Acknowledgment: The authors thank Ms. Reiko Matsuura for technical assistance.

REFERENCES

Abe K, Kawagoe J, Aoki M, Kogure K, Itoyama Y (1998) Stress protein inductions after brain ischemia. *Cell Mol Neurobiol* 18:709–719

Akaji K, Suga S, Fujino T, Mayanagi K, Inamasu J, Horiguchi T, Sato S, Kawase T (2003) Effect of intra-ischemic hypothermia on the expression of c-Fos and c-Jun, and DNA binding activity of AP-1 after focal cerebral ischemia in rat brain. *Brain Res* 975:149–157

Barone FC, White RF, Spera PA, Ellison J, Currie RW, Wang X, Feuerstein GZ (1998) Ischemic preconditioning and brain tolerance: temporal histological and functional outcomes, protein synthesis requirement, and interleukin-1 receptor antagonist and early gene expression. *Stroke* 29:1937–1950

Bernaudin M, Tang Y, Reilly M, Petit E, Sharp FR (2002) Brain genomic response following hypoxia and re-oxygenation in the neonatal rat. Identification of genes that might contribute to hypoxia-induced ischemic tolerance. *J Biol Chem* 277:39728–39738

Borsello T, Clarke PG, Hirt L, Vercelli A, Repici M, Schorderet DF, Bogousslavsky J, Bonny C (2003) A peptide inhibitor of c-Jun N-terminal kinase protects against excitotoxicity and cerebral ischemia. *Nat Med* 9:1180–1186

Brondello JM, Brunet A, Pouyssegur J, McKenzie FR (1997) The dual specificity mitogen-activated protein kinase phosphatase-1 and -2 are induced by the p42/p44MAPK cascade. *J Biol Chem* 272: 1368–1376

Brondello JM, Pouyssegur J, McKenzie FR (1999) Reduced MAP kinase phosphatase-1 degradation after p42/p44MAPK-dependent phosphorylation. *Science* 286:2514–2517.

Catania MV, Copani A, Calogero A, Ragonese GI, Condorelli DF, Nicoletti F (1999) An enhanced expression of the immediate early gene, Egr-1, is associated with neuronal apoptosis in culture. *Neuroscience* 91:1529–1538

Chen J, Zhu RL, Nakayama M, Kawaguchi K, Jin K, Stetler RA, Simon RP, Graham SH (1996) Expression of the apoptosis-effector gene, Bax, is up-regulated in vulnerable hippocampal CA1 neurons following global ischemia. *J Neurochem* 67:64–71

Chen J, Graham SH, Nakayama M, Zhu RL, Jin K, Stetler RA, Simon RP (1997) Apoptosis repressor genes Bcl-2 and Bcl-x-long are expressed in the rat brain following global ischemia. *J Cereb Blood Flow Metab* 17:2–10

Chen J, Nagayama T, Jin K, Stetler RA, Zhu RL, Graham SH, Simon RP (1998) Induction of caspase-3-like protease may mediate delayed neuronal death in the hippocampus after transient cerebral ischemia. *J Neurosci* 18:4914–4928

Chen ZY, Shie J, Tseng C (2000) Up-regulation of gut-enriched kruppel-like factor by interferon-gamma in human colon carcinoma cells. FEBS Lett 477:67–72

Cho S, Park EM, Kim Y, Liu N, Gal J, Volpe BT, Joh TH (2001) Early c-Fos induction after cerebral ischemia: a possible neuroprotective role. *J Cereb Blood Flow Metab* 21:550–556

Chudin E, Walker R, Kosaka A, Wu SX, Rabert D, Chang TK, Kreder DE (2002) Assessment of the relationship between signal intensities and transcript concentration for Affymetrix GeneChip arrays. *Genome Biol* 3:RESEARCH0005

Dragunow M, Beilharz E, Sirimanne E, Lawlor P, Williams C, Bravo R, Gluckman P (1994) Immediate-early gene protein expression in neurons undergoing delayed death, but not necrosis, following hypoxic-ischaemic injury to the young rat brain. *Brain Res Mol Brain Res* 25:19–33

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci* U S A 95:14863–14868

Franklin CC, Kraft AS (1997) Conditional expression of the mitogen-activated protein kinase (MAPK) phosphatase MKP-1 preferentially inhibits p38 MAPK and stress-activated protein kinase in U937 cells. *J Biol Chem* 272:16917–16923

Friddle CJ, Koga T, Rubin EM, Bristow J (2000) Expression profiling reveals distinct sets of genes altered during induction and regression of cardiac hypertrophy. *Proc Natl Acad Sci* U S A 97:6745–6750

Furuta Y, Uehara T, Nomura Y (2003) Correlation between delayed neuronal cell death and selective decrease in phosphatidylinositol 4-kinase expression in the CA1 subfield of the hippocampus after transient forebrain ischemia. *J Cereb Blood Flow Metab* 23: 962–971

Gonzalez-Zulueta M, Feldman AB, Klesse LJ, Kalb RG, Dillman JF, Parada LF, Dawson TM, Dawson VL (2000) Requirement for nitric oxide activation of p21(ras)/extracellular regulated kinase in neuronal ischemic preconditioning. *Proc Natl Acad Sci* U S A 97:436–441

Gu Z, Jiang Q, Zhang G, Cui Z, Zhu Z (2000) Diphosphorylation of extracellular signal-regulated kinases and c-Jun N-terminal protein kinases in brain ischemic tolerance in rat. *Brain Res* 860:157–160

Gu Z, Jiang Q, Zhang G (2001) Extracellular signal-regulated kinase and c-Jun N-terminal protein kinase in ischemic tolerance. *Neuroreport* 12:3487–3491

Herdegen T, Leah JD (1998) Inducible and constitutive transcription factors in the mammalian nervous system: control of gene expression by Jun, Fos and Krox, and CREB/ATF proteins. *Brain Res Brain Res Rev* 28:370–490

Hermann DM, Mies G, Hossmann KA (1999) Biochemical changes and gene expression following traumatic brain injury: Role of spreading depression. *Restor Neurol Neurosci* 14:103–108

Honkaniemi J, Sharp FR (1996) Global ischemia induces immediate-early genes encoding zinc finger transcription factors. *J Cereb Blood Flow Metab* 16:557–565

Honkaniemi J, States BA, Weinstein PR, Espinoza J, Sharp FR (1997) Expression of zinc finger immediate early genes in rat brain after permanent middle cerebral artery occlusion. *J Cereb Blood Flow Metab* 17:636–646

Huang RP, Fan Y, deBelle I, Ni Z, Matheny W, Adamson ED (1998) Egr-1 inhibits apoptosis during the UV response: correlation of cell survival with Egr-1 phosphorylation. *Cell Death Differ* 5:96–106

Irving EA, Bamford M (2002) Role of mitogen- and stress-activated kinases in ischemic injury. *J Cereb Blood Flow Metab* 22:631–647

Ishii M, Hashimoto S, Tsutsumi S, Wada Y, Matsushima K, Kodama T, Aburatani H (2000) Direct comparison of GeneChip and SAGE on the quantitative accuracy in transcript profiling analysis. *Genomics* 68:136–143

Jang CW, Chen CH, Chen CC, Chen JY, Su YH, Chen RH (2002) TGF-beta induces apoptosis through Smad-mediated expression of DAP- kinase. *Nat Cell Biol* 4:51–58

Jin K, Mao XO, Eshoo MW, Nagayama T, Minami M, Simon RP, Greenberg DA (2001) Microarray analysis of hippocampal gene expression in global cerebral ischemia. *Ann Neurol* 50:93–103

Jin K, Nagayama T, Mao X, Kawaguchi K, Hickey RW, Greenberg DA, Simon RP, Graham SH (2002) Two caspase-2 transcripts are expressed in rat hippocampus after global cerebral ischemia. *J Neurochem* 81:25–35

Kariko K, Harris VA, Rangel Y, Duvall ME, Welsh FA (1998) Effect of cortical spreading depression on the levels of mRNA coding for putative neuroprotective proteins in rat brain. *J Cereb Blood Flow Metab* 18:1308–1315

Kawahara N, Ruetzler CA, Klatzo I (1995) Protective effect of spreading depression against neuronal damage following cardiac arrest cerebral ischaemia. *Neurol Res* 17:9–16

Kawahara N, Mishima K, Higashiyama S, Taniguchi N, Tamura A, Kirino T (1999) The gene for heparin-binding epidermal growth factor-like growth factor is stress-inducible: its role in cerebral ischemia. *J Cereb Blood Flow Metab* 19:307–320

Keyvani K, Witte OW, Paulus W (2002) Gene expression profiling in perilesional and contralateral areas after ischemia in rat brain. *J Cereb Blood Flow Metab* 22:153–160

Kim YD, Sohn NW, Kang C, Soh Y (2002) DNA array reveals altered gene expression in response to focal cerebral ischemia. *Brain Res Bull* 58:491–498

Kinouchi H, Sharp FR, Chan PH, Koistinaho J, Sagar SM, Yoshimoto T (1994) Induction of c-fos, junB, c-jun, and hsp70 mRNA in cortex, thalamus, basal ganglia, and hippocampus following

middle cerebral artery occlusion. *J Cereb Blood Flow Metab* 14:808–817

Kirino T (1982) Delayed neuronal death in the gerbil hippocampus following ischemia. *Brain Res* 239:57–69

Kirino T, Tsujita Y, Tamura A (1991) Induced tolerance to ischemia in gerbil hippocampal neurons. *J Cereb Blood Flow Metab* 11: 299–307

Kirino T (2002) Ischemic tolerance. *J Cereb Blood Flow Metab* 22:1283–1296.

Kleihues P, Hossmann KA (1971) Protein synthesis in the cat brain after prolonged cerebral ischemia. *Brain Res* 35:409–418

Lee CK, Weindruch R, Prolla TA (2000) Gene-expression profile of the ageing brain in mice. *Nat Genet* 25:294–297

Li H, Kolluri SK, Gu J, Dawson MI, Cao X, Hobbs PD, Lin B, Chen G, Lu J, Lin F, Xie Z, Fontana JA, Reed JC, Zhang X (2000) Cytochrome c release and apoptosis induced by mitochondrial targeting of nuclear orphan receptor TR3. *Science* 289:1159–1164

Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 14:1675–1680

Lockhart DJ, Winzeler EA (2000) Genomics, gene expression and DNA arrays. *Nature* 405:827–836

Lu P, Nakorchevskiy A, Marcotte EM (2003) Expression deconvolution: A reinterpretation of DNA microarray data reveals dynamic changes in cell populations. *Proc Natl Acad Sci U S A* 100: 10370–10375

MacManus JP, Linnik MD (1997) Gene expression induced by cerebral ischemia: an apoptotic perspective. *J Cereb Blood Flow Metab* 17: 815–832

Maehara K, Oh-Hashi K, Isobe KI (2001) Early growth-responsive-1-dependent manganese superoxide dismutase gene transcription mediated by platelet-derived growth factor. *Faseb J* 15:2025–2026

Maher P (2001) How protein kinase C activation protects nerve cells from oxidative stress-induced cell death. *J Neurosci* 21:2929–2938

Marte BM, Downward J (1997) PKB/Akt: connecting phosphoinositide 3-kinase to cell survival and beyond. *Trends Biochem Sci* 22: 355–358

Masuyama N, Oishi K, Mori Y, Ueno T, Takahama Y, Gotoh Y (2001) Akt inhibits the orphan nuclear receptor Nur77 and T-cell apoptosis. *J Biol Chem* 276:32799–32805

Murphy TC, Woods NR, Dickson AJ (2001) Expression of the transcription factor GADD153 is an indicator of apoptosis for recombinant chinese hamster ovary (CHO) cells. *Biotechnol Bioeng* 75: 621–629

Namura S, Iihara K, Takami S, Nagata I, Kikuchi H, Matsushita K, Moskowitz MA, Bonventre JV, Alessandrini A (2001) Intravenous administration of MEK inhibitor U0126 affords brain protection against forebrain ischemia and focal cerebral ischemia. *Proc Natl Acad Sci U S A* 98:11569–11574

Nishi H, Nishi KH, Johnson AC (2002) Early Growth Response-1 gene mediates up-regulation of epidermal growth factor receptor expression during hypoxia. *Cancer Res* 62:827–834

Nowak TS (1990) Protein synthesis and the heart shock/stress response after ischemia. *Cerebrovasc Brain Metab Rev* 2:345–366

Paxinos G, Watson C (1986) The rat brain in stereotaxic coordinates. San Diego: Academic Press

Pero R, Lembo F, Di Vizio D, Boccia A, Chieffi P, Fedele M, Pierantoni GM, Rossi P, Iuliano R, Santoro M, Viglietto G, Bruni CB, Fusco A, Chiariotti L (2001) RNF4 is a growth inhibitor expressed in germ cells but not in human testicular tumors. *Am J Pathol* 159: 1225–1230

Pulsinelli WA, Brierley JB, Plum F (1982) Temporal profile of neuronal damage in a model of transient forebrain ischemia. *Ann Neurol* 11:491–498

Raghavendra Rao VL, Bowen KK, Dhodda VK, Song G, Franklin JL, Gavva NR, Dempsey RJ (2002) Gene expression analysis of spontaneously hypertensive rat cerebral cortex following transient focal cerebral ischemia. *J Neurochem* 83:1072–1086

Rampon C, Jiang CH, Dong H, Tang YP, Lockhart DJ, Schultz PG, Tsien JZ, Hu Y (2000) Effects of environmental enrichment on gene expression in the brain. *Proc Natl Acad Sci U S A* 97: 12880–12884

Schneider A, Martin-Villalba A, Weih F, Vogel J, Wirth T, Schwaninger M (1999) NF-kappaB is activated and promotes cell death in focal cerebral ischemia. *Nat Med* 5:554–559

Scholl FA, McLoughlin P, Ehler E, de Giovanni C, Schafer BW (2000) DRAL is a p53-responsive gene whose four and a half LIM domain protein product induces apoptosis. *J Cell Biol* 151:495–506

