Partial catalytic Cys oxidation of human GAPDH [version 1; peer review: awaiting peer review]

Andrea Lia¹², Adam Dowle ¹³, Chris Taylor³, Angelo Santino ², Pietro Roversi ¹

¹Leicester Institute of Chemical and Structural Biology and Department of Molecular and Cell Biology, University of Leicester, Henry Wellcome Building, Lancaster Road, LE1 7HB, UK
²Institute of Sciences of Food Production, C.N.R. Unit of Lecce, ia Monteroni, Lecce, 73100, Italy
³Bioscience Technology Facility Department of Biology, University of York, Wentworth Way, York, YO10 5DD, UK

Abstract

Background: n-Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) catalyses the reversible NAD⁺-dependent oxidative phosphorylation of n-glyceraldehyde-3-phosphate to 1,3-diphospho-n-glycerate in both glycolysis and gluconeogenesis.

Methods: Four distinct crystal structures of human n-Glyceraldehyde-3-phosphate dehydrogenase (HsGAPDH) have been determined from protein purified from the supernatant of HEK293F human epithelial kidney cells.

Results: X-ray crystallography and mass-spectrometry indicate that the catalytic cysteine of the protein (HsGAPDH Cys152) is partially oxidised to cysteine S-sulfonic acid. The average occupancy for the Cys152-S-sulfonic acid modification over the 20 crystallographically independent copies of Hs GAPDH across three of the crystal forms obtained is 0.31±0.17.

Conclusions: The modification induces no significant structural changes on the tetrameric enzyme, and only makes aspecific contacts to surface residues in the active site, in keeping with the hypothesis that the oxidising conditions of the secreted mammalian cell expression system result in Hs GAPDH catalytic cysteine S-sulfonic acid modification and irreversible inactivation of the enzyme.

Keywords

GAPDH, cysteine S-sulfonic acid, X-ray crystallography, mass-spectrometry, moonlight enzyme
Corresponding author: Pietro Roversi (pr159@leicester.ac.uk)

Author roles: Lia A: Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Dowle A: Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Taylor C: Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Santino A: Conceptualization, Funding Acquisition, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; Roversi P: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the Wellcome Trust through a LISCBI Wellcome Trust ISSF award [204801; P.R.] and a Wellcome Trust Seed Award in Science [214090; P.R.]. A.L. is the recipient of a PhD Studentship from the Agricultural and Forestry Department (DAFNE) of the University of Tuscia (Viterbo, Italy) and funding from the Italian National Research Council (to A.S.). The York Centre of Excellence in Mass Spectrometry was created thanks to a major capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC grants [EP/K039660/1; EP/M028127/1]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2020 Lia A et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Lia A, Dowle A, Taylor C et al. Partial catalytic Cys oxidation of human GAPDH [version 1; peer review: awaiting peer review] Wellcome Open Research 2020, 5:114 https://doi.org/10.12688/wellcomeopenres.15893.1

First published: 01 Jun 2020, 5:114 https://doi.org/10.12688/wellcomeopenres.15893.1
Introduction
Mammalian HEK293F cells are routinely used in conjunction with secreted protein expression vectors for recombinant protein production. They owe their popularity to ease of handling, robust growth rate, excellent transfectability, high capacity for recombinant protein expression, low-cost media requirements and low levels of secreted contaminants. Of course, in absence of secretion of the desired recombinant protein at hand, contaminants are the only proteins present in recombinant protein expression systems. In the course of a research effort aimed at the purification of recombinant Chaetomium thermophilum ERAD mannosidase CHTM1P from the supernatant of HEK293F cells, we purified, crystallised and determined crystal structures of human n-Glyceraldehyde-3-phosphate dehydrogenase (HsGAPDH, EC 1.2.1.12). GAPDH catalyses the reversible NAD+-dependent oxidative phosphorylation of n-glyceraldehyde-3-phosphate to 1,3-diphospho-n-glycerate in both glycolysis and gluconeogenesis. This reaction occurs at an important transition point in glycolysis between the enzymatic steps that consume and generate ATP. Several studies indicate that the enzyme has pleiotropic functions independent of its canonical role in glycolysis. In addition to the somatic cells HsGAPDH isoform, the human genome encodes a testis-specific isoform HsGAPDHS (G3PT_HUMAN (Uniprot O14556)), which is expressed only in the post-meiotic period of spermatogenesis. In multiple mammalian species the sperm isoyme shares about 70% amino acid identity with the somatic isoyme and plays a role in anchoring GAPDHS to the fibrous sheath in the principal piece of the sperm flagellum.

GAPDH was one of the first enzymes to be crystallised and one of the first enzymes whose structure was determined by X-ray crystallography. A number of studies have reported the GAPDH catalytic Cys carrying oxidised post-translational modifications. In particular, cysteine sulfenic acid (CSX, see Figure 1A) has been observed in the structure of rabbit muscle GAPDH (of course, by X-ray diffraction alone, this modification may be difficult to distinguish from S-hydroxycysteine (CSO, see Figure 1C)). Cysteine-S-sulfenic acid (CSD, see Figure 1B) has been observed in structures of rat GAPDHS and E. coli GAPDH. Cysteine-S-sulfonic acid (CSU, see Figure 1D) modifies the catalytic Cys in a structure of Streptococcus pneumoniae GAPDH.

