



(86) Date de dépôt PCT/PCT Filing Date: 2008/02/22
(87) Date publication PCT/PCT Publication Date: 2008/08/28
(85) Entrée phase nationale/National Entry: 2009/07/27
(86) N° demande PCT/PCT Application No.: US 2008/054805
(87) N° publication PCT/PCT Publication No.: 2008/103962
(30) Priorités/Priorities: 2007/02/22 (US60/891,196);
2007/11/13 (US60/987,752); 2008/01/28 (US61/024,170)

(51) Cl.Int./Int.Cl. *C12Q 1/68* (2006.01)
(71) Demandeur/Applicant:
GENENTECH, INC., US
(72) Inventeurs/Inventors:
DIEHL, LAURI, US;
FLANAGAN, KENNETH, US;
MO, LIAN, US
(74) Agent: SMART & BIGGAR

(54) Titre : METHODE DE DETECTION D'UNE MALADIE INTESTINALE INFLAMMATOIRE
(54) Title: METHODS FOR DETECTING INFLAMMATORY BOWEL DISEASE

(57) **Abrégé/Abstract:**

The present invention provides for a method of detecting the presence of inflammatory bowel disease in gastrointestinal tissues or cells of a mammal by detecting increased expression of LY6 genes in the tissues or cells of the mammal relative to a control.



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
28 August 2008 (28.08.2008)

PCT

(10) International Publication Number
WO 2008/103962 A3

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2008/054805

(22) International Filing Date:
22 February 2008 (22.02.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/891,196 22 February 2007 (22.02.2007) US
60/987,752 13 November 2007 (13.11.2007) US
61/024,170 28 January 2008 (28.01.2008) US

(71) Applicant (*for all designated States except US*): **GENEN-TECH, INC.** [US/US]; 1 Dna Way, South San Francisco, CA 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **DIEHL, Lauri** [US/US]; 1460 Hollidale Court, Los Altos, CA 94024 (US). **FLANAGAN, Kenneth** [US/US]; 1972 11th Ave, San Francisco, CA 94116 (US). **MO, Lian** [US/US]; 420 Lian, 420 James Rd #9, Palo Alto, CA 94306 (US).

(74) Agents: **DREGER, Ginger, R.** et al.; Goodwin Procter LLP, 135 Commonwealth Drive, Menlo Park, CA 94025 (US).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

(88) Date of publication of the international search report:
18 December 2008

(54) Title: METHODS FOR DETECTING INFLAMMATORY BOWEL DISEASE

(57) Abstract: The present invention provides for a method of detecting the presence of inflammatory bowel disease in gastrointestinal tissues or cells of a mammal by detecting increased expression of LY6 genes in the tissues or cells of the mammal relative to a control.



WO 2008/103962 A3

METHODS FOR DETECTING INFLAMMATORY BOWEL DISEASE**FIELD OF THE INVENTION**

[0001] The present invention relates to gene expression profiles in inflammatory bowel disease pathogenesis. This discovery finds use in the detection and diagnosis of inflammatory bowel disease.

BACKGROUND OF THE INVENTION

[0002] Inflammatory bowel disease (IBD), a chronic inflammatory disorder of the gastrointestinal tract suffered by approximately one million patients in the United States, is made up of two major disease groups: ulcerative colitis (UC) and Crohn's Disease (CD). In both forms of IBD, intestinal microbes may initiate the disease in genetically susceptible individuals. UC is often restricted to the colon, while CD typically occurs in the ileum of the small intestine and in the colon. (Podolsky, D.K., N. Engl. J. Med. 347:417-429 (2002). Gene expression profiling of tissue from IBD patients has provided some insight into possible targets for therapy and/or diagnosis (see, for example, Dieckgraefe, B.K. et al., Physiol. Genomics 4:1-11 (2000); Lawrance I.C. et al., Hum Mol Genet. 10:445-456 (2001); Dooley T.P. et al., Inflamm. Bowel Dis. 10:1-14 (2004); and Uthoff S.M., Int J Oncol. 19:803-810 (2001)). Further investigations of gene dysregulation in patients experiencing inflammatory bowel disease include, for example, Lawrance, I.C. et al., who disclosed distinctive gene expression profiles for several genes in UC and CD (Lawrance, I.C. et al., Human Mol. Genetics 10(5):445-456 (2001)). Uthoff, S.M.S. et al. disclosed the identification of candidate genes for UC and CD using micro array analysis (Uthoff, S.M.S. et al., Int'l. J. Oncology 19:803-810 (2001)). Dooley, T.P. et al. disclosed correlation of gene expression in IBD with drug treatment for the disorder (Dooley, T.P. et al., Inflamm. Bowel Dis. 10(1):1-14 (2004)).

[0003] There is a need for the identification of additional biological markers of inflammatory bowel disease for use in diagnosis of this chronic disease. The present disclosure fills that need.

[0004] The entire contents of all references cited herein are hereby incorporated by reference.

SUMMARY OF THE INVENTION

[0005] Disclosed herein is the unique finding that members of the LY6 superfamily of genes are upregulated on the surface of intestinal epithelial cells (IEC) in models of murine colitis and in intestinal tissue of human patients experiencing IBD, which genes are not

expressed on healthy IEC. The majority of LY6 family members are GPI-anchored cell surface glycoproteins with broad distribution on cells of hematopoietic origin, and more limited expression on non-hematopoietic cells. Though widely used as markers of differentiation of immune cells (Sunderkotter, C. et al., *J. Immunol.* 172:4410-4417 (2004)), the functions that the LY6 family possesses have been difficult to elucidate (Shevach, E.M. and P.E. Korthy, *Immunol. Today* 10:195-200 (1989)). Reports have shown that LY6 molecules are involved in a diverse array of functions including T cell activation (Zhang, Z.X. et al., *Eur. J. Immunol.* 32:1584-1592 (2002) and Henderson, S.C. et al., *J. Immunol.* 168:118-126 (2002), olfaction (Chou, J.H. et al., *Genetics* 157:211-224 (2001) and cellular adhesion (Jaakkola, I. et al., *J. Immunol.* 170:1283-1290 (2003)).

[0006] In the broadest sense, the invention provides for a method of detecting increased expression of genes of the human LY6 gene family in intestinal tissue in intestinal tissue from a first mammal experiencing an intestinal disorder relative to a control mammal. In a more directed sense, the method is expected to be applicable to the diagnosis of disorders related to intestinal disorders associated with human LY6H, LYPD1, LYPD3, and LYPD5 expression, which disorders include without limitation inflammatory bowel disease (IBD), such as ulcerative colitis (UC) and Crohn's Disease (CD). In one embodiment, the method of the invention is useful to detect responders and nonresponders of IBD therapeutic treatment. In one embodiment, the IBD is ulcerative colitis (UC). In one embodiment, the IBD is Crohn's Disease (CD). In one embodiment, the intestinal tissue is colon tissue. In one embodiment, the colon tissue is sigmoid colon. In one embodiment, LY6H, LYPD1, LYPD3 and/or LYPD5 gene expression is increased in intestinal tissue (such as colon tissue) in an IBD, UC or CD mammal relative to normal intestinal (such as normal colon tissue) of a mammal not experiencing IBD, CD or UC. In an embodiment, the LY6H gene comprises the nucleic acid of SEQ ID NO:1 and encodes the LY6H polypeptide comprising SEQ ID NO:2. In an embodiment, the LYPD1 gene comprises the nucleic acid of SEQ ID NOS:3 or 4 and encodes the LYPD1 polypeptide comprising SEQ ID NO:5. In an embodiment, the LYPD3 gene comprises the nucleic acid of SEQ ID NO:6 and encodes the LYPD3 polypeptide comprising SEQ ID NO:7. In an embodiment, the LYPD5 gene comprises the nucleic acid of SEQ ID NOS:8 or 9 and encodes the LYPD5 polypeptide comprising SEQ ID NO:10.

[0007] In one embodiment, the method of the invention comprises obtaining a tissue sample from a test mammal suspected of experiencing an intestinal disorder, contacting the tissue with a detectable agent that interacts with LY6H, LYPD1, LYPD3 and/or LYPD5

protein or with nucleic acid encoding LY6H, LYPD1, LYPD3 and/or LYPD5 and determining the level of LY6H, LYPD1, LYPD3 and/or LYPD5 expression relative to a control tissue. In one embodiment increased expression of LY6H, LYPD1, LYPD3 and/or LYPD5 relative to control is indicative of IBD in the test mammal. In one embodiment, increased expression of LY6H, LYPD1, LYPD3 and/or LYPD5 in test intestinal tissue relative to control intestinal tissue is indicative of UC in the test mammal. In one embodiment, increased expression of LY6H, LYPD1, LYPD3 and/or LYPD5 in test intestinal tissue relative to control intestinal tissue is indicative of CD in the test mammal. In one embodiment the tissue or cells from the test and control mammal are from the colon.

[0008] In one embodiment, LY6H, LYPD1, LYPD3 and/or LYPD5 expression is determined by detection of gene expression, such as by detection of mRNA encoding LY6H, LYPD1, LYPD3 and/or LYPD5 in a tissue sample or cells. In an embodiment, a control sample is a sample of tissue or cells of the same tissue or cell type obtained from a mammal known not to be experiencing a gastrointestinal disorder, such as IBD, UC or CD. In an embodiment, a control sample is a universal standard comprising RNA from several normal tissues or from multiple cell lines. In microarray analysis, such universal standards are useful for monitoring and controlling intra- and inter-experimental variation. In one embodiment, a universal standard (or Universal Reference RNA (URR)) is prepared as provided in Novoradovskaya, N. et al., (2004) BMC Genomics 5:20, which reference is hereby incorporated by reference in its entirety. In one embodiment, for use as a control in microarray analysis of mouse RNA, the URR is a Universal Mouse Reference RNA from Stratagene® (catalog #740100, Stratagene®, La Jolla, CA). In one embodiment, for use as a control in microarray analysis of human RNA, the URR is a Universal Human Reference RNA from Stratagene® (catalog #740000). In one embodiment, for use as a control in microarray analysis of rat RNA, the URR is a Universal Rat Reference RNA from Stratagene® (catalog #740200). In one embodiment, where the RNA is mouse RNA, the cell lines from which total RNA is extracted comprise cell lines derived from embryo, embryo fibroblast, kidney, liver hepatocyte, lung alveolar macrophage, B-lymphocyte, T-lymphocyte (thymus), mammary gland, muscle myoblast, skin, and testis. In one embodiment, where the RNA is human RNA, the cell lines from which total RNA is extracted comprise cell lines derived from mammary gland adenocarcinoma, liver hepatoblastoma, cervix adenocarcinoma, embryonal carcinoma or testis, brain glioblastoma, melanoma, liposarcoma, histiocytic lymphoma (macrophage, histocyte), T lymphoblast lymphoblastic leukemia, B lymphocyte plasmacytoma melanoma.

In one embodiment where the RNA is rat RNA, the cell lines from which total RNA is extracted comprise cell lines derived from blood basophilic leukemia, blood T-lymphocyte lymphoma, blood B-lymphoblast hybridoma, brain glioma, embryo yolk sac carcinoma, embryo normal fibroblast, normal kidney, liver hepatoma, lung normal alveolar macrophage, lung normal alveolar type II, mammary gland adenocarcinoma, muscle myoblast, normal skin, and testis leydig cell tumor.

[0010] In one aspect, the invention concerns an article of manufacture comprising a container and a composition of matter contained within the container, wherein the composition of matter comprises a nucleic acid encoding LY6H, LYPD1, LYPD3 and/or LYPD5 or their complements, and/or an anti- LY6H, LYPD1, LYPD3 and/or LYPD5 antibody or antibodies, or anti- LY6H-, LYPD1-, LYPD3- and/or LYPD5-binding fragment thereof, wherein the nucleic acids and/or antibodies are detectable. In one embodiment, the composition of matter comprises detecting agents for detecting nucleic acid binding, such as without limitation LY6H-, LYPD1-, LYPD3- and/or LYPD5-encoding nucleic acids or their complements, to LY6H, LYPD1, LYPD3 and/or LYPD5 nucleic acid in a tissue sample of a test mammal suspected of experiencing an intestinal disorder. In one embodiment, the compositions of matter comprises detecting agents for detecting antibody binding to, for example, LY6H, LYPD1, LYPD3 and/or LYPD5 in a tissue sample of a test mammal suspected of experiencing an intestinal disorder. In one embodiment, the antibody of the composition is detectably labeled. In one embodiment, the antibody of the composition is detectable by a second antibody, which second antibody is detectable or detectably labeled. The article may further optionally comprise a label affixed to the container, or a package insert included with the container, that refers to the use the LY6H, LYPD1, LYPD3 and/or LYPD5 nucleic acid or its complement and/or the anti-LY6H, anti-LYPD1, anti-LYPD3 and/or anti-LYPD5 antibody or LY6H, LYPD1, LYPD3 and/or LYPD5 binding fragment thereof in the detection of increased expression of LY6H, LYPD1, LYPD3 and/or LYPD5 in intestinal tissue, including without limitation, colon tissue. In an embodiment, the intestinal disorder is IBD. In an embodiment the intestinal disorder is UC or CD. In an embodiment the LYPD1 polypeptide and the anti-LYPD1 antibody is an antibody as disclosed in US7,157,558 and US7,144,990, respectively.

[0011] In one aspect, the present invention concerns a method of diagnosing the presence of an intestinal disorder in a mammal, comprising detecting the level of expression of a gene encoding LY6H, LYPD1, LYPD3 and/or LYPD5 polypeptide (a) in a test sample of tissue or cells obtained from said mammal, and (b) in a control sample of known normal cells

from a mammal not experiencing an intestinal disorder of the same tissue origin or type, wherein a higher level of expression of the LY6H, LYPD1, LYPD3 and/or LYPD5 polypeptide in the test sample, as compared to the control sample, is indicative of the presence of an intestinal disorder in the mammal from which the test sample was obtained. In an embodiment, the intestinal disorder is IBD. In an embodiment, the IBD is UC. In an embodiment, the IBD is CD. In an embodiment, the detecting is by contacting an antibody to LY6H, LYPD1, LYPD3 and/or LYPD5 polypeptide, or binding fragment of the antibody, with the test and control samples and determining the relative amount of antibody-polypeptide complex formation. A higher level of antibody-polypeptide complex formation in the test sample relative to the control sample is indicative of intestinal disorder, such as IBD, UC or CD, in the test mammal. The antibody of the invention is detectably labeled or, alternatively, the antibody is detected by subsequent binding of a second antibody which is detectable.

[0012] In yet a further embodiment, the present invention concerns a method of diagnosing the presence of an intestinal disorder in a mammal, comprising (a) contacting a test sample comprising tissue or cells obtained from the test mammal with an oligonucleotide that hybridizes at high stringency to LY6H, LYPD1, LYPD3 and/or LYPD5 nucleic acid (or its complement) or an antibody that binds specifically to LY6H, LYPD1, LYPD3 and/or LYPD5 polypeptide and (b) detecting the formation of a complex between the oligonucleotide or antibody and the LY6H, LYPD1, LYPD3 and/or LYPD5 nucleic acid (or its complement) or LY6H, LYPD1, LYPD3 and/or LYPD5 polypeptide, respectively, in the test sample, wherein the formation of more of such complex in the test sample relative to a control sample is indicative of the presence of an intestinal disorder (such as IBD, UC or CD) in the test mammal. In one embodiment, the intestinal disorder is IBD. In one embodiment, the disorder is UC. In one embodiment the disorder is CD. In one embodiment the tissue of the test and control mammals is colon tissue. Optionally, the LY6H, LYPD1, LYPD3 and/or LYPD5 polypeptide binding antibody or LY6H, LYPD1, LYPD3 and/or LYPD5 gene hybridizing oligonucleotide employed by the method of the invention is detectable, detectably labeled, attached to a solid support, or the like, and/or the test sample of tissue or cells is obtained from an individual suspected of experiencing an intestinal disorder, wherein the disorder is IBD, such as without limitation, UC or CD.

[0013] In yet a further embodiment, the present invention concerns the use of (a) a LY6H, LYPD1, LYPD3 and/or LYPD5 polypeptide, (b) a nucleic acid encoding a LY6H, LYPD1, LYPD3 and/or LYPD5 polypeptide or a vector or host cell comprising the nucleic

acid of (a), (c) an anti-LY6H, LYPD1, LYPD3 and/or LYPD5 polypeptide antibody, or (d) a LY6H, LYPD1, LYPD3 and/or LYPD5-binding oligopeptide, in the preparation of a medicament useful for the diagnostic detection of an intestinal disorder, including without limitation, IBD CD or UC, in an intestinal tissue of a mammal, including without limitation colon tissue.

[0014] In one aspect, the invention comprises a method of detecting a therapeutic drug response in a mammal treated with an IBD therapeutic agent, wherein the method comprises determining LY6H, LYPD1, LYPD3 and/or LYPD5 expression in gastrointestinal tissue of a test mammal relative to a control gastrointestinal tissue of a control mammal, where a higher level of expression of LY6H, LYPD1, LYPD3 and/or LYPD5 in a test tissue relative to a control tissue indicates a disease state or continuation of the disease state. A difference in LY6H, LYPD1, LYPD3 and/or LYPD5 expression in the test tissue that is not significantly higher than normal control expression levels or are within a range of normal expression levels for LY6H, LYPD1, LYPD3 and/or LYPD5 in a population of mammals indicates improvement or resolution of the intestinal disorder, which improvement or resolution may be attributed to the therapeutic agent. In one embodiment, a therapeutic response is determined when the levels of expression of LY6H, LYPD1, LYPD3 and/or LYPD5 in gastrointestinal or colon tissues or cells of the mammal treated with a therapeutic agent are different (expression is more similar to normal control, i.e., LY6H, LYPD1, LYPD3 and/or LYPD5 levels are lower than LY6H, LYPD1, LYPD3 and/or LYPD5 expression levels were in the mammal prior to treatment).

[0015] Yet further embodiments of the present invention will be evident to the skilled artisan upon a reading of the present specification.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] **Figures 1A and 1B** depict the nucleic acid sequence (SEQ ID NO:1) encoding human LY6H polypeptide and the amino acid sequence of human LY6H polypeptide (SEQ ID NO:2).

[0017] **Figures 2A and 2B** depict nucleic acid sequences (SEQ ID NOS:3 and 4) encoding the human LYPD1 polypeptide and the amino acid sequence of human LYPD1 polypeptide shown in **Figure 2C** (SEQ ID NO:5).

[0018] **Figures 3A and 3B** depict the nucleic acid sequence (SEQ ID NO:6) encoding human LYPD3 polypeptide and the amino acid sequence of human LYPD3 polypeptide (SEQ ID NO:7).

[0019] **Figures 4A and 4B** depict nucleic acid sequences (SEQ ID NOS:8 and 9) encoding human LYPD5 polypeptide and the amino acid sequence of human LYPD5 polypeptide shown in **Figure 4C** (SEQ ID NO:10).

[0020] **Figures 5A and 5B** depict the nucleic acid sequence (SEQ ID NO:11) encoding human LY6D polypeptide and the amino acid sequence of human LY6D polypeptide (SEQ ID NO:12).

[0021] **Figures 6A and 6B** depict the nucleic acid sequence (SEQ ID NO:13) encoding human LY6E polypeptide and the amino acid sequence of human LY6E polypeptide (SEQ ID NO:14).

[0022] **Figures 7A and 7B** depict the nucleic acid sequence (SEQ ID NO:15) encoding human LYPD2 polypeptide and the amino acid sequence of human LYPD2 polypeptide (SEQ ID NO:16).

[0023] **Figures 8A-8H** depict sequences of GLG-1 (ESL-1) molecules: (A-B) Accession No. U64791, nucleic acid sequence (SEQ ID NO:17) encoding human GLG-1 (ESL-1) polypeptide (SEQ ID NO:18); (C-D) Accession No. NM_012201, nucleic acid sequence (SEQ ID NO:19) encoding human GLG-1 (ESL-1) polypeptide (SEQ ID NO:20); (E-F) Accession No. AK172806, nucleic acid sequence (SEQ ID NO:21) encoding human GLG-1 (ESL-1) polypeptide (SEQ ID NO:22); and Accession No. AK131501, nucleic acid sequence (SEQ ID NO:23) encoding human GLG-1 (ESL-1) polypeptide (SEQ ID NO:24).

[0024] **Figures 9A and 9B** depict the nucleic acid sequence (SEQ ID NO:25) encoding murine LY6A polypeptide and the amino acid sequence of murine LY6A polypeptide (SEQ ID NO:26).

[0025] **Figures 10A and 10B** depict the nucleic acid sequence (SEQ ID NO:27) encoding murine LY6C polypeptide and the amino acid sequence of murine LY6C polypeptide (SEQ ID NO:28).

[0026] **Figures 11A and 11B** depict the nucleic acid sequence (SEQ ID NO:29) encoding murine LY6D polypeptide and the amino acid sequence of murine LY6D polypeptide (SEQ ID NO:30).

[0027] Figures 12A and 12B depict the nucleic acid sequence (SEQ ID NO:31) encoding murine LY6E polypeptide and the amino acid sequence of murine LY6E polypeptide (SEQ ID NO:32).

[0028] Figures 13A and 13B depict the nucleic acid sequence (SEQ ID NO:33) encoding murine LY6F polypeptide and the amino acid sequence of murine LY6F polypeptide (SEQ ID NO:34).

[0029] Figures 14A and 14B depict the nucleic acid sequence (SEQ ID NO:35) encoding murine LY6I polypeptide and the amino acid sequence of murine LY6I polypeptide (SEQ ID NO:36).

[0030] Figures 15A and 15B depict the nucleic acid sequence (SEQ ID NO:37) encoding murine LY6K polypeptide and the amino acid sequence of murine LY6K polypeptide (SEQ ID NO:38).

[0031] Figures 16A and 16B depict the nucleic acid sequence (SEQ ID NO:45) encoding murine LYPD3 polypeptide and the amino acid sequence of murine LYPD3 polypeptide (SEQ ID NO:46).

[0032] Figures 17A and 17B depict the nucleic acid sequence (SEQ ID NO:47) encoding murine LY6H polypeptide and the amino acid sequence of murine LY6H polypeptide (SEQ ID NO:48).

[0033] Figures 18A and 18B depict the nucleic acid sequence (SEQ ID NO:49) encoding murine LYPD1 polypeptide and the amino acid sequence of murine LYPD1 polypeptide (SEQ ID NO:50).

[0034] Figures 19A and 19B depict the nucleic acid sequence (SEQ ID NO:51) encoding murine LYPD2 polypeptide and the amino acid sequence of murine LYPD2 polypeptide (SEQ ID NO:52).

[0035] Figures 20A and 20B depict the nucleic acid sequence (SEQ ID NO:53) encoding murine LY6g5c polypeptide and the amino acid sequence of murine LY6g5c polypeptide (SEQ ID NO:54).

[0036] Figures 21A and 22B depict the nucleic acid sequence (SEQ ID NO:55) encoding murine LY6g6c polypeptide and the amino acid sequence of murine LY6g6c polypeptide (SEQ ID NO:56).

[0037] Figures 22A and 22B depict the nucleic acid sequence (SEQ ID NO:57) encoding murine SLURP2/LYNX1 polypeptide and the amino acid sequence of murine SLURP2/LYNX1 polypeptide (SEQ ID NO:58).

[0038] **Figure 23** shows that LY6 family members are upregulated in IEC in murine models of colitis. IEC in both the IL10^{-/-} (Figure 23A) and CD45RB^{Hi} transfer colitis model (Figure 23B) were isolated by LCM and RNA was purified. Microarray analysis was performed and analyzed as described in the Examples. Numbers represent the mean of the fold change compared to a universal standard RNA of colitic mice over healthy mice. Numbers below the heatmap indicate the inflammation score of the individual mouse.

[0039] **Figures 24A-24D** show that surface expression of LY6 molecules is upregulated on IEC in the IL10^{-/-} model of colitis. Wild type (Figure 24A) or IL10^{-/-} mice (Figure 24B) were stained for surface expression of LY6A (green, with DAPI counterstain). Similarly, wild type (Figure 24C) or IL10^{-/-} mice (Figure 24D) were stained for surface expression of LY6C.

[0040] **Figures 25A-25I** show that surface expression of LY6A and LY6C are upregulated in response to inflammatory cytokines, particularly IFN γ . YAMC cells were treated with the indicated cytokine for 15 hours and stained for surface expression of LY6C (Figure 25A) and LY6A (Figure 25B). YAMC cells were cultured for 15 hours in the presence of increasing doses of IFN γ and analyzed by flow cytometry for expression of LY6C (Figure 25C) and LY6A (Figure 25D). IFN γ stimulated YAMC cells were collected at various time points, as indicated, and analyzed by flow cytometry for expression of LY6C (Figure 25E) and LY6A (Figure 25F). IL-22 upregulated expression of both LY6C (Figure 25G) and LY6A (Figure 25H). Levels of both LY6A and LY6C were upregulated in the murine IEC line, CMT93 in response to treatment with IFN γ (Figure 25I).

[0041] **Figures 26A-26E** Lipid raft depletion results in an inhibition of LY6C-mediated chemokine production. Cholesterol depleted (dark bars) or non-depleted (open bars) YAMC cells were incubated with plate-bound nati-KLH or anti-LY6C as indicated for 15 hours. RNA was collected and expression levels of CXCL2, CXCL5, and CCL7 were determined (Figures 26A-26C). Surface levels of LY6A (Figure 26D) and LY6C (Figure 26E) were decreased in response to cholesterol depletion.

[0042] **Figures 27A-27D** show that crosslinking of LY6C, but not LY6A, induces upregulation of surface expression of LY6A and LY6C. YAMC cells were incubated for 24 hours on plates coated with anti-KLH control, anti-LY6A or anti-LY6C and analyzed by flow cytometry for expression of LY6C (Figure 27A) or LY6A (Figure 27B). Cells were pretreated for 12 hours with 100 U/ml of IFN γ and similarly plated on antibody coated plates and analyzed for expression of LY6C (Figure 27C) or LY6A (Figure 27D).

[0043] **Figures 28A-28C** show that crosslinking LY6C, but not LY6A, induces secretion of chemokines. Figure 28A: YAMC cells were preincubated or not, as indicated, with 100 U/ml of IFN γ for 15 hours and cultured on plates coated with 10 μ g/ml of anti-LY6A (black bars) or anti-LY6C (hatched bars) or anti-KLH control (open bars). RNA was isolated at 24 (left), 48 (center) and 72 (right) hours and analyzed for expression of CXCL5 or CCL7 (A). Data indicates mean $\pm$ SD of the fold change (as determined by $2^{-\Delta\Delta C_t}$ method) compared to untreated, isotype crosslinked cells. Figure 28B: Supernatants were collected at 48 hours in cells crosslinked, as above, with 1, 5 or 10 μ g/ml (as indicated) of antibody and CXCL5 secretion into the supernatant was determined by ELISA. * <0.05 . Figure 28C: Levels of both CXCL5 and CXCL2 in response to LY6C crosslinking were diminished when LY6C levels were knocked down with siRNA.

[0044] **Figures 29A-29B** show that IEC in colitis possess a similar chemokine gene expression pattern. IEC in both the IL10 $^{-/-}$ (Figure 29A) and CD45RB ^{Hi} transfer colitis model (Figure 29B) were isolated by LCM and RNA was purified. Microarray analysis was performed and analyzed as described in the Examples. Numbers represent the mean of the fold change compared to the universal standard RNA of colitic mice over healthy mice. Numbers below the heatmap indicate the inflammation score of the individual mouse.

[0045] **Figures 30A-30C** show that expression of human LY6 family genes is upregulated in colon cells treated with cytokines. Human Colo-205 cells were treated with the indicated cytokines, or combinations of cytokines, for 18 or 24 hours. The fold increase in expression of human LY6H (Figure 30A), human LYPD3 (Figure 30B), and human LYPD5 (Figure 30C) are shown relative to human β -actin control.

[0046] **Figures 31A-31B** show that patients with Crohn's Disease have elevated levels of LYPD1 (Figure 31A) and LYPD5 (Figure 31B) in the colon. Tissue samples from human IBD patients were obtained and LYPD1 and LYPD5 gene expression was determined. Statistically significant increases in expression of LYPD1 and LYPD5 were observed in inflamed tissue of CD patients. A statistically significant increase in expression of LYPD5 was also observed in inflamed tissue of UC patients. Y-axis values reflect gene expression relative to a universal RNA standard.

[0047] **Figures 32A and 32B** shows (A) untransfected COS cells, and (B) COS cells transfected with GLG-1 (ESL-1) polypeptide and stained with LYPD5-Fc protein.

[0048] **Figure 33A** depicts the structure of GLG-1 or ESL-1 and various fragments suitable for characterizing the binding of LYPD5 and **Figure 33B** shows the results of a co-immunoprecipitation study characterizing the binding of LYPD5 and an LYPD5 ligand.

[0049] **Figure 34A** depicts the structure of GLG-1 or ESL-1 and various fragments suitable for characterizing the binding of LYPD5 and **Figure 34B** shows the results of a co-immunoprecipitation study characterizing the binding of LYPD5 and an LYPD5 ligand.

[0050] **Figure 35A** depicts the structure of GLG-1 or ESL-1 and various fragments suitable for characterizing the binding of LYPD5 and **Figure 35B** shows the results of a co-immunoprecipitation study characterizing the binding of LYPD5 and an LYPD5 ligand.

[0051] **Figures 36A and 36B** depict the nucleic acid sequence (SEQ ID NO: 68) encoding human integrin, beta 7, and the amino acid sequence of human integrin, beta 7 polypeptide (SEQ ID NO: 69).

DETAILED DESCRIPTION OF THE INVENTION

Definitions

[0052] “Inflammatory Bowel Disease” or “IBD” is used interchangeably herein to refer to diseases of the bowel that cause inflammation and/or ulceration and includes without limitation Crohn’s disease and ulcerative colitis.

[0053] “Crohn’s disease (CD)” or “ulcerative colitis (UC)” are chronic inflammatory bowel diseases of unknown etiology. Crohn’s disease, unlike ulcerative colitis, can affect any part of the bowel. The most prominent feature Crohn’s disease is the granular, reddish-purple edematous thickening of the bowel wall. With the development of inflammation, these granulomas often lose their circumscribed borders and integrate with the surrounding tissue. Diarrhea and obstruction of the bowel are the predominant clinical features. As with ulcerative colitis, the course of Crohn’s disease may be continuous or relapsing, mild or severe, but unlike ulcerative colitis, Crohn’s disease is not curable by resection of the involved segment of bowel. Most patients with Crohn’s disease require surgery at some point, but subsequent relapse is common and continuous medical treatment is usual.

[0054] Crohn’s disease may involve any part of the alimentary tract from the mouth to the anus, although typically it appears in the ileocolic, small-intestinal or colonic-anorectal regions. Histopathologically, the disease manifests by discontinuous granulomas, crypt abscesses, fissures and aphthous ulcers. The inflammatory infiltrate is mixed, consisting of lymphocytes (both T and B cells), plasma cells, macrophages, and neutrophils. There is a

disproportionate increase in IgM- and IgG-secreting plasma cells, macrophages and neutrophils.

[0055] Anti-inflammatory drugs sulfasalazine and 5-aminosalicylic acid (5-ASA) are useful for treating mildly active colonic Crohn's disease and is commonly prescribed to maintain remission of the disease. Metronidazole and ciprofloxacin are similar in efficacy to sulfasalazine and appear to be particularly useful for treating perianal disease. In more severe cases, corticosteroids are effective in treating active exacerbations and can even maintain remission. Azathioprine and 6-mercaptopurine have also shown success in patients who require chronic administration of corticosteroids. It is also possible that these drugs may play a role in the long-term prophylaxis. Unfortunately, there can be a very long delay (up to six months) before onset of action in some patients.

[0056] Antidiarrheal drugs can also provide symptomatic relief in some patients. Nutritional therapy or elemental diet can improve the nutritional status of patients and induce symptomatic improvement of acute disease, but it does not induce sustained clinical remissions. Antibiotics are used in treating secondary small bowel bacterial overgrowth and in treatment of pyogenic complications.

[0057] "Ulcerative colitis (UC)" afflicts the large intestine. The course of the disease may be continuous or relapsing, mild or severe. The earliest lesion is an inflammatory infiltration with abscess formation at the base of the crypts of Lieberkühn. Coalescence of these distended and ruptured crypts tends to separate the overlying mucosa from its blood supply, leading to ulceration. Symptoms of the disease include cramping, lower abdominal pain, rectal bleeding, and frequent, loose discharges consisting mainly of blood, pus and mucus with scanty fecal particles. A total colectomy may be required for acute, severe or chronic, unremitting ulcerative colitis.

[0058] The clinical features of UC are highly variable, and the onset may be insidious or abrupt, and may include diarrhea, tenesmus and relapsing rectal bleeding. With fulminant involvement of the entire colon, toxic megacolon, a life-threatening emergency, may occur. Extraintestinal manifestations include arthritis, pyoderma gangrenosum, uveitis, and erythema nodosum.

[0059] Treatment for UC includes sulfasalazine and related salicylate-containing drugs for mild cases and corticosteroid drugs in severe cases. Topical administration of either salicylates or corticosteroids is sometimes effective, particularly when the disease is limited to the distal bowel, and is associated with decreased side effects compared with systemic use.

Supportive measures such as administration of iron and antidiarrheal agents are sometimes indicated. Azathioprine, 6-mercaptopurine and methotrexate are sometimes also prescribed for use in refractory corticosteroid-dependent cases.

[0060] As used herein, “LY6 gene family member” or “LY6 gene superfamily member” is used interchangeably herein to refer to a gene having homology to members of the LY6 gene family, the majority of which gene family members are GPI-anchored cell surface glycoproteins with broad distribution on cells of hematopoietic origin and more limited expression on non-hematopoietic cells. Members of this gene family are used as markers of differentiation of immune cells (Sunderkotter, C. et al., J. Immunol. 172:4410-4417 (2004)). Genes of the LY6 family have been examined (Shevach, E.M. and P.E. Korthy, Immunol. Today 10:195-200 (1989)) and functions include T cell activation (Zhang, Z.X. et al., Eur. J. Immunol. 32:1584-1592 (2002) and Henderson, S.C. et al., J. Immunol. 168:118-126 (2002), olfaction (Chou, J.H. et al., Genetics 157:211-224 (2001) and cellular adhesion (Jaakkola, I. et al., J. Immunol. 170:1283-1290 (2003)). Members of the LY6 gene family include without limitation members of the mammalian LY6 gene family, such as the LY6 family genes of mouse or human. As use here, “LY6 gene” refers to a LY6 gene family member and “LY6 polypeptide” refers to the polypeptide encoded by a LY6 gene. Murine LY6 gene family members include, without limitation, LY6A (NM_010738, nucleic acid SEQ ID NO:25 which encodes polypeptide SEQ ID NO:26), LY6C (NM_010741, nucleic acid SEQ ID NO:27 which encodes polypeptide SEQ ID NO:28), LY6D (NM_003695, nucleic acid SEQ ID NO:29 which encodes polypeptide SEQ ID NO:30), LY6E (NM_002346, nucleic acid SEQ ID NO:31 which encodes polypeptide SEQ ID NO:32), LY6F (NM_008530, nucleic acid SEQ ID NO:33 which encodes polypeptide SEQ ID NO:34), LY6I (NM_020498, nucleic acid SEQ ID NO:35 which encodes polypeptide SEQ ID NO:36), and LY6K (NM_017527, nucleic acid SEQ ID NO:37 which encodes polypeptide SEQ ID NO:38). Human LY6 gene family members include, without limitation, LY6H (NM_002347, nucleic acid SEQ ID NO:1 which encodes polypeptide SEQ ID NO:2), LYPD1 (NM_144586, nucleic acid SEQ ID NOS:3 or 4 which encodes polypeptide SEQ ID NO:5), LYPD3 (NM_014400, nucleic acid SEQ ID NO:6 which encodes polypeptide SEQ ID NO:7), LYPD5 (NM_182573, nucleic acid SEQ ID NOS:8 or 9 which encodes polypeptide SEQ ID NO:10), LY6D (NM_003695, nucleic acid SEQ ID NO:11 which encodes polypeptide SEQ ID NO:12), LY6E (NMNM_002346, nucleic acid SEQ ID NO:13 which encodes polypeptide SEQ ID NO:14), LYPD2 (NM_205545, nucleic acid SEQ ID NO:15 which encodes polypeptide SEQ ID NO:16). In embodiments, the polynucleotide of

each LY6 gene family member disclosed herein comprises at least 15, at least 25, at least, at least 50, at least 100, at least 250, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, or at least 2040 contiguous nucleotides of SEQ ID NOs:1, 3, 4, 6, 8, 9, 11, 13, 15, 25, 27, 29, 31, 33, 35, 37, 45, 47, 49, 51, 53, 55, or 57, or the LY6 gene family member polynucleotide comprises SEQ ID NOS:1, 3, 4, 6, 8, 9, 11, 13, 15, 25, 27, 29, 31, 33, 35, 37, 45, 47, 49, 51, 53, 55, or 57. In one embodiment, a polynucleotide that binds a LY6 gene family member polynucleotide (SEQ ID NOs:1, 3, 4, 6, 8, 9, 11, 13, 15, 25, 27, 29, 31, 33, 35, 37, 45, 47, 49, 51, 53, 55, or 57), or fragment thereof, has at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% or 100% sequence identity with the LY6 polypeptide or fragment thereof. In one embodiment, the LY6 gene family member polypeptide comprises at least 10, at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 225, at least 250, at least 275, at least 300, or at least 325, at least contiguous amino acids of SEQ ID NOs:2, 5, 7, 10, 12, 14, 26, 28, 30, 32, 34, 36, 38, 46, 48, 50, 52, 54, 56, or 58, or the LY6 gene family polypeptide comprises SEQ ID NOs:2, 5, 7, 10, 12, 14, 26, 28, 30, 32, 34, 36, 38, 46, 48, 50, 52, 54, 56, or 58).

[0061] A “native sequence polypeptide” of any of the LY6 gene family members comprises a polypeptide having the same amino acid sequence as the corresponding LY6 gene family member polypeptide derived from nature. Such native sequence LY6 polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term “native sequence LY6 polypeptide” specifically encompasses naturally-occurring truncated or secreted forms of the specific LY6 polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In one specific aspect, the native sequence LY6 polypeptides disclosed herein are mature or full-length native sequence polypeptides corresponding to the sequences in Figures 1-7 and SEQ ID NOs:2, 5, 7, 10, 12, 14, 26, 28, 30, 32, 34, 36, 38, 46, 48, 50, 52, 54, 56, or 58.

[0062] As used herein, a “LY6 polypeptide variant” means a LY6 polypeptide, preferably biologically active forms thereof, as defined herein, having at least about 80% amino acid sequence identity with a full-length native sequence LY6 polypeptide sequence, as disclosed herein, and variant forms thereof lacking the signal peptide, an extracellular domain, or any other fragment of a full length native sequence LY6 polypeptide such as those referenced herein. Such variant polypeptides include, for instance, polypeptides wherein one

or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. In a specific aspect, such variant polypeptides will have at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to a full-length native sequence LY6 polypeptide sequence polypeptide, as disclosed herein, and variant forms thereof lacking a signal peptide, an extracellular domain, or any other fragment of a full length native sequence LY6 polypeptide such as those disclosed herein.

[0063] "Percent (%) amino acid sequence identity" with respect to a LY6 polypeptide sequence identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific LY6 polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

[0064] As used herein "LY6 variant polynucleotide" or "LY6 variant nucleic acid sequence," or "LY6 gene" refers to a nucleic acid molecule which encodes a LY6 gene family member polypeptide, preferably biologically active forms thereof, as defined herein, and which have at least about 80% nucleic acid sequence identity with a nucleotide acid sequence

encoding a full-length native sequence LY6 polypeptide sequence identified herein, or any other fragment of the respective full-length LY6 polypeptide sequence as identified herein (such as those encoded by a nucleic acid that represents only a portion of the complete coding sequence for a full-length LY6 polypeptide). Ordinarily, such variant polynucleotides will have at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% nucleic acid sequence identity with a nucleic acid sequence encoding the respective full-length native sequence LY6 polypeptide sequence or any other fragment of the respective full-length LY6 polypeptide sequence identified herein. Such variant polynucleotides do not encompass the native nucleotide sequence.

[0065] Ordinarily, such variant polynucleotides vary at least about 50 nucleotides in length from the native sequence polypeptide, alternatively the variance can be at least about 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length.

[0066] "Percent (%) nucleic acid sequence identity" with respect to a LY6 gene polypeptide-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the LY6 gene nucleic acid sequence of interest, respectively, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is

registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

[0067] In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program=s alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "REF-DNA", wherein "REF-DNA" represents a hypothetical LY6 gene-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "REF-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent different hypothetical nucleotides. Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program.

[0068] In other embodiments, LY6 gene variant polynucleotides are nucleic acid molecules that encode LY6 polypeptide, respectively, and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length LY6 polypeptide, respectively, as disclosed herein. Such variant polypeptides may be those that are encoded by such variant polynucleotides.

[0069] “Isolated”, when used to describe the various LY6 polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, such polypeptides will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Such isolated polypeptides includes the corresponding polypeptides *in situ* within recombinant cells, since at least one component of the LY6 polypeptide from its natural environment will not be present. Ordinarily, however, such isolated polypeptides will be prepared by at least one purification step.

[0070] An “isolated” LY6 polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. Any of the above such isolated nucleic acid molecule is other than in the form or setting in which it is found in nature. Any such nucleic acid molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells.

[0071] The term “control sequences” refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

[0072] Nucleic acid is “operably linked” when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, “operably linked” means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If

such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

[0073] As used herein "expression" as applied to gene expression, refers to transcription of a gene encoding a protein to produce mRNA as well as translation of the mRNA to produce the protein encoded by the gene. Thus, increased or decreased expression refers to increased or decreased transcription of a gene and/or increased or decreased translation of mRNA resulting from transcription.

[0074] "Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

[0075] "Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50EC; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42EC; or (3) overnight hybridization in a solution that employs 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42EC, with a 10 minute wash at 42EC in 0.2 x SSC (sodium chloride/sodium citrate) followed by a 10 minute high-stringency wash consisting of 0.1 x SSC containing EDTA at 55EC.

[0076] "Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor

Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The ordinarily skilled artisan will recognize how to adjust the temperature, ionic strength, *etc.* as necessary to accommodate factors such as probe length and the like.

[0077] The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising an LY6 polypeptide, or LY6 polypeptide binding agent fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with the activity of the polypeptide to which it is fused. The tag polypeptide preferably also is sufficiently unique so that such antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

[0078] "Active" or "activity" for the purposes herein refers to form(s) of polypeptides which retain a biological and/or an immunological activity of native or naturally-occurring polypeptide, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring polypeptide other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring polypeptide, and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring polypeptide. An active polypeptide, as used herein, is an antigen that is differentially expressed, either from a qualitative or quantitative perspective, in IBD tissue, relative to its expression on similar tissue that is not afflicted with IBD.

[0079] The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native polypeptide disclosed herein. Suitable antagonist molecules specifically include antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native polypeptides, peptides, antisense oligonucleotides, small organic molecules, *etc.* Methods for identifying antagonists may comprise contacting such a polypeptide, including a cell

expressing it, with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with such polypeptide.

[0080] “Treating” or “treatment” or “alleviation” refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the progression of a disease. Treatment may also refer the modification of the progression of an IBD.

[0081] “Diagnosing” refers to the process of identifying or determining the distinguishing characteristics of a disease including without limitation IBD, UC and/or Crohn’s Disease. The process of diagnosing is sometimes also expressed as staging or disease classification based on severity or disease progression as well as on location (such as, for example, location within or along the gastrointestinal tract at which inflammation and/or altered gene expression is found).

[0082] Subjects in need of diagnosis include those already experiencing with aberrant LY6 expression as well as those prone to having or those in whom aberrant LY6 expression is to be prevented. Accordingly, an aspect of the invention is the detection of a therapeutic drug response in a mammal treated with a therapeutic agent for the treatment of IBD, wherein the method comprises determining the LY6 expression in gastrointestinal tissue of a test mammal relative to a control and determining that the LY6 expression levels are within not significantly different from normal control expression levels. In one embodiment, a therapeutic response is determined when the levels of expression of LY6 of the mammal treated with a therapeutic agent are different (expression is more similar to normal control, i.e., LY6 expression levels are lower than LY6 expression levels were in the mammal prior to treatment).

[0083] The above parameters for assessing successful treatment and improvement in the disease are readily measurable by routine procedures familiar to a physician. For IBD therapy, efficacy can be measured, for example, by assessing the time to disease progression (TTP) and/or determining the response rate (RR). Biopsies may be taken to assess gene expression and observe histopathology of gastrointestinal tissue from the patient. The invention described herein relating to the process of prognosing and/or diagnosing involves the determination and evaluation of LY6 gene expression upregulation.