Schwarz DA, Barry G, Mackay KB, Manu F, Naeve GS, Vana AM, Verge G, Conlon PJ, Foster AC, Maki RA (2002) Identification of differentially expressed genes induced by transient ischemic stroke. *Brain Res Mol Brain Res* 101:12–22

Sommer C, Gass P, Kiessling M (1995) Selective c-JUN expression in CA1 neurons of the gerbil hippocampus during and after acquisition of an ischemia-tolerant state. *Brain Pathol* 5:135–144

Song G, Cechvala C, Resnick DK, Dempsey RJ, Rao VL (2001) GeneChip analysis after acute spinal cord injury in rat. *J Neurochem* 79:804–815

Soriano MA, Tessier M, Certa U, Gill R (2000) Parallel gene expression monitoring using oligonucleotide probe arrays of multiple transcripts with an animal model of focal ischemia. *J Cereb Blood Flow Metab* 20:1045–1055

Stanton LW, Garrard LJ, Damm D, Garrick BL, Lam A, Kapoun AM, Zheng Q, Protter AA, Schreiner GF, White RT (2000) Altered patterns of gene expression in response to myocardial infarction. *Circ Res* 86:939–945

Sugino T, Nozaki K, Takagi Y, Hattori I, Hashimoto N, Moriguchi T, Nishida E (2000) Activation of mitogen-activated protein kinases after transient forebrain ischemia in gerbil hippocampus. *J Neurosci* 20:4506–4514

Tang Y, Lu A, Aronow BJ, Sharp FR (2001) Blood genomic responses differ after stroke, seizures, hypoglycemia, and hypoxia: blood genomic fingerprints of disease. *Ann Neurol* 50:699–707

Truettner J, Busto R, Zhao W, Ginsberg MD, Perez-Pinzon MA (2002) Effect of ischemic preconditioning on the expression of putative neuroprotective genes in the rat brain. *Brain Res Mol Brain Res* 103:106–115

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 98:5116–5121

Wang H, Zhan Y, Xu L, Feuerstein GZ, Wang X. (2001) Use of suppression subtractive hybridization for differential gene expression in stroke: discovery of CD44 gene expression and localization in permanent focal stroke in rats. *Stroke* 32:1020–1027

Whitfield PC, Williams R, Pickard JD (1999) Delayed induction of JunB precedes CA1 neuronal death after global ischemia in the gerbil. Brain Res 818:450–458

Xia Z, Dickens M, Raingeaud J, Davis RJ, Greenberg ME (1995) Opposing effects of ERK and JNK-p38 MAP kinases on apoptosis. *Science* 270:1326–1331

Zhang Y, Widmayer MA, Zhang B, Cui JK, Baskin DS (1999) Suppression of post-ischemic-induced fos protein expression by an antisense oligonucleotide to c-fos mRNA leads to increased tissue damage. *Brain Res* 832:112–117

Zhao Y, Zhang ZY (2001) The mechanism of dephosphorylation of extracellular signal-regulated kinase 2 by mitogen-activated protein kinase phosphatase 3. *J Biol Chem* 276:32382–32391

Oncogene (2004) 23, 726–734
© 2004 Nature Publishing Group All rights reserved 0950-9232/04 $25.00
www.nature.com/onc

STAT3 and MITF cooperatively induce cellular transformation through upregulation of c-*fos* expression

Akiko Joo[1], Hiroyuki Aburatani[2], Eiichi Morii[3], Hideo Iba[4] and Akihiko Yoshimura*,[1]

[1]*Division of Molecular and Cellular Immunology, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan;* [2]*Division of Genome Science, Research Center for Advanced Science and Technology, The University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8904, Japan;* [3]*Department of Pathology, Osaka University Medical School, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan;* [4]*Institute of Medical Science, The University of Tokyo, Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan*

The signal transducer and activator of transcription (STAT) family proteins are transcription factors critical in mediating cytokine signaling. Among them, STAT3 is frequently activated in a number of human cancers and transformed cell lines and is implicated in tumorigenesis. However, although constitutively activated STAT3 mutant (STAT3C) leads to cellular transformation, its transformation potential such as colony-forming activity in soft-agar is much weaker than that of v-src. To identify tumorigenic factors that cooperatively induce cellular transformation with STAT3C, we screened the retroviral cDNA library. We found that the microphthalmia-associated transcription factor (MITF), an essential transcription factor for melanocyte development and pigmentation, induces anchorage-independent growth of NIH-3T3 cells in cooperation with STAT3C. Microarray analysis revealed that c-*fos* is highly expressed in transformants expressing STAT3C and MITF. Promoter analysis and chromatin immunoprecipitation assay suggested that both STAT3 and MITF can cooperatively upregulate the c-*fos* gene. In addition, the transformation of NIH-3T3 cells by both MITF and STAT3C was significantly suppressed by a dominant-negative AP-1 retrovirus. These data indicate that MITF and STAT3 cooperatively induce c-*fos*, resulting in cellular transformation.

Oncogene (2004) **23**, 726–734. doi:10.1038/sj.onc.1207174

Keywords: MITF; STAT3; c-*fos*; cellular transformation

Introduction

The signal transducer and activator of transcription (STAT) family proteins were identified in the last decade as transcription factors essential for mediating virtually all cytokine signaling (Darnell, 1997; Stark *et al.*, 1998). These proteins become activated through tyrosine phosphorylation. In addition to their central roles in normal cell signaling, recent studies have demonstrated that constitutively activated STAT signaling, especially STAT3, directly contributes to oncogenesis (Bromberg and Darnell, 2000). For example, all *src*-transformed cell lines exhibit constitutively activated STAT3 (Yu *et al.*, 1995), and dominant-negative STAT3 suppresses *src* transformation without having any effect on *ras* transformation (Turkson *et al.*, 1998). More directly, Bromberg *et al.* (1999) demonstrated that a constitutively activated form of STAT3, STAT3C, which has two substituted cysteine residues within the C-terminal loop of the SH2 domain, resulting in a spontaneous transcriptionally active dimer, causes cellular transformation scored by colony formation in soft-agar and tumor formation in nude mice. Thus, the activated STAT3 molecule by itself can mediate cellular transformation. Extensive surveys of primary tumors and cell lines derived from tumors have indicated that an inappropriate activation of STAT3 occurs at a surprisingly high frequency in a wide variety of human cancers (Bowman *et al.*, 2000). However, until now, mutations in the STAT3 gene have not been identified in these cancers, hence it remained to be determined how endogenous STAT3 is constitutively activated and what kinds of genes are involved in tumorigenicity induced by constitutively activated STAT3.

The microphthalmia-associated transcription factor (MITF) is a basic helix–loop–helix leucine zipper (b-HLH-Zip) transcription factor that plays a critical role in the differentiation of various cell types, including neural crest-derived melanocytes, mast cells, osteoclasts, and optic cup-derived retinal pigment epithelium. MITF mutations in humans produce auditory–pigmentary syndromes, such as Waardenburg syndrome type IIa and Tietz syndrome, characterized by mast cell defects, inner ear problems, and abnormal, patchy pigmentation of the hair and skin. In mice, the *mi* allele protein with the deletion of 216R in the basic region is known as a dominant-negative form through the sequestration of wild-type partners in non-DNA-binding dimmers. In addition to the complete absence of melanocytes, MITF dominant-negative mutants exhibit osteopetrosis (Kitamura *et al.*, 2002). MITF consists of at least five

*Correspondence: A Yoshimura; E-mail: yakihiko@bioreg.kyushu-u.ac.jp
Received 24 June 2003; revised 4 September 2003; accepted 5 September 2003

isoforms, including MITF-A, MITF-B, MITF-C, MITF-H, and MITF-M, and MITF-M is the melanocyte-specific type (Tachibana, 1997; Udono *et al.*, 2000; Shibahara *et al.*, 2001). MITF regulates the expression of melanocyte differentiation markers, including tyrosinase, tyrosinase-related protein, and dopachrome tautomerase (DCT), all of which are required for pigmentation (Carreira *et al.*, 2000). MITF is one of the genes involved in tumor growth and the metastasis of melanoma (Vachtenheim *et al.*, 2001, Nyormoi and Bar-Eli, 2003). However, transcriptional target genes of MITF that regulate melanoma tumorigenicity or metastasis have not yet been elucidated. Moreover, since MITF alone has low or no oncogenic activity, a cofactor(s) that cooperatively functions with MITF may be necessary for the transformation of melanocytes.

In this study, we first demonstrated that STAT3 and MITF cooperatively induce cellular transformation *in vitro*. We also identified c-*fos* as a target gene of STAT3 and MITF using microarray analysis. The induction of the c-*fos* gene is necessary for the anchorage-independent growth of NIH-3T3 cells transformed with STAT3C and MITF. Our study provides a novel role of STAT3 in melanocyte proliferation and tumor growth of melanoma.

Results

Screening for STAT3C cofactors for cellular transformation

We and others have shown that NIH-3T3 cells expressing STAT3C or wild-type STAT3, which is activated by the type C hepatitis virus (HCV) core protein, possess a colony-forming potential in soft-agar and tumorigenicity in nude mice (Yoshida *et al.*, 2002). However, the number and size of the colonies and tumor size by the expression of active STAT3 are much smaller than those of NIH-3T3 cells transformed with v-src (Bromberg *et al.*, 1999; Yoshida *et al.*, 2002). Therefore, the constitutive activation of STAT3 may not be sufficient for full transformation. With this in mind, we screened cofactors that induce full transformation in cooperation with activated STAT3 by using retrovirus cDNA transfer (Kitamura *et al.*, 1995). NIH-3T3 cells expressing STAT3C (STAT3C-3T3) were infected with the HeLa cell retroviral cDNA library (2×10^6 independent clones) and plated into soft-agar medium. After 3 weeks of incubation, two large colonies were formed and the integrated cDNAs were recovered by PCR and sequenced. One colony contained MITF cDNA with N-terminal 104 amino acids deletion compared with MITF-M (ΔN-MITF), and the other colony included full-length granulin cDNA that has been shown to induce colony formation in soft-agar in NIH-3T3 cells (Zanocco-Marani *et al.*, 1999). The ATG of the exon 3 of the MITF gene was utilized as the first AUG codon in ΔN-MITF (Figure 1a). N-terminal truncation resulted in a missing N-terminal glutamine-rich region, but ΔN-MITF retained DNA-binding and transactivation domains.

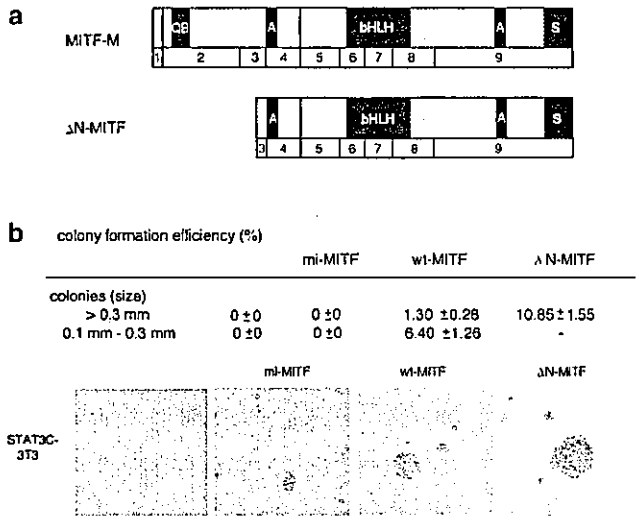


Figure 1 Transforming potential with a combination of MITF and STAT3C. (a) Structures of MITF-M and our screening clone, ΔN-MITF. The numbers shown under MITF isoforms indicate exons. The glutamine-rich basic region (QB), the transcriptional activation domain (A), the bHLH-LZ structure, and the serine-rich domain (S) are indicated. (b) STAT3C-transformed 3T3 (STAT3C-3T3) cells were infected with the pMX empty vector, pMX-mi-MITF, wt-MITF, or ΔN-MITF and plated into soft-agar medium. On day 21, colonies were counted and photographed

MITF-induced anchorage-independent growth in cooperation with STAT3C

The inserted cDNAs subcloned into the retroviral vector, pMX-IRES-EGFP, were introduced into parental NIH-3T3 cells or STAT3C-3T3 cells and then plated into soft-agar medium. N-terminal-truncated MITF induced the cellular transformation of STAT3C-3T3, but not parental NIH-3T3 cells, whereas granulin cDNA induced anchorage-independent cell growth in both NIH-3T3 cells and STAT3C-3T3 cells (Figure 1b and data not shown). Therefore, ΔN-MITF has the potential to induce the anchorage-independent growth of NIH-3T3 cells in cooperation with STAT3C. We also found that full-length (wt-) MITF could lead to anchorage-independent growth of NIH-3T3 cells in cooperation with STAT3C (Figure 1b). However, ΔN-MITF showed greater colony-forming activity, both in size and number, than wt-MITF.

We then compared the cellular morphology of transfectants. It has been reported that the forced expression of MITF in NIH-3T3 cells results in refractile cell morphology, which resembled dendritic cells and melanocytes (Tachibana, 1997). We also observed that ΔN-MITF-infected NIH-3T3 cells showed dendritic cell-like morphological changes (Figure 2a). However, as shown in Figure 2b, STAT3C-3T3 cells expressing wt-MITF or ΔN-MITF displayed some of the morphological changes associated with fibroblast transformation, that is, elongated shape and rounding.