In this paper, we present and discuss mass spectrometry and crystallographic evidence supporting partial oxidation of the catalytic Cys residue of HsGAPDH purified from the supernatant of a HEK293F cell culture. A fraction of the molecules shows Cys-sulfonic acid instead of Cys in the active site. The four HsGAPDH crystal structures we describe (which we label as P2₁ forms A-D, see Table 1) represent novel HsGAPDH monoclinic polymorphs and are the first HsGAPDH crystal structures.

Figure 1. Cys oxidised modifications and GAPDH PDB entries that carry them at the catalytic Cys152. A: S-Oxy-Cysteine aka Cysteine sulfenic acid (CSX, PDB ID 1J0X); B: S-Cysteinesulfinic acid aka S-sulfinocysteine aka 3-Sulfino-L-alanine (CSD, PDB IDs 2VYN and 2VYV); C: S-Oxy-Cysteine (CSO, not described in any GAPDH structures); D: Cysteine-S-sulfonic acid (CSU, PDB ID 5M6D and this work)
Table 1. Human GAPDH crystal structures.

<table>
<thead>
<tr>
<th>Isoform (Uniprot)</th>
<th>PDB ID</th>
<th>Polymorph (Z)</th>
<th>Resolution (Å)</th>
<th>Tetramer PG</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>G3P_HUMAN (P04406)</td>
<td>6YN</td>
<td>P2, Form A (16)</td>
<td>1.54</td>
<td>1</td>
<td>This work</td>
</tr>
<tr>
<td>G3P_HUMAN (P04406)</td>
<td>6YNE</td>
<td>P2, Form B (8)</td>
<td>1.85</td>
<td>1</td>
<td>This work</td>
</tr>
<tr>
<td>G3P_HUMAN (P04406)</td>
<td>6YNF</td>
<td>P2, Form C (16)</td>
<td>2.39</td>
<td>1</td>
<td>This work</td>
</tr>
<tr>
<td>3GPD</td>
<td>C2 Form A (8)</td>
<td>3.50</td>
<td>2</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>4WNC</td>
<td>P2, Form E (16)</td>
<td>1.99</td>
<td>1</td>
<td>21,22</td>
<td></td>
</tr>
<tr>
<td>4WNI</td>
<td>P2, Form F (16)</td>
<td>2.30</td>
<td>1</td>
<td>21,22</td>
<td></td>
</tr>
<tr>
<td>6IQ6</td>
<td>P2, Form F (16)</td>
<td>2.29</td>
<td>1</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>6ADE</td>
<td>I222 (24)</td>
<td>3.15</td>
<td>2 and 222</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>G3PT_HUMAN (O14556)</td>
<td>3H9E, 3PFW, 5C7O</td>
<td>C2 Form B (8)</td>
<td>1.72, 2.15, 1.73</td>
<td>2</td>
<td>24,10</td>
</tr>
<tr>
<td>5C7O</td>
<td>P3,21 (12)</td>
<td>1.86</td>
<td>2</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

carrying a Cysteine-S-sulfonic acid modification of the catalytic cysteine Cys152. Crystal form A is the highest resolution HsGAPDH structure determined to date (1.52 Å).

Methods
Cloning
The cloning of the Chaetomium thermophilum ERAD mannosidase ChtHM1 (the product of the gene CTHT_0058730, Uniprot entry G0SCX7_CHATD) in the ChtHM1<sub>so-tsg</sub>-pHLsec vector for secreted mammalian cell expression is described in detail in the Open Laboratory Notebook page (see Extended data<sup>25</sup>). Briefly, the DNA encoding ChtHM1<sub>so-tsg</sub> was PCR amplified from a commercially obtained plasmid encoding the full length gene, and inserted into the pHLsec vector<sup>2</sup> using ligation independent cloning: the ChtHM1<sub>so-tsg</sub> DNA insert and the AgeI/KpnI linearised pHLsec DNA were mixed in 3:1 molar ratio; 0.06 pmol of ChtHM1<sub>so-tsg</sub> DNA (122 ng) and 0.02 pmol of AgeI/KpnI linearised pHLsec DNA (60 ng). To this DNA, 10 µL of Gibson Assembly MasterMix (2X) (New England Bioscience E2611L) were added and the total volume made 20 µL with deionised water. The mix was heated at 50 °C for 1 hour.

