nature genetics

FGF9 monomer–dimer equilibrium regulates extracellular matrix affinity and tissue diffusion

Masayo Harada[1,2], Hirotaka Murakami[3,11], Akihiko Okawa[3,11], Noriaki Okimoto[4,11], Shuichi Hiraoka[1,10,11], Taka Nakahara[5,10,11], Ryogo Akasaka[6,11], Yo-ichi Shiraishi[7,11], Noriyuki Futatsugi[4], Yoko Mizutani-Koseki[1], Atsushi Kuroiwa[7], Mikako Shirouzu[6], Shigeyuki Yokoyama[6,8], Makoto Taiji[4], Sachiko Iseki[5], David M Ornitz[9] & Haruhiko Koseki[1]

The spontaneous dominant mouse mutant, Elbow knee synostosis (*Eks*), shows elbow and knee joint synosotsis, and premature fusion of cranial sutures. Here we identify a missense mutation in the *Fgf9* gene that is responsible for the *Eks* mutation. Through investigation of the pathogenic mechanisms of joint and suture synostosis in *Eks* mice, we identify a key molecular mechanism that regulates FGF9 signaling in developing tissues. We show that the *Eks* mutation prevents homodimerization of the FGF9 protein and that monomeric FGF9 binds to heparin with a lower affinity than dimeric FGF9. These biochemical defects result in increased diffusion of the altered FGF9 protein (FGF9^{Eks}) through developing tissues, leading to ectopic FGF9 signaling and repression of joint and suture development. We propose a mechanism in which the range of FGF9 signaling in developing tissues is limited by its ability to homodimerize and its affinity for extracellular matrix heparan sulfate proteoglycans.

The fibroblast growth factors (FGFs) are widely expressed in developing and adult tissues and have diverse functions in organogenesis, tissue repair, nervous system control, metabolism and physiological homeostasis[1]. In humans and mice, the 22 FGF ligands are expressed in a spatiotemporally regulated manner and mediate signals through seven different isoforms of FGF receptors (FGFRs)[1]. The pharmacologic potential of FGF ligands has been highlighted by identification of gain-of-function mutations in genes encoding FGFRs 1–3 in individuals with chondrodysplasia and craniosynostosis syndromes[2,3]. These human diseases identify essential roles of FGF signaling not only in development but also in homeostasis of bones and joints.

Given these clinical, genetic and biochemical studies in humans and mice, the coordinated development of bones and joints seems to rely on precise FGFR signaling. This suggests that spatiotemporal constraints on FGF signaling are a prerequisite for appropriate functions *in vivo* and are indeed modulated at several distinct levels. First, the expression of FGF ligands is spatiotemporally restricted. Among the 22 FGF ligands, FGF2, FGF4, FGF7, FGF8, FGF9, FGF10, FGF17 and FGF18 are expressed in the limb bud and developing skeleton[4–6]. Of these, loss-of-function mutations have demonstrated that FGF2, FGF9 and FGF18 are involved in chondrogenesis and/or osteogenesis[7–10]. Induction of chondrodysplastic phenotypes by overexpression of FGF9 in mice also shows its ability to affect chondrogenesis[11]. Other elements implicated in FGF signaling are the heparan sulfate proteoglycans (HSPGs). Genetic studies in mice and *Drosophila melanogaster* suggest that HSPGs regulate the distribution and receptor binding of FGF ligands[12,13]. Finally, structural analyses of FGF9 suggest that it may form homodimers that could affect its ability to signal[14,15]. Because FGF9 homodimerization occludes several critical receptor binding sites, an autoinhibitory mechanism may function to modulate FGF9-dependent signal transduction. However, a functional demonstration of this proposed mechanism is lacking.

We have previously reported that a dominant mouse mutant, Elbow knee synostosis (*Eks*), shows radiohumeral and tibiofemoral synostosis, craniosynostosis (**Supplementary Fig. 1** online) and lung hypoplasia[16]. In this study, we identify a missense mutation that replaces Asn143 with threonine in the *Fgf9* gene in *Eks* mutant mice. We designate this mutant allele as *Fgf9*^{Eks} and show that this mutation

[1]RIKEN Research Center for Allergy and Immunology, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. [2]Department of Immunology and [3]Department of Orthopaedic Surgery, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan. [4]RIKEN Advanced Science Institute, Computational Systems Biology Research Group, 61-1 Ono-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0046, Japan. [5]Section of Molecular Craniofacial Embryology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8549, Japan. [6]RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. [7]Division of Biological Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan. [8]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [9]Department of Developmental Biology, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110, USA. [10]Present addresses: Department of Systems Biomedicine, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya-ku, Tokyo 157-8535, Japan (S.H.) and Section of Developmental and Regenerative Dentistry, School of Life Dentistry at Tokyo, The Nippon Dental University, 1-9-20 Fujimi, Chiyoda-ku, Tokyo 102-8159, Japan (T.N.). [11]These authors contributed equally to this work. Correspondence should be addressed to H.K. (koseki@rcai.riken.jp).

Received 20 June 2008; accepted 22 December 2008; published online 15 February 2009; doi:10.1038/ng.316

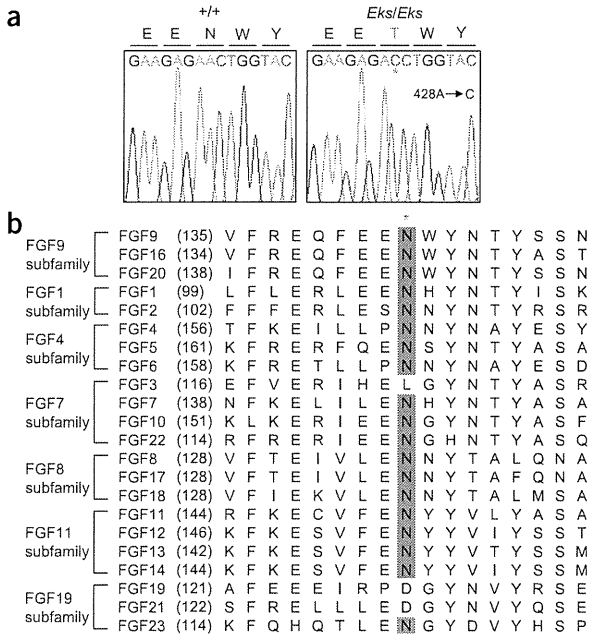

Figure 1 Missense mutation in the *Fgf9* gene of *Eks* mice. (**a**) Nucleotide sequence of the *Fgf9* cDNA derived from +/+ and *Eks*/*Eks* mice. *Eks* mutants have an A-to-C substitution at position 428, which results in the replacement of Asn143 with threonine. The *Eks* missense mutation is indicated by the asterisk, and the corresponding amino acid is shown in purple. (**b**) Structure-based sequence alignment of human FGFs. The amino acid sequence surrounding the N143T substitution in FGF9^{Eks} and that of its corresponding domain in other human FGF family proteins are aligned on the basis of sequence identity. The Asn143 residue in FGF9 is highly conserved among most FGF proteins (purple box). The asterisk denotes the site of *Eks* mutation.

prevents homodimerization of FGF9, consequently decreasing the affinity of FGF9 for heparin. As a result, FGF9^{Eks} is more diffusible in developing tissues, leading to ectopic FGF9 signaling in the prospective joints and sutures where it functions to repress development. Molecular-dynamics calculations suggest that the reduction in FGF9 affinity for heparin is due to the predominance of the monomeric form rather than to changes in its intrinsic affinity for heparin. We thus propose a mechanism in which the range of FGF9 signaling is limited through regulation of its affinity for HSPGs, which is at least in part controlled by the FGF9 monomer–dimer equilibrium. These observations could have far-reaching implications for the pharmacologic manipulation of FGF signaling under a variety of circumstances and in a wide range of tissues.

