BIOINFORMATICS APPLICATIONS NOTE Vol. 28 no. 19 2012, pages 2551–2552
doi:10.1093/bioinformatics/bts468

Databases and ontologies

Advance Access publication July 25, 2012

RedoxDB—a curated database for experimentally verified protein oxidative modification

Ming-an Sun, Yejun Wang, Han Cheng, Qing Zhang, Wei Ge and Dianjing Guo*

School of Life Sciences, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China

Associate Editor: Alfonso Valencia

ABSTRACT

Summary: Redox regulation and signaling, which are involved in various cellular processes, have become one of the research focuses in the past decade. Cysteine thiol groups are particularly susceptible to post-translational modification, and their reversible oxidation is of critical role in redox regulation and signaling. With the tremendous improvement of techniques, hundreds of redox proteins along with their redox-sensitive cysteines have been reported, and the number is still fast growing. However, until now there is no database to accommodate the rapid accumulation of information on protein oxidative modification. Here we present RedoxDB—a manually curated database for experimentally validated redox proteins. RedoxDB (version 1.0) consists of two datasets (A and B, for proteins with or without verified modified cysteines, respectively) and includes 2157 redox proteins containing 2203 cysteine residues with oxidative modification. For each modified cysteine, the exact position, modification type and flanking sequence are provided. Additional information, including gene name, organism, sequence, literature references and links to UniProt and PDB, is also supplied. The database supports several functions including data search, blast and browsing. Bulk download of the entire dataset is also available. We expect that RedoxDB will be useful for both experimental studies and computational analyses of protein oxidative modification.

Availability: The database is freely available at: http://biocomputer.bio.cuhk.edu.hk/RedoxDB.

Contact: djguo@cuhk.edu.hk

Supplementary Information: Supplementary data are available at *Bioinformatics* Online.

Received on February 17, 2012; revised on July 8, 2012; accepted on July 23, 2012

1 INTRODUCTION

Oxidative stress represents an imbalance between the reactive oxygen species (ROS) and a biological system's ability to detoxify the reactive intermediates or to repair the resulting damage. Many diseases, including type II diabetes, cancer, neurodegenerative diseases and cardiovascular disease, are associated with oxidative stress (Sarsour *et al*., 2009). ROS has been previously regarded as unwanted by-products of aerobic metabolism. However, under normal condition, ROS can regulate the structure and function of proteins and act as important signaling molecule in various cellular processes (Imlay, 2003), and many key cellular processes are indicated to be redox-sensitive. With the rapid development of 'redox proteomics' that enables high-throughput detection of redox proteins, redox regulation and signaling have become the research focus in recent years (Chouchani *et al*., 2011; Forman *et al*., 2010).

Cysteine thiol groups are particularly susceptible to oxidation by ROS, reactive nitrogen species and other electrophilic molecules. These thiol groups can be reversibly oxidized to intra- and intermolecular disulfide bonds (S-S), sulfenic acid intermediate (SOH), sulfinic acid (SO_2H), S-nitrosothiols (S-NO) and S-glutathione (S-SG). Those oxidative products can be reduced back by cellular reductants such as thioredoxin (Trx) and glutathione (Grx) under certain conditions (Meyer *et al*., 2009; Reddie and Carroll, 2008). This reversible oxidation plays critical roles in redox regulation and signaling (Antelmann and Helmann, 2011; Forman *et al*., 2010), and represents another common type of protein post-translational modification apart from phosphorylation (Chiarugi and Buricchi, 2007).

Protein oxidative modification events are traditionally identified by case-by-case studies which usually accompanied by site-directed mutagenesis experiment. With the development of proteomics techniques in the past decade, it becomes possible to identify hundreds of redox-sensitive proteins in one single experiment (Chouchani *et al*., 2011; Lindahl *et al*., 2011; Weerapana *et al*., 2010). Currently, hundreds of redox proteins have been identified, and the number is still fast growing (Chouchani *et al*., 2011). For many of these redox proteins, the corresponding cysteines undergoing oxidative modification have also been experimentally determined. Meanwhile, several computational studies for cysteine oxidative modification began to emerge in recent years (Fomenko *et al*., 2007; Marino and Gladyshev, 2009; Sanchez *et al*., 2008). However, to our knowledge, until now there is no database to accommodate the rapid accumulation of information on protein oxidative modification.

Here, we describe RedoxDB, a manually curated database of experimentally verified protein oxidative modification. Annotations are provided at both sequence and residue levels. The database supports several functions including data search, blast, browsing and bulk download, which enables users to search and retrieve data easily.

DATABASE CONTENT

RedoxDB (version 1.0) consists of two datasets. Dataset A includes proteins that the modified cysteines have been determined and dataset B includes other redox proteins. All the data are collected from published literatures and are experimentally verified by various research groups. These redox proteins are

*To whom correspondence should be addressed.

Table 1. Summary of different types of cysteine oxidative modification lodged in RedoxDB (version 1.0)

Modification type	Dataset A		Dataset B
	No. of protein	No. of cysteine	No. of protein
Reversible disulfide	342	651	103
S-nitrosylation	511	674	211
S-glutathionylation	174	243	154
Sulphenation	107	102	59
Sulphination	35	38	0
Others or unknown	280	515	424
Total	1277	2203	880

derived from 123 different species including mammals, plants, bacteria, viruses and others.