Protein expression
A volume of 450 µL of HEK293F cells at a concentration of 10<sup>6</sup> cells/mL suspended in Gibco FreeStyle 293 Media (Thermo Fisher Scientific 12338018) was transfected with the ChtHM1<sub>so-tsg</sub>-pHLsec vector, using the FreeStyle MAX 293 expression system (Thermo Fisher K900010). Briefly, 1 µg of DNA was used per mL of culture: the DNA vector was initially dissolved in 45 mL of phosphate-buffered saline (PBS: 0.01 M phosphate buffer pH 7.4, 0.0027 M potassium chloride and 0.137 M sodium chloride; PBS tablets, Sigma Aldrich P4417) and vortexed vigorously for 3 seconds; 1.8 mL of a filter-sterilised solution of 0.5 mg/mL polyethyleneimine (PEI) was added to the PBS/DNA solution and vortexed vigorously for 3 seconds; the mixture was incubated at room temperature for 20 minutes; the DNA/PEI mixture was added to the cell culture; the cell culture was made 5 µM kifunensine (an inhibitor of endoplasmic reticulum and Golgi mannosidases, Cayman Chemical 109944-15-2). The cell culture was split into three 500 mL Erlenmeyer flasks with 0.2 µm vent caps (Corning), with 150 mL of culture in each flask, and incubated in an orbital shaker incubator at 37°C, shaking at 120 rpm, under a 5% CO<sub>2</sub> atmosphere. The cells’ supernatant was harvested 4 days post-transfection by centrifuging at 4,000 x g for 5 minutes.

Protein purification
The cell supernatant was made 1x PBS by addition of the appropriate volume of 10x PBS stock (obtained by dissolving five PBS tablets (Sigma Aldrich P4417) in 200 mL of deionised water). The pH was adjusted to 7.4 and the solution filter-sterilised through a 0.2 µm bottle-top filter. Nickel immobilised metal affinity chromatography (IMAC) was used as the first step of purification. A 5 mL HisTrap HP column (GE Healthcare 17-5248-01) was equilibrated with five column volumes (CV) of binding buffer (50 mM sodium phosphate pH 7.5, 300 mM NaCl, 1 mM TCEP (Sigma Aldrich C4706)). The cell culture supernatant was loaded onto the column at room temperature at a flow rate of 3 mL/min. After a 5CV wash in binding buffer, the column was fitted to an ÄKTA purifier fast protein liquid chromatography (FPLC) machine in a 4°C cabinet and further washed with 6 mL of elution buffer (50 mM sodium phosphate pH 7.5, 500 mM imidazole (Honeywell Fluka 56750), 200 mM NaCl).

The wash fractions from the IMAC run (total volume 4.5 mL) were pooled and concentrated to 2 mL in a polyethersulfone spin concentrator of 10 kDa MW cut-off (Thermo Fisher Scientific 88513). The concentrated sample was filtered through a 0.22 µm spin filter and loaded on a 2 mL loop connected to an ÄKTA purifier FPLC machine in a 4°C cabinet. The sample...
was injected on a size exclusion chromatography (SEC) S200 16/60 column (GE Healthcare 28-9893-35) equilibrated in filter-sterilised and degassed 20 mM HEPES pH 7.5, 100 mM NaCl and 1mM TCEP (Sigma Aldrich C4706) buffer, and run down the column at 0.4mL/min flow rate, collecting 1 mL elution fractions. 

HsGAPDH-containing fractions were pooled and concentrated as described above to a volume of 70 µL, and protein concentration measured by loading 1.5 µL of sample on a NanoDrop 1000 spectrophotometer (Thermo Scientific). The absorbance was OD$_{280}^{280}$=15.0, equivalent to a concentration of 18.8 mg/mL (calculated from $\epsilon_{280}=0.7963 (mg/mL)^{-1} cm^{-1}$).

A volume of 30 µL of the HsGAPDH obtained from the S200 SEC run was diluted to 500 µL in filter-sterilised and degassed buffer HEPES (Sigma Aldrich H3375) 20 mM pH7.5, NaCl 100 mM, TCEP 1mM and injected onto a Superdex S200 10/300 column equilibrated in the same buffer. The mass of the sample was detected on elution with an 18-angle multiangle light scattering detector (Dawn HELEOS II) coupled with a differential Refractive Index detector (Optilab T-rEX) (Wyatt Technology). Figure 2 illustrates the results of the run. The three fractions C2-C3-C4 eluting between 12.5 and 14 mL were pooled and concentrated to a volume of 45 µL - and a concentration of 3.0 mg/mL.

Protein crystallisation

Vapour diffusion crystallisation 200 nL sitting drops were set up using a Mosquito crystallisation robot (SPT Labtech), mixing HsGAPDH in mother liquor:protein ratios 1:1 (drop 1) and 1:2 (drop 2), with the 96-conditions MORPHEUS crystallisation screen (Molecular Dimensions$^{26,27}$), and the drops were left equilibrating at 18°C. P$_2$, Forms A and B: the protein from the S200 SEC run was set up for crystallisation at a concentration of 18.8 mg/mL. Form A: these crystals grew in drop 2 equilibrated against condition H9 of the MORPHEUS crystallisation screen (Molecular dimensions$^{26,27}$): 0.1 M Amino acids solution (DL-Glutamic acid, DL-Alanine, Glycine, DL-Lysine, DL-Serine); 0.1 M Buffer System 3 (Tris (base), bicine pH 8.5); 30% v/v Precipitant Mix 1 (40% v/v PEG 500MME, 20% w/v PEG 20000). Form B: these crystals grew in drop 2 equilibrated against condition F9 of the MORPHEUS crystallisation screen$^{26,27}$: 0.12M monosaccharides solution (D-Glucose, D-Mannose, D-Galactose, L-Fucose, D-Xylose, N-Acetyl-D-Glucosamine); 0.1M Buffer System 3 (Tris (base), bicine pH 8.5) 30% v/v Precipitant Mix 1 (40% v/v PEG 500MME, 20% w/v PEG 20000).