RESULTS

Eks is caused by a missense mutation in *Fgf9*

The *Eks* mutation was mapped between the polymorphic markers *D14Mit62* and *D14Mit5* on mouse chromosome 14 (ref. 16). Among 169 genes located in this interval, *Fgf9* seemed a likely candidate for the *Eks* mutation because FGF9 is a ligand for FGFR2c and FGFR3c[17] and is expressed in the developing limbs, cranial sutures and developing lung[8,18,19]. Sequence analysis of *Fgf9* cDNA from homozygous *Eks* mice revealed an A-to-C substitution at position 428, which resulted in the replacement of Asn143 with threonine (**Fig. 1a**). Notably, the Asn143 residue in FGF9 is highly conserved among most FGF proteins (**Fig. 1b**) and is predicted to be a critical amino acid residue for homodimerization and receptor binding[14,15].

We used a genetic approach to determine whether the N143T substitution in *Fgf9* was responsible for the *Eks* phenotype. We observed a mendelian pattern of inheritance of the mutation among 976 offspring of *Eks* heterozygous (*Fgf9*$^{Eks/+}$) matings, with heterozygous mice showing mild skeletal phenotypes and homozygous *Fgf9*$^{Eks/Eks}$ littermates showing severe skeletal phenotypes. The *Eks* phenotype and the mutation in *Fgf9* cosegregated in all cases. The absence of recombination between *Eks* and *Fgf9* among nearly 2,000 meiotic events provides strong evidence that the *Eks* mutation is allelic with *Fgf9*.

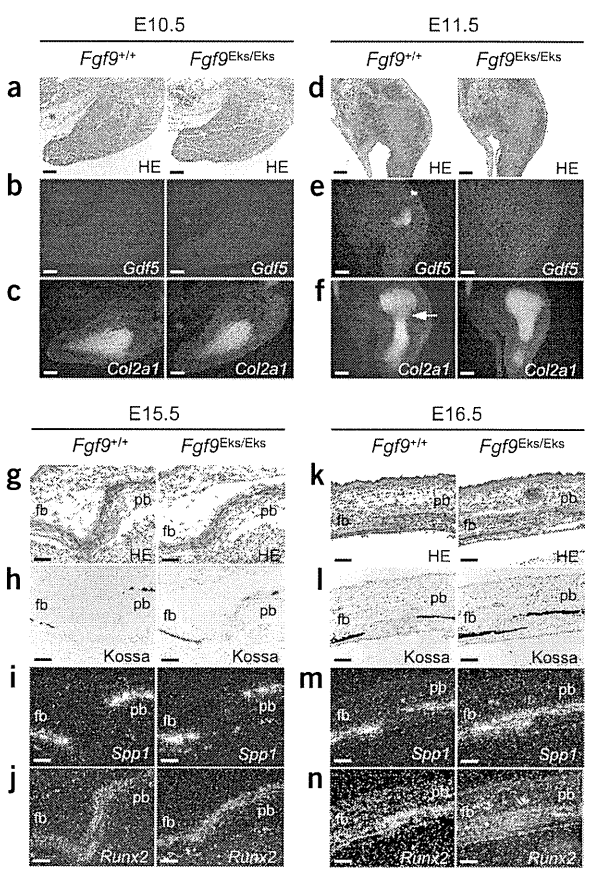

Figure 2 Synostotic phenotypes in *Fgf9*$^{Eks/Eks}$ mice. (**a**–**f**) Defects in early specification of prospective elbow joints in *Fgf9*$^{Eks/Eks}$ embryos. Hematoxylin and eosin staining (**a**,**d**) and *in situ* detection of *Gdf5* (**b**,**e**) and *Col2a1* (**c**,**f**) in the forelimb buds of *Fgf9*$^{+/+}$ and *Fgf9*$^{Eks/Eks}$ embryos at E10.5 and E11.5. In *Fgf9*$^{+/+}$ embryos, there was *Gdf5* expression at the prospective elbow joint (**e**, left), which was demarcated as the gap of *Col2a1* expression (**f**, arrow), at E11.5. In *Fgf9*$^{Eks/Eks}$ embryos, there was not *Gdf5* expression at the prospective elbow joint (**e**, right). Scale bars, 100 μm. (**g**–**n**) Ectopic osteogenesis at the coronal sutures in *Fgf9*$^{Eks/Eks}$ fetuses. Hematoxylin and eosin staining (**g**,**k**) and von Kossa staining (**h**,**l**) and *in situ* detection of *Spp1* (**i**,**m**) and *Runx2* (**j**,**n**) in the coronal suture of *Fgf9*$^{+/+}$ and *Fgf9*$^{Eks/Eks}$ fetuses at E15.5 and E16.5. Note the ectopic ossification in the suture of *Fgf9*$^{Eks/Eks}$ at E16.5. fb, frontal bone; pb, parietal bone. Scale bars, 100 μm.

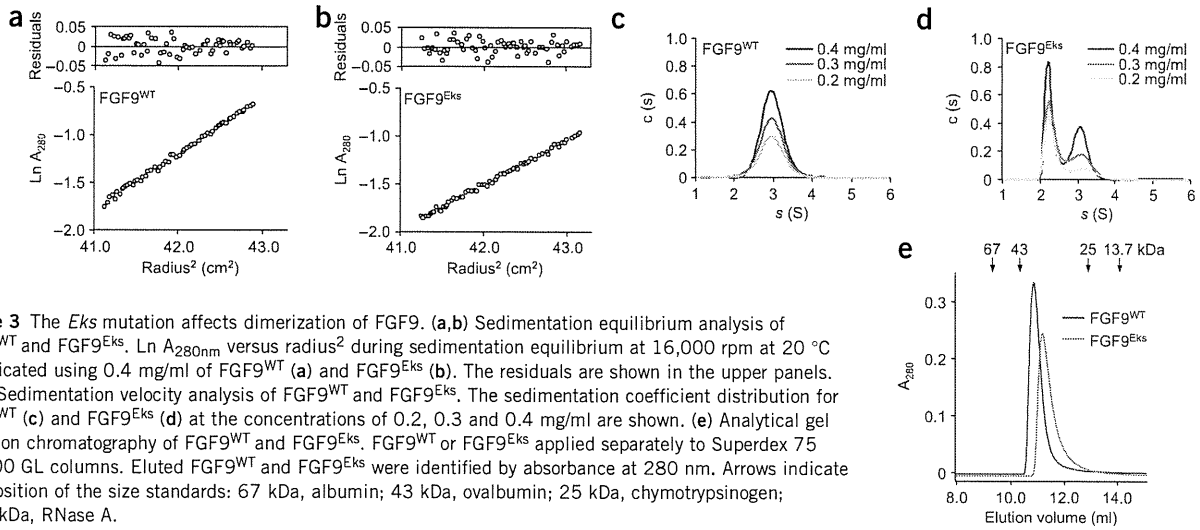

Figure 3 The *Eks* mutation affects dimerization of FGF9. (**a**,**b**) Sedimentation equilibrium analysis of $FGF9^{WT}$ and $FGF9^{Eks}$. Ln A_{280nm} versus radius^2 during sedimentation equilibrium at 16,000 rpm at 20 °C is indicated using 0.4 mg/ml of $FGF9^{WT}$ (**a**) and $FGF9^{Eks}$ (**b**). The residuals are shown in the upper panels. (**c**,**d**) Sedimentation velocity analysis of $FGF9^{WT}$ and $FGF9^{Eks}$. The sedimentation coefficient distribution for $FGF9^{WT}$ (**c**) and $FGF9^{Eks}$ (**d**) at the concentrations of 0.2, 0.3 and 0.4 mg/ml are shown. (**e**) Analytical gel filtration chromatography of $FGF9^{WT}$ and $FGF9^{Eks}$. $FGF9^{WT}$ or $FGF9^{Eks}$ applied separately to Superdex 75 10/300 GL columns. Eluted $FGF9^{WT}$ and $FGF9^{Eks}$ were identified by absorbance at 280 nm. Arrows indicate the position of the size standards: 67 kDa, albumin; 43 kDa, ovalbumin; 25 kDa, chymotrypsinogen; 13.7 kDa, RNase A.

