Table 3
IC_{50} values of inhibitors

	Inhibitors			
Tumor cells	iCRT14 (μM)	PNU74654 (μM)	DZNeP (μM)	Olaparib (μM)
mES #1	5.90	2.08	0.68	7.07
mES #5	5.61	6.79	10.30	2.36
mES #33	0.76	1.96	10.95	17.50
hES_EWS	1.71	2.98	13.46	0.86
hES_KH	7.41	6.05	15.87	2.70
hCCS_KAS	2.16	3.16	16.58	28.85
hOS_U2OS	14.79	3.42	19.33	40.42

Probes), anti-mouse CD99 (a gift of Dietmar Vestweber, Max Planck Institute for Molecular Biomedicine, Muenster, Germany), anti-COL2A (Millipore), anti-S100 (Dako), anti-COL10 (SLS), anti-CD57 (Sigma-Aldrich), anti-NGFR (Millipore), anti–β-catenin (Becton Dickinson), anti-nestin (Chemicon), and anti-myosin (Nichirei). Immunofluorescent images were photographed with a Zeiss LSM 710 laser scanning microscope with a ×40 objective (Zeiss) and LSM Software ZEN 2009 (Zeiss).

Western blotting. Western blot analysis was performed using lysates of whole tumor tissues as described previously (54).

RT-PCR and real-time quantitative RT-PCR. Total RNA extraction, reverse transcription, and RNA quantification were performed according to methods described previously (54). Conventional RT-PCR and real-time quantitative RT-PCR were performed by using a Gene Amp 9700 thermal cycler (Applied Biosystems) and a 7500 Fast Real-Time PCR System (Applied Biosystems), respectively. The sequences of the oligonucleotide primers are shown in Supplemental Excel File 6.

Luciferase assay. A 1,340-bp genomic DNA fragment upstream from the murine *Gdf5* exon 1 was amplified by PCR using the following primers: forward (5′-TTCTATAATCCTACTCTGTAG-3′) and reverse (5′-CTGAAAATAACTCGTTCTTG-3′). The fragment was inserted into the pGL4.10 vector (Promega) and transfected into eSZ, eGP, eSyR, or trunk cells using Lipofectamine 2000 (Invitrogen). Luciferase assays were performed as described previously (54).

In vitro differentiation assay. Cells were plated at 2×10^5 cells per well in 6-well plates and cultured in growth medium. Adipogenic, chondrogenic, osteogenic, myogenic, and neurogenic differentiation assays were conducted according to the methods previously described (55–57).

Microarray analysis. GeneChip analysis was conducted to determine gene expression profiles. A per cell normalization method was applied to eSZ and eGP samples (58). Briefly, cellular lysates were prepared with RLT buffer (QIAGEN). After RNA cocktails were added to the cell lysates according to the amount of DNA, total RNA was extracted using the RNeasy Mini Kit (QIAGEN). The murine Genome 430 2.0 Array (Affymetrix) was hybridized with aRNA probes generated from eSZ and eGP cells and murine Ewing's sarcoma tissue. After staining with streptavidin-phycoerythrin conjugates, arrays were scanned using an Affymetrix GeneChip Scanner 3000 and analyzed using Affymetrix GeneChip Command Console Software (Affymetrix) and GeneSpring GX 11.0.2 (Agilent Technologies) as described previously (59). The expression data for eSZ and eGP cells were converted to mRNA copy numbers per cell by the Percellome method, quality controlled, and analyzed using Percellome software (58). GSEA was performed using GSEA-P 2.0 software (60).

Data comparisons and clustering between murine and human microarray data sets. The microarray data from 10 murine Ewing's sarcoma samples were compared with human microarray data sets. Data from the ONCOMINE database (https://www.oncomine.org/) were accessed in June 2011. Five microarray studies containing 117 tumor samples that were analyzed using Human Genome U133A Array (Affymetrix) were queried for gene expression. CEL files from E-MEXP-353 (61), E-MEXP-1142 (62), GSE6481 (63), GSE7529 (64), GSE21122 (65), GSE6461 (66), GSE42548 (67), GSE23972 (68), GSE20196 (69), and GSE10172 (70) were downloaded. The probe sets of the human U133A array were translated into 23,860 murine 430 2.0 arrays by the translation function of GeneSpring using Entrez Gene ID to make a novel common platform. Hierarchical clustering was achieved using log-transformed data and the following procedure. For the initial statistical analysis, 13,026 genes that showed a "present" or "marginal" call in at least 24 of a total of 32 human Ewing's sarcoma samples were selected. Then, 12,340 probes were selected by 1-way ANOVA ($P < 0.05$) analysis. Finally, 1,819 probes that showed >2-fold differences of expression in at least 3 tumor types were selected. With these 1,819 probes, hierarchical clustering was performed using the average linkage method and the Pearson's centered measurements. In addition, a probe set consisting of the 2,000 sequences that were the most altered in expression in human and mouse round cell tumors (Ewing's sarcoma, neuroblastoma, poorly differentiated synovial sarcoma, and malignant lymphoma) was used to distinguish each tumor from the other 3 using a fold-change analysis. Then, the frequencies of these 2,000 probes were compared between mouse Ewing's sarcoma and 4 human tumor types and between human Ewing's sarcoma and 4 mouse tumor types to find the closest tumor type using similar entities from GeneSpring.

ChIP. A total of 5×10^6 cells per immunoprecipitation were cross-linked with 10% formaldehyde for 10 minutes at room temperature. Histone immunoprecipitation was performed with anti-histone antibodies targeted against H3K9/K14Ac, H3K4/me3, H3K27/me3, total H3 (Cell Signaling Technologies), or H3K9/me3 (Millipore) preconjugated to protein G magnetic beads. Immunoprecipitated DNA was amplified with primers specific for each region. Sequences are shown in Supplemental Excel File 6.

Cre/loxP-mediated gene silencing. eSZ cells were transduced with a floxed *EWS-FLI1* retrovirus, and Ewing's sarcoma cells were obtained from a subcutaneous tumor developed in a nude mouse. Tumor cells were transduced with pMSCV-Cre-puro retrovirus in vitro. Senescence-associated β-galactosidase expression was detected using a Senescence Detection Kit (Biovision) 4 days after transduction of the retrovirus.

siRNA interference studies. For knockdown of *FLI1*, *Dkk2*, *Catnb*, *Prkcb1*, *Ezh2*, *Igf1*, *Dkk1*, and *Erg*, siRNAs were purchased from QIAGEN. The list of siRNAs is shown in Supplemental Excel File 7. siRNAs were introduced into mouse Ewing's sarcoma cells according to the manufacturer's protocol. Knockdown efficiencies were confirmed by Western blotting using anti-FLAG (Sigma-Aldrich), anti-ERG and anti-PKC β1 (Santa Cruz Biotechnology), anti–β-catenin (Becton Dickinson), and anti-EZH2 (Cell Signaling Technologies) or RT-PCR.

