332–333 *Nucleic Acids Research, 1998, Vol. 26, No. 1*

Current status of the SWISS-2DPAGE database

Christine Hoogland*, Jean-Charles Sanchez[1], Luisa Tonella[1], Amos Bairoch[2], Denis F. Hochstrasser[1] and Ron D. Appel

Laboratoire d'Imagerie Moléculaire et Bioinformatique, Division d'Informatique Médicale, Hôpital Cantonal Universitaire, 24, rue Micheli-du-Crest, CH-1211 Genève 14, Switzerland, [1]Central Clinical Chemistry Laboratory, Geneva University Hospital, Geneva, Switzerland and [2]Department of Medical Biochemistry, University of Geneva, Geneva, Switzerland

Received September 26, 1997; Accepted October 3, 1997

ABSTRACT

The SWISS-2DPAGE database (http://www.expasy.ch/ch2d/ch2d-top.html) consists of two-dimensional polyacrylamide gel electrophoresis images, as well as textual descriptions of the proteins that have been identified on them. The current release contains 15 reference maps from human biological samples, as well as from *Saccharomyces cerevisiae*, *Escherichia coli* and *Dictyostelium discoideum* origin. These reference maps have 2088 identified spots, corresponding to 410 separate protein entries in the database, in addition to virtual entries for each SWISS-PROT sequence.

INTRODUCTION

SWISS-2DPAGE is a database containing experimental data available from two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) (1). Identification of proteins on 2-D PAGE maps is actually obtained by microsequencing, immunoblotting, amino acid composition analysis, gel comparison, and peptide mass fingerprinting using mass spectrometry or by combination of these technics (2). The SWISS-2DPAGE database thus assembles textual description for the proteins identified, including mapping procedures, physiological and pathological informations, experimental data (isoelectric point, molecular weight, amino acid composition) and bibliographical references, in addition to the various 2-D PAGE images showing the protein locations. Cross-references are provided to Medline, to other 2-D PAGE databases (ECO2DBASE, HSC-2DPAGE, YEPD) and to SWISS-PROT, which provides many links to other molecular databases (EMBL, GenBank, PROSITE, OMIM, etc.).

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 (3), in addition to specific lines dedicated to the 2-D PAGE data:

```
ID   P17_DICDI; STANDARD; 2DG.
AC   P34121;
DT   01-SEP-1997 (REL. 06, CREATED)
DT   01-SEP-1997 (REL. 06, LAST UPDATE)
DE   CYTOSKELETAL P17 PROTEIN (COACTOSIN) (CYCLIC AMP-REGULATED
DE   PROTEIN P16).
OS   DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
OC   EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; SARCODINA;
OC   RHIZOPODA; EUMYCETOZOA; DICTYOSTELIA.
MT   DICTYSLUG.
IM   DICTYSLUG.
RN   [1]
RP   MAPPING ON GEL
RX   MEDLINE; 97295287.
RA   YAN J.X., TONELLA L., SANCHEZ J.-C., WILKINS M.R., PACKER
RA   N.H., GOOLEY A.A., HOCHSTRASSER D.F., WILLIAMS K.L;
RL   ELECTROPHORESIS 18:491-497(1997).
2D   -!- MASTER: DICTYSLUG;
2D   -!-   PI/MW: SPOT 2D-000SEW=5.07/14887;
2D   -!-   AMINO ACID COMPOSITION: SPOT 2D-000SEW: B=11.30,
2D         S=10.80, Z=12.10, G=11.20, T=3.90, H=0.90, Y=4.60,
2D         A=9.40, P=2.30, R=3.40,M=1.40, V=9.20, I=3.60,
2D         L=6.30, F=2.30, K=7.30 ;
2D   -!-   MAPPING: AMINO ACID COMPOSITION ANALYSIS [1].
DR   SWISS-PROT; P34121; P17_DICDI
//
```

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

(i) the master line (MT) lists the reference maps where the entry has been identified; (ii) the images line (IM) lists the 2-D PAGE images available for the entry; (iii) the 2D lines group different topics including the mapping procedure, spot co-ordinates, protein amino acid composition, 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 6 of SWISS-2DPAGE (September 1997) contains 15 reference maps from human: cells or tissues (kidney, liver, lymphoma, platelet cells, red blood 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