Constitutive activation of STAT3 in melanoma cells

We then examined STAT3 activation in melanoma cells in which MITF plays an important role in transformed

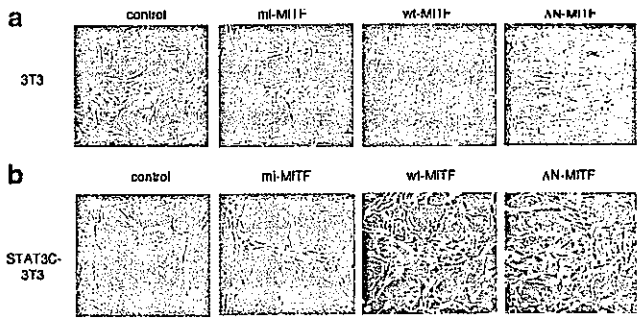


Figure 2 Cytology of MITF-infected 3T3 and STAT3C-3T3 cells. 3T3 and STAT3C-3T3 were infected with the pMX empty vector, pMX-mi-MITF, wt-MITF, or ΔN-MITF. wt-MITF and ΔN-MITF induced morphological change in both 3T3 (a) and STAT3C-3T3 (b)

phenotypes. As shown in Figure 3a, some melanoma cell lines, B16F10, G361, MMAc, and HMV-II, showed constitutive phosphorylation of STAT3. We examined whether MITF induced the constitutive activation of STAT3. Immunoblotting with an anti-phosphorylated STAT3-specific antibody revealed that phosphorylation occurred in tyrosine 705 (Y705) of STAT3 in STAT3C-3T3 cells (Figure 3b). wt-MITF and ΔN-MITF did not affect phosphorylation states of STAT3 in NIH-3T3 cells (Figure 3b). As shown in Figure 3c, MITF had little effect on or rather suppressed STAT3-dependent APRE-luciferase activity. These data indicate that STAT3 is often constitutively activated in melanoma cells, but the mechanism is probably independent of MITF expression. Furthermore, MITF-transactivation activity was not affected by STAT3C (data not shown). Therefore, we speculated that an oncogenic target gene(s) could be induced by the cooperative action of STAT3 and MITF.

Microarray screening for target genes of STAT3C and MITF

To identify target genes of MITF and STAT3, a microarray-based screen was undertaken. Total RNA was isolated from ΔN-MITF-infected NIH-3T3 (ΔN-MITF-3T3), STAT3C-3T3, and ΔN-MITF-infected STAT3C-3T3 (ΔN-MITF/STAT3C-3T3) cells and subjected to Affymetrix microarray analysis (about 12000 genes). As summarized in Figure 4a, seven genes in ΔN-MITF/STAT3C-3T3 cells were identified as more than 10-fold upregulated genes compared with ΔN-MITF-3T3 and STAT3C-3T3 cells. Most of the genes were mast cell or melanocyte-specific genes and chemokines, and the upregulation of these genes was confirmed by RT–PCR analysis (Figure 4b). Among these seven genes, the upregulation of c-*fos* is particularly interesting because c-*fos* is a component of the AP-1 transcription factor and known to be an oncogene. The functions of AP-1, composed of Fos family proteins (c-Fos, Fra-1, Fra-2, and FosB) and Jun family proteins (c-Jun, JunB, and JunD), were shown to play important roles not only in normal cell growth but also in several transformed cells induced by oncogenes (Ui *et al*., 2000). Therefore,

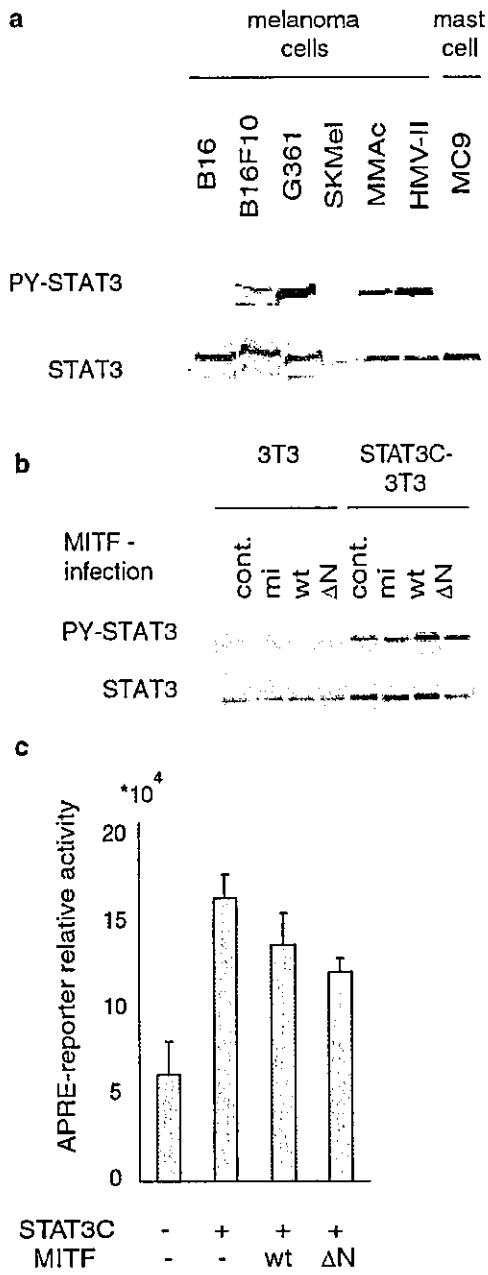
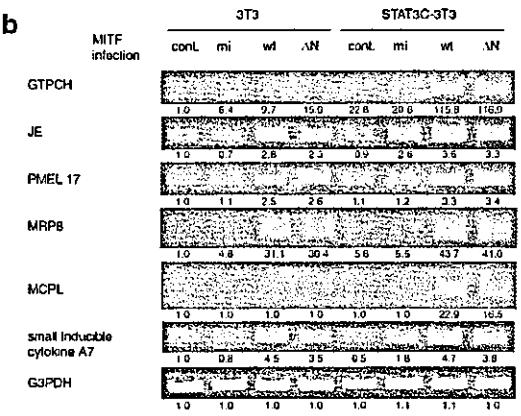


Figure 3 MITF and STAT3C do not directly activate each other. (a and b) Phosphorylation of STAT3 was detected by Western blotting with anti-phosphorylated Tyr705 of the STAT3-specific antibody. Lysate from cell lines of melanoma cells and mast cells (a) or 3T3 and STAT3C-3T3 cells infected with the pMX empty vector, pMX-mi-MITF, wt-MITF, or ΔN-MITF (b) were examined. (c) HEK293 cells were transfected with a plasmid mixture containing the APRE-luciferase reporter gene (0.04 μg) and the β-galactosidase gene (0.1 μg). To examine the MITF-dependent APRE-luciferase activity, cDNA of STAT3C (0.2 μg) and MITF (0.1 μg) was also introduced. Data normalized with the β-galactosidase activity from triplicate experiments are shown

we confirmed the upregulation of the c-*fos* gene by Northern blotting. As shown in Figure 4c, c-*fos* was consecutively expressed in ΔN-MITF/STAT3C-3T3 cells, but was not detected in quiescent 3T3 cells. We also detected the endogenously high expression of c-*fos*

a Target genes of ΔN MITF and STAT3C

Acc. Number	gene
L09737	GTP cyclohydrolase 1
M19681	platelet-derived growth factor-inducible protein (JE)
U14133	pmel 17
M83218	intracellular calcium binding protein (MRP-8)
M57401	mast cell protease-like protein
X70058	small inducible cytokine A7
V00727	c-fos oncogene



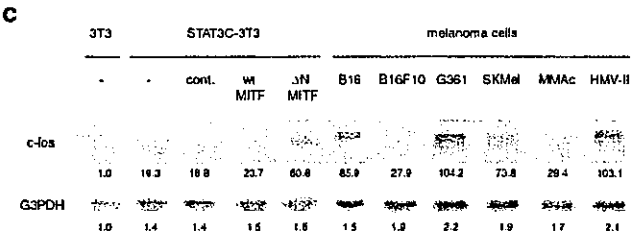


Figure 4 Expression of genes upregulated by MITF and STAT3C. (a) In microarray analysis, seven genes in 3T3 cells expressing both ΔN-MITF and STAT3C were found as >10-fold upregulated genes compared with STAT3C-3T3 or ΔN-MITF single transfectants. (b) Upregulation in six of the genes detected by microarray analysis. The mRNA expression level was evaluated by the RT–PCR method in 3T3 or STAT3C-3T3 cells infected with the pMX empty vector, pMX-mi-MITF, wt-MITF, or ΔN-MITF. The intensity of the PCR band was quantified with NIH image software. (c) Northern blotting analysis of the c-*fos* oncogene. 3T3 cells, STAT3C-3T3 cells uninfected or infected with empty vector, wt-, or ΔN-MITF viruses, and six kinds of melanoma cell lines were examined with RNA probes of c-*fos* and G3PDH

in several melanoma cell lines (Figure 4c), suggesting that the constitutive expression of c-*fos* contributes to the oncogenesis of melanoma.

*STAT3 and MITF directly bind to the promoter region of c-*fos

It has been shown that the c-*fos* proximal promoter region contains a single STAT3-binding site in the SIE region (Shibuya *et al*., 1994). We also noticed many MITF-binding motifs (CANNTG; E-box) (Tsujimura *et al*., 1996) in the promoter region (Figure 5a). To confirm that c-*fos* activation was cooperatively induced by MITF and STAT3C, a reporter gene assay using c-*fos* promoter luciferase constructs (Hatakeyama *et al*., 1992) was carried out. The transcriptional activity of c-*fos* was significantly increased by the transient expression of wt- or ΔN-MITF and STAT3C in HEK293 cells. MITF-induced c-*fos* promoter activation was further stimulated by leukemia inhibitory factor (LIF), which activates endogenous STAT3 (Figure 5b and data not shown).

Next, we examined the region of the c-*fos* promoter responsible for the interaction of MITF and STAT3. The c-*fos* promoter construct contains five potential MITF-binding sites. Using mutated or truncated forms of the c-*fos* promoter, we found that the SIE region is important for activation by STAT3 and an MITF-binding motif in the SRE region is essential for promoter activation by MITF (Figure 5c).

To confirm the direct binding of STAT3 and MITF to the c-*fos* promoter region, DNA-binding assay (Figure 5d) as well as chromatin immunoprecipitation (ChIP) assay (Figure 5e) were performed. First, nuclear extracts from 293T cells transfected with Myc-tagged ΔN-MITF and STAT3C or from cells stimulated with or without LIF were incubated with beads conjugated with oligonucleotides of the human c-*fos* promoter sequence, including the SIE and the SRE (55 mer). As shown in Figure 5d, MITF (lanes 2, 4, and 6) as well as both phosphorylated endogenous STAT3 (lanes 3 and 4) and STAT3C (lanes 5 and 6) bound to the oligonucleotides of the c-*fos* promoter region *in vitro*. Non-phosphorylated STAT3 without LIF stimulation (lanes 1 and 2) did not bind to the oligonucleotide beads, suggesting a specific interaction of activated STAT3 and the DNA.

For ChIP assay (Figure 5e), the crosslinked chromatins from wt-MITF/STAT3C-3T3 cells and ΔN-MITF/STAT3C-3T3 cells as well as melanomas (G361 and HMV-II) in which STAT3 was consecutively phosphorylated were immunoprecipitated with STAT3- or MITF-specific antibodies. The crosslinked protein was then removed from DNA by proteolysis. Finally, the immunoprecipitated DNA was analysed by PCR to detect the c-*fos* promoter region. As shown in Figure 5e, the anti-STAT3 antibody and the anti-MITF antibody precipitated the c-*fos* promoter SIE region and the MITF-binding motif in the SRE region, respectively.

Dominant-negative mutant of AP-1 inhibited cellular transformation of MITF/STAT3C-3T3

To investigate the contribution of the c-*fos* gene in the anchorage-independent growth of MITF/STAT3C-3T3 cells, we introduced a dominant-negative mutant of AP-1 into these cells. We used a retrovirus carrying SupJunD-1, which has an N-terminal deletion of the transactivation domain of c-Jun. We have shown that the SupJunD-1 virus suppresses the transactivation activity of AP-1 and inhibits colony formation in soft agar of various types of tumor cells (Ui *et al*., 2000). As shown in Figure 6, the dominant-negative AP-1

