

The SWISS-2DPAGE database: what has changed during the last year

Christine Hoogland*, Jean-Charles Sanchez¹, Luisa Tonella¹, Amos Bairoch², Denis F. Hochstrasser¹ and Ron D. Appel

Swiss Institute of Bioinformatics, c/o Medical Informatics Division, Geneva University Hospital, 24 rue Micheli-du-Crest, 1211 Geneva 14, Switzerland, ¹Central Clinical Chemistry Laboratory, Geneva University Hospital, Geneva, Switzerland and ²Department of Medical Biochemistry, University of Geneva, Geneva, Switzerland

Received October 1, 1998; Revised October 7, 1998; Accepted October 14, 1998

ABSTRACT

SWISS-2DPAGE (<http://www.expasy.ch/ch2d/>) is an annotated two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) database established in 1993. The current release contains 21 reference maps from human and mouse biological samples, as well as from *Saccharomyces cerevisiae*, *Escherichia coli* and *Dictyostelium discoideum* origin. These reference maps now have 2480 identified spots, corresponding to 528 separate protein entries in the database, in addition to virtual entries for each SWISS-PROT sequence. During the last year, the SWISS-2DPAGE has undergone major changes. Six new maps have been added, and new functions to access the data have been provided through the ExpASY server. Finally, an important change concerns the database funding source.

INTRODUCTION

The SWISS-2DPAGE database collects data on proteins identified on various two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) maps (1). Protein identification on 2-D PAGE maps is done by various techniques (2), including gel comparison, microsequencing, immunoblotting, amino acid composition analysis, peptide mass fingerprinting using mass spectrometry, or a combination of these techniques. The core of the SWISS-2DPAGE database consists of the description of the proteins identified, including mapping procedures, physiological and pathological information, experimental data (isoelectric point, molecular weight, amino acid composition, peptide mass) and bibliographical references, in addition to the 2-D PAGE images showing the protein locations. Cross-references are provided to Medline, to other 2-D PAGE databases [ECO2DBASE (3), HSC-2DPAGE (4), YEPD (5)], and to SWISS-PROT (6), which provides many links to other molecular databases [EMBL (7), GenBank (8), PROSITE (9), MIM (10), etc.]. The SWISS-2DPAGE database is maintained by the Swiss

Institute of Bioinformatics, in collaboration with the Central Clinical Chemistry Laboratory of the Geneva University Hospital.

FORMAT

The protein entries in SWISS-2DPAGE are text files structured to be readable by human as well as by computer programs. Each entry is composed of defined lines, used to record various kinds of data (Fig. 1). For standardization purposes, the format of SWISS-2DPAGE entries is similar to the SWISS-PROT database (6), in addition to specific lines dedicated to 2-D PAGE data: (i) the master line (MT) lists the reference maps where the entry has been identified; (ii) the image line (IM) lists the 2-D PAGE images available for the entry; (iii) the 2D lines group different topics including the mapping procedure, spot coordinates, protein amino acid composition, peptide masses from mass spectrometry, protein expression levels and modifications. On the ExpASY Web server (see availability section) the data image associated with a protein entry displays the experimental location of the protein on the chosen map, in addition to a theoretical region computed from the protein sequence (Fig. 2).

CURRENT CONTENT

Release 8 of SWISS-2DPAGE (October 1998) contains 21 reference maps from human: cells or tissues (kidney, liver, lymphoma, platelet cells, red blood cells, colorectal epithelia cells), body fluids (cerebrospinal fluid, plasma), culture cells (erythroleukemia cell line, hepatoblastoma carcinoma derived cells, hepatoblastoma carcinoma derived cell line secreted proteins, macrophage like cell line, promyelocytic leukemia derived cells), from mouse cells or tissues (liver, gastrocnemius muscle, Pancreatic islet cells, epididymal fat pad), and from *Saccharomyces cerevisiae*, *Escherichia coli* (for two pI ranges: one from 3.5 to 10 and one from 4.5 to 5.5) and *Dictyostelium discoideum* origin. Table 1 gives detailed descriptions for each of these maps, including creation and release dates, number of detected spots, number of identified spots, and number of distinct protein entries. There is a total of 528 protein entries. In addition,