P$_2$, Forms C and D: the protein from the size exclusion chromatography multiangle light scattering (SEC-MALS) run was set up for crystallisation at a concentration of 3.0 mg/mL. Both form C and from D grew in drop 1 equilibrated against the F1 condition of the MORPHEUS crystallisation screen$^{26,27}$: 0.12 M monosaccharides solution (D-Glucose, D-Mannose, D-Galactose, L-Fucose, D-Xylose, N-Acetyl-D-Glucosamine); 0.1 M Buffer System 3 (Tris (base), bicine pH 6.5) 30% v/v Precipitant Mix 1 (40% v/v PEG 500MME, 20% w/v PEG 20000).

X-ray data diffraction collection and processing

All crystals were cryo-cooled by plunging them into liquid nitrogen. X-ray diffraction data were collected at beamline I03 of the Diamond Light Source in Harwell, England, UK, with $^{28}$.
an X-ray beam of wavelength $\lambda=0.97622$ Å and size 80x20 μM. Other data collection parameters are listed in Table 2. X-ray diffraction data were processed with the autoPROC suite of programs version 1.0.5\textsuperscript{31} with the following command line options: process -h5 Datapath/Data_master.h5 -M HighResCutOnCChalf -d DataOutputDirectory. An alternative Open Source suite that would enable X-ray diffraction data processing in equivalent ways is xia2\textsuperscript{28} running DIALS\textsuperscript{30}.

Structure determination and refinement

P2\textsubscript{1}, Form A: initial structure factor phases were computed by molecular replacement, searching for eight copies of the HsGAPDH monomer in PDB ID 1U8F in space group P2\textsubscript{1}, using the program CCP4-Molrep version 11.7.02\textsuperscript{31} with all default parameters. The eight copies of the HsGAPDH monomer are arranged in two tetramers in the asymmetric unit. Initial automated water addition and positional and individual B-factor refinement were carried out in autoBUSTER, version 2.10.3\textsuperscript{32}. An alternative Open Source piece of software that would enable refinement against X-ray diffraction data in equivalent ways is Vagabond\textsuperscript{33}. autoBUSTER was run with the following command line options: refine -m Data. staraniso_alldata-unique.mtz -p Model.pdb -autoncs -Seq GAPDH.seq -d OutputDir -l XPE. grade_PDB_ligand.cif ExcludeBadContacts=EXCLUDE *|203:*|1203:*' AutomaticallyFormfactorCorrection=6YN. Refinement statistics are reported in Table 3. P2\textsubscript{1}, Forms B, C and D: initial structure factor phases were computed by molecular replacement, searching with the P2\textsubscript{1} Form A tetramer (PDB ID 6YN) in space group P2\textsubscript{1}, placing one tetramer per asymmetric unit in P2\textsubscript{1} Form B and D and two tetramers per asymmetric unit in P2\textsubscript{1} Form C, using the program CCP4-Molrep\textsuperscript{34} version 11.7.02, with all default parameters. The same refinement protocol and restraints were used a described for P2\textsubscript{1} Form A, but with additional external secondary structure restraints \textsuperscript{35} to the highest resolution P2\textsubscript{1} Form A structure, using the extra command line flag to -ref 6YND.pdb when running autoBUSTER, version 2.10.3\textsuperscript{32}. An alternative Open Source piece of software that would enable refinement against X-ray diffraction data in equivalent ways is Vagabond\textsuperscript{33}. Crystal forms B, C and D were deposited in the PDB with ID codes 6YN, 6YNF and 6YNH, respectively. Refinement statistics are reported in Table 3.