Pharmacological experiments with specific inhibitors. Mouse Ewing's sarcoma cells were treated with MEK1 inhibitor U0126 (Cell Signaling Technologies) in vitro. Both mouse and human Ewing's sarcoma cell lines were treated with WNT/β-catenin inhibitors, iCRT14 and PNU74654 (Tocris Bioscience); an EZH2 inhibitor, DZNeP (Cayman Chemical); or a PARP1 inhibitor, olaparib (Selleckchem), both in vitro and in vivo. Inhibition of ERG phosphorylation was examined by Western blotting using anti-P-ERK1/2 and anti-ERK1/2 (Cell Signaling

Technologies). For in vivo experiments, 1×10^6 tumor cells were transplanted subcutaneously into nude mice, and the mice were treated with specific inhibitors when the tumor diameter reached 5 mm. All the inhibitors were dissolved in 0.2% DMSO, and they were administered by intraperitoneal injection 3 times per week.

Cell cycle assay. Single-cell suspensions were permeabilized with 0.1% Triton X-100 in PBS, and 50 mg/ml propidium iodide and 1 mg/ml RNAse A were added. The cell suspensions were then analyzed by using a FACSCalibur flow cytometer and ModFit software (Becton Dickinson).

Cloning retroviral integration sites. Retroviral integration sites of individual mouse Ewing's sarcoma were isolated by inverse PCR, sequenced, and mapped as described previously (71).

Accession numbers. The microarray data sets are accessible through the NCBI Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo), with accession numbers GSE32615 and GSE32618.

Statistics. Continuous distributions were compared with 2-tailed Student's *t* test. Survival analysis was performed using the Kaplan-Meier life table method, and survival between groups was compared with the log-rank test. The 2-proportion z test was used to evaluate the significance of differences in the matched probe sets between 2 tumor types. All *P* values were 2 sided, and a *P* value of less than 0.05 was considered significant.

Study approval. Animals were handled in accordance with the guidelines of the animal care committee at the Japanese Foundation for Cancer Research, which gave ethical approval for these studies.

Acknowledgments

We thank W. Kurihara and T. Koui for helping with the microarray analyses; S.J. Baker (St Jude Children's Research Hospital, Memphis, Tennessee, USA) for EWS-FLI1 cDNA; and D. Vestweber (Max Planck Institute for Molecular Biomedicine, Muenster, Germany) for anti-mouse CD99. The research was funded by Grant-in-Aid for Scientific Research on Priority Areas Cancer 17013086 from Ministry of Education, Culture, Sports, Science and Technology and Grant-in-Aid for Young Scientists 23791672 from Japan Society for the Promotion of Science.

Received for publication July 29, 2013, and accepted in revised form April 10, 2014.

Address correspondence to: Takuro Nakamura, 3-8-31 Ariake, Koto-ku, Tokyo 135-8550, Japan. Phone: 81.3.3570.0462; Fax: 81.3.3570.0463; E-mail: takuro-ind@umin.net.

1. Bernstein M, et al. Ewing's sarcoma family of tumors: current management. *Oncologist.* 2006;11(5):503–519.
2. Ewing J. Diffuse endothelioma of bone. *Proc New York Pathol Soc.* 1921;21:17–24.
3. Hu-Lieskovan S, Zhang J, Wu L, Shimada H, Schofield DE, Triche TJ. EWS-FLI1 fusion protein up-regulates critical genes in neural crest development and is responsible for the observed phenotype of Ewing's family of tumors. *Cancer Res.* 2005;65(11):4633–4644.
4. Tirode F, Laud-Duval K, Prieur A, Delorme B, Cahrbord P, Delattre O. Mesenchymal stem cell features of Ewing tumors. *Cancer Cell.* 2007;11(5):421–429.
5. Delattre O, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature.* 1992;359(6391):162–165.
6. Zucman J, et al. Combinatorial generation of variable fusion proteins in the Ewing family of tumours. *EMBO J.* 1993;12(12):4481–4487.
7. Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ, Denny CT. A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat Genet.* 1994;6(2):146–151.
8. Ordonez JL, Osuna D, Herrero D, de Alava E, Madoz-Gurpide J. Advances in Ewing's sarcoma research: where are we now and what lies ahead? *Cancer Res.* 2009;69(18):7140–7150.
9. Riggi N, et al. Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal progenitor cells. *Cancer Res.* 2005;65(24):11459–11468.
10. Castillero-Trejo Y, Eliazer S, Xiang L, Richardson JA, Ilaria RL Jr. Expression of the EWS/FLI-1 oncogene in murine primary bone-derived cells results in EWS/FLI-1-dependent, Ewing sarcoma-like tumors. *Cancer Res.* 2005;65(19):8698–8705.
11. Arndt CAS, Crist WM. Common musculoskeletal tumors of childhood and adolescence. *N Engl J Med.* 1999;341(5):342–352.
12. Iwamoto M, et al. 2007. Transcription factor ERG and joint and articular cartilage formation during mouse limb and spine skeletogenesis. *Dev Biol.* 2007;305(1):40–51.
13. Koyama E, et al. A distinct cohort of progenitor cells participates in synovial joint and articular cartilage formation during mouse limb skeletogenesis. *Dev Biol.* 2008;316(1):62–73.
14. Mundy C, et al. Synovial joint formation requires local Ext1 expression and heparin sulfate production in developing mouse embryo limbs and spine. *Dev Biol.* 2011;351(1):70–81.
15. Ambros IM, Ambros PF, Strehl S, Kovar H, Gadner H, Salzer-Kuntschik M. MIC2 is a specific marker for Ewing's sarcoma and peripheral primitive neuroectodermal tumors. Evidence for a common histogenesis of Ewing's sarcoma and peripheral primitive neuroectodermal tumors from MIC2 expression and specific chromosome aberration. *Cancer.* 1991;67(7):1886–1893.
16. Bixel G, Kloep S, Butz S, Petri B, Engelhardt B, Vestweber D. Mouse CD99 participates in T-cell recruitment into inflamed skin. *Blood.* 2004;104(10):3205–3213.
17. Vortkamp A, Lee K, Lanske B, Segre GV, Kronenberg HM, Tabin CJ. Regulation of rate of cartilage differentiation by Indian Hedgehog and PTH-related protein. *Science.* 1996;273(5275):613–622.
18. Sohaskey ML, Yu J, Diaz MA, Plaas AH, Harland RM. JAWS coordinates chondrogenesis and synovial joint positioning. *Development.* 2008;135(13):2215–2220.
19. Vijayaraj P, et al. Erg is a crucial regulator of endocardial-mesenchymal transformation during cardiac valve morphogenesis. *Development.* 2012; 139(21):3973–3985.
20. Storm EE, Kingsley DM. GDF5 coordinates bone and joint formation during digit development. *Dev Biol.* 1999;209(1):11–27.
21. St-Jacques B, Hammerschmidt M, McMahon AP. Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation. *Genes Dev.* 1999;13(16):2072–2086.
22. Mak KK, Kronenberg HM, Chuang PT, Mackem S, Yang Y. Indian hedgehog signals independently of PTHrP to promote chondrocyte hypertrophy. *Development.* 2008;135(11):1947–1956.
23. Toomey EC, Schiffman JD, Lessnick SL. Recent advances in the molecular pathogenesis of Ewing's sarcoma. *Oncogene.* 2010;29(32):4504–4516.
24. James CG, Stanton LA, Agoston H, Ulici V, Underhill TM, Beier F. Genome-wide analyses of gene expression during mouse endochondral ossification. *PLoS One.* 2010;5(1):e8693.
25. Honsei N, Ikuta T, Kawana K, Kaneko Y, Kawajiri K. Participation of nuclear localization signal 2 in the 3′-ETS domain of FLI1 in nuclear translocation of various chimeric EWS-FLI1 oncoproteins in Ewing tumor. *Int J Oncol.* 2006;29(3):689–693.
26. Miyagawa Y, et al. EWS/ETS regulates the expression of the Dickkopf family in Ewing family tumor cells. *PLoS One.* 2009;4(2):e4634.
27. Surdez D, et al. Targeting the EWSR1-FLI1 oncogene-induced protein kinase PKC-β abolishes Ewing sarcoma growth. *Cancer Res.* 2012;72(17):4494–4503.
28. Richter GH, et al. EZH2 is a mediator of EWS/FLI1 driven tumor growth and metastasis blocking endothelial and neuro-ectodermal differentiation. *Proc Natl Acad Sci U S A.* 2009;106(13):5324–5329.
29. Nishimori H, et al. The Id2 gene is a novel target of transcriptional activation by EWS-ETS fusion proteins in Ewing family tumors. *Oncogene.* 2002;21(54):8302–8309.
30. Smith R, et al. Expression profiling of EWS/FLI1 identifies NKX2.2 as a critical target gene in Ewing's sarcoma. *Cancer Cell.* 2006;9(5):405–416.
31. Kinsey M, Smith R, Lessnick SL. NR0B1 is required for the oncogenic phenotype mediated by EWS/FLI in Ewing's sarcoma. *Mol Cancer Res.* 2006;4(11):851–859.
32. Abaan OD, Levenson A, Khan O, Furth PA, Uren A, Toretsky JA. PTPL1 is a direct transcriptional target of EWS-FLI1 and modulates Ewing's sarcoma tumorigenesis. *Oncogene.* 2005;24(16):2715–2722.
33. Wakahara K, et al. EWS-Fli1 up-regulates expression of the Aurora A and Aurora B kinases. *Mol Cancer Res.* 2008;6(12):1937–1945.
34. Luo W, Gangwal K, Sankar S, Boucher KM, Thomas D, Lessnick SL. GSTM4 is a microsatellite-containing EWS/FLI target involved in Ewing's sarcoma oncogenesis and therapeutic resistance. *Oncogene.* 2009;28(46):4126–4132.
35. Fuchs B, Inwards C, Scully SP, Janknecht R. hTERT is highly expressed in Ewing's sarcoma and activated by EWS-ETS oncoproteins. *Clin Orthop Relat Res.* 2004;426(426):64–68.
36. Watanabe G, et al. Induction of tenascin-C by tumor-specific EWS-ETS fusion genes. *Genes Chromosomes Cancer.* 2003;36(3):224–232.
37. Deneen B, Hamidi H, Denny CT. Functional analysis of the EWS/ETS target gene uridine phosphorylase. *Cancer Res.* 2003;63(14):4268–4274.
38. Hahm KB, et al. Repression of the gene encoding the TGF-beta type II receptor is a major target of the EWS-FLI1 oncoprotein. *Nat Genet.* 1999; 23(2):222–227.
39. Dohjima T, Ohno T, Banno Y, Nozawa Y, Wen-yi Y, Shimizu K. Preferential down-regulation of phospholipase C-beta in Ewing's sarcoma cells transfected with antisense EWS-Fli-1. *Br J Cancer.* 2000;82(1):16–19.
40. Scotlandi K, et al. Insulin-like growth factor I receptor-mediated circuit in Ewing's sarcoma/peripheral