* To whom correspondence should be addressed. Tel: +41 22 372 62 80; Fax: +41 22 372 61 98; Email: christine.hoogland@dim.hcuge.ch

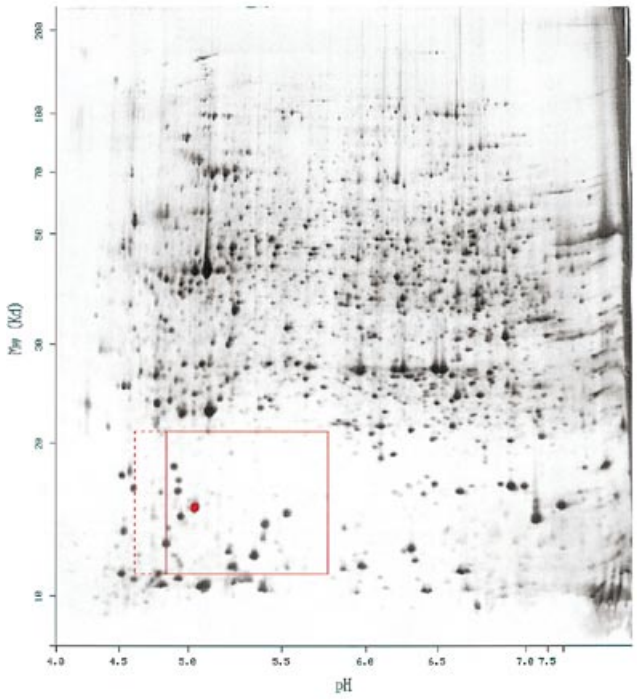

Figure 2. The protein from Figure 1 shown on the DICTYSLUG reference map.

line, promyelocytic leukemia derived cells), and from *Saccharomyces cerevisiae*, *Escherichia coli* and *Dictyostelium discoideum* origin. Table 1 gives detailed description 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 410 protein entries. In addition, there are as many virtual entries as protein sequences in the SWISS-PROT database.

User manual and release notes are also included in this release, as well as experimental protocols for 2-D PAGE and post-separation analysis (including photographs of each procedure).

AVAILABILITY

The most efficient and user-friendly way to interactively browse in SWISS-2DPAGE is through the ExPASy World Wide Web molecular biology server (URL is http://www.expasy.ch/). The SWISS-2DPAGE top page (URL: http://www.expasy.ch/ch2d/ch2d-top.html) provides several text searches, and displays results with active links to other databases. 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/).

Table 1. Content of SWISS-2DPAGE

Map	Creation	Last modification	Detected spots	Identified spots	No. of entries
CSF	AUG-93	JUN-97	1664	309	30
DICTYSLUG	SEP-97	SEP-97	3164	25	16
ELC	AUG-93	JUN-97	2144	35	19
ECOLI	AUG-95	SEP-97	2364	200	178
HEPG2	AUG-93	SEP-97	2862	98	45
HEPG2SP	AUG-93	JAN-96	1734	155	25
HL60	SEP-97	SEP-97	3164	26	17
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
YEAST	FEB-95	JUN-97	1940	101	62

Abbreviations as follows: CSF, cerebrospinal fluid; DICTYSLUG, *Dictyostelium discoideum*; ELC, erythroleukemia cell line; ECOLI, *Escherichia coli*; 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; YEAST, *Saccharomyces cerevisiae*.

ACKNOWLEDGEMENTS

The authors thank S. Frutiger, G. Hughes, M. Wilkins, N. Packer, A. Gooley, K. Williams, B. Walsh, M. Molloy, K. Ou and B. Herbert for their assistance in protein identification, J. Yan for providing Dicty data, C. Sarto for Kidney data, J. Carron for HL60 data, C. Pasquali for *E.coli* data and F. Ravier, S. Jaccoud, I. Demalte and V. Rouge for their technical assistance.

REFERENCES

1 Herbert,B.R., Sanchez,J.-C. and Bini,L. (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, Heidelberg, New York, pp. 13–33.
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, Heidelberg, New York, pp. 35–64.
3 Bairoch,A. and Apweiler,R. (1998) *Nucleic Acids Res.*, **26**, 38–42.