Fgf9$^{Eks/Eks}$ mice phenocopy *Fgfr* gain-of-function mutants

Eks is a dominant mutation and joint synostosis and premature fusion of sutures have also been reported in mice expressing gain-of-function alleles of *Fgfr1* (ref. 20) and *Fgfr2c* (*Fgfr2*^{C342Y})[21]. Furthermore, homozygous *Fgf9* null mutants (*Fgf9*$^{-/-}$) show rhizomelia but not joint or suture synostosis[8]. Thus, we hypothesized that *Fgf9*^{Eks} encodes a gain-of-function mutation. To test this possibility, we first investigated whether there were phenotypic similarities between *Eks* mutants and gain-of-function mutants for *Fgfr1* and *Fgfr2*^{C342Y}. As initiation of elbow joint development was primarily impaired in *Fgfr1* gain-of-function transgenic mice[20], we examined radiohumeral joint development in *Fgf9*$^{Eks/Eks}$ mice (**Fig. 2a–f**). *Gdf5* (ref. 22) and *Col2a1* (ref. 23) demarcate the prospective elbow joint and cartilaginous condensation, respectively. *Gdf5* expression in the prospective joint space was observed as early as embryonic day E11.5 in *Fgf9*$^{+/+}$ control mice, but was completely absent in *Fgf9*$^{Eks/Eks}$ mice (**Fig. 2e**). Analysis of the prospective cartilage revealed a gap of *Col2a1* expression at the prospective elbow joint in E11.5 wild-type embryos (**Fig. 2f**, left). The gap of *Col2a1* expression was absent in *Fgf9*$^{Eks/Eks}$ mice (**Fig. 2f**, right). In summary, ectopic chondrocyte differentiation in the prospective elbow joint of *Fgf9*$^{Eks/Eks}$ mice seems very similar to that seen in transgenic mice that ectopically express an activated *Fgfr1* kinase domain in the presumptive joint field[20].

Premature fusion of coronal sutures in *Fgfr2*^{C342Y} mice results from excess osteogenic differentiation within the coronal suture mesenchyme[21]. To determine whether *Fgf9*$^{Eks/Eks}$ mice had similar histological features, we examined mineralization and the expression of the early osteoblast differentiation markers *Spp1* (ref. 24) and *Runx2* (ref. 25) in the coronal suture (**Fig. 2g–n**). At E15.5, both wild-type and *Fgf9*$^{Eks/Eks}$ mice showed similar coronal suture histology (**Fig. 2g–j**). However, at E16.5 von Kossa staining revealed considerably more overlap of the frontal and parietal bones in *Fgf9*$^{Eks/Eks}$ mice compared to *Fgf9*$^{+/+}$ mice (**Fig. 2l**). *Spp1* expression domains, which demarcate preosteoblasts and osteoblasts, showed wide separation of the frontal and parietal bones in *Fgf9*$^{+/+}$ mice; however, there was overlap in the *Fgf9*$^{Eks/Eks}$ mice (**Fig. 2m**). *Runx2* is highly expressed in immature osteoblasts at the leading edge of the frontal and parietal bones in *Fgf9*$^{+/+}$ mice (**Fig. 2n**, left). In E16.5 *Fgf9*$^{Eks/Eks}$ mice, the intensity of *Runx2* expression in the coronal suture was lower than in *Fgf9*$^{+/+}$ mice (**Fig. 2n**, right), suggesting premature differentiation of the osteoblasts at the presumptive suture. These phenotypes reflect abnormal osteogenic differentiation within the coronal suture mesenchyme and indicate that the defect in suture development occurs before E16.5. Taken together, these observations suggest that the $FGF9^{Eks}$ altered protein mediates excess FGFR signals within the prospective joints and sutures, preventing joint formation and promoting the fusion of cranial sutures.

The *Eks* mutation impairs homodimerization of FGF9

The predicted involvement of the Asn143 residue in homodimerization and receptor activation[14,15] suggests that changes in these processes may account for the apparent gain-of-function activity of the $FGF9^{Eks}$ altered protein. Homodimerization of FGF9 has been proposed to occlude receptor binding sites and consequently mediate an autoinhibitory mechanism for FGF9 signaling. We thus hypothesized that the *Eks* mutation might impair the autoinhibition. To test this possibility, we compared the degree of homodimerization of wild-type ($FGF9^{WT}$) and $FGF9^{Eks}$ proteins by analytical ultracentrifugation. $FGF9^{WT}$ and $FGF9^{Eks}$ were expressed in *Escherichia coli* and purified by column chromatography (**Supplementary Methods** and **Supplementary Fig. 2** online).

The molecular mass and association constant of $FGF9^{WT}$ and $FGF9^{Eks}$ were determined by sedimentation equilibrium centrifugation using the purified recombinant proteins (**Fig. 3a**,**b**). The estimated average molecular masses of $FGF9^{WT}$ and $FGF9^{Eks}$ were 39,264 and 32,929 Da, respectively, whereas the calculated monomeric molecular masses were 20,090 and 20,077 Da, respectively. These data suggest that $FGF9^{WT}$ primarily exists as a dimer, whereas $FGF9^{Eks}$ exists in a monomer–dimer equilibrium. The calculated association constants of $FGF9^{WT}$ and $FGF9^{Eks}$ were 10.4 μM^{-1} and 0.119 μM^{-1}, respectively. We further measured the sedimentation coefficient of $FGF9^{Eks}$ by sedimentation velocity centrifugation. The overlay plots of c(s)-sedimentation coefficient distributions show that $FGF9^{WT}$ has a unimodal peak at 3.0 S for a single ideal species, whereas $FGF9^{Eks}$ has bimodal peaks (2.2 S and 3.1 S) for two ideal species (**Fig. 3c**,**d**). This observation suggests that $FGF9^{WT}$ is present primarily as a dimer, whereas $FGF9^{Eks}$ exists primarily as a monomer. These interpretations are consistent with the retarded elution of $FGF9^{Eks}$

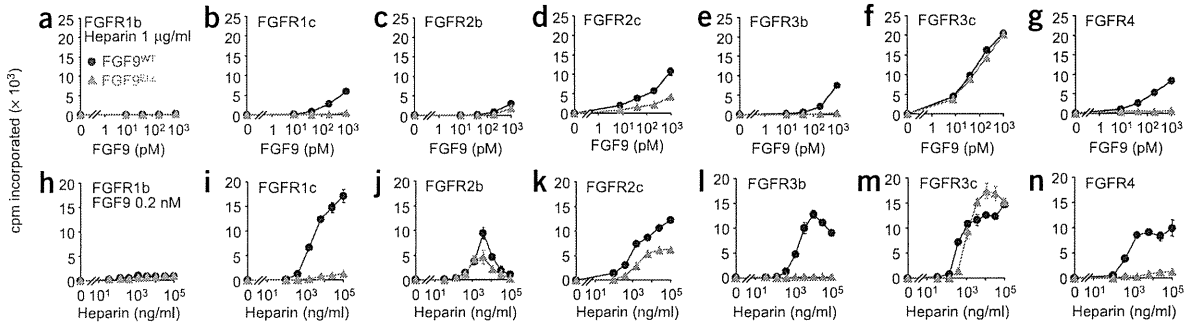

Figure 4 The *Eks* mutation affects the mitogenic activity of FGF9. (**a–g**) Dose-dependent changes in mitogenic activity of $FGF9^{WT}$ and $FGF9^{Eks}$. BaF3 cells expressing exogenous FGFR1b, 1c, 2b, 2c, 3b, 3c or 4 were treated with increasing concentrations of $FGF9^{WT}$ or $FGF9^{Eks}$ in the presence of 1 μg/ml heparin. Cell proliferation was determined by [^{3}H]thymidine incorporation after 36 h in culture. (**h–n**) Heparin-dependent changes in mitogenic activity of $FGF9^{WT}$ and $FGF9^{Eks}$. BaF3 cells expressing the respective FGFR were treated with increasing concentrations of heparin in the presence of 0.2 nM $FGF9^{WT}$ or $FGF9^{Eks}$. Cell proliferation was determined as above. Data are represented as mean ± s.e.m. of triplicate assays. These results are representative of at least two independent experiments.

relative to $FGF9^{WT}$ on a gel filtration column (**Fig. 3e**). Therefore, $FGF9^{Eks}$ is defective in homodimer formation.