neuroectodermal tumor: a possible therapeutic target. *Cancer Res.* 1998;56(20):4570–4574.
41. Balamuth NJ, Womer RB. Ewing's sarcoma. *Lancet Oncol.* 2010;11(2):184–192.
42. Niehrs C. Function and biological roles of the Dickkopf family of Wnt modulators. *Oncogene.* 2006;25(57):7469–7481.
43. Mikheev AM, Mikheeva SA, Liu B, Cohen P, Zarbl H. A functional genomics approach for the identification of putative tumor suppressor genes: Dickkopf-1 as suppressor of HeLa cell transformation. *Carcinogenesis.* 2004;25(1):47–59.
44. Uren A, Wolf V, Sun YF, Azari A, Rubin JS, Toretsky JA. Wnt/Frizzled signaling in Ewing sarcoma. *Pediatr Blood Cancer.* 2004;43(3):243–249.
45. Garnett MJ, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature.* 2012;483(7391):570–575.
46. Deneen B, Denny CT. Loss of p16 pathways stabilizes EWS/FLI1 expression and complements EWS/FLI1 mediated transformation. *Oncogene.* 2001;20(46):6731–6741.
47. Sohn EJ, Li H, Reldy K, Beers LF, Christensen BL, Lee SB. EWS/FLI1 oncogene activates caspase 3 transcription and triggers apoptosis in vivo. *Cancer Res.* 2010;70(3):1154–1163.
48. Riggi N, et al. EWS-FLI-1 modulates miRNA145 and SOX2 expression to initiate mesenchymal stem cell reprogramming toward Ewing sarcoma cancer stem cells. *Genes Dev.* 2010;24(9):916–932.
49. Grier HE. The Ewing family of tumors: Ewing's sarcoma and primitive neuroectodermal tumors. *Pediatr Clin North Am.* 1997;44(4):991–1004.
50. Soliman H, Ferrari A, Thomas D. Sarcoma in the young adult population: An international view. *Sem Oncol.* 2009;36(3):227–236.
51. Martins AS, et al. A pivotal role for heat shock protein 90 in Ewing sarcoma resistance to anti-insulin-like growth factor 1 receptor treatment: in vitro and in vivo study. *Cancer Res.* 2008;68(15):6260–6270.
52. Glade Bender JL, et al. Phase I pharmacokinetic and pharmacodynamics study of pazopanib in children with soft tissue sarcoma and other refractory solid tumors: a children's oncology group phase I consortium report. *J Clin Oncol.* 2013;31(24):3034–3043.
53. Jin G, et al. Trib1 and Evi1 cooperate with Hoxa and Meis1 in myeloid leukemogenesis. *Blood.* 2007;109(9):3998–4005.
54. Kawamura-Saito M, et al. Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. *Hum Mol Genet.* 2006;15(13):2125–2137.
55. Pittenger MF, et al. Multilineage potential of adult human mesenchymal stem cells. *Science.* 1999;284(5411):143–147.
56. Sakaguchi Y, Sekiya I, Yagishita K, Muneta T. Comparison of human stem cells derived from various mesenchymal tissues: superiority of synovium as a cell source. *Arthritis Rheum.* 2005;52(8):2521–2529.
57. Shiota M, et al. Isolation and characterization of bone marrow-derived mesenchymal progenitor cells with myogenic and neuronal properties. *Exp Cell Res.* 2007;313(5):1008–1023.
58. Kanno J, et al. "Per cell" normalization method for mRNA measurement by quantitative PCR and microarrays. *BMC Genomics.* 2006;7:64.
59. Fujino T, et al. Function of EWS-POU5F1 in sarcomagenesis and tumor cell maintenance. *Am J Pathol.* 2010;176(4):1973–1982.
60. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for gene set enrichment analysis. *Bioinformatics.* 2007; 23(23):3251–3253.
61. Henderson SR, et al. A molecular map of mesenchymal tumors. *Genome Biol.* 2005;6(9):R76.
62. Schaefer KL, et al. Microarray analysis of Ewing's sarcoma family of tumours reveals characteristic gene expression signatures associated with metastasis and resistance to chemotherapy. *Eur J Cancer.* 2008;44(5):699–709.
63. Nakayama R, et al. Gene expression analysis of soft tissue sarcomas: characterization and reclassification of malignant fibrous histiocytoma. *Mod Pathol.* 2007;20(7):749–759.
64. Albino D, et al. Identification of low intratumoral gene expression heterogeneity in neuroblastic tumors by genome-wide expression analysis and game theory. *Cancer.* 2008;113(6):1412–1422.
65. Barretina J, et al. Subtype-specific genomic alterations define new targets for soft-tissue sarcoma therapy. *Nat Genet.* 2010;42(8):715–721.
66. Haldar M, et al. A conditional mouse model of synovial sarcoma: insights into a myogenic origin. *Cancer Cell.* 2007;11(4):375–388.
67. Teitz T, et al. Th-MYCN mice with caspase-8 deficiency develop advanced neuroblastoma with bone marrow metastasis. *Cancer Res.* 2013;73(13):4086–4097.
68. Jeannet R, et al. Oncogenic activation of the Notch1 gene by deletion of its promoter in Ikaros-deficient T-ALL. *Blood.* 2010;115(25):5443–5454.
69. Nakayama R, et al. Gene expression profiling of synovial sarcoma: distinct signature of poorly differentiated type. *Am J Surg Pathol.* 2010;34(11):1599–1607.
70. Masque-Soler N, Szczepanowski M, Kohler CW, Spang R, Klapper W. Molecular classification of mature aggressive B-cell lymphoma using digital multiplexed gene expression on formalin-fixed paraffin-embedded biopsy specimens. *Blood.* 2013; 122(11):1985–1986.
71. Tanaka M, Jin G, Yamazaki Y, Takahara T, Takuwa M, Nakamura T. Identification of candidate cooperative genes of the Apc mutation in transformation of the colon epithelial cell by retroviral insertional mutagenesis. *Cancer Sci.* 2008;99(5):979–985.