*To whom correspondence should be addressed. Tel: +41 22 372 62 81; Fax: +41 22 372 61 98; Email: christine.hoogland@isb-sib.ch

```

ID  ARGT_ECOLI; STANDARD; 2DG.
AC  P09551;
DT  01-OCT-1998 (REL. 08, CREATED)
DT  01-OCT-1998 (REL. 08, LAST UPDATE)
DE  LYSINE-ARGININE-ORNITHINE-BINDING PERIPLASMIC PROTEIN
DE  (LAO-BINDING PROTEIN).
OS  ESCHERICHIA COLI, AND SALMONELLA TYPHIMURIUM.
OC  PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY
OC  ANAEROBIC RODS; ENTEROBACTERIACEAE.
MT  ECOLI4.5-5.5.
IM  ECOLI4.5-5.5.
RN  [1]
RP  MAPPING ON GEL
RA  TONELLA L., WALSH B., SANCHEZ J.-C., OU K., WILKINS M.,
RA  TYLER M., FRUTIGER S., GOOLEY A., APPEL R.D., BAIROCH A.,
RA  HOOGLAND C., HUGUES G., WILLIAMS K., HOCHSTRASSER D.;
RL  SUBMITTED (OCT-1998) TO THE SWISS-2DPAGE DATABASE.
2D  -!- MASTER: ECOLI4.5-5.5;
2D  -!- PI/MW: SPOT 2D-0015AB=5.17/26881;
2D  -!- MAPPING: MASS FINGERPRINTING [1].
2D  -!- PEPTIDE MASSES: 1140.99; 1776.18; 2196.32; TRYPSIN.
CC -----
CC This SWISS-2DPAGE entry is copyright the Swiss Institute of
CC Bioinformatics. There are no restrictions on its use by
CC non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for
CC commercial entities requires a license agreement (See
CC http://www.isb-sib.ch/announce/ or send an email to
CC license@isb-sib.ch).
CC -----
DR  SWISS-PROT; P09551; ARGT_ECOLI.
//

```

Figure 1. Example of a protein entry from SWISS-2DPAGE.

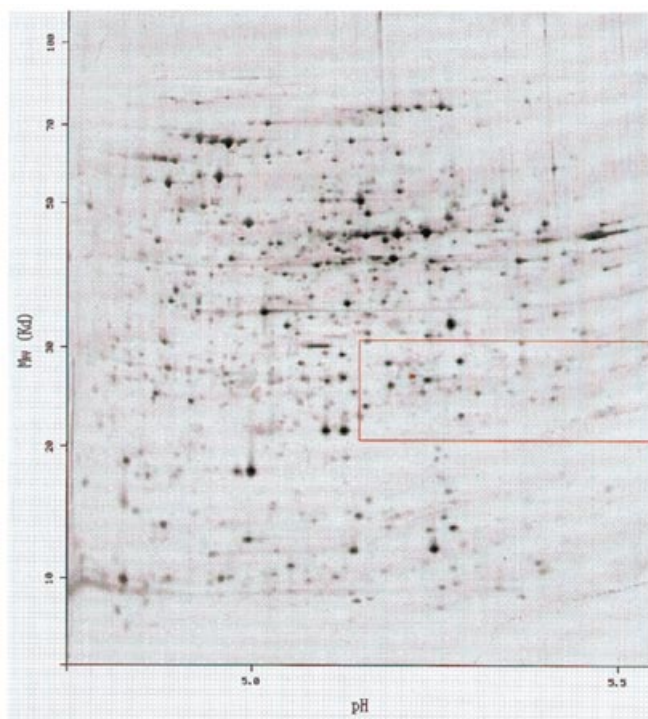


Figure 2. The protein from Figure 1 shown on the ECOLI4.5-5.5 reference map.

there are as many virtual entries as protein sequences in the SWISS-PROT database.

Table 1. Content of SWISS-2DPAGE

Map	Creation	Last modification	Detected spots	Identified spots	Number of entries
CEC	JAN-98	JAN-98	3077	56	40
CSF	AUG-93	JUN-97	1664	309	30
ELC	AUG-93	JUN-97	2144	35	19
HEPG2	AUG-93	SEP-97	2862	98	45
HEPG2SP	AUG-93	JAN-96	1734	155	25
HL60	SEP-97	SEP-97	3164	26	17
HUMAN KIDNEY	SEP-97	SEP-97	2896	42	27
LIVER	AUG-93	SEP-97	2413	138	68
LYMPHOMA	AUG-93	JUN-97	1890	60	33
PLASMA	AUG-93	SEP-97	1966	626	70
PLATELET	AUG-93	JUN-97	2193	41	18
RBC	AUG-93	SEP-97	1800	190	33
U937	AUG-93	JUN-97	895	42	32
DICTYSLUG	SEP-97	SEP-97	3164	25	16
ECOLI	AUG-95	JAN-98	2364	206	180
ECOLI4.5-5.5	OCT-98	OCT-98	1025	98	68
YEAST	FEB-95	JUN-97	1940	101	62
EPF	SEP-98	SEP-98	3540	87	52
MOUSE ISLETS	SEP-98	SEP-98	2528	7	2
LIVER	SEP-98	SEP-98	2822	139	73
MUSCLE	SEP-98	SEP-98	2638	61	36

Abbreviations as follows: CEC, colorectal epithelia cells; CSF, cerebrospinal fluid; ELC, erythroleukemia cell line; HEPG2, hepatoblastoma carcinoma derived cells; HEPG2SP, hepatoblastoma carcinoma derived cell line secreted proteins; HL60, promyelocytic leukemia derived cells; RBC, red blood cells; U937, macrophage like cell line; DICTYSLUG, *D.discoideum*; ECOLI, *E.coli*; ECOLI4.5-5.5, *E.coli* with pI range from 4.5 to 5.5; YEAST, *S.cerevisiae*; EPF, epididymal fat pad; ISLETS, pancreatic islet cells; MUSCLE, gastrocnemius muscle.

