao Q, Huang H, Dong J, Wu Q, Zhao Y, Qian X, Dai J (2017) 2-Hydroxyisobutyrylation on histone H4K8 is regulated by glucose homeostasis in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 114:8782-8787

Huang H, Luo Z, Qi S, Huang J, Xu P, Wang X, Gao L, Li F, Wang J, Zhao W, Gu W, Chen Z, Dai L, Dai J, Zhao Y (2018) Landscape of the regulatory elements for lysine 2-hydroxyisobutyrylation pathway. Cell Res 28:111-125

Jeffers V, Sullivan WJ Jr (2012) Lysine acetylation is widespread on proteins of diverse function and localization in the protozoan parasite *Toxoplasma gondii.* Eukaryotic Cell 11:735-742

494 Jensen KD, Hu K, Whitmarsh RJ, Hassan MA, Julien L, Lu D, Chen L, Hunter CA, Saeij JP
495 (2013) *Toxoplasma gondii* rhoptry 16 kinase promotes host resistance to oral infection and
496 intestinal inflammation only in the context of the dense granule protein GRA15. Infect
497 Immun 81:2156-2167
498 Lamarque MH, Roques M, Kong-Hap M, Tonkin ML, Rugarabamu G, Marq JB, Penarete-
499 Vargas DM, Boulanger MJ, Soldati-Favre D, Lebrun M (2014) Plasticity and redundancy
500 among AMA-RON pairs ensure host cell entry of *Toxoplasma* parasites. Nat Commun
501 5:4098
502 Li X, Hu X, Wan Y, Xie G, Li X, Chen D, Cheng Z, Yi X, Liang S, TanF (2014) Systematic
503 identification of the lysine succinylation in the protozoan parasite *Toxoplasma gondii*. J
504 Proteome Res 13:6087-6095
505 Lindsay DS, Dubey JP, Blagburn BL, Toivio-Kinnucan M (1991) Examination of tissue cyst
506 formation by *Toxoplasma gondii* in cell cultures using bradyzoites, tachyzoites, and
507 sporozoites. J Parasitol 77:126-132
508 Liu WG, Xu XP, Chen J, Xu QM, Luo SL, Zhu XQ (2016) MIC16 gene represents a potential
509 novel genetic marker for population genetic studies of *Toxoplasma gondii*. BMC Microbiol
510 16:101
511 Meng X, Xing S, Perez LM, Peng X, Zhao Q, Redoña ED, Wang C, Peng Z (2017) Proteome-
512 wide analysis of lysine 2-hydroxyisobutyrylation in developing rice (*Oryza sativa*) seeds.
513 Sci Rep 7:17486
514 Montoya JG, Liesenfeld O (2004) Toxoplasmosis. Lancet 363:1965-1976
515 Mouveaux T, Oria G, Werkmeister E, Slomianny C, Fox BA, Bzik DJ, Tomavo S (2014)

516 Nuclear glycolytic enzyme enolase of *Toxoplasma gondii* functions as a transcriptional
517 regulator. PLoS One 9:e105820
518 Nie LB, Liang QL, Du R, Elsheikha HM, Han NJ, Li FC, Zhu XQ (2020) Global proteomic
519 analysis of lysine malonylation in *Toxoplasma gondii*. Front Microbiol 11:776.
520 Nitzsche R, Günay-Esiyok Ö, Tischer M, Zagoriy V, Gupta N (2017) A plant/fungal-type
521 phosphoenolpyruvate carboxykinase located in the parasite mitochondrion ensures
522 glucose-independent survival of *Toxoplasma gondii*. J Biol Chem 292:15225-15239
523 Ong YC, Reese ML, Boothroyd JC (2010) *Toxoplasma* rhoptry protein 16 (ROP16) subverts
524 host function by direct tyrosine phosphorylation of STAT6. J Biol Chem 285:28731-40
525 Picchi GF, Zulkievicz V, Krieger MA, Zanchin NT, Goldenberg S, de Godoy LM (2017) Post-
526 translational modifications of *Trypanosoma cruzi* canonical and variant histones. J
527 Proteome Res 16:1167-1179
528 Rosowski EE, Saeij JP (2012) *Toxoplasma gondii* clonal strains all inhibit STAT1
529 transcriptional activity but polymorphic effectors differentially modulate IFNγ induced
530 gene expression and STAT1 phosphorylation. PLoS One 7:e51448
531 Rougier S, Montoya JG, Peyron F (2017) Lifelong persistence of *Toxoplasma* cysts: A
532 questionable dogma? Trends Parasitol 33(2):93-101
533 Shen B, Sibley LD (2014) *Toxoplasma* aldolase is required for metabolism but dispensable for
534 host-cell invasion. Proc Natl Acad Sci U S A 111:3567-3572
535 Shukla A, Olszewski KL, Llinás M, Rommereim LM, Fox BA, Bzik DJ, Xia D, Wastling J,
536 Beiting D, Roos DS, Shanmugam D (2018) Glycolysis is important for optimal asexual
537 growth and formation of mature tissue cysts by *Toxoplasma gondii*. Int J Parasitol