Integrative knowledge management to enhance pharmaceutical R&D

Maria Marti-Solano, Ewan Birney, Antoine Bril, Oscar Della Pasqua, Hiroaki Kitano, Barend Mons, Ioannis Xenarios and Ferran Sanz

Information technologies already have a key role in pharmaceutical research and development (R&D), but achieving substantial advances in their use and effectiveness will depend on overcoming current challenges in sharing, integrating and jointly analysing the range of data generated at different stages of the R&D process.

Maria Marti-Solano and Ferran Sanz are at the IMIM and Universitat Pompeu Fabra, Dr. Aiguader 88, 08003 Barcelona, Spain.
Ewan Birney is at the European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, Cambridgeshire CB10 1SD, UK.
Antoine Bril is at the Institut de Recherches Internationales Servier, 53 rue Carnot, 92284 Suresnes Cedex, France.
Oscar Della Pasqua is at GlaxoSmithKline, Stockley Park West, Uxbridge, Middlesex UB11 1BT, UK.
Hiroaki Kitano is at the Systems Biology Institute, 5-6-9 Shirokanedai, Minato, Tokyo 108–0071, Japan; and the RIKEN Center for Integrative Medical Sciences, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama City, Kanagawa 230–0045, Japan.
Barend Mons is at the Netherlands Bioinformatics Centre, 260 NBIC, 6500 HB Nijmegen, The Netherlands; and the Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands.
Ioannis Xenarios is at the Swiss Institute of Bioinformatics, Quartier Sorge - Batiment Genopode, 1015 Lausanne, Switzerland.
Correspondence to F.S.
e-mail: ferran.sanz@upf.edu
doi:10.1038/nrd4290

The explosion in the accumulation of biomedical data resulting from technological advances such as next-generation sequencing, coupled with progress in information technologies, offers new opportunities to use such information for pharmaceutical R&D. In parallel, the increasingly collaborative nature of many R&D projects has highlighted the importance of effective information sharing between project partners and, from a wider perspective, with the overall biomedical community.

Initiatives such as the European Innovative Medicines Initiative (IMI) are supporting the development of strategies aimed at improving translational knowledge management in biomedical sciences, as well as overcoming barriers in information sharing. In this article, we highlight the key points emerging from a debate on 'Translational Knowledge Management in Pharmaceutical R&D', held in Brussels in July 2013, that involved representatives from a range of collaborative projects in the field, along with other experts and stakeholders (see Supplementary information S1 (box) for details). Challenges that were discussed include the complexity and heterogeneity of biomedical data, the need to establish relevant, widely accepted and openly available data standards and the lack of integration of knowledge from different disciplines and stages of the R&D process.

Key issues and challenges

Data evaluation and integration. Given the complexity of most biological questions, combining data from multiple levels (molecular, cellular, tissue and others), disciplines (molecular and systems biology, medicinal chemistry, preclinical and clinical pharmacology, and others) and sources may be needed to develop information resources and computational models that are useful for addressing pharmaceutical R&D questions effectively. Indeed, the issue of managing data arising from different sources is becoming increasingly important as changes in the R&D ecosystem mean that information is not necessarily generated in-house by large pharmaceutical companies but is derived from external organizations such as academic institutions, biotechnology companies or contract research organizations.

The heterogeneity of biomedical data is a major challenge. In many cases, a key determinant of data usefulness and reusability is the availability of additional information to evaluate the relevance and quality of a particular data set. Being able to track the data source and to retrieve information on the context in which the data were generated (for example, experimental conditions and the model organisms used) is crucial to assess whether the data are suitable to address a particular research question and, moreover, whether data from different sources can be meaningfully combined. In addition, to perform large-scale statistical analyses and generate useful models from biomedical data, it is necessary to have not only information on the positive results but also on the negative or discarded ones.

Ontologies and standards. Developing, disseminating and promoting the wide-scale adoption of appropriate biomedical ontologies and data standards is also crucial in allowing data integration. The definition and usage of standards for the characterization of experimental assays, for the collection of clinical information, for model description and, in general, for the representation of different types of metadata, will facilitate true data interoperability — that is, a meaningful and accurate exchange, integration and joint exploitation of biomedical data[1]. In order to promote the broad adoption of such standards, it is important that they are open and reasonably flexible, and that key stakeholders and user communities are engaged with their development. In some areas, these aims may be best served by a bottom-up approach, whereas others will require a coordinated top-down initiative.

An example of a high-impact top-down initiative is the recent commitment by the European Medicines Agency (EMA), the Japanese Pharmaceutical & Medical Devices Agency (PMDA) and the US Food and Drug Administration (FDA) to require data standards developed by the Clinical Data Interchange Standards Consortium (CDISC).

Data and information sharing. Another of the major challenges for the effective integration of data relevant to pharmaceutical R&D is that data, associated information and knowledge often remain siloed[2]. Moreover, there are communication barriers between researchers from different disciplines. So, not only the data but also relevant details about it need to be shared. This requires assessment of the relevance of data, taking into account the underlying clinical or biomedical research question and adapting the information to suit its use across different areas of expertise. This could be achieved by identifying the crucial information and tailoring the key messages that are required for decision-making at different stages of drug discovery and development, while promoting the traceability of data. This is helped by training and involving specific professionals (knowledge engineers and data scientists) who are capable of bridging information silos and facilitating communication and collaboration[3].