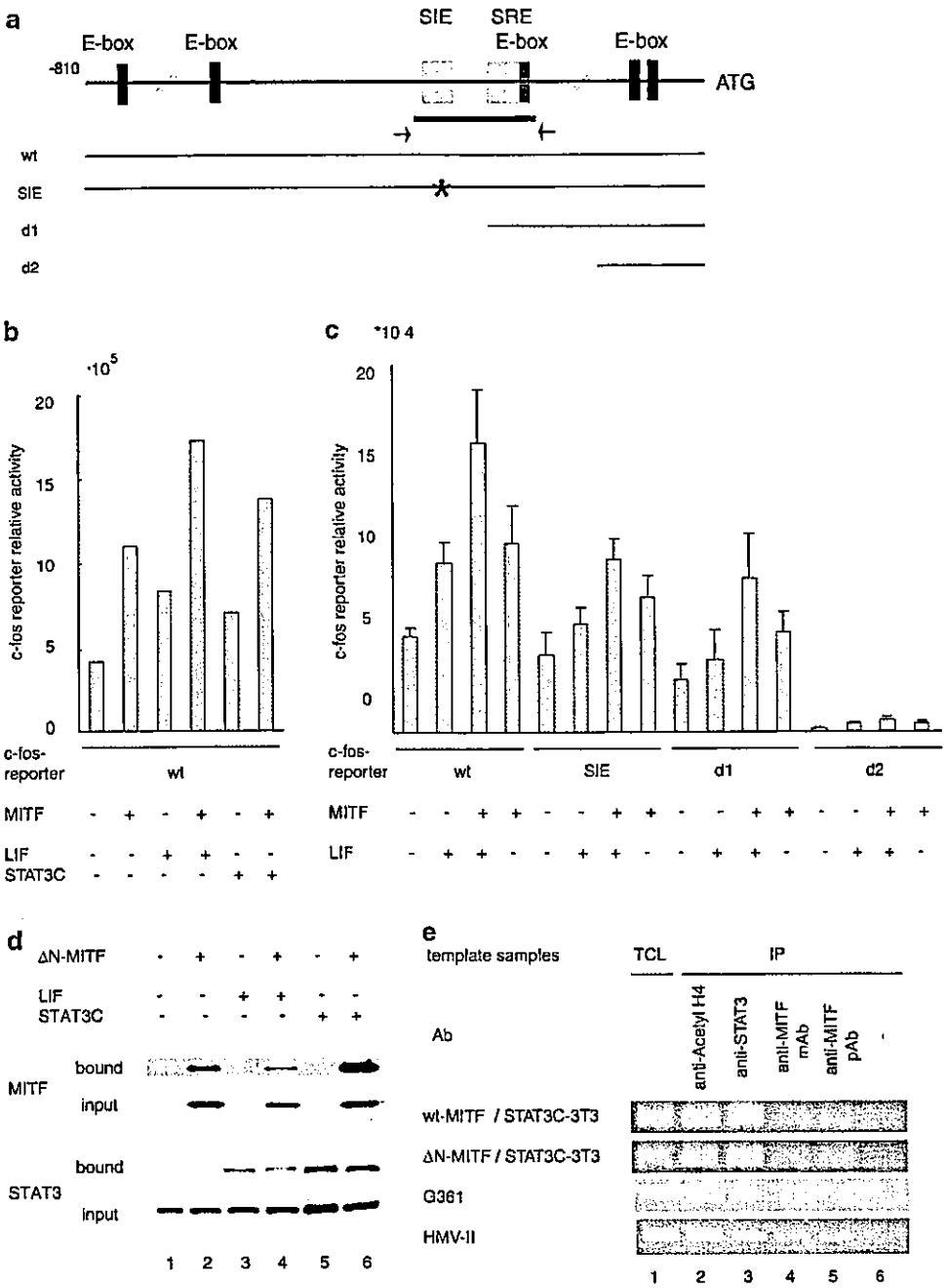


Figure 5 MITF and STAT3C directly binds to the promoter region of c-fos. (a) A diagram of the c-*fos*-luciferase reporter gene containing five E-boxes, potential MITF-binding motifs, and the SIE region known as a binding site of STAT3. Asterisks indicate mutations in SIE, and two deletion mutants (d1 and d2) are shown. The bold bar shows the region used for the DNA-binding assay and arrows show the PCR primers for ChIP assay. (b and c) *Reporter assay*. HEK293 cells were transfected with a plasmid mixture containing the c-*fos*-luciferase reporter gene (2 ng), the β-galactosidase gene (0.1 μg), STAT3C (0.2 μg), and the wt-MITF plasmid (b, 200 ng, and c, 10 ng). After transfection, cells were incubated in the presence or absence of 10 ng/ml LIF for 8 h, and cell extracts were prepared. Data normalized with the β-galactosidase activity from triplicate experiments are shown. (d) *DNA-binding assay*. 293T cells were transfected with the pcDNA3-Myc-ΔN-MITF and pRcCMV-STAT3C and stimulated with 10 ng/ml LIF for 6 h, and the nuclear extracts were prepared. The DNA-binding proteins bound to the oligonucleotide-conjugated beads were analysed by immunoblotting with anti-STAT3 and anti-Myc antibodies. (e) ChIP assay was performed to determine the binding of MITF and STAT3 to the promoter region of c-*fos*. wt- or ΔN-MITF-infected STAT3C-3T3 was determined *in vitro*, and the melanoma cells G361 and HMV-II were examined *in vivo*. Chromatin complexes were immunoprecipitated with anti-acetylated histone H4, anti-STAT3, and monoclonal or polyclonal anti-MITF antibodies

significantly suppressed the anchorage-independent growth of wt- or ΔN-MITF-infected STAT3C-3T3 cells. These data support our notion that c-*fos* induction is one of the mechanisms of increased anchorage-independent growth of MITF/STAT3C-expressing cells.

colony formation efficiency (%)

	retrovirus infection	wt-MITF	ΔN-MITF
colonies (size)			
> 0.3 mm	-	1.30 ± 0.28	10.85 ± 1.55
	cont.	1.11 ± 0.41	6.83 ± 3.31
	dnAP-1	0.03 ± 0.05	1.33 ± 0.90
0.1 mm - 0.3 mm	-	6.40 ± 1.26	-
	cont.	4.98 ± 1.88	-
	dnAP-1	0.84 ± 0.88	-

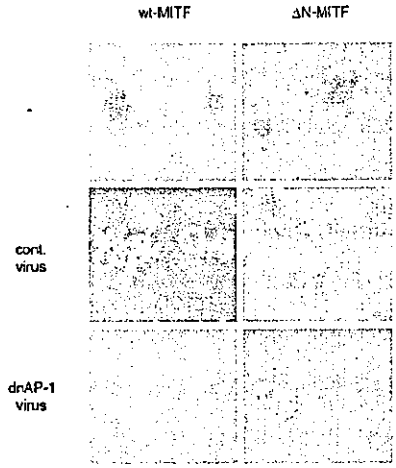


Figure 6 Expression of dominant-negative AP-1 inhibits anchorage-independent growth of MITF/STAT3C-expressing cells. wt- or ΔN-MITF-infected STAT3C-3T3 cells were infected with retrovirus carrying dominant-negative AP-1 (dnAP-1) or an empty vector by the retrovirus at moi = 10. Cells were selected with puromycin and then plated into soft-agar medium. On day 21, colonies were counted and photographed. Data from triplicate experiments are shown

Discussion

In this study, we identified MITF as a collaborative factor of STAT3 for cellular transformation. Many factors have been shown to interact with and activate (or in some cases inactivate) STAT3. For example, we have shown that the HCV core protein directly interacts with STAT3 and induces phosphorylation and activation (Yoshida *et al*., 2002). Nakayama *et al*. (2002) reported that a nuclear zinc-finger protein EZI enhances the nuclear retention and transactivation of STAT3. PIAS3 (Levy *et al*., 2002), cyclinD (Matsui *et al*., 2002), and GRIM-19 (Lufei *et al*., 2003) also interact directly with STAT3, but inhibit transcriptional activity. Most of these factors are isolated as a physical binding protein with STAT3. In this study, by using functional expression screening, we showed that STAT3 and MITF interact functionally, but not physically. A common target of STAT3 and MITF is found to be c-*fos*, which may participate in transformation by constitutively active STAT3 and MITF. Our functional cloning strategy using retroviral cDNA transfer will provide an additional candidate for the cofactor that modulates the function of STAT3. Furthermore, in this study we demonstrated that the microarray technique is a powerful tool to identify a target gene that is cooperatively induced by two distinct classes of transcription factors.

MITF consists of at least five isoforms, MITF-A, MITF-B, MITF-C, MITF-H, and MITF-M, differing at their N-termini and expression patterns (Tachibana, 1997; Udono *et al*., 2000; Shibahara *et al*., 2001). However, the clone (ΔN-MITF) we obtained from the HeLa cell library has an N-terminal deletion, which is shorter than other reported forms. Since mRNA corresponding to ΔN-MITF has not been reported, we speculate that ΔN-MITF is a product of the incomplete elongation of cDNA by reverse transcriptase. Nevertheless, ΔN-MITF seems to be a more potent inducer for a transformed phenotype of NIH-3T3 cells (Figure 1) and c-*fos* induction (Figure 4). In addition to the previously characterized acidic activation domain necessary for melanocyte differentiation, a second potential activation domain is shown to be located between amino acids 140 and 185 (Mansky *et al*., 2002), and ΔN-MITF contains these regions. Therefore, the N-terminal, with about 100 amino acids of MITF-M, may be a negative regulatory domain. This region contains a glutamine-rich basic region (QB), but the function of this region has not been elucidated. Most notably, Ser 73 of MITF-M is a predicted MAP kinase-phosphorylation site and is implicated in p300/CBP recruitment (Hemesath *et al*., 1998) and reduced MITF protein stability (Kim *et al*., 2003). Since ΔN-MITF lacks this Ser 73, dN-MITF may be more stable than wt-MITF.

Several types of functional interactions between STAT3 and MITF have been proposed. The protein inhibitor of activated STAT3 (PIAS3) has been shown to bind to a b-HLH-Zip domain of MITF, resulting in the suppression of MITF-induced transcriptional activity (Levy *et al*., 2002). However, it has been reported that STAT3 does not interfere, either *in vitro* or *in vivo*, with the interaction between PIAS3 and MITF. This finding is consistent with our data showing that STAT3 does not affect MITF transcriptional activity (data not shown) and MITF does not interfere with STAT3 transcriptional activity (Figure 3c). Recently, Kamaraju *et al*. (2002) reported that IL-6 receptor/IL-6 chimera induces a loss of melanogenesis preceded by a sharp decrease in MITF mRNA and gene promoter activity in B16/F10.9 melanoma cells. IL6RIL6 stimulates gp130, leading to the rapid activation of STAT3, which downregulates Pax3, a paired homeodomain factor regulating MITF mRNA levels and the development of melanocytes. Therefore, in this case, MITF downregulation by STAT3 is indirect, and the mechanism of Pax3 downregulation by STAT3 is not clear. The Pax3 downregulation in IL6RIL6-induced F10.9 cells leads to growth arrest and transdifferentiation to a glial cell phenotype. Therefore, the genetic interaction between MITF and STAT3 seems to be complicated and probably different among cell types.

We found that the c-*fos* gene is a common target of STAT3 and MITF, which probably contributes to transformation. Microarray analysis also revealed that

the c-*fos* gene is strongly upregulated by MITF in primary melanocytes (McGill *et al.*, 2002). In agreement with this notion, STAT3 was frequently phosphorylated in several human melanoma cell lines (Figure 3a). However, while B16 and SKMel cells are negative for STAT3 phosphorylation, the same cells do express c-*fos*. Apparently, STAT3-independent mechanism of c-*fos* induction is present in these melanoma cells. Finding such mechanisms of c-*fos* expression in these cells may provide a new clue for understanding the role of c-*fos* for the generation of melanomas.

Promoter analysis and ChIP assay indicated that STAT3 and MITF can bind to the c-*fos* promoter and activate this promoter independently. Reporter gene analysis suggested that the effect of MITF and STAT3 on the c-*fos* promoter are additive rather than cooperative (Figure 5b). This is consistent with our observation that there is no direct physical interaction between MITF and STAT3. Nevertheless, high levels of c-*fos* expression and transformation of NIH-3T3 cells are dependent on both STAT3 and MITF. In addition to promoter activation, another mechanism of c-*fos* expression, such as stabilization of c-*fos* mRNA, may present in the cooperative effect of STAT3 and MITF. Furthermore, a previous report demonstrated that wt- and mi-MITF bind to the c-Fos protein. In particular, mi-MITF prevents c-Fos from being transported to the nucleus, and this may be a reason for the osteopetrosis of mi/mi mutant mice (Sato *et al.*, 1999). wt-MITF was also shown to bind to c-Jun and enhance the transactivation of the MMCP-7 gene, while complexes of mi-MITF and c-Jun were predominantly found in the cytoplasm and suppressed transactivation (Ogihara *et al.*, 2001). Thus, MITF could be involved in AP-1 activation not only by inducing c-*fos* but also by interacting with AP-1 itself. The fact that a dominant-negative AP-1 (SupJunD) suppressed colony formation of transformants expressing both MITF and STAT3C suggests that AP-1 can be a therapeutic target of melanoma.

In addition to c-*fos*, we found several interesting genes that are cooperatively induced by STAT3 and MITF (Figure 4a). GTP cyclohydrase-1 and mast cell proteases are probably involved in melanocyte and mast cell functions. PMEL17 is also known as an MITF-inducible pigment cell-specific gene and a melanosomal matrix protein, which may function as a structural protein in melanogenesis (Kobayashi *et al.*, 1994). JE and A7 are chemokines that are induced by immunoglobulin (Ig) E plus antigen stimulation in mast cells. (Burd *et al.*, 1989; Kulmburg *et al.*, 1992; Ong *et al.*, 1993). Therefore, STAT3/MITF target genes are strongly linked to the function of mast cells and melanocytes, which constitutively express MITF. This finding implies that STAT3 (or other STATs) may be involved in mast cell and melanocyte functions by inducing genes in cooperation with MITF. It is interesting that IL3, which activates STAT5, is a mast cell growth factor *in vitro*. The functions of STATs in mast cells and melanocytes are under investigation.

Materials and methods

Cell culture and transfection

NIH3T3 and HEK293 cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% CS. STAT3C-transformed 3T3 (STAT3C-3T3) cells (Yoshida *et al.*, 2002) were cultured in DMEM supplemented with 10% CS containing 0.8 mg/ml G418. PLAT-E, a packaging cell line (Morita *et al.*, 2000) was maintained in DMEM with 10% FCS containing blasticidin S (10 mg/ml) and puromycin (1 mg/ml). Melanoma cell lines were cultured in the following medium containing 10% FCS, B16 and B16F10 in RPMI 1640, G361 and SKMel28 in MEM, HMV-II in Ham's F12, and MMAc in DMEM. MC9, a mast cell line, was cultured in RPMI 1640 supplemented with 5% FCS, 50 μM 2-mercaptoethanol, and 6 ng/ml mouse recombinant IL-3. Cells were transfected by the calcium phosphate method with Cell Phect (Amersham) or by the lipofection method with FuGENE 6 (Roche). For retrovirus-mediated gene expression, NIH3T3 cells were infected with the retroviruses produced by PLAT-E as reported (Sasaki *et al.*, 2003). B16, B16F10, G361, and SKMel 28 cell lines were kindly provided by Dr Yonemitsu (Kyushu University, Japan).

Library screening

Retroviruses containing the human HeLa retroviral cDNA library (CLONTECH, Palo Alto, CA, USA) were produced from PLAT-E and 1×10^6 STAT3C-3T3 cells were infected with 10 ml of the retroviral supernatant containing 10 mg/ml polybrene. After 48 h of virus transduction, STAT3C-3T3 cells were seeded in soft-agar medium. Within 3 weeks, two colonies were picked up and expanded, and genomic DNA was isolated. Integrated cDNAs were recovered by PCR using a pair of primers (FWD, 5′-AGCCCTCACTCCTTCTCTAG, and REV, 5′-ATGGCGTTACTTAAGCTAGCTTGCCAAACCTAC) and sequenced.