| Table 2. Human GAPDH crystal X-ray diffraction data collection parameters and data processing statistics. Values in parentheses refer to the highest resolution shell. |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|
| PDB ID                         | Form A          | Form B          | Form C          | Form D          |
| PDB dist., d\textsubscript{max} (mm, Å) | 6YN | 6YN | 6YN | 6YN |
| Det. flux (photons/s)           | 8.84x10\textsuperscript{11} | 8.88x10\textsuperscript{11} | 8.85x10\textsuperscript{11} | 3.88x10\textsuperscript{12} |
| Transmission                    | 25% | 25% | 25% | 100% |
| Number of images                | 3,600 | 3,600 | 3,600 | 3,600 |
| Oscillation range (°)           | 0.10 | 0.10 | 0.1 | 0.25 |
| Exposure time (s)               | 0.05 | 0.05 | 0.03 | 0.013 |
| Space Group                     | P2\textsubscript{1} | P2\textsubscript{1} | P2\textsubscript{1} | P2\textsubscript{1} |
| Cell edges: a,b,c (Å)           | 81.88, 124.45, 141.99 | 81.79, 124.65, 79.64 | 87.14, 111.43, 135.94 | 87.02, 111.30, 69.74 |
| Cell angle β (°)                | 99.38 | 117.04 | 96.02 | 98.33 |
| Resolution Range (Å)            | 93.04-1.52 (1.71-1.52) | 72.85-1.85 (2.05-1.85) | 135.18-2.39 (2.74-2.39) | 86.10-2.62 (2.89-2.62) |
| R\textsubscript{merge}           | 0.08 (1.07) | 0.226 (1.488) | 0.27 (1.28) | 0.40 (2.47) |
| R\textsubscript{free}            | 0.09 (1.16) | 0.245 (1.618) | 0.29 (1.38) | 0.41 (2.54) |
| Observations                    | 2,029,459 (100,354) | 568,580 (26,433) | 418,478 (20,849) | 496,488 (24,336) |
| Unique observations             | 295,270 (14,763) | 81,465 (4,074) | 59,243 (2,962) | 28,207 (1,409) |
| Average I/σ(I)                  | 11.7 (1.7) | 6.5 (1.5) | 6.5 (1.6) | 9.0 (1.5) |
| Completeness                    | 69.4 (12.2) | 67.6 (12.8) | 58.2 (8.7) | 71.3 (14.0) |
| Multiplicity                     | 6.9 (6.8) | 7.0 (6.5) | 7.1 (7.0) | 17.6 (17.3) |
| CC1/2                           | 0.997 (0.608) | 0.991 (0.456) | 0.986 (0.591) | 0.990 (0.464) |
Table 3. Human GAPDH crystal structures refinement statistics. Values in parentheses refer to the highest resolution shell, unless otherwise specified.

<table>
<thead>
<tr>
<th></th>
<th>Form A</th>
<th>Form B</th>
<th>Form C</th>
<th>Form D</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB ID</td>
<td>6YND</td>
<td>6YNE</td>
<td>6YNF</td>
<td>6YNH</td>
</tr>
<tr>
<td>Space group (Z)</td>
<td>P2₁</td>
<td>P2₁</td>
<td>P2₁</td>
<td>P2₁</td>
</tr>
<tr>
<td>Resolution range</td>
<td>1.63-1.52</td>
<td>1.98-1.85</td>
<td>2.62-2.39</td>
<td>2.72-2.62</td>
</tr>
<tr>
<td>Reflex. s working set</td>
<td>280,585</td>
<td>77,398</td>
<td>56,303</td>
<td>25,751</td>
</tr>
<tr>
<td>Reflex. s free set</td>
<td>14,686</td>
<td>4,059</td>
<td>2,940</td>
<td>1,048</td>
</tr>
<tr>
<td>R, R_{ref}</td>
<td>0.18,0.199</td>
<td>0.190,0.207</td>
<td>0.179,0.219</td>
<td>0.176,0.215</td>
</tr>
<tr>
<td>Ramachandran fav.</td>
<td>97.4%</td>
<td>97.5%</td>
<td>97.0%</td>
<td>97.5%</td>
</tr>
<tr>
<td>Ramachandran allow.</td>
<td>99.7%</td>
<td>99.7%</td>
<td>99.7%</td>
<td>99.7%</td>
</tr>
<tr>
<td>Tetramers</td>
<td>(A,E,H)</td>
<td>(B,D,F,G)</td>
<td>(A,C,E,H)</td>
<td>(B,D,F,G)</td>
</tr>
<tr>
<td>Occ. Cys-SO₂H</td>
<td>A:0.13</td>
<td>C:0.09</td>
<td>E:0.13</td>
<td>G:0.15</td>
</tr>
<tr>
<td>Prot.(Wat.) Atoms</td>
<td>20,428</td>
<td>10,108</td>
<td>20,152</td>
<td>14,686</td>
</tr>
<tr>
<td>(B)ₜₜₜₜ (B)ₜₜₜₜ (Å²)</td>
<td>27.26</td>
<td>29.40</td>
<td>39.74</td>
<td>43.25</td>
</tr>
</tbody>
</table>

Mass spectrometry

Protein was purified by 1D-PAGE before Coomassie staining and excision for proteomic analysis. Protein was digested in-gel with the addition of 20 ng sequencing grade trypsin (Promega V5111). No reduction or alkylation was performed during digestion to help preserve the native state of oxidation. Resulting peptides were analysed over a 1 h liquid chromatography–mass spectrometry acquisition with elution from a 50 cm EasyNano C18 column as detailed in Figure 1B onto a Thermo Orbitrap Fusion Tribrid mass spectrometer. Spectra were acquired in data dependent acquisition mode with sequential fragmentation of all selected peptide precursors using both higher-energy collisional dissociation and electrontransfer dissociation (ETD). All spectra were acquired in the Orbitrap mass analyser with internal calibration from the ETD reagent.