Annotations at both sequence and residue levels (if available) are provided. RedoxDB provides annotations including the exact position, modification type and flanking sequence (10 amino acids) for each modified cysteine. Additional information, such as gene name, organism, sequence as well as literature references with the PubMed links, is supplied. Cross-reference to the UniProt protein sequence database is also available (Wu *et al*., 2006). Moreover, links to PDB database are also provided when the 3D structure data are available (Westbrook *et al*., 2002).

RedoxDB includes cysteines undergoing different types of oxidative modifications, such as reversible disulfide formation, sulphenation, sulphination, *S*-glutathionylation and *S*-nitrosylation. Cysteines that could form reversible disulfides or be *S*-nitrosylated constitute the major part of data, as summarized in Table 1.

DATABASE INTERFACE AND TOOLS

RedoxDB is implemented using Apache, MySQL5.0, PHP and PERL and runs under Ubuntu system. It supports several functions including data searching, browsing and retrieving. Bulk download of the entire dataset is also amenable.

RedoxDB can be searched using GENENAME, UNIPROT_ID, PDB_ID or ORGANISM as keywords, and the result can be filtered by modification type and experimental source. The user can also choose to include dataset B or not. The search output includes general annotations for all the matched entries (Fig. 1A). Another page shows the detailed annotation for each entry (Fig. 1B–D). RedoxDB can also be searched by BLAST function (Altschul *et al*., 1997) using query sequences. The output shows general information about blast result and provides links to the matched entries.

RedoxDB also provides an interface for data browsing. The number of redox proteins and modified cysteines are summarized for each species. Additional links are also provided for retrieving data by species.

The entire dataset in RedoxDB can be bulk downloaded in different format. We also welcome scientists in the redox research community to share their research result via the database.

DISCUSSION

RedoxDB is the first extensive database that provides information about protein oxidative modification verified by experimental studies. RedoxDB not only includes data from case-by-case studies but also integrates high-throughput data from proteomics studies. We expect that RedoxDB will be useful for both experimental studies and computational analyses of protein oxidative modification.

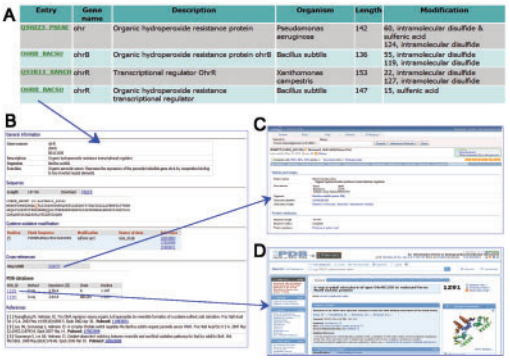

Fig. 1. A snapshot for searching RedoxDB using 'ohr' as query. (**A**) Tabular results for all the matched entries. (**B**) Main display page with detailed annotation and links to (**C**) UniprotKB and (**D**) PDB database.

Funding: The work is funded by grant HKBU1/CRF/10 and partially by a grant from the Institute of Plant Molecular Biology and Agricultural Biotechnology at CUHK.

Conflict of Interest: none declared.

REFERENCES

Altschul,S.F. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.

Antelmann,H. and Helmann,J.D. (2011) Thiol-based redox switches and gene regulation. *Antioxid. Redox Signal.*, **14**, 1049–1063.

Chiarugi,P. and Buricchi,F. (2007) Protein tyrosine phosphorylation and reversible oxidation: two cross-talking posttranslation modifications. *Antioxid. Redox Signal.*, **9**, 1–24.

Chouchani,E.T. *et al.* (2011) Proteomic approaches to the characterization of protein thiol modification. *Curr. Opin. Chem. Biol.*, **15**, 120–128.

Fomenko,D.E. *et al.* (2007) High-throughput identification of catalytic redox-active cysteine residues. *Science (New York, NY)*, **315**, 387–389.

Forman,H.J. *et al.* (2010) Signaling functions of reactive oxygen species. *Biochemistry*, **49**, 835–842.

Imlay,J.A. (2003) Pathways of oxidative damage. *Annu. Rev. Microbiol.*, **57**, 395–418.

Lindahl,M. *et al.* (2011) The disulfide proteome and other reactive cysteine proteomes: analysis and functional significance. *Antioxid. Redox Signal.*, **14**, 2581–2642.

Marino,S.M. and Gladyshev,V.N. (2009) A structure-based approach for detection of thiol oxidoreductases and their catalytic redox-active cysteine residues. *PLoS Comput. Biol.*, **5**, e1000383.

Meyer,Y. *et al.* (2009) Thioredoxins and glutaredoxins: unifying elements in redox biology. *Annu. Rev. Genet.*, **43**, 335–367.

Reddie,K.G. and Carroll,K.S. (2008) Expanding the functional diversity of proteins through cysteine oxidation. *Curr. Opin. Chem. Biol.*, **12**, 746–754.

Sanchez,R. *et al.* (2008) Prediction of reversibly oxidized protein cysteine thiols using protein structure properties. *Protein Sci.*, **17**, 473–481.

Sarsour,E.H. *et al.* (2009) Redox control of the cell cycle in health and disease. *Antioxid. Redox Signal.*, **11**, 2985–3011.

Weerapana,E. *et al.* (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature*, **468**, 790–795.

Westbrook,J. *et al.* (2002) The Protein Data Bank: unifying the archive. *Nucleic Acids Res.*, **30**, 245–248.

Wu,C.H. *et al.* (2006) The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.*, **34**, D187–D191.