[0084] “Mammal” or “mammalian subject” for purposes of the treatment of, alleviating the symptoms of or diagnosis of a IBD refers to any animal classified as a mammal,

including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, ferrets, *etc.* Preferably, the mammal is human.

[0085] Administration “in combination with” one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

[0086] “Carriers” as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN7, polyethylene glycol (PEG), and PLURONICS7.

[0087] By “solid phase” or “solid support” is meant a non-aqueous matrix to which a polypeptide, nucleic acid, antibody or LY6 binding agent can adhere or attach. Examples of solid phases encompassed herein include those formed partially or entirely of glass (*e.g.*, controlled pore glass), polysaccharides (*e.g.*, agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (*e.g.*, an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

[0088] A “liposome” is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

[0089] A “small molecule” or “small organic molecule” is defined herein to have a molecular weight below about 500 Daltons.

[0090] An “effective amount” of an antagonist agent is an amount sufficient to bring about a physiological effect, such as without limitation to inhibit, partially or entirely, function of gene or its encoded protein. An “effective amount” may be determined empirically and in a routine manner, in relation to this purpose.

[0091] The term “therapeutically effective amount” refers to an antagonist or other drug effective to “treat” a disease or disorder in a subject or mammal. In the case of IBD, the therapeutically effective amount of the drug will restore aberrant LY6 expression to normal physiological levels; reduce gastrointestinal inflammation; reduce the number of gastrointestinal lesions; and/or relieve to some extent one or more of the symptoms associated with IBD, UC and/or CD. See the definition herein of “treating”.

[0092] A “growth inhibitory amount” of an antagonist is an amount capable of inhibiting the growth of a cell, especially tumor, *e.g.*, cancer cell, either *in vitro* or *in vivo*. For purposes of inhibiting neoplastic cell growth, such an amount may be determined empirically and in a routine manner.

[0093] A “cytotoxic amount” of an antagonist is an amount capable of causing the destruction of a cell, especially a proliferating cell, *e.g.*, cancer cell, either *in vitro* or *in vivo*. For purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

[0094] The term “antibody” is used in the broadest sense and specifically covers, for example, anti- LY6 monoclonal antibodies (including antagonist and neutralizing antibodies), anti- LY6 antibody compositions with polyepitopic specificity, polyclonal antibodies, single chain anti- LY6 antibodies, multispecific antibodies (*e.g.*, bispecific) and antigen binding fragments (see below) of all of the above enumerated antibodies as long as they exhibit the desired biological or immunological activity. The term “immunoglobulin” (Ig) is used interchangeably with antibody herein.

[0095] An “isolated” antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody *in situ* within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

[0096] The basic 4-chain antibody unit is a heterotetrameric glycoprotein composed of two identical light (L) chains and two identical heavy (H) chains (an IgM antibody consists of 5 of the basic heterotetramer unit along with an additional polypeptide called J chain, and therefore contain 10 antigen binding sites, while secreted IgA antibodies can polymerize to form polyvalent assemblages comprising 2-5 of the basic 4-chain units along with J chain). In the case of IgGs, the 4-chain unit is generally about 150,000 daltons. Each L chain is linked to an H chain by one covalent disulfide bond, while the two H chains are linked to each other by one or more disulfide bonds depending on the H chain isotype. Each H and L chain also has regularly spaced intrachain disulfide bridges. Each H chain has at the N-terminus, a variable domain (V_H) followed by three constant domains (C_H) for each of the α and γ chains and four C_H domains for μ and ϵ isotypes. Each L chain has at the N-terminus, a variable domain (V_L) followed by a constant domain (C_L) at its other end. The V_L is aligned with the V_H and the C_L is aligned with the first constant domain of the heavy chain (C_{H1}). Particular amino acid residues are believed to form an interface between the light chain and heavy chain variable domains. The pairing of a V_H and V_L together forms a single antigen-binding site. For the structure and properties of the different classes of antibodies, see, *e.g.*, *Basic and Clinical Immunology*, 8th edition, Daniel P. Stites, Abba I. Terr and Tristram G. Parslow (eds.), Appleton & Lange, Norwalk, CT, 1994, page 71 and Chapter 6.

[0097] The L chain from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains. Depending on the amino acid sequence of the constant domain of their heavy chains (C_H), immunoglobulins can be assigned to different classes or isotypes. There are five classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, having heavy chains designated α , δ , ϵ , γ , and μ , respectively. The γ and α classes are further divided into subclasses on the basis of relatively minor differences in C_H sequence and function, *e.g.*, humans express the following subclasses: IgG1, IgG2, IgG3, IgG4, IgA1, and IgA2.

[0098] The term "variable" refers to the fact that certain segments of the variable domains differ extensively in sequence among antibodies. The V domain mediates antigen binding and define specificity of a particular antibody for its particular antigen. However, the variability is not evenly distributed across the approximately 110-amino acid span of the variable domains. Instead, the V regions consist of relatively invariant stretches called framework regions (FRs) of 15-30 amino acids separated by shorter regions of extreme variability called "hypervariable regions" that are each 9-12 amino acids long. The variable

domains of native heavy and light chains each comprise four FRs, largely adopting a β -sheet configuration, connected by three hypervariable regions, which form loops connecting, and in some cases forming part of, the β -sheet structure. The hypervariable regions in each chain are held together in close proximity by the FRs and, with the hypervariable regions from the other chain, contribute to the formation of the antigen-binding site of antibodies (see Kabat *et al.*, *Sequences of Proteins of Immunological Interest*, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)). The constant domains are not involved directly in binding an antibody to an antigen, but exhibit various effector functions, such as participation of the antibody in antibody dependent cellular cytotoxicity (ADCC).

[0099] The term “hypervariable region” when used herein refers to the amino acid residues of an antibody which are responsible for antigen-binding. The hypervariable region generally comprises amino acid residues from a “complementarity determining region” or “CDR” (*e.g.* around about Kabat residues 24-34 (L1), 50-56 (L2) and 89-97 (L3) in the V_L , and around about Kabat residues 31-35B (H1), 50-65 (H2) and 95-102 (H3) in the V_H (Kabat *et al.*, *Sequences of Proteins of Immunological Interest*, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)) and/or those residues from a “hypervariable loop” (*e.g.* around about Chothia residues 26-32 (L1), 50-52 (L2) and 91-96 (L3) in the V_L , and 26-32 (H1), 52A-55 (H2) and 96-101 (H3) in the V_H (Chothia and Lesk J. Mol. Biol. 196:901-917 (1987))).

[0100] The term “monoclonal antibody” as used herein refers to an antibody from a population of substantially homogeneous antibodies, *i.e.*, the individual antibodies comprising the population are identical and/or bind the same epitope(s), except for possible variants that may arise during production of the monoclonal antibody, such variants generally being present in minor amounts. Such monoclonal antibody typically includes an antibody comprising a polypeptide sequence that binds a target, wherein the target-binding polypeptide sequence was obtained by a process that includes the selection of a single target binding polypeptide sequence from a plurality of polypeptide sequences. For example, the selection process can be the selection of a unique clone from a plurality of clones, such as a pool of hybridoma clones, phage clones or recombinant DNA clones. It should be understood that the selected target binding sequence can be further altered, for example, to improve affinity for the target, to humanize the target binding sequence, to improve its production in cell culture, to reduce its immunogenicity *in vivo*, to create a multispecific antibody, *etc.*, and that an antibody comprising the altered target binding sequence is also a monoclonal antibody of this invention.

In contrast to polyclonal antibody preparations which typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody of a monoclonal antibody preparation is directed against a single determinant on an antigen. In addition to their specificity, the monoclonal antibody preparations are advantageous in that they are typically uncontaminated by other immunoglobulins. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies to be used in accordance with the present invention may be made by a variety of techniques, including, for example, the hybridoma method (*e.g.*, Kohler *et al.*, *Nature*, 256:495 (1975); Harlow *et al.*, *Antibodies: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling *et al.*, in: *Monoclonal Antibodies and T-Cell Hybridomas* 563-681, (Elsevier, N.Y., 1981)), recombinant DNA methods (see, *e.g.*, U.S. Patent No. 4,816,567), phage display technologies (see, *e.g.*, Clackson *et al.*, *Nature*, 352:624-628 (1991); Marks *et al.*, *J. Mol. Biol.*, 222:581-597 (1991); Sidhu *et al.*, *J. Mol. Biol.* 338(2):299-310 (2004); Lee *et al.*, *J. Mol. Biol.* 340(5):1073-1093 (2004); Fellouse, *Proc. Nat. Acad. Sci. USA* 101(34):12467-12472 (2004); and Lee *et al.* *J. Immunol. Methods* 284(1-2):119-132 (2004), and technologies for producing human or human-like antibodies in animals that have parts or all of the human immunoglobulin loci or genes encoding human immunoglobulin sequences (see, *e.g.*, WO 1998/24893; WO 1996/34096; WO 1996/33735; WO 1991/10741; Jakobovits *et al.*, *Proc. Natl. Acad. Sci. USA*, 90:2551 (1993); Jakobovits *et al.*, *Nature*, 362:255-258 (1993); Bruggemann *et al.*, *Year in Immuno.*, 7:33 (1993); U.S. Patent Nos. 5,545,806; 5,569,825; 5,591,669 (all of GenPharm); 5,545,807; WO 1997/17852; U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; and 5,661,016; Marks *et al.*, *Bio/Technology*, 10: 779-783 (1992); Lonberg *et al.*, *Nature*, 368: 856-859 (1994); Morrison, *Nature*, 368: 812-813 (1994); Fishwild *et al.*, *Nature Biotechnology*, 14: 845-851 (1996); Neuberger, *Nature Biotechnology*, 14: 826 (1996); and Lonberg and Huszar, *Intern. Rev. Immunol.*, 13: 65-93 (1995).

[0101] "Chimeric" antibodies (immunoglobulins) have a portion of the heavy and/or light chain identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity

(U.S. Patent No. 4,816,567; and Morrison et al., Proc. Natl. Acad. Sci. USA 81:6851-6855 (1984)). Humanized antibody as used herein is a subset of chimeric antibodies.

[0102] "Humanized" forms of non-human (*e.g.*, murine) antibodies are chimeric antibodies which contain minimal sequence derived from non-human immunoglobulin. For the most part, humanized antibodies are human immunoglobulins (recipient or acceptor antibody) in which hypervariable region residues of the recipient are replaced by hypervariable region residues from a non-human species (donor antibody) such as mouse, rat, rabbit or nonhuman primate having the desired specificity, affinity, and capacity. In some instances, Fv framework region (FR) residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues which are not found in the recipient antibody or in the donor antibody. These modifications are made to further refine antibody performance such as binding affinity. Generally, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable loops correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin sequence although the FR regions may include one or more amino acid substitutions that improve binding affinity. The number of these amino acid substitutions in the FR are typically no more than 6 in the H chain, and in the L chain, no more than 3. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see Jones *et al.*, *Nature* 321:522-525 (1986); Reichmann *et al.*, *Nature* 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992).

[0103] "Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, and Fv fragments; diabodies; linear antibodies (see U.S. Patent No. 5,641,870, Example 2; Zapata *et al.*, *Protein Eng.* 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

[0104] Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. The Fab fragment consists of an entire L chain along with the variable region domain of the H chain (VH), and the first constant domain of one heavy chain (C_{H1}). Each Fab fragment is monovalent with respect to antigen binding, *i.e.*, it has a single antigen-binding site. Pepsin treatment of an antibody yields a single large F(ab')₂ fragment

which roughly corresponds to two disulfide linked Fab fragments having divalent antigen-binding activity and is still capable of cross-linking antigen. Fab= fragments differ from Fab fragments by having additional few residues at the carboxy terminus of the C_H1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')₂ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

[0105] The Fc fragment comprises the carboxy-terminal portions of both H chains held together by disulfides. The effector functions of antibodies are determined by sequences in the Fc region, which region is also the part recognized by Fc receptors (FcR) found on certain types of cells.

[0106] "Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This fragment consists of a dimer of one heavy- and one light-chain variable region domain in tight, non-covalent association. From the folding of these two domains emanate six hypervariable loops (3 loops each from the H and L chain) that contribute the amino acid residues for antigen binding and confer antigen binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

[0107] "Single-chain Fv" also abbreviated as "sFv" or "scFv" are antibody fragments that comprise the V_H and V_L antibody domains connected into a single polypeptide chain. Preferably, the sFv polypeptide further comprises a polypeptide linker between the V_H and V_L domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in *The Pharmacology of Monoclonal Antibodies*, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994); Borrebaeck 1995, *infra*.

[0108] As used herein "LY6 binding polypeptide" is an oligopeptide that binds, preferably specifically, to a LY6 polypeptide, ligand or signaling component, respectively, or a LY6 binding portion or fragment thereof. Such oligopeptides may be chemically synthesized using known oligopeptide synthesis methodology or may be prepared and purified using recombinant technology. Such oligopeptides are usually at least about 5 amino acids in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24,

25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length or more. Such oligopeptides may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening oligopeptide libraries for oligopeptides that are capable of specifically binding to a polypeptide target are well known in the art (see, *e.g.*, U.S. Patent Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publication Nos. WO 84/03506 and WO84/03564; Geysen *et al.*, Proc. Natl. Acad. Sci. U.S.A., 81:3998-4002 (1984); Geysen *et al.*, Proc. Natl. Acad. Sci. U.S.A., 82:178-182 (1985); Geysen *et al.*, in Synthetic Peptides as Antigens, 130-149 (1986); Geysen *et al.*, J. Immunol. Meth., 102:259-274 (1987); Schoofs *et al.*, J. Immunol., 140:611-616 (1988), Cwirla, S. E. *et al.* Proc. Natl. Acad. Sci. USA, 87:6378 (1990); Lowman, H.B. *et al.* Biochemistry, 30:10832 (1991); Clackson, T. *et al.* Nature, 352: 624 (1991); Marks, J. D. *et al.*, J. Mol. Biol., 222:581 (1991); Kang, A.S. *et al.* Proc. Natl. Acad. Sci. USA, 88:8363 (1991), and Smith, G. P., Current Opin. Biotechnol., 2:668 (1991).

[0109] An LY6 antagonist (*e.g.*, antibody, polypeptide, oligopeptide or small molecule) “which binds” a target antigen of interest, *e.g.* LY6 is one that binds the target with sufficient affinity so as to be a useful diagnostic, prognostic and/or therapeutic agent in targeting a cell or tissue expressing the antigen, and does not significantly cross-react with other proteins. The extent of binding to a non-desired marker polypeptide will be less than about 10% of the binding to the particular desired target, as determinable by common techniques such as fluorescence activated cell sorting (FACS) analysis or radioimmunoprecipitation (RIA).

[0110] Moreover, the term “specific binding” or “specifically binds to” or is “specific for” a particular LY6 polypeptide or an epitope on a particular LY6 polypeptide target means binding that is measurably different from a non-specific interaction. Specific binding can be measured, for example, by determining binding of a molecule compared to binding of a control molecule, which generally is a molecule of similar structure that does not have binding activity. For example, specific binding can be determined by competition with a control molecule that is similar to the target, for example, an excess of non-labeled target. In this case, specific binding is indicated if the binding of the labeled target to a probe is competitively inhibited by excess unlabeled target. In one embodiment, such terms refer to

binding where a molecule binds to a particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope. Alternatively, such terms can be described by a molecule having a K_d for the target of at least about 10^{-4} M, 10^{-5} M, 10^{-6} M, 10^{-7} M, 10^{-8} M, 10^{-9} M, 10^{-10} M, 10^{-11} M, 10^{-12} M, or greater.

[0111] A gastrointestinal cell or tissue that "overexpresses" LY6 if that cell or tissue is shown to have increased nucleic acid encoding LY6 in a cells or if that cell or tissue over produces and secretes LY6 protein, compared to a normal gastrointestinal cell or tissue of the same tissue type. Such overexpression may result from gene amplification or by increased transcription or translation. Various diagnostic or prognostic assays are known that measure altered expression levels resulting in increased or decreased levels at the cell surface or increased or decreased levels of secreted protein and include without limitation immunohistochemistry assay using anti- LY6 antibodies, FACS analysis, *etc.* Alternatively, the levels of LY6 encoding nucleic acid or mRNA can be measured in the cell, *e.g.*, via fluorescent *in situ* hybridization using a nucleic acid based probe corresponding to a LY6-encoding nucleic acid or the complement thereof; (FISH; see WO98/45479 published October, 1998), Southern blotting, Northern blotting, or polymerase chain reaction (PCR) techniques, such as real time quantitative PCR (RT-PCR). Alternatively, LY6 polypeptide overexpression is determinable by measuring shed antigen in a biological fluid such as serum, *e.g.*, using antibody-based assays (see also, *e.g.*, U.S. Patent No. 4,933,294 issued June 12, 1990; WO91/05264 published April 18, 1991; U.S. Patent 5,401,638 issued March 28, 1995; and Sias *et al.*, J. Immunol. Methods 132:73-80 (1990)). In addition to the above assays, various *in vivo* assays are available to the skilled practitioner. For example, one may expose cells within the body of the patient to an antibody which is optionally labeled with a detectable label, *e.g.*, a radioactive isotope, and binding of the antibody to cells in the patient can be evaluated, *e.g.*, by external scanning for radioactivity or by analyzing a biopsy taken from a patient previously exposed to the therapeutic agent.

[0112] As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (*i.e.*, is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of

a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

[0113] The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody, oligopeptide or other organic molecule so as to generate a "labeled" antibody, oligopeptide or other organic molecule. The label may be detectable by itself (*e.g.* radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

[0114] The term "cytotoxic agent" as used herein refers to a substance that inhibits or prevents the function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes (*e.g.*, At²¹¹, I¹³¹, I¹²⁵, Y⁹⁰, Re¹⁸⁶, Re¹⁸⁸, Sm¹⁵³, Bi²¹², P³² and radioactive isotopes of Lu), chemotherapeutic agents, enzymes and fragments thereof such as nucleolytic enzymes, antibiotics, and toxins such as small molecule toxins or enzymatically active toxins of bacterial, fungal, plant or animal origin, including fragments and/or variants thereof, and the various antitumor or anticancer agents disclosed below. Other cytotoxic agents are described below. A tumoricidal agent causes destruction of tumor cells.

[0115] A "chemotherapeutic agent" or "therapeutic agent" is a chemical compound useful in the treatment of a disorder or disease. Examples of chemotherapeutic or therapeutic agents for the treatment of IBD include without limitation anti-inflammatory drugs sulfasalazine and 5-aminosalicylic acid (5-ASA); metroidazole and ciprofloxacin are similar in efficacy to sulfasalazine and appear to be particularly useful for treating perianal disease; in more severe cases, corticosteroids are effective in treating active exacerbations and can even maintain remission; azathioprine, 6-mercaptopurine, and methotrexate have also shown success in patients who require chronic administration of cortico steroids; antidiarrheal drugs can also provide symptomatic relief in some patients; nutritional therapy or elemental diet can improve the nutritional status of patients and induce symptomatic improvement of acute disease; antibiotics are used in treating secondary small bowel bacterial overgrowth and in treatment of pyogenic complications. IBD chemotherapeutic agents further include biologicals and other agents as follows: anti-beta7 antibodies (see, for example, WO2006026759), anti-alpha4 antibodies (such as ANTEGEN®), anti-TNF antibody (REMICADE®) or non-protein compounds including without limitation 5-ASA compounds ASACOL®, PENTASA™, ROWASA™, COLAZAL™, and other compounds such as Purinethol and steroids such as

prednisone. Examples of chemotherapeutic agents for the treatment of cancer include hydroxyureataxanes (such as paclitaxel and doxetaxel) and/or anthracycline antibiotics; alkylating agents such as thiotepa and CYTOXAN⁷ cyclophosphamide; alkyl sulfonates such as busulfan, improsulfan and piposulfan; aziridines such as benzodopa, carboquone, meturedopa, and uredopa; ethylenimines and methylamelamines including altretamine, triethylenemelamine, trietylenephosphoramide, triethylenethiophosphoramide and trimethylolomelamine; acetogenins (especially bullatacin and bullatacinone); delta-9-tetrahydrocannabinol (dronabinol, MARINOL⁷); beta-lapachone; lapachol; colchicines; betulinic acid; a camptothecin (including the synthetic analogue topotecan (HYCAMTIN⁷), CPT-11 (irinotecan, CAMPTOSAR⁷), acetylcamptothecin, scoplectin, and 9-aminocamptothecin); bryostatin; callystatin; CC-1065 (including its adozelesin, carzelesin and bizelesin synthetic analogues); podophyllotoxin; podophyllinic acid; teniposide; cryptophycins (particularly cryptophycin 1 and cryptophycin 8); dolastatin; duocarmycin (including the synthetic analogues, KW-2189 and CB1-TM1); eleutherobin; pancratistatin; a sarcodictyin; spongistatin; nitrogen mustards such as chlorambucil, chlornaphazine, cholophosphamide, estramustine, ifosfamide, mechlorethamine, mechlorethamine oxide hydrochloride, melphalan, novembichin, phenesterine, prednimustine, trofosfamide, uracil mustard; nitrosureas such as carmustine, chlorozotocin, fotemustine, lomustine, nimustine, and ranimustine; antibiotics such as the enediyne antibiotics (e. g., calicheamicin, especially calicheamicin gammaII and calicheamicin omegaII (see, *e.g.*, Agnew, Chem Intl. Ed. Engl., 33: 183-186 (1994))); dynemicin, including dynemicin A; an esperamicin; as well as neocarzinostatin chromophore and related chromoprotein enediyne antiobiotic chromophores), aclacinomysins, actinomycin, authramycin, azaserine, bleomycins, cactinomycin, carabycin, carminomycin, carzinophilin, chromomycins, dactinomycin, daunorubicin, detorubicin, 6-diazo-5-oxo-L-norleucine, ADRIAMYCIN⁷ doxorubicin (including morpholino-doxorubicin, cyanomorpholino-doxorubicin, 2-pyrrolino-doxorubicin and deoxydoxorubicin), epirubicin, esorubicin, idarubicin, marcellomycin, mitomycins such as mitomycin C, mycophenolic acid, nogalamycin, olivomycins, peplomycin, potfiromycin, puromycin, quelamycin, rodorubicin, streptonigrin, streptozocin, tubercidin, ubenimex, zinostatin, zorubicin; anti-metabolites such as methotrexate and 5-fluorouracil (5-FU); folic acid analogues such as denopterin, methotrexate, pteropterin, trimetrexate; purine analogs such as fludarabine, 6-mercaptopurine, thiamiprine, thioguanine; pyrimidine analogs such as ancitabine, azacitidine, 6-azauridine, carmofur, cytarabine, dideoxyuridine, doxifluridine,

enocitabine, floxuridine; androgens such as calusterone, dromostanolone propionate, epitiostanol, mepitiothane, testolactone; anti- adrenals such as aminoglutethimide, mitotane, trilostane; folic acid replenisher such as frolinic acid; aceglatone; aldophosphamide glycoside; aminolevulinic acid; eniluracil; amsacrine; bestabucil; bisantrene; edatraxate; defofamine; demecolcine; diaziquone; elfornithine; elliptinium acetate; an epothilone; etoglucid; gallium nitrate; hydroxyurea; lentinan; lonidainine; maytansinoids such as maytansine and ansamitocins; mitoguazone; mitoxantrone; mopidanmol; nitraerine; pentostatin; phenamet; pirarubicin; losoxantrone; 2-ethylhydrazide; procarbazine; PSK7 polysaccharide complex (JHS Natural Products, Eugene, OR); razoxane; rhizoxin; sizofiran; spirogermanium; tenuazonic acid; triaziquone; 2,2',2''-trichlorotriethylamine; trichothecenes (especially T-2 toxin, verracurin A, roridin A and anguidine); urethan; vindesine (ELDISINE⁷, FILDESIN⁷); dacarbazine; mannomustine; mitobronitol; mitolactol; pipobroman; gacytosine; arabinoside ("Ara-C"); thiotepa; taxoids, *e.g.*, TAXOL⁷ paclitaxel (Bristol-Myers Squibb Oncology, Princeton, N.J.), ABRAXANETM Cremophor-free, albumin-engineered nanoparticle formulation of paclitaxel (American Pharmaceutical Partners, Schaumburg, Illinois), and TAXOTERE⁷ doxorubicin (Rhône-Poulenc Rorer, Antony, France); chlorambucil; gemcitabine (GEMZAR⁷); 6-thioguanine; mercaptopurine; methotrexate; platinum analogs such as cisplatin and carboplatin; vinblastine (VELBAN⁷); platinum; etoposide (VP-16); ifosfamide; mitoxantrone; vincristine (ONCOVIN⁷); oxaliplatin; leucovorin; vinorelbine (NAVELBINE⁷); novantrone; edatrexate; daunomycin; aminopterin; ibandronate; topoisomerase inhibitor RFS 2000; difluoromethylornithine (DMFO); retinoids such as retinoic acid; capecitabine (XELODA⁷); pharmaceutically acceptable salts, acids or derivatives of any of the above; as well as combinations of two or more of the above such as CHOP, an abbreviation for a combined therapy of cyclophosphamide, doxorubicin, vincristine, and prednisolone, and FOLFOX, an abbreviation for a treatment regimen with oxaliplatin (ELOXATINTM) combined with 5-FU and leucovorin.

[0116] The term "cytokine" is a generic term for proteins released by one cell population which act on another cell as intercellular mediators. Examples of such cytokines are lymphokines, monokines, and traditional polypeptide hormones. Included among the cytokines are growth hormone such as human growth hormone, N-methionyl human growth hormone, and bovine growth hormone; parathyroid hormone; thyroxine; insulin; proinsulin; relaxin; prorelaxin; glycoprotein hormones such as follicle stimulating hormone (FSH), thyroid stimulating hormone (TSH), and luteinizing hormone (LH); hepatic growth factor; fibroblast

growth factor; prolactin; placental lactogen; tumor necrosis factor- α and - β ; mullerian-inhibiting substance; mouse gonadotropin-associated peptide; inhibin; activin; vascular endothelial growth factor; integrin; thrombopoietin (TPO); nerve growth factors such as NGF- β ; platelet-growth factor; transforming growth factors (TGFs) such as TGF- α and TGF- β ; insulin-like growth factor-I and -II; erythropoietin (EPO); osteoinductive factors; interferons such as interferon - α , - β , and - γ ; colony stimulating factors (CSFs) such as macrophage-CSF (M-CSF); granulocyte-macrophage-CSF (GM-CSF); and granulocyte-CSF (G-CSF); interleukins (ILs) such as IL-1, IL-1a, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; a tumor necrosis factor such as TNF- α or TNF- β ; and other polypeptide factors including LIF and kit ligand (KL). As used herein, the term cytokine includes proteins from natural sources or from recombinant cell culture and biologically active equivalents of the native sequence cytokines.

[0117] The term “package insert” is used to refer to instructions customarily included in commercial packages of therapeutic products, that contain information about the indications, usage, dosage, administration, contraindications and/or warnings concerning the use of such therapeutic products.

[0118] “Epithelia,” “epithelial” and “epithelium” refer to the cellular covering of internal and external body surfaces (cutaneous, mucous and serous), including the glands and other structures derived therefrom, *e.g.*, corneal, esophageal, epidermal, and hair follicle epithelial cells. Other exemplary epithelial tissue includes: olfactory epithelium - the pseudostratified epithelium lining the olfactory region of the nasal cavity, and containing the receptors for the sense of smell; glandular epithelium - the epithelium composed of secreting cells squamous epithelium; squamous epithelium - the epithelium comprising one or more cell layers, the most superficial of which is composed of flat, scalelike or platelike cells. Epithelium can also refer to transitional epithelium, like that which is characteristically found lining hollow organs that are subject to great mechanical change due to contraction and distention, *e.g.*, tissue which represents a transition between stratified squamous and columnar epithelium.

[0119] The “growth state” of a cell refers to the rate of proliferation of the cell and/or the state of differentiation of the cell. An “altered growth state” is a growth state characterized by an abnormal rate of proliferation, *e.g.*, a cell exhibiting an increased or decreased rate of proliferation relative to a normal cell.

[0120] The term “LY6” or “LY6 polypeptide” is used herein to refer generically to any of the mammalian homologs of the mammalian LY6 gene family. The term “LY6” may be used to describe protein or nucleic acid.

[0121] The term “overexpression” as used herein, refers to cellular gene expression levels of a tissue that is higher than the normal expression levels for that tissue. The term “underexpression” as used herein, refers to cellular gene expression levels of a tissue that is lower than the normal expression levels for that tissue. In either case, the higher or lower expression is significantly different from normal expression under controlled conditions of the study.

[0122] A “control” includes a sample obtained for use in determining base-line or normal expression or activity in a mammal that is not experiencing IBD. Accordingly, a control sample may be obtained by a number of means including from tissue or cells not affected by inflammation and/or IBD, UC or CD (as determined by standard techniques); non-IBD cells or tissue e.g., from cells of a subject not experiencing IBD; from subjects not having an IBD, Crohn’s disease, or ulcerative colitis disorder; from subjects not suspected of being at risk for an IBD, CD or UC; or from cells or cell lines derived from such subjects. A control also includes a previously established standard. For assays, such as mRNA assays, including microarray assays, a control may be a universal control. Such universal control refers to RNA expression information of a particular LY6 gene obtained from RNA isolated from a mixture of healthy tissues or from a mixture of cell lines derived from various tissues such as, without limitation, universal reference RNAs disclosed herein. Accordingly, any test or assay conducted according to the invention may be compared with the established standard and it may not be necessary to obtain a control sample for comparison each time.

Table 1

```

/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int  _day[26][26] = {
/*   A  B  C  D  E  F  G  H  I  J  K  L  M  N  O  P  Q  R  S  T  U  V  W  X  Y  Z */
/* A */ { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */ { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */ { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */ {-1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */ { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ { _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M},
0, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M,
/* P */ { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */ {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ { 0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

Table 1 (cont. 1)

```

/*
*/
#include <stdio.h>
#include <ctype.h>

#define MAXJMP      16      /* max jumps in a diag */
#define MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define JMPS        1024    /* max jmps in an path */
#define MX          4       /* save if there's at least MX-1 bases since last jmp */

#define DMAT         3      /* value of matching bases */
#define DMIS         0      /* penalty for mismatched bases */
#define DINS0        8      /* penalty for a gap */
#define DINS1        1      /* penalty per base */
#define PINS0        8      /* penalty for a gap */
#define PINS1        4      /* penalty per residue */

struct jmp {
    short      n[MAXJMP];    /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
};
/* limits seq to 2^16 -1 */

struct diag {
    int      score;          /* score at last jmp */
    long     offset;         /* offset of prev block */
    short    ijmp;           /* current jmp index */
    struct jmp jp;           /* list of jmps */
};

struct path {
    int      spc;             /* number of leading spaces */
    short    n[JMPS];         /* size of jmp (gap) */
    int      x[JMPS];         /* loc of jmp (last elem before gap) */
};

char        *ofile;          /* output file name */
char        *namex[2];       /* seq names: getseqs( ) */
char        *prog;           /* prog name for err msgs */
char        *seqx[2];        /* seqs: getseqs( ) */
int          dmax;            /* best diag: nw( ) */
int          dmax0;           /* final diag */
int          dna;             /* set if dna: main( ) */
int          endgaps;         /* set if penalizing end gaps */
int          gapx, gapy;      /* total gaps in seqs */
int          len0, len1;      /* seq lens */
int          ngapx, ngapy;    /* total size of gaps */
int          smax;            /* max score: nw( ) */
int          *xbm;            /* bitmap for matching */
long         offset;          /* current offset in jmp file */
struct diag *dx;              /* holds diagonals */
struct path  pp[2];           /* holds path for seqs */

char        *calloc( ), *malloc( ), *index( ), *strcpy( );
char        *getseq( ), *g_calloc( );

```

Table 1 (cont. 2)

```

/* Needleman-Wunsch alignment program
*
* usage: progs file1 file2
* where file1 and file2 are two dna or two protein sequences.
* The sequences can be in upper- or lower-case and may contain ambiguity
* Any lines beginning with ';', '>' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
* Output is in the file "align.out"
*
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
#include "nw.h"
#include "day.h"

static _dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static _pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)
    int    ac;
    char    *av[];
{
    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0;                /* 1 to penalize endgaps */
    ofile = "align.out";        /* output file */

    nw();                       /* fill in the matrix, get the possible jumps */
    readjumps();                /* get the actual jumps */
    print();                    /* print stats, alignment */

    cleanup(0);                 /* unlink any tmp files */
}

```

Table 1 (cont. 3)

```

/* do the alignment, return best score: main( )
* dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
* pro: PAM 250 values
* When scores are equal, we prefer mismatches to any gap, prefer
* a new gap to extending an ongoing gap, and prefer a gap in seqx
* to a gap in seq y.
*/
nw( )
{
    char      *px, *py;          /* seqs and ptrs */
    int        *ndely, *dely;    /* keep track of dely */
    int        ndelx, delx;      /* keep track of delx */
    int        *tmp;            /* for swapping row0, row1 */
    int        mis;              /* score for each type */
    int        ins0, ins1;       /* insertion penalties */
    register   id;               /* diagonal index */
    register   ij;               /* jmp index */
    register   *col0, *col1;     /* score for curr, last row */
    register   xx, yy;           /* index into seqs */

    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;          /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

    /* fill in match matrix
    */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
        */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
}

```

Table 1 (cont. 4)

...nw

```

for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
    mis = col0[yy-1];
    if (dna)
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
    else
        mis += _day[*px-'A'][*py-'A'];

    /* update penalty for del in x seq;
     * favor new del over ongong del
     * ignore MAXGAP if weighting endgaps
     */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (col0[yy] - ins0 >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] -= ins1;
            ndely[yy]++;
        }
    } else {
        if (col0[yy] - (ins0+ins1) >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
     * favor new del over ongong del
     */
    if (endgaps || ndelx < MAXGAP) {
        if (col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0+ins1) >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score; we're favoring
     * mis over any del and delx over dely
     */

```

Table 1 (cont. 5)

```

id = xx - yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
    coll[yy] = mis;
else if (delx >= dely[yy]) {
    coll[yy] = delx;
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writejumps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
        dx[id].jp.n[ij] = ndelx;
        dx[id].jp.x[ij] = xx;
        dx[id].score = delx;
    }
    else {
        coll[yy] = dely[yy];
        ij = dx[id].ijmp;
        if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
            dx[id].ijmp++;
            if (++ij >= MAXJMP) {
                writejumps(id);
                ij = dx[id].ijmp = 0;
                dx[id].offset = offset;
                offset += sizeof(struct jmp) + sizeof(offset);
            }
            dx[id].jp.n[ij] = -ndely[yy];
            dx[id].jp.x[ij] = xx;
            dx[id].score = dely[yy];
        }
        if (xx == len0 && yy < len1) {
            /* last col
            */
            if (endgaps)
                coll[yy] -= ins0+ins1*(len1-yy);
            if (coll[yy] > smax) {
                smax = coll[yy];
                dmax = id;
            }
        }
    }
    if (endgaps && xx < len0)
        coll[yy-1] -= ins0+ins1*(len0-xx);
    if (coll[yy-1] > smax) {
        smax = coll[yy-1];
        dmax = id;
    }
    tmp = col0; col0 = coll; coll = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)coll);
}

```

...nw

Table 1 (cont. 6)

```

/*
 *
 * print( ) -- only routine visible outside this module
 *
 * static:
 * getmat( ) -- trace back best path, count matches: print( )
 * pr_align( ) -- print alignment of described in array p[]: print( )
 * dumpblock( ) -- dump a block of lines with numbers, stars: pr_align( )
 * nums( ) -- put out a number line: dumpblock( )
 * putline( ) -- put out a line (name, [num], seq, [num]): dumpblock( )
 * stars( ) -- put a line of stars: dumpblock( )
 * stripname( ) -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC      3
#define P_LINE 256      /* maximum output line */
#define P_SPC    3      /* space between name or num and seq */

extern  _day[26][26];
int     olen;           /* set output line length */
FILE    *fx;           /* output file */

print( )
{
    int     lx, ly, firstgap, lastgap;    /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align( );
}

```

Table 1 (cont. 7)

```

/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                                getmat
{
    int      lx, ly;                /* "core" (minus endgaps) */
    int      firstgap, lastgap;     /* leading trailing overlap */

    int      nm, i0, i1, siz0, siz1;
    char      outx[32];
    double    pct;
    register  n0, n1;
    register char *p0, *p1;

    /* get total matches, score
     */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
        else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
    }

    /* pct homology:
     * if penalizing endgaps, base is the shorter seq
     * else, knock off overhangs and take shorter core
     */
    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
    pct = 100.*((double)nm)/((double)lx);
    fprintf(fx, "\n");
    fprintf(fx, "<%d match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);
}

```

Table 1 (cont. 8)

```

fprintf(fx, "<gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, " (%d %s%s)",
        ngapx, (dna)? "base":"residue", (ngapx == 1)? "" : "s");
    fprintf(fx, "%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy);
    if (gapy) {
        (void) sprintf(outx, " (%d %s%s)",
            ngapy, (dna)? "base":"residue", (ngapy == 1)? "" : "s");
        fprintf(fx, "%s", outx);
    }
    if (dna)
        fprintf(fx,
            "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
            smax, DMAT, DMIS, DINS0, DINS1);
    else
        fprintf(fx,
            "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
            smax, PINS0, PINS1);
    if (endgaps)
        fprintf(fx,
            "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
            lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
    else
        fprintf(fx, "<endgaps not penalized\n");
}

static      nm;          /* matches in core -- for checking */
static      lmax;        /* lengths of stripped file names */
static      ij[2];       /* jmp index for a path */
static      nc[2];       /* number at start of current line */
static      ni[2];       /* current elem number -- for gapping */
static      siz[2];
static char  *ps[2];      /* ptr to current element */
static char  *po[2];      /* ptr to next output char slot */
static char  out[2][P_LINE]; /* output line */
static char  star[P_LINE]; /* set by stars() */
/*
 * print alignment of described in struct path pp[]
 */
static
pr_align( )
{
    int      nn;          /* char count */
    int      more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(nameex[i]);
        if (nn > lmax)
            lmax = nn;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
}

```

...getmat

pr_align

Table 1 (cont. 9)

```

for (nn = nm = 0, more = 1; more; ) {
    for (i = more = 0; i < 2; i++) {
        /*
         * do we have more of this sequence?
         */
        if (!*ps[i])
            continue;

        more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
        }
        else if (siz[i]) { /* in a gap */
            *po[i]++ = '-';
            siz[i]--;
        }
        else { /* we're putting a seq element
                */
            *po[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);
            po[i]++;
            ps[i]++;

            /*
             * are we at next gap for this seq?
             */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                 * we need to merge all gaps
                 * at this location
                 */
                siz[i] = pp[i].n[ij[i]++];
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i]++];
            }
            ni[i]++;
        }
    }
    if (++nn == olen || !more && nn) {
        dumpblock( );
        for (i = 0; i < 2; i++)
            po[i] = out[i];
        nn = 0;
    }
}

/*
 * dump a block of lines, including numbers, stars: pr_align( )
 */
static
dumpblock( )
{
    register i;
    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';
}

```

...pr_align

dumpblock

Table 1 (cont. 10)

...dumpblock

```

(void) putc('\n', fx);
for (i = 0; i < 2; i++) {
    if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
        if (i == 0)
            nums(i);
        if (i == 0 && *out[1])
            stars( );
        putline(i);
        if (i == 0 && *out[1])
            fprintf(fx, star);
        if (i == 1)
            nums(i);
    }
}

/*
 * put out a number line: dumpblock( )
 */
static
nums(ix)
int ix; /* index in out[] holding seq line */
{
    char nline[P_LINE];
    register i, j;
    register char *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
            i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock( )
 */
static
putline(ix)
int ix;
{

```

nums

putline

Table 1 (cont. 11)

```

int          i;
register char *px;

for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
    (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
    (void) putc(' ', fx);

/* these count from 1:
 * ni[] is current element (from 1)
 * nc[] is number at start of current line
 */
for (px = out[ix]; *px; px++)
    (void) putc(*px&0x7F, fx);
(void) putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock( )
 */
static
stars( )
{
    int          i;
    register char *p0, *p1, cx, *px;

    if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';

    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
                cx = '*';
                nm++;
            }
            else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                cx = '.';
            else
                cx = ' ';
        }
        else
            cx = ' ';
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';
}

```

...putline

stars

Table 1 (cont. 12)

```

/*
 * strip path or prefix from pn, return len: pr_align( )
 */
static
stripname(pn)
    char    *pn;    /* file name (may be path) */
{
    register char    *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}

```

stripname

Table 1 (cont. 13)

```

/*
 * cleanup( ) -- cleanup any tmp file
 * getseq( ) -- read in seq, set dna, len, maxlen
 * g_calloc( ) -- calloc( ) with error checkin
 * readjumps( ) -- get the good jumps, from tmp file if necessary
 * writejumps( ) -- write a filled array of jumps to a tmp file: nw( )
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";      /* tmp file for jumps */
FILE    *fj;

int      cleanup( );                      /* cleanup tmp file */
long     lseek( );

/*
 * remove any tmp file if we blow
 */
cleanup(i)                                cleanup
{
    int    i;
    if (fj)
        (void) unlink(jname);
    exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                          getseq
{
    char    *file;      /* file name */
    int      *len;      /* seq len */

    char      line[1024], *pseq;
    register char    *px, *py;
    int      natgc, tlen;
    FILE      *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc( ) failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
    Page 1 of nwsubr.c

```

Table 1 (cont. 14)

```

py = pseq + 4;
*len = tlen;
rewind(fp);

while (fgets(line, 1024, fp)) {
    if (*line == ';' || *line == '<' || *line == '>')
        continue;
    for (px = line; *px != '\n'; px++) {
        if (isupper(*px))
            *py++ = *px;
        else if (islower(*px))
            *py++ = toupper(*px);
        if (index("ATGCU", *(py-1)))
            natgc++;
    }
}
*py++ = '\0';
*py = '\0';
(void) fclose(fp);
dna = natgc > (tlen/3);
return(pseq+4);
}

char *
g_calloc(msg, nx, sz)
char *msg;          /* program, calling routine */
int nx, sz;          /* number and size of elements */
{
    char *px, *calloc( );

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
        if (*msg) {
            fprintf(stderr, "%s: g_calloc( ) failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
    }
    return(px);
}

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main( )
 */
readjmps( )
{
    int fd = -1;
    int siz, i0, i1;
    register i, j, xx;

    if (fj) {
        (void) fclose(fj);
        if ((fd = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open( ) %s\n", prog, jname);
            cleanup(1);
        }
    }
    for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
        while (1) {
            for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                ;

```

Table 1 (cont. 15)

...readjumps

```

        if (j < 0 && dx[dmax].offset && fj) {
            (void) lseek(fd, dx[dmax].offset, 0);
            (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
            (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
            dx[dmax].ijmp = MAXJMP-1;
        }
        else
            break;
    }
    if (i >= JMPS) {
        fprintf(stderr, "%s: too many gaps in alignment\n", prog);
        cleanup(1);
    }
    if (j >= 0) {
        siz = dx[dmax].jp.n[j];
        xx = dx[dmax].jp.x[j];
        dmax += siz;
        if (siz < 0) { /* gap in second seq */
            pp[1].n[i1] = -siz;
            xx += siz;
            /* id = xx - yy + len1 - 1
             */
            pp[1].x[i1] = xx - dmax + len1 - 1;
            gapy++;
            ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
            siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
            i1++;
        }
        else if (siz > 0) { /* gap in first seq */
            pp[0].n[i0] = siz;
            pp[0].x[i0] = xx;
            gapx++;
            ngapx += siz;
/* ignore MAXGAP when doing endgaps */
            siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
            i0++;
        }
    }
    else
        break;
}

/* reverse the order of jumps
 */
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}
}
nwsubr.c

```

Table 1 (cont. 16)

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw( )
 */
writejumps(ix)                                     writejumps
{
    int ix;
    char *mktemp( );
    if (!fj) {
        if (mktemp(jname) < 0) {
            fprintf(stderr, "%s: can't mktemp( ) %s\n", prog, jname);
            cleanup(1);
        }
        if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
    }
    (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}

```

Table 2

Reference	XXXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXYYYYYYY	(Length = 12 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the reference polypeptide)

= 5 divided by 15 = 33.3%

Table 3

Reference	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXYYYYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the reference polypeptide)

= 5 divided by 10 = 50%

Table 4

Reference-DNA	NNNNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the reference-DNA nucleic acid sequence)

= 6 divided by 14 = 42.9%

Table 5

Reference-DNA	NNNNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the reference-DNA nucleic acid sequence)

= 4 divided by 12 = 33.3%

Diagnostic Methods of the Invention

[0123] It is further contemplated that use of therapeutic agents for IBD may be specifically targeted to disorders where the affected tissue and/or cells exhibit increased LY6 expression relative to control. Accordingly, it is contemplated that the detection of increased LY6 expression may be used to detect IBD, such as CD or UC, in the gastrointestinal tissue of a mammal and/or to identify tissues and disorders that will particularly benefit from treatment with an IBD therapeutic agent, including a chemotherapeutic agent, useful in ameliorating IBD, UC and/or CD in a human patient.