$FGF9^{Eks}$ mediates less potent signaling via several FGFRs

To examine whether the *Eks* mutaion altered the binding of FGF9 to FGFRs by impairing the autoinhibitory mechanisms, we compared the ability of $FGF9^{WT}$ and $FGF9^{Eks}$ to activate FGFRs by assaying the mitogenic activity of both proteins on BaF3 cells expressing individual FGF receptors[17]. FGFR-expressing BaF3 cell lines were treated with increasing concentrations of purified recombinant FGF9 in the presence of 1 μg/ml heparin. Compared to $FGF9^{WT}$, $FGF9^{Eks}$ showed less activity on cells expressing any of the FGFRs except FGFR3c, where $FGF9^{Eks}$ showed equivalent activity (**Fig. 4a–g**). To test the ability of heparin to enhance FGF9 activity, we treated the BaF3 cell lines with increasing concentrations of heparin in the presence of 0.2 nM $FGF9^{WT}$ or $FGF9^{Eks}$. $FGF9^{Eks}$ also showed a decreased heparin-dependent mitotic response on all FGFRs except FGFR3c, where $FGF9^{Eks}$ showed higher activity in the presence of high concentrations of heparin (**Fig. 4h–n**). Because $FGF9^{Eks}$ does not mediate excess signaling via FGFRs, other properties of the altered protein must be responsible for the phenotype of the *Eks* mice.

The *Eks* mutation lowers FGF9 affinity for heparin

The decreased heparin-dependent mitogenic activity of $FGF9^{Eks}$ suggested that its affinity for heparin may be reduced. Heparin is functionally very similar to heparan sulfate, which is present in most tissues in the form of HSPGs. HSPGs function to modulate FGFR activation directly, by mediating FGF–FGFR interactions, and indirectly, by binding FGF ligands and regulating their diffusion through the extracellular matrix and thus their access to distant FGFRs[1,12,13,26]. As the gain-of-function property of $FGF9^{Eks}$ may not involve direct interaction with the FGFRs, we hypothesized that its decreased affinity for heparin might allow increased diffusion and thus bioavailability in

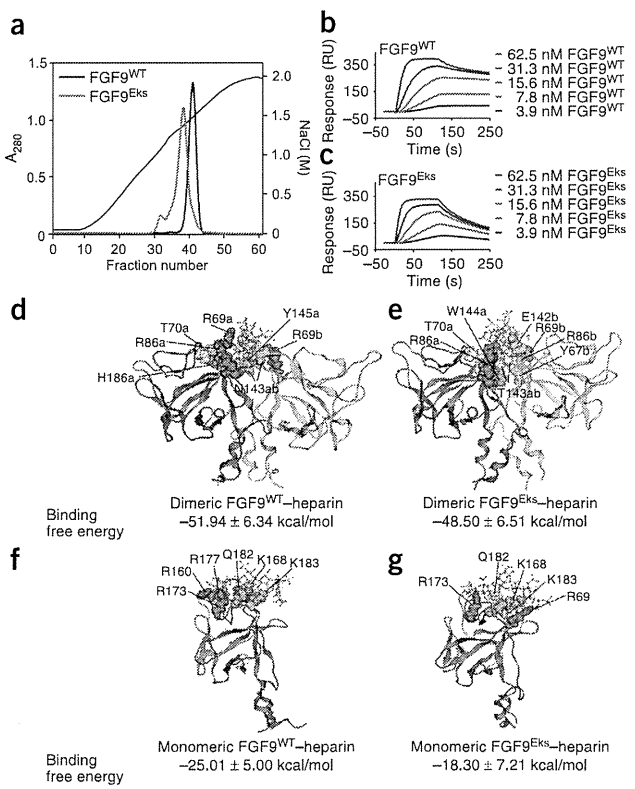

Figure 5 The *Eks* mutation reduces FGF9 affinity for heparin by impairing its dimerization. (**a**) Chromatographic analysis of the affinities of $FGF9^{WT}$ and $FGF9^{Eks}$ for heparin. $FGF9^{WT}$ or $FGF9^{Eks}$ were loaded onto a heparin-conjugated agarose column and eluted with a linear gradient of NaCl from 120 mM to 2.0 M (black line). Elution profiles of $FGF9^{WT}$ and $FGF9^{Eks}$ were determined by monitoring absorbance at 280 nm. (**b,c**) Surface plasmon resonance analysis of the affinities of $FGF9^{WT}$ and $FGF9^{Eks}$ for heparin. Sensorgrams indicating the interaction of $FGF9^{WT}$ (**b**) and $FGF9^{Eks}$ (**c**) with immobilized heparin were determined at different concentrations. The biosensor chip response is indicated on the *y* axis (RU) as a function of time (*x* axis) at room temperature. (**d–g**) The most probable solution structures of dimeric $FGF9^{WT}$–heparin (**d**), dimeric $FGF9^{Eks}$–heparin (**e**), monomeric $FGF9^{WT}$–heparin (**f**) and monomeric $FGF9^{Eks}$–heparin (**g**) complexes deduced by molecular-dynamics simulations. Heparin and protein residues that form important hydrogen bonds are drawn in ball and stick and space-filling modes. The single-letter amino acid code, residue number and chain code are indicated for each of these residues. Computed binding free energy of each complex is shown under the respective illustrated structure. Data are represented as mean ± s.d. of energies obtained from 200 molecular-dynamics snapshots in respective molecular-dynamics trajectory.

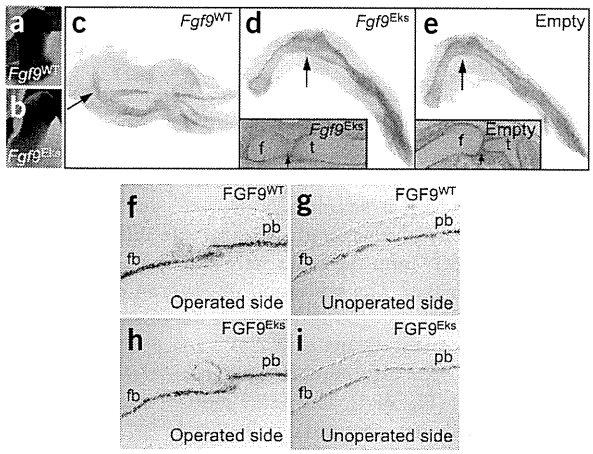

Figure 6 FGF9^{Eks} can inhibit joint and suture development as well as FGF9^{WT}. (**a–e**) Inhibition of knee joint development induced by ectopic expression of *Fgf9*^{WT} and *Fgf9*^{Eks}. Hindlimb buds of Hamburger-Hamilton stage 17 chickens were infected with RCAS-*Fgf9*^{WT}, RCAS-*Fgf9*^{Eks} or empty RCAS virus. (**a,b**) *Fgf9* expression was examined by *in situ* hybridization 2 d after infection. (**c–e**) Respective knee joints (arrows) were examined by Alcian blue staining and hematoxylin and eosin staining of sections through the knee joint (insets in **d** and **e**) 5 d after infection. f, femur; t, tibia. (**f–i**) Inhibition of coronal suture development by the ectopic presence of FGF9^{WT} and FGF9^{Eks}. FGF9^{WT} or FGF9^{Eks} beads were implanted onto the coronal suture in E15 mice, and the *Spp1* expression was examined by *in situ* hybridization 24 h after implantation. On the operated sides with FGF9^{WT} (**f**) and FGF9^{Eks} (**h**) bead implants, there was overlap of *Spp1* expression in the frontal and parietal bones, which was not seen on the unoperated sides (**g,i**). fb, frontal bone; pb, parietal bone.