Product ion spectra were searched against the expected sequence of *Hs*GAPDH (Uniprot P04406) appended to a custom in-house database. All pieces of software were run with default parameters, with the exception that propionamide (+71.037114) and dehydro (-1.007825 Da) were added as manually defined variable modifications. The analysis was carried out using Mascot version 2.6.1 (Matrix Science), PEAKS Studio X+10.5 (Bioinformatics Solutions Inc.) and Byonic™ 3.0 (Protein Metrics Inc) search engines (Figure 3A and Figure 4A). The same data analysis was also carried out with freeware, using a combination of MSConvert version 3.0, SearchGui version 3.3.18 and PeptideShaker version 1.16.45 (Figure 3B and Figure 4B). The resulting peptide assignments are equivalent to the ones obtained with the commercial pieces of software, see Figure 3 and Figure 4. Searching specified a precursor tolerance of 3 ppm and a fragment ion tolerance of 0.02 Da. Initial searches included variable modification of: S-Oxy Cysteine (Figure 1A), Cysteine S-sulphinic acid (Figure 1B), and Cysteine S-sulphonic acid (Figure 1D). Searches were then expanded in PEAKS to include 313 of the most frequently observed proteomic modifications and in Byonic™ to include disulfide bonding and wildcard mass addition to Cys.

Peptide spectra matches were filtered to p-values (or equivalents) of <0.05. Individual spectra were manually inspected for site localisation or ambiguity before reporting. Global protein coverage of over 90% was achieved. Individual post-translational modification (PTM) modified spectra were manually inspected for site localisation specificity and potential PTM ambiguity before reporting.

Results

Since the first crystal structures of the lobster GAPDH enzyme several crystal structures of prokaryotic and eukaryotic GAPDHs have been determined. At the time of this writing, the PDB contains more than 142 GAPDH entries, 11 of which are of human GAPDH isoforms (see Table 1). All human GAPDH crystal structures described so far contain tetramers, sitting in crystal sites of three different symmetries: point group 1 (four copies of the monomer all slightly different from each other); point group 2 (dimer of dimers); and point group 222 (four identical copies in a tetramer of exact 222 symmetry).
**Figure 3.** Product ion spectrum of GAPDH-derived tryptic peptide **IISNASCTTNCLAPLAK** lacking Cys modification. Product ions are annotated with theoretical b- and y- ions (fragment ions that appear to extend from the amino- or carboxy-terminus of a peptide, respectively) from the peptide assignment IISNASCTTNCLAPLAK. 

A: data analysed using Mascot™, PEAKS and Byonic™ search engines, with all default parameters.

B: data analysed using MSConvert, SearchGui and PeptideShaker.

*HsGAPDH* crystals of form A and form B grew from the *HsGAPDH* protein sample purified from the supernatant of HEK293F cells by two chromatography steps (IMAC+SEC); crystals of forms B and form C grew from the same *HsGAPDH* protein sample after an extra SEC-MALS chromatography step, see Figure 2. Our *HsGAPDH* crystal structures all contain tetramers of point group 1 (two tetramers per asymmetric unit in forms A and C, and one tetramer per asymmetric units in forms B and D). Overall, the four novel crystal forms contain 24 crystallographically independent observations of the *HsGAPDH* molecule, adding to the 29 ones already present in the PDB at the time of this writing (see Table 1). The overall rmsd is 0.170 Å across these 24 crystallographically independent copies of *HsGAPDH*.

The protein sample from which crystal forms C and D were grown was analysed by mass spectrometry. In addition to peptides corresponding to unmodified Cys152 (see Figure 3), we detected ions derived from fragmentation of the Cys152-containing tryptic peptide **IISNASCTTNCLAPLAK**, which suggests Cys152 S-sulfonic acid modification (Figure 4). Annotation is complicated by the presence of a further modification at Cys156 with a mass addition of +71.036 Da, as directly observed in the fragmentation spectrum by the mass spacing between the y-6 and y-7 ions (Figure 4). The +71.036 Da mass addition at Cys156 is best annotated as acrylamide adduct (propionamide) imparted during PAGE separation rather than being a native modification at this residue. The measured mass of the peptide was 1869.8623 Da and the theoretical mass for the unmodified peptide is 1718.8695 Da. Including the annotated acrylamide adduct, this leaves 79.9568 Da unaccounted for, which is an extremely close match to S-sulfonic acid modification of Cys (+79.9568 Da). A potential ambiguity for the assignment is the similarity in mass of phosphorylation (+79.966331); however, phosphorylation would equate to mass error of 9.5 mDa as opposed to just 0.018 mDa for S-sulfonic acid. Unlike in the case of the acrylamide adduct, the +79.9568 Da mass addition cannot be observed in any of the individual b- or y-ions in the fragmentation spectrum (Figure 4), implying complete neutral loss of the moiety between precursor selection and fragmentation. Although neutral loss is frequently observed with Ser and Thr phosphorylation, it is rarely seen with 100% efficiency, further suggesting sulfation of Cys152 as the most credible assignment. Relative quantification of modified versus unmodified peptide forms is not possible based on the mass spectrometry data alone, due to unknown but likely significant differences in relative response resulting from the changes in the physico-chemical properties of the Cys152-containing ions upon Cys-S-sulfation.
Figure 4. Product ion spectrum of GAPDH-derived tryptic peptide IISNASCTTNCLAPLAK. A precursor equating to a mass of 1869.8623 Da was selected and fragmented by HCD. Resulting ions are annotated with theoretical b- and y-ions from the peptide assignment IISNASC(+79.9568 Da)TTNC(+71.036 Da)LAPLAK using the Byonic™ search engine. Annotation of acrylamide adduct modification (+71.036 Da) at Cys156 is implicitly localised from the m/z difference between the product ions y6 and y7. Cys-S-sulfonic acid modification at Cys152 is implied from the difference in mass between the precursor selected and the resulting y11-15 m/z values observed as neutral mass products labelled with “~”. The measured neutral mass difference is 79.9568 Da, equating to a mass error of 0.018 mDa for S-sulfonic acid modification. A: data analysed using Mascot, PEAKS and Byonic™ search engines, with all default parameters. B: data analysed using MSConvert, SearchGui and PeptideShaker.