[0124] In preferred embodiments, LY6 expression levels are detected, either by direct detection of the gene transcript or by detection of protein levels or activity. Transcripts may be detected using any of a wide range of techniques that depend primarily on hybridization or probes to the LY6 mRNA transcripts, to cDNAs synthesized therefrom, or to DNA where LY6 gene amplification is present. Well known techniques include Northern blotting, reverse-transcriptase PCR and microarray analysis of transcript levels. Methods for detecting LY6 protein levels include Western blotting, immunoprecipitation, two-dimensional polyacrylamide gel electrophoresis (2D SDS-PAGE - preferably compared against a standard wherein the position of the LY6 proteins has been determined), and mass spectroscopy. Mass spectroscopy may be coupled with a series of purification steps to allow high-throughput identification of many different protein levels in a particular sample. Mass spectroscopy and 2D SDS-PAGE can also be used to identify post-transcriptional modifications to proteins including proteolytic events, ubiquitination, phosphorylation, lipid modification, etc. LY6 activity may also be assessed by analyzing binding to substrate DNA or in vitro transcriptional activation of target promoters. Gel shift assay, DNA footprinting assays and DNA-protein crosslinking assays are all methods that may be used to assess the presence of a protein capable of binding to Gli binding sites on DNA. J Mol. Med 77(6):459-68 (1999); Cell 100(4): 423-34 (2000); Development 127(19): 4923-4301 (2000).

[0125] In certain embodiments, LY6 transcript levels are measured, and diseased or disordered tissues showing significantly elevated LY6 levels relative to control are treated with an IBD therapeutic compound. Accordingly, LY6 expression levels are a powerful diagnostic measure for determining whether a patient is experiencing IBD and whether that patient should receive an IBD therapeutic agent.

Antibody Compositions for Use in the Methods of the InventionA. Anti-LY6 Antibodies

[0126] In one embodiment, the present invention provides the use of anti-LY6 antibodies, which may find use herein as therapeutic, diagnostic and/or prognostic agents in determining the existence, severity of and/or prognosing the disease course of an inflammatory bowel disease such as UC. Exemplary antibodies that may be used for such purposes include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies. The term “antibodies” sometimes also include antigen-binding fragments. Anti-LY6 antibodies are available commercially, such as for example, from R&D Systems, Minneapolis, MN. Antibodies that bind specifically to LY6 as antigen may be obtained commercially or prepared by standard methods known in the art of antibody and protein chemistry for use in the method of the invention. Antibodies to LYPD1 are disclosed, for example in USUS7,144,990, the disclosure patent is hereby incorporated by reference in its entirety.

1. Polyclonal Antibodies

[0127] Polyclonal antibodies are preferably raised in animals by multiple subcutaneous (sc) or intraperitoneal (ip) injections of the relevant antigen and an adjuvant. It may be useful to conjugate the relevant antigen (especially when synthetic peptides are used) to a protein that is immunogenic in the species to be immunized. For example, the antigen can be conjugated to keyhole limpet hemocyanin (KLH), serum albumin, bovine thyroglobulin, or soybean trypsin inhibitor, using a bifunctional or derivatizing agent, *e.g.*, maleimidobenzoyl sulfosuccinimide ester (conjugation through cysteine residues), N-hydroxysuccinimide (through lysine residues), glutaraldehyde, succinic anhydride, SOCl_2 , or $\text{R}^1\text{N}=\text{C}=\text{NR}$, where R and R^1 are different alkyl groups.

[0128] Animals are immunized against the antigen, immunogenic conjugates, or derivatives by combining, *e.g.*, 100 μg or 5 μg of the protein or conjugate (for rabbits or mice, respectively) with 3 volumes of Freund's complete adjuvant and injecting the solution intradermally at multiple sites. One month later, the animals are boosted with 1/5 to 1/10 the original amount of peptide or conjugate in Freund's complete adjuvant by subcutaneous injection at multiple sites. Seven to 14 days later, the animals are bled and the serum is assayed for antibody titer. Animals are boosted until the titer plateaus. Conjugates also can be made in recombinant cell culture as protein fusions. Also, aggregating agents such as alum are suitably used to enhance the immune response.

2. Monoclonal Antibodies

[0129] Monoclonal antibodies may be made using the hybridoma method first described by Kohler et al., *Nature*, 256:495 (1975), or may be made by recombinant DNA methods (U.S. Patent No. 4,816,567).

[0130] In the hybridoma method, a mouse or other appropriate host animal, such as a hamster, is immunized as described above to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the protein used for immunization. Alternatively, lymphocytes may be immunized *in vitro*. After immunization, lymphocytes are isolated and then fused with a myeloma cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, *Monoclonal Antibodies: Principles and Practice*, pp. 59-103 (Academic Press, 1986)).

[0131] The hybridoma cells thus prepared are seeded and grown in a suitable culture medium which medium preferably contains one or more substances that inhibit the growth or survival of the unfused, parental myeloma cells (also referred to as fusion partner). For example, if the parental myeloma cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the selective culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine (HAT medium), which substances prevent the growth of HGPRT-deficient cells.

[0132] Preferred fusion partner myeloma cells are those that fuse efficiently, support stable high-level production of antibody by the selected antibody-producing cells, and are sensitive to a selective medium that selects against the unfused parental cells. Preferred myeloma cell lines are murine myeloma lines, such as those derived from MOPC-21 and MPC-11 mouse tumors available from the Salk Institute Cell Distribution Center, San Diego, California USA, and SP-2 and derivatives *e.g.*, X63-Ag8-653 cells available from the American Type Culture Collection, Manassas, Virginia, USA. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, *J. Immunol.*, 133:3001 (1984); and Brodeur *et al.*, *Monoclonal Antibody Production Techniques and Applications*, pp. 51-63 (Marcel Dekker, Inc., New York, 1987)).

[0133] Culture medium in which hybridoma cells are growing is assayed for production of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunosorbent assay (ELISA).

[0134] The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis described in Munson *et al.*, *Anal. Biochem.*, 107:220 (1980).

[0135] Once hybridoma cells that produce antibodies of the desired specificity, affinity, and/or activity are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods (Goding, *Monoclonal Antibodies: Principles and Practice*, pp.59-103 (Academic Press, 1986)). Suitable culture media for this purpose include, for example, D-MEM or RPMI-1640 medium. In addition, the hybridoma cells may be grown *in vivo* as ascites tumors in an animal e.g., by i.p. injection of the cells into mice.

[0136] The monoclonal antibodies secreted by the subclones are suitably separated from the culture medium, ascites fluid, or serum by conventional antibody purification procedures such as, for example, affinity chromatography (*e.g.*, using protein A or protein G-Sepharose) or ion-exchange chromatography, hydroxylapatite chromatography, gel electrophoresis, dialysis, *etc.*

[0137] DNA encoding the monoclonal antibodies is readily isolated and sequenced using conventional procedures (*e.g.*, by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as *E. coli* cells, simian COS cells, Chinese Hamster Ovary (CHO) cells, or myeloma cells that do not otherwise produce antibody protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. Review articles on recombinant expression in bacteria of DNA encoding the antibody include Skerra *et al.*, *Curr. Opinion in Immunol.*, 5:256-262 (1993) and Plückthun, *Immunol. Revs.* 130:151-188 (1992).

[0138] In a further embodiment, monoclonal antibodies or antibody fragments can be isolated from antibody phage libraries generated using the techniques described in McCafferty *et al.*, *Nature*, 348:552-554 (1990). Clackson *et al.*, *Nature*, 352:624-628 (1991) and Marks *et al.*, *J. Mol. Biol.*, 222:581-597 (1991) describe the isolation of murine and human antibodies, respectively, using phage libraries. Subsequent publications describe the production of high affinity (nM range) human antibodies by chain shuffling (Marks *et al.*, *Bio/Technology*, 10:779-783 (1992)), as well as combinatorial infection and *in vivo* recombination as a strategy for constructing very large phage libraries (Waterhouse *et al.*, *Nuc. Acids. Res.* 21:2265-2266 (1993)). Thus, these

techniques are viable alternatives to traditional monoclonal antibody hybridoma techniques for isolation of monoclonal antibodies.

[0139] The DNA that encodes the antibody may be modified to produce chimeric or fusion antibody polypeptides, for example, by substituting human heavy chain and light chain constant domain (C_H and C_L) sequences for the homologous murine sequences (U.S. Patent No. 4,816,567; and Morrison, *et al.*, Proc. Natl Acad. Sci. USA, 81:6851 (1984)), or by fusing the immunoglobulin coding sequence with all or part of the coding sequence for a non-immunoglobulin polypeptide (heterologous polypeptide). The non-immunoglobulin polypeptide sequences can substitute for the constant domains of an antibody, or they are substituted for the variable domains of one antigen-combining site of an antibody to create a chimeric bivalent antibody comprising one antigen-combining site having specificity for an antigen and another antigen-combining site having specificity for a different antigen.

3. Human and Humanized Antibodies

[0140] The anti- LY6 antibodies useful in the practice of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (*e.g.*, murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones *et al.*, Nature, 321:522-525 (1986); Riechmann *et al.*, Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

[0141] Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers (Jones *et al.*, Nature, 321:522-525 (1986); Riechmann *et al.*, Nature, 332:323-327 (1988); Verhoeyen *et al.*, Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

[0142] The choice of human variable domains, both light and heavy, to be used in making the humanized antibodies is very important to reduce antigenicity and HAMA response (human anti-mouse antibody) when the antibody is intended for human therapeutic use. According to the so-called "best-fit" method, the sequence of the variable domain of a rodent antibody is screened against the entire library of known human variable domain sequences. The human V domain sequence which is closest to that of the rodent is identified and the human framework region (FR) within it accepted for the humanized antibody (Sims *et al.*, J. Immunol. 151:2296 (1993); Chothia *et al.*, J. Mol. Biol., 196:901 (1987)). Another method uses a particular framework region derived from the consensus sequence of all human antibodies of a particular subgroup of light or heavy chains. The same framework may be used for several different humanized antibodies (Carter *et al.*, Proc. Natl. Acad. Sci. USA, 89:4285 (1992); Presta *et al.*, J. Immunol. 151:2623 (1993)).

[0143] It is further important that antibodies be humanized with retention of high binding affinity for the antigen and other favorable biological properties. To achieve this goal, according to a preferred method, humanized antibodies are prepared by a process of analysis of the parental sequences and various conceptual humanized products using three-dimensional models of the parental and humanized sequences. Three-dimensional immunoglobulin models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable

three-dimensional conformational structures of selected candidate immunoglobulin sequences. Inspection of these displays permits analysis of the likely role of the residues in the functioning of the candidate immunoglobulin sequence, *i.e.*, the analysis of residues that influence the ability of the candidate immunoglobulin to bind its antigen. In this way, FR residues can be selected and combined from the recipient and import sequences so that the desired antibody characteristic, such as increased affinity for the target antigen(s), is achieved. In general, the hypervariable region residues are directly and most substantially involved in influencing antigen binding.

[0144] Various forms of a humanized anti-LY6 antibody antibodies are contemplated. For example, the humanized antibody may be an antibody fragment, such as a Fab, which is optionally conjugated with one or more cytotoxic agent(s) in order to generate an immunoconjugate. Alternatively, the humanized antibody may be an intact antibody, such as an intact IgG1 antibody.

[0145] As an alternative to humanization, human antibodies can be generated. For example, it is now possible to produce transgenic animals (*e.g.*, mice) that are capable, upon immunization, of producing a full repertoire of human antibodies in the absence of endogenous immunoglobulin production. For example, it has been described that the homozygous deletion of the antibody heavy-chain joining region (J_H) gene in chimeric and germ-line mutant mice results in complete inhibition of endogenous antibody production. Transfer of the human germ-line immunoglobulin gene array into such germ-line mutant mice will result in the production of human antibodies upon antigen challenge. See, *e.g.*, Jakobovits *et al.*, Proc. Natl. Acad. Sci. USA, 90:2551 (1993); Jakobovits *et al.*, Nature, 362:255-258 (1993); Bruggemann *et al.*, Year in Immuno. 7:33 (1993); U.S. Patent Nos. 5,545,806, 5,569,825, 5,591,669 (all of GenPharm); 5,545,807; and WO 97/17852.

[0146] Alternatively, phage display technology (McCafferty *et al.*, Nature 348:552-553 (1990)) can be used to produce human antibodies and antibody fragments *in vitro*, from immunoglobulin variable (V) domain gene repertoires from unimmunized donors. According to this technique, antibody V domain genes are cloned in-frame into either a major or minor coat protein gene of a filamentous bacteriophage, such as M13 or fd, and displayed as functional antibody fragments on the surface of the phage particle. Because the filamentous particle contains a single-stranded DNA copy of the phage genome, selections based on the functional properties of the antibody also result in selection of the gene encoding the antibody exhibiting those properties. Thus, the phage

mimics some of the properties of the B-cell. Phage display can be performed in a variety of formats, reviewed in, *e.g.*, Johnson, Kevin S. and Chiswell, David J., *Current Opinion in Structural Biology* 3:564-571 (1993). Several sources of V-gene segments can be used for phage display. Clackson *et al.*, *Nature*, 352:624-628 (1991) isolated a diverse array of anti-oxazolone antibodies from a small random combinatorial library of V genes derived from the spleens of immunized mice. A repertoire of V genes from unimmunized human donors can be constructed and antibodies to a diverse array of antigens (including self-antigens) can be isolated essentially following the techniques described by Marks *et al.*, *J. Mol. Biol.* 222:581-597 (1991), or Griffith *et al.*, *EMBO J.* 12:725-734 (1993). See, also, U.S. Patent Nos. 5,565,332 and 5,573,905.

[0147] As discussed above, human antibodies may also be generated by *in vitro* activated B cells (see U.S. Patents 5,567,610 and 5,229,275).

4. Antibody fragments

[0148] In certain circumstances there are advantages of using antibody fragments, rather than whole antibodies. The smaller size of the fragments allows for rapid clearance, while retaining similar antigen binding specificity of the corresponding full length molecule, and may lead to improved access to solid tumors.

[0149] Various techniques have been developed for the production of antibody fragments. Traditionally, these fragments were derived via proteolytic digestion of intact antibodies (see, *e.g.*, Morimoto *et al.*, *Journal of Biochemical and Biophysical Methods* 24:107-117 (1992); and Brennan *et al.*, *Science*, 229:81 (1985)). However, these fragments can now be produced directly by recombinant host cells. Fab, Fv and scFv antibody fragments can all be expressed in and secreted from *E. coli*, thus allowing the facile production of large amounts of these fragments. Antibody fragments can be isolated from the antibody phage libraries discussed above. Alternatively, Fab'-SH fragments can be directly recovered from *E. coli* and chemically coupled to form F(ab')₂ fragments (Carter *et al.*, *Bio/Technology* 10:163-167 (1992)). According to another approach, F(ab')₂ fragments can be isolated directly from recombinant host cell culture. Fab and F(ab')₂ fragment with increased *in vivo* half-life comprising a salvage receptor binding epitope residues are described in U.S. Patent No. 5,869,046. Other techniques for the production of antibody fragments will be apparent to the skilled practitioner. In other embodiments, the antibody of choice is a single chain Fv fragment (scFv). See WO 93/16185; U.S. Patent No. 5,571,894; and U.S. Patent No. 5,587,458. Fv and sFv are the only species with intact combining sites that are devoid of constant regions; thus, they are

suitable for reduced nonspecific binding during in vivo use. sFv fusion proteins may be constructed to yield fusion of an effector protein at either the amino or the carboxy terminus of an sFv. See *Antibody Engineering*, ed. Borrebaeck, supra. The antibody fragment may also be a "linear antibody", e.g., as described in U.S. Patent 5,641,870 for example. Such linear antibody fragments may be monospecific or bispecific.

5. Bispecific Antibodies

[0150] Bispecific antibodies are antibodies that have binding specificities for at least two different epitopes. Exemplary bispecific antibodies may bind separate antigens or bind to two different epitopes of a particular LY6 polypeptide described herein. Other such antibodies may combine the above LY6 binding site with a binding site for another protein. Where the bispecific antibody is useful in the diagnostic method of the invention, the second antibody arm may bind a detectable polypeptide. Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g., F(ab')₂ bispecific antibodies).

[0151] Methods for making bispecific antibodies are known in the art. Traditional production of full length bispecific antibodies is based on the co-expression of two immunoglobulin heavy chain-light chain pairs, where the two chains have different specificities (Millstein *et al.*, Nature 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of 10 different antibody molecules, of which only one has the correct bispecific structure. Purification of the correct molecule, which is usually done by affinity chromatography steps, is rather cumbersome, and the product yields are low. Similar procedures are disclosed in WO 93/08829, and in Traunecker *et al.*, EMBO J. 10:3655-3659 (1991).

[0152] According to a different approach, antibody variable domains with the desired binding specificities (antibody-antigen combining sites) are fused to immunoglobulin constant domain sequences. Preferably, the fusion is with an Ig heavy chain constant domain, comprising at least part of the hinge, C_H2, and C_H3 regions. It is preferred to have the first heavy-chain constant region (C_H1) containing the site necessary for light chain bonding, present in at least one of the fusions. DNAs encoding the immunoglobulin heavy chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host cell. This provides for greater flexibility in adjusting the mutual proportions of the three polypeptide fragments in embodiments when unequal ratios of the three polypeptide

chains used in the construction provide the optimum yield of the desired bispecific antibody. It is, however, possible to insert the coding sequences for two or all three polypeptide chains into a single expression vector when the expression of at least two polypeptide chains in equal ratios results in high yields or when the ratios have no significant affect on the yield of the desired chain combination.

[0153] In a preferred embodiment of this approach, the bispecific antibodies are composed of a hybrid immunoglobulin heavy chain with a first binding specificity in one arm, and a hybrid immunoglobulin heavy chain-light chain pair (providing a second binding specificity) in the other arm. It was found that this asymmetric structure facilitates the separation of the desired bispecific compound from unwanted immunoglobulin chain combinations, as the presence of an immunoglobulin light chain in only one half of the bispecific molecule provides for a facile way of separation. This approach is disclosed in WO 94/04690. For further details of generating bispecific antibodies see, for example, Suresh *et al.*, Methods in Enzymology 121:210 (1986).

[0154] According to another approach described in U.S. Patent No. 5,731,168, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the C_H3 domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (*e.g.*, tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (*e.g.*, alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

[0155] Bispecific antibodies include cross-linked or "heteroconjugate" antibodies. For example, one of the antibodies in the heteroconjugate can be coupled to avidin, the other to biotin. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360, WO 92/200373, and EP 03089). Heteroconjugate antibodies also find use in the present method of the invention by providing multiple (either different or the same) detectable markers on each antibody for improved assay detection. Heteroconjugate antibodies may be made using any convenient cross-linking methods. Suitable cross-linking agents are well known in the art, and are disclosed in U.S. Patent No. 4,676,980, along with a number of cross-linking techniques.

[0156] Techniques for generating bispecific antibodies from antibody fragments have also been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan *et al.*, Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent, sodium arsenite, to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

[0157] Recent progress has facilitated the direct recovery of Fab'-SH fragments from *E. coli*, which can be chemically coupled to form bispecific antibodies. Shalaby *et al.*, J. Exp. Med. 175: 217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

[0158] Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny *et al.*, J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger *et al.*, Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a V_H connected to a V_L by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by

the use of single-chain Fv (sFv) dimers has also been reported. See Gruber et al., J. Immunol., 152:5368 (1994).

[0159] Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).

6. Multivalent Antibodies

[0160] A multivalent antibody may be internalized (and/or catabolized) faster than a bivalent antibody by a cell expressing an antigen to which the antibodies bind. The antibodies of the present invention can be multivalent antibodies (which are other than of the IgM class) with three or more antigen binding sites (*e.g.* tetravalent antibodies), which can be readily produced by recombinant expression of nucleic acid encoding the polypeptide chains of the antibody. The multivalent antibody can comprise a dimerization domain and three or more antigen binding sites. The preferred dimerization domain comprises (or consists of) an Fc region or a hinge region. In this scenario, the antibody will comprise an Fc region and three or more antigen binding sites amino-terminal to the Fc region. The preferred multivalent antibody herein comprises (or consists of) three to about eight, but preferably four, antigen binding sites. The multivalent antibody comprises at least one polypeptide chain (and preferably two polypeptide chains), wherein the polypeptide chain(s) comprise two or more variable domains. For instance, the polypeptide chain(s) may comprise VD1-(X1)_n-VD2-(X2)_n-Fc, wherein VD1 is a first variable domain, VD2 is a second variable domain, Fc is one polypeptide chain of an Fc region, X1 and X2 represent an amino acid or polypeptide, and n is 0 or 1. For instance, the polypeptide chain(s) may comprise: VH-CH1-flexible linker-VH-CH1-Fc region chain; or VH-CH1-VH-CH1-Fc region chain. The multivalent antibody herein preferably further comprises at least two (and preferably four) light chain variable domain polypeptides. The multivalent antibody herein may, for instance, comprise from about two to about eight light chain variable domain polypeptides. The light chain variable domain polypeptides contemplated here comprise a light chain variable domain and, optionally, further comprise a CL domain.

7. Effector Function Engineering

[0161] It may be desirable to modify the antibody of the invention with respect to effector function, *e.g.*, so as to enhance antigen-dependent cell-mediated cytotoxicity (ADCC) and/or complement dependent cytotoxicity (CDC) of the antibody. This may be achieved by introducing one or more amino acid substitutions in an Fc region of the antibody. Alternatively or additionally, cysteine residue(s) may be

introduced in the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med. 176:1191-1195 (1992) and Shopes, B. J. Immunol. 148:2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff et al., Cancer Research 53:2560-2565 (1993). Alternatively, an antibody can be engineered which has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design 3:219-230 (1989). To increase the serum half life of the antibody, one may incorporate a salvage receptor binding epitope into the antibody (especially an antibody fragment) as described in U.S. Patent 5,739,277, for example. As used herein, the term "salvage receptor binding epitope" refers to an epitope of the Fc region of an IgG molecule (*e.g.*, IgG₁, IgG₂, IgG₃, or IgG₄) that is responsible for increasing the *in vivo* serum half-life of the IgG molecule.

8. Immunoconjugates

[0162] The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, a growth inhibitory agent, a toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate) and/or a detectable label.

a. Chemotherapeutic agents

[0163] Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolacca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ²¹²Bi, ¹³¹I, ¹³¹In, ⁹⁰Y, and ¹⁸⁶Re. Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as

dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutarelddehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

[0164] Conjugates of an antibody and one or more small molecule toxins, such as a calicheamicin, maytansinoids, a trichothene, and CC1065, and the derivatives of these toxins that have toxin activity, are also contemplated herein.

B. LY6 Binding Oligopeptides

[0165] LY6 binding oligopeptides of the present invention are oligopeptides that bind, preferably specifically, to a LY6 polypeptide as described herein. LY6 binding oligopeptides may be chemically synthesized using known oligopeptide synthesis methodology or may be prepared and purified using recombinant technology. LY6 binding oligopeptides are usually at least about 5 amino acids in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length or more, wherein such oligopeptides that are capable of binding, preferably specifically, to a LY6 polypeptide as described herein. LY6 binding oligopeptides may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening oligopeptide libraries for oligopeptides that are capable of specifically binding to a polypeptide target are well known in the art (see, *e.g.*, U.S. Patent Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publication Nos. WO 84/03506 and WO84/03564; Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 81:3998-4002 (1984); Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 82:178-182 (1985); Geysen et al., in Synthetic Peptides as Antigens, 130-149 (1986); Geysen et al., J. Immunol. Meth., 102:259-274 (1987); Schoofs et al., J. Immunol., 140:611-616 (1988), Cwirla, S. E. et al. (1990) Proc. Natl. Acad. Sci. USA, 87:6378; Lowman, H.B. et al. (1991) Biochemistry, 30:10832;

Clackson, T. et al. (1991) *Nature*, 352: 624; Marks, J. D. et al. (1991), *J. Mol. Biol.*, 222:581; Kang, A.S. et al. (1991) *Proc. Natl. Acad. Sci. USA*, 88:8363, and Smith, G. P. (1991) *Current Opin. Biotechnol.*, 2:668).

[0166] In this regard, bacteriophage (phage) display is one well known technique which allows one to screen large oligopeptide libraries to identify member(s) of those libraries which are capable of specifically binding to a polypeptide target. Phage display is a technique by which variant polypeptides are displayed as fusion proteins to the coat protein on the surface of bacteriophage particles (Scott, J.K. and Smith, G. P. (1990) *Science* 249: 386). The utility of phage display lies in the fact that large libraries of selectively randomized protein variants (or randomly cloned cDNAs) can be rapidly and efficiently sorted for those sequences that bind to a target molecule with high affinity. Display of peptide (Cwirla, S. E. et al. (1990) *Proc. Natl. Acad. Sci. USA*, 87:6378) or protein (Lowman, H.B. et al. (1991) *Biochemistry*, 30:10832; Clackson, T. et al. (1991) *Nature*, 352: 624; Marks, J. D. et al. (1991), *J. Mol. Biol.*, 222:581; Kang, A.S. et al. (1991) *Proc. Natl. Acad. Sci. USA*, 88:8363) libraries on phage have been used for screening millions of polypeptides or oligopeptides for ones with specific binding properties (Smith, G. P. (1991) *Current Opin. Biotechnol.*, 2:668). Sorting phage libraries of random mutants requires a strategy for constructing and propagating a large number of variants, a procedure for affinity purification using the target receptor, and a means of evaluating the results of binding enrichments. U.S. Patent Nos. 5,223,409, 5,403,484, 5,571,689, and 5,663,143.

[0167] Although most phage display methods have used filamentous phage, lambdoid phage display systems (WO 95/34683; U.S. 5,627,024), T4 phage display systems (Ren et al., *Gene*, 215: 439 (1998); Zhu et al., *Cancer Research*, 58(15): 3209-3214 (1998); Jiang et al., *Infection & Immunity*, 65(11): 4770-4777 (1997); Ren et al., *Gene*, 195(2):303-311 (1997); Ren, *Protein Sci.*, 5: 1833 (1996); Efimov et al., *Virus Genes*, 10: 173 (1995)) and T7 phage display systems (Smith and Scott, *Methods in Enzymology*, 217: 228-257 (1993); U.S. 5,766,905) are also known.

[0168] Many other improvements and variations of the basic phage display concept have now been developed. These improvements enhance the ability of display systems to screen peptide libraries for binding to selected target molecules and to display functional proteins with the potential of screening these proteins for desired properties. Combinatorial reaction devices for phage display reactions have been developed (WO 98/14277) and phage display libraries have been used to analyze and control bimolecular

interactions (WO 98/20169; WO 98/20159) and properties of constrained helical peptides (WO 98/20036). WO 97/35196 describes a method of isolating an affinity ligand in which a phage display library is contacted with one solution in which the ligand will bind to a target molecule and a second solution in which the affinity ligand will not bind to the target molecule, to selectively isolate binding ligands. WO 97/46251 describes a method of biopanning a random phage display library with an affinity purified antibody and then isolating binding phage, followed by a micropanning process using microplate wells to isolate high affinity binding phage. The use of *Staphylococcus aureus* protein A as an affinity tag has also been reported (Li *et al.* (1998) Mol Biotech., 9:187). WO 97/47314 describes the use of substrate subtraction libraries to distinguish enzyme specificities using a combinatorial library which may be a phage display library. A method for selecting enzymes suitable for use in detergents using phage display is described in WO 97/09446. Additional methods of selecting specific binding proteins are described in U.S. Patent Nos. 5,498,538, 5,432,018, and WO 98/15833.

[0169] Methods of generating peptide libraries and screening these libraries are also disclosed in U.S. Patent Nos. 5,723,286, 5,432,018, 5,580,717, 5,427,908, 5,498,530, 5,770,434, 5,734,018, 5,698,426, 5,763,192, and 5,723,323.

[0170] In aspect, the present invention concerns ligands for the LYPD5 polypeptide. Figure 32 demonstrates this showing untransfected COS cells (A) and COS cells transfected with GLG-1 and stained with LYPD5-Fc protein. In one embodiment, the ligand for LYPD5 is the golgi complex localized glycoprotein 1 (GLG-1) or E-selectin 1 (ESL-1) polypeptide as shown in SEQ ID NOS:18, 20, 22, or 24, encoded by the nucleic acid shown as SEQ ID NOS:17, 19, 21, or 23, respectively. In another embodiment, the polynucleotide encoding a GLG-1 polypeptide comprises at least 15, at least 25, at least, at least 50, at least 100, at least 250, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2040, at least 2090, at least 2150, at least 2200, at least 2300, at least 2400, at least 2500, at least 2600, at least 2700, at least 2800, at least 2900, at least 3000, at least 3100, at least 3200, at least 3300, at least 3400, at least 3500, at least 3600, at least 3700, or at least 3720 contiguous nucleotides of SEQ ID NOS 17, 19, 21, or 23, or the polynucleotide encoding a GLG-1 comprises SEQ ID NOS 17, 19, 21, or 23. In one embodiment, a polynucleotide that binds a polynucleotide encoding a GLG-1 (SEQ ID NOS:17, 19, 21, or 23), or fragment thereof, has at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 99% or 100% sequence identity with the GLG-1 polypeptide or fragment

thereof. In one embodiment, the GLG-1 polypeptide comprises at least 10, at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 225, at least 250, at least 275, at least 300, at least 325, at least 350, at least 400, at least 450, at least 500, at least 550, at least 600, at least 650, at least 700, at least 750, at least 800, at least 850, at least 900, at least 950, at least 1000, at least 1050, at least 1100, at least 1150, or at least 1200 contiguous amino acids of SEQ ID NOs: 18, 20, 22, or 24, or the GLG-1 polypeptide comprises SEQ ID NOs: 18, 20, 22, or 24. GLG-1 or ESL-1 is expressed on neutrophils, believe to be involved in extravasation of neutrophils into tissues, and thought to play an important role in inflammation (see Hidalgo et al. (2007) Immunity, 26(4): 477-489 incorporated herein by reference in its entirety). GLG-1 or ESL-1 has 14 cysteine rich GLG1 domains. The extracellular domain (ECD) is lengthy and as described below, variants or fragments of the GLG-1 ECD were found to have the ability to bind LYPD5.

[0171] In another embodiment, the LYPD5 ligand is a variant or fragment a GLG-1 or ESL-1 molecule described herein. As shown in Figure 33A-B, GLG-1 or ESL-1 may be viewed as fragments 1, 2, 3, and 4 and as described in Example 11, any one of the 4 fragments are sufficient for LYPD5 binding.

[0172] In another embodiment, the LYPD5 ligand is a variant or fragment of GLG-1 or ESL-1 that is a single GLG-1 domain. As shown in Figure 34A-B, GLG-1 is made up of multiple GLG-1 domains and as described in Example 11, single GLG-1 domains are sufficient for LYPD5 binding.

[0173] In another embodiment, the LYPD5 ligand is a variant or fragment of GLG-1 or ESL-1 that is specific for LYPD5. As shown in Figure 35A-B, GLG-1 includes domains 26-114, domain 115, and domain 150 and as described in Example 11, domain 115 binds LYPD5 but domains 26-114 does not bind LYPD5.

[0174] The present invention contemplates variants of GLG-1 in the same manner it contemplates variants for LY6 family members.

C. Polypeptide Variants

[0175] In addition to the polypeptides, antibodies and LY6 binding polypeptides described herein, it is contemplated that variants of such molecules can be prepared for use with the invention herein. Such variants can be prepared by introducing appropriate nucleotide changes into the encoding DNA, and/or by synthesis of the desired antibody or polypeptide. Those skilled in the art will appreciate that amino acid changes

may alter post-translational processes of these molecules, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

[0176] Variations in amino acid sequence can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the amino acid sequence that results in a change in the amino acid sequence as compared with the native sequence. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the amino acid sequence of interest. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the amino acid sequence of interest with homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, *i.e.*, conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

[0177] Fragments of the various polypeptides are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native antibody or protein. Such fragments which lack amino acid residues that are not essential for a desired biological activity are also useful with the disclosed methods.

[0178] The above polypeptide fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating such fragments by enzymatic digestion, *e.g.*, by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding the desired fragment fragment by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, such fragments

share at least one biological and/or immunological activity with the corresponding full length molecule.

[0179] In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened in order to identify the desired variant.

Table 6

Original Residue	Exemplary Substitutions	Preferred Substitutions
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Asp; Lys; Arg	Gln
Asp (D)	Glu; Asn	Glu
Cys (C)	Ser, Ala	Ser
Gln (Q)	Asn; Glu	Asn
Glu (E)	Asp, Gln	Asp
Gly (G)	Pro; Ala	Ala
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe; Norleucine	Leu
Leu (L)	Norleucine; Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Trp; Leu; Val; Ile; Ala; Tyr	Leu
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Val; Ser	Ser
Trp (W)	Tyr; Phe	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala; Norleucine	Leu

[0180] Substantial modifications in function or immunological identity of the LY6 polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

- (1) hydrophobic: Norleucine, Met, Ala, Val, Leu, Ile;
- (2) neutral hydrophilic: Cys, Ser, Thr; Asn; Gln
- (3) acidic: Asp, Glu;
- (4) basic: His, Lys, Arg;
- (5) residues that influence chain orientation: Gly, Pro; and
- (6) aromatic: Trp, Tyr, Phe.

[0181] Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

[0182] The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis (Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)), cassette mutagenesis (Wells et al., Gene, 34:315 (1985)), restriction selection mutagenesis (Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)) or other known techniques can be performed on the cloned DNA to produce the anti-LY6 molecule.

[0183] Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant (Cunningham and Wells, Science, 244:1081-1085 (1989)). Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions (Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)). If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

[0184] Any cysteine residue not involved in maintaining the proper conformation of the LY6 polypeptide also may be substituted, generally with serine, to improve the oxidative stability of the molecule and prevent aberrant crosslinking. Conversely, cysteine bond(s) may be added to such a molecule to improve its stability (particularly where the antibody is an antibody fragment such as an Fv fragment).

[0185] A particularly preferred type of substitutional variant involves substituting one or more hypervariable region residues of a parent antibody (*e.g.*, a humanized or human antibody). Generally, the resulting variant(s) selected for further development will have improved biological properties relative to the parent antibody from which they are generated. A convenient way for generating such substitutional variants involves affinity maturation using phage display. Briefly, several hypervariable region sites (*e.g.*, 6-7 sites) are mutated to generate all possible amino substitutions at each site. The antibody variants thus generated are displayed in a monovalent fashion from filamentous phage particles as fusions to the gene III product of M13 packaged within each particle. The phage-displayed variants are then screened for their biological activity (*e.g.*, binding affinity) as herein disclosed. In order to identify candidate hypervariable region sites for modification, alanine scanning mutagenesis can be performed to identify hypervariable region residues contributing significantly to antigen binding. Alternatively, or additionally, it may be beneficial to analyze a crystal structure of the antigen-antibody complex to identify contact points between the antibody and target polypeptide. Such contact residues and neighboring residues are candidates for substitution according to the techniques elaborated herein. Once such variants are generated, the panel of variants is subjected to screening as described herein and antibodies with superior properties in one or more relevant assays may be selected for further development.

[0186] Nucleic acid molecules encoding amino acid sequence variants of LY6 polypeptides are prepared by a variety of methods known in the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of a native sequence or an earlier prepared variant.

D. Modifications of Polypeptides

[0187] Polypeptides and/or antibodies that have been covalently modified may also be suitable for use within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of such antibodies and polypeptides with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of such antibodies and polypeptides. Derivatization with bifunctional agents is useful, for instance, for crosslinking the preceding molecules to a water-insoluble support matrix or surface for use in purification.

Commonly used crosslinking agents include, *e.g.*, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

[0188] Other modifications include deamidation of glutamyl and asparagyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains (T.E. Creighton, *Proteins: Structure and Molecular Properties*, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)), acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

[0189] Another type of covalent modification of the polypeptides or antibodies comprises altering the native glycosylation pattern of the antibody or polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the respective native sequence. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

[0190] Glycosylation of antibodies and other polypeptides is typically either N-linked or O-linked. N-linked refers to the attachment of the carbohydrate moiety to the side chain of an asparagine residue. The tripeptide sequences asparagine-X-serine and asparagine-X-threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of the carbohydrate moiety to the asparagine side chain. Thus, the presence of either of these tripeptide sequences in a polypeptide creates a potential glycosylation site. O-linked glycosylation refers to the attachment of one of the sugars N-acylgalactosamine, galactose, or xylose to a hydroxyamino acid, most commonly serine or threonine, although 5-hydroxyproline or 5-hydroxylysine may also be used.

[0191] Addition of glycosylation sites may be accomplished by altering the amino acid sequence such that it contains one or more of the above-described tripeptide sequences (for N-linked glycosylation sites). The alteration may also be made by the

addition of, or substitution by, one or more serine or threonine residues to the sequence of the original such antibody or polypeptide (for O-linked glycosylation sites). Such antibody or polypeptide sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the preceding amino acid sequences at preselected bases such that codons are generated that will translate into the desired amino acids.

[0192] Another means of increasing the number of carbohydrate moieties is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

[0193] Removal of carbohydrate moieties may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

[0194] Another type of covalent modification comprises linking to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337. The LY6 polypeptide may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization (for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively), in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules), or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, 16th edition, Oslo, A., Ed., (1980).

[0195] Modifications forming chimeric molecules results from fusions of one polypeptide to another, heterologous polypeptide or amino acid sequence are contemplated for use with the present methods.

[0196] In one embodiment, such a chimeric molecule comprises a fusion of a polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of such antibody or polypeptide. The presence of such epitope-tagged

forms of such antibodies or polypeptides can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables such antibodies or polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 (Field *et al.*, *Mol. Cell. Biol.*, 8:2159-2165 (1988)); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto (Evan *et al.*, *Molecular and Cellular Biology*, 5:3610-3616 (1985)); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky *et al.*, *Protein Engineering*, 3(6):547-553 (1990)). Other tag polypeptides include the Flag-peptide (Hopp *et al.*, *BioTechnology*, 6:1204-1210 (1988)); the KT3 epitope peptide (Martin *et al.*, *Science*, 255:192-194 (1992)); an α -tubulin epitope peptide (Skinner *et al.*, *J. Biol. Chem.*, 266:15163-15166 (1991)); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth *et al.*, *Proc. Natl. Acad. Sci. USA*, 87:6393-6397 (1990)).

[0197] In an alternative embodiment, the chimeric molecule may comprise a fusion of a polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an “immunoadhesin”), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a preceding antibody or polypeptide in the place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH₂ and CH₃, or the hinge, CH₁, CH₂ and CH₃ regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

E. Preparation of Polypeptides

[0198] The description below relates primarily to production of polypeptides by culturing cells transformed or transfected with a vector containing nucleic acid such antibodies, polypeptides and oligopeptides. The term “polypeptides” may include antibodies, polypeptides and oligopeptides. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare such antibodies, polypeptides and oligopeptides. For instance, the appropriate amino acid sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, *e.g.*, Stewart *et al.*, Solid-Phase Peptide Synthesis, W.H. Freeman Co.,

San Francisco, CA (1969); Merrifield, J. *Am. Chem. Soc.*, 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of such antibodies, polypeptides or oligopeptides may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the desired product.

1. Isolation of DNA Encoding a Polypeptide

[0199] DNA encoding a polypeptide may be obtained from a cDNA library prepared from tissue believed to possess such antibody, polypeptide or oligopeptide mRNA and to express it at a detectable level. Accordingly, DNA encoding such polypeptides can be conveniently obtained from a cDNA library prepared from human tissue, a genomic library or by known synthetic procedures (*e.g.*, automated nucleic acid synthesis).

[0200] Libraries can be screened with probes (such as oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). Alternatively, PCR methodology may be used. [Sambrook *et al.*, *supra*; Dieffenbach *et al.*, PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

[0201] Techniques for screening a cDNA library are well known in the art. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook *et al.*, *supra*.

[0202] Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across

the full-length sequence can be determined using methods known in the art and as described herein.

[0203] Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook *et al.*, supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

[0204] Host cells are transfected or transformed with expression or cloning vectors described herein for LY6 polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: A Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook *et al.*, supra.

[0205] Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, CaCl_2 , CaPO_4 , liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook *et al.*, supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw *et al.*, Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen *et al.*, J. Bact., 130:946 (1977) and Hsiao *et al.*, Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, *e.g.*, polybrene, polyornithine, may also be used. For various techniques for

transforming mammalian cells, see Keown *et al.*, Methods in Enzymology, 185:527-537 (1990) and Mansour *et al.*, Nature, 336:348-352 (1988).

[0206] Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, *e.g.*, *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, *e.g.*, *Salmonella typhimurium*, *Serratia*, *e.g.*, *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (*e.g.*, *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA* ; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan^r*; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan^r*; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, *e.g.*, PCR or other nucleic acid polymerase reactions, are suitable.

[0207] Full length antibody, antibody fragments, and antibody fusion proteins can be produced in bacteria, in particular when glycosylation and Fc effector function are not needed, such as when the therapeutic antibody is conjugated to a cytotoxic agent (*e.g.*, a toxin) and the immunoconjugate by itself shows effectiveness in tumor cell destruction. Full length antibodies have greater half life in circulation. Production in *E. coli* is faster and more cost efficient. For expression of antibody fragments and polypeptides in bacteria, see, *e.g.*, U.S. 5,648,237 (Carter *et. al.*), U.S. 5,789,199 (Joly *et al.*), and U.S.

5,840,523 (Simmons *et al.*) which describes translation initiation region (TIR) and signal sequences for optimizing expression and secretion, these patents incorporated herein by reference. After expression, the antibody is isolated from the *E. coli* cell paste in a soluble fraction and can be purified through, *e.g.*, a protein A or G column depending on the isotype. Final purification can be carried out similar to the process for purifying antibody expressed in suitable cells (*e.g.*, CHO cells).

[0208] In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for vectors encoding desired polypeptides. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Flier *et al.*, Bio/Technology, 9:968-975 (1991)) such as, *e.g.*, *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt *et al.*, J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilarum* (ATCC 36,906; Van den Berg *et al.*, Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *Yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna *et al.*, J. Basic Microbiol., 28:265-278 [1988]); *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case *et al.*, Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, *e.g.*, *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance *et al.*, Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn *et al.*, Gene, 26:205-221 [1983]; Yelton *et al.*, Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylotropic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

[0209] Suitable host cells for the expression of glycosylated polypeptide production are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells, such as cell cultures of cotton, corn, potato, soybean, petunia, tomato, and tobacco. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts

such as *Spodoptera frugiperda* (caterpillar), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (fruitfly), and *Bombyx mori* have been identified. A variety of viral strains for transfection are publicly available, e.g., the L-1 variant of *Autographa californica* NPV and the Bm-5 strain of *Bombyx mori* NPV, and such viruses may be used as the virus herein according to the present invention, particularly for transfection of *Spodoptera frugiperda* cells.

[0210] However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure. Examples of useful mammalian host cell lines are monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham *et al.*, J. Gen. Virol. 36:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells/-DHFR (CHO, Urlaub *et al.*, Proc. Natl. Acad. Sci. USA 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod. 23:243-251 (1980)); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Mather *et al.*, Annals N.Y. Acad. Sci. 383:44-68 (1982)); MRC 5 cells; FS4 cells; and a human hepatoma line (Hep G2).

[0211] Host cells are transformed with the above-described expression or cloning vectors for desired polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

3. Selection and Use of a Replicable Vector

[0212] The nucleic acid (e.g., cDNA or genomic DNA) encoding the respective LY6 polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence.

Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

[0213] The desired polypeptide may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the DNA encoding the mature sequence that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

[0214] Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

[0215] Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, *e.g.*, ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, *e.g.*, the gene encoding D-alanine racemase for *Bacilli*.

[0216] An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up nucleic acid encoding the desired protein, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub *et al.*, Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7

[Stinchcomb *et al.*, Nature, 282:39 (1979); Kingsman *et al.*, Gene, 7:141 (1979); Tschemper *et al.*, Gene, 10:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

[0217] Expression and cloning vectors usually contain a promoter operably linked to the nucleic acid sequence encoding the desired amino acid sequence, in order to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang *et al.*, Nature, 275:615 (1978); Goeddel *et al.*, Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the *tac* promoter [deBoer *et al.*, Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the desired protein sequence.