tissues. To address this possibility, we first measured FGF9–heparin affinity by heparin affinity chromatography (**Fig. 5a**). FGF9^{WT} was eluted from heparin-conjugated agarose with 1.50 M NaCl as a single peak. In contrast, most FGF9^{Eks} was eluted at 1.38 M NaCl and a small fraction eluted at 1.10 M NaCl.

We next measured the kinetic constants for the FGF9^{Eks}–heparin interaction using surface plasmon resonance analysis (**Fig. 5b,c** and **Supplementary Table 1** online). The resulting sensorgrams were used for kinetic parameter determination by globally fitting the experimental data to a 1:1 interaction model. The association rate constant (k_a) of FGF9^{Eks} was slightly greater than that of FGF9^{WT}, whereas the dissociation rate constant (k_d) of FGF9^{Eks} was 18-fold greater than that of FGF9^{WT}. The dissociation constants (K_D) for FGF9^{WT} and FGF9^{Eks} were 0.71 ± 0.02 nM and 5.24 ± 0.03 nM, respectively, representing an 86% decrease in affinity of the FGF9^{Eks} protein for heparin.

Dimerization of FGF9 confers heparin affinity

The above studies indicate the *Fgf9*^{Eks} mutation concurrently affects monomer–dimer equilibrium and affinity for heparin. We thus went on to address whether the N143T substitution directly affects the affinity of FGF9 for heparin or whether it directly affects homodimerization and secondarily affects heparin affinity. However, direct biochemical measurements of the affinity of the two species for heparin are not possible because monomeric and dimeric forms of FGF9 are in equilibrium. We therefore analyzed the configuration of heparin-binding domains in monomeric and dimeric FGF9 using molecular-dynamics simulations and calculated the binding free energy between FGF9 and heparin using the molecular mechanics Poisson-Boltzmann/surface area (MM-PBSA) method. It is well known that the binding free energies calculated by this method show good qualitative but not quantitative agreement with experimental observations[27].

To model the heparin binding affinity of FGF9, we carried out molecular-dynamics simulations of 2:2 FGF9^{WT}–heparin and 2:2 FGF9^{Eks}–heparin complexes based on a 2:2:2 FGF2-FGFR1-heparin crystal structure (Protein Data Bank (PDB) ID: 1FQ9)[28]. The conformations of two heparin oligosaccharides in each complex were influenced by strong electrostatic repulsions, resulting in the exclusion of one heparin oligosaccharide from the complex (data not shown). This analysis suggested that 2:2 FGF9–heparin complexes would be unstable. In contrast, molecular-dynamics simulations of 2:1 FGF9^{WT}–heparin, 2:1 FGF9^{Eks}–heparin, 1:1 FGF9^{WT}–heparin and 1:1 FGF9^{Eks}–heparin complexes suggested that these complexes are stable (**Fig. 5d–g**). Molecular-dynamics simulations of dimeric FGF9–heparin complexes did not show a big difference in heparin-binding free energies for 2:1 FGF9^{WT}–heparin (dimeric FGF9^{WT}–heparin) and 2:1 FGF9^{Eks}–heparin (dimeric FGF9^{Eks}–heparin) complexes (**Fig. 5d,e**). This is due to the strong interaction between the negatively charged heparin oligosaccharide chain and the array of basic amino acid residues located in the heparin binding site near the groove of the dimer interface in both dimeric complexes. In addition, flexibility of the heparin oligosaccharide chain would help to maintain electrostatic interactions. Similarly, there was little difference in the heparin-binding free energies in 1:1 FGF9^{WT}–heparin (monomeric FGF9^{WT}–heparin) and 1:1 FGF9^{Eks}–heparin (monomeric FGF9^{Eks}–heparin) complexes (**Fig. 5f,g**). This is also due to heparin oligosaccharide chain flexibility, the strong negative charge of the heparin oligosaccharide and the presence of several basic amino acid residues in the heparin binding site. Therefore, the *Eks* mutation does not seem to influence the heparin binding affinity of either the dimeric or the monomeric FGF9–heparin complex. Because the heparin-binding free energies for dimeric FGF9 (**Fig. 5d,e**) were smaller than those for monomeric FGF9–heparin for both FGF9^{WT} and FGF9^{Eks} (**Fig. 5f,g**), the reduced binding affinity to heparin of the FGF9^{Eks} protein is most likely due to the shift in the monomer–dimer equilibrium toward the monomer. In summary, the *Eks* mutation primarily affects homodimerization of FGF9 and only secondarily affects heparin affinity.

FGF9^{Eks} is potentially hyperdiffusible in tissues

Heparin–FGF2 interactions have previously been shown to regulate the diffusibility of FGF2 (refs. 26,29). We hypothesized that the diffusibility of FGF9^{Eks} in tissues would be increased because of its lower affinity for heparin, leading to ectopic localization outside of the normal signaling domain and, consequently, ectopic activation of FGFRs. However, this model can only be considered if the following two prerequisites are met: first, *Fgf9* and *Fgfrs* are expressed in the proximity of the prospective elbow and knee joints and coronal sutures and, second, the increased diffusibility of FGF9^{Eks} is dominant over its decreased ability to activate FGFRs.

We first examined the expression of *Fgf9* and *Fgfr1*, *Fgfr2* and *Fgfr3* in the forelimb buds in E10.5 and E11.5 mice. *Fgf9* was expressed in migrating myoblasts, both in *Fgf9*$^{+/+}$ and *Fgf9*$^{Eks/Eks}$ mice (**Supplementary Fig. 3a,b,i,j** online). At E10.5, *Fgfr1*, *Fgfr2* and *Fgfr3* were expressed diffusely in the limb bud mesenchyme, overlapping the expression domain of *Col2a1* in both *Fgf9*$^{+/+}$ and *Fgf9*$^{Eks/Eks}$ tissues (**Supplementary Fig. 3c–h**). At E11.5, *Fgfr1* expression was excluded from the cartilaginous condensation, whereas *Fgfr2* and *Fgfr3*

expression was observed mainly in this location (**Supplementary Fig. 3k–p**). Therefore, mesenchymal cells in the prospective elbow joint express FGFRs.

Previous reports showed that *Fgf9* is expressed in the developing frontal and parietal bones, particularly strongly at the rims of the bones[18]. *Fgfr1*, *Fgfr2* and *Fgfr3* are expressed within and around the developing frontal and parietal bone domains[30]. Thus, the first prerequisite was validated.