48(12):955-968

Silmon de Monerri NC, Yakubu RR, Chen AL, Bradley PJ, Nieves E, Weiss LM, Kim K (2015) The ubiquitin proteome of *Toxoplasma gondii* reveals roles for protein ubiquitination in cell-cycle transitions. Cell Host Microbe 18:621-633

Sugi T, Tu V, Ma Y, Tomita T, Weiss LM (2017) *Toxoplasma gondii* requires glycogen phosphorylase for balancing amylopectin storage and for efficient production of brain cysts. mBio 8:e01289-17

Sun HJ, Liu XW, Li FF, Li W, Zhang J, Xiao ZX, Shen LL, Li Y, Wang FL, Yang JG (2017) First comprehensive proteome analysis of lysine crotonylation in seedling leaves of *Nicotiana tabacum*. Sci Rep 7:3013

Treeck M, Sanders JL, Elisa JE, Boothroyd JC (2011) The phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites' boundaries. Cell Host Microbe 10:410-419

Walsh CT, Garneau-Tsodikova S, Gatto GJ Jr (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. Angew Chem Int Ed Engl 44:7342-7372.

Wang K, Peng ED, Huang AS, Xia D, Vermont SJ, Lentini G, Lebrun M, Wastling JM, Bradley PJ (2016) Identification of novel o-linked glycosylated *Toxoplasma* proteins by vicia villosa lectin chromatography. PLoS One 11:e0150561

Wu Q, Ke L, Wang C, Fan P, Wu Z, Xu X (2018) Global analysis of lysine 2-hydroxyisobutyrylome upon SAHA treatment and its relationship with acetylation and crotonylation. J Proteome Res 17:3176-3183

Xue B, Jeffers V, Sullivan WJ, Uversky VN (2013) Protein intrinsic disorder in the acetylome

of intracellular and extracellular *Toxoplasma gondii.* Mol Biosyst 9:645-657

Yamamoto M, Standley DM, Takashima S, Saiga H, Okuyama M, Kayama H, Kubo E, Ito H, Takaura M, Matsuda T, Soldati-Favre D, Takeda K (2009) A single polymorphic amino acid on *Toxoplasma gondii* kinase ROP16 determines the direct and strain-specific activation of stat3. J Exp Med 206:2747-2760

Yakubu RR, Weiss LM, Silmon de Monerri NC (2018) Post-translational modifications as key regulators of apicomplexan biology: insights from proteome-wide studies. Mol Microbiol 107:1-23

Yin D, Jiang N, Zhang Y, Wang D, Sang X, Feng Y, Chen R, Wang X, Yang N, Chen Q (2019) global lysine crotonylation and 2-hydroxyisobutyrylation in phenotypically different *Toxoplasma gondii* parasites. Mol Cell Proteomics 18:2207-2224

Yu Z, Ni J, Sheng W, Wang Z, Wu Y (2017) Proteome-wide identification of lysine 2-hydroxyisobutyrylation reveals conserved and novel histone modifications in *Physcomitrella patens*. Sci Rep 7:15553

Legends to Figures:

Figure 1 Proteome-wide identification of lysine 2-hydroxyisobutyrylated sites in *Toxoplasma gondii*. **(A)** Western blot analysis of tachyzoite lysate probed with anti-2-hydroxyisobutyrylated antibodies. **(B)** Distribution of lysine 2-hydroxyisobutyrylated sites per protein. **(C)** Distribution of 2-hydroxyisobutyrylated peptides based on the peptide length. **(D)** The frequency of K_{hib} sites on the 2-hydroxyisobutyrylated protein detected in *T. gondii*.

Figure 2 Gene ontology (GO) classification of the identified 2-hydroxyisobutyrylated proteins based on biological processes **(A)**, cellular components **(B)**, molecular functions **(C)**, and subcellular localization **(D)**.

Figure 3 Characterization of the motifs of lysine 2-hydroxyisobutyrylated peptides. **(A)** the 2-hydroxyisobutyrylated peptide motifs and conserved lysine 2-hydroxyisobutyrylated sites. The height of each letter represents the frequency of the amino acid residue at this position. The middle "K" refers to the lysine 2-hydroxyisobutyrylated sites. **(B)** Heatmap of the amino acid compositions around the lysine 2-hydroxyisobutyrylated sites. The different colors represent the frequency of 20 amino acids surrounding the lysine 2-hydroxyisobutyrylated sites. Dark red and peacock blue show enrichment and depletion, respectively.

Figure 4 Functional enrichment analysis of the 2-hydroxyisobutyrylated proteins in *Toxoplasma gondii* according to the categories of **(A)** cellular component. **(B)** molecular function. **(C)** biological process.

Figure 5 Enrichment analysis of lysine 2-hydroxyisobutyrylated proteins. (A) Protein domain enrichment analysis of lysine 2-hydroxyisobutyrylated proteins. (C) KEGG enrichment analysis of lysine 2-hydroxyisobutyrylated proteins.

Figure 6 PPI networks of lysine 2-hydroxyisobutyrylated proteins in *Toxoplasma gondii*.