[0218] Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman *et al.*, J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess *et al.*, J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

[0219] Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

[0220] DNA Transcription in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, *e.g.*, the actin

promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

[0221] Transcription of a DNA encoding the desired polypeptide may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the coding sequence of the preceding amino acid sequences, but is preferably located at a site 5' from the promoter.

[0222] Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding the respective antibody, polypeptide or oligopeptide described in this section.

[0223] Still other methods, vectors, and host cells suitable for adaptation to the synthesis of the respective antibody, polypeptide or oligopeptide in recombinant vertebrate cell culture are described in Gething *et al.*, Nature, 293:620-625 (1981); Mantei *et al.*, Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Culturing the Host Cells

[0224] The host cells used to produce the LY6 polypeptide may be cultured in a variety of media. Commercially available media such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), (Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) are suitable for culturing the host cells. In addition, any of the media described in Ham *et al.*, Meth. Enz. 58:44 (1979), Barnes *et al.*, Anal. Biochem. 102:255 (1980), U.S. Pat. Nos. 4,767,704; 4,657,866; 4,927,762; 4,560,655; or 5,122,469; WO 90/03430; WO 87/00195; or U.S. Patent Re. 30,985 may be used as culture media for the host cells. Any of these media may be supplemented as necessary

with hormones and/or other growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleotides (such as adenosine and thymidine), antibiotics (such as GENTAMYCIN drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

5. Detecting Gene Amplification/Expression

[0225] Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

[0226] Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies suitable for the present method may be prepared against a native sequence polypeptide or oligopeptide, or against exogenous sequence fused to DNA and encoding a specific antibody epitope of such a polypeptide or oligopeptide.

6. Protein Purification

[0227] Polypeptides may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (*e.g.* Triton-X 100) or by enzymatic cleavage. Cells employed in expression of the preceding can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

[0228] It may be desirable to purify the preceding from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the desired molecules. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular antibody, polypeptide or oligopeptide produced for the claimed methods.

[0229] When using recombinant techniques, the LY6 polypeptide can be produced intracellularly, in the periplasmic space, or directly secreted into the medium. If such molecules are produced intracellularly, as a first step, the particulate debris, either host cells or lysed fragments, are removed, for example, by centrifugation or ultrafiltration. Carter *et al.*, Bio/Technology 10:163-167 (1992) describe a procedure for isolating antibodies which are secreted to the periplasmic space of *E. coli*. Briefly, cell paste is thawed in the presence of sodium acetate (pH 3.5), EDTA, and phenylmethylsulfonylfluoride (PMSF) over about 30 min. Cell debris can be removed by centrifugation. Where the antibody is secreted into the medium, supernatants from such expression systems are generally first concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. A protease inhibitor such as PMSF may be included in any of the foregoing steps to inhibit proteolysis and antibiotics may be included to prevent the growth of adventitious contaminants.

[0230] Purification can occur using, for example, hydroxylapatite chromatography, gel electrophoresis, dialysis, and affinity chromatography, with affinity chromatography being the preferred purification technique. The suitability of protein A as an affinity ligand depends on the species and isotype of any immunoglobulin Fc domain that is present in the antibody. Protein A can be used to purify antibodies that are based on human $\gamma 1$, $\gamma 2$ or $\gamma 4$ heavy chains (Lindmark *et al.*, J. Immunol. Meth. 62:1-13 (1983)). Protein G is recommended for all mouse isotypes and for human $\gamma 3$ (Guss *et al.*,

EMBO J. 5:15671575 (1986)). The matrix to which the affinity ligand is attached is most often agarose, but other matrices are available. Mechanically stable matrices such as controlled pore glass or poly(styrenedivinyl)benzene allow for faster flow rates and shorter processing times than can be achieved with agarose. Where the antibody comprises a C_H3 domain, the Bakerbond ABXJresin (J. T. Baker, Phillipsburg, NJ) is useful for purification. Other techniques for protein purification such as fractionation on an ion-exchange column, ethanol precipitation, Reverse Phase HPLC, chromatography on silica, chromatography on heparin SEPHAROSEJ chromatography on an anion or cation exchange resin (such as a polyaspartic acid column), chromatofocusing, SDS-PAGE, and ammonium sulfate precipitation are also available depending on the antibody to be recovered.

[0231] Following any preliminary purification step(s), the mixture comprising the antibody of interest and contaminants may be subjected to low pH hydrophobic interaction chromatography using an elution buffer at a pH between about 2.5-4.5, preferably performed at low salt concentrations (*e.g.*, from about 0-0.25M salt).

H. Pharmaceutical Formulations

[0232] Therapeutic formulations ("therapeutic agent") used in accordance with the present invention may be prepared for storage by mixing the therapeutic agent(s) having the desired degree of purity with optional pharmaceutically acceptable carriers, excipients or stabilizers (Remington: The Science of Practice of Pharmacy, 20th edition, Gennaro, A. et al., Ed., Philadelphia College of Pharmacy and Science (2000)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients, or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as acetate, Tris, phosphate, citrate, and other organic acids; antioxidants including ascorbic acid and methionine; preservatives (such as octadecyldimethylbenzyl ammonium chloride; hexamethonium chloride; benzalkonium chloride, benzethonium chloride; phenol, butyl or benzyl alcohol; alkyl parabens such as methyl or propyl paraben; catechol; resorcinol; cyclohexanol; 3-pentanol; and m-cresol); low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, histidine, arginine, or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; tonicifiers such as

trehalose and sodium chloride; sugars such as sucrose, mannitol, trehalose or sorbitol; surfactant such as polysorbate; salt-forming counter-ions such as sodium; metal complexes (*e.g.*, Zn-protein complexes); and/or non-ionic surfactants such as TWEEN7, PLURONICS7 or polyethylene glycol (PEG). The antibody preferably comprises the antibody at a concentration of between 5-200 mg/ml, preferably between 10-100 mg/ml.

[0233] The formulations herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. For example, in addition to the preceding therapeutic agent(s), it may be desirable to include in the formulation, an additional antibody, *e.g.*, a second such therapeutic agent, or an antibody to some other target such as a growth factor that affects the growth of the glioma. Alternatively, or additionally, the composition may further comprise a chemotherapeutic agent, cytotoxic agent, cytokine, growth inhibitory agent, anti-hormonal agent, and/or cardioprotectant. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

[0234] The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington: The Science and Practice of Pharmacy, *supra*.

[0235] Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semi-permeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT7 (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid.

[0236] The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Methods for the Diagnosis and/or Treatment of Inflammatory Bowel Disease

[0237] To determine LY6 expression in gastrointestinal tissue or cells of a mammal, such as a mammal experiencing IBD, various diagnostic assays are available. In one embodiment, LY6 polypeptide overexpression may be analyzed by RT-PCR, in-situ hybridization, microarray analysis, and/or immunohistochemistry (IHC). Fresh, frozen and/or parafin embedded tissue sections from a gastrointestinal biopsy (such as from the colon or, more specifically, the sigmoid colon) from a mammal (such as without limitation a human) may be subjected to the RT-PCR, in situ hybridization, microarray analysis and/or IHC assay.

[0238] Alternatively, or additionally, FISH assays such as the INFORM7 (sold by Ventana, Arizona) or PATHVISION7 (Vysis, Illinois) may be carried out on formalin-fixed, paraffin-embedded tissue to determine the extent (if any) of LY6 expression and/or upregulation in a tissue sample or biopsy.

[0239] LY6 expression may be evaluated using an *in vivo* diagnostic assay, *e.g.*, by administering a molecule (such as an antibody, oligopeptide or organic molecule) which binds the molecule to be detected and is tagged with a detectable label (*e.g.*, a radioactive isotope or a fluorescent label) and externally scanning the patient for localization of the label.

[0240] Currently, depending on the stage of the IBD, treatment involves one or a combination of the following therapies: surgery to remove affected bowel tissue, administration of therapeutic agents, including without limitation chemotherapy; dietary changes, and lifestyle management. Therapeutic agents or chemotherapeutic agents useful in the treatment of IBD are known in the art and representative therapeutic and chemotherapeutic agents are disclosed herein.

[0241] In particular, combination therapy with palictaxel and modified derivatives (see, *e.g.*, EP0600517) is contemplated. The preceding antibody, polypeptide, oligopeptide or organic molecule will be administered with a therapeutically effective dose of the chemotherapeutic agent. In another embodiment, such antibody, polypeptide, oligopeptide or organic molecule is administered in conjunction with chemotherapy to enhance the activity and efficacy of the chemotherapeutic agent, *e.g.*, paclitaxel. The

Physicians= Desk Reference (PDR) discloses dosages of these agents that have been used in treatment of various cancers. The dosing regimen and dosages of these aforementioned chemotherapeutic drugs that are therapeutically effective will depend on the particular cancer being treated, the extent of the disease and other factors familiar to the physician of skill in the art and can be determined by the physician.

[0242] Therapeutic agents or chemotherapeutic agents are administered to a human patient, in accord with known methods, such as intravenous administration, *e.g.*, as a bolus or by continuous infusion over a period of time, by intracranial, intracerebrospinal, intra-articular, intrathecal, intravenous, intraarterial, subcutaneous, oral, topical, or inhalation routes.

[0243] The present invention provides methods that involve a diagnostic step and a therapeutic treatment step. In one embodiment, the present invention provides methods of detecting inflammatory bowel disease (IBD) in a mammalian subject that include the steps of (1) detecting the level of expression of a nucleic acid or a gene encoding a LY6 polypeptide (a) in a test sample of tissue or cells obtained from the subject, and (b) in a control sample where a higher level of expression of the LY6 nucleic acid or gene in the test sample, as compared to the control sample, indicates the presence of an IBD in the subject from which the test sample was obtained; and (2) administering to the subject an effective amount of an IBD therapeutic agent. In one embodiment, the IBD therapeutic agent is an antagonist of another IBD-associated molecule. The present invention contemplates various IBD-associated molecules that are differentially expressed in IBD. In one embodiment, the IBD-associated molecule is a molecule that is differentially expressed in an IBD. In another embodiment, the IBD-associated molecule is over-expressed in an IBD. In yet another embodiment, the over-expressed IBD-associated molecule is an integrin. In one other embodiment, the IBD-associated molecule is integrin, beta 7 (ITGB2) (see WO 2006/026759, which is incorporated herein by reference in its entirety) The term "IBD therapeutic agent" as used herein refers to an antagonist of an IBD-associated molecule. In one embodiment, the IBD therapeutic agent is an antagonist of an integrin. In another embodiment, the IBD therapeutic agent is an antagonist of ITGB7. In yet another embodiment, the IBD therapeutic agent is an antagonist of the polypeptide shown as SEQ ID NO: 69 encoded by the nucleic acid sequence shown as SEQ ID NO: 68.

J. Articles of Manufacture and Kits

[0244] For diagnostic applications, the article of manufacture comprises a container and a label or package insert on or associated with the container indicating a use for detecting and expression of LY6 (such as, without limitation LY6, LYPD1, LYPD3, and/or LYPD5) in a gastrointestinal tissue or cell of a mammal. In one embodiment, the mammal is a human. In one embodiment, the tissue or cell is gastrointestinal tissue or cell. In one embodiment, detecting includes quantitation relative to a control sample. In an embodiment, the container, label or package insert indicates that the gastrointestinal tissue or cells are from colon of a mammal. In an embodiment, the container, label or package insert indicates that increased LY6 expression relative to a control sample is indicative of IBD, including without limitation CD and/or UC, in the mammal. Suitable containers include, for example, bottles, vials, syringes, *etc.* The containers may be formed from a variety of materials such as glass or plastic. Additionally, the article of manufacture may further comprise a second container comprising a buffer or other reagent (such as detectable label) useful for carrying out the detection. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, and dyes.

[0245] For isolation and purification of LY6 polypeptide, the kit can contain the LY6-binding reagent coupled to beads (*e.g.*, sepharose beads). Kits can be provided which contain such molecules for detection and quantitation of LY6 polypeptide *in vitro*, *e.g.*, in an ELISA or a Western blot. As with the article of manufacture, the kit comprises a container and a label or package insert on or associated with the container. The container holds a composition comprising at least one such LY6 binding antibody, oligopeptide or organic molecule useable with the invention. Additional containers may be included that contain, *e.g.*, diluents and buffers, control antibodies. The label or package insert may provide a description of the composition as well as instructions for the intended *in vitro* or diagnostic use.

K. Sense and Anti-Sense LY6-Encoding Nucleic Acids

[0246] Molecules that would be expected to bind to nucleic acids encoding an LY6 gene include sense and antisense oligonucleotides, which comprise a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target LY6 mRNA or DNA sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of the LY6 DNA or its complement. The

ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (*Cancer Res.* 48:2659, 1988) and van der Krol *et al.* (*BioTechniques* 6:958, 1988).

[0247] The sense and/or antisense oligonucleotides hybridizable to a LY6 gene are useful, for example, for detecting the presence of LY6 DNA or mRNA in a tissue or cell sample gastrointestinal tissue or cells of mammal according to the invention. The sense and/or antisense compounds used in accordance with this invention may be conveniently and routinely made through the well-known technique of solid phase synthesis. Equipment for such synthesis is sold by several vendors including, for example, Applied Biosystems (Foster City, Calif.). Any other means for such synthesis known in the art may additionally or alternatively be employed. It is well known to use similar techniques to prepare oligonucleotides such as the phosphorothioates and alkylated derivatives. The compounds of the invention may also be admixed, encapsulated, conjugated or otherwise associated with other molecules, molecule structures or mixtures of compounds, as for example, liposomes, receptor targeted molecules, oral, rectal, topical or other formulations, for assisting in uptake, distribution and/or absorption. Patents that teach the preparation of such uptake, distribution and/or absorption assisting formulations include, but are not limited to, U.S. Pat. Nos. 5,108,921; 5,354,844; 5,416,016; 5,459,127; 5,521,291; 5,543,158; 5,547,932; 5,583,020; 5,591,721; 4,426,330; 4,534,899; 5,013,556; 5,108,921; 5,213,804; 5,227,170; 5,264,221; 5,356,633; 5,395,619; 5,416,016; 5,417,978; 5,462,854; 5,469,854; 5,512,295; 5,527,528; 5,534,259; 5,543,152; 5,556,948; 5,580,575; and 5,595,756, each of which is herein incorporated by reference.

[0248] Sense and antisense oligonucleotides include without limitation primers and probes useful in PCR, RT-PCR, hybridization methods, in-situ hybridization, and the like.

[0249] Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

[0250] Antisense or sense RNA or DNA molecules are generally at least about 5 nucleotides in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term “about” means the referenced nucleotide sequence length plus or minus 10% of that referenced length.

EXAMPLES

[0251] The following nonlimiting examples are provided for illustrative purposes and are not intended to limit the scope of the invention. Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cell lines identified in the following examples, and/or throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA.

Example 1: Materials and Methods

[0252] *Reagents, cells and mice:* IFN γ , TNF α , and IL1 β were obtained from Peprotech™ (Rocky Hill, NJ). IFN α was obtained from Hycult Biotechnology™ (The Netherlands). For crosslinking experiments, anti-KLH control antibody, anti-LY6A (clone E13-161.7 or D7) were obtained from Pharmingen™ (San Diego, CA). Anti-LY6C (clone HK1.4) was obtained from Southern Biotech™ (Birmingham, AL).

[0253] Chronic CD45RB^{high} transfer colitis was induced as described previously in SCID mice on a Balb/c background (Powrie, F. et al., (1994) *Immunity* 1:553-562). IL10^{-/-} mice (Kuhn, R. et al., (1993) *Cell* 75:263-274) on a 129 background, which develop spontaneous colitis, were sacrificed between 11 and 13 weeks of age. Colons were snap frozen in OCT until used in experiments as described. Proximal colon, middle colon, distal colon and rectum were scored using a scale of 0-5 (0 = normal bowel, 5 = severe disease). Scores were summed to achieve a total colitis severity score for each animal.

[0254] The young adult mouse colonocyte (YAMC) cell line (provided by Robert Whitehead, Vanderbilt University Medical Center, Nashville, TN) was derived from the Immortomouse™ a transgenic animal containing a temperature-sensitive T-antigen (tsTag) under the control of an interferon- γ -dependent promoter, as previously described (Whitehead, R.H. et al, (1993) *Proc Natl Acad Sci U S A* 90:587-591). YAMC cells proliferate under permissive conditions of 32°C in the presence of 5 units/ml IFN- γ (Peprotech™, New Jersey), but no longer proliferate upon removal of IFN- γ at 37°C (nonpermissive conditions).

[0255] YAMC cells were cultured in RPMI containing 5% FBS, 2mM L-glutamine, penicillin/streptomycin, 5 U/ml IFN γ and N-2 supplement (Invitrogen™, Carlsbad, CA). Cells were cultured under non-permissive conditions for 24 hours prior to experiments, and for the duration of experimentation.

[0256] CMT93 cells were obtained from ATCC (ATCC Number® CCL-223™, ATCC, Manassas, VA) cultured in DMEM containing 10% FBS, 2mM L-glutamine, and penicillin/streptomycin.

[0257] *Laser capture microscopy and RNA purification:* 10-12 μ m sections were applied to LCM membrane slides (Molecular Machines™, Glattbrugg, Switzerland). Slides were subjected to an abbreviated H&E stain (total time of about five minutes) before crypt epithelial cells were histologically identified and dissected using an MMI Cellcut™ microscope (Molecular Machines, Glattbrugg, Switzerland). RNA was purified from the dissected cells using the Arcturus™ Picopure™ RNA purification kit and manufacturer's protocols (Arcturus™, Sunnyvale, CA) and quantified using the NanoDrop ND-1000™ Spectrophotometer (NanoDrop Technologies™, Wilmington, DE).

[0258] *Microarray hybridization and data analysis:* The quantity and quality of input total RNA samples was determined using ND-1000 spectrophotometer (NanoDrop™ Technologies, Montchanin, DE) and Bioanalyzer 2100™ (Agilent™ Technologies, Palo Alto, CA), respectively. The method for preparation of Cy-dye labeled cRNA and array hybridization was provided by Agilent™ Technologies (Palo Alto, CA). Briefly, total RNA sample was converted to double-stranded cDNA and then to labeled cRNA using a Low RNA Input Fluorescent Linear Amplification™ Kit (Agilent™, Product# 5184-3523). The labeled cRNA was purified using RNeasy™ mini kit (Qiagen™, San Diego, CA) and then quantified using ND-1000™ spectrophotometer (Nanodrop™ Technologies). Cy-dye incorporation was determined by running the

labeled cRNA on a Novex™ TBE-Urea gel (Invitrogen™, Carlsbad, CA) followed by gel scanning on a Typhoon™ scanner (GE Healthcare™, Piscataway, NJ). To determine the amount of Cy-dye fluorescent counts, the gel images were analyzed using ImageQuant™ software (GE Healthcare™). Approximately 500,000 counts of Cy-dye labeled cRNA was fragmented and hybridized to the Agilent's whole mouse genome array as described in Agilent's *In situ* Hybridization kit-plus (Agilent™, Product# 5184-3568). LCM samples were labeled with Cy5 dye and hybridized against Cy3 dye labeled universal mouse reference (Stratagene™, La Jolla, CA). Following hybridization, the arrays were washed, dried with acetonitrile and scanned on the Agilent™ DNA microarray scanner. The array image files were analyzed using Agilent™'s Feature Extraction™ software 7.5 and further data analysis was performed using Resolver™ (Merck™, Seattle, WA).

[0259] Data was analyzed using Rosetta Resolver™ software (Rosetta Biosoftware™, Seattle, WA). Briefly, healthy and colitic samples were grouped separately and probes that passed two-tailed anova ($p < 0.05$) were selected. These probes were analyzed further for probes that demonstrated a two fold or greater change in colitic samples versus healthy samples.

[0260] *Real time quantitative RT-PCR:* RT-PCR was performed on extracted RNA using Taqman™ Gold™ RT-PCR kit and reagents (Applied Biosystems™, Foster City, CA). All samples were run with gene specific primers using 5'-FAM and 3'-TAMRA labeled internal probes. Analysis was performed compared to housekeeping gene, SPF31, specific primers by the $2^{-\Delta\Delta C_t}$ method as described (Livak, K.J., and T.D. Schmittgen (2001) *Methods* 25:402-408). Primers and probes were either designed using Primer3™ software (Rozen, S., and H. Skaletsky (2000) *Methods Mol Biol* 132:365-386) or obtained commercially (Applied Biosystems™). Primers and probes used for these assays were the following, shown in the 5'-3' direction:

LY6A:

Sense: CTT ACC CAT CTG CCC TCC TA	(SEQ ID NO:39)
Antisense: CCT CCA TTG GGA ACT GCT AC	(SEQ ID NO:40)
Probe: TCC TGT TGC CAG GAA GAC CTC TGC	(SEQ ID NO:41)

LY6C:

Sense: ACT TCC TGC CCA GCA GTT AC	(SEQ ID NO:42)
Antisense: GGC ACT GAC GGG TCT TTA GT	(SEQ ID NO:43)

Probe: CTG CCG CGC CTC TGA TGG AT

(SEQ ID NO:44)

[0261] *Immunofluorescent staining:* Frozen tissues were cut into 5 μ m sections and stained with biotinylated anti-LY6C (Southern Biotech™, Birmingham, AL) or anti-SCA-1 at 2.5 ng/ml (R&D Systems™, Minneapolis, MN). Slides were washed and labeled with Alexa Fluor™ 488 conjugated streptavidin, mounted using Prolong Gold™ with DAPI (Invitrogen™, Carlsbad, CA) and visualized by confocal microscopy.

[0262] *Crosslinking LY6 molecules:* The ability of crosslinked LY6 polypeptide to effect chemokine production was tested by incubating YAMC cells with plate-bound anti-LY6C or anti-KLH (control) antibodies and measuring the production of chemokines CXCL2, CXCL5, and CCL7. Because lipid raft formation in the cell membrane is required for crosslinking, chemokine production was tested under conditions of normal raft formation (non-cholesterol depletion) and under conditions of cholesterol depletion.

[0263] For crosslinking using plate-bound antibody, 100 μ l of anti-LY6C or anti-KLH (control) antibody at 5 μ g/mL concentration was added to a 96 well plate, or 2 mL were added to a 60 mm² dish and allowed to bind to the plate for 15 hours at 4 °C. YAMC cells, grown in cholesterol depleting or non-depleting conditions (as provided in Example 5, herein) were incubated with the plate-bound antibodies for 15 hours at 32°C under cholesterol non-depleting conditions RNA was collected and expression levels of CXCL2, CXCL5, and CCL7 were determined. The assay is further described and results are shown in Example 5, herein.

[0264] *siRNA inhibition:* Individual siRNA directed against murine LY6C were obtained from Dharmacon (Lafayette, CO). SiRNA was transfected into YAMC cells using lipofectamine 2000 (Invitrogen) and standard protocols. 72 hours after transfection, cells were collected to determine knockdown efficiency. One siRNA was chosen for crosslinking experiments based on superior knockdown efficiency (95% inhibition by quantitative RT-PCR).

[0265] *CXCL5 secretion:* Supernatants were collected at the indicated time point from stimulated cells and cytokine CXCL5 concentrations were determined by ELISA using a commercially available kit from R&D Systems™ and manufacturer's protocols. The level of detection was 15 pg/ml of CXCL5.

[0266] *Cholesterol depletion:* YAMC cells were cultured for 72 hours in serum free medium at 37°C in the presence of 4 µM lovastatin and 250 µM mevalonate (Sigma). Cells were plated and maintained in lovastatin and mevalonate throughout the experiment.

[0267] *Statistics:* Student's t test was used for comparison between groups (* indicates $p < 0.05$).

Example 2: Gene expression patterns of IEC are altered during colitis

[0268] Studies have indicated that gene expression patterns of IEC are significantly altered in mouse models of colitis, as well as human IBD (Fahlgren, A., et al. (2004) *Clin Exp Immunol* 137:379-385; Brand, S. et al. (2006) *Am J Physiol Gastrointest Liver Physiol* 290:G827-838; Ruiz, P.A. et al. (2005) *J Immunol* 174:2990-2999). In this example, evaluated gene expression patterns in IEC of healthy and colitic mice were examined in order to illuminate novel genes and pathways altered in IBD.

[0269] The identification of genes involved in the immunopathology of IBD was sought by evaluating intestinal epithelial cells (IEC) from the CD45RB^{Hi} T cell transfer colitis mouse model as well as the IL10^{-/-} mouse model, both of which result from Th1 dysregulation and share many features of human Crohn's disease (Elson, C.O. et al. (2005) *Immunol Rev* 206:260-276; Bouma, G., and W. Strober (2003) *Nat Rev Immunol* 3:521-533). Laser capture microdissection (LCM) was used to isolate crypt IEC from the colons of healthy and colitic mice in the two models of murine IBD. RNA was extracted from these samples and analyzed by microarray technology as described herein in Example 1. The gene expression profile of IEC of colitic mice in the transfer colitis model identified 1770 probes with > 2 fold expression changes compared to control mice, while the IL10^{-/-} model identified 1140 probes. Overlapping in both models, there were 540 probes with >2 fold changes in expression, corresponding to approximately 400 different genes (data not shown).

Example 3: Pathways and genes affected in IEC during colitis

[0270] Of the approximately 400 genes affected in both models, genes involved in antigen presentation, TLR signaling and cell migration were overrepresented (Table 7). In Table 7, numbers represent the mean with standard deviation of the fold change compared to universal standard RNA of colitic mice over healthy mice in either the IL10^{-/-} model of colitis or the CD45RB^{Hi} model of colitis, as indicated. The results

indicated that some IEC expressed genes show altered expression patterns in murine models of IBD. Many of these genes, including TLR2, CCL7, CXCL5 and ICAM-1 have been described previously as having increased epithelial expression during colitis (Breider, M.A. et al. (1997) *Vet Pathol* 34:598-604; Uguccioni, M. et al. (1999) *Am J Pathol* 155:331-336; Z'Graggen, K. et al. (1997) *Gastroenterology* 113:808-816; Singh, J.C. et al. (2005) *Am J Physiol Gastrointest Liver Physiol* 288:G514-524), suggesting that the gene expression pattern obtained in these microarrays are an accurate reflection of the biology of IEC in colitis.

Table 7

Fold change (p value):

Cell migration IL10 -/- model CD45RBhi model

CXCL1	+3.89 (<0.0001)	+2.09 (0.00066)
CXCL5	+21.82 (<0.0001)	+23.34 (<0.0001)
CXCL13	+3.01 (<0.0001)	+2.85 (<0.0001)
CCL6	-3.47 (<0.0001)	-2.5 (<0.0001)
CCL7	+4.2 (<0.0001)	+5.54 (0.00026)
CCL11	-3.43 (<0.0001)	-3.6 (0.00607)

TLR signaling

TLR2	+2.15 (<0.0001)	+2.68 (<0.0001)
Fos	+3.64 (<0.0001)	+2.03 (<0.0001)
LBP	+2.34 (<0.0001)	+2.57(<0.0001)
NFKBIA	+2.37 (<0.0001)	+2.15(<0.0001)

Antigen

presentation

H2-D1	+2.77 (<0.0001)	+2.23 (<0.0001)
HLA-A	+2.83(<0.0001)	+2.40(<0.0001)
HLA-B	+2.71(<0.0001)	+2.44(<0.0001)
HLA-E	+2.31(<0.0001)	+2.34(<0.0001)
ICAM-1	+2.51(<0.0001)	+2.587(<0.0001)
PSMB8	+8.10(<0.0001)	+3.09(<0.0001)
PSMB9	+6.61(<0.0001)	+2.72(<0.0001)
TAP1	+4.05(<0.0001)	+4.10(<0.0001)
TAP2	+2.08(<0.0001)	+2.18(<0.0001)

[0271] IEC can function as non-professional APC (Snoeck, V. et al., (2005) *Microbes Infect* 7:997-1004; and Shao, L et al., (2005) *Immunol Rev* 206:160-176), and the gene expression pattern obtained in these microarrays indicate that these functions are enhanced during colitis by upregulation in genes associated with antigen processing, such

as LMP7 and TAP1, as well as MHC class I and II genes which would serve to enhance presentation of antigens on the surface of the IEC.

[0272] The microarray data supports the concept that colitic IEC attract immune cells to the colon through altered chemokine expression, and may present antigen to infiltrating T cells by upregulating expression of genes associated with antigen presentation.

Example 4: Expression of LY6 family members is strongly upregulated on the surface of colitic IEC

[0273] Members of the mouse LY6 family of molecules were overrepresented in number as well as degree of upregulation in both the transfer colitis mouse model and the IL10^{-/-} mouse model (Figures 23A and 23B). These results were confirmed by real-time quantitative RT-PCR of pooled and amplified IEC RNA in the transfer colitis model (data not shown). Expression of the LY6 family members was unique to the disease state, so no healthy mice expressed appreciable levels of any of these LY6 family members.

[0274] While expression of murine LY6 molecules on the surface of cells of hematopoietic origin is known, expression on IEC has not been previously described (Bamezai, A. (2004) *Arch Immunol Ther Exp (Warsz)* 52:255-266; and Rock, K.L. et al. (1989) *Immunol Rev* 111:195-224). Expression of murine LY6A and LY6C is detectable on many non-epithelial cells present within the colon, such as T cells and granulocytes. Immunofluorescent staining was performed for both murine LY6A and LY6C on healthy and colitic colons. Levels of murine LY6A and LY6C were minimal or absent on the surface of healthy IEC (Figure 24A and 24C, respectively). Expression of both murine LY6A and LY6C was detectable on the surface of IEC throughout the colons of colitic mice (Figure 24B and 24D, respectively). There was no evidence of polarization of either LY6A or LY6C, and staining was present on both the apical and basolateral membranes, making LY6 molecules potentially accessible to ligands on either surface. These results indicate that the microarray analysis results showing upregulation of murine LY6A and LY6C in murine colitic models was not due to the influx of contaminating immune cells.

Example 4: Transcription of LY6 genes is stimulated by inflammatory cytokines

[0275] LY6 expression on T cells is induced and enhanced by both type I and type II IFNs (Khodadoust, M.M., K.D. Khan, and A.L. Bothwell. 1999. Complex regulation of Ly-6E gene transcription in T cells by IFNs. *J Immunol* 163:811-819).

Furthermore, expression of a number of cytokines, is elevated in the colon during active colitis (Niessner, M., and B.A. Volk. 1995. Altered Th1/Th2 cytokine profiles in the intestinal mucosa of patients with inflammatory bowel disease as assessed by quantitative reversed transcribed polymerase chain reaction (RT-PCR). *Clin Exp Immunol* 101:428-435).

[0276] To determine if cytokines present during colitis affect transcription of LY6 family members in IEC, we treated YAMC cells, a conditionally immortalized murine IEC line, with IL-1 β , IFN α , TNF α , IFN γ or the combination of TNF α and IFN γ and analyzed the transcription of all identified murine LY6 genes by real-time quantitative RT-PCR (Table 8). Briefly, mRNA levels of the indicated LY6 family member in IEC was determined by real time quantitative RT-PCR after 15 hours of treatment with the indicated cytokine. Number represents the fold change (determined by $2^{-\Delta\Delta C_t}$ method) versus the untreated, media control. *, P<0.05 versus media control. †, p<0.05 versus IFN γ treated cells. The following LY6 family members were tested, but not detected in samples, regardless of treatment: LY6K, Lypd3, Lypd4, Lypd5, LY6g5b, Ly6g6d, Ly6g6e, Slurp1. The results indicate that IEC upregulate LY6 family members in response to inflammatory cytokines.

Table 8

	Media	IL1 β	TNF α	IFN α	IFN γ	IFN γ & TNF α
Ly6A	1.0	1.8*	2.2*	2.8*	33.1*	65.4*†
Ly6C	1.0	1.6*	1.2*	2.4*	65.6*	63.6*
Ly6D	1.0	2.7*	2.1*	1.5*	1.0	0.9
Ly6E	1.0	1.4*	1.5*	2.1*	1.9*	2.9*†
Ly6F	1.0	2.5*	0.6*	7.1*	108.2*	169.7*†
Ly6H	1.0	1.0	1.1	1.2	3.7*	1.4†
Lypd1	1.0	1.5*	2.1*	1.3*	1.3*	2.9*†
Lypd2	1.0	0.1*	ND	0.4*	0.1*	ND
Ly6g5c	1.1	1.3	0.8	0.9	1.3	1.1
Ly6g6c	1.0	0.7	0.7	0.6	0.6*	0.3*†
Slurp2/Lynx1	1.1	0.7	0.4	0.7	1.5	0.4*

[0277] While many of the LY6 family members were not detected in either the presence or absence of inflammatory cytokines, we detected a strong upregulation in the transcription of murine LY6A, LY6C and LY6F in response to the majority of the cytokines tested, as well as more moderate upregulation of murine LY6E, LY6H and LYPD1 in response to some cytokines tested. However, IFN γ was by far the most potent cytokine in inducing LY6 upregulation. Furthermore, TNF α enhanced the effects of IFN γ

on the expression of LY6A, LY6F, LY6E and LYPD1. Similar upregulation of LY6 family members were seen in another murine IEC line, CMT93 (data not shown).

[0278] To examine the surface expression of LY6 family members in response to cytokines, YAMC cells were exposed to the above cytokines and analyzed by flow cytometry for expression of murine LY6A and LY6C, for which commercial antibodies are available, as described herein in Example 1. High levels of murine LY6A was expressed on YAMC cells even in the absence of added cytokines (Figure 25B, media). Expression of murine LY6C (Figure 25A, media) was considerably lower than expression of LY6A.

[0279] IL-1 β and TNF α induced slight increases in the surface expression of both murine LY6A and LY6C, in agreement with the RNA expression (Figures 25A and 25B). A more moderate increase in expression was noted when IFN α was added to the cells, while IFN γ induced dramatic increases in surface expression of both LY6A and LY6C (Figure 25A and 25B). Surface protein expression closely mirrored RNA expression. Th2 cytokines, such as IL4, IL10 or IL13 had no effect on surface expression of either LY6A or LY6C (data not shown).

[0280] Induction of both LY6A (Figure 25D) and LY6C (Figure 25C) by IFN γ was dose dependent. Doses as low as 6.25 units/ml of IFN γ resulted in detectable increases in both LY6 molecules by flow cytometry. Furthermore, the increase in both LY6A (Figure 25F) and LY6C (Figure 25E) surface expression became evident between 2 and 4 hours after IFN γ treatment, and steadily increased for at least 24 hours after IFN γ treatment. This data indicates that relatively low concentrations of IFN γ are sufficient to increase surface expression of LY6 molecules within hours.

[0281] There is evidence that IL-22, which is secreted primarily from activated T cells, functions through the IL-22R complex, present on IEC to promote cytokine production and an inflammatory phenotype (Brand, S.F. et al. Am J Physiol Gastrointest Liver Physiol 290:G827-838 (2006)). Furthermore, IL-22 is involved in the immunopathogenesis of Crohn's Disease. To examine whether IL-22 affects LY6 molecule expression on murine IEC, YAMC cells were cultured in the presence of IL-22 and analyzed for expression of LY6C (Figure 25G) and LY6A (Figure 25H). Both LY6 molecules were substantially increased in the presence of IL-22 at comparable levels to the induction seen after treatment with IFN γ .

[0282] To ensure that the upregulation of LY6 molecules was not specific to the YAMC cell line, RNA levels of murine LY6A and LY6C in the murine colonic

epithelial tumor cell line CMT93 was examined. Levels of both murine LY6A and murine LY6C were upregulated upon treatment with IFN γ (Figure 25I). Though the levels of upregulation of LY6 molecules were more modest in CMT93 cells, flow cytometry analysis indicated that levels were quite high even in untreated cells (data not shown), which is likely a result of the tumor phenotype of CMT93 cells.

[0283] This data supports the data obtained by real time quantitative RT-PCR in confirming that IEC upregulate LY6 family members in response to inflammatory cytokines.

Example 5: LY6 stimulation of IEC is associated with lipid raft formation

[0284] As GPI-anchored proteins, LY6 family members do not possess a unique intracellular domain associated with traditional outside-in signaling. Rather, they are present within lipid raft microdomains (Bohuslav, J. et al. *Eur J Immunol* 23:825-831(1993)). However, it has been suggested that cross-linking of LY6 family members on the surface of cells results in redistribution of other cell surface molecules as well as reorganization of lipid raft structures, suggesting a mechanism by which LY6 molecules can affect signal transduction and downstream cellular functions (Simons, K. et al., *Nat Rev Mol Cell Biol* 1:31-39 (2000)).

[0285] Few ligands for LY6 proteins have been identified to date, and no ligand for LY6A or LY6C is currently known (Paret, C. et al., (2005) *Int J Cancer* 115:724-733; Apostolopoulos, J. et al., (2000) *Immunity* 12:223-232; and Classon, B.J. (2001) *Trends Immunol.* 22:126-127). Cholesterol is required to maintain lipid raft integrity. (Simons, K., et al. *J Clin Invest* 110:597-603 (2002)), and depletion of cholesterol is often used to inhibit lipid raft biosynthesis *in vitro* (von Tresckow, B. et al. *J Immunol* 172:4324-4331 (2004)).

[0286] To analyze whether lipid raft reorganization occurs in IEC in response to LY6 crosslinking, YAMC cells were grown in cholesterol-depleting conditions (conditions under which lipid rafts are depleted from cells) and cholesterol non-depleting conditions (conditions permissive for lipid raft formation). For cholesterol depleting conditions, YAMC cells were grown in the absence of serum and in the presence of 4 μ M lovastatin and 0.25 mM mevalonate (Sigma Chemical Co., St. Louis, MO) for 72 hours at 37 °C. The same growth conditions were use for YAMC cells under cholesterol non-depleting conditions, except that no lovastatin or mevalonate were added to the growth medium. Cells were then lifted and LY6C was crosslinked as described above in

Example 1. RNA was collected and expression levels of CXCL2, CXCL5, and CCL7 were determined.

[0287] The results of these studies indicated that lipid raft depletion results in an inhibition of LY6C-mediated chemokine production. Figures 26A-26C show that cholesterol depleted (dark bars) YAMC cells produced less chemokine than cells that were not depleted of cholesterol (open bars). Cholesterol depletion affected chemokine production in control anti-KLH stimulated groups, irrespective of LY6C stimulation, however the response was minimal and not in a consistent direction. To examine whether cholesterol depletion globally affected cell viability, we measured cell death, by 7AAD exclusion, and determined that cholesterol depletion did not significantly affect the viability of the YAMC cells (92% viability versus 86% in the cholesterol depleted cells, data not shown). Surface expression of both LY6A (Figure 26D) and LY6C (Figure 26E) were both significantly lower in cholesterol depleted YAMC cells, suggesting that plasma membrane cholesterol levels and lipid raft integrity affect the levels of LY6 expression on the surface of cells. This data suggests that lipid raft integrity, influenced by cholesterol biosynthesis, allows for the expression of LY6 molecules on the surface, and is potentially involved in the LY6C mediated induction of chemokines. Thus, the enhancement of chemokine production mediated by interaction of LY6C polypeptides in the cell membrane requires the presence of lipid rafts on the cell surface.

Example 6: Crosslinking LY6C results in increased surface expression of LY6 molecules

[0288] It has been reported that crosslinking LY6C on the surface of T cells results in shedding of LY6C (Jaakkola, I. et al. (2003) *J Immunol* 170:1283-1290). However, unlike T cells, when murine LY6C was crosslinked on the surface of IEC, no shedding of either LY6A or LY6C occurred (Figure 27A and 27B, respectively). To the contrary, in the absence of IFN γ , surface expression levels of both LY6A and LY6C were increased on IEC with crosslinked LY6C, but not LY6A. When IEC were preincubated with IFN γ , much of this effect was abolished (Figure 27C), however a slight upregulation of LY6A was still detected (Figure 27D).

[0289] These data indicate a positive feedback loop whereby stimulation through LY6C on IEC results in increased surface expression of LY6 molecules.

Example 7: Stimulation of LY6A results in increased secretion of chemokines

[0290] Functions for LY6 molecules have not been fully elucidated. To examine the role of LY6 molecules in the immunopathology of colitis, stimulation of LY6 molecules was studied for affects on the transcription and secretion of chemokines from IEC.

[0291] To analyze production of chemokines from IEC in response to crosslinking of murine LY6 molecules, YAMC cells, either pretreated with IFN γ or untreated, were cultured on plates coated with either anti-KLH control antibody, anti-LY6A or anti-LY6C. Twenty four hours later, mRNA from these cells was obtained and analyzed by quantitative RT-PCR for expression of CCL2, CCL4, CCL5, CCL7, CCL8, CCL25, CXCL1, CXCL2, CXCL5, CXCL10, CXCL12 and CX3CL1, which are chemokines that have been implicated in colitis (Table 9) (Papadakis, K.A. (2004) *Curr Allergy Asthma Rep* 4:83-89; Banks, C. et al., (2003) *J Pathol* 199:28-35; and Papadakis, K.A., and S.R. Targan (2000) *Inflamm Bowel Dis* 6:303-313). The assay was performed under non-permissive growth conditions (37°C in the absence of IFN γ) to rule out the possibility of increased proliferation of IEC in response to IFN γ stimulation.

Table 9

Pre-treatment >	Media			IFN		
Crosslink >	Anti-KLH	Anti-LY6A	Anti-LY6C	Anti-KLH	Anti-LY6A	Anti-LY6C
CCL2	1.01 (0.14)	-1.38 (0.13)	8.81 (0.72)	3.53 (0.21)	2.83 (0.21)	16.12 (0.56)
CCL4	1.31 (1.14)	1.83 (0.49)	3.15 (1.17)	5.35 (0.48)	3.65 (0.69)	16.12 (0.56)
CCL5	1.00 (0.10)	1.09 (0.02)	3.31 (0.15)	2.73 (0.13)	2.76 (0.23)	10.13 (0.27)
CCL7	1.00 (0.11)	1.06 (0.08)	3.37 (0.15)	2.82 (0.44)	1.39 (0.33)	5.81 (0.51)
CCL8	1.03 (0.30)	2.05 (0.37)	12.78 (3.14)	74.22 (8.94)	74.44 (9.81)	110.44 (3.36)
CCL25	1.01 (0.16)	1.06 (0.11)	1.16 (0.00)	1.38 (0.36)	1.32 (0.15)	1.46 (0.19)
CXCL1	1.00 (0.07)	-3.17 (0.11)	11.58 (0.12)	-1.13 (0.10)	-1.64 (0.17)	13.36 (0.35)
CXCL2	1.33 (1.10)	ND	21.81 (3.13)	14.30 (3.30)	10.95 (3.05)	113.20 (16.23)
CXCL5	1.08 (0.53)	ND	118.45 (65.14)	1.70 (1.15)	ND	150.99 (55.50)
CXCL10	1.00 (0.05)	1.02 (0.06)	5.11 (0.19)	5.68 (0.31)	5.22 (0.22)	12.22 (0.51)
CXCL12	1.01 (0.14)	1.12 (0.05)	-1.99 (0.14)	-1.11 (0.05)	-1.23 (0.21)	-3.02 (0.06)
CX3CL1	1.00 (0.08)	-1.18 (0.15)	1.92 (0.07)	2.21 (0.11)	1.87 (0.16)	3.22 (0.42)

[0292] Cells pretreated with IFN γ showed upregulation of many of these chemokine genes (see Media, Anti-KLH group versus IFN γ , Anti-KLH group of Table 9). However, with the exception of an upregulation of CCL8 and a downregulation of CXCL1, anti-LY6A stimulated YAMC cells showed similar gene expression patterns as anti-KLH stimulated YAMC cells. However, YAMC cells stimulated with anti-LY6C showed increased expression of all chemokines analyzed except for CCL25, which remained essentially unchanged, and CXCL12, which was downregulated in response to LY6C stimulation. While the increased gene expression of chemokines induced by LY6C crosslinking was not dependent upon IFN γ , cells pretreated with IFN γ showed increased expression of chemokines versus cells that had not been pretreated with IFN γ .

[0293] To analyze the kinetics of chemokine induction induced by murine LY6C stimulation, 96 well plates were coated with anti-KLH antibody or either anti-LY6A or anti-LY6C monoclonal antibodies. YAMC cells, either pretreated or not with IFN γ , were added for 24, 48 or 72 hours. At the indicated time point RNA was collected for quantitative RT-PCR analysis and supernatants were collected for ELISA.

[0294] Within 24 hours, a spike in transcription of both CXCL5 and CCL7 was detected on cells with crosslinked LY6C, but not LY6A (Figure 28A). Increased expression of CXCL5 and CCL7 diminished over time but was still detectable after 72 hours in culture. Though IFN γ was not required to enhance chemokine transcription, IFN γ acted synergistically with LY6C stimulation in inducing transcription of both CXCL5 and CCL7 at early time points.