FGF9^{Eks} is able to inhibit joint and suture development

To address the second prerequisite for our model, we examined the inhibitory effects of FGF9^{WT} and FGF9^{Eks} on joint development by ectopically expressing them in the chicken limb bud using a replication component retroviral vector (RCAS) transduction. RCAS-*Fgf9*^{WT}, RCAS-*Fgf9*^{Eks} or empty RCAS virus was used to infect the prospective hindlimb bud region in the lateral plate mesoderm. FGF9^{WT} and FGF9^{Eks} were expressed throughout the hindlimb buds (**Fig. 6a,b**). Ectopic expression of both *Fgf9*^{WT} and *Fgf9*^{Eks} caused knee joint fusion (**Fig. 6c,d**), whereas no abnormalities were induced by the empty vector (**Fig. 6e**). Therefore, FGF9^{Eks} retains inhibitory effects on joint development as well as FGF9^{WT}. It is notable that skeletal defects induced by the expression of FGF9^{WT} were widespread, whereas those mediated by FGF9^{Eks} were limited to the prospective joint regions. This is consistent with our finding that FGF9^{Eks} preferentially activates FGFR3c (**Fig. 4**), which is expressed in the bone anlagen (**Supplementary Fig. 3**), whereas FGF9^{WT} is expected to activate all of the mesodermally expressed FGFRs.

To examine the inhibitory effects of FGF9^{WT} and FGF9^{Eks} on suture development, we implanted FGF9-soaked AffiGel-Blue beads in the coronal suture of normal mouse fetal skulls around the initiation stage of the suture defect, at E15, by *ex utero* surgery. We first confirmed that approximately equal amounts of FGF9^{WT} and FGF9^{Eks} were loaded in each AffiGel-Blue bead and that the diffusion rates of FGF9^{WT} and FGF9^{Eks} from the beads were almost identical (**Supplementary Fig. 4** online). The expression of *Spp1*, an early osteoblast differentiation marker upregulated by FGF9, was examined 24 h after *in utero* bead placement. Grafts of FGF9^{WT} and FGF9^{Eks} beads also promoted ectopic *Spp1* expression at the leading edges of the frontal and parietal bones (**Fig. 6f–i**). This FGF9-induced ectopic expression of *Spp1* resembled that observed in the *Fgf9*$^{Eks/Eks}$ coronal suture (**Fig. 2m**, right). Therefore, ectopic expression of either FGF9^{Eks} or FGF9^{WT} within the suture inhibits suture development.

FGF9^{Eks} is more diffusible than FGF9^{WT} in developing tissues

To examine whether the diffusibility of FGF9^{Eks} in mesenchymal tissue is increased in comparison with FGF9^{WT}, we measured the diffusibility of FGF9^{WT} and FGF9^{Eks} in the skull following bead implantation (**Fig. 7a–e**). Because FGF9 upregulates *Spp1* expression, we could measure the area of high *Spp1* expression as an indication of

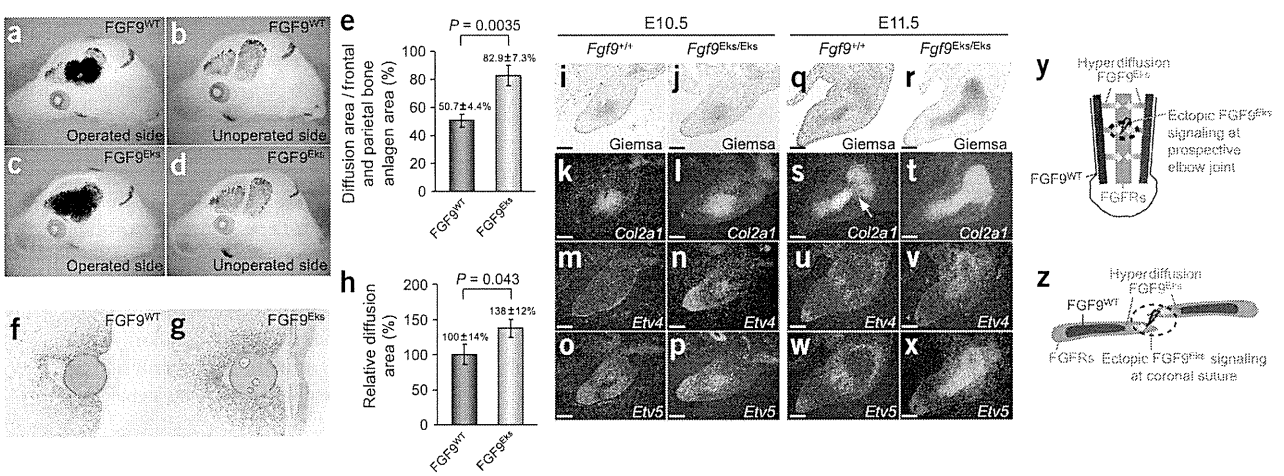

Figure 7 Ectopic FGF9^{Eks} signaling due to its hyperdiffusibility. (**a–e**) Increased diffusibility of FGF9^{Eks} in the skull bone anlagen. FGF9^{WT} or FGF9^{Eks} beads were implanted onto the coronal suture at E15 mice and *Spp1* expression was examined by whole-mount *in situ* hybridization 24 h after implantation. On the operated sides with FGF9^{WT} (**a**) and FGF9^{Eks} (**c**) bead implants, we observed well-defined intense signals in the frontal and parietal bone anlagen around the implants, which were not seen on the unoperated side (**b,d**). This domain with intense *Spp1* signals reflects diffusibility of exogenous FGF9 proteins. We therefore compared diffusibility of FGF9^{WT} and FGF9^{Eks} based on the induced expression domain of *Spp1* (**e**). The diffusion areas (%) in the frontal and parietal bone anlagen area were estimated from the area ratio of the intense *Spp1* expression against the frontal and parietal bone anlagen. Data are represented as mean ± s.e.m. of six operations. FGF9^{Eks} is more diffusible than FGF9^{WT}. Significance was determined by two-tailed Student's *t*-test. (**f–h**) Increased diffusibility of FGF9^{Eks} in the forelimb bud. FGF9^{WT} or FGF9^{Eks} beads were implanted into forelimb buds of *Fgf9*$^{-/-}$ embryos of E10.5 mice. Diffusion of exogenous FGF9^{WT} (**f**) and FGF9^{Eks} (**g**) after 2 h was immunodetected using a FGF9 antibody. (**h**) The diffusion area of FGF9^{Eks} and FGF9^{WT} was measured at the level of the equator of the beads. Data are represented as mean (FGF9^{WT} = 100%) ± s.e.m. of four (FGF9^{WT}) or five (FGF9^{Eks}) operations. FGF9^{Eks} is also more diffusible than FGF9^{WT} in limb buds. Significance was determined by one-tailed Student's *t*-test. (**i–x**) The downstream target genes of FGF signaling, *Etv4* and *Etv5*, are expressed ectopically in the prospective elbow joint in *Fgf9*$^{Eks/Eks}$ mice. Counterstaining with Giemsa (**i,j,q,r**) and *in situ* detection of *Col2a1* (**k,l,s,t**), *Etv4* (**m,n,u,v**) and *Etv5* (**o,p,w,x**) in the forelimb buds of *Fgf9*$^{+/+}$ and *Fgf9*$^{Eks/Eks}$ embryos at E10.5 and E11.5. There was ectopic expression of *Etv4* (**n,v**) and *Etv5* (**p,x**), in the cartilaginous condensation including the prospective elbow joint position, which was demarcated as the gap of *Col2a1* expression (**s**, arrow), in *Fgf9*$^{Eks/Eks}$, which were not seen in *Fgf9*$^{+/+}$ mice (**m,o,u,w**). Scale bars, 100 μm. (**y**) A model for the pathogenic mechanism underlying elbow joint synosotsis in *Fgf9*$^{Eks/Eks}$ mice. In *Fgf9*$^{Eks/Eks}$ mice, ectopic FGF9 signaling due to hyperdiffusion of FGF9^{Eks} at the prospective elbow joint may inhibit the initiation of joint development. (**z**) A model for the pathogenic mechanism underlying premature fusion of the coronal suture in *Fgf9*$^{Eks/Eks}$ mice. In *Fgf9*$^{Eks/Eks}$ mice, ectopic FGF9 signaling due to hyperdiffusion of FGF9^{Eks} at the coronal suture may promote ectopic osteogenesis and subsequently induce premature fusion of the suture.