Colony formation in soft-agar

The MITF-infected 3T3 and STAT3C-3T3 cells were seeded into 35-mm dishes in suspensions of 0.36% Agar noble (Difco) in DMEM supplemented with 10% CS on top of a bed of 0.72% Agar noble in the same complete medium. The MITF-infected STAT3C-3T3 cells were inoculated with a retrovirus of dominant-negative AP-1 or an empty vector (Ui *et al.*, 2000), selected with puromycin, and plated into soft-agar medium. The cultures were incubated for 21 days and then the colonies were counted and photographed.

Construction of MITF

The cDNAs of wt-, ΔN-, and mi-MITF were subcloned into a bicistronic retrovirus vector pMX-IRES-EGFP (Nosaka *et al.*, 1999) or a pcDNA3 expression vector.

Retrovirus of dominant-negative AP-1

The control virus (pBabe-IRESpuro) and the SupJunD-1 virus (pBabe-SupJunD-1-IRESpuro) as a dominant-negative AP-1 were generated as described (Ui *et al.*, 2000). After infection, STAT3C-3T3 cells were cultured with 2 mg/ml puromycin.

High-density oligonucleotide microarray analysis

RNA was extracted by standard methods. Cells were lysed directly in their Petri dishes in TRIzol reagent (Invitrogen, Carlsbad, CA, USA) and total RNA was isolated according to

the manufacturer's instructions. cRNA preparation and microarray hybridization were carried out according to the supplier's instructions (Affymetrix, Santa Clara, CA, USA), using Genechip HG-U95Av2. Scanned output files were analysed by the probe level analysis package, Microarray Suite MAS 5.0 (Affymetrix, Santa Clara, CA, USA). The signal for each of these genes was determined from the 'probe set' in use for this gene and by the probe level analysis method provided by Affymetrix software.

Northern Blotting analysis and RT–PCR

Total RNA (10 μg) extracted from cells using TRIzol was evaluated by Northern blotting analysis with digoxigenin-labeled antisense RNA of c-*fos* and G3PDH labeled with the DIG RNA Labeling Kit (Roche). To determine the microarray data, 1 μg of total RNA was reverse-transcribed with the Reverse Transcribed Kit (Roche) and RT–PCR was performed with primers as follows: mouse MITF, 5′-GGAACAGCAAC-GAGCTAAGG and 5′-TGATGATCCGATTCACCAGA; human MITF, 5′-AGAACAGCAACGCGCAAAAGAAC and TGATGATCCGATTCACCAAATCTG; mouse GTP cyclohydrorase, 5′-GGCTGCTTACTCGTCCATTC and AG GTGATGCTCACACATGGA; mouse JE, AGGTCCCTGT CATGCTTCTG and TCTGGACCCATTCCTTCTTG; mouse PMEL17 CAGGGGTCTAACTGCTGGAG and TT CGGAGGTTTAGGACCAGA; mouse MRP8, GGAAAT CACCATGCCCTCTA and TGGCTGTCTTTGTGAGAT GC; mouse MARC, TCTGTGCCTGCTGCTCATAG and CTTTGGAGTTGGGGTTTTCA; and mouse MCPL, GCA CTTCTCTTGCCTTCTGG and TGTGCAGCAGTCATCA CAAA.

Western blotting analysis

Cells were lysed in a lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5% or 1% NP-40, 1 mM EDTA, 1 mM vanadate, 50 mM NaF, 1 mM DTT, 0.01 mM APMSF) and centrifuged at 12 000*g* for 10 min. The supernatants were resolved on SDS–PAGE of 10% gels, blotted, and immunostained with an anti-phospho STAT3 (Tyr705) polyclonal antibody (Cell Signaling Technology) and an anti-STAT3 polyclonal antibody (Santa Cruz Biotechnology).

Luciferase assay

An APRE-luciferase reporter gene for STAT3 activity and a reporter gene construct containing MITF-responsive mMCP-6 (mouse mast cell protease) luciferase have been described previously (Morii *et al*., 1996; Yasukawa *et al*., 1999). A c-*fos* promoter–reporter gene consisting of the c-*fos* promoter region containing SIE, SRE, and several potential MITF-binding motifs (Tsujimura *et al*., 1996) is also described (Shibuya *et al*., 1994; Kawahara *et al*., 1995; Kim *et al*., 1998). Luciferase assays were carried out using the dual-luciferase reporter system (Promega). The expression plasmid of STAT3-C in pRcCMV was kindly provided by Dr JE Darnell Jr (Bromberg *et al*., 1999).

Nuclear extract preparation and oligonucleotide-binding assay

293T cells (1.5×10^7) were transfected with Myc-ΔN-MITF and STAT3C were stimulate with LIF for 6 h. Cells were collected and resuspended in 0.4 ml of buffer A (10 mM HEPES-KOH (pH 7.8), 10 mM KCl, 0.1 mM EDTA, 0.4% NP-40, 1 mM DTT, 0.5 mM APMSF, 2 μg/ml leupeptin, and 1 mM Na_3VO_4). After a brief vortexing and centrifugation, the supernatant was discarded and the nuclei-containing pellet was resuspended in 0.05 ml of buffer C (50 mM HEPES-KOH (pH 7.8), 420 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 2% glycerol, 1 mM DTT, 0.5 mM APMSF, 2 μg/ml leupeptin, and 1 mM Na_3VO_4) at 4°C for 30 min. The suspension was pelleted by centrifugation and the supernatants were collected and stored at −80°C until use. Oligonucleotide-conjugated beads were prepared using the DNA-binding protein purification kit (Roche) according to the manufacturer's protocol. In brief, the sense and antisense oligonucleotide of the human c-*fos* promoter sequence, including the SIE and SRE regions (5′-C AGTTCCCGTCAATCCCTCCCCCCTTACACAGGATGTC CATATTAGGACATCTGC-3′) were annealed, and they were ligated with the washed streptavidin magnetic particles for 30 min at 25°C. After ligation was carried out, the particles were washed using a magnetic separator and mixed with 100 μg of the nuclear extract for 30 min at 25°C. Then, the particles were washed and the DNA-binding proteins were eluted. The elutates were pelleted with trichloroacetic acid solution, and analysed by immunoblotting with anti-STAT3 and anti-Myc antibodies.

ChIP assay

For ChIP assay, 243T cells (5×10^6) were fixed with 1% formaldehyde for 10 min at 37°C, washed with PBS, and lysed in ChIP lysis buffer (Upstate Biotechnology, Lake Placid, NY, USA). DNA was sonicated by pulsing five times. Anti-acetylated histone H4 (Upstate), anti-STAT3 (Sigma), or anti-MITF monoclonal antibody (Oncogene), and anti-MITF polyclonal antibody (Santa Cruz Biotechnology) was added (5 μg–20 μg per immunoprecipitation) and incubated overnight. Protein A–agarose beads (Upstate Biotechnology) were added for 1 h and then washed once each with a low-salt buffer, high-salt buffer, and LiCl buffer. They were then washed twice with a TE buffer. The beads were eluted with 0.1 M NaHCO3 and 1% SDS, and crosslinks were reversed at 65°C. DNA was ethanol precipitated in the presence of 20 μg glycogen. PCR was carried out using primers specific to the promoter region of mouse c-*fos* (FWD, 5′-TCTGCCTTTCCCGCCTCCCC, and REV, 5′-GGCCGTGGAAACCTGCTGAC) for NIH3T3, and human c-fos (FWD, 5′-CCCGACCTCGGGAACAAGGG, and REV, 5′-ATGAGGGGTTTCGGGGATGG) for HMV-II, or (FWD, 5′-TCTCATTCTGCGCCGTTCCC, and REV, 5′-GGCCGTGGAAACCTGCTGAC) for G361.

Acknowledgements
We thank Dr T Yoshida (Kurume University) for STAT3C-3T3, Dr Yonemitsu (Kyushu University) for melanoma cell lines, H Meguro for oligonucleotide microarray analysis, Y Kawabata for technical assistance, and N Arifuku for manuscript preparation. This work was supported by special grants-in-aid from the Ministry of Education, Science, Technology, Sports, and Culture of Japan, the Japan Health Science Foundation, the Human Frontier Science Program, the Japan Research Foundation for Clinical Pharmacology, Haraguchi Memorial foundation, and the Uehara Memorial Foundation. AJ is supported by the fellowship from Japan Society for Promotion of Science.

References

Bowman T, Garcia R, Turkson J and Jove R. (2000). *Oncogene*, **19**, 2474–2488.

Bromberg J and Darnell Jr JE. (2000). *Oncogene*, **19**, 2468–2473.

Bromberg JF, Wrzeszczynska MH, Devgan G, Zhao Y, Pestell RG, Albanese C and Darnell Jr JE. (1999). *Cell*, **98**, 295–303.

Burd PR, Rogers HW, Gordon JR, Martin CA, Jayaraman S, Wilson SD, Dvorak AM, Galli SJ and Dorf ME. (1989). *J. Exp. Med.*, **170**, 245–257.

Carreira S, Liu B and Goding CR. (2000). *J. Biol. Chem.*, **275**, 21920–21927.

Darnell Jr JE. (1997). *Science*, **277**, 1630–1635.

Hatakeyama M, Kawahara A, Mori H, Shibuya H and Taniguchi T. (1992). *Proc. Natl. Acad. Sci. USA*, **89**, 2022–2026.

Hemesath TJ, Price ER, Takemoto C, Badalian T and Fisher DE. (1998). *Nature*, **391**, 298–301.

Kamaraju AK, Bertolotto C, Chebath J and Revel M. (2002). *J. Biol. Chem.*, **277**, 15132–15141.

Kawahara A, Minami Y, Miyazaki T, Ihle JN and Taniguchi T. (1995). *Proc. Natl. Acad. Sci. USA*, **92**, 8724–8728.

Kim DW, Cheriyath V, Roy AL and Cochran BH. (1998). *Mol. Cell. Biol.*, **18**, 3310–3320.

Kim DS, Hwang ES, Lee JE, Kim SY, Kwon SB and Park KC. (2003). *J. Cell Sci.*, **116**, 1699–1706.

Kitamura Y, Morii E, Jippo T and Ito A. (2002). *Mol. Immunol.*, **38**, 1173.

Kitamura T, Onishi M, Kinoshita S, Shibuya A, Miyajima A and Nolan GP. (1995). *Proc. Natl. Acad. Sci. USA*, **92**, 9146–9150.

Kobayashi T, Urabe K, Orlow SJ, Higashi K, Imokawa G, Kwon BS, Potterf B and Hearing VJ. (1994). *J. Biol. Chem.*, **269**, 29198–29205.

Kulmburg PA, Huber NE, Scheer BJ, Wrann M and Baumruker T. (1992). *J. Exp. Med.*, **176**, 1773–1778.

Levy C, Nechushtan H and Razin E. (2002). *J. Biol. Chem.*, **277**, 1962–1966.

Lufei C, Ma J, Huang G, Zhang T, Novotny-Diermayr V, Ong CT and Cao X. (2003). *EMBO J.*, **22**, 1325–1335.

Mansky KC, Marfatia K, Purdom GH, Luchin A, Hume DA and Ostrowski MC. (2002). *J. Leukocyte Biol.*, **71**, 295–303.

Matsui T, Kinoshita T, Hirano T, Yokota T and Miyajima A. (2002). *J. Biol. Chem.*, **277**, 36167–36173.

McGill GG, Horstmann M, Widlund HR, Du J, Motyckova G, Nishimura EK, Lin YL, Ramaswamy S, Avery W, Ding HF, Jordan SA, Jackson IJ, Korsmeyer SJ, Golub TR and Fisher DE. (2002). *Cell*, **109**, 707–718.

Morii E, Tsujimura T, Jippo T, Hashimoto K, Takebayashi K, Tsujino K, Nomura S, Yamamoto M and Kitamura Y. (1996). *Blood*, **88**, 2488–2494.

Morita S, Kojima T and Kitamura T. (2000). *Gene Therapy*, **7**, 1063–1066.

Nakayama K, Kim KW and Miyajima A. (2002). *EMBO J.*, **21**, 6174–6184.

Nosaka T, Kawashima T, Misawa K, Ikuta K, Mui AL and Kitamura T. (1999). *EMBO J.*, **18**, 4754–4765.

Nyormoi O and Bar-Eli M. (2003). *Clin. Exp. Metastasis*, **20**, 251–263.

Ogihara H, Morii E, Kim DK, Oboki K and Kitamura Y. (2001). *Blood*, **97**, 645–651.

Ong EK, Griffith IJ, Knox RB and Singh MB. (1993). *Gene*, **134**, 235–240.

Sasaki A, Inagaki-Ohara K, Yoshida T, Yamanaka A, Sasaki M, Yasukawa H, Koromilas AE and Yoshimura A. (2003). *J. Biol. Chem.*, **278**, 2432–2436.

Sato M, Morii E, Takebayashi-Suzuki K, Yasui N, Ochi T, Kitamura Y and Nomura S. (1999). *Biochem. Biophys. Res. Commun.*, **254**, 384–387.

Shibahara S, Takeda K, Yasumoto K, Udono T, Watanabe K, Saito H and Takahashi K. (2001). *J. Invest. Dermatol. Symp. Proc.*, **6**, 99–104.

Shibuya H, Kohu K, Yamada K, Barsoumian EL, Perlmutter RM and Taniguchi T. (1994). *Mol. Cell. Biol.*, **14**, 5812–5819.

Stark GR, Kerr IM, Williams BR, Silverman RH and Schreiber RD. (1998). *Annu. Rev. Biochem.*, **67**, 227–264.

Tachibana M. (1997). *Pigment Cell Res.*, **10**, 25–33.

Tsujimura T, Morii E, Nozaki M, Hashimoto K, Moriyama Y, Takebayashi K, Kondo T, Kanakura Y and Kitamura Y. (1996). *Blood*, **88**, 1225–1233.

Turkson J, Bowman T, Garcia R, Caldenhoven E, De Groot RP and Jove R. (1998). *Mol. Cell. Biol.*, **18**, 2545–2552.

Udono T, Yasumoto K, Takeda K, Amae S, Watanabe K, Saito H, Fuse N, Tachibana M, Takahashi K, Tamai M and Shibahara S. (2000). *Biochim. Biophys. Acta*, **1491**, 205–219.