Active involvement of the community is the basis for successful integrative knowledge management in R&D. This may require the creation of regulations to generate confidence about data sharing and the promotion of information formats that facilitate it. It is possible to harness the know-how of the research community by throwing down scientific challenges, involving the community in knowledge extraction, annotation and filtering[4]. Furthermore, the availability of novel communication channels (such as social networks) and the increasing availability of individual genetic information owing to the decrease in sequencing costs will affect current research models and could make patient participation in the R&D process more active.

Data sustainability. Ensuring sustainability of data repositories in the long term — including their maintenance as well as their updating and evolution in a changing environment — should be a key component of collaborative projects. Making data accessible to a wider community and enabling their reuse beyond the project that generates them will be crucial to optimize the use of resources by minimizing replication. We therefore believe that data sharing should become a requirement when public funding is involved. However, to empower the scientific community to effectively reuse data, it has to be enriched with relevant metadata, and published data should be converted into appropriate formats for integrative knowledge management[5]. Similarly, computational models that are developed to enable data representation must be maintained and updated, as well as systematically characterized by appropriate metadata that allow further appraisal of their predictive performance and possible biases.

Recommendations

To maximize the effectiveness of efforts to integrate data in pharmaceutical R&D, it is vital to develop strategies and processes to ensure that:

- Protocol design and data collection focus on questions that are relevant for decision-making at the different stages of pharmaceutical R&D
- Relevant standards for the characterization of data, methods and models are identified (or developed if unavailable) and their use is promoted and facilitated by dissemination and training
- Data and models are annotated with enough detail regarding their provenance and quality to allow a critical assessment of their suitability for reuse
- Data sharing is understood as a responsibility (especially when data are derived from public funding), and the community participates in the promotion and recognition of data sharing, as well as in knowledge extraction and management
- Sustainability of data collections is considered a key component in the life cycle of collaborative projects

The dissemination and adoption of these principles among stakeholders — and in particular across the scientific community, clinical researchers and policy makers — is essential to establish knowledge management strategies in pharmaceutical R&D that efficiently exploit the increasing availability of novel biomedical data and learn from previous experience, thus enabling a more efficient search for innovative, effective and safe medicines.

1. Sansone, S. A. *et al.* Toward interoperable bioscience data. *Nature Genet.* **44**, 121–126 (2012).
2. Cases, M. *et al.* Improving data and knowledge management to better integrate health care and research. *J. Intern. Med.* **274**, 321–328 (2013).
3. Beck, T. *et al.* Knowledge engineering for health: a new discipline required to bridge the "ICT gap" between research and healthcare. *Hum. Mutat.* **33**, 797–802 (2012).
4. Kitano, H. *et al.* Social engineering for virtual 'big science' in systems biology. *Nature Chem. Biol.* **7**, 323–326 (2011).
5. Williams, A. J. *et al.* Open PHACTS: semantic interoperability for drug discovery. *Drug Discov. Today* **17**, 1188–1198 (2012).

Acknowledgements

The debate on "Translational Knowledge Management in Pharmaceutical R&D" was held in Brussels on 11–12 July 2013, within the framework of the European INBIOMEDvision CSA (funded by the EU FP7 under grant agreement no. 270107) and with the collaboration of Innovative Medicines Initiative (IMI). In addition to the authors, A. Martin, R. Bellazzi, N. Blomberg, A. Brookes, R. Campbell, B. de Bono, I. Dix, G. Ecker, Y. Guo, B. Hardy, M. Hofmann-Apitius, J. Houwing-Duistermaat, L. D. Hudson, B. Kisler, J. Kleinjans, V. Maojo, J. Mestres, J.L. Oliveira, M. Pastor, T. Rijnders, S. Scollen, T. Steger-Hartmann, P. Ruch, T. Slater, G. O. Strawn, M. Sundgren, J. van der Lei and M. Goldman were speakers at the debates and are acknowledged for their valuable input and support for the ideas in this paper, which represent the consensus view of all the participants (see Supplementary information S1 (box) for their affiliations).

Competing interests

The authors declare no competing interests.

SUPPLEMENTARY INFORMATION

See online article: S1 (box)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

Advanced Biomedical Engineering
3: 50-58, 2014.

Original Paper

A Versatile Platform for Multilevel Modeling of Physiological Systems: SBML-PHML Hybrid Modeling and Simulation

Yoshiyuki ASAI,[*, #] Takeshi ABE,[*] Hideki OKA,[**] Masao OKITA,[***] Ken-ichi HAGIHARA,[***] Samik GHOSH,[†††] Yukiko MATSUOKA,[†††] Yoshihisa KURACHI,[†] Taishin NOMURA,[††] Hiroaki KITANO[*, †††]

Abstract Specialized languages used for describing computational models in the field of systems biology and physiology, such as Systems Biology Markup Language (SBML), CellML, and Physiological Hierarchy Markup Language (PHML), have been devised to enhance effective model reuse and sharing among researchers for developing large, multilevel models. Each language has its own specialty. By combining two of these languages, i. e. SBML for illustrating subcellular phenomena and PHML for expressing supracellular dynamics, a novel technology has been developed to describe models of multilevel biophysiological systems. For practical use of the aforementioned languages, consolidated software applications providing intuitive graphical user interfaces are necessary. Starting from 2011, a versatile platform called PhysioDesigner has been developed for multilevel modeling of physiological systems based on PHML. In this article, we focus on the newly developed distinguishing features of PhysioDesigner and PHML for the development of multilevel biophysiological models using SBML-PHML hybridization.

Keywords : multilevel modeling, SBML-PHML hybrid modeling, PhysioDesigner.

Adv Biomed Eng. **3**: pp. 50-58, 2014.

1. Introduction

In order to integrate the massive amounts of data generated by clinical and laboratory studies with simulation results, computable mathematical models are themselves increasing in size and complexity [1]. For this reason, it is essential that pieces of models are able to be shared and re-used, in the same manner of building blocks. To promote effective collaboration for building large-scale models, fundamental tools that support these activities should be consolidated.

To enhance model sharing, several pioneering efforts have been undertaken. For example, XML-based descriptive language formats used to describe the dynamics of biological and physiological systems, such as SBML [2], CellML [3], NeuroML [4], and PHML, have been proposed. The main objective in developing these languages was to establish a common communication foundation for enhancing exchange of models among collaborators.

To use these languages effectively for multilevel modeling of physiological systems, application supports are essential. Many applications have been developed and published and listed on websites, such as software applications for SBML [5], NeuroML [6] and CellML [7]. For example, CellDesigner [8] is a versatile modeling tool of biochemical networks based on SBML. Although many applications employ single languages, some tools support more than one. For example, VCell, an environment for virtual cell modeling and simulation [9], and JSim, a simulation framework that natively uses a modeling language called Mathematical Modeling Language (MML) [10], can import SBML and CellML models. However, in general, these applications convert the imported models into their own native languages. Or, while they support multiple languages, each of the models parsed by such applications must be written in single language, which can be either SBML or others.