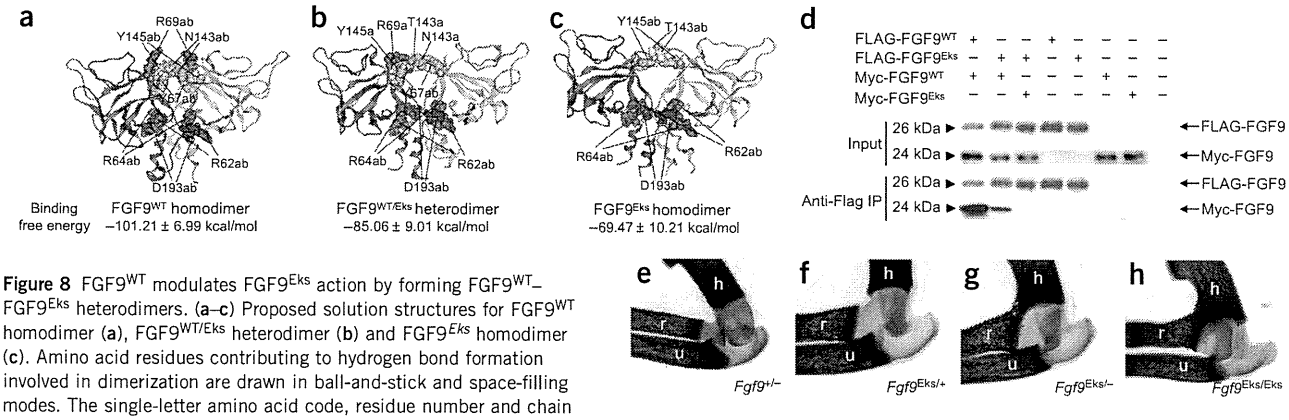

Figure 8 FGF9^{WT} modulates FGF9^{Eks} action by forming FGF9^{WT}–FGF9^{Eks} heterodimers. (**a–c**) Proposed solution structures for FGF9^{WT} homodimer (**a**), FGF9$^{WT/Eks}$ heterodimer (**b**) and FGF9^{Eks} homodimer (**c**). Amino acid residues contributing to hydrogen bond formation involved in dimerization are drawn in ball-and-stick and space-filling modes. The single-letter amino acid code, residue number and chain code are indicated for each of these residues. Computed binding free energy of each dimer is shown under the respective illustrated structure. Data are represented as mean ± s.d. of energies obtained from 200 molecular-dynamics snapshots in respective molecular-dynamics trajectory. (**d**) FGF9^{WT} is capable of forming dimers with FGF9^{Eks}. The expression of FGF9^{WT} homodimers, FGF9$^{WT/Eks}$ heterodimers and FGF9^{Eks} homodimers was analyzed using immunoprecipitation and protein blots. The expression vectors for FLAG- or Myc-tagged FGF9^{WT} and FGF9^{Eks} were transfected into COS7 cells and culture supernatants were subjected to immunoprecipitation and protein blot analysis. (**e–h**) Less severe elbow joint synostosis in *Fgf9*$^{Eks/+}$ than *Fgf9*$^{Eks/-}$. Forelimbs from *Fgf9*$^{+/-}$, *Fgf9*$^{Eks/+}$, *Fgf9*$^{Eks/-}$ and *Fgf9*$^{Eks/Eks}$ embryos at E17.5 were stained with Alcian blue and Alizarin red. Synostotic change is restricted to the cartilaginous component in *Fgf9*$^{Eks/+}$ embryos, whereas it is extended to the bony component in *Fgf9*$^{Eks/-}$ and *Fgf9*$^{Eks/Eks}$ embryos. h, humerus; r, radius; u, ulna.

the distance over which FGF9 exerts its effects. Implantation of FGF9^{Eks} beads resulted in a larger area of *Spp1* expression (high *Spp1* expression area / Frontal and parietal bone anlagen area = 82.9 ± 7.3%) (± s.e.m.) compared with beads loaded with FGF9^{WT} (50.7 ± 4.4%; $P = 0.0035$), suggesting that the altered protein diffused more effectively through the developing tissue.

Next, we investigated the diffusibility of FGF9^{Eks} in forelimb buds. FGF9^{WT}- or FGF9^{Eks}-soaked AffiGel-Blue beads were grafted into the dorsal and central forelimb bud region of *Fgf9*$^{-/-}$ embryos around the initiation stage of the joint defect, at E10.5. FGF9 protein released from the beads into mesenchymal tissue 2 h after implantation was detected by immunohistochemistry using FGF9 antibodies on sections (**Fig. 7f–h**). This analysis showed that FGF9^{Eks} permeated through the limb bud mesenchyme to a greater extent (relative disffusion area = 138 ± 12%) (± s.e.m.) than FGF9^{WT} (100 ± 14%; $P = 0.043$), supporting the hypothesis that FGF9^{Eks} is more diffusible than FGF9^{WT}.

Ectopic FGF signaling in the prospective joint of *Eks* mice

To examine whether diffusion of endogenous FGF9^{Eks} is increased in comparison to FGF9^{WT}, we determined the degree of activation of FGFRs in the prospective elbow joint of *Fgf9*$^{Eks/Eks}$ mice. As a readout for FGFR signaling, we examined the expression of *Etv4* and *Etv5*, both of which are known to be transcriptionally activated by FGF signaling, in the forelimb buds[31] (**Fig. 7i–x**). In wild-type E10.5 limbs, we did not observe the intensive expression of *Etv4* or *Etv5* within the region undergoing cartilaginous condensation demarcated by *Col2a1* expression (**Fig. 7m,o**). However, in E10.5 *Fgf9*$^{Eks/Eks}$ limbs, we found ectopic expression of both *Etv4* and *Etv5* in the cartilaginous condensation (**Fig. 7n,p**). At E11.5, *Etv4* and *Etv5* were expressed in the myoblasts and cells surrounding the cartilaginous condensation in wild-type mice (**Fig. 7u,w**), whereas in *Fgf9*$^{Eks/Eks}$ mice, we observed clear expression of *Etv4* and *Etv5* in the cartilaginous condensation including the prospective elbow joint position (**Fig. 7v,x**). These results demonstrate ectopic FGF signaling in the prospective elbow joint in *Fgf9*$^{Eks/Eks}$ mice. Because the ectopic expression domain of *Etv4* and *Etv5* in *Fgf9*$^{Eks/Eks}$ mice was not consistent with the *Fgf9* expression domain (**Supplementary Fig. 3a,b,i,j**), this outcome is likely due to increased diffusion of FGF9^{Eks} protein over a larger area than FGF9^{WT}.

From these results, we propose a mechanism of elbow joint synostosis in *Fgf9*$^{Eks/Eks}$ mice in which FGF9^{Eks} produced by myoblasts diffuses beyond its normal range and ectopically activates FGFRs in the prospective elbow joint, preventing joint formation (**Fig. 7y**). Similarly, FGF9^{Eks} produced at the rims of the frontal and parietal bones diffuses beyond its normal area and ectopically activates FGFRs in the coronal suture mesenchyme, promoting the fusion of coronal sutures (**Fig. 7z**).