Ui M, Mizutani T, Takada M, Arai T, Ito T, Murakami M, Koike C, Watanabe T, Yoshimatsu K and Iba H. (2000). *Biochem. Biophys. Res. Commun.*, **278**, 97–105.

Vachtenheim J, Novotna H and Ghanem G. (2001). *J. Invest. Dermatol.*, **117**, 1505–1511.

Yasukawa H, Misawa H, Sakamoto H, Masuhara M, Sasaki A, Wakioka T, Ohtsuka S, Imaizumi T, Matsuda T, Ihle JN and Yoshimura A. (1999). *EMBO J.*, **18**, 1309–1320.

Yoshida T, Hanada T, Tokuhisa T, Kosai K, Sata M, Kohara M and Yoshimura A. (2002). *J. Exp. Med.*, **196**, 641–653.

Yu CL, Meyer DJ, Campbell GS, Larner AC, Carter-Su C, Schwartz J and Jove R. (1995). *Science*, **269**, 81–83.

Zanocco-Marani T, Bateman A, Romano G, Valentinis B, He ZH and Baserga R. (1999). *Cancer Res.*, **59**, 5331–5340.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computational Biology and Chemistry
ELSEVIER
Computational Biology and Chemistry 28 (2004) 401–407
www.elsevier.com/locate/compbiolchem

Database Note

Development of *Ki*Bank, a database supporting structure-based drug design

Junwei Zhang[a,*], Masahiro Aizawa[a], Shinji Amari[a], Yoshio Iwasawa[b], Tatsuya Nakano[c], Kotoko Nakata[c]

[a] *Collaborative Research Center of Frontier Simulation Software for Industrial Science, Institute of Industrial Science, University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan*
[b] *AdvanceSoft Corporation, NISSAY Bldg. 4F, 16-2 Samon-cho, Shinjyuku-ku, Tokyo 160-0017, Japan*
[c] *Division of Safety Information on Drug, Food and Chemicals, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan*

Received 1 September 2004; received in revised form 13 September 2004; accepted 15 September 2004

Abstract

*Ki*Bank is a database of inhibition constant (K_i) values with 3D structures of target proteins and chemicals. K_i values were accumulated from peer-reviewed literature searched via PubMed. The 3D structure files of target proteins were originally from Protein Data Bank (PDB), while the 2D structure files of the chemicals were collected together with the K_i values and then converted into 3D ones. In *Ki*Bank, the chemical and protein 3D structures with hydrogen atoms were optimized by energy minimization and stored in MDL MOL and PDB format, respectively.

*Ki*Bank is designed to support structure-based drug design. It provides structure files of proteins and chemicals ready for use in virtual screening through automated docking methods, while the K_i values can be applied for tests of docking/scoring combinations, program parameter settings, and calibration of empirical scoring functions. Additionally, the chemical structures and corresponding K_i values in *Ki*Bank are useful for lead optimization based on quantitative structure–activity relationship (QSAR) techniques.

*Ki*Bank is updated on a daily basis and is freely available at http://kibank.iis.u-tokyo.ac.jp/. As of August 2004, *Ki*Bank contains 8000 K_i values, over 6000 chemicals and 166 proteins covering the subtypes of receptors and enzymes.
© 2004 Elsevier Ltd. All rights reserved.

Keywords: Target protein; Inhibition constant; Database; Drug design

1. Introduction

Since last decade, structure-based drug design (SBDD) has become a mature discipline of medicinal chemistry (Anderson, 2003; Böhm, 1996; Klebe, 2000; Nakata, 2002), and development of computer technologies to calculate molecular properties, of combinatorial chemistry and abundant data on target proteins coming from human genome research have opened new opportunities and feasible approaches for drug discovery (Bailey and Brown, 2001; Kirkpatrick et al., 1999).

Estimation of the binding affinity of novel chemicals to target proteins is a critical procedure in computational approaches to drug design including SBDD. The strategies that can be applied for this purpose fall in to two major categories: a target-based approach and a ligand-based approach. Recently, some researchers have combined both of these approaches in an automated unbiased procedure (Dean et al., 2004; Loew et al., 1993; Sippl, 2002b). The former can be used if the 3D structure of the binding site is known as is the case of SBDD. In practice, in silico screening of chemical databases is widely applied to find lead candidates for target proteins. Each of the reported methods has two steps, docking and scoring (Ewing et al., 2001; Goodsell et al., 1996; Jones et al., 1997; Rarey et al., 1996). Although several scoring methods for estimating binding affinity have been documented, it

* Corresponding author. Tel.: +81 3 5452 6622; fax: +81 3 5452 6623.
E-mail address: cho@fsis.iis.u-tokyo.ac.jp (J. Zhang).

1476-9271/$ – see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiolchem.2004.09.003

is not yet clear which docking/scoring combinations will provide the best accuracy. Therefore, before beginning to screen the entire chemical database, it is necessary to test docking/scoring combinations and program parameter settings by a test screening of a reduced database including known ligands. Experimental binding affinity values are also needed for the calibration of most scoring functions (Bissantz et al., 2000; Schneider and Böhm, 2002). On the other hand, traditional quantitative structure–activity relationship (QSAR) and modern 3D QSAR techniques are widely used in the ligand-based approach when the target structure is unknown (Akamatsu, 2002; Loew et al., 1993; Yoo et al., 2000). In the past few years, QSAR techniques have also been used in combination with structure-based methods (Lozano et al., 2000; Sippl, 2002b; Vaz et al., 1998). The QSAR techniques are based on experimental structure–activity relationships, and thus require large amount of experimental data.

Generally, to perform SBDD, 3D structures of target proteins are definitely needed, while experimental data is indispensable for accurate estimation of the binding affinities of newly designed chemicals. Although a database including such data will exactly facilitate this drug discovery approach, to our knowledge, it has been nonexistent up to now. Therefore, we developed *Ki*Bank providing K_i values and 3D structure files of chemicals and proteins ready for use in SBDD.

2. Methods

2.1. Creating KiBank

*Ki*Bank is a PostgreSQL database consisting of three knowledge areas — binding affinity data, chemical data and protein data (Aizawa et al., 2004) (Fig. 1).

As the binding affinity data, inhibition constant (K_i) values were accumulated from peer-reviewed literature searched via PubMed by using the name of a target protein (e.g., androgen receptor) and "K_i" as search terms. Articles published from 1985, with the majority from around the year 2000, were then selected for input into *Ki*Bank through intensive reading. For some of the nuclear receptors, supplementary search was performed by using the names of target proteins and "inhibition constant" or "binding affinity". Other than K_i values, experimental details such as species, tissue, tracer (substrate), pH, and buffer were also extracted, while the PMID (PubMed ID) was used to indicate the origin of the data set.

The chemical data consist of chemical name, molecular weight, and structure file. The 2D chemical structure files in the literature were first reproduced with ChemDraw Ultra (version 7.0, CambridgeSoft Co., USA) and then converted into 3D structures with Chem3D Ultra (version 7.0, CambridgeSoft Co., USA). Finally, the 3D chemical structures with hydrogen atoms were optimized by energy minimization and stored in MDL MOL format.

The protein data comprise nucleotide sequence, amino acid sequence, and structure file. Currently noted drug targets were selected as protein entries (Alexander et al., 2001). Three-dimensional structure files of target proteins were initially downloaded from the Protein Data Bank (PDB) (Berman et al., 2000) and then completed by using side chain rotamer prediction (Andrec et al., 2002). The protein structures with hydrogen atoms added were optimized by energy minimization and stored in PDB format. The pages of protein data were linked to source websites to provide up to date information, for amino acid sequence data, to the Protein Information Resource (PIR) (McGarvey et al., 2000) and Swiss Prot (Boeckmann et al., 2003); for cDNA sequence data, to

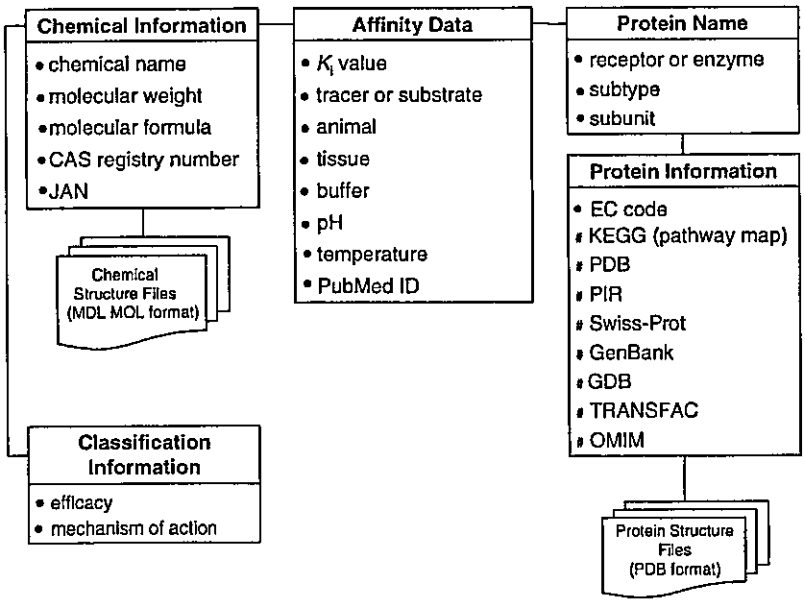


Fig. 1. A schematic illustration of the contents of *Ki*Bank. JAN: Japanese Accepted Names for Pharmaceuticals; #: the linked resources for protein information; GDB: the genome database; TRANSFAC: the transcription factor database; OMIM: online Mendelian inheritance in man; PIR: protein information resource; PDB: protein data bank; KEGG: Kyoto encyclopedia of genes and genomes.

GenBank (Benson et al., 2003); and for human gene data, to the genome database (GDB) (Cuticchia, 2000). Protein information also contains links to transcription factor data by TRANSFAC (Matys et al., 2003), to single nucleotide polymorphism data by dbSNP (Sherry et al., 2001), to human gene and genetic disorder data by OMIM (Hamosh et al., 2002) and to pathway maps by KEGG (Kanehisa et al., 2002).

2.2. Accessing and searching KiBank

*Ki*Bank has a Java-based web interface at http://kibank.iis.u-tokyo.ac.jp/. Internet Explorer 6.0 or later is recommended, and installation of MDL Chime (MDL Information Systems Inc., USA) is essential to view the 3D images of chemicals and proteins. The "Search *Ki*Bank" button on the top page will lead users to the search page of *Ki*Bank. From this page, for instance, users can access a page of binding affinity data by first selecting a protein name from the pull down menu, and then checking the binding affinity data and clicking the protein search button. The pages of protein data, chemical data, and 3D structure images are linked from the page of binding affinity data (Fig. 2). The chemical data and protein data are also searchable from the search page by selecting the names in each given list.

On the page of binding affinity data, a check box was placed at the left side of each record. Users can select and see structure files, K_i values, and experimental details of chemicals they are interested in by checking these boxes and clicking on the "view checked structure" button at the bottom of this page (Fig. 2A and F).

2.3. Assessing the drug-like property of the chemicals with structure files in KiBank

To identify the suitability of using the chemicals having 3D structure files available in *Ki*Bank for drug discovery, the drug-like properties, namely, molecular weight, lipophilicity, and hydrogen bonding potential, were calculated by using molecular operating environment (MOE, Chemical Computing Group Inc., Canada). The results were judged with Lipinski's "rules of five" (Lipinski, 2000).

3. Results and discussion

As of August 2004, *Ki*Bank contains 166 proteins covering the subtypes of receptors and enzymes, over 6000 chemicals and 8000 K_i values.

A drug is effective when it binds more specifically and tightly to the target protein against natural ligands in a competitive fashion (McIlwain, 1986). Thus consideration of chemicals' binding competitiveness is important for the computational approaches to drug design. Experimental binding affinities are reported as the inhibition constant (K_i), relative binding affinity (RBA), inhibition concentration for 50% of the organism exposed (IC_{50}), and dissociation constant (K_D). We collected K_i values because they are not theoretically influenced by the property of the tracers (substrates) used in the experiments. Thus K_i values reported in different literature for the same target protein are comparable, and more identical experimental situations make the comparison more accurate.

Our purpose of searching PubMed is to collect enough and useful K_i values, not all of them from the publications for input into *Ki*Bank. Success in doing it revealed effectiveness of our literature retrieval method. Actually, if "K_i" was used as a search term, it would be searched in All Fields of PubMed including abstract. According to our experience, if the heart of the reported work was concerning about ligand binding affinity and K_i values were obtained as results, a wording containing "K_i" would almost appeared in the abstract. Therefore, most of the needed literature was included in the search results, and intensive reading of the abstracts helped us to take out the inappropriates. For some of the nuclear receptors, in order to increase the number of K_i values, additional search was performed by using the name of the target protein and "inhibition constant" or "binding affinity" as search terms, and only a small amount of literature was found.

Some other well-built databases concerning lead finding and optimization in drug discovery were investigated before the development of *Ki*Bank. Among which, RDB (Nakata et al., 1999) and TTD (Chen et al., 2002) provides the target protein information. Integrity (Prous Science, S.A., Barcelona, Spain) is a fully integrated knowledge area related to therapeutics and drug R&D, and the K_i database (Roth et al., 2000) currently deposits K_i values for drugs and drug candidates active in the central nervous system. To our knowledge, however, none of these databases was developed with the aim to support structure-based drug design. If used for virtual screening, it would be difficult to select and extract data from these databases. Alternatively, *Ki*Bank was developed to provide a wide scope of easily sorted and ready-to-use data sets, including target protein information, chemicals structures, and K_i values, all of which are freely available at all times.