PHML, which is a successor to *insilico*ML [11], was developed relatively recently compared with the languages mentioned above. It was developed in parallel with PhysioDesigner [12] (**Fig. 1**), a platform on which users can build mathematical models of multilevel physiological systems with a graphical user interface. Models built with PhysioDesigner are written in PHML, which is effective at explicitly describing the hierarchical structure of physiological systems. The main elements in a PHML model are called modules. Modules form a tree structure to express the hierarchical structure of phys-

This study was presented at the Symposium on Biomedical Engineering 2013, Fukuoka, September, 2013.
Received on July 16, 2013; revised on November 1, 2013 and December 18, 2013; accepted on February 17, 2014.

[*] Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan.
[**] RIKEN Brain Science Institute, Neuroinformatics Japan Center, Wako, Japan.
[***] Graduate School of Information Science and Technology, Osaka University, Suita, Japan.
[†] Graduate School of Medicine, Osaka University, Suita, Japan.
[††] Graduate School of Engineering Science, Osaka University, Toyonaka, Japan.
[†††] The Systems Biology Institute, Tokyo, Japan.
[#] 1919-1 Tancha, Onna-son, Okinawa, Japan.
E-mail: yoshiyuki.asai@oist.jp

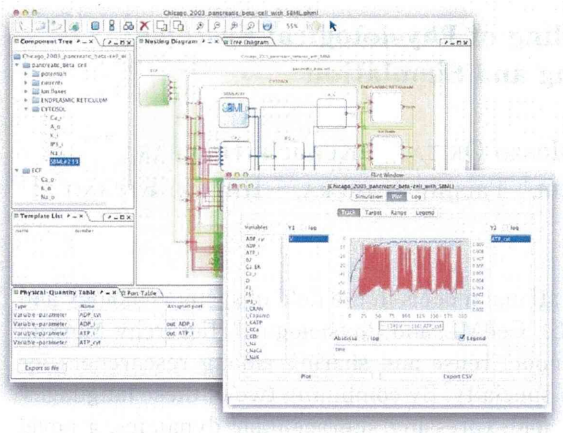

Fig. 1 Snapshots of PhysioDesigner and Flint.

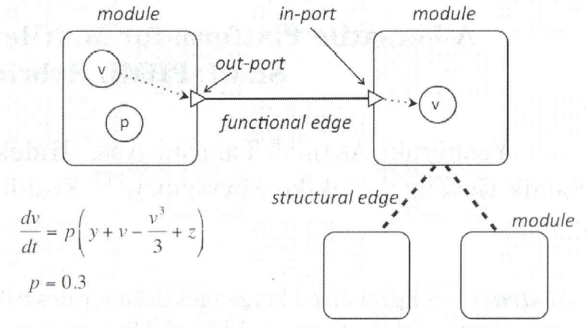

Fig. 2 Scheme of a PHML model. Modules are fundamental elements for constructing a model in PHML. Each module (shown as a rounded rectangle) is quantitatively characterized by physical quantities (shown as encircled letters). Relationships among modules are explicitly represented by edges (solid line: functional edge, dashed line: structural edge). A module can have offspring modules forming a tree-like structure representing the hierarchical structure seen in physiological systems. The value of a physical quantity can be exported from a module through an out-port (small triangle on the right edge of the module) to another module, which enters the module through an in-port (small triangle on the left edge of the module). The equation and parameter at the bottom left corner are examples of the definitions of physical quantities in the module on the left.

iological systems, and they form a network to imitate the functional relationship among physiological components. Each module is quantitatively characterized by the defining physical quantities inside.

PHML itself was originally designed to present hierarchical structures of physiological systems, including even subcellular phenomena. However, because SBML is a language dedicated to describing subcellular phenomena such as signal transduction and protein-protein interactions, it is better to use SBML to describe subcellular phenomena. We can then benefit from existing SBML resources such as the BioModels database[13]. To edit the SBML part, we can also use CellDesigner or other applications that are dedicated to manipulating SBML, which is a considerably better method than re-implementing a function to edit SBML on PhysioDesigner.

For describing computable models of multilevel physiological systems including subcellular and supracellular phenomena, a novel technology that incorporates SBML and PHML has been developed to leverage the advantages from both languages. In the PHML framework, it is possible to integrate an SBML model into a functional network of modules of a PHML model by embedding the SBML model into one of the modules. Then, the module containing SBML represents a model of subcellular phenomena as modeled by the SBML model. This process is not merely simple embedment of SBML models into a PHML model at the model description level, but is also a computable hybridization accompanied by numerical simulation with Flint, a simulator developed concomitantly with PhysioDesigner.

In this article, we introduce a new technique developed on PhysioDesigner for SBML-PHML hybrid modeling and simulation.

2. PhysioDesigner Overview

PhysioDesigner is a versatile platform that supports modeling and simulation of physiological systems with multiple spatiotemporal levels. The current version as of October 2013 is 1.0 beta6, and it is available at http://physiodesigner.org. Development of PhysioDesigner began in 2011, with the inheritance of all features from *insilico*IDE [14-16], and the tool is posted at http://www.physiome.jp[17].

The main components in a PHML model are called modules, which represent biological and physiological elements (**Fig. 2**). Multiple modules can be defined as a single module at one level above. For example, many cells form a tissue. These modules at the lower level represent physiological entities that are more precise in spatial scale and more detailed in logical scale. By this nested representation of modules forming the module tree structure, hierarchical structures of physiological systems are explicitly expressed in a model.

Each module is quantitatively characterized by several physical quantities such as states defining system dynamics, and variable and static parameters. The dynamics such as ordinary/partial differential equations, and functions of physical quantities are defined by mathematical equations using physical quantities. To define physical quantities in a module, it is often necessary to refer to values of physical quantities defined in other modules. The value of a physical quantity can be exported from a module through an out-port. Then, the numerical information is carried from the out-port to an in-port of a destination module which is pointed by a functional edge linking them. The value arriving at the in-port can be used to define a value of a physical quantity in the destination module. In this sense, modules form a large functional network in a model.

The concept of capsulation, or making a package of a

physiological function, was introduced to PHML to enhance sharing and reuse of models or their components. Capsulation is an operation that involves the encapsulation of an arbitrary number of modules acting together as a certain physiological function by a capsule module. All edge connections to (from) encapsulated modules from (to) the outside of the capsule must pass through the capsule module once in order to secure the independence of the encapsulated modules. Namely, the capsule module acts as an interface or gateway for all modules in the capsule. With this isolation of modules, it becomes easier to reuse the encapsulated modules in another part of the model or in other models.

Simulations of PHML models are conducted by the simulator Flint[12, 16], which is being developed concurrently with PhysioDesigner. Flint was rebranded from *insilico* Sim [18, 19], and is also available at http://physiodesigner.org. One of the features of Flint is that it can execute simulation of SBML models as well as PHML models, using the SBML ODE Solver Library (SOSlib) to parse SBML.

3. SBML and PHML Hybrid Modeling

3.1 Concept and Basic Usage

Because PHML is designed to represent the hierarchical structure of physiological phenomena, it is possible to describe a model that includes integration of subcellular and supracellular phenomena. However, instead of modeling subcellular phenomena with PHML, a novel modeling method of hybridizing SBML and PHML has been developed.