AVAILABILITY

The most efficient and user-friendly way to interactively browse in SWISS-2DPAGE is through the ExPASy World Wide Web molecular biology server at URL: <http://www.expasy.ch/>. The SWISS-2DPAGE home page at URL: <http://www.expasy.ch/ch2d/> provides several textual and graphical queries, and displays results with active links to other databases. A SWISS-2DPAGE entry may be reachable by keyword search (protein name, accession number, description or authors) or by selecting a spot on one of the 2-D PAGE maps. A full text search is now also available, allowing boolean operators (and, or, not) to restrict queries. Two new features have also been added to facilitate visualisation and differentiation of spots: (i) when clicking on a spot in one of the 2-D PAGE maps, the 2D line describing this spot in the corresponding protein entry will be highlighted in green; (ii) when displaying a protein entry, links are provided from each described spot to the corresponding 2-D PAGE map, in which the spot will be highlighted in green (in contrast to the other spots for that protein displayed in red).

It is also possible to get a local copy of SWISS-2DPAGE using anonymous ftp (file transfer protocol) from the ExPASy FTP server (<ftp://ftp.expasy.ch/databases/swiss-2dpage/>).

Since September 1998, a system of an annual subscription fee for commercial users of the SWISS-2DPAGE database has been implemented. Increased data flow and especially automation of proteome analysis have created a requirement for resources, which cannot be addressed in full by public funding. It has therefore been decided to adopt the solution chosen for the SWISS-PROT database (6), namely to ask non-academic users to financially contribute to the maintenance of the database (for details, see URL <http://www.expasy.ch/announce/>).

ACKNOWLEDGEMENTS

The authors thank S. Frutiger, G. Hughes, M. Wilkins, A. Gooley, K. Williams, B. Walsh, K. Ou and P.-A. Binz for their assistance in protein identification, M. Raymond for providing colorectal data, J. Yan for Dicty data, C. Sarto for kidney data, R. Joubert-Carron for HL60 data, C. Pasquali for *E.coli* data, M.

Cawthorne for mouse data, and F. Ravier, S. Jaccoud, I. Demalte and V. Rouge for their technical assistance.

REFERENCES

- 1 Appel,R.D, Sanchez,J-C., Bairoch,A., Golaz,O., Miu,M., Vargas,R. and Hochstrasser,D. (1993) *Electrophoresis*, **14**, 1232–1238.
- 2 Wilkins,M.R. and Gooley,A.A. (1997) In Wilkins,M.R., Williams,K.L., Appel,R.D. and Hochstrasser,D.F. (Eds), *Proteome Research: New Frontiers in Functional Genomics*. Springer Verlag, Berlin, pp. 35–64.
- 3 VanBogelen,R.A., Abshire,K.Z., Moldover,B., Olson,E.R. and Neidhardt,F.C. (1997) *Electrophoresis*, **18**, 1243–1251.
- 4 Corbett,J.M., Wheeler,C.H., Baker,C.S., Yacoub,M.H. and Dunn,M.J. (1994) *Electrophoresis*, **15**, 1459–1465.
- 5 Payne,W.E. and Garrels,J.I. (1997) *Nucleic Acids Res.*, **25**, 57–62.
- 6 Bairoch,A. and Apweiler,R. (1998) *Nucleic Acids Res.*, **26**, 38–42.
- 7 Stoesser,G., Moseley,M.A., Sleep,J., McGowran,M., Garcia-Pastor,M. and Sterk,P. (1998) *Nucleic Acids Res.*, **26**, 8–15.
- 8 Benson,D.A., Boguski,M.S., Lipman,D.J., Ostell,J. and Ouellette B.F.F. (1998) *Nucleic Acids Res.*, **26**, 1–7.
- 9 Bairoch,A., Bucher,P. and Hofmann,K. (1997) *Nucleic Acids Res.*, **25**, 217–221.
- 10 Pearson,P., Francomano,C., Foster,P., Bocchini,C., Li,P. and McKusick,V.A. (1994) *Nucleic Acids Res.*, **22**, 3470–3473.