[0295] In parallel with the gene expression, supernatants of LY6C, but not LY6A, crosslinked cells contained significantly higher concentrations of CXCL5 at 48 hours (Figure 28B). The effect was dose dependent, and detectable with as little as 1 μ g/ml of coated anti-LY6C. Like transcription, secretion of CXCL5 was enhanced when cells were pretreated with IFN γ , but IFN γ was not required for the effect. Increased secretion of CXCL5 was noted at both the 24 and 72 hour time points as well.

[0296] To ensure that LY6C was involved in the observed upregulation of chemokines, we used siRNA to knockdown LY6C. LY6C transcript was inhibited by 95% in the absence of IFN γ and about 90% in the presence of IFN γ by real time quantitative RT-PCR which corresponded to significantly lower levels of LY6C on the surface of the YAMC cells (data not shown). Cells with decreased levels of LY6C on the surface showed a diminished response to LY6C crosslinking with regard to transcription of chemokines (Figure 28C). Secretion of CXCL5 was markedly inhibited by knocking down LY6C as well (data not shown).

[0297] These results indicate that crosslinking of LY6C, but not LY6A, on the surface of IEC results in increased secretion of chemokines.

Example 8: IEC in vivo show a similar chemokine gene expression to LY6C stimulated cells

[0298] The above data establishes a model whereby IEC stimulated through murine LY6C significantly upregulate expression of chemokine genes.

[0299] Analyzing the microarray data from laser capture microdissected IEC in murine models of colitis, the expression of the same 12 chemokine genes in healthy

and colitic mice in the two murine models of colitis was examined to determine if the chemokines stimulated by LY6C crosslinking *in vitro* correlate with the chemokines secreted by IEC *in vivo* (Figures 29A and 29B). Though the expression pattern is not identical to the upregulation of chemokines resulting from LY6C stimulation, expression of CXCL5, which was the most highly upregulated chemokine gene in *in vitro* studies, was also the highest upregulated chemokine in murine models of colitis. We saw significant upregulation in expression of CXCL1, CXCL10, CCL5 and CCL7 in both models of colitis. In addition, we saw upregulation of CCL4 and CCL8 in the transfer colitis model or the IL10 $-/-$ model, respectively.

[0300] Interestingly, the only chemokine that was down-regulated as a result of murine LY6C stimulation *in vitro*, CXCL12 was also the only one of these chemokines downregulated *in vivo*.

Example 9: Expression of human LY6 genes in colon cells

[0301] Expression of human LY6H, LYPD1, LYPD3, and LYPD5 in a human colon cell line, Colo 205 cells (a cell line derived from human colon carcinoma, ATCCTM accession number CCL-222TM), was examined. Human Colo 205 cells were treated with the cytokines IFN- γ , LPS, TNF α , IFN- γ + TNF α , IFN- γ + LPS, or LPS + TNF α (all at 100 ng/ml, except LPS at 1 ug/ml) for 18 hours (LYPD3) or 24 hours (LY6H or LYPD5). RNA was collected and purified and expression of the indicated LY6 family member was determined by quantitative RT-PCR using reagents from Applied BiosystemsTM according to manufacturer's instructions. Primers and probes used for RT-PCR analysis were the following:

LYPD1:

Sense: CAT GAT CCT CCG AAT CTG GT (SEQ ID NO:59)

Antisense: AGC ACA GAA CAG AGG GGC TA (SEQ ID NO:60)

Probe: ATA CGG CCA ATG TCA CAA CA (SEQ ID NO:61)

LYPD3:

Sense: ACT TCC TGT TCC CAC CAC TG (SEQ ID NO:62)

Antisense: AGA GGA CAA GCG GAG AGA CA (SEQ ID NO:63)

Probe: TTC TGG CAG GGG TGT TCT AG (SEQ ID NO:64)

LY6H:

Sense: AGC AGC AGC AGG AAG GAT (SEQ ID NO:65)

Antisense: AAA AGT GCC GCT TAA CGA AG (SEQ ID NO:66)

Probe: CAA GAT GTG TGC TTC CTC CTG CGA (SEQ ID NO:67)

[0302] LYPD5 primers and probes were purchased from Applied Biosystems™ (catalog number HS00289062_m1).

[0303] The results plotted in Figures 30A-30C indicate fold increases in expression of these human LY6 genes relative to human B-actin control. Significant increases in expression of human LY6H, LYPD3, and LYPD5 were observed following treatment with the indicated cytokines.

Example 10: Expression of human LY6 genes in colon biopsy tissue

[0304] To further investigate the source of the increased LYPD1 and LYPD5 expression in the colon of patients with CD and UC, was undertaken in a cohort of biopsies of patients with UC, CD and controls. Microarray analysis for LYPD1 expression using RNA extracted from colon biopsies showed statistically increased expression in inflamed colon tissue of CD patients (Figure 31A). In the UC and CD biopsies taken from the colon, statistically increased LYPD5 expression was observed in inflamed UC and CD patients (Figure 31B). This was not observed in the non-inflamed control biopsies.

[0305] Expression of human LY6H in terminal ileum biopsies of inflamed IBD tissue was analyzed relative to control (non-IBD) terminal ileum biopsies using RT-PCR (Taqman™) analysis. Human LY6H expression was at least 1.5 fold greater in inflamed IBD biopsies relative to control.

[0306] Human LYPD3 expression in inflamed UC colon biopsies was upregulated and less than 2 fold greater in inflamed IBD biopsies relative to control.

[0307] The results of these examples demonstrated expression of LY6 molecules on the surface of IEC, and further indicated that expression is unique to IEC in the context of inflammation. Furthermore, surface expression levels of LY6A and LY6C were high on IEC of colitic mice, and nearly universal throughout the colon. As molecules both specific to the diseased state, and ubiquitously expressed during disease, detection of human LY6 gene or polypeptide expression, particularly human LY6H, LYPD1, LYPD3, and LYPD5, is a useful method for detecting IBD, including UC and/or CD in humans. Additionally, the method of detecting human LY6 expression is useful for diagnosing IBD, UC and/or CD in a human and monitoring response to IBD therapeutic agents.

[0308] In the Examples disclosed herein, the functional significance of LY6 expression in IEC was demonstrated. YAMC cells were strongly positive for LY6A, and expressed lower levels of LY6C. However, upon stimulation with a number of cytokines present within the colon during colitis, including IL-1 β , TNF α , IFN α , and in particular IL-22 and IFN γ , expression levels of both LY6 molecules were greatly enhanced. YAMC cells pretreated with IFN γ to upregulate expression of LY6 molecules, were a useful *in vitro* model to analyze functional significance for LY6 expression.

[0309] The conditionally immortalized nature of the YAMC cells comes from MHC II promoter driven expression of the SV40 large T antigen; low levels (2.5-5 U/ml) of IFN γ are used to drive proliferation of these cells (Whitehead, R.H. et al. (1993) *Proc Natl Acad Sci U S A* 90:587-591; Whitehead, R.H., and J.L. Joseph. (1994) *Epithelial Cell Biol* 3:119-125). YAMC cells are often used as an *in vitro* model for cytokine treatments of murine IEC (Mei, J.M. et al. (2000) *Faseb J* 14:1188-1201; Yan, F., and D.B. Polk (2002) *J Biol Chem* 277:50959-50965). The SV40 large T antigen that these cells contained is temperature sensitive, and non-functional at 37°C. All experiments performed herein involved IFN γ treatment under these non-permissive conditions. In addition, YAMC cells were serum starved (and IFN γ starved) at 37°C for 24 hours prior to experiments. Under such conditions, effects indicating residual T antigen expression, such as proliferation of cells, were not observed. As a result, effects of IFN γ treatment were due to inherent effects of IFN γ rather than effects stemming from driving expression of the T antigen. Furthermore, the upregulation of LY6 family members was detected in a second murine cell line, CMT93, confirming that this effect is broadly applicable to IEC.

[0310] Furthermore, IFN γ was not unique among cytokines for inducing LY6 molecules as modest upregulation of LY6 expression was noted after treatment with TNF α , IL-1 β and, IL-22. The upregulation of LY6 molecules on IEC in response to IL-22 is interesting in light of recent data demonstrating a potential role for IL-22 in Crohn's Disease (Wolk, K., et al. *J Immunol* 178:5973-5981 (2007)). Though homology between mouse and human LY6 molecules are often complicated, there is evidence to suggest that the upregulation of LY6 molecules is not restricted to mice. Previous studies in rats have suggested upregulation of LY6 molecules in the small intestine in colitis models, and it has been suggested that such expression is involved in inflammation, cell/cell interactions as well as signaling within the rat IEC (Baksheev, L. et al. *J Gastroenterol* 41:1041-1052 (2006)).

[0311] The data described above indicates that there is a possibility that lipid raft integrity is involved in LY6C mediated signal transduction in IEC. This implies that disruption of lipid rafts might serve to attenuate downstream affects of LY6C stimulation both by downregulating LY6C expression and disrupting the structural components of LY6C signaling. Recently, it has been determined that cholesterol depletion of IEC with statins inhibits proinflammatory gene expression through NF- κ B modulation (Lee, J. et al., *Int Immunopharmacol* 7:241-248 (2007)). Furthermore, statins have been effective therapeutics in murine models of colitis (Naito, Y., et al. *Int J Mol Med* 17:997-1004 (2006)). The mechanism linking lipid raft motility and NF- κ B blockade remain undetermined, but our data suggests that activation through LY6C could be one hypothesis to explain the mechanism of action.

[0312] In this study, we identify LY6 molecules as a potential upstream switch in the expression of chemokine genes. Crosslinking of the LY6C receptor with monoclonal antibodies resulted in dramatic upregulation of nearly all chemokines analyzed, including CXCL5. We further confirmed that CXCL5 secretion is greatly enhanced in LY6C crosslinked IEC. It is interesting that even though both LY6A and LY6C are anchored to the cell surface by a GPI moiety, and despite higher levels of expression of LY6A than LY6C on the surface of IEC, that the downstream effects on chemokine secretion are seen with LY6C crosslinking and not consistently with LY6A crosslinking.

Example 11: Identification of a ligand for LYPD5

[0313] In this study, a search for ligands of LYPD5 was performed through techniques well known to those of ordinary skill in the art, namely by expression cloning of about 14,000 human genes under CMV promoter into COS cells. Pools of 100 genes were transfected into the COS cells grown in 140 wells on 12 well plates. Following transfection, the cells were stained with LYPD5-Fc protein (see Figure 32). Wells with positive staining were identified and individual clones were transfected into COS cells. A single well expressing a single protein, GLG-1 (ESL-1) was identified as a ligand for LYPD5. GLG-1 is characterized by a lengthy extracellular domain (ECD), a transmembrane domain and a cytoplasmic domain. A series of co-immunoprecipitation studies were conducted using techniques known to those of ordinary skill in the art to assess the ability of various regions of the GLG-1 ECD to bind LYPD5. It was found that variants or fragments of the GLG-1 ECD (see Figures 33-35) were able to serve as a

ligand for LYPD5. Figure 33B shows the results of co-immunoprecipitation studies using Fragments 1, 2, 3, or 4 as depicted in Figure 33A and demonstrates that any one of the fragments is sufficient for LYPD5 binding.

[0314] In addition, a GLG-1 ECD domain by itself was found to be sufficient for LYPD5 binding. As shown in Figure 33, GLG-1 is made up of multiple GLG-1 domains and single GLG-1 domains can bind LYPD5. Figure 34B shows the results of a co-immunoprecipitation demonstrating that Fragments 1, 2, 3, and 4, as well as single GLG-1 domains 115, 150, 215, 538, 609, 670, 729, and 858 (as shown in Figure 34A) were able to bind LYPD5.

[0315] Through another co-immunoprecipitation study, binding was shown to be specific based on fragments of LYPD5 (see Figure 35A) in which LYPD5 was found not to bind BAP negative control, an FN14 negative control was not found to bind GLG-1 fragment 2, the human GLG-1 domain 115 binds LYPD5, domain 115 is not always expressed at detectable levels but still pulls down LYPD5, and the fraction of human GLG-1 fragment 1 that lacks domain 115 (residues 26-114) does not bind LYPD5 (Figure 35B).

[0316] The “*” in Figures 34A and 35A indicates a potential fucosylation site.

[0317] Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, the descriptions and examples should not be construed as limiting the scope of the invention. The disclosures of all patent and scientific literatures cited herein are expressly incorporated in their entirety by reference.

WHAT IS CLAIMED IS:

1. A method of detecting inflammatory bowel disease (IBD) in a mammal, comprising detecting the level of expression of a gene encoding a LY6 polypeptide (a) in a test sample of tissue or cells obtained from said mammal, and (b) in a control sample, wherein a higher level of expression of the LY6 nucleic acid or polypeptide in the test sample, as compared to the control sample, is indicative of the presence of IBD in the mammal from which the test sample was obtained.
2. The method of claim 1, wherein the tissue or cells of the test sample are from the gastrointestinal tract of the mammal.
3. The method of claim 2, wherein the tissue or cells of the test sample are from the colon of the mammal.
4. The method of claim 1, wherein the control sample is a sample of normal non-IBD tissue or cells of the same tissue origin or type, or multiple samples of non-IBD tissue or cells of the same tissue origin or type the expression levels of which are averaged, or a universal control representing gene expression in multiple samples of healthy, normal tissue of the same species.
5. The method of claim 1, wherein the tissue or cells of the test sample are inflamed.
6. The method of claim 1, wherein the tissue or cells of the test sample are not inflamed.
7. The method of claim 1, comprising:
 - (a) contacting the test sample with a detectable agent that specifically binds a polynucleotide that encodes LY6 polypeptide or fragment thereof; (b) contacting the control sample with the detectable agent; and (c) detecting the formation of a complex between the agent and the polynucleotide of the test sample and the control sample, wherein the formation of less complex in the test sample relative to the control sample is indicative of the presence of IBD in the mammal.

8. The method of claim 7, wherein the polynucleotide comprises the nucleic acid sequence of SEQ ID NO:1, 3, 5, 7, or a fragment thereof comprising at least 15 contiguous nucleotides of SEQ ID NO:1, 3, 5, or 7.
9. The method of claim 7, wherein the tissue or cells of the test sample are from the gastrointestinal tract of the mammal.
10. The method of claim 9, wherein the tissue or cells of the test sample are from the colon of the mammal.
11. The method of claim 10, wherein the wherein the control sample is a sample of normal non-IBD tissue or cells of the same tissue origin or type, or multiple samples of non-IBD tissue or cells of the same tissue origin or type the expression levels of which are averaged, or a universal control representing gene expression in multiple samples of healthy, normal tissue of the same species.
12. The method of claim 1, wherein the tissue or cells of the test sample are inflamed.
13. The method of claim 1, wherein the tissue or cells of the test sample are not inflamed.
14. The method of claim 7, wherein the agent is a second polynucleotide that hybridizes to a polynucleotide having the sequence SEQ ID NO:1, 3, 5, 7 or its complement.
15. The method of claim 14, wherein the second polynucleotide comprises a detectable label or attached to a solid support.
16. The method of claim 15, wherein the detectable label is directly detectable.
17. The method of claim 16, wherein the detectable label is indirectly detectable.
18. The method of claim 14, wherein the detectable label is a fluorescent label.

19. The method of claim 7, wherein the method is in situ hybridization assay.
20. The method of claim 7, wherein the method is real time polymerase chain reaction (RT-PCR) assay.
21. The method of claim 7, wherein the method is real time polymerase chain reaction (RT-PCR) assay.
22. The method of claim 1, comprising:
 - (a) contacting the test sample with a detectable agent that specifically binds a LY6 polypeptide or fragment thereof; (b) contacting the control sample with the detectable agent; and (c) detecting the formation of a complex between the agent and the polypeptide of the test sample and the control sample, wherein the formation of less complex in the test sample relative to the control sample is indicative of the presence of IBD in the mammal.
23. The method of claim 22, wherein the LY6 polypeptide comprises SEQ ID NO:2, 4, 6, 8 or a fragment thereof comprising at least 10 contiguous amino acids of SEQ ID NO:2, 4, 6, or 8.
24. The method of claim 22, wherein the tissue or cells of the test sample are from the gastrointestinal tract of the mammal.
25. The method of claim 24, wherein the tissue or cells of the test sample are from the colon of the mammal.
26. The method of claim 7 or 22, wherein the tissue or cells of the test sample are inflamed.
27. The method of claim 7 or 22, wherein the tissue or cells of the test sample are not inflamed.

28. The method of claim 22, wherein the agent is an antibody or binding fragment thereof.
29. The method of claim 28, wherein the antibody or binding fragment thereof comprises a detectable label.
30. The method of claim 29, wherein the detectable label is directly detectable.
31. The method of claim 29, wherein the detectable label is indirectly detectable.
32. The method of claim 29, wherein the detectable label is a fluorescent label or a radiolabel.
33. The method of claim 1, claim 7 or claim 22, wherein the test sample of tissue or cells is obtained from a mammal suspected of experiencing IBD.
34. The method of claim 1, claim 7, or claim 22, wherein the test sample of tissue or cells is obtained from a mammal suspected of experiencing UC.
35. The method of claim 1, claim 7, or claim 22, wherein the tissues or cells of the mammal have been contacted with a therapeutic agent, and wherein the level of LY6 expression is indicative of the presence or absence of a response to the therapeutic agent in the tissue or cells of the mammal.
36. The method of claim 1, claim 7, or claim 22, wherein the tissues or cells of the mammal have been contacted with a therapeutic agent, wherein the detecting is a second or subsequent detecting, and wherein the level of LY6 expression is indicative of the presence or absence of a response to the therapeutic agent in the tissue or cells of the mammal.
37. The method of claim 1, claim 7, or claim 22, wherein the increase in LY6 expression in the test sample is at least 1.5 fold, at least 2 fold, at least 3 fold, at least 5 fold, at least 6 fold, at least 7 fold, at least 8 fold, at least 9 fold, or at least 10 fold greater than the control sample.

cgcgtctcggctgcgttccccgaaagacgaggctgcgcccggattccggtccgcaggagaccgaagggcacagctccccgcgccg
cgcacgccgcccagccccggagtgcggacacccccgggatgcttgcgccccagaggacccgcgccccaaagccccgcgccgcccc
aggccccacccggagcatgctgcctgcagccatgaagggcctcggcctggcgctgctggccgtcctgctgtgctcggcgcccgtcatgg
cctgtggtgccaggactgcaccctgaccaccaactccagccattgcaccccaaagcagtgccagccgtccgacacggtgtgtgccagtgt
ccgaatcaccgatcccagcagcagcaggaaggatcactcgggaacaagatgtgtgcctcctcctgtgacttcgttaagcgacacttttctc
agactatctgatgggggtttattaactctgggatcttaaggctgacgtggactgctgcgagaaggatttgtgcaatggggcggcaggggcag
ggcacagccccctgggccctggccgggggggctcctgctcagcctggggcctgccctcctctgggctggggccctgatgtctcctcctccca
cggggcttctgagcttgctccccctgagcctgtggctgccctctccccagcctggcgtggctggggctgggggcagccttgcccagctcc
gtggctgtggcctgtggctctcactcctccccgacgtgaagcctccctgtctctccgagctctgagtcacagcagctggacatctcca
ggaaaccaggccatctgggcaggaggcctggggatgagggtggggggggacccccaggtcccggagggggaagtgaagcaacagcc
cagctggaagggcgtcttctgcggag aaataaagtcacttttgagtcctgagaaaaaaaaa

Figure 1A

MLPAAMKGLGLALLAVLLCSAPAHGLWCQDCTLTTNSSHCTPKQCQPSDTVCAVRIT
DPSSSRKDHSVNMKMCASSCDFVKRHHFFSDYLMGFINSKILKVDVDCCEKDLNKAAGA
GHSPWALAGGLLLSLGPALLWAGP

Figure 1B

agggcggtgt caatgcaccc tccagcgggtg cgcgcagggc ggagaaggga gggcgggccg ggcaagtgag acagttaagg
 cagtgtccc accacacccc caccagatt ggccacgcc agctgggtct tgacagaagg ccttcgcgga ggaagagggg
 gcacagctgc acaggacacc ctacggagcc tgcgggcgtg gaactttgcc aggcgcacgg gaacgcgcgc ccttctgtc
 agcctcccgg ggcgccaggc tcccgcggcc cgcagcggga cagcctcagt tgtgtgggtt ggaccagtc gctggggtag
 cgaccagtc tggaaggcgc agaggacgtg gactggggag gctgccttc tatgtcgaa gggccagccg ggcacgcagt
 cctcagaccc tagtcgcac ccggcaggtc cccacggcac ctgctgcgc ctcctgcgc ctecccaac cteccatct
 cagaaaacta ccagttctt cccgcccccc ggcgccccct tcccaggaac gtgcggaggc gggagaagag gaagacagga
 aggggggtgg gatgtgaagc gaccgtcca gcctccccg cccgccacc ccacccaac tcggcagccg tcacgtgatg
 cctggagtgg gaggtgggga gaaaaggcga gactttgtg ggtgctccg atcgccagta gtctctcag tctcagccgc
 caactccgga ggcgcggtgc tcggcccggg agcgcgagcg ggaggagcag agaccgcag ccgggagccc gagcgcgggc
 gatgcaggct ccgcgagcgg cacctgcggc tctctaagc tacgaccgtc gtctccgcg cagcagcgc ggccccagca
 gcctcgagc ccacagccgc tgcagccggg gcagcctcc ctgctgtgc ctcctctgat gcgcttgccc tctccggcc
 ccgggactcc gggagaatgt gggctctagg catcgcgga acttttgcg gattgttct gcttcaggc ttgcgctgc
 aaatccagt ctaccagtgt gaagaattcc agctgaacaa cgactgtcc tccccgagt tcattgtgaa ttgcaggtg
 aacgttcaag acatgtgta gaaagaagt atggagcaaa gtgccgggat catgtaccgc aagtcctgt catcatcagc
 ggctgtctc atgcctctg ccgggtacca gtcctctgc tcccaggga aactgaact agttgcac agctgtgca
 acacctct ttgtaacgg ccaaggcca agaaaaggg aagtctgcc tcggccctca ggccagggt ccgcaccacc
 atcctgttc tcaaattagc cctctctcg gcacactgt gaagctgaag gagatgccac cccctctgc attgtctc cagccctgc
 cccaacccc ccacctccct gactgagttt ctctgggtg tcttttatt ctgggtagg agcgggagtc cgtgttctt ttgttctg
 tgcaataat gaaagagtc ggtaaagcat tctgaataaa ttcagcctga ctgaatttc agtatgtact tgaaggaag
 aggtggagt aaagttacc ccatgtctg tgtaaccgga gtcaaggcca ggctggcaga gtcagtcctt agaagtcact
 gaggtgggca tctgccttt gtaaagcctc cagtgtccat tccatccctg atgggggcat agtttgagac tgcagagtga
 gactgacgt ttcttagggc tggagggcca gtcccaactc aaggctccct cgttgacat tcaaactca tgcctctgaa aaccattctc
 tgcagcagaa ttggtggtt tcgcgctga gttgggtct agtgactga gactcaatga ctgggactta gactggggct
 cggcctcgt ctgaaaagt ctaagaaaa tctctcagt tctcctgca gaggactgg gccgggacgc gaagagcaac
 gggcgctgca caaagcgggc gctgtcgggt gtggagtgc catgtacgc caggcgctc tctggttg cgtgtgcag
 cgacaggcgg cagcacagca cctgcacgaa caccgcccga aactgctgc aggacaccgt gtacaggagc gggttgatga
 ccgagctgag gtagaaaaac gtctccgaga aggggaggag gatcatgtac gcccggaagt aggacctgt ccagtcgtgc
 ttgggtttgg ccgcagccat gatctccga atctggttg gcatccagca tacggccaat gtcacaaca tcagccctgg
 gcagacacga gcaggaggga gagacagaga aaagaaaaac acagcatgag aacacagtaa atgaataaaa ccataaata
 tttagccct ctgtctgtg ctactggcc aggaatggt accaatttt cagtgttga cttgacagct tctttgcca caagcaagag
 agaattaac actgtttcaa acccggggga gttggctgt taaagaaag accattaat gcttagaca gtgtattat accagttgat
 gtctgttaatt ttaaaaaaa tgtttcatt ggtgttgt tgcgtatcca gaaagcagtt catgttatcc ata

Figure 2A

gatgcctgga gtgggaggtg gggagaaaag gcgagacttt tgtgggtgct cccgatcgcc agtagttcct tcagtctcag
ccgccaactc cggaggcgcg gtgctcggcc cgggagcgcg agcgggagga gcagagaccc gcagccggga gcccagcgcg
gggcgatgca ggctccgcga gcggcacctg cggctcctct aagctacgac cgtcgtctcc gcggcagcag cgcgggcccc
agcagcctcg gcagccacag ccgctgcagc cggggcagcc tccgctgctg tcgcctcctc tgatgcgctt gccctctccc
ggccccggga ctccgggaga atgtgggtcc taggcatcgc ggcaactttt tgcggattgt tcttgcttcc aggctttgcg
ctgcaaatec agtgctacca gtgtgaagaa ttccagctga acaacgactg ctctccccc gagttcattg tgaattgcac
ggtgaacgtt caagacatgt gtcagaaaga agtgatggag caaagtgccg ggatcatgta ccgcaagtc tgtgcatcat
cagcggcctg tctcatcgcc tctgccgggt accagtcctt ctgctcccca gggaaactga actcagtttg catcagctgc
tgcaaacacc ctctttgtaa cggggccaagg cccaagaaaa ggggaagtgc tgcctcggcc ctccaggccag ggctccgcac
caccatctg ttctcaaat tagccctctt ctggcacac tctgaagct gaaggagatg ccacccctc ctgcattgtt ctccagccc
tcgccccaa cccccacct cctgagtga gtttctctg ggtgtcctt tattctgggt agggagcggg agtccgtgtt ctctttgtt
cctgtgcaa laatgaaaga gctcggtaaa gcattctgaa taaattcagc ctgactgaat ttctagatg tacttgaagg
aaggaggtgg agtgaaagt caccacctg tctgtgtaac cggagtcaag gccaggctgg cagagtcagt ccttagaagt
cactgaggtg ggcatctgcc tttgtaaaag cctccagtgt ccattccatc cctgatgggg gcatagttg agactgcaga
gtgagagtga cgtttctta gggctggagg gccagttccc actcaaggct cctcgttg acattcaaac ttcatgctcc
tgaaaacat tctctgcagc agaattggct ggttcgcgc ctgagttggg ctctagtac tcgagactca atgactggga
cttagactgg ggctcggcct cgctctgaaa agtgcttaag aaaatctct cagttctct tgcagaggac tggcgccggg
acggaagag caacgggcgc tgcacaaagc gggcgctgtc ggtggtggag tgcgcatgta cgcgcaggcg ctctcgtg
ttggcgtgct gcagcgacag gcggcagcac agcacctgca cgaacacccg ccgaaactgc tgcgaggaca ccgtgtacag
gagcgggtg atgaccgagc tgaggtagaa aaacgtctcc gagaaggga ggaggatcat gtacgcccgg aagtaggacc
tcgtccagtc gtgcttgggt ttggccgag ccatgatcct ccgaatctgg ttgggcatcc agcataggg caatgtcaca
acaatcagcc ctgggcagac acgagcagga gggagagaca gagaaaagaa aaacacagca tgagaacaca gtaaatgaat
aaaaccataa aatatttagc ccctctgtt tgtgcttact ggccaggaaa tggaccaat tttcagtggt tggactgac agcttcttt
gccacaagca agagagaatt taacactgt tcaaaccgg gggagttggc tgtgttaaag aaagaccatt aaatgcttta gacagt

Figure 2B

MWVLGIAATFCGLFLLPGFALQIQCYQCEEFQLNNDCCSPEFIVNCTVNVQDMCQKEV
MEQSAGIMYRKSCASSAACLIASAGYQSFCSPGKLNSVCISCCNTPLCNGPRPKKRGSSA
SALRPGLRTTILFLKLALFSAHC

Figure 2C

aaggctgggg ttgcctgggg cgaggttact catcctgggc tcaggtaaga gggcccgagc tcggaggcgg cacatccagg
ggggacgcca agggagcagg acggagccat ggaccccgcc aggaaagcag gtgcccaggc catgatctgg actgcaggct
ggctgctgct gctgctgctt cgcggaggag cgcaggccct ggagtgtac agctgcgtgc agaaagcaga tgacggatgc
tccccgaaca agatgaagac agtgaagtgc gcgcccggcg tggacgtctg caccgaggcc gtggggggcg tggagaccat
ccacggacaa ttctcgctgg cagtgcgggg ttgcggttcg ggactccccg gcaagaatga ccgcggcctg gatcttcacg
ggcttctggc gtcatccag ctgcagcaat gcgtcagga tcgtgcaac gccaagctca acctcacctc gcgggcgctc
gacccggcag gtaatgagag tgcatacccg cccaacggcg tggagtgtgta cagctgtgtg ggcctgagcc gggaggcgtg
ccagggtaca tcgcccgccg tcgtgagctg ctacaacgcc agcgatcatg tctacaaggg ctgcttcgac ggcaacgtca
ccttgacggc agctaattgt actgtgtcct tgcctgtccg gggctgtgtc caggatgaat tctgcactcg ggatggagta
acaggcccag ggttcacgt cagtggctcc tgttccagg ggtcccgtg taactctgac ctccgcaaca agacctactt
ctccccga atcccacccc ttgtccggt gcccccca gagccacga ctgtggcctc aaccacatct gtcaccactt
ctacctggc cccagtgaga cccacatcca ccacaaacc catgccagcg ccaaccagtc agactccgag acagggagta
gaacacgagg cctcccggga tgaggagccc aggttgactg gaggcgccgc tggccaccag gaccgcagca attcagggca
gtatcctgca aaaggggggc cccagcagcc ccataataa ggctgtgtgg ctcccacagc tggattggca gcccttctgt
tgccgtggc tctgtgtgtc ctactgtgag cttctccacc tggaaattc cctctacct acttctctgg cctgggtac cctctctc
atcacttct gtcccacca ctggactggg ctggcccagc cctgtttt ccaacattcc ccagtatccc cagcttctgc tgcgtggtt
tgcggctttg ggaaataaaa taccgttgta tatattctgc cagggtgtgt ctagctttt gaggacagct cctgtatcct tctatcctt
gtctctccgc ttgtcctct gtgatgttag gacagagtga gagaagtcag ctgtcacggg gaaggtgaga gagaggatgc
taagttcct actcacttc tctagccag cctggacttt ggagcgtggg gtgggtggga caatggctcc cactctaag
cactgcctcc cctactccc gcatttttg ggaatcggt cccatatgt cttccttact agactgtgag ctctcgagg
gcagggaccg tgccttatgt ctgtgtgtga tcagttctg gcacataaat gccacaataa agatttaatt actttgaaaa aaaaaaaaaa
aaaaaaaaa aaaaaaaaaa aaaaa

Figure 3A

MDPARKAGAQAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAP
GVDVCTEAVGAVETIHGQFSLAVRGCGSLPGKNDRGLDLHGLLAIFIQLQQCAQDRCN
AKLNLTSRALDPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGC
FDGNVTLTAANVTVSLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYF
SPRIPPLVRLPPPEPTTVASTTSVTTSTSAPVRPTSTTKPMPAPTSQTPRQGVEHEASRDEE
PRLTGGAAGHQDRSNSGQYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

Figure 3B

gatccgcttt ggcgatccca gtgattcttg ggttccgcgt gtagtttcgg aaggagacat cgaagcaggg cgaggcgcag
agggcggtgc ggactcatgc ccagtcggc agtgcggggt cccaagccct gcagtgtac agctttgagc acacctactt
tgcccccttt gacctcaggg ccatgaagct gccagcatc tctgtcctc atgagtgtt tgaggctatc ctgtctctgg
acaccgggta tcgcgcgccg gtgacctgg tgcggaagggt ctgtggacc gggcctcctg cgggccagac gcaatcgaac
gcggacgcgc tgccgccaga ctactcgggt gtgcgcgggt gcacaactga caaatgcaac gccacctca tgactcatga
cgccctcccc aacctgagcc aagcaccga cccgccgacg ctacgcggcg ccgagtgtta cgcctgtatc ggggtccacc
aggatgactg cgctatcggc aggtcccgac gagtcagtg tcaccaggac cagaccgct gcttcagggt caatggcaga
atgacagtg gcaattctc agtcctgtg tacatcagaa cctgccaccg gccctcctgc accaccgagg gcaccaccag
ccccctggaca gccatcgacc tcaggggctc ctgtgtgag gggtagctt gcaacaggaa atccatgacc cagccctca
ccagtgttc agccaccacc cctccccgag cactacaggt cctggccctg ctctccag tctctgtgt ggtggggctc
tcagcataga ccgccccctc aggatgtgtg ggacagggt cacacacctc attctgtgt ctccagcccc tatcatatg
ctactggaa aatgatgtta aagtaagaat tgcactctg tctctgtgc ctccatctc tccgcccc gtccccaca acctggccaa
cagtactgga agaaactgga cacagtcacc agcatccccg gggaggggcaa aacagccatg tcgtgccccg atgaagagca
attctgatca cagtgttac tcatgagca ccaggcaggc accaggcacc ccataacacg gcttctgtg ctctccctcc
agagcctgtc gcagctctag gagggagcta tacaatgatg tctttattag tgcacatg agaagcccaa taagcagtat
gccctaacag ttagtaggac aggtcttgga gctaagctgc atgggtcaa atccagctc caccattcag cctgcagaga
ccatgagcga gttacttaag ccaggctctg gagtaagct gcatgggtc aaatcccagc tccagcttc agcctacaga
gacatgggt gagttactta agccaggctc tggagctaag ctgcatgggt tcaaatccca gctccacat tcagcctga
gagactgtgg gtgagttact tgagctctc gtgccaatat ttctcactc ataaggtgga ggtgaaaata aactctataa catgacaaga
actacttcac agtagttgca gtgaggatc aacgagatga acatttagta cttgggacac agcagtggcc cagtgtaat
gggctacttg tcataagccc taagtcacag gtcaacaaac tgagaggcaa aagcacttg ttgagctgt gtatctagt
agtatggatt cagggaccag attcccagcc ccacgaactg ctaagcaacc ccacctcta aacacatgag tgccgattaa
cttcacagaa aaacacacaa ggcaaagttc agcgagggtga aattctccaa gctataaaga tcagggaaga ctcttgag
gaattcacc ttgagcaaaa tctaaagga tcaatagtag ctggcaaaaa gaagcaggag gaagcgcatt ctaggtagag
gagacagcct ggacaaaggt ctgagggagg aaggagcaca aggagtgcag gacacttca tgagtgcagg acacttcat
aactgcatga actcataga gatgggatcc tttagcatgt tctctgtgca catgctgac catgttctt cacatgttt ttgccattg
atcttccag caactcagt agagaagcaa aaaagtaagt tgcactctg tattgtctga atgttgtgt ctccccaaaa tcatcttt
gaaacctaat taccaaagt atattactgg gaggtggggc cttgggagg tggtagatc atgagggtgg agccccatg
aataggatta gtgcccttat aaaagaggcc ctggagagct gccttcccc tccaccaca tgagaacaca gccagcaggt
gcctataagc aagaaagtgg gtctcacca gccatgaat ctgctggtgc attgattgca gacttccag actccagagc
tatgagacat aaattctgt tgtgtataag ccaaaaaaaaa aaaaaaaaaa

Figure 4A

gattcttggg ttccgcgcgt agtttcggaa ggagacatcg aagcagggcg aggcgcagag ggcgttgccg actcatgccc
 cagtcggcag tgcgggggtcc caagccctgc agtgctacag CtttgagCaC acctactttg gccccttga cctcagggcc
 atgaagctgc ccagcatctc ctgtctcat gagtgtttg aggctatcct gtctctggac accgggtatc gcgcgccggt
 gaccctgggt cggaagggt gctggaccgg gcctcctgcg ggccagacgc aatcgaaccc ggacgcgctg ccgccagact
 actcgggtgt gcgcggctgc acaactgaca aatgcaacgc ccacctatg actcatgacg cctccccaa cctgagccaa
 gcacccgacc cgccgacgt cagcggcgcc gagtgtacg cctgtatcgg ggtccaccag gatgactgcg ctatcggcag
 gtcccgacga gtccagtgc accaggacca gaccgcctgc ttccagggca gtggcagaat gacagttggc aatttctcag
 tcctgtgta catcagaacc tgccaccggc cctcctgcac caccgagggc accaccagcc cctggacagc catcgacctc
 cagggtcct gctgtgaggg gtacctctgc aacaggaaat ccatgaccca gcccttcacc agtgcttcag ccaccacccc
 tccccagca ctacaggtcc tggccctgct cctcccagtc ctctgtctgg tggggctctc agcatagacc gcccctccag
 gatgtgggg acagggtca cacacctat tctgtctgt tcagcccta tcacatagct cactggaaaa tgatgttaa
 gtaagaattg cactcctgc cctcggcct tccatctctc ctgccctgt gcccacaac ctggccaaca gtactggaag
 aaactggaca cagtcaccag catcccaggg gagggcaaaa cagccatgtc gtgccctgat gaagagcaat tctgatcaca
 gctgttactc actgagcacc agccaggcac caggcacccc ataacacggc ttctgtgtc ctcttcag agcctgtcgc
 agctctaggg gggagctata caatgatgtc ttatttagtg tcatcatgag aagcccaata agcagtatgc cctaacagtt
 agtaggccag gctctggagc taagctgcat gggttcacat cccagctcca ccattcagcc tgcagagacc atgagcgagt
 tacttaagcc aggctctgga gctaagctgc atgggttcaa atcccagtc cagcattcag cctacagaga ccatgggtga
 gttacttaag ccaggctctg gagctaagct gcatgggttc aaatcccagc tccaccattc agcctgcaga gactgtgggt
 gagttacttg agctctctgt gccaatattt tctacctat aaggtggagg tgaaaataaa ctctataaca tgacaagaac tacttcacag
 tagttgcagt gaggattcaa cgagatgaac atttagtact tgggacacag cagtggcca gtataaatgg gctactgtc
 ataagcccta agtcacaggt caacaaactg agaggtaaaa gcacttggtt gagcttgtgt atctagttag tatggattca
 gggaccagat tcccagcccc acgaactgt aagcaacccc acctcctaaa cacatgagt ccgattaact tcacagaaaa
 acacacaagg caaagttcag cgaggtgaaa ttccaagc tataaagatc agggaagact tctggagga attcaccctt
 gagcaaaatc cttaaaggatc aatagtagct ggcaaaaaga agcaggagga agcacattt aggtagagga gacagcctgg
 acaaaggtct gagggaggaa ggaacacaag gagtgcagga cacttcata actgcatgaa ctcatagag atgggatcct
 ttagcatgtt ctctgtgcac atgcttgacc atgttcttc acatgcttt tgccacttga tcttccagc aactcagtga gagaagcaaa
 aaagtaagtt gcatcctg

Figure 4B

MKLPSISCPHECFEAILSLDTGYRAPVTLVRKGCWTGPPAGQTQSNADALPPDYSVVRG
CTTDKCNAHLMTHDALPNLSQAPDPPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQ
TACFQGNRMTVGNFSVPVYIRTCHRPSCCTTEGTTSPWTAIDLQGSCCEGYLCNRKSMT
QPFTSASATTPPRALQVLALLLPVLLLVGLSA

Figure 4C

ctgagaggaa gttttatctg tgcagccett ctctgaggat ggacacttct cacactacaa agtcctgttt gctgattctt cttgtggccc
tactgtgtgc agaaagagct cagggactgg agtggtacca gtgctatgga gtcccatttg agacttcttg cccatcaatt
acctgcccci accctgatgg agtctgtgtt actcaggagg cagcagttat tgtggattct caaacaagga aagtaaagaa
caatctttgc ttacccatct gccctcctaa tattgaaagt atggagatcc tgggtactaa ggtcaacgtg aagacttcct gttgccagga
agacctctgc aatgtagcag ttcccaatgg aggcagcacc tggaccatgg caggggtgct tctgttcagc ctgagctcag
tctctctgca gaccttgcct tgatggctct cccaatgacc tccacccttg tcttttctc ctcatgtgca acaattcttc ctggagccct
ctagtatga attatgagtt atagaagctc caaggtggga gtagtgtgtg aaataccatg tttgccttt atagcccctg ctgggtaggt
aggtgctcta atctctcta gggctttcaa gtctgtactt cctagaatgt cattttgttg tggattgctg ctcatgaccc tggaggcaca
cagccagcac agtgaagagg cagaattcca aggtattatg ctatccat ccacacataa gtatctgggg tctgcaatg
ttcccatg tctctgaat gtcccctgt tgagtccaat aaacccttg ttcccctaaa aaaaaaaaaa aa

Figure 5A

MDTSHTTKSCLLILLVALLCAERAQGLECYQCYGVVPFETSCPSITCPYPDGVCVTQEAAV
IVDSQTRKVKNLCLPICPPNIESMEILGTKVNVKTSCCQEDLCNVAVPNGGSTWTMAG
VLLFSLSSVLLQTLL

Figure 5B

gctccggcca gccgcggtcc agagcgcgcg aggttcgggg agctccgcca ggctgctggt acctgcgtcc gcccggcgag
 caggacaggc tgccttggtt tgtgacctcc aggcaggacg gccatcctct ccagaatgaa gatcttcttg ccagtgtctg
 tggctgccct tctgggtgtg gagegagcca gctcgctgat gtgcttctcc tgcctgaacc agaagagcaa tctgtactgc
 ctgaagccga ccaatctgtc cgaccaggac aactactgcg tgactgtgtc tgctagtgcc ggcattggga atctcgtgac
 atttggccac agcctgagca agacctgttc cccggcctgc cccatcccag aaggcgtaa tgttggtgtg gcttccatgg
 gcatcagctg ctgccagagc ttctgtgca atttcagtgc ggccgatggc gggctgcggg caagcgtcac cctgctgggt
 gccgggctgc tctgagcct gctgccggcc ctgctgcgtt ttggccctg accgcccaga ccctgtcccc cgatcccca
 gctcaggaag gaaagcccag cctttcttg atcccacagt gtatgggagc cctgactcc tcacgtgcct gatctgtgcc
 ctltgtccca ggtcaggccc accccctgca cctccacctg cccagcccc tgcctctgcc caagtgggcc agctgccctc
 acttctgggg tggatgatgt gaccttctt gggggactgc ggaaggagc agggttccct ggagtcttac ggtccaacat
 cagaccaagt cccatggaca tctgacagg gtcccaggg agaccgtgtc agtagggatg tgtgcctggc tgtgtacgtg
 ggtgtgcagt gcacgtgaga gcacgtggcg gcttctgggg gccatgtttg gggaggagg tgtgccagca gctggagag
 cctcagccc ttagccccc tgcctggca cagctgcatg cactcaagg gcagccttg ggggttggg ttctgccac
 ttccgggtct aggccctgcc caaatccagc cagtctgcc ccagcccacc cccacattgg agccctcctg ctgcttgggt
 gcctcaaata aatacagatg tcccc

Figure 6A

MKIFLPVLLAALLGVERASSLMCFSCLNQKSNLYCLKPTICSDQDNYCVTVSASAGIGNL
VTFGHSLSKTCSPACPIPEGVNVGVASMGISCCQSFLCNFSAADGGLRASVTLLGAGLLL
SLLPALLRFGP

Figure 6B

ccagtctgtc gccacctcac ttggtgtctg ctgtccccgc caggcaagcc tggggtgaga gcacagagga gtgggcccggg
accatgcggg ggacgcggct ggcgctcctg gcgctggtgc tggctgcctg cggagagctg gcgccggccc tgcgctgcta
cgtctgtccg gagcccacag gagtgtcgga ctgtgtcacc atcgccacct gcaccaccaa cgaaacctatg tgcaagacca
cactctactc ccgggagata gtgtaccctt tccaggggga ctccacggtg accaagtcct gtgccagcaa gtgtaagccc
tcggatgtgg atggcatcgg ccagaccctg cccgtgtcct gctgcaatac tgagctgtgc aatgtagacg gggcgcccg
tctgaacagc ctccactgcg gggccctcac gtcctccca ctcttgagcc tccgactgta gagtccccgc ccacccccat
ggcctatgc ggcccagccc cgaatgcctt gaagaagtgc ccctgcacc aggaaaaaaa aaaaaaaaaa

Figure 7A

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREI
VYPFQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLL
PLLSLRL