SBML-PHML hybrid modeling is achieved by embedding a whole SBML model into a PHML module (**Fig. 3**). The module then represents the biological system expressed by SBML. The SBML model is integrated into a tree structure expressing the hierarchical structure of multilevel biophysiological systems, and also in a functional network composed of modules.

A limitation is that one module can include only one SBML model. However, obviously one PHML model can have multiple modules that contain an SBML model each. Hence, this function can be utilized not only in subcellular-supracelluar multilevel modeling, but also to create a PHML-based network of multiple SBML models expressing phenomena in a single cell.

3.2 Value Exchange between SBML and PHML

To functionally integrate an SBML model into a PHML network, numerical information must be exchanged between them during a simulation. The main players in an SBML model carrying numerical information (or representing biological entities) are called species, which are used mainly to express, for example, the concentration of ions and molecules that take part in one or more reactions. To define a reaction, parameters such as velocity constant as well as species are used. In a module containing an SBML model, associations between physical quantities and species or parameters have to be defined to form a bridge between the SBML moiety and the PHML moiety.

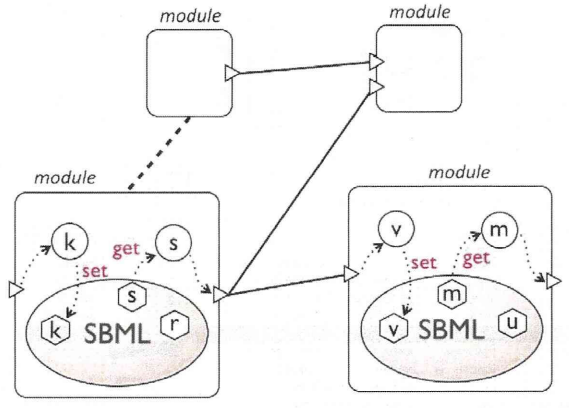

Fig. 3 Schematic diagram of an SBML-PHML hybrid model. A module of PHML may contain a whole SBML model represented by an oval in a module. By making associations between physical quantities in the module (letters in circles) and species/parameters in the SBML model (letters in hexagons), the SBML model is functionally integrated into the network of modules. There are two ways to accomplish this. One is a "get" action, which is used to retrieve the values from the SBML portion and carry them to the PHML portion. The other is a "set" action used to override the values or dynamics originally defined in the SBML portion by those defined in the PHML portion. Dotted arrows indicate the flow of values.

These are two-way actions. One is the "get" action, which converts the value defined in a species or parameter in the SBML model to a physical quantity. Then the other physical quantities in the module can utilize the numerical information defined in the SBML model via the physical quantities of the "get" action. This is similar to associating a physical quantity with an in-port to receive a value carried to it.

The other action is similar, but in the opposite direction, i.e. "set" action. A physical quantity originating in the PHML part with a "set" definition can affect the SBML part by overriding the original definition of species or parameters in the SBML model, without direct modification of the SBML model itself. Even if the species have dynamics originally defined in the SBML model, it is completely overridden by that defined in the physical quantity, replacing the value at every step during the course of a simulation. By this interpretation rule of the definition of the bridge between species/parameter and physical quantity, the SBML model can be effectively involved in the model.

The process of embedding an SBML model in a module is assisted by a PhysioDesigner interface shown in **Fig. 4**. PhysioDesigner employs a dialog box in which a list of the species involved in the SBML model is shown, so that users can define the interaction between SBML species and PHML physical quantities. The direction of the action ("get" or "set") and the associated physical quantity can be selected by combo-boxes.

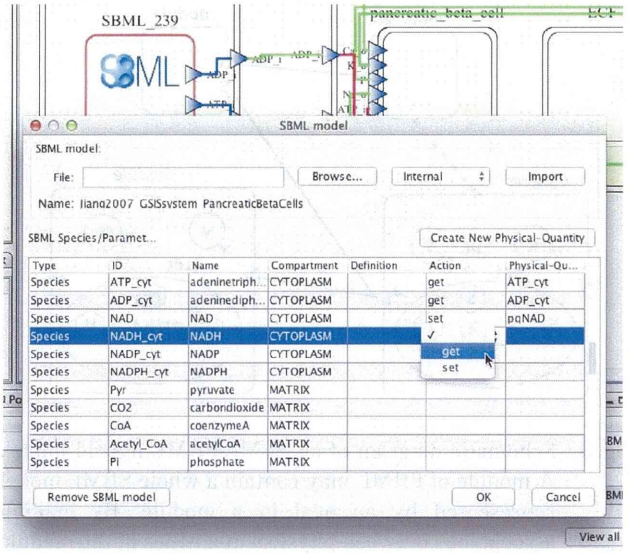

Fig. 4 Dialog of PhysioDesigner to support embedment of an SBML model into a module. It shows a list of species included in the SBML model. The two right-most columns provide an interactive interface to select the action (get or set) and a physical quantity to be associated.

3.3 Simulation of Hybrid Models

Flint is capable of parsing and simulating SBML models as well as PHML models. In particular, it is possible to take into account the SBML models embedded in a PHML model. Flint uses the SBML ODE Solver Library (SOSlib) [20] to extract formulas as abstract syntax trees (AST). SOSlib converts reaction rules or assignments in an SBML model to ODEs and algebraic equations, and events to conditional statements.

After extracting these ASTs, Flint detects species and parameters from which values should be transferred to physical quantities according to the bridge definition with a "get" action, or detects species and parameters which should be overridden by physical quantities according to the definition with a "set" action. Interpretation of a "get" action is simple. The value defined by a species or a parameter is transferred to a physical quantity. The value does not need to be constant. In the case of a "set" action, if the target is a parameter, the interpretation is still rather simple. The value of the parameter is replaced by that of the corresponding physical quantity. If the target is a species, we need to be careful not to introduce any inconsistency into the reaction network described in the SBML model. In our framework, the interpretation of the hybridization of SBML and PHML is carried out after extracting equations from the SBML side. At this stage, the dynamics of each species is described by a single ODE derived by interpreting the reaction formulae. The derivative of the species is replaced by the definition of a physical quantity without harming the consistency of the logic modeled in the other part of the SBML model.

Once the ASTs are merged into other formulae coming from the PHML part, they are sent together to the next stage to generate the bytecode for the execution of a simulation. Hence, Flint can handle elements defined in SBML Level 2, which is supported by the latest SOSlib. In other words, this is a limitation of Flint in supporting SBML. Note that to solve equations, Flint does not call the solver API of SOSlib. Instead, it creates a bytecode including the numerical integration algorithm implemented by Flint.

4. Examples of Hybrid Modeling

4.1 A Simple Example

A simple example of SBML-PHML hybrid modeling, basically a caricature model of the disposition of carboxydichlorofluroscein in hepatocyte, is illustrated in this section [21]. Only two or three players are extracted from the complicated signal transduction pathways originally proposed for this model (**Fig. 5A**). The model contains carboxydichlorofluroscein diacetate (CDFDA) in a hepatocyte, which is hydrolyzed to carboxydichlorofluroscein (CDF). The initial concentrations for CDFDA and CDF are set at 10 and 0 μM, respectively. By the hydrolytic reaction, the concentration of CDFDA decreases and CDF increases (**Fig. 5B**), which is simulated by Flint.