As described in Section 2, the accumulated 2D chemical structures need to first be converted into 3D ones, followed by the addition of hydrogen atoms and energy minimization. Currently, these manipulations are completed for 2742 chemicals. The drug-like properties of these chemicals, which were mostly reported as analogues of the lead candidates, were analyzed. The results were judged by using Lipinski's "rules of five". Briefly, among 71% of the chemicals, $\log P \leq 5$; among 87% of the chemicals, molecular weight ≤ 500; among 97% of the chemicals, hydrogen bond donors ≤ 5; and among 99% of the chemicals, hydrogen bond acceptors ≤ 10. Furthermore, fully 65% of the chemicals met all of these criteria, indicating that these chemicals could be used as candidates or reference data for drug discovery (Fig. 3).

Most of the recent docking programs such as GOLD (Jones et al., 1997) and AutoDock (Goodsell et al., 1996) consider the ligand to be a flexible molecule and dock it

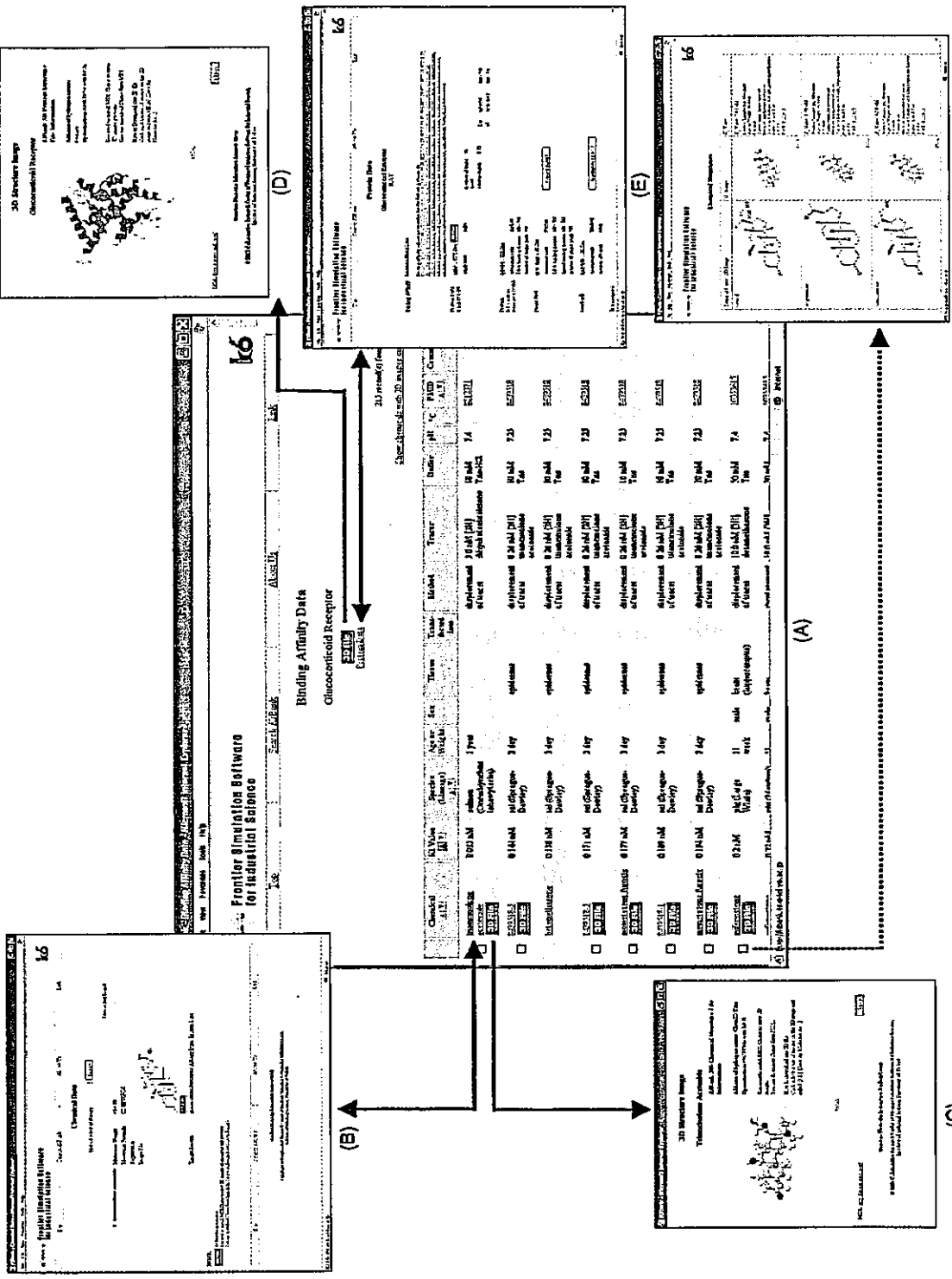


Fig. 2. An example of *Ki*Bank searching results: glucocorticoid receptor and related information: (A)–(F) show the pages of binding affinity data, chemical data, chemical 3D structure, protein 3D structure, protein data, and view checked chemicals, respectively. The arrows indicate the linkage of the pages.

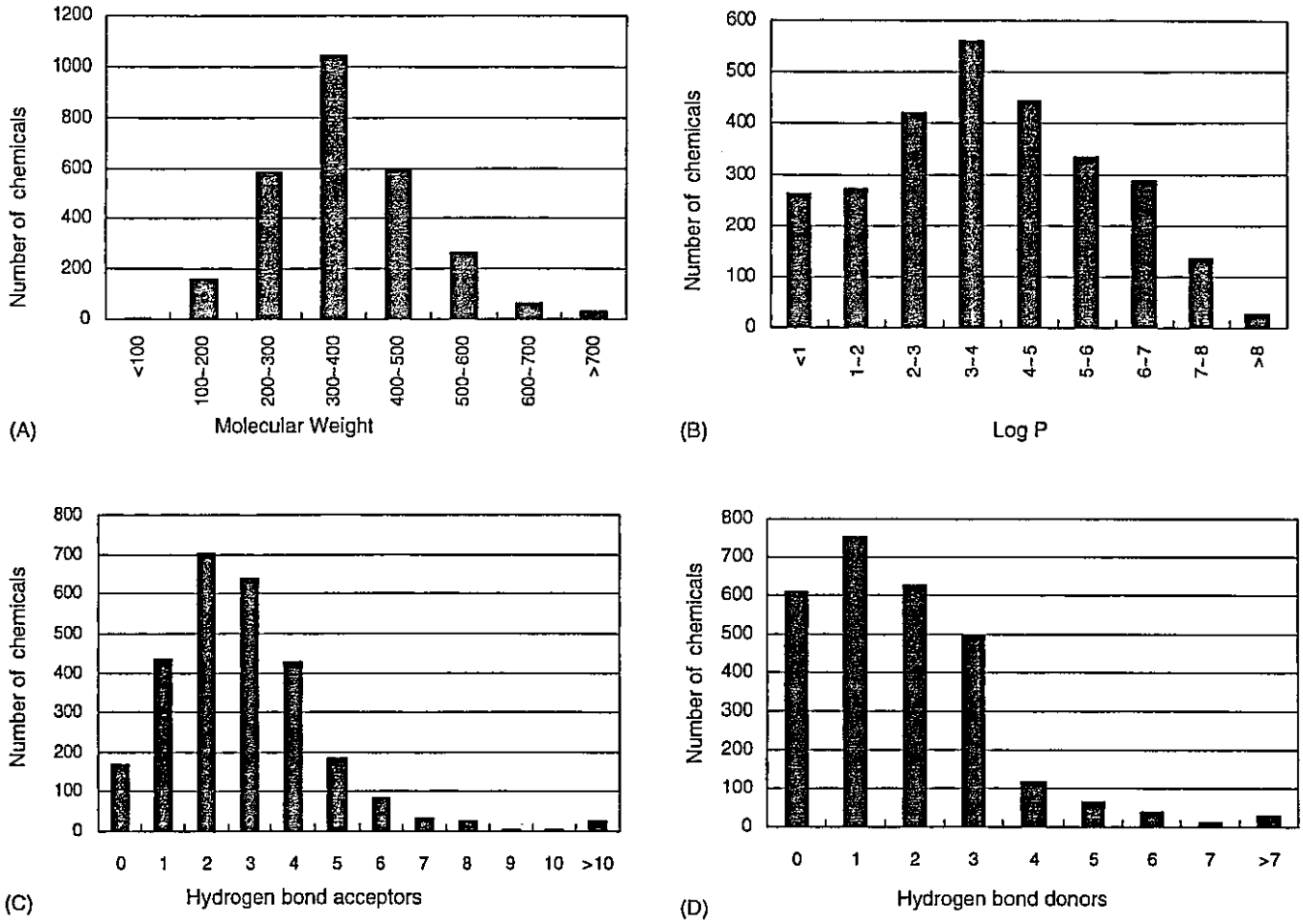


Fig. 3. Number distribution of the 2742 chemicals having 3D structure files in *Ki*Bank according to the drug-like properties: (A) molecular weight; (B) lipophilicity, expressed as log partition coefficient; (C) and (D) H-bonding potential, expressed as the number of H-bond acceptors and donors, respectively.

to the target protein binding site. Once the ligand has been docked, it should be scored according to the tightness of the target–ligand interaction (Bissantz et al., 2000). Experimental data are needed for the calibration of some scoring functions and parameter settings prior to the screening (Schneider and Böhm, 2002). K_i values and the structure files of proteins and chemicals in KiBank are useful in this procedure.

The QSAR approach is a rational approach to lead optimization when the structure of the target is not known. The underlying premise of QSAR is that there is a relationship between the biological and pharmacological activities of a compound and its structural, physical and chemical properties (Selassie et al., 2002). As an extension of the QSAR approach, comparative molecular field analysis (CoMFA), which is based on a congeneric series of molecules, is a very widely used 3D QSAR technique. Originally, the CoMFA approach was developed for ligand data sets where the structure of the target protein is unknown. In the past few years, however, it has also been applied to cases where the corresponding target structure has been determined. The superimposition of ligands derived from molecular docking runs is taken as the starting point for a CoMFA study. If one already has a series of molecules and their corresponding binding affinities, then the 3D QSAR equation may provide a valuable method to forecast affinity of further analogues (Kellogg and Semus, 2003; Sippl, 2002a, 2002b). In *Ki*Bank, there is growing number of chemicals (named using a PMID and the number of the chemical in the publication, e.g., 9990464-7), which are analogues synthesized based on the structure of lead candidates. These sets of data, the chemical 2D and 3D structures and known properties (K_i values), are thought to be useful for QSAR and 3D QSAR approaches. Moreover, to make it easy for users to compare the structures of lead candidates or substituents and to reference the K_i values, a check box was placed at the left side of each record on the page of binding affinity data (Fig. 2A and F). For example, if users want to study the relationship between the K_i values and structures of substituents, they can first sort the K_i value records by PMID, then check the records from the same literature, and so obtain the wanted set of data as it was described in the methods. Similarly, if users checked the records of lead candidates from different literature, they can compare various types of lead structures for the ligand candidates to the same drug target.

Furthermore, any candidate ligands that were designed using a combination of in silico screening and QSAR techniques need to undergo post-screening selection or identification before being used for experimental tests. This provides a very active research field, and some researchers attempted to carry it out by calculating the binding energy with ab initio fragment molecular orbital (FMO) method, a quantum mechanical calculation based on docking strategy (Fukuzawa et al., 2003; Komeiji et al., 2004; Nakano et al., 2000, 2002). Structure files and experimental data in *Ki*Bank are useful for testing such algorithms, and, if such a program were validated, our data could also be used to estimate the K_i values of newly designed ligands, through the relationship between the K_i values and calculated results.

*Ki*Bank is updated on a daily basis and will be expanded to include ion channels and transporters in the near future. New data of the structures of targets and chemicals and K_i values available from experimental studies will be added into *Ki*Bank as fast as possible, while supplementary system functions will be added in order to make it a more user-friendly database for SBDD.

Acknowledgements

This research was done under the "Frontier Simulation Software for Industrial Science (FSIS)" project supported by the IT program of the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT), and was partly supported by the Toxico-proteomics project fund from the Japanese Ministry of Health, Labour and Welfare.

References

Aizawa, M., Onodera, K., Zhang, J., Amari, S., Iwasawa, Y., Nakano, T., Nakata, K., 2004. *Ki*Bank: a database for computer-aided drug design based on protein–chemical interaction analysis. Yakugaku Zasshi 124, 613–619.

Akamatsu, M., 2002. Current state and perspectives of 3D-QSAR. Curr. Top. Med. Chem. 2, 1381–1394.

Alexander, S.P.H., Mathie, A., Peters, J.A., 2001. TiPS Nomenclature Supplement, 12th ed., Elsevier Current Trends, London.

Anderson, A.C., 2003. The process of structure-based drug design. Chem. Biol. 10, 787–797.

Andrec, M., Harano, Y., Jacobson, M.P., Friesner, R.A., Levy, R.M., 2002. Complete protein structure determination using backbone residual dipolar couplings and side chain rotamer prediction. J. Struct. Funct. Genom. 2, 103–111.

Bailey, D., Brown, D., 2001. High-throughput chemistry and structure-based design: survival of the smartest. Drug Discov. Today 6, 57–59.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Wheeler, D.L., 2003. GenBank. Nucl. Acids Res. 31, 23–27.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucl. Acids Res. 28, 235–242.

Bissantz, C., Folkers, G., Rognan, D., 2000. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. J. Med. Chem. 43, 4759–4767.

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S., Schneider, M., 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003 31, 365–370.

Böhm, H.J., 1996. Computational tools for structure-based ligand design. Prog. Biophys. Mol. Biol. 66, 197–210.

Chen, X., Ji, Z.L., Chen, Y.Z., 2002. TTD: therapeutic target database. Nucl. Acids Res. 30, 412–415.

Cuticchia, A.J., 2000. Future vision of the GDB human genome database. Human Mutat. 15, 62–67.

Dean, P.M., Lloyd, D.G., Todorov, N.P., 2004. De novo drug design: integration of structure-based and ligand-based methods. Curr. Opin. Drug Discov. Develop. 7, 347–353.

Ewing, T.J., Makino, S., Skillman, A.G., Kuntz, I.D., 2001. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. J. Comput. Aid. Mol. Des. 15, 411–428.