Figure 7B

aagatggcgg cgtgtggacg tgtacggagg atgttccgct tgcggcggc gctgcatctg ctgctgctat tcgcgcccg
ggccgagaaa ctccccggcc atggcgtcca cagccagggc cagggtcccg gggccaactt tgtgtccttc gtagggcagg
ccggaggcgg cggccccggc ggtcagcagc tgcctcagct gcttcagtc tgcagcttc agcagcaaca gcagcagcag
caacagcaac agcagcttca gccgccgcag ccgccttcc cggcgggtgg gcctccggcc cggcgggggag gagcgggggc
tggtgggggc tggaagctgg cggaggaaga gtctgcagg gaggacgtga ccgcgtgtg ccctaagcac acctggagca
acaacctggc ggtgctcgag tgcctgcagg atgtgaggga gcctgaaaat gaaatttctt cagactgcaa tcattgttg
tggaattata agctgaacct aactacagat cccaaattg aatctgtggc cagagaggtt tgcaaacta ctataacaga
gattaaagaa tgtgctgatg aaccggttgg aaaagggtac atggttccct gcttagtgga tcaccgaggc aacatcactg
agtatcagtg tcaccagtac attaccaaga tgacggccat cattttagt gattaccgtt taatctgttg ctcatggat gactgcaaaa
atgacatcaa cattctgaaa tgtggcagta ttggcttg agaaaaggat gcacattcac aaggtaggtt ggtatcatgc
ttggagaaag gcctggtgaa agaagcagaa gaaagagaac ccaagattca agtttctgaa ctctgcaaga aagccattct
ccgggtggct gagctgcat cggatgactt tcacttagac cggcatttat attttgcttg ccgagatgat cgggagcgtt tttgtgaaa
tacacaagct ggtgagggca gagggtataa gtgcctctt aaccataaat ttgaagaatc catgagtga aagtgtcgag
aagcacttac aaccgcca aagctgattg ccaggatta taaagtcagt tattcattgg ccaaactctg taaaagtgc
ttgaagaaat accggtgcaa tgtggaaaac ctccgcgat cgcgtgaagc caggctctcc tacttgtaa tgtgcctgga
gtcagctga cacagagggc gacaagtcag cagttagtgc cagggggaga tcttgatta ccgacgcatg ttgatggaag
actttctct gagccctgag atcatcctaa gctgtcgggg ggagattgaa caccattgtt ccggattaca tcgaaaaggg
cggaccctac actgtctgat gaaagtagtt cgaggggaga aggggaacct tggaatgaac tgccagcagg cgcttcaaac
actgattcag gagactgacc ctggtgcaga ttaccgcat gatcgagctt tgaatgaagc ttgtgaatct gtaatccaga
cagcctgcaa acatataaga tctggagacc caatgatctt gtcgtgcctg atggaacatt tatacacaga gaagatggtg
gaagactgtg aacaccgtct cttagagctg cagtatttca tctccggga ttggaagctg gaccctgtcc tgtaccgcaa
gtgccaggga gacgcttct gtcttgcca caccacggt tggaatgaga ccagtgaatt tatgcctcag ggagctgtgt tctctgtt
atacagacac gcctaccgca ctgaggaaca gggaaggagg ctctcacggg agtgccgagc tgaagtcaa aggatcctac
accagcgtgc catggatgtc aagctggatc ctgccctcca ggataagtgc ctgattgatc tgggaaaatg gtgcagtga
aaaacagaga ctggacagga gctggagtgc cttcaggacc atctggatga cttggtggtg gagtgtagag atatagttg
caacctact gagttagaat cagaggatat tcaaataaga gccttgctga tgagagcctg tgagccata attcagaact
tctgccaga tgtggcagat aaccagatag actctgggga cctgatggag tgtctgatac agaacaaca ccagaaggac
atgaacgaga agtgtgcat cggagtacc cactccagc tgggtcagat gaaggattt cggtttctt acaagttta aatggcctgc
aaggaggacg tgttgaagct ttgcccac ataaaaaga aggtggacgt ggtgatctgc ctgagcacga ccgtgcgcaa
tgacactctg caggaagcca aggagcacag ggtgtccctg aagtgccga ggcagctccg tgtggaggag ctggagatga
cggaggacat ccgcttgag

Figure 8A

ccagatctat acgaagcctg caagagtgac atcaaaaact tctgttccgc tgtgcaatat ggcaacgctc agattatcga
atgtctgaaa gaaaacaaga agcagctaag caccgcgtgc caccaaaaag tatttaagct gcaggagaca gagatgatgg
accagagct agactacacc ctcatgaggg tctgcaagca gatgataaag aggttctgtc cggaagcaga ttctaaaacc
atgttgcagt gcttgaagca aaataaaaac agtgaattga tggatcccaa atgcaaacag atgataacca agcgccagat
caccagaac acagattacc gcttaaacc catgttaaga aaagcctgta aagctgacat tctaaattc tgtcacgta
tctgactaa ggccaaggat gattcagaat tagaaggaca agtcatctct tgcctgaagc tgagatatgc tgaccagcgc
ctgtcttcag actgtgaaga ccagatccga atcattatcc aggagtcgc cctggactac cgcctggatc ctgagctcca
gctgcactgc tcagacgaga tctccagtct atgtgctgaa gaagcagcag cccaagagca gacaggtcag gtggaggagt
gcctcaaggt caacctgctc aagatcaaaa cagaattgtg taaaaggaa gtgctaaaca tgctgaagga aagcaaagca
gacatctttg ttgaccggt acttcatact gcttgtgcc tggacattaa acaccactgc gcagccatca cccctggccg
cgggctcaa atgtctgtc tcatggaagc actggaggat aagcgggtga ggttacagcc cgagtcaaaa aagcgctca
atgaccgat tgagatgtgg agttacgcag caaagggtgc cccagcagat ggcttctctg atcttgccat gcaagtaatg
acgtctccat ctaagaacta cattctctct gtgatcagtg ggagcatctg tatattgttc ctgattggcc tgatgtgtgg acggtacc
aagcgagtga cacgagagct caaggacagg tagagccacc ttgaccacca aaggaactac ctatccagtg cccagtttgt
acagccctct tgtatagcat cccactcac ctgctcttc tcagaagtga caccaacccc gtgtagagc attagcagat
gtccactgcg ttgtccatc cagcctccac tctgttccat ggtgtctcc tctctccac cgtgcagcag cagcagctgg
tcgtgggggt tactgccttt gttggcaaa cttgggttta cctgcctgta gacaagctc tctcatacca acagaacttc cggacttcc
agaaccaact cacctgacct gcaactcaaa ggcttttta agaaaaccac caaaaaaaaa a

Figure 8A (continued)

MAACGRVRRMFRLSAALHLLLFAAGAEKLPGHGVHSQQGQPGANFVSFVGQAGGGG
PAGQQLPQLLQSSQLQQQQQQQQQQQLQPPQPPFPAGGPPARRGGAGAGGGWKLAE
EESCREDVTRVCPKHTWSNNLAVLECLQDVREPENEISSDCNHLLWNYKLNLTDPKFE
SVAREVCKSTITEIKECADEPVGKGYMVSCLVDRHGNITEYQCHQYITKMTAIFSDYRLI
CGFMDDCKNDINILKCGSIRLGEKDAHSQGEVVSCLEKGLVKEAEEREPIQVSELCKK
AILRVAELSSDDFHLDRLHYFACRDDRERFCENTQAGEGRVYKCLFNHKKFEESMSEKCR
EALTTRQKLIAQDYKVSYSLAKSCKSDLKKYRCNVENLPRSREARLSYLLMCLESAVHR
GRQVSSECQGEMLDYRRMLMEDFSLSPEILSCRGEIEHHCSGLHRKGRTLHCLMKVVR
GEKGNLGMNCQQALQTLIQETDPGADYRIDRALNEACESVIQTACKHIRSGDPMILSCL
MEHLYTEKMVEDCEHRLLELQYFISRDWKLDPVLYRKCCQGDASRLCHTHGWNSETSEF
MPQGAVFSCLYRHAYRTEEQGRRLSRECRAEVQRILHQRAMDVKLDPALQDKCLIDLG
KWCSEKTETGQELECLQDHLDDLVECRDIVGNLTELESEDIQIEALLMRACEPIIQNFC
HDVADNQIDSGDLMECLIQNKHKQDMNEKCAIGVTHFQLVQMKDFRFSYKFKMACKE
DVLKLCPNIKKKVDVVICLSTTVRNDTLQEAKEHRVSLKCRRLRVEELEMTEDIRLEP
DLYEACKSDIKNFCSAVQYGNAQIIECLKENKKQLSTRCHQKVFKLQETEMMDPELDYT
LMRVCKQMIKRFCPEADSKTMLQCLKQNKNSLMDPKCKQMITKRQITQNTDYRLNP
MLRKACKADIPKFCHGILTKAKDDSELEGQVISCLKLRYADQRLSSDCEDQIRIIQESAL
DYRLDPQLQLHCSDEISSLCAEEAAAQEQTGQVEECLKVNLLKIKTELCKKEVLNMLKE
SKADIFVDPVLHTACALDIKHHCAAITPGRGRQMSCLMEALEDKRVRLQPECKKRLND
RIEMWSYAAKVAPADGFSDLAMQVMTSPSKNYILSVISGSICILFLIGLMCGRITKRVTRE
LKDR

Figure 8B

gcgtcgagct cgccgcggac tcaagatggc ggcgtgtgga cgtgtacgga ggatgttccg ctgtcggcg gcgctgcac
 tgctgctgct attcgcgcc ggggccgaga aactccccgg ccagggcgic cacagccagg gccaggggtcc cggggccaac
 tttgttcct tcgtaggga ggccggaggc ggcggcccg cgggtcagca gctgccccag ctgcctcagt catcgagct
 tcagcagcaa cagcagcagc agcaacagca acagcagcct cagccgccgc agccgcctt cccggcgggt gggcctccgg
 cccggcgggg aggagcgggg gctggtgggg gctggaagct ggcggaggaa gagtcctgca gggaggacgt gacccgcgtg
 tgcctaagc acacctggag caacaacctg gcggtgctcg agtcctgca ggatgtgagg gacctgaaa atgaaattc
 ttcagactgc aatcattgt tgtggaatta taagctgaac ctaactacag atcccaaatt tgaatctgtg gccagagagg ttgcaaatc
 tactataaca gagattaaag aatgtgctga tgaaccggtt ggaaaaggtt acatggttct ctgcttggtg gatcaccgag
 gcaacatcac tgagtacag tgcaccagt acattacca gatgacggcc atcatttta gtgattaccg ttaatctgt ggcttcattg
 atgactgcaa aatgacatc aacattctga aatgtggcag tttcggctt ggagaaaagg atgcacattc acaaggtgag
 gtggtatcat gcttgagaa aggcctggtg aaagaagcag aagaaagaga acccaagatt caagttctg aactctgcaa
 gaaagccatt ctccgggtgg ctgagctgc atcgatgac ttctactag accggcatt atatttctg tgcgagatg
 atcgggagcg ttttgtgaa aatacacaag ctggtgaggg cagagtgtat aagtcctct ttaaccataa attgaagaa
 tccatgagtg aaaagtgtc agaagcatt acaaccgcc aaaagctgat tgcacagat tataaagtca gttattcatt
 ggccaaatcc tgtaaaagt acttgaagaa ataccggtgc aatgtggaaa acctccgcg atcgctgaa gccaggctct
 cctacttgtt aatgtgctg gagtcagctg tacacagagg gcgacaagtc agcagtgagt gccaggggga gatgctggat
 taccgacgca tgttgatgga agactttct ctgagccctg agatcactc aagctgtcgg ggggagattg aacaccattg
 ttccgatta catgaaaag ggcggaccct aactgtctg atgaaagtag ttcagggga gaaggggaac ctggaatga
 actgccagca ggcgttcaa aactgattc aggagactga cctggtgca gattaccgca ttgatcgagc ttgaatgaa
 gcttgtgaat ctglaatcca gacagcctgc aaacatata gatctggaga ccaatgatc ttgtcgtgcc tgatggaaca ttatacaca
 gagaagatgg tagaagactg tgaacaccgt ctcttagagc tgcagtattt catctcccg gattggaagc tggaccctgt
 cctgtaccgc aagtccagg gagacgctc tctcttgc cacaccacg gtggaatga gaccagtga tttatgcctc
 agggagctgt gttctctgt ttatacagac acgcctaccg cactgaggaa cagggaagga ggctctcac ggagtccga
 gctgaagtc aaaggatcct acaccagcgt gccatggatg tcaagctgga tctgccctc caggataagt gcctgattga
 tctgggaaaa tgggtcagtg agaaaacaga gactggacag gagctggagt gccttcagga ccatctggat gacttggtg
 tggagtgtag agatatagt ggcaacctca ctgagttaga atcagaggat attcaaata aagccttct gatgagagcc
 tgtgagcca taattcagaa ctctgccac gatgtggcag ataaccagat agactctggg gacctgatg agtgtctgat
 acagaacaaa caccagaagg acatgaacga gaagtgtgcc atcgagatta cccactcca gctggtgcag atgaaggatt
 ttcggtttc

Figure 8C

ttacaagttt aaaatggcct gcaaggagga cgtgttgaag ctttgcctaa acataaaaaa gaaggtggac gtggtgatct
 gcctgagcac gaccgtgcgc aatgacactc tgcaggaagc caaggagcac aggggtgtccc tgaagtgccg caggcagctc
 cgtgtggagg agctggagat gacggaggac atccgcttgg agccagatct atacgaagcc tgcaagagtg acatcaaaaa
 cttctgttcc gctgtgcaat atggcaacgc tcagattatc gaatgtctga aagaaaacaa gaagcagcta agcaccgct
 gccaccaaaa agtatttaag ctgcaggaga cagagatgat ggaccagag ctagactaca ccctcatgag ggtctgcaag
 cagatgataa agaggttctg tccggaagca gattctaaaa ccatgttgca gtgcttgaag caaaataaaa acagtgaatt
 gatggatccc aaatgcaaac agatgataac caagcgccag atcacccaga acacagatta ccgcttaaac cccatgttaa
 gaaaagcctg taaagctgac attcctaaat tctgtcacgg tctctgact aaggccaagg atgattcaga attagaagga
 caagtcact cttgcctgaa gctgagatat gctgaccagc gccgtgttc agactgtgaa gaccagatcc gaatcattat
 ccaggagtcc gccctggact accgcctgga tctcagctc cagctgact gctcagacga gatctccagt ctatgtgctg
 aagaagcagc agccaagag cagacaggtc aggtggagga gtgcctcaag gtcaacctgc tcaagatcaa aacagaattg
 tgtaaaaagg aagtgtctaa catgctgaag gaaagcaaag cagacatctt tgttgaccg gtactcata ctgcttgtgc
 cctggacatt aaacaccact gcgcagccat caccctggc cgcgggcgc aaatgtcctg tctcatggaa gcactggagg
 ataagcgggt gaggttacag cccgagtga aaaagcgct caatgaccg attgagatgt ggagttacgc agcaaagggt
 gcccagcag atggcttctc tgatcttgc atgcaagtaa tgacgtctc atctaagaac tacattctct ctgtgatcag tgggagcatc
 tgtatattgt tctgattgg cctgatgtgt ggacggatca ccaagcgagt gacacgagag ctcaaggaca ggctacaata
 caggtcagag acaatggctt ataaaggttt agtgtgtct caggatgtga caggcagtc agcctgacct ttctgcacac
 tccagacaaa ctcccagac aagtccttt gtgcctctac gtggagaggg tgtggaaagt tatcacatta aaagatggag
 gatttaaaaa aaaaaaaaaa aaaaaaaaaa aaagaaaaaa aaaaaaaaaa

Figure 8C (continued)

MAACGRVRRMFRLSAAHLLLLFAAGAEKLPGQGVHSQGQGPANFVSFVGQAGGGG
PAGQQLPQLPQSSQLQQQQQQQQQQQQPPQPPFPAGGPPARRGGAGAGGGWKLAE
EESCREDVTRVCPKHTWSNNLAVLECLQDVREPENEISSDCNHLLWNYKLNLTDPKFE
SVAREVCKSTITEIKECADEPVGKGYMVSCLVDHRGNITEYQCHQYITKMTAIFSDYRLI
CGFMDDCKNDINILKCGSIRLGEKDAHSQGEVVSCLEKGLVKEAEEREPIQVSELCKK
AILRVAELSSDDFHLDRHLYFACRDDRERFCENTQAGEGRVYKCLFNHKFEESMSEKCR
EALTTRQKLIQAQDYKVSYSLAKSCKSDLKKYRCNVENLPRSREARLSYLLMCLESVHR
GRQVSSECQGEMLDYRRMLMEDFSLSPeiILSCRGEIEHHCSGLHRKGRTLHCLMKVVR
GEKGNLGMNCQQALQTLIQETDPGADYRIDRALNEACESVIQTACKHIRSGDPMILSCL
MEHLYTEKMVEDCEHRLLELQYFISRDWKLDPVLYRKCCQGDASRLCHTHGWNSETSEF
MPQGAVFSCLYRHAYRTEEQGRRLSRECRAEVQRILHQRAMDVKLDPALQDKCLIDLG
KWCSEKTETGQELECLQDHLDDLVECRDIVGNLTELESEDIQIEALLMRACEPIIQNFC
HDVADNQIDSGDLMECLIQNKHKQKDMNEKCAIGVTHFQLVQMKDFRFSYKFKMACKE
DVLKLCPNIKKKVDVVICLSTTVRNDTLQEAKEHRVSLKCRRLRVEELEMTEDIRLEP
DLYEACKSDIKNFCSAVQYGNAQIIECLKENKKQLSTRCHQKVFKLQETEMMDPELDYT
LMRVCKQMIKRFCPEADSKTMLQCLKQNKNSLMDPKCKQMITKRQITQNTDYRLNP
MLRKACKADIPKFCHGILTKAKDDSELEGQVISCLKLRYADQRLSSDCEDQIRIIQESAL
DYRLDPQLQLHCSDEISSLCAEEAAAQEQTGQVEECLKVNLLKIKTELCKKEVLNMLKE
SKADIFVDPVLHTACALDIKHHCAAITPGRGRQMSCLMEALEDKRVRLQPECKKRLND
RIEMWSYAAKVAPADGFSDLAMQVMTSPSKNYILSVISGSICILFLIGLMCGRITKRVTRE
LKDRLQYRSETMAYKGLVWSQDVTGSPA

Figure 8D

caaaaacttc tgttccgctg tgcaatatgg caacgctcag attatcgaat gtctgaaaga aaacaagaag cagctaagca
 cccgctgcca ccaaaaagta tttaagctgc aggagacaga gatgatggac ccagagctag actacaccct catgagggtc
 tgtaagcaga tgataaagag gttctgtccg gaagcagatt ctaaaacat gtgacagtgc ttgaagcaaa ataaaaacag
 tgaattgatg gatcccaaat gcaaacagat gataaccaag cgccagatca ccagaacac agattaccgc ttaaacccca
 tgtaagaaa agcctgtaaa gctgacattc cttaaattctg tcacggtatc ctgactaagg ccaaggatga ttcagaatta
 gaaggacaag tcattctctg cctgaagctg agatatgctg accagcgcct gtcttcagac tgtgaagacc agatccgaat
 cattatccag gagtccgccc tggactaccg cctggatcct cagctccagc tgcactgctc agacgagatc tccagtctat
 gtgctgaaga agcagcagcc caagagcaga caggtcaggt ggaggagtgc ctcaaggta acctgctcaa gatcaaaaca
 gaattgtgta aaaaggaagt gctaaacatg ctgaaggaaa gcaaagcaga catctttgtt gaccgggtac ttcatactgc
 ttgtgccctg gacattaaac accactgcgc agccatcacc cctggccgcg ggcgtcaaat gtctgtctc atggaagcac
 tggaggataa gcgggtgagg ttacagcccg agtgcaaaaa gcgcctcaat gaccggattg agatgtggag ttacgcagca
 aaggtggccc cagcagatgg ctctctgat ctggccatgc aagtaatgac gtctccatct aagaactaca ttctctctgt gatcagtggg
 agcatctgta taltgttctt gattggcctg atgtgtggac ggatcaccaa gcgagtga caagagctca aggacaggta
 gagccacctt gaccaccaa ggaactacct atccagtgc cagttgtac agccctcttg tatagcatcc ccactacct
 cgctctctc agaagtga ccaaccccg gttagagcat tagcagatgt ccactgcgtt gtcccatcca gcctccactc
 gtgctcatgg tgtctctc ctctcaccg tgcagcagca gcagctggc gctggggta ctgcctttgt ttggcagact tggttacct
 gcctgtatgc aagtctctc cataccaaca gaactccgg tacttcaga accaactcac ctgacctga actcaaaggc tttttaaga
 aaaccaccaa aaaaaaaaaat tttttaag aaaaaaatgt atatagtaac gcctctctc caggcttgat ttgggcaatg gggttatgc
 ttcatatga ctgtgtaaaa caaagacagg acttgaggga gaagcacacc acccagtgtg ccatgactga ggtgtctgt
 tcattctca gaagcgcctt ggggcctcgc cagggccgtg gtcttcaccg aggcgtgggt gggcagccgt tcccaggct
 gtgtggggtc ctgcttctt ctgctgagac agtgacgctt tccagttcc accctaatca gccactgctg gtcacagccc
 cacagccatg ggtattctg tggtctctc gcttcattga agcaaagcat gagccttct agacaagggc agctggggag
 gggaaggga cggaagttg tgaagtgaa cagtcctcc atctgcactg agaggctgga tctgagtc cggggcagca
 ggatcccagg aacctctc ctccaggga gcacaggact cagccgtgc tggaccggc ctgctgagc tacagtcact
 ctggaagctc tgcgttcat caggaggcag gactgtggc ggagggggtc ttgaagatgg gtgtggggag cagtgggtca
 ggaagtggga gccagaggtt tgactcactt tgctttattt ttcaggctac aatacaggtc agagacaatg gcttataaag gtttagtgtg
 gtctcaggat gtgacaggca gtccagcctg accttctgc acactccaga caaactccc agacaagctc ctttgtcct
 ctacgtggag aggggtgtga aagttatcac attaaaagat ggaggattaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa

Figure 8E

MMDPELDYTLMRVCKQMIKRFCPEADSKTMLQCLKQNKNSELMDPKCKQMITKRQIT
QNTDYRLNPMLRKACKADIPKFCHGILTKAKDDSELEGQVISCLKLRYADQRLSSDCED
QIRIIQESALDYRLDPQLQLHCSDEISSLCAEEAAAQEQTGQVEECLKVNLLKIKTELCK
KEVLNMLKESKADIFVDPVLHTACALDIKHHCAAITPGRGRQMSCLMEALEDKRVRLQ
PECKKRLNDRIEMWSYAAKVAPADGFSDLAMQVMTSPSKNYILSVISGSICILFLIGLMC
GRITKRVTRELKDR

Figure 8F

aatttagcca gctgtggtgg cacatacctg taatcccagc tacttgagag actgaggcag gagaatcact tgaaccggga
gatggagttt gcagttagcc gagatgggtgc cactgtactc cagcctgggt gacagagcaa gactctgtct ccaaaaaaaaa
aaaaaaaaaa aaaaaaaga catttcaagc tggaagattt gggtccctaa ctttgagcct agctctttca ttaaagtaat aataaaagta
gaactctaca ttatataat ggttttgact ttccaaagtg atttcacat ctacagcagtc ctgtgaagga cttaaataagg tgtttcaggg
tagacttggc attgtgtttt gcaaagaagg tccaaggcca tgcagctatt tggtagacaga attgaaagta aagcctgatt ctcttctgc
aaggcgactt tgctatctag aagccagggt cactagacaa gatgcagtc acaataaagt ctccagaaca tatgacatct
ccagcctaaa ccaagctcac cttccatgc tggctccctc atgcagacgg aggacatccg ctggagcca gatctatacg
aagcctgcaa gagtgacatc aaaaacttct gtccgctgt gcaatatggc aacgctcaga ttatcgaatg tctgaaagaa
aacaagaagc agctaagcac ccgctgccac caaaaagtat ttaagctgca ggagacagag atgatggacc cagagctaga
ctacaccctc atgaggggtc gcaagcagat gataaagagg ttctgtccgg aagcagattc taaaaccatg ttgagtgct
tgaagcaaaa taaaaacagt gaattgatgg atcccaaag caaacagatg ataaccaagc gccagatcac ccagaacaca
ggtaagatct tggcttggct ctctggccc cgtggagtat ctgaaaagga attcagtggtc tgtagagtga cctgtcaca
ctcccagggc ttgttgcct gggaatttta agggaggagt ctgagttaa gcagggcctt cctcctttga ggagcatcca
gaaaaatgga gggagagtca ggggagagag gaggccacaa gaaccagaaa actgccctaa aagaacgttc agaaggaatc
aggccggcag tcttggaaa gaaaaatcta gaaattcaat aaaacttcat gagtgtgcca ggagaatga cgggtaatct
gattcggaac agaaacattt cacctctgag ttggaagacc tctgaagta atggtcacag tgagtggat attgtattc ttttcagt
tttcaaaaag tgtctgtat ggggaagggt gctgatgtcc cttgattt tctgaggact ctttagagta ttggagtctg caaaaaccc
cgcagagtag aaagattcct gaggacctcc agaagtactc gtaacaagt catattgtc attaaaaaca gttagtgag
agctcagtaa atgtttattg aatagataaa tccatgggtg tagtcatgat cattgacata atatgctccc tttaggaagg tggatatcta
aaaatgtgtg aatcaggtgg aatgtttgt cacatgctca ctgcttcta ctctagatta ccgcttaaac cccatgttaa gaaaagcctg
taaagctgac attcctaaat tctgtcacgg tatcctgact aaggccaagg atgattcaga attagaagga caagtcacti ctgcctgaa
gctgagatal gctgaccagc gcctgtcttc agactgtgaa gaccagatcc gaatcattat ccaggagtcc gccctggact
accgctgga tctcagctc cagctgcaat gctcagacga ggtgggattt gcgtgcaaaa ctggttacgc acagagctgc
tcagagaagt ttccactgga gaaaagtgt ttactttctc tccctcagc cgtgaatgat ctggtgaatt gaaggccatc ttctaggctc
tccatggtct gcattctgt tctttgtaac actgaattca acttggcatt agtcttgaca ctctaaagcg ttgtccata tttctgtt
gaacaagggt gttcttcat tatagctctc tgtaaatttg ttctccctt ctcttattc tggatggtaa acccaagacc tgccagaaag
ataaaagtgc tttagctgg gcacgggtggc tcacgcctgt aatcccaaca ctttgggagg ccaaggaggg tggatcatct
gaggtcagga gttcaagacc agcctggcta acatggagaa atctgtctct actaaaaata caaaaaatta gccaggcgtg
gtggcgtgca ccagtaatct cagctactca ggaggctgag gcaggagaat cacttgaacc cgggaggcgg tgggtgcagt
gagctgagat catgccactg cccccagcc tgggcgacag aggaagactc tgtctc

Figure 8G

MQTEDIRLEPDLYEACKSDIKNFCSAVQYGNAQIIECLKENKKQLSTRCHQKVFKLQETE
MMDPELDYTLMRVCKQMIKRFCPEADSKTMLQCLKQNKNSELMDPKCKQMITKRQIT
QNTGKILAWLSWPRGVSEKEFSGCRVTCSNSQGFVAWEF

Figure 8H

ctgagaggaa gttttatctg tgcagccctt ctctgaggat ggacacttct cacactacaa agtcctgttt gctgattctt ctgtggccc
tactgtgtgc agaaagagct cagggactgg agtgttacca gtgctatgga gtccatttg agacttcttg cccatcaatt
acctgccctt accctgatgg agtctgtgtt actcaggagg cagcagttat tgtggattct caaacaagga aagtaaagaa
caatctttgc ttacccatct gccctcctaa tattgaaagt atggagatcc tgggtactaa ggtcaacgtg aagacttcct gttgccagga
agacctctgc aatgtagcag ttcccaatgg aggcagcacc tggaccatgg caggggtgct tctgttcagc ctgagctcag
tcctctgca gacctgtctc tgatggctct cccaatgacc tccaccttg tccttttctc ctcatgtgca acaattcttc ctggagccct
ctagtatga attatgagtt atagaagctc caaggtggga gtagtgtgtg aaataccatg tttgccttt atagcccctg ctgggtaggt
aggtgctcta atcctctcta gggctttcaa gtctgtactt cctagaatgt cattttgttg tggattgctg ctcatgacct tggaggcaca
cagccagcac agtgaagagg cagaattcca aggtattatg ctatccat ccacacataa gtatctgggg tcctgcaatg
ttcccatg taccctgaat gtcccctgt tgagtccaat aaacctttg ttcccctaaa aaaaaaaaaa aa

Figure 9A

MDTSHTTKSCLLILLVALLCAERAQGLECYQCYGVVPFETSCPSITCPYPDGVCVTQEAAV
IVDSQTRKVKNNLCLPICPPNIESMEILGTKVNVKTSCCQEDLCNVAVPNGGSTWTMAG
VLLFSLSSVLLQTLL

Figure 9B

atctgacaga acttgccact gtgcctgcaa cttgtctga gaggaaccct tctctgagga tggacacttc tcacactaca
aagtcctgtg tgctcattct tctgtggcc ctactgtgtg cagaaagagc tcagggactg cagtgtacg agtgctatgg
agtgccaatt gagacttcct gccagcagt lacctgccgc gcctctgatg gattctgcat tgctcaaac atagaattga
ttgaggactc tcaaagaagg aaactaaaga cccgtcagt ctttcttc tgcctgctg gtgtgccaat cagggatcct
aacatcaggg agaggacttc ctgttcagc gaagacctct gcaatgcagc agttcccact gcaggtagca cctggaccat
ggcaggggtg cttctgttca gcctgagctc agtcgtcctg cagacctgc tctgatggc cttccaatga cccccacct tttcttta
tctcatgtg caaccactct ttctggagt cctctagtga caaattatat gttatagaag gtccaatgtg gggatagtgt gtggaacacc
ctgttcacc ttatagccc ctgctgggta agtgcccgac tcctctctag ggcttcaaa tctgtacttc ttgcaatgcc attagtgt
ggattctat tcttgccct ggaggcatgt ggccagcaca tgcaacaggc agtattcaa ggtattatag taccaccatc
cacacataag tatctgggt cctgcagggt tccatgtat gcctgtcaat gaccctgtt gaggccaata aaagctttgt tctccagcc
aaaaaaaaa aaaaaaaaaa aa

Figure 10A

MDTSHTTKSCVLILLVALLCAERAQGLQCYECYGVPIETSCPAVTCRASDGFCIAQNIELI
EDSQRRLKTRQCLSFCPAGVPIRDPNIRERTSCCSEDLNAAVPTAGSTWTMAGVLLFS
LSSVVLQTLL

Figure 10B

gcccaccccc gccagcccc tgcctataag gccttggcaa tgcaggggcc cgcactgctc ccagacgaca tcagagatga
ggacagcatt gctgctcctt gcagccctgg ctgtggctac agggccagcc cttacctgc gctgccacgt gtgcaccagc
tcagcaact gcaagcattc tgttgtctgc ccggccagct ctcgttctg caagaccacg aacacagtgg agcctctgag
ggggaatctg gtgaagaagg actgtgcgga gtctgcaca ccagctaca ccctgcaagg ccaggtcagc agcggcacca
gctccacca gtgtgccag gaggacctgt gcaatgagaa gctgcacaac gctgcacca cccgcaccgc cctcgcccac
agtgcctca gcctggggct ggccctgagc ctctggccg tcctcttagc cccagcctg tgacctccc cccagggaag
gcccctcatg ccttcttc ctttctctg gggattccac acctcttc cccagccgca acgggggtgc caggagcccc
aggctgaggg ctccccgaa agtctgggac caggtccagg tgggcatgga atgtgatga cttggagcag gcccacaga
ccccacagag gatgaagcca cccacagag gatgcagccc ccagctgcat ggaagggtga ggacagaagc cctgtggatc
cccgatttc acactcttc tgtttgtg ccgtttatt ttgtactcaa atctctacat ggagataaat gatttaaacc agaaaa

Figure 11A

MRTALLLLAALAVATGPALTLRCHVCTSSSNCKHSVVCPASSRFCKTTNTVEPLRGNLV
KKDCAESCTPSYTLQGQVSSGTSSSTQCCQEDLCNEKLHNAAPTRTALAHSALSGLALS
LLAVILAPSL

Figure 11B

gctccggcca gccgcgggcc agagcgcgcg aggttcgggg agctccgcca ggctgctggt acctgcgtcc gcccggcgag
 caggacaggc tgccttggtt tctgacctcc aggcaggacg gccatctctt ccagaatgaa gatcttcttg ccagtgtctc
 tggctgccct tctgggtgtg gagcgagcca gctcgtgat gtgcttctcc tgcctgaacc agaagagcaa tctgtactgc
 ctgaagccga ccatctgctc cgaccaggac aactactgcg tgactgtgtc tgctagtgcc ggcatggga atctcgtgac
 atttgccac agcctgagca agacctgttc cccggcctgc cccatcccag aaggcgtcaa tgttggtgtg gcttccatgg
 gcatcagctg ctgccagagc ttctgtgca atttcagtgc ggccgatggc gggctgcggg caagcgtcac cctgctgggt
 gccgggctgc tgcctgagct gctgccggcc ctgctgcggt ttggccctg accgccaga ccctgtccc cgatcccca
 gctcaggaag gaaagcccag cccttcttg atcccacagt gtatgggagc cctgactcc tcacgtgcct gatctgtgcc
 ctggtccca ggtcagggcc accccctgca cctccacctg cccagcccc tgcctctgcc caagtgggccc agctgccctc
 acttctgggg tggatgatgt gaccttctt gggggactgc ggaaggagc agggttccct ggagtcttac ggtccaacat
 cagaccaagt cccatggaca tgcctgacagg gtccccagg agacctgtc agtaggatg tgtgcctggc tgtgtacgtg
 ggtgtgcagt gcacgtgaga gcacgtggcg gcttctgggg gccatgttg gggaggagg tgtgccagca gcctggagag
 cctcagtccc tctagccccc tgcctggca cagctgcatg cacttcaagg gcagccttg ggggttgggg ttctgccac
 tccgggtct aggcctgcc caaatccagc cagtctgcc .ccagcccacc cccacattgg agccctctg ctgcttgggt
 gcctcaaata aatacagatg tccc

Figure 12A

MKIFLPVLLAALLGVERASSLMCFSCLNQKSNLYCLKPTICSDQDNYCVTVSASAGIGNL
VTFGHSLSKTCSPACPIPEGVNVGVASMGISCCQSFLCNFSAADGGLRASVTLLGAGLLL
SLLPALLRFGP

Figure 12B

actgtgcctg caacctggc agagaggaag taaggactgg tgcaggagg gagctgctag gttgatctg tgcagccctt
ctccaaggat ggacagttgt cacactacaa agtcctgtgt actcatcctt cttgtgggcc tattgtgtgc agaaagagct
caggggctgg agtgcataa ctgcctggga gttcacttg gaattgcctg caaatcaatt acctgcccct acccigatgc
agtctgcatt tctcagcagg tagaacttat tgtggactct caaagaagga aagtaaagaa caaactctgc ttcccttct gccctgctaa
tcttgaaaat atggagatcc tgggtactac tgtcaacgtg aatacttct gttgcaagga agacctctgc aatgcaccat ttccactgg
aggcagcacc tggaccatga caaggggtgt tctgttaaat ctgggctcgg tcttctgca gaccttgctg taaaaggctc
tccaaggac ctccaccctt gttgtttat cctcatttg aactattct tcttgagcc cttagtgtat gaattatgag atattgaagc
tccaagggtg gagtagtgt tttggaatac gttgtttcaa ctttatagcc cctgcttgg aaatgcccc ctcctctcta ggaattcaa
atatgtactt cctagaatgc cathttgtg tggcttgcta atcttgccc tggaggccc tggctagcag agggtagagg
cagaattcca aggtattaag ccatcaccat ccacacataa gtgtctgagg ttctgcagga ttctcatgta tgcggcttta tgcaccttg
ttgagtcctaa taaaccttt gttctcc

Figure 13A

MDSCHTTKSCVLILLVVLLCAERAQGLECYNCLGVSLGIACKSITCPYPDAVCISQQVELI
VDSQRRKVKNKLCFPFCPANLENMEILGTTVNVNTSCCKEDLCNAPFSTGGSTWTMTR
VLLNLGSVFLQTLL

Figure 13B

ctgcagccag gtctgagagg aagtaaggac tgggtgcagg agggagctgc taggtgacaa agggaagaac cctcaggata
gggctgtggt gggagtgaga ttaggaaaga agagctgggt ggggtggtgga tgagagaagt aggcagacat gtattcctca
gggaaagctg ttagaggggt tggagggagg gaatattgga tggctgagcc gtgtgagagc ccaggggtgt gatcaggggt
ctattaactg gctccaactt ccaagggttt atctgtgcag cccttctcca aggatggaca ctctcacga gataaagtcc tgtgtgctga
tccttctgt gaccctactc tgtgcagaaa gagctcaggg actggagtgt taccagtgt atggagtccc atttgagact tctgccccat
cattacctg cccctacct gatggattct gtgtgtctca ggaggaagaa ttattgcaa actctcaaag aaagagagta
aagagccgtt ctgccatcc ttctgccct gatgaaattg aaaagaagt taccctggat ctaacacca agatgaatat ttctgttgc
caggaagacc tctgcaatgc agcagttccc actggaggca gctcctggac cacggcaggg gtgcttctgt tcagcctggg
ctcagtcctc ctgcagacc tgatgtgatg gtccccaccc

Figure 14A

MDTSHEIKSCVLILLVTLLCAERAQGLECYQCYGVPPFETSCPSFTCPYPDGFCAQEEEFI
ANSQRKRVKSRSCHPFCPDEIEKKFILDPN TKMNISCCQEDLCNAAVPTGGSSWTTAGV
LLFSLGSLVLLQTLM

Figure 14B

gttatcagag gtgagcccggt gctcttcagc ggagaagatc ccctacctgg ccgccggcca ctttctgtgg gccgtggggt
 cctcaaggag acggcccttg ggctcagggg ctgcgtttcc acacgcgcct ttcccagggc tcccgcgccc gttcctgcct
 ggccgccggc cgctccaaca gcagcacaag gcgggactca gaaccggcgt tcagggccgc cagcggccgc gaggccctga
 gatgaggctc caaagacccc gacaggcccc ggccgggtggg aggcgcgcgc cccggggcgg gcggggctcc ccctaccggc
 cagacccggg gagaggcgcg cggaggctgc gaaggttcca gaaggcggg gagggggcgc cgcgcgctga ccctccctgg
 gcaccgctgg ggacgatggc gctgctgcc ttgctgctgg tctggccct accgcgggtg tggacagacg ccaacctgac
 tgcgagacaa cgagatccag aggactcca gcgaacggac gagggtgaca atagagtgtg gtgtcatgtt tgtgagagag
 aaaacacttt cgagtgccag aaccaagga ggtgcaaatg gacagagcca tactgcgtta tagcggccgt gaaaatattt
 ccacgtttt tcatggttc gaagcagtgc tccgtggtt gtgcagcgt ggagagaccc aagccagagg agaagcgggt
 tctctggaa gagcccatgc cttctttta cctcaagtgt tgtaaaattc gctactgcaa ttagagggg ccacctatca actcatcgt
 gttcaaagaa tatgctggga gcatgggtga gagctgtggt gggctgtggc tggccatcct cctgctgctg gcctccattg
 cagccggcct cagcctgtct tgagccacgg gactgccaca gactgagcct tccggagcat ggactcgtc cagaccgttg
 tcacctgttg cattaaactt gtttctgtt gattacctct tggttgact tcccagggtc ttgggatggg agagtgggga tcaggtgcag
 ttggctctta accctcaagg gttctttaac tcacattcag aggaagtcca gatctctga gtagtgattt tggtgacaag ttttctct
 tgaaatcaaa ccttgtaact catttattgc tgatggccac tctttcctt gactccctc tgcctctgag ggcttcagta ttgatgggga
 gggaggccta agtaccactc atggagagta tgtgctgaga tgcttccgac ctttcaggtg acgcaggaac actgggggag
 tctgaatgat tgggggtgaag acatccctgg agtgaaggac tctcagcat ggggggcagt ggggcacacg ttagggctgc
 ccccatcca gtggtggagg cgctgtggat ggctgctttt cctcaacctt tctaccaga ttccaggagg cagaagataa
 ctaattgtgt tgaagaaact tagacttcac ccaccagctg gcacaggtgc acagattcat aaattcccac acgtgtgtgt tcaacatctg
 aaacttaggc caagtagaga gcatcagggt aaatggcgtt catttctctg ttaagatgca gccatccatg gggagctgag
 aaatcagact caaagtcca ccaaaaacaa atacaagggg acttcaaaag ttcacgaaaa aattgaatta aaagataaaa attaa

Figure 15A

MRLQRPRQAPAGGRRAPRGGRGSPYRPDPGRGARRLRRFQKGGEGAPRADPPWAPLGT
MALLALLVVALPRVWTDANLTARQRDPEDSQRTDEGDNRVWCHVCERENTFECQNP
RRCKWTEPYCVIAAVKIFPRFFMVAKQCSAGCAAMERPKPEEKRFLLPEPMPFFYLKCC
KIRYCNLEGPPINSSVFKEYAGSMGESCGGLWLAILLLLASIAAGLSLS

Figure 15B

ggagagagca ggacacagct atggatgccg ccaggagagg agatacacag ccagtgatgt ggaccaccgg atggctgttg
 ctgctgccgc ttctgtgtg tgaaggagcg caagccctggagtgtacag ctgcgtgcag aaggcggacg atggatgctc
 tccgcacagg atgaagacagtcaaatgtgg tcccgggggtg gacgtctgta ccgaggccgt gggagcggta
 gagaccatccacgggcaatt ctctgtggcg gtgcggggct gcggttccgg aatcccgggc aagaacgaccgcggactgga
 ccttcacggg ctctggcct tctttcagct acagcagtgc tccgaggaccgatgcaacgc caaactcaac ctcaatttgc
 gaggcctcaa ccctgcaggc aatgagagtgcatatgagcc taacgggtgca gagtgtctaca gctgtgtggg tctgagccgc
 gagaagtgccagggctccat gccgccggtc gtgaactgct acaacgccag tggccgtgtc tacaagggctgcttcgatgg
 taacgtcacc ctgacggcag ccaacgtgac cgtgtcctta cctgtccgaggctgcgtcca ggacgagacc tgcacccggg
 atggggtgac gggtcagga ttacactcagcggctcttg ctgtcagggc ccccgctgta acgccgacct tcgcaacaag
 acctatttctccccctgaat cccacccta gtctgtgtgc cccctccaac caccgcagcc ccatccactcgggcccagaa
 ctctccagc acgacctta cagcagcccc aaccaagacc acctccatcatcaagcccac cacagcccaa gccagccaca
 ctctccca tgaatggat ctgaagtcatacaggaaga gggggcgctg ttgagtggag gtgctgcggg ccatggaggt
 actgcgggccatggaggtgc tgcgggccac caagaccgca gcaatatgga gaagtatcca ggaaagggtg gggcccagat
 cccagctaaa ggaggctctg gcactctagg gtcttggtg tctgcagttc tgttgactgt ggttgctggc gcgatgctgt
 gaatgtctca tctgaaaag tccatctcac ttgtttccc tggccccgtg gtaccaactc ttccatttc tcaattgact ggactggctc
 cgccccatc cttcagcatt ctcaattccg actgcactgg ttgcagctt cggaaaacag tcctctgtt taaatattcc gctcgggtgg
 ccctactttt ttgatgcggc cacagcattc cccctgatgg tgaccaggac agagggaaga gacgtctact ggctgagaga
 ggcccagaga gtccacggca agcctcctct tcccgtttc ctgaccaggc tggaagatga ccaggcaggt agacaatgga
 tccatcctcc gagcactgtg cttgcctggc acattgtgcg gaaatctggt cgcctgtctt ccttaggaga ctgtgaacaa
 ctclacaaca gggcttgtc tctggcctct ctatgtgtc tgtctggcac aggaagggtgt caataaagat ttagttactt tgaatagtga
 gttaactaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa

Figure 16A

MDAARRGDTQPVMWTTGWLLLLPLLLCEGAQALECYSCVQKADDGCSPHRMKTVC
GPGVDVCTEAVGAVETIHGQFSVAVRGCGSGIPGKNDRGLDLHGLLAFFQLQQCSEDR
CNAKLNLTLRGLNPAGNESAYEPNGAECYSCVGLSREKCQGSMPPVVNCYNASGRVY
KGCDFDGNVTLTAANVTVSLPVRGCVQDETCTRDGVTGPGFTLSGSCCQGPRCNADLRN
KTYFSPRIPLVLLPPPTTAAPSTRAQNSSSTTSTAAPTSTTSIIKPTTAQASHTSPHEMDLE
VIQEEGASLSGGAAGHGGTAGHGGAAGHQDRSNMEKYPGKGGAQIPAKGGSGTLGSW
LSAVLLTVVAGAML