Based on the above SBML model, we expand the model on PhysioDesigner by adding a component representing extracellular CDFDA that can diffuse passively into the cell. Extracellular CDFDA is added as a physical quantity and a module in PHML creating a hybrid model (**Fig. 5C**).

From the mathematical point of view, the original dynamics of the intracellular CDFDA/CDF concentration are described by the following ordinary differential equations (ODEs).

$$\frac{d}{dt}[\mathrm{CDFDA_{in}}] = -k_1[\mathrm{CDFDA_{in}}], \quad (1)$$

$$\frac{d}{dt}[\mathrm{CDF_{in}}] = k_1[\mathrm{CDFDA_{in}}]. \quad (2)$$

where $[\mathrm{CDFDA}_{in}]$ and $[\mathrm{CDF}_{in}]$ represent the concentrations of intracellular CDFDA and CDF, respectively. $[\mathrm{CDFDA}_{in}]$ decreases monotonically and $[\mathrm{CDF}_{in}]$ increases. Adding extracellular CDFDA that diffuses into the hepatocyte modifies the ODEs as follows:

$$\frac{d}{dt}[\mathrm{CDFDA_{ex}}] = -k_2[\mathrm{CDFDA_{ex}}], \quad (3)$$

$$\frac{d}{dt}[\mathrm{CDFDA_{in}}] = -k_1[\mathrm{CDFDA_{in}}] + k_2[\mathrm{CDFDA_{ex}}], \quad (4)$$

$$\frac{d}{dt}[\mathrm{CDF_{in}}] = k_1[\mathrm{CDFDA_{in}}]. \quad (5)$$

where [CDFDAex] represents the concentration of extracellular CDFDA. After modification, $[\mathrm{CDFDA}_{in}]$ no longer decreases monotonically, but shows a single peak depending on the reaction velocity constants k_1 and k_2 (**Fig. 5D**).

The first step of SBML-PHML hybrid modeling for this example is to create two modules on PhysioDesigner. One is a hepatocyte module consisting of the dynamics of intracellular CDFDA/CDF, in which the SBML model is

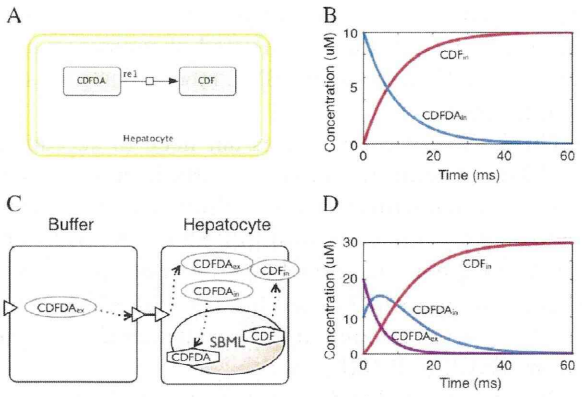

Fig. 5 A simple example of SBML-PHML hybrid modeling. A. Two-species model on CellDesigner. B. Simulation result of the simple model containing two species. The concentration of CDFDA decreases exponentially (blue curve), whereas that of CDF increases asymptotically (red curve). C. Schematic representation of the hybrid model of this simple example. D. Simulation results of the hybrid model from Flint. The concentration of external CDFDA starts from 20 μM and decreases (violet curve), and that of internal CDFDA has a unimodal peak (blue curve).

imported. The other is to implement the extracellular buffer, consisting of extracellular CDFDA. Once the SBML model is imported into the hepatocyte module, a variable-type physical quantity is created, and a bridge with a "get" action between the physical quantity and a species representing CDF in SBML are established to monitor its dynamics. In addition, the dynamics of intracellular CDFDA in the SBML model have to be overridden. For this, a state type physical quantity is created to implement an ODE shown in Eq. 4. The bridge between the state type physical quantity and the species with the "set" action should be defined.

Next, the buffer module has to be implemented with a state type physical quantity defined by Eq. 3. Additionally, a relationship between two modules has to be defined by linking an edge between them to transport the value of extracellular CDFDA, since the value is used in Eq. 4 in the hepatocyte module (**Fig. 5C**). **Figure 5D** presents the simulation result of the extended model, showing a monotonic decrease in extracellular CDFDA, a monotonic increase in intracellular CDF, and a unimodal increase followed by decrease in intracellular CDFDA.

4.2 Realistic Example

Let us observe another more realistic example involving insulin secretion from pancreatic β-cells. It is known that pancreatic β-cells exhibit complex and periodic spike-burst activity in response to an elevated concentration of extracellular glucose. There is a model that reproduces membrane-potential-level dynamics, called the Chicago model [22], which includes membrane potential, ATP/ADP concentrations, and various ionic currents such as sodium, potassium, and calcium. One of the components is an ATP-sensitive K^+ current, the channel of which is inhibited by high ATP concentration, resulting in membrane depolarization followed by an influx of Ca^{2+} and exocytosis of insulin granules. The model mainly focuses on the electrical mechanism of burst generation, in which oscillation of Ca^{2+} concentration, in particular, plays an important role. However, it does not pay much attention to the biochemical mechanism of glucose metabolism and ATP generation by the TCA cycle in mitochondria, although ATP concentration plays an important role in triggering insulin secretion. There is another model written in SBML that represents the glucose-stimulated insulin secretion network of pancreatic β-cells [23], which includes many entities relating to glycolysis, the TCA cycle, the respiratory chain, NADH shuttles, and the pyruvate cycle. This model, however, does not include membrane potential and ionic currents.

Using the above two models, we can create SBML-PHML hybrid multilevel models. Because the dynamics of ATP and ADP in the network model written in SBML (**Fig. 6A**) are described more carefully than those in the Chicago model (written in PHML), it is worthwhile to spool them up in the Chicago model. For this, the physical quantities representing ATP and ADP in the Chicago model are removed, and a module importing the network model is introduced (**Fig. 6B**). Then, two-variable parameter type physical quantities representing ATP and ADP concentrations are created in the module, and bridges from ATP and ADP species in the network model to those new physical quantities are established to get the values from the SBML portion to the PHML portion. By using these two physical quantities instead of the original in the Chicago model, the network model is effectively integrated through ATP and ADP dynamics, and the hybrid model using Flint exhibits a periodic burst of the membrane potential (**Fig. 6C**).

5. Linkage with Other Software

PhysioDesigner cannot edit SBML models embedded in modules. Because several kinds of software dedicated to SBML, such as CellDesigner, are available, it is not worthwhile to re-implement the same function in PhysioDesigner. To communicate with other software, we leverage the Garuda platform (http://www.garuda-alliance.org/), which is a new software developed to provide a means for systems biology tools to interoperate seamlessly. PhysioDesigner and Flint comply with the Garuda Alliance [24]. PhysioDesigner can extract an embedded SBML model from the PHML model, and send it to CellDesigner to browse or edit via the Garuda platform. Of course, it is also possible to receive an SBML model and import it into a module not only from CellDesigner but also from other software that can send out an SBML model via the Garuda platform.

6. Discussion

We have demonstrated a new function for creating SBML-PHML hybrid models developed on the existing versatile platform, PhysioDesigner. SBML is suitable for describing