Figure 5. Fo-Fc difference crystal electron density map around HsGAPDH Cys152 residues. Representative HsGAPDH monomers in crystal form C (PDB ID 6YNF) are shown: A: monomer B; B: monomer E; C: monomer G. The 2.5 σ contour of the unbiased Fo-Fc map is depicted as a green mesh (the map was calculated with phases from the model before the Cys-S-sulfonic acid modification was built). Residue HsGAPDH 152 was modelled and refined as a 0.28:0.72 superposition of Cys:Cys-S-sulfonic acid. C atoms: cyan; N atoms: blue; O atoms: red; S atoms: yellow. Selected non-bonding interactions between the modified Cys and neighbouring residues are represented as white dashed lines with distances in Å. The Figure was made in PyMOL. An alternative Open Source piece of software is USF Chimera.
Peaks suggesting Cys152 S-sulfonic acid modification are also visible in the initial crystal electron density Fo-Fc difference maps of crystal forms A, C and D, contoured at 2.5σ level or higher, in close proximity to the Cys152 γ sulfur atom of most HsGAPDH chains in the asymmetric unit (see Figure 5). The orientation of the Cys-S-sulfonic acid moiety in the active site is not unique, but it varies from chain to chain across the three crystal forms. Figure 6A illustrates the modification in the eight copies of HsGAPDH in crystal form C. The Cys152 S-sulfonic acid chain makes only loose non-bonding contacts to residues in the active site (see also Figure 5). Modelling of the Cys152 residue as a superposition of Cys and Cys-S-sulfonic acid with variable occupancies in crystal forms A, C and D enables estimation of the extent to which the crystals contain the Cys152 modification. The average occupancy for the Cys152-S-sulfonic acid modification over these 20 crystallographically independent copies of HsGAPDH is 0.31±0.17 (see Table 3). The HsGAPDH monomer with the largest occupancy for the modification (Occ$_{Cys-SO_3^-}$=0.72) is monomer H in crystal form C. This predominantly Cys-S-sulfonated HsGAPDH monomer superposes with an rmsd$_{C_{\alpha}}$=0.423 Å (over 334 C$_{\alpha}$ atoms) with the Streptococcus pneumoniae GAPDH monomer carrying the same modification$^{17}$, see Figure 6B. When comparing the modified structure to the one of HsGAPDH with unmodified Cys152 (PDB ID 1U8F) no significant local changes to active site side chains are observed either (rmsd$_{C_{\alpha}}$=0.127 over 333 C$_{\alpha}$ atoms), see Figure 6C. None of the HsGAPDH molecules in the crystals contain NAD$^+$ nor any bound ligand in the active site.

$^{1}$ Judging from the crystal electron density difference maps, crystal form B does not seem to have significant Cys152 S-sulfonic acid modification, so in this crystal form residue 152 was modelled as Cys in all chains.

---

Figure 6. Cys152-S-sulfonic acid modification of HsGAPDH. N atoms: blue; O atoms: red; S atoms: yellow. A: the Cys152-S-sulfonic acid modification in the 8 copies of HsGAPDH in crystal form C. The GAPDH monomer is in surface representation and coloured blue to red from N- to C-terminus. The inset shows the Cys152 S-sulfonic acid chains adopting slightly different conformations in different molecules in the asymmetric unit. B: superposition of the active site of Cys152-S-sulfonic acid modified HsGAPDH (chain H in crystal form C, green C atoms) with two copies of Cys152-S-sulfonic acid modified Streptococcus pneumoniae GAPDH (SpGAPDH, PDB ID 5M6D, chain “A”, white C atoms; chain “B”, yellow C atoms). C: superposition of the active site of Cys152-S-sulfonic acid modified HsGAPDH (chain H in crystal form C, green C atoms) with native HsGAPDH (chain “O” of PDB ID 1U8F, cyan C atoms). The side chain of the catalytic Cys152 of native HsGAPDH in PDB ID 1U8F was modelled and refined as a superposition two alternate conformations. The Figure was made in PyMOL$^{47}$. An alternative Open Source piece of software is USF Chimera$^{48}$. 