Figure 16B

tcgtccccc cgcctgccc gccgtgagc ccggagtgcg gacaccccag ggatgcctgc gcccagagg accccgcct
 gcagcccccg cgcctcttc aggcctatc ggagcatgctgcctgcagcc atgaagagcc tcggtctggc gctgctggcc
 ttgttctct gcccctcgccggcccatggc ctgtggtgcc aggactgcac cctggccaat tccagccatt
 gcgtccgaagcagtgccag cccaccgata ccgtttgtgc cagcgtgcgg atcaccgacc ccagcagcagcaggaaggat
 cattctgtga acaagatgtg tgcttctcc tgcgacttcg itaagcggcactttttctca gactatctga tggggttcat taactctggg
 atcttaaaag tcgacgtggactgctgcgag aaagatttgt gcaacggggc atcggtcgca ggacgcagcc
 cctgggccctggctgggggg ctctgtctca gcctggggcc tgctcttctc tgggctgggc cctaagaccctccctccct
 cctgctgggc ttggagctt gtcccctaag cctgttgctg cccctcccagcctggcctg gctggggctg ggacagcaag
 gggttggcat caaggtctga ggctctcaaccctccctagat gtgagtgagc ctctccgtt tctccaccag ctccatatcc
 caagcagctgaatatctcca ggagtccaga catctggca ggaagctggg gtagggggga gggggagggaaggactga
 gacctccag gtctccaagg ggaggaggt caagccaggg acagcccaacagcctggcct gaggggcatt aactacagag
 aaataaagtc acttctgagt cttgtgaaaaaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa

Figure 17A

MPAPQRT PACSPRASFRPYRSMLPAAMKSLGLALLALLCPSPAHLWCQDCTLANSSH
CAPKQCQPTDTVCASVRITDPSSSRKDHSVNMKMCASSCDFVKRHHFFSDYLMGFINSGLK
SRRGLLRERFVQRGIGRRTQPLGPGWGAPVQLGAWLFFGLGPKTPPSLLVGFGACPLSL
LVPLPSLAWLGLGQQGFGIKV

Figure 17B

agaagaggcg agactttttt ggggtgctccg gatcgccagt agttcttcaa gcctcagcag ccaactcctc cggaggcgct
 gcgctccgcc ccagggagcg cgaatccaag gagcctgggaccagcctctg ggagcccccg gcgcggggcga tgcggggcgcc
 gcggggcgaca cctgcgggctcctctcggtgg cagccgtcgc ttgggcggca gcagcgcgag cctcggcagc
 ctggcagctactgtcgccg cggccagaac agcctccgct gcggctcgtgg tctctgatgc tcttgcccgtcccggccct
 gccgatccgg gaggatgtgg gttctcggca tcgcagcaac ttttgcggattgttctggc ttccagggtt ggcgctgcaa
 attcagtgt accagtgtga agaattccagctgaacaacg attgtctatc cctgagttc atcgtaaatt gcaccgtgaa
 cgttcaagacatgtgtcaga aagaagtgt ggagcaaagt gctgggatca tgtaccggaa gtcgtgtgcatcgtcagcag
 cctgtctcat tgcctcagct gggtaccagt ccttctgttc cctgggaaactgaactccg tgtgcatcag ctgctgcaac acccctctt
 gcaatgggcc gaggcccaagaagagaggca gctctgcctc ggccatcagg ccagggttc tcaccactct
 cctgttcttccacttagccc tctgcttggc aactgtctga agctaaagga gatgccaacc cctgctgcctcacctgtctg gcccttcgtc
 tctcaccttc ccgagtctct tctgggtgtc ctttattctgggtagacaa gggagtctt ttgtccctc ctttaagta acgcaagatt
 gccgtgcacaaacttttg taagctctga accaattcat tctgaatttc tgtgtgtagt tgaagaaaaagcatggagc agaaagtcca
 gacctcca tcccaatctg gtaaccacc gccaaaggta gcctggaaga accagccctt agaagtcatt gagatacgca
 tctgccttc ccaaagcctt gagcttccat tctgtcccag taggagtcac agtctattca gagactgtctg ctgctgtaag gtaactttgc
 ttttgcggga ggggagagcc agtttcggct caaggcttct gaacttgcca ttcatacttc ctgctctgt aaactatttt ctgggggtgga
 cccagctggt ttggtctctg agccagtctg tgggtactca ggactcaagg gctggggctt agcctctcca ggcttggcct
 cagtctgaaa agtgcttaag aaaaccttgt tagttctct ggaggaagag ttactgcgcc gggaggctag gaagatgagg
 gggctgcggg ctgagctggt gctgtccttg gtggagatga agcgggcacg ctggcgcttc tcttggttgg catgctgcag
 agtcaggcgg cagcagagca cctgccagaa caccttcgg aactgtctgag aggacacgtt gtagaggaga gggttgacca
 cagagctgag gtagaagaag gtatcagaga agggcaggag gatcatgtat gccctgaagt acgttctggt ccagtcattg
 ttgggttttg ctgcagccat gatccgtcgg atctgattgg gcatccaaca cacggccaac gtcaccacaa tcagtccttg
 caggcaagaa caggagagaa aaggagacgg ggagagaaac agcatgagaa caaaaataaa taaataaaaa cccataaaat
 attaagcccc ttggttctgt tgcctacttg ccgagaaacg gtaccaatct ttcagctctg tgcctgtcgg ctctttttg ccaactggcaa
 aggagaattt aatgtctct caagctcagg ggacttggct atgttaaaaa gcgttaaatg ctttcgacag tgtatttata cttacggctg
 cctgttaatt ttcaaatgt ttcatgtt gctcgtgtat ccagaaaata tctcacgttg gccaaaa

Figure 18A

MWVLGIAATFCGLFWLPGLALQIQCYQCEEFQLNNDCSSPEFIVNCTVNVQDMCQKEV
MEQSAGIMYRKSCASSAACLIASAGYQSFCSPGKLNSVCISCCNTPLCNGPRPKKRGSSA
SAIRPGLLTLLFFHLALCLAHC

Figure 18B

ggcaggcctg agtgaggacc tcgacatgc aggggacctg gatggtgctg ttggcactga tattgggcac ctgcggggag
cttgctatgg ccttacagtg ctacacctgt gcgaatcctgtgagtgcac caactgtgtc accaccaccc actgccacat
caatgaaacc atgtgcaagactacgctcta ctccctggag attgttttcc ctttcctggg ggactccacg gtgaccaagtcctgcgccag
caagtgtgag ccttcggatg tggatggcat tgggcaaacc cggccagtgtcctgctgcaa ttctgacctg tgcaacgtgg
atggggcacc cagcctgggc agtcctgggtggcctgctcct tgccttgga cttttcttgc tcttgggtgt cctgctgtaa
agccatggccatctagctcc actcccttgt ccctgacatc ccagttccct aatgcctaga agaaatacaatggccatctg
caaaaaaaaa aaaaaaaaaa aaaaaa

Figure 19A

MQGTWMVLLALILGTFGELAMALQCYTCANPVSASNCVTTTHCHINETMCKTTLYSLEI
VFPFLGDSTVTKSCASKCEPSDVDGIGQTRPVSCCNSDLCNVDGAPSLGSPGGLLLALAL
FLLLGVLL

Figure 19B

atgcttttta tggcaggccc tgcagccagc tggccctga ggccctggg actccatggc gtccccaag ccttgtgtgc
tgtccttta acagtgtgg tcatgaagac cttggtctcggtgatacca agctcgagga ccttcacct cagtcctcc cactaaaca
gtacctgaattgctaccgat gtctgtgga gaccgaagag ctgggggtgcc tcctgggggc tgacacctgcctgacacctc
tgggcagcag ctgtgtcacc ctgcacataa agaacagcag cggttttaatgtcatggtga gcgactgcta cagcaaggag
cagatgggcc attgttcata taccgtgtctccccgggtgt ttggctttg gatattctat cagtgttgct tcctggattt
ctgaacaatccggacaaca gaaagaatag catgcactag

Figure 20A

MLFMAGPAASWSLRPLGLHGVPQALCAVLLTVLVMKTLVLGDTKLEDLHPQSLPLNKY
LNCYRCLLETEELGCLLGSDTCLTPLGSSCVTLHIKNSSGFNVMVSDCYSKEQMVHCSY
TRASPVFGFWIFYQCCFLDFCNNPDNRKNSMH

Figure 20B

agagctggag acctgggaat ctgctgtcaa ctgctggggc tgtggacatt ctcaggacce tcacatgaa acacctcctg
ttgtcaccc tgtctgcct actctactgc tgggtctcagctgatactcg atgtcactcc tgctacaaag tcctgtgtg gggctgtgtg
gatgccagtcctgccgcct ggagccgggc cacaatgcc tgacaacaaa cgtgtacctt gggaagatgtgggtttctc
taacctgcgc tgcggcacac cagaagagcc ttgtcgggag gtcttcaacgaaaccaacca taagctgggc ctgaactaca
acaccacctg ctgtgacaag gataactgtaacagcccggc tccacggccc acaccgcac tggccctcat ctcctcacc
tccttggctggcctcggcct ctggttattg cattgagact agtccatgg ctacaatctt accacctgctatagcctgag cttttctcc
tgtgtctca gagctccagc ttccagaat ctctctctcccacccct tcttctgaag atcatgtccc tagtcctata ccatttatt
catgggactgtacctggagt ggcctttcta gccaccgctc ctctcctca ctgtcacct tccactccattccaccaca cacagacaca
cagacacaca gacacaaaga cacacacaca cacacacacacacacaca cccagtcctt tccatttcc ttctagaaca
ctctacctcc tccactggccactgaaaggc tccctcctt ggacgcacac tgtgtgcct ctgggatcta agtctggaagaactcctgtc
ttgtctccag ggagtgttc caaaaggcgc tggcctcatt gcatgggcctggcttaccag accctctgt tgtcccttc tatcttgaga
aataaacatc agtgtctaat

Figure 21A

MKHLLLLTLSALLYCWVSADTRCHSCYKVPVLGCVDRQSCRLEPGHKCLTTNVYLGK
MWVFSNLRCGTPEEPCREVFNETNHKLGLNYNTTCCDKDNCNSPAPRPTPALALISLTS
LAGLGLWLLH

Figure 21B

gcctacttgg cctgcctgcg atgcggtacc aacaccgcac gaagtgtgta cagattccca gttagacagc aggagggacc
 tgggagcggc caggggggatg ttttatctct aagagaccaagagctcaggc agggcttctg tgccctgctt cctccctggc
 ttgagctgga tcttgaccagctgctgacc tctgttcac tctggcactg cctcacgtc tccgtcatga cccatctgctcagtgctc
 ctggtggccc tgatgggcct gcctgtggcc caggctctgg agtgccacgtgtgtgcctac aatggagaca actgcttcaa
 acccatgcgc tgcccagcca tggccacctactgtatgacc acacgaactt acttcacccc ataccggatg aaggtgagga
 agtcctgtgtcccagctgc ttgaaaccg tgtacgatgg ctattccaag catgcatctg ccacctctgttgccagtac tacctctgca
 acggtgctgg ctttgcctacc ccggtgacct tggcccttggtcccagcactc ctagctacct tctggagctt gctgtaaagc
 tccgttcccc aagccagatccactcaaacg caacactctc aaaaaacaca gtttccctct ctctcccaat tcactccaccaacgctctt
 ccttctgaca ctctcaact accacgaggt cccatggcta cctacgaaagaactgatggc atccagatac ctactccaa ggtcatttc
 agaaggctga catgtggacctgtaatgtgc ccacccatgg gtggggcagg ctgggcttct cctctacca
 agatcaggggcatctgggag aatgtttatg gaggaggggt catcactcaa gtcaaggagc actgattgatagaattagt
 agccaaactc caccttcaga accctgcctc agtctacca gtagaggatgggtctgctag aggtgagggg aggagagcgg
 cggagaataa cgagctggct agaagcagag aaagactcag cagggtctgc tccgaagatc agcgcggctt gccagagcaa
 atgtgatgtg gaagccatgt gaggaagccc ttgtcattt ccacttatct gaggaactct gccagacctg atgttgggat
 agccattggc caagggttcc tagcaacggc gtcatttcca taggacctg aaatccctcc agccccagct cagcaggccc
 cttgacctcc actacagtcc ttattcaca caccagctgc tgggccttga agttggcagg gacttgggag caggtgacct
 atgctatttt ttgtctggcc tgttattctg ggcattggcaa gaagggatca gacgcaggtc agagcagggc agtagggcga
 ctgagacagg gaaacagact tcagccagtg gcttcccagg tccgtaggc agtcctaca tcttcagtc tctgttaca
 tccccgggag acaaatafac agggagccaa gccgagtgt aggtgatgac tgccctgtgaa gtctattgtg gccacagact
 gctgggtacc aagtctcagg agaaccacgc ctagatttag gagacacaga tctgccttcc atgcagtgtg gctgtccttg
 ggagccttac catgctctct aactagtacc tcaactcaca tgtcactgag gaaccccta acactggccc agcccagggg
 tccggatgct ggccaatgtc catggagtgg gactacccat ggagagtct tgggtcatca catcacaat gttttattcc
 aacctcccag tggtagagac tccggacaca aaggtccatc ctggggacct tcttctggt tctaggcaga cctgaactct
 gtctgctgct agagctgatg tggttttccg cctcagtttc ctctccggg gataggccac cggaggattt gggaggggtg
 ggagggcatc ctgctgatgg gctcggcag gttctcagga acaggaacgg gcggggcttt agtacacagg tgagtgggt
 gggaactggc ccggagctga ggagacactg actgggcaga gggaagatga gtctcaaggg agggcaggaa aaggagggg
 gagcgcgcat gcacatgtgc actcagtga ggctacagag cccaaaaggc agcactggct gtggtgtccc ctgaggccca
 ggcaagatgc taggaggaag ccaatgctgc cccacctga gtcacatgg aacatgcaca ccaccagcag cagcagcaag
 cattgagact gacctgtgga cgccataggg cactggcaag gagggtcaga ggcgggtccc tgactcagtg ggtgaggccc
 gggaacatt atctgttac cctgcgtgtg caagatcatt gtccccagct

Figure 22A

agatggcgtc ctcaacaaa actgagagga gccccagttc aggtcctccc tcctaccaca aggggggtgt gtggaggagg
 cttgattgcc cttggagaag caccggtact gcagagctgg gggccagctt cttcatctg tgtctagaca ccgaccagat
 aggccccaca gtggcaacac tgccacacag ccctacaaga agccctgtgc ctagctagca cagagcccca aaaggtgctc
 aattaataca gggccaagcc tgccagtggg ggggatgcag attaggggaa cagacccaga tggcctgtcc tgaaccctgt
 ctgggggtgt gtgatgagca tctgtctagc ccactgcagg tggctctaca cactccacaa cagttctgca aaagtgtatg
 aggtggtcat tactgcgcc ctctcacagg taaaggcact gaggcacgga ggagtgaggc acttcatttt cctgggccat
 tcaactttcc aggaccaaca cattcaacta tgggtactac tccaatagct ggggttcttt gaggctgggc cccctgaaga
 tgatagtggc ttcataacc agagaatttc agagtgcagt gttgtaggag cctatgaacc tgaaatgtca gaactggagg
 tttagggggc tgaggggtag gccaggggtg tctggcccct tgtgtggaga cagagagaga gggaacatgg gatggggtag
 tagagagaag tgcaaaggag cgtcagcctt tctcagggt aatgctgtca gggacgaggg ctcaagcctg tgagtgtct
 cacactgtga taaacagtgg cccctcaaca cagacggtgt ccagagtggc cggcagtggg tatctagagt tgcaatctgg
 aagcctcttg gtagtactg gagagaggcc gcttgatggg acagcaccaa atgtgtgtgc ttctgtggga tgtgaggaag
 ctgggtcagc gcatgaagcc aaagcgtcct tcagagcaga ggggtggctg gtctagtcca ccagagacaa gctatccagt
 gagagtata ctctgccacc gtctctgtga ttacctacc ccaaagcaga cggggacggg atgcagagca cccgtgtctt
 catcttctgc ggcaagcacg tgagttcaca ttctgaaact ctagaagat ttccaggagt ggggtgtgcc ttgtcttgg tgcattgta
 ctctctggca agcaccgtgg catcccgcag cactgagtga cctgggctcc tcaagccatc tcattgtga aatgacagt
 ccagtaccct ctacgtggc tcttgaggc ctgtgcatgg ggtctgcaca gaggaggccc ccaaactatg catggacgga
 cacgtgatgc ctacacttc ccttggttgt gtctctgcca accccaggct ctaccacgc aaggaaatga aatccacttt
 tatgacacat ctccctccc cagccagctc cattaccta tatgccaggg tggctccctt caatgtctgt ccccatgg atgaataaac
 aagcgaagga ca

Figure 22A (continued)

MTHLLTVFLVALMGLPVAQALECHVCAYNQDNCFKPMRCPAMATYCMTTRTYFTPYR
MKVRKSCVPSCFETVYDGYSKHASATSCCQYYLCNGAGFATPVTALVPALLATFWSL
L

Figure 22B

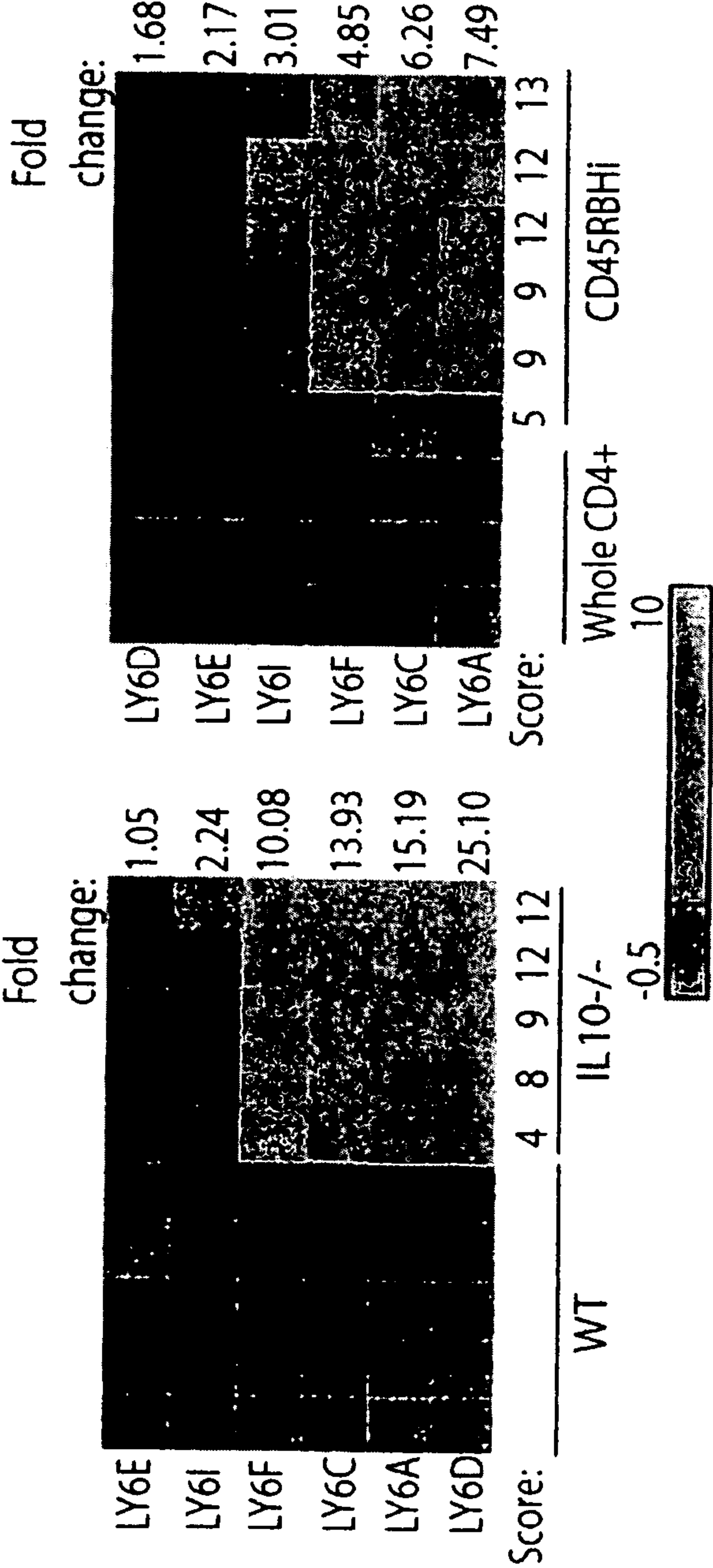


Figure 23 A

Figure 23 B

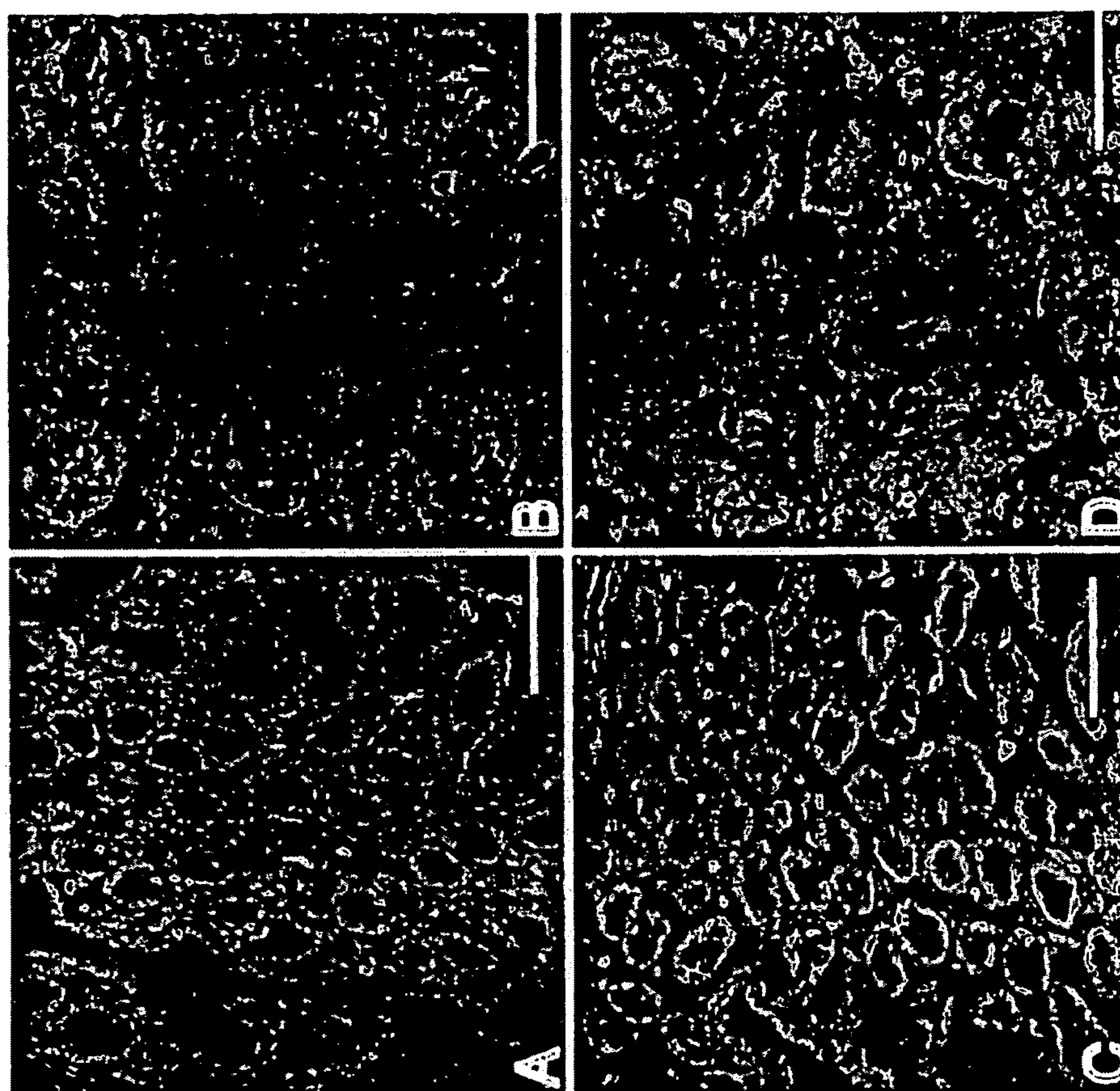


Figure 24

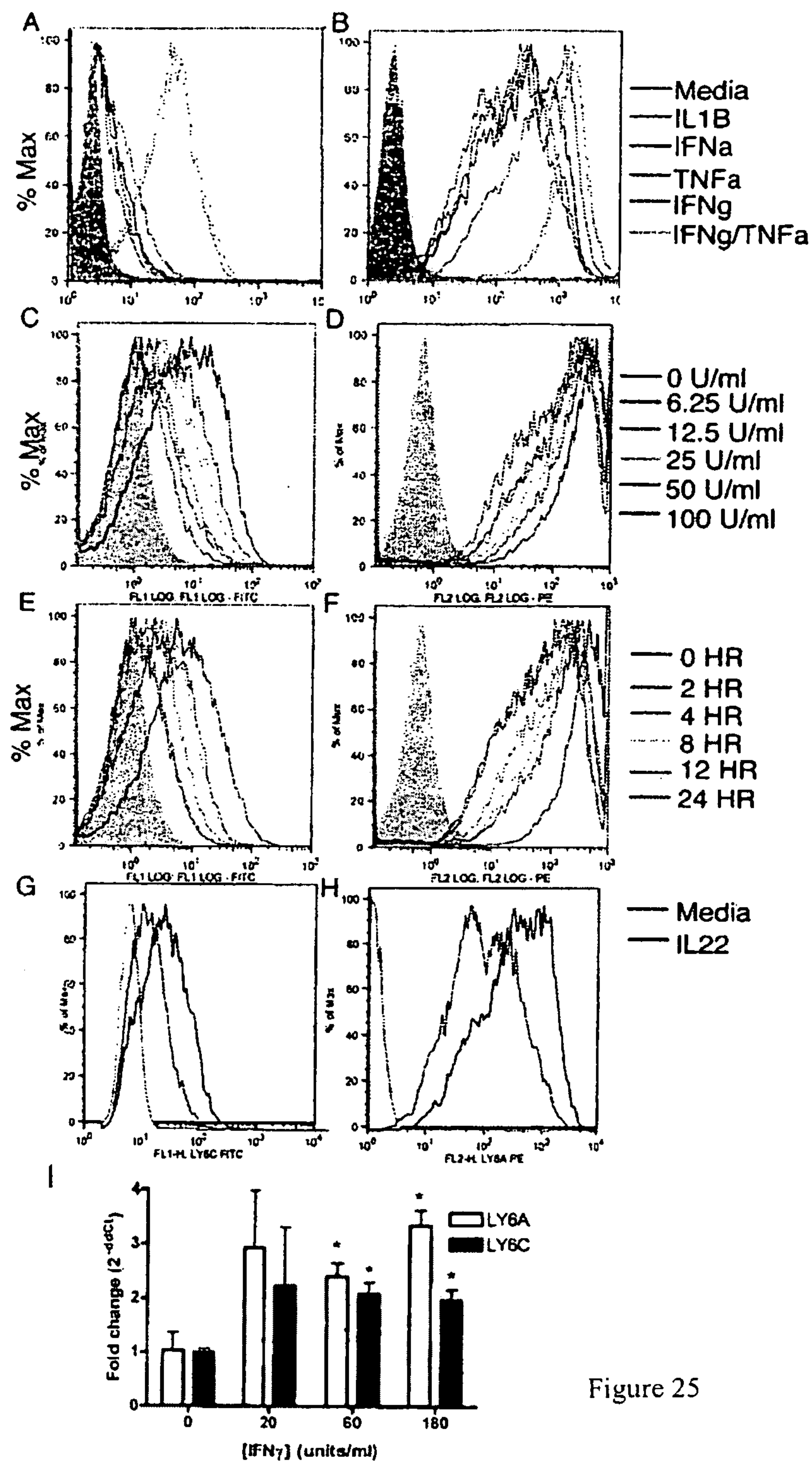
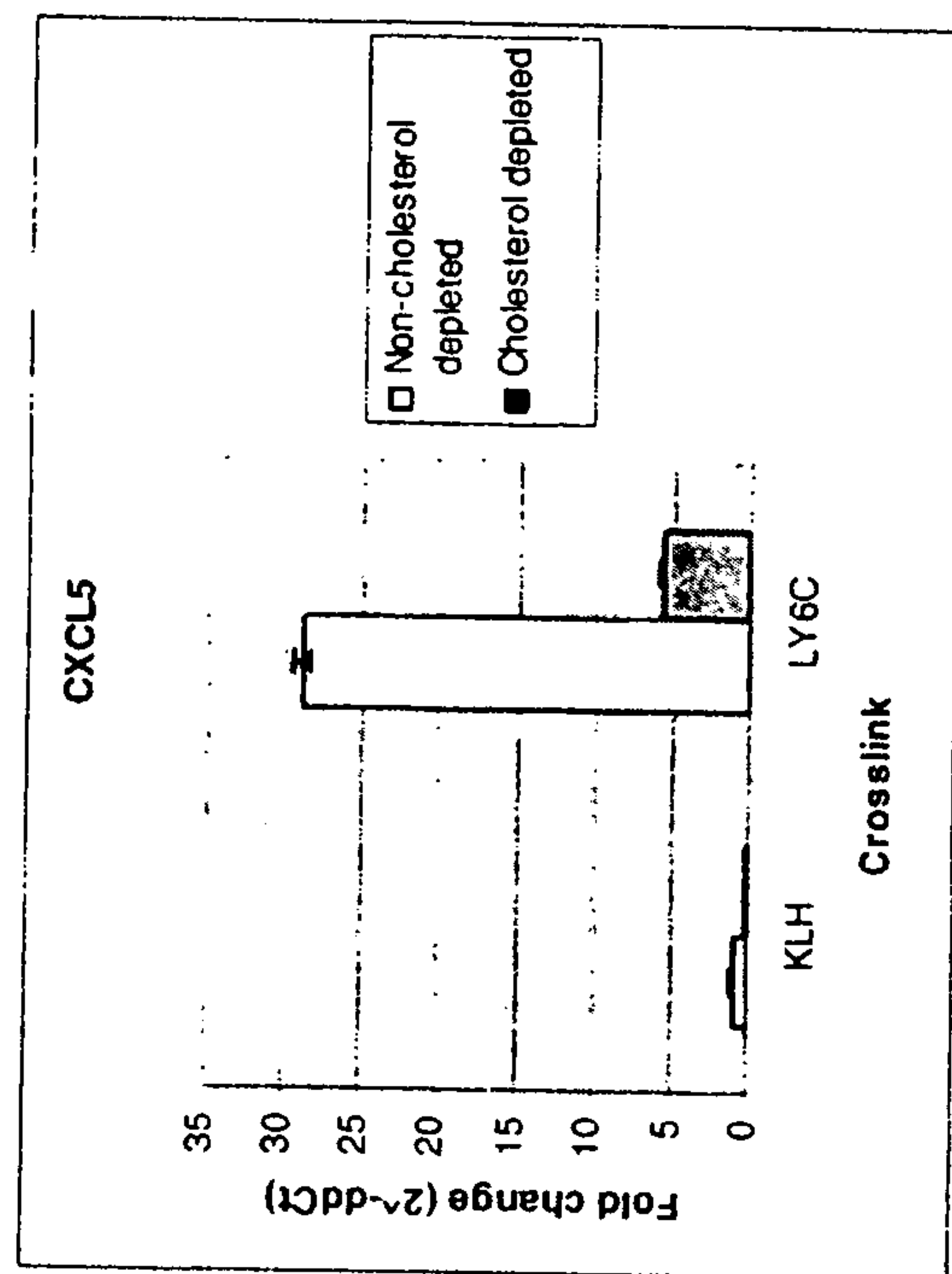
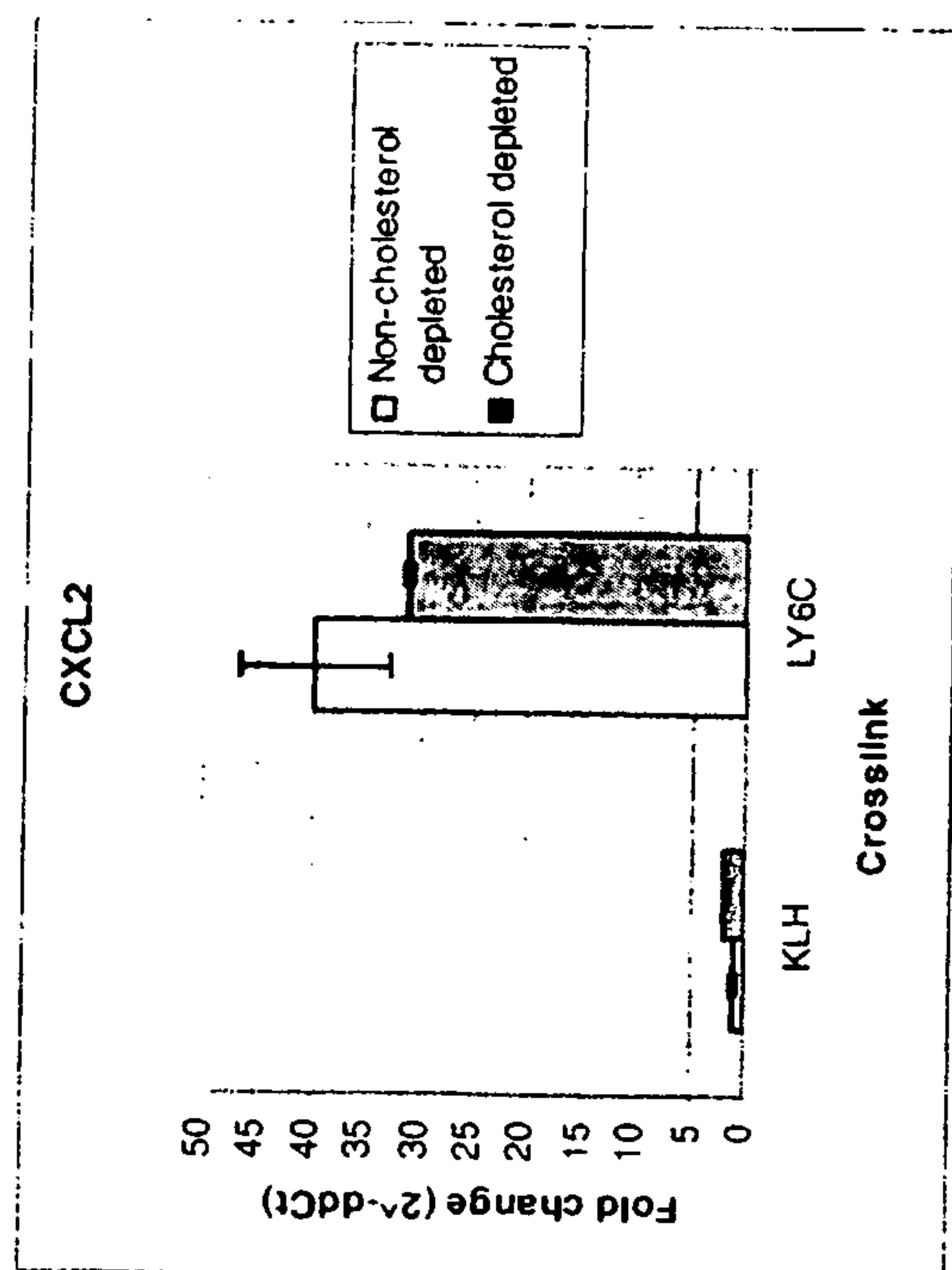


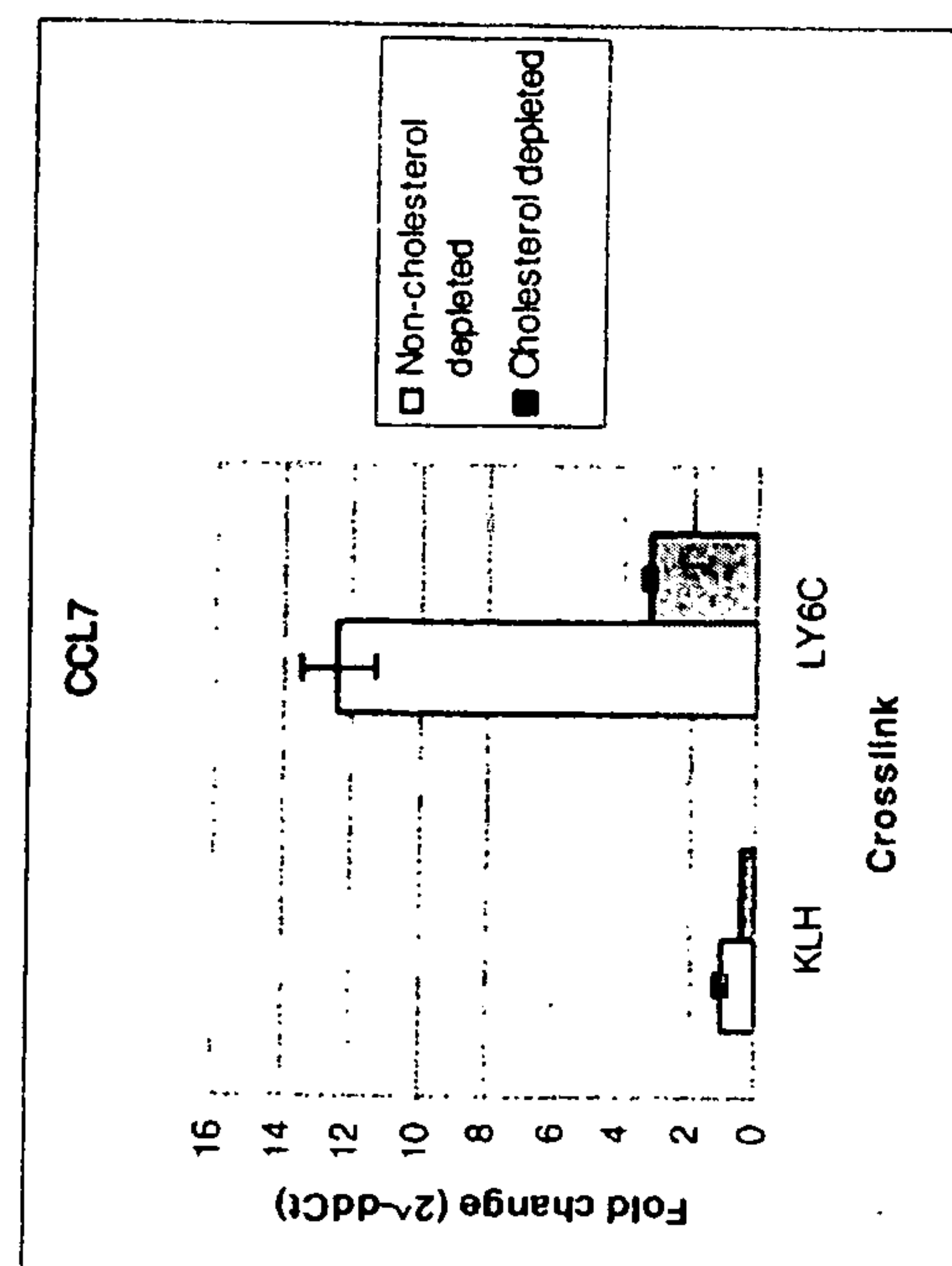
Figure 25



B



A



C

Figure 26

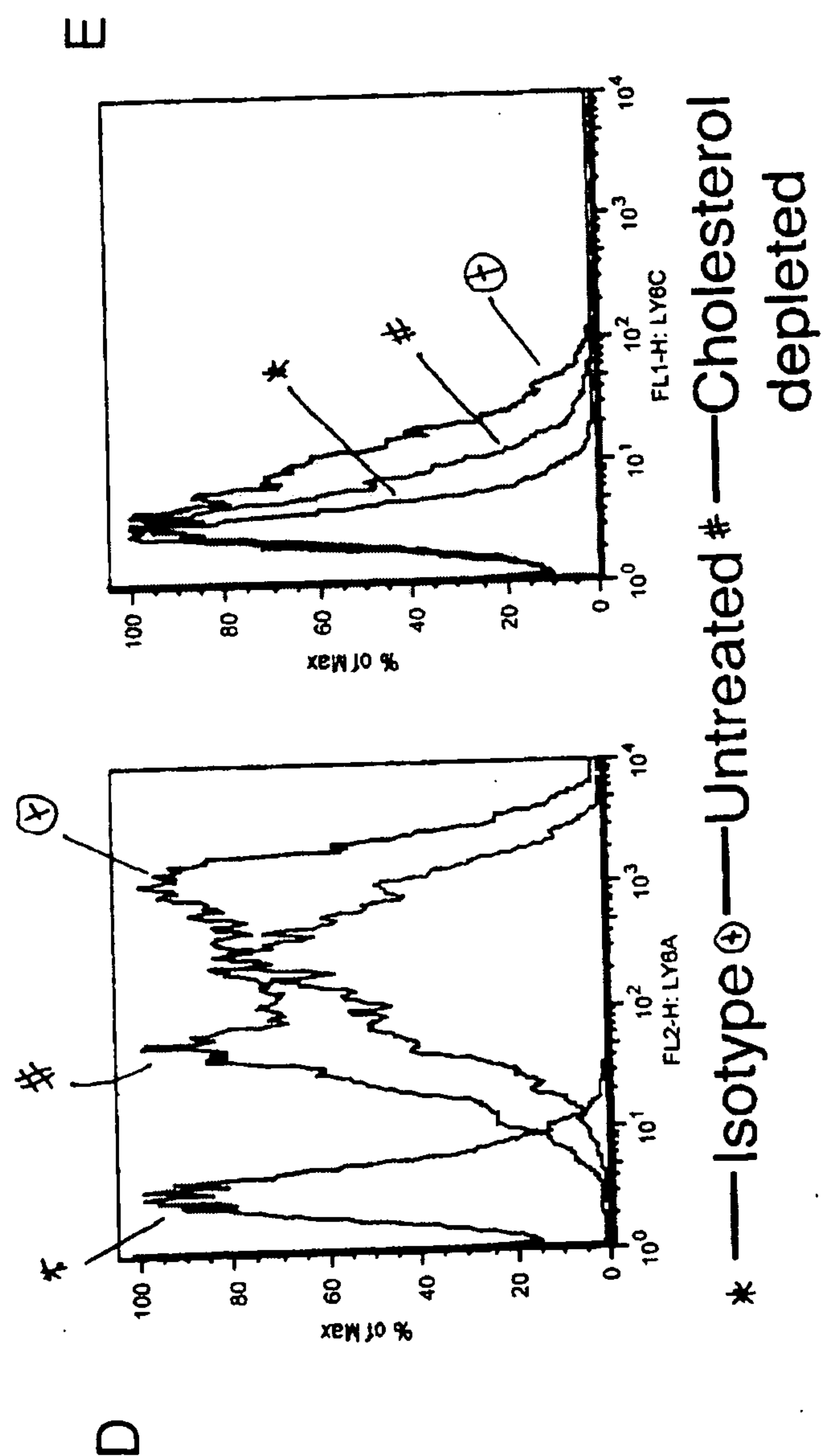


Figure 26 (cont)