Fukuzawa, K., Kitaura, K., Nakata, K., Kaminuma, T., Nakano, T., 2003. Fragment molecular orbital study of the binding energy of ligands to the estrogen receptor. Pure Appl. Chem. 75, 2405–2410.

Goodsell, D.S., Morris, G.M., Olson, A.J., 1996. Automated docking of flexible ligands: applications of AutoDock. J. Mol. Recog. 9, 1–5.

Hamosh, A., Scott, A.F., Amberger, J., Bocchini, C., Valle, D., McKusick, V.A., 2002. Online Mendelian inheritance in man (OMIM), a knowledgebase of human genes and genetic disorders. Nucl. Acids Res. 30, 52–55.

Jones, G., Willett, P., Glen, R.C., Leach, A.R., Taylor, R., 1997. Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol. 267, 727–748.

Kanehisa, M., Goto, S., Kawashima, S., Nakaya, A., 2002. The KEGG databases at GenomeNet. Nucl. Acids Res. 30, 42–46.

Kellogg, G.E., Semus, S.F., 2003. 3D QSAR in modern drug design. EXS, 223–241.

Kirkpatrick, D.L., Watson, S., Ulhaq, S., 1999. Structure-based drug design: combinatorial chemistry and molecular modeling. Comb. Chem. High Throughput Screen. 2, 211–221.

Klebe, G., 2000. Recent developments in structure-based drug design. J. Mol. Med. 78, 269–281.

Komeiji, Y., Inadomi, Y., Nakano, T., 2004. PEACH 4 with ABINIT-MP: a general platform for classical and quantum simulations of biological molecules. Comput. Biol. Chem. 28, 155–161.

Lipinski, C.A., 2000. Drug-like properties and the causes of poor solubility and poor permeability. J. Pharmacol. Toxicol. Meth. 44, 235–249.

Loew, G.H., Villar, H.O., Alkorta, I., 1993. Strategies for indirect computer-aided drug design. Pharm. Res. 10, 475–486.

Lozano, J.J., Pastor, M., Cruciani, G., Gaedt, K., Centeno, N.B., Gago, F., Sanz, F., 2000. 3D-QSAR methods on the basis of ligand–receptor complexes. Application of COMBINE and GRID/GOLPE methodologies to a series of CYP1A2 ligands. J. Comput. Aid. Mol. Des. 14, 341–353.

Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E., Kel-Margoulis, O.V., Kloos, D.U., Land, S., Lewicki-Potapov, B., Michael, H., Munch, R., Reuter, I., Rotert, S., Saxel, H., Scheer, M., Thiele, S., Wingender, E., 2003. TRANSFAC: transcriptional regulation, from patterns to profiles. Nucl. Acids Res. 31, 374–378.

McGarvey, P.B., Huang, H., Barker, W.C., Orcutt, B.C., Garavelli, J.S., Srinivasarao, G.Y., Yeh, L.S., Xiao, C., Wu, C.H., 2000. PIR: a new resource for bioinformatics. Bioinformatics 16, 290–291.

McIlwain, H., 1986. The origin and use of the terms competitive and non-competitive in interactions among chemical substances in biological systems. Essays Biochem. 22, 158–186.

Nakano, T., Kaminuma, T., Sato, T., Akiyama, Y., Uebayasi, M., Kitaura, K., 2000. Fragment molecular orbital method: application to polypeptides. Chem. Phys. Lett. 318, 614–618.

Nakano, T., Kaminuma, T., Sato, T., Fukuzawa, K., Akiyama, Y., Uebayasi, M., Kitaura, K., 2002. Fragment molecular orbital method:

use of approximate electrostatic potential. Chem. Phys. Lett. 351, 475–480.

Nakata, K., 2002. Theoretical approach to endocrine disruptors. Front Biosci. 7, c68–c73.

Nakata, K., Takai, T., Kaminuma, T., 1999. Development of the receptor database (RDB): application to the endocrine disruptor problem. Bioinformatics 15, 544–552.

Rarey, M., Kramer, B., Lengauer, T., Klebe, G., 1996. A fast flexible docking method using an incremental construction algorithm. J. Mol. Biol. 261, 470–489.

Roth, B.L., Lopez, E., Patel, S., Kroeze, W.K., 2000. The multiplicity of serotonin receptors: uselessly diverse molecules or an embarrassment of riches? Neuroscientist 6, 252–262.

Schneider, G., Böhm, H.J., 2002. Virtual screening and fast automated docking methods. Drug Discov. Today 7, 64–70.

Selassie, C.D., Mekapati, S.B., Verma, R.P., 2002. QSAR: then and now. Curr. Top. Med. Chem. 2, 1357–1379.

Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., Sirotkin, K., 2001. dbSNP: the NCBI database of genetic variation. Nucl. Acids Res. 29, 308–311.

Sippl, W., 2002a. Binding affinity prediction of novel estrogen receptor ligands using receptor-based 3D QSAR methods. Bioorg. Med. Chem. 10, 3741–3755.

Sippl, W., 2002b. Development of biologically active compounds by combining 3D QSAR and structure-based design methods. J. Comput. Aid. Mol. Des. 16, 825–830.

Vaz, R.J., McLean, L.R., Pelton, J.T., 1998.xy]phenyl]-3-(7-am idino-2-naphthyl)propanoic acid hydrochloride and some analogs to factor Xa using a comparative molecular field analysis J. Comput. Aid. Mol. Des. 12, 99–110.

Yoo, S.E., Kim, S.K., Lee, S.H., Kim, N.J., Lee, D.W., 2000. The conformation and activity relationship of benzofuran type of angiotensin II receptor antagonists. Bioorg. Med. Chem. 8, 2311–2316.

—Regular Articles—

*Ki*Bank：創薬のためのタンパク質—化合物相互作用解析支援データベース

愛澤昌宏,*,[a] 小野寺賢司,[a] 張　軍衛,[a] 甘利真司,[a]
岩澤義郎,[b] 中野達也,[c] 中田琴子[c]

*Ki*Bank: A Database for Computer-Aided Drug Design Based on Protein-Chemical Interaction Analysis

Masahiro AIZAWA,*,[a] Kenji ONODERA,[a] Junwei ZHANG,[a] Shinji AMARI,[a]
Yoshio IWASAWA,[b] Tatsuya NAKANO,[c] and Kotoko NAKATA[c]

Collaborative Research Center of Frontier Simulation Software for Industrial Science, Institute of Industrial Science, University of Tokyo,[a] 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan, AdvanceSoft Corporation,[b] NISSAY Yotsuya Bldg. 4F, 16-2 Samon-cho, Shinjuku-ku, Tokyo 160-0017, Japan, and Division of Safety Information on Drug, Food and Chemicals, National Institute of Health Sciences,[c] 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

(Received April 7, 2004; Accepted June 23, 2004; Published online June 24, 2004)

*Ki*Bank is a database for computer-aided drug design and consists of binding affinities and chemical and target protein structures. Each chemical or protein structure with hydrogen atoms added was optimized by energy minimization and stored in PDB or MDL MOL file format, so that the structural data can be directly used for *in silico* binding studies. To describe the extent of inhibition, the inhibition constant (K_i) value is used to simplify comparisons of strengths among chemical-protein bindings. As of April 2004, *Ki*Bank contained 142 proteins, over 5000 chemicals, and over 6000 binding affinity values that were published in peer-reviewed journals. The binding affinity values are currently mostly for membrane and nuclear receptors but are soon being expanded to other drug targets. *Ki*Bank is updated daily and can be accessed on the Web at http://kibank.iis.u-tokyo.ac.jp/at no charge.

Key words——inhibition constant (K_i) database; protein structure; chemical structure; computer-aided drug design; internet

緒　　言

近年のコンピュータ技術の進歩はめざましく，一昔前のスーパーコンピュータと同程度の性能のコンピュータを，パーソナルコンピュータとして個人で利用できるようになってきている．並列計算技術の進歩と相まって，コンピュータの高速化はますます進んでいる．このため，例えば第一原理(量子力学)に基づいた化合物とタンパク質の相互作用解析のような，かつては不可能であった複雑な計算も実用化の段階にまできている．

医薬品のような生理活性を有する化合物の多くは，受容体，酵素，トランスポーター等の生体内タンパク質を標的として結合し，その活性を発現する．したがって医薬品の開発は，標的タンパク質に対していかに高い特異性を持つ化合物を作ることができるかにかかっていると言っても過言ではない．医薬品が標的とするタンパク質は約500種類あると言われており，[1] 新たに合成したすべての化合物とすべてのタンパク質の組み合わせについて従来の方法でスクリーニングを実施することは，例えそれが比較的簡単な *in vitro* スクリーニングであったとしても，膨大な時間とコストがかかってしまい非現実的である．しかし，仮想的なスクリーニングシステムをコンピュータ上（*in silico*）に構築することができれば，*in vitro* よりもはるかに高速にスクリーニングを行うことが可能となる．このようなComputer-aided Drug Designと呼ばれるコンピュータを使用した医薬品開発の重要性はますます増大している．

[a] 東京大学生産技術研究所計算科学技術連携研究センター，[b] アドバンスソフト株式会社，[c] 国立医薬品食品衛生研究所安全情報部
e-mail: aizawa@fsis.iis.u-tokyo.ac.jp

そこで，本研究では *in silico* スクリーニングに基づいた医薬品開発を支援するためのデータベースを構築することを目的とした．

方　　法

1. データベースの構築　データベース管理ソフトウェアとして PostgreSQL バージョン 7.2.3 Cygwin 版を採用した．データベースサーバには IBM-PC/AT 互換機（CPU: Intel Pentium 4, 2.00 GHz, RAM: 1 GB, オペレーティングシステム (OS)：Windows XP Professional (Microsoft Co., U.S.A.)）を用いた．ここに Cygwin (Red Hat, Inc., U.S.A.) をインストールし，この Cygwin に PostgreSQL をインストールしてデータベースシステムを構築した．

*Ki*Bank は「化合物と受容体の結合親和性に関するデータ」，「化合物とその三次元座標に関するデータ」，「タンパク質とその三次元座標に関するデータ」の 3 つの基幹と成るデータとそれらを補足するデータから構成される (Fig. 1)．化合物とタンパク質の結合親和性を示す値としては IC_{50}，結合阻害定数 (K_i)，相対的結合能 (Relative Binding Affinity, RBA) 等が広く用いられているが，*Ki*Bank では結合親和性の指標として結合実験で得られた K_i 値を採用した．

K_i 値は，「受容体名」と「Ki」をキーワードとして PubMed[2] で文献検索を行い，なるべく新しいものから受容体結合実験で得られた K_i 値を実験条件データとともにデータベースに格納した．また，IC_{50} 値と K_D 値が得られている場合には，計算によって K_i 値を算出し，参考データとして格納した．さらに，一部，Binding Affinity Database[3] 中の情報も利用した．

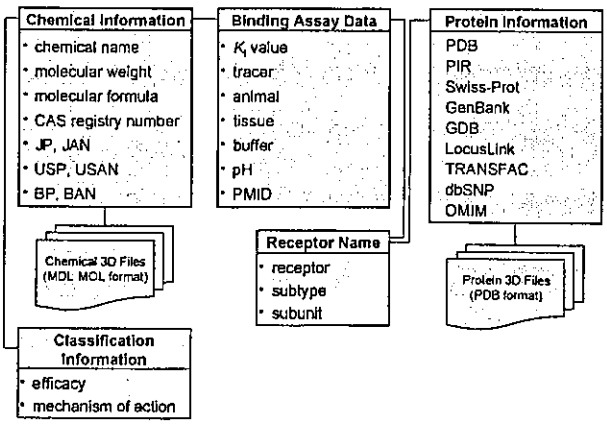


Fig. 1. A Schematic Diagram of the *Ki*Bank Database Structure

JP: Japanese Pharmacopoeia, JAN: Japanese Accepted Names for Pharmaceuticals, USP: United States Pharmacopoeia, USAN: United States Adopted Names, BP: British Pharmacopoeia, BAN: British Approved Names, PDB: Protein Data Bank, PIR: Protein Information Resource, GDB: The Genome Database, TRANSFAC: The Transcription Factor DataBase, OMIM: Online Mendelian Inheritance in Man.

受容体と化合物の構造に関するデータは Receptor Database[4] と日本医薬品一般名称データベース[5]を利用した．

2. 化合物三次元座標データの作成　化合物の三次元座標データは，ChemDraw Ultra (CambrigeSoft Co., U.S.A.) バージョン 7.0 で作成した構造式を Chem3D Ultra バージョン 7.0 (CambrigeSoft Co., U.S.A.) で読み込み，立体構造に変換したのちに，統合計算化学システム Molecular Operating Environment (MOE, Chemical Computing Group Inc., Canada) を使用して MMFF94s 力場[6]で構造最適化を行い，MDL MOL 形式で保存した．

3. タンパク質三次元座標データの作成　タンパク質の三次元座標データは，Protein Data Bank (PDB)[7]から基となる座標データをダウンロードし，不足している水素原子を Reduce[8] で補ったのちに，付加した水素原子に対して AMBER94 力場[9]で構造最適化を行い，PDB 形式で保存した．

PDB からダウンロードした座標データ中に原子座標の欠落がある場合には，MOE の残基補完機能を用いて，欠落している原子座標の補完を行い，側鎖の自由回転の自由度については，MOE のロータマーサーチ機能[10-12]を用いて安定な側鎖のコンフォメーションを決定した．

4. データベースへのアクセス　*Ki*Bank を一般に公開するための Web サーバを構築した．HTTP サーバソフトウェアとして Apache 2.0.46 (Apache Software Foundation, U.S.A.) と Tomcat 4.1 (Apache Software Foundation, U.S.A.) を採用した．サーバマシンは IBM-PC/AT 互換機 (CPU: Dual Pentium II Xeon, 450 MHz, RAM: 1 GB, OS: Windows 2000 Professional (Microsoft Co., U.S.A.)) を使用した．

Web ページは，ユーザーが *Ki*Bank 内のデータを検索できるように Java (Sun Microsystems, Inc., U.S.A.) によって Java Server Pages 又は Java Serv-