---

Page 10 of 13
Discussion and conclusions

GAPDH is mainly a cytoplasmic enzyme, but it has been reported to moonlight as a cell surface and/or extracellular enzyme in healthy human cells\textsuperscript{60,61}, in human cells subjected to stresses such as micronutrient starvation, hypoxia, infection, and cancer\textsuperscript{62–64}, and to act as an extracellular effector in fungi and bacteria\textsuperscript{33–62}. In our study, HsGAPDH was purified from the supernatant of HEK293F cells treated with the mannosidase inhibitors kifunensine and kept for 4 days at 37°C after transfection with a DNA plasmid for recombinant expression of an endoplasmic reticulum glycoprotein. Although kifunensine is toxic to cells only in concentrations higher than the one used (5μM), during the 4 days between transfection and harvesting of the supernatant the experimental conditions are likely harsher than the basal levels of oxidative stress in the HEK293F cell culture media\textsuperscript{63}.

Stable oxidation of the GAPDH catalytic cysteine has been observed in a number of studies, for example to Cys sulfenic acid mediated by thiolate oxidation\textsuperscript{64}; the modification is reversible but it confers the modified cysteine electrophilic properties similar to the ones of sulfenyl halides\textsuperscript{65}. For example, GAPDH catalytic cysteine modification to sulfenic acid converts the enzyme from dehydrogenase to an acyl phosphatase\textsuperscript{66–68}. In the presence of excess oxidant, further oxidation of sulfenic acid to sulfinic acid or sulfonic acid is possible. However, in contrast to sulfenic acid, higher oxidation states of catalytic Cys residues lead to enzyme inactivation and Cys S-sulfonation is generally considered an irreversible form associated with protein misfolding, degradation, and pathology\textsuperscript{69}. In particular, S-sulfonated GAPDH has been reported to translocate to subcellular domains where it does not normally occur, where it may stimulate a “gain of function” that could provoke apoptosis\textsuperscript{70}.

It is unclear whether the Cys S-sulfonic acid modification we observe in HsGAPDH purified from the supernatant of HEK293F cells represents a physiological response of the cells to the recombinant secreted protein expression system conditions, or whether it is simply the outcome of a protracted oxidative environment on the pool of extracellular enzyme instead. GAPDH allostery is mediated by intra-dimer contacts (the active site of one dimer subunit is in close proximity to residues in the neighbouring subunit). The observation that HsGAPDH tetramers pack in the crystals in ordered fashion even in presence of partial Cys-S-sulfonic acid modification, together with the overall structural similarity of the Cys S-sulfonated and native monomeric enzymes, suggests that the secreted HsGAPDH is constantly inactivated by catalytic Cys S-sulfonation in the expression system oxidising extracellular environment.

Data availability

Underlying data

X-ray data of forms A, B, C and D at Protein Data Bank, Accession numbers 6YND, 6YNE, 6YNF and 6YNH:
https://identifiers.org/pdb:6YND
https://identifiers.org/pdb:6YNE
https://identifiers.org/pdb:6YNF
https://identifiers.org/pdb:6YNH

Mass spectrometry datasets at MassIVE, Accession number MSV000085325: https://identifiers.org/massive:MSV000085325

Zenodo: Crystal structure determination of Cys-S-sulfonated HsGAPDH from protein purified from the supernatant of HEK293F cells. https://doi.org/10.5281/zenodo.3817277\textsuperscript{25}

This project contains the following underlying data:
• Gel-Figure3.jpeg (original unedited gel image for Figure 2)

Extended data

Zenodo: Crystal structure determination of Cys-S-sulfonated HsGAPDH from protein purified from the supernatant of HEK293F cells. https://doi.org/10.5281/zenodo.3817277\textsuperscript{25}

This project contains the following extended data:
• Copy of the Open Laboratory Notebook (PDF)
• Original unedited gel images for Open Laboratory Notebook (JPG, JPEG and PNG files)

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgements

Siyu Wang and Louise Fairall helped with HEK293F cells expression. John W.R. Schwabe gave us access to the SEC-MALS instrument and Christopher J. Millard helped with the SEC-MALLS experiment. Trevor Greenhough coordinates the Midlands Diamond BAG. Neil Paterson helped with the handling of samples at beamline I03 of the Diamond Light Source in Harwell, England, UK. Tony Larson gave us access to resources at the York Centre of Excellence in Mass Spectrometry. Thanks are due to Damir Handzar and Matjaz Hren for their help with retrieval of original un-annotated SDS-PAGE gels from the SciNote backups.

References