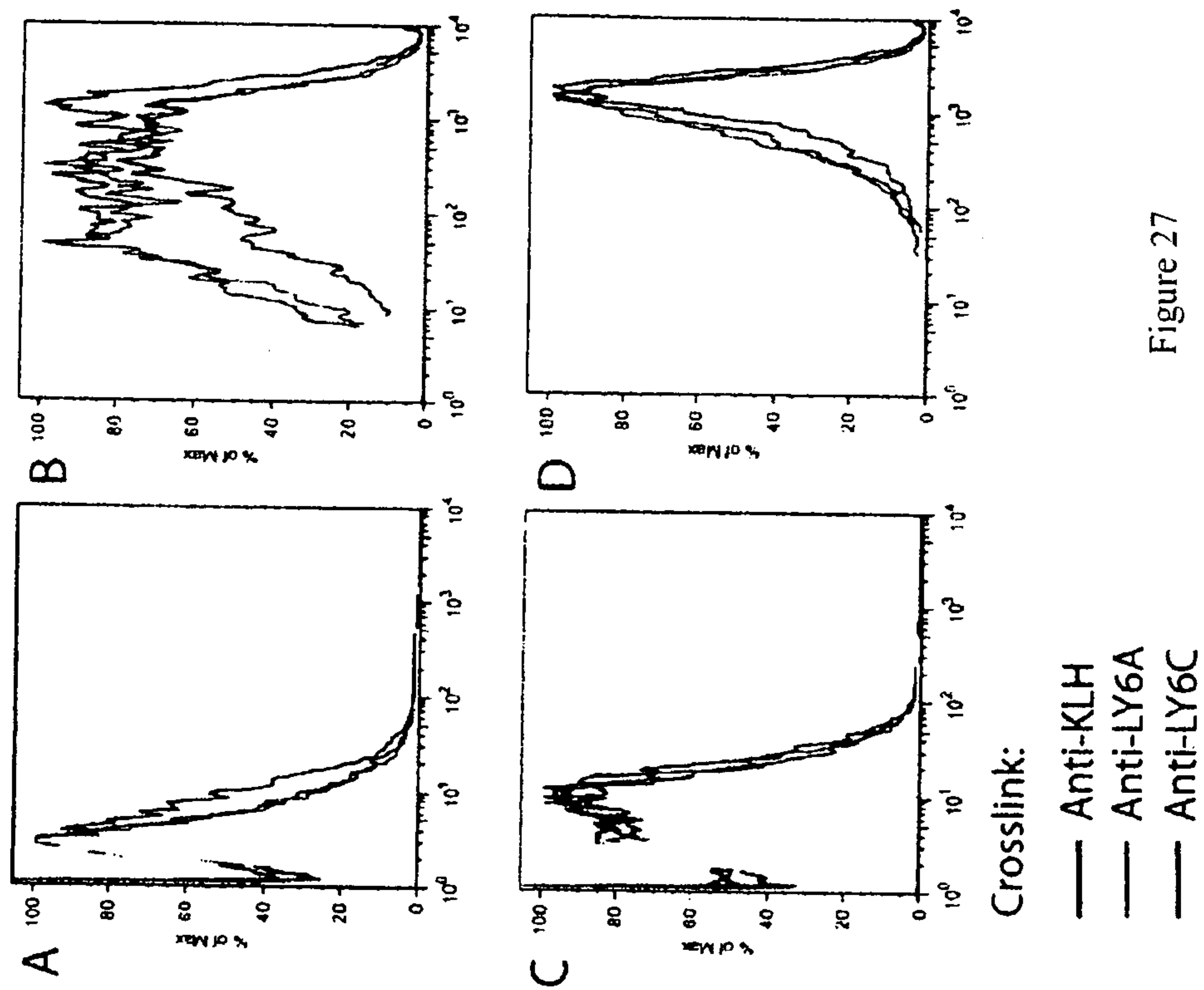


Figure 27

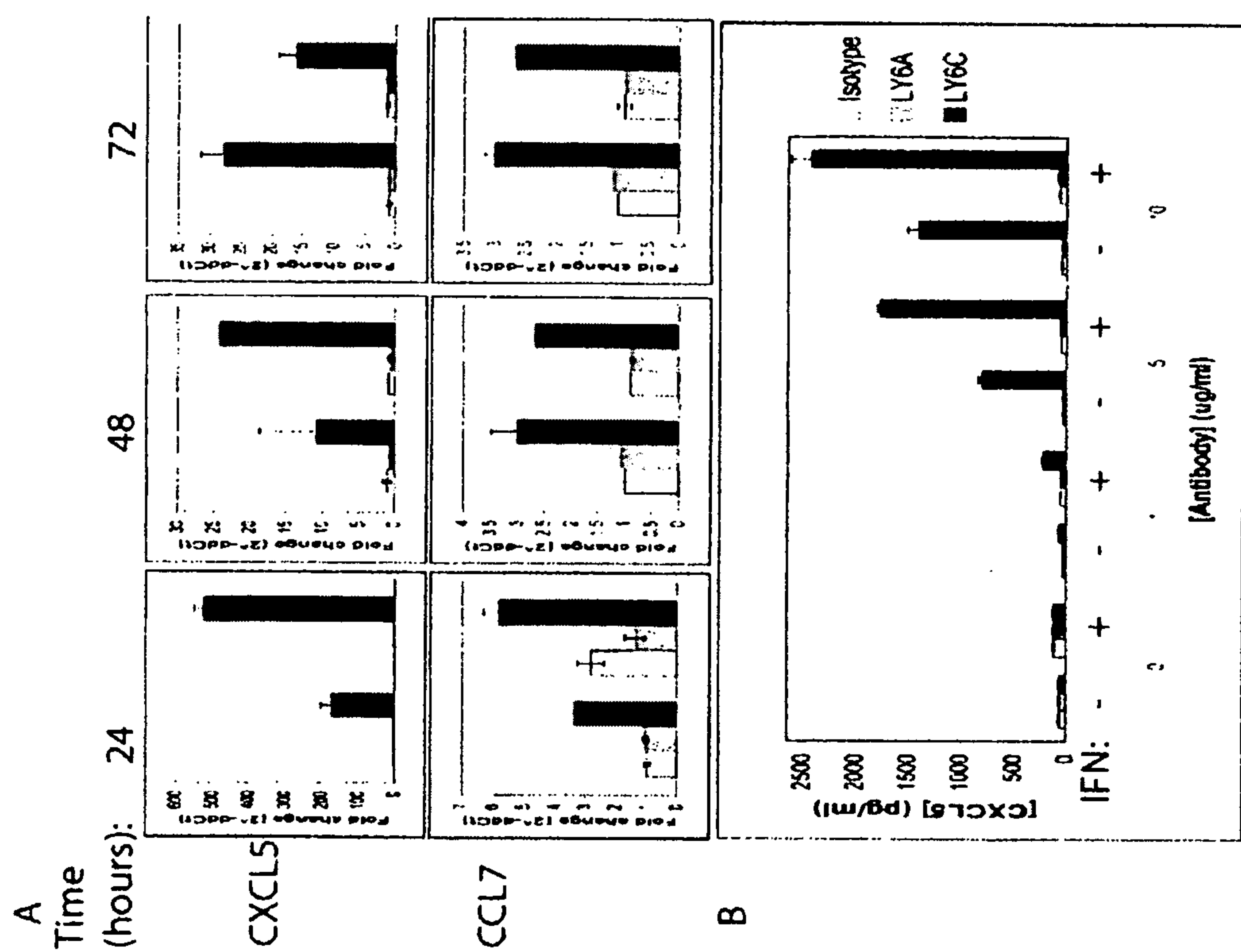


Figure 28

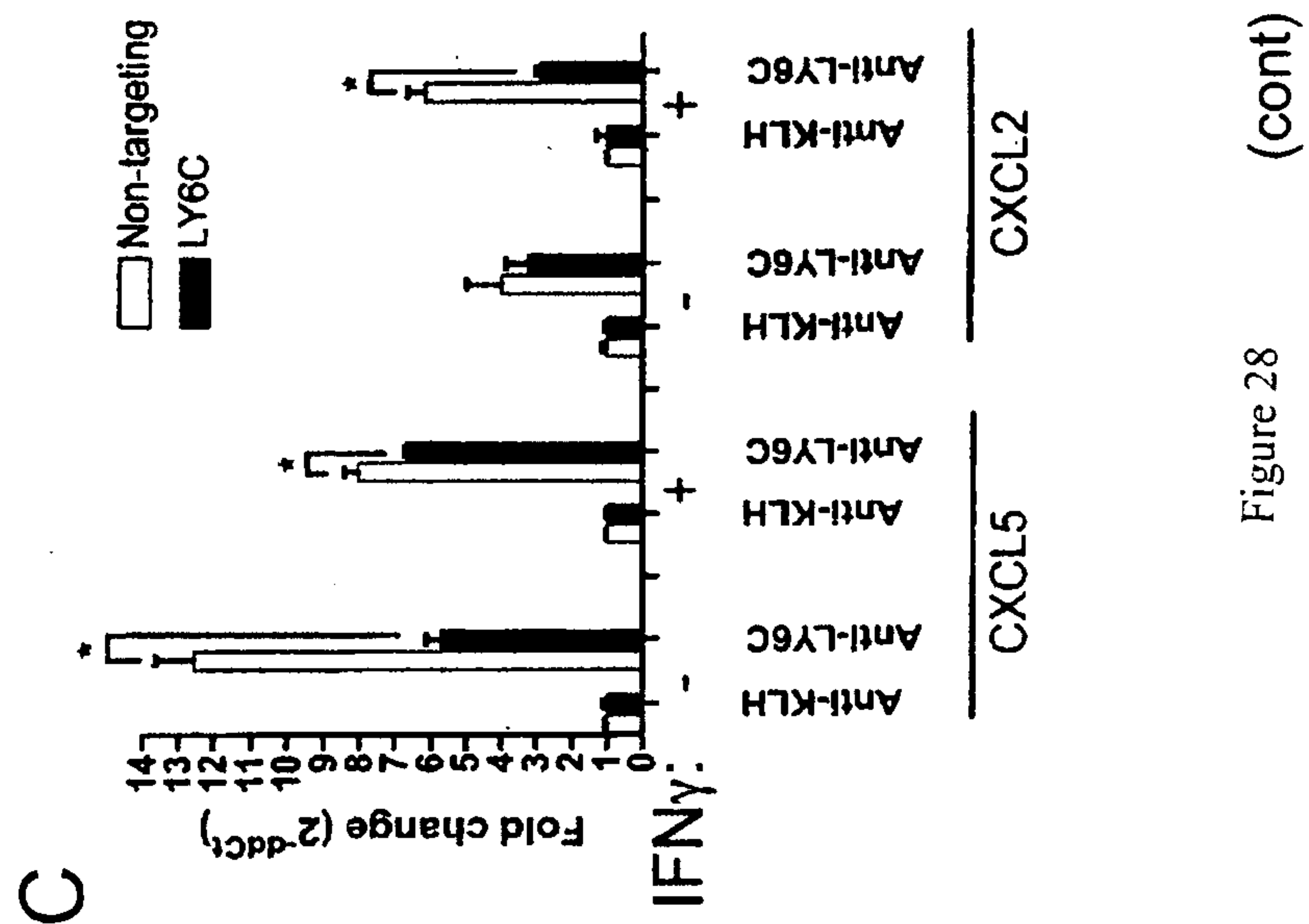


Figure 28 (cont)

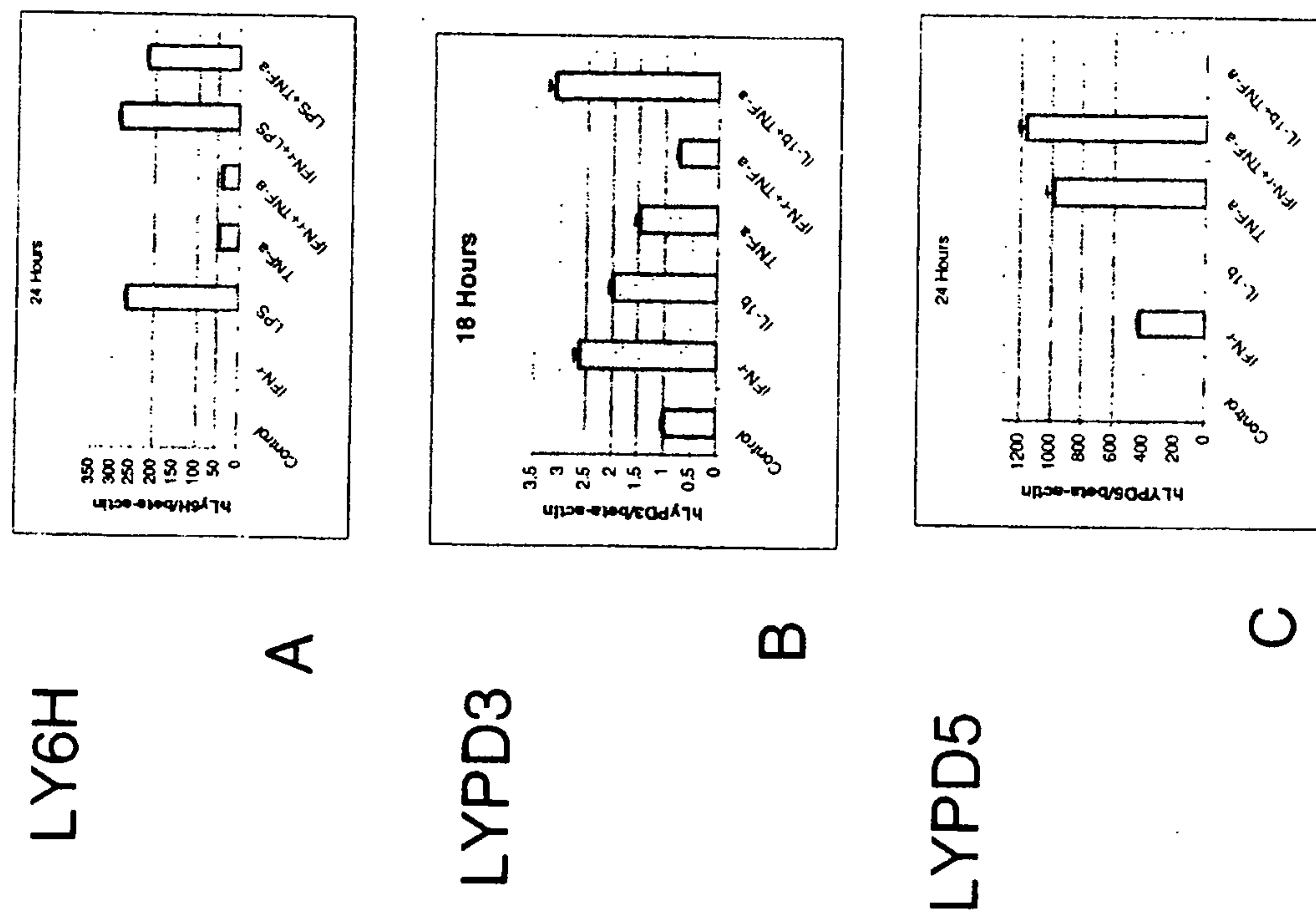


Figure 30

LYPD1

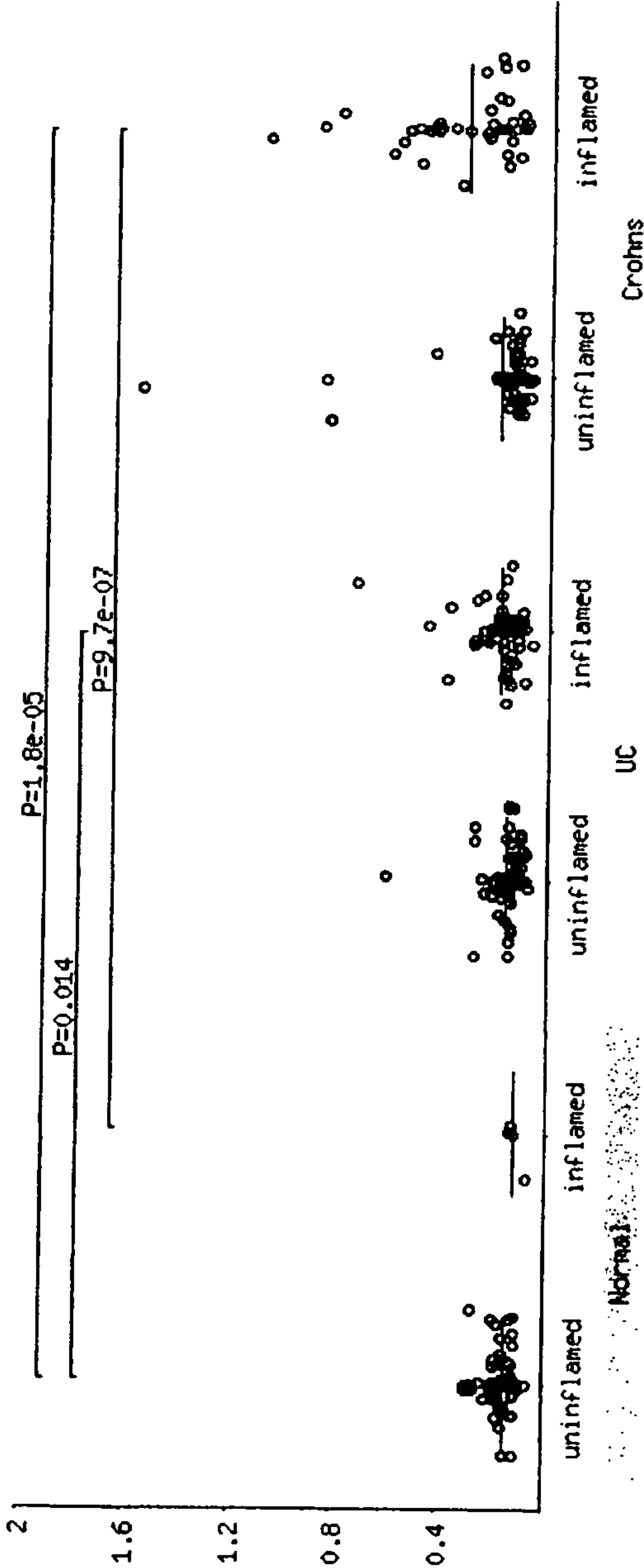


Figure 31 A

LYPD5

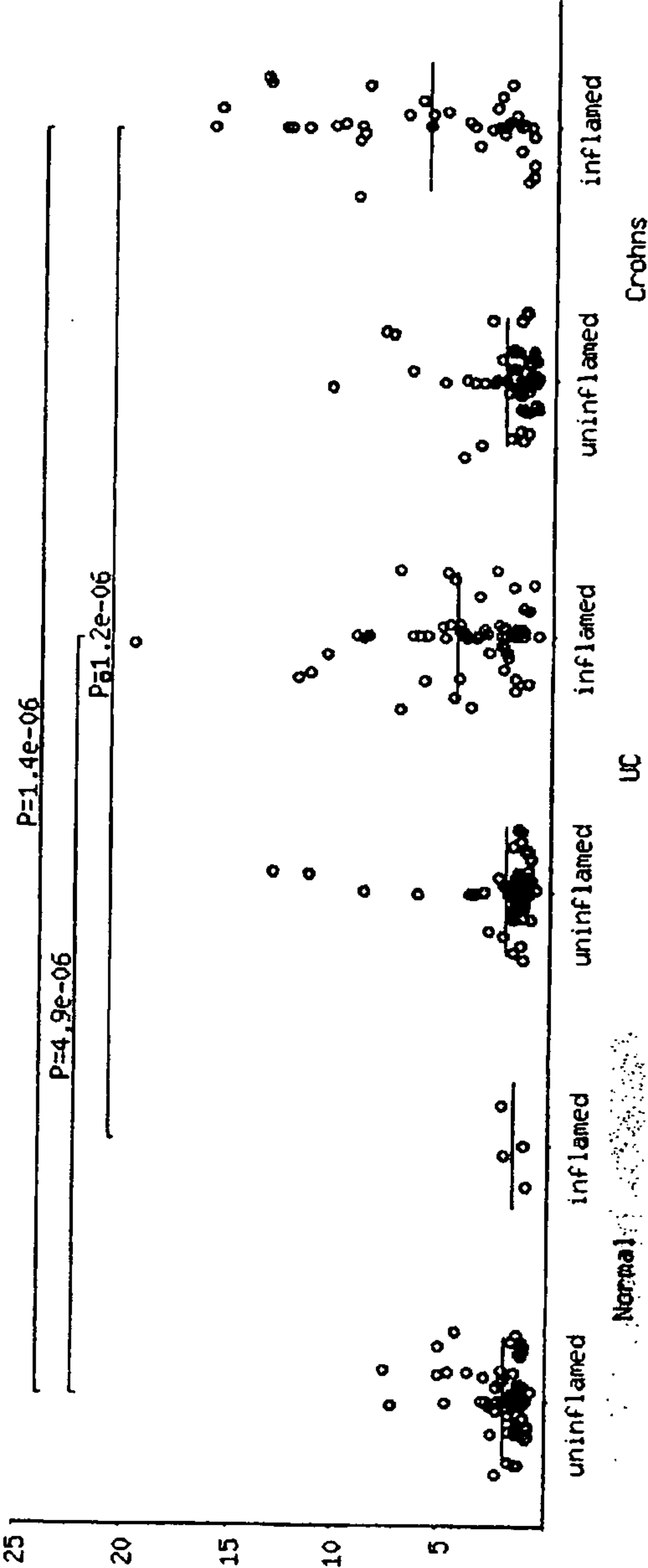


Figure 31 B

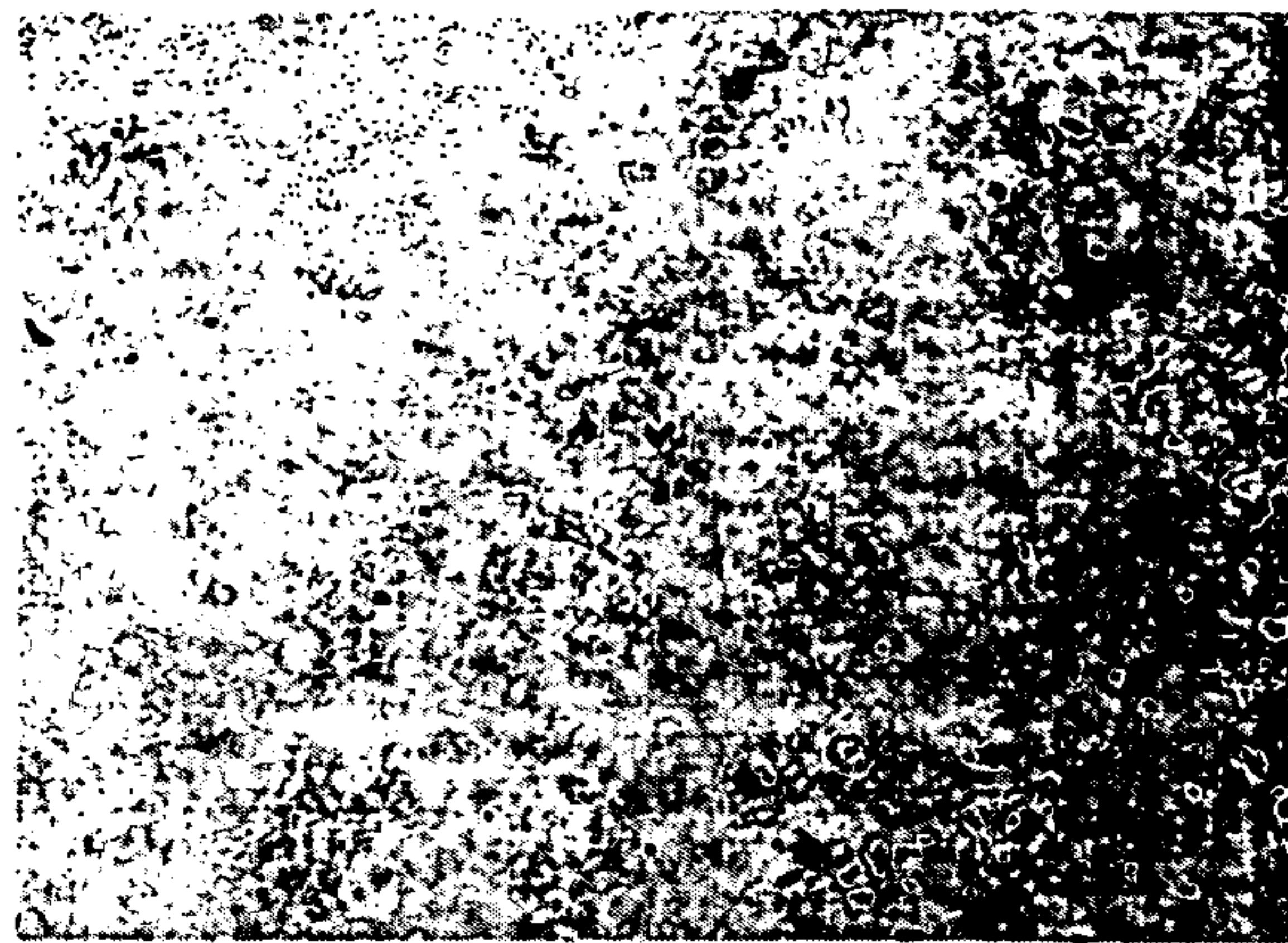


Figure 32A



Figure 32B

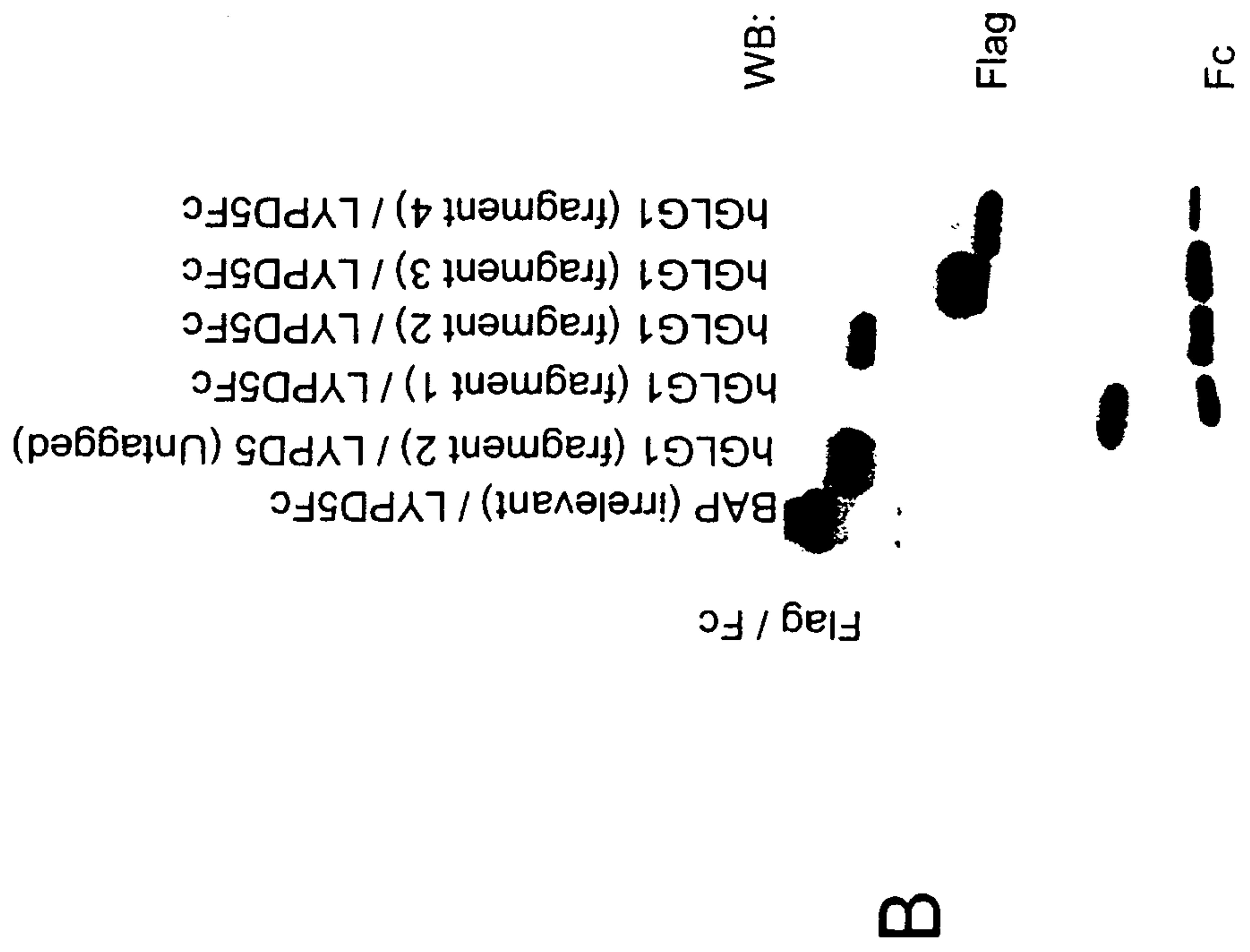
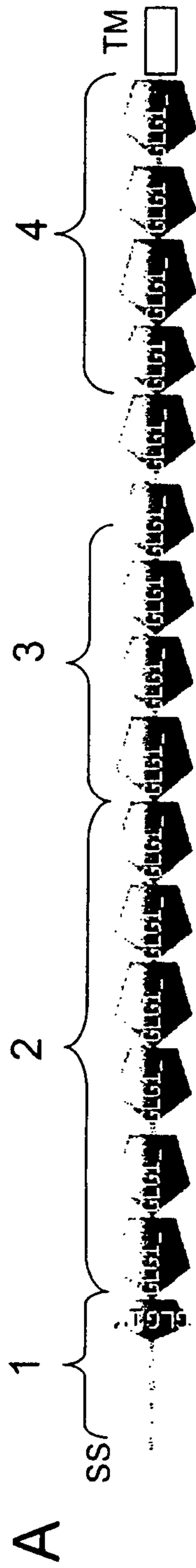


Figure 33

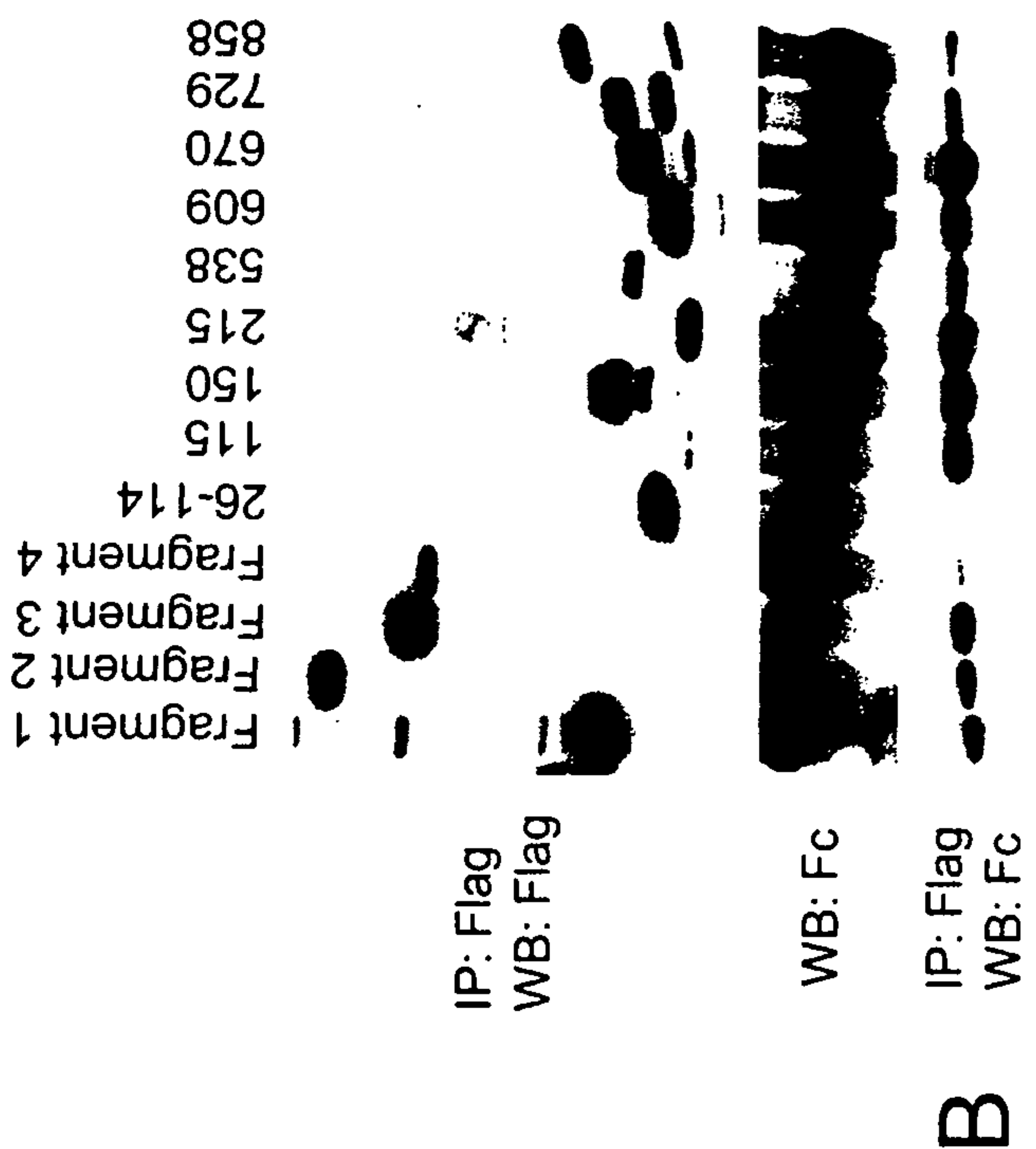
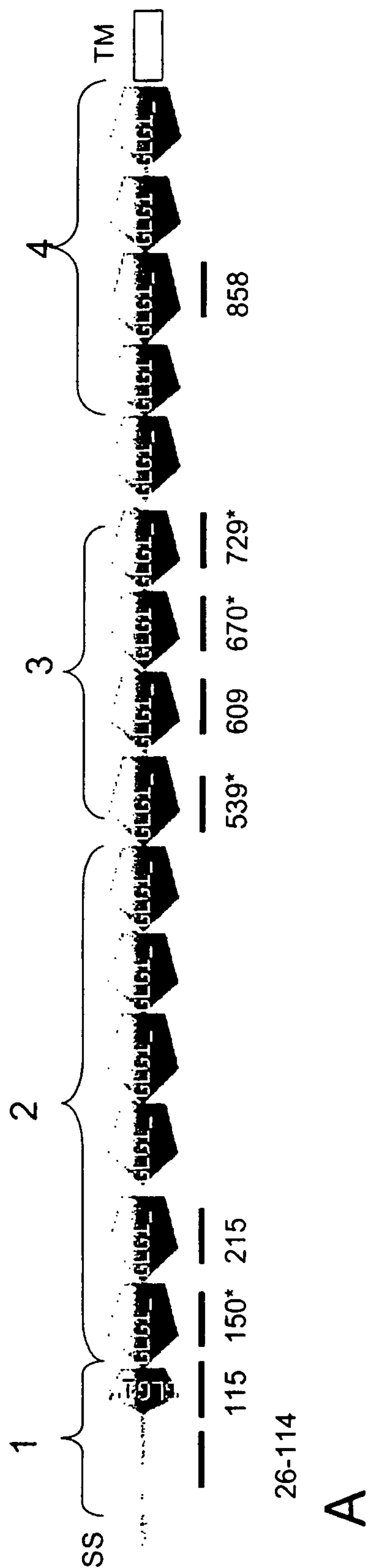


Figure 34

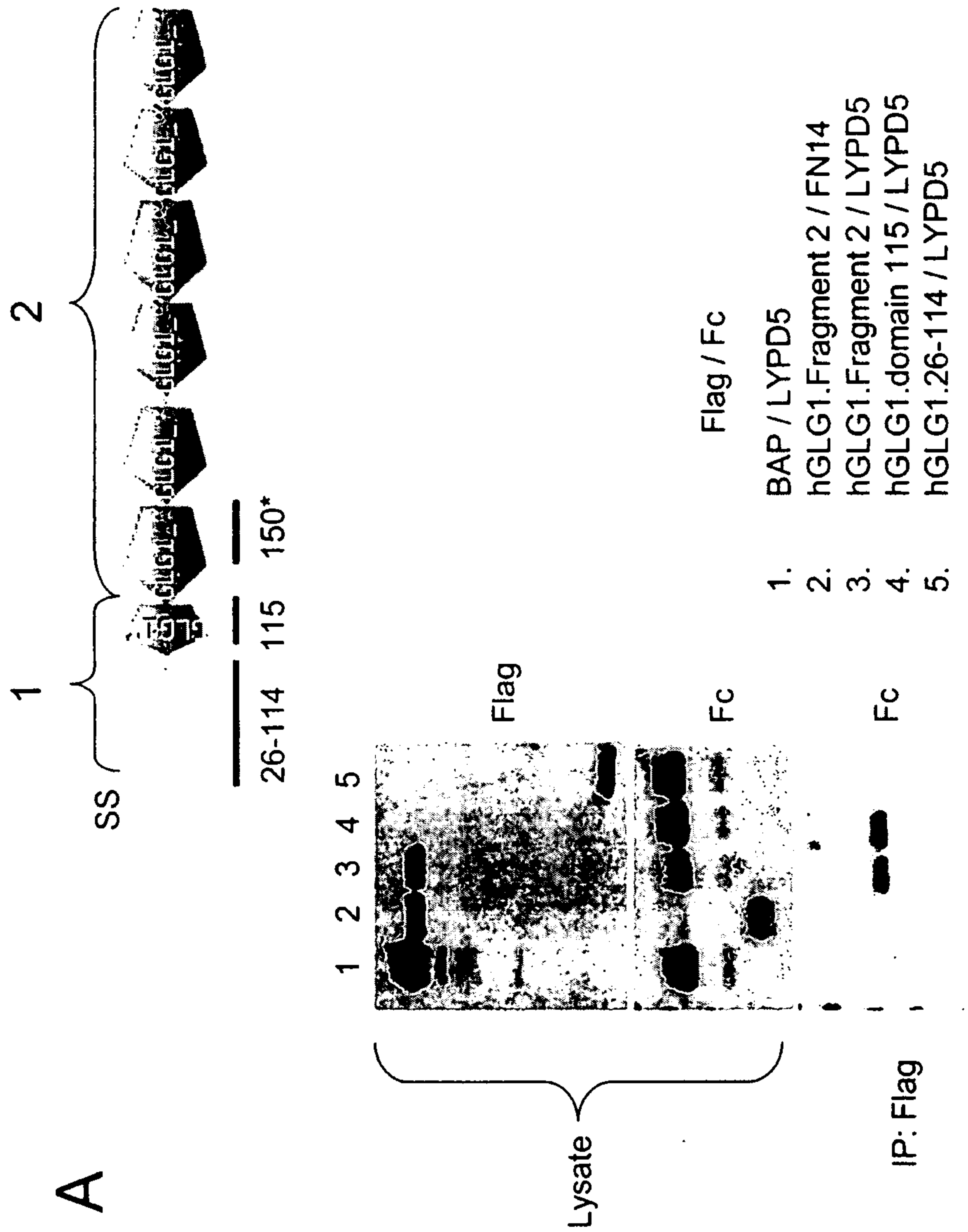


Figure 35

cgttgctgtc gctctgcacg cacctatgtg gaaactaaag cccagagaga aagtctgact
 tgccccacag ccagttagtg actgcagcag caccagaatc tggctgttt cctgtttggc
 tcttctacca ctacggcttg ggatctcggg catggtggct ttgccaatgg tcttgtttt
 gctgctggtc ctgagcagag gtgagagtga attggacgcc aagatcccat ccacagggga
 tgccacagaa tggcggaatc ctacactgtc catgctgggg tctgccagc cagccccctc
 ctgccagaag tgcacctct cacaccccag ctgtgcatgg tgcaagcaac tgaacttcac
 cgcgtcggga gaggcggagg cgcggcgctg cgcccgacga gaggagctgc tggctcagg
 ctgcccgtg gaggagctgg aggagccccg cggccagcag gaggtgctgc aggaccagcc
 gctcagccag ggcgcccgcg gagagggtgc caccagctg gcgcccgcagc gggtcgggt
 cacgtgcgg cctggggagc cccagcagct ccaggtccgc ttcctcgtg ctgagggata
 cccggtggac ctgtactacc ttatggacct gagctactcc atgaaggacg acctggaacg
 cgtgcgccag ctggggcacg ctctgctggc ccggctgcag gaagtcacc acctgtgcg
 cattggtttt ggttctttg tggacaaaac ggtgctgccc ttgtgagca cagtaccctc
 caaactgcgc caccctgcc ccaccggct ggagcgtgc cagtcacat tcagcttca
 ccatgtctg tccctgacgg gggacgcaca agccttcgag cgggaggtgg ggcgccagag
 tgtgccggc aatctggact cgcctgaagg tggcttcgat gccattctgc aggctgcact
 ctgccaggag cagattggct ggagaaatgt gtcccggctg ctggtgtca ctccagacga
 cacattccat acagctgggg acgggaagtt gggcggcatt tcatgcccga gtgatgggca
 ctgccacttg gacagcaatg gctctacag tcgcagcaca gatttgact accttctgt
 gggtcagga gccagggc tctctgcagc aaatatccag cccatcttg ctgtcaccag
 tgccgcactg cctgtctacc aggagctgag taaactgatt cctaagtctg cagtgggga
 gctgagtga gactccagca acgtgttaca gctcatcatg gatgcttata atagcctgtc
 ttcaccgtg accttgaac actcttact cctcctggg gtccacatt ctacgaatc
 ccagtgtgag ggtcctgaga agaggagggg taaggctgag gatcgaggac agtgcaacca
 cgtccgaatc aaccagacgg tgaattctg ggtttctc caagccacc actgcctccc
 agagcccat ctctgaggg tccgggccc tggcttctca gaggagctga ttgtggagt
 gcacacgtg tgtactga attgcagtga caccagccc caggctccc actgcagtga
 tggccaggga cactacaat gtggtgatg cagctgtgcc cctggccgc taggtcggct
 ctgtgagtgc tctgtggcag agctgtctc cccagacctg gaatctgggt gccgggctcc
 caatggcaca gggccccgt gcagtggaaa ggtcactgt caatgtggac gctgcagctg
 cagtggacag agctctgggc atctgtgcga gtgtgacgat gccagctgtg agcgacatga
 gggcatcctc tgcggaggct ttggtcgtg ccaatgtgga gtatgtact gtcagccaa
 ccgcacgggc agagcatgcg aatgcagtgg ggacatggac agttgcatca gtcccagggg
 agggctctgc agtgggcatg gacgtgcaa atgcaaccgc tgccagtgtg ttgacggcta
 ctatggtgct ctatgcgacc aatgcccagg ctgcaagaca ccatgcgaga gacaccggga
 ctgtgcagag tgtggggcct tcaggactgg cccactggcc accaactgca gtacagcttg
 tgccatacc aatgtgacc tggccttggc ccctatctg gatgatggct ggtgcaaaga
 gcggaccctg gacaaccagc tgttcttct ctgtgtggag gatgaccca gaggcacgg
 cgtgctcaga gtgagacccc aagaaaagg agcagaccac acgcaggcca ttgtgctggg
 ctgcgtaggg ggcacgtgg cagtggggct ggggctgtg ctggttacc ggctctcgt
 ggaaatctat gaccgccgg aatacagtc ctttgagaag gacgagcaac aactcaactg
 gaagcaggac agtaatcctc tctacaaaag tgccatcacg accacatca atcctcgtt

Figure 36A

tcaagaggca gacagtccca ctctctgaag gagggaggga cacttaccca aggctcttct
ccttgaggga cagtgggaac tggaggggtga gaggaagggt gggctctgtaa gaccttggtgta
ggggactaat tcactggcga ggtgcggcca ccaccctact tcattttcag agtgacaccc
aagagggctg ctteccatgc ctgcaacctt gcatccatct gggctacccc acccaagtat
acaataaagt cttaccticag aaaaaaaaaa aaaaaaaaaa

Figure 36A (continued)

MVALPMVLVLLLVLSRGESELD AKIPSTGDATEWRNPHLSMLGSCQPAPSCQKCILSHP
SCAWCKQLNFTASGEAEARRCARREELLARGCPLEELEEPRGQQEV LQDQPLSQGARG
EGATQLAPQVRVRLRPGE PQQLQVRFLRAEGYPVDLYYLMDSL SYSMKDDL ERVRQL
GHALLVRLQEVTHSVRIGFGSFVDKTVLPFVSTVPSKLRHPCPTRLERCQSPFSFHHVLSL
TGDAQAFEREVGRQSVSGNLDSP EGGFDAILQAALCQEQIGWRNVSRLLVFTSDDTFHT
AGDGKLG GIFMPSDGHCHLDSNGLYSRSTEFDYPSVGQVAQALSAANI QPIFAV TSAAL
PVYQELSKLIPKSAVGELSEDSSNVVQLIMDAYNSLSSTVTLEHSSLPPGVHISYESQCEG
PEKREGKAEDRGQCNHVRINQTVTFWVSLQATHCLPEPHLLRLRALGFSEELIVELHTLC
DCNCSDTQPQAPHCS DGQGH LQCGVCSCAPGRLGRLCECSVAELSSPDLESGCRAPNGT
GPLCSGKGHCQCGRCS CSGQSSGHLCECDDASCERHEGILCGGFGR CQCQGVCHCHANR
TGRACECSGDM DSCISPEGGLCSGHGRCKCNRCQCLDGYYGALCDQCPGCKTPCERHR
DCAECGAFRTGPLATNCSTACAHTNVT LALAPILDDGWCKERTLDNQLFFFLVEDDAR
GTVVLRVRPQEKGADHTQAIVLGCVGGIVAVGLGLVLAYRLSVEIYDRREYSRFEKEQQ
QLNWKQDSNPLYKSAITTTINPRFQEAD SPTL

Figure 36B