Genetic validation of the hyperdiffusible model

A prediction of our model is that the severity of synostotic phenotypes will correlate with a shift in the equilibrium of FGF9 from dimer toward monomer. By molecular-dynamics simulations, we estimated that FGF9^{WT} homodimer, FGF9$^{WT/Eks}$ heterodimer and FGF9^{Eks} homodimer have ten, eight, six intermonomer hydrogen bonds, respectively (**Fig. 8a–c** and **Supplementary Table 2** online). From these results, we calculated the binding free energy of the FGF9$^{WT/Eks}$ heterodimer to be between those of the FGF9^{WT} and FGF9^{Eks} homodimers (**Fig. 8a–c**), suggesting that FGF9$^{WT/Eks}$ heterodimers are more stable than FGF9^{Eks} homodimers. If our model is correct, FGF9^{WT} should interfere with FGF9^{Eks} action by forming FGF9$^{WT/Eks}$ heterodimers. In other words, skeletal phenotypes due to the *Eks* mutation should be alleviated by the expression of FGF9^{WT}.

We first sought experimental evidence that the FGF9$^{WT/Eks}$ heterodimer was substantially more stable than the FGF9^{Eks} homodimer. We addressed this issue by immunoprecipitation and protein blot analysis after tagging FGF9^{WT} and FGF9^{Eks} with Flag or Myc peptides. Flag-FGF9^{WT} or Flag-FGF9^{Eks} was overexpressed together with Myc-FGF9^{WT} or Myc-FGF9^{Eks} in COS7 cells and the culture supernatants were subjected to the immunoprecipitation and protein blot analysis (**Fig. 8d**). We readily detected Flag-FGF9^{WT}–Myc-FGF9^{WT} dimers. Flag-FGF9^{Eks}–Myc-FGF9^{WT} dimers were detected at lower level; however, Flag-FGF9^{Eks}–Myc-FGF9^{Eks}

dimers did not form under these conditions. These data suggest that $FGF9^{WT}$ and $FGF9^{Eks}$ can form heterodimers that are more stable than $FGF9^{Eks}$ homodimers.

We finally examined whether the elbow synostosis caused by $FGF9^{Eks}$ could be alleviated by the expression of $FGF9^{WT}$. We thus compared the severity of elbow synostosis in $Fgf9^{Eks/+}$ and compound heterozygous mutants ($Fgf9^{Eks/-}$) relative to that of $Fgf9^{+/-}$ and $Fgf9^{Eks/Eks}$ mice (**Fig. 8e–h**). Elbow joint formation was not affected in $Fgf9^{+/-}$ mice (**Fig. 8e**), whereas elbow synostosis in $Fgf9^{Eks/Eks}$ mice involved both cartilaginous and bony components (**Fig. 8h**). In contrast, the synostosis in $Fgf9^{Eks/+}$ mice was limited to the cartilaginous component (**Fig. 8f**), whereas the involvement of the bony component in $Fgf9^{Eks/-}$ mice was similar to that of $Fgf9^{Eks/Eks}$ mice (**Fig. 8g**). Therefore, elbow synostosis in $Fgf9^{Eks/-}$ is more severe than in $Fgf9^{Eks/+}$. These findings strongly support our model that the monomer–dimer status of FGF9 influences its affinity for HSPGs and, consequently, its distribution in developing tissues.

DISCUSSION

In the present study, we identified a missense mutation in the *Fgf9* gene that is responsible for the *Eks* mutant phenotype, which includes elbow and knee joint synostosis and craniosynostosis. We further demonstrate that the N143T substitution in FGF9 favors formation of the monomeric form of FGF9, which binds to heparin with a lower affinity than dimeric FGF9. The decreased affinity for heparin leads to increased diffusion of the altered protein in developing tissues, resulting in ectopic FGF9 signaling. We propose that regulation of the monomer–dimer equilibrium of FGF9, and potentially of other FGFs, and its affinity for HSPGs is a mechanism that functions to shape FGF9 concentration gradients in developing tissues. We further propose that these biochemical properties of FGF9 restrict its signaling activity to limited skeletal domains. Data presented here and in previous studies indicate that low FGF signaling in the presumptive joint space is necessary for the formation of the joint space and maintenance of an open suture[20,21]. Common usage of receptor binding and homodimerization sites of FGF9 could be at least in part instrumental for local modulation of FGF9 signaling activity.

Homodimerization is suggested to be a common feature of the FGF9/16/20 subfamily[32] and of FGF2 (refs. 33,34). It is not known to what extent homodimerization affects the activity of other FGFs. The discovery that a mutation in *Fgf9* can affect homodimerization, affinity for heparin and biological activity suggests that pharmacological agents that affect FGF homodimerization could be useful tools to modulate its activity.

METHODS

Detection of a mutation in the *Fgf9* gene. To identify the mutation responsible for the *Eks* mutant phenotype, we surveyed the cDNA sequence of *Fgf9* from normal (+/+), heterozygous (*Eks*/+) and homozygous (*Eks/Eks*) mice through reverse transcription-PCR (RT-PCR) and direct sequencing of RT-PCR products.

For genotyping of the *Eks* allele, genomic DNA spanning the *Eks* mutation was amplified by PCR using specific primers (**Supplementary Table 3** online). PCR products were digested with the diagnostic *BsrI* restriction enzyme. Wild-type mice show two bands of 147 bp and 42 bp, whereas the *Eks* allele shows one band of 189 bp (**Supplementary Fig. 5** online).

Skeletal and histological preparations. Bones and cartilage of E17.5 fetuses were stained with Alizarin red and Alcian blue as previously described[35]. For histological preparations, tissues were fixed in 4% paraformaldehyde, embedded in paraffin, sectioned at 5 μm, and stained with hematoxylin and eosin and von Kossa.

***In situ* hybridization.** *In situ* hybridization of paraffin sections was done as previously described[36], using radiolabeled antisense RNA for *Gdf5* (MGI: 95688), *Col2a1* (MGI: 88452), *Spp1* (MGI: 98389), *Runx2* (MGI: 99829), *Fgf9* (MGI: 104723), *Fgfr1* (MGI: 95522), *Fgfr2* (MGI: 95523), *Fgfr3* (MGI: 95524), *Etv4* (MGI: 99423) and *Etv5* (MGI: 1096867). *In situ* hybridization after bead implantation in fetal skulls was done as previously described[24].

$FGF9^{WT}$ and $FGF9^{Eks}$ expression and purification. $FGF9^{WT}$ and $FGF9^{Eks}$ expression and purification were done as described in **Supplementary Methods.**

Analytical ultracentrifugation. All analytical ultracentrifuge experiments were done on a Beckman Coulter XL-I analytical ultracentrifuge. We diluted the samples in 25 mM ammonium acetate buffer (pH 5.5) containing 120 mM NaCl. The partial specific volumes were estimated as 0.7317 ml/g ($FGF9^{WT}$) or 0.7322 ml/g ($FGF9^{Eks}$) by SEDNTERP software. All experiments were done at 20 °C and the absorbance wavelength was 280 nm. We carried out sedimentation equilibrium experiments with six channel centerpieces, with loading concentrations of 0.8, 0.4 and 0.2 mg/ml. Data were obtained at 12, 14 and 16 k rpm for $FGF9^{WT}$ or at 14, 16 and 18 k rpm for $FGF9^{Eks}$. A total equilibration time of 16 h was used for each speed with scans taken at 12, 14 and 16 h. We analyzed the sedimentation equilibrium data using the Beckman XL-A/XL-I Data Analysis software. All nine datasets (three speeds, three concentrations) were fitted together by 'self association model' calculation. Sedimentation velocity experiments were carried out with double sector centerpieces. Protein concentrations were 0.4, 0.3 or 0.2 mg/ml. We scanned the absorbance data 100 times every 5 min at 40 k rpm. The measurements data were analyzed by SEDFIT software.

Analytical gel filtration chromatography. Purified $FGF9^{WT}$ and $FGF9^{Eks}$ (100 μl of 2 mg/ml) were loaded onto a Superdex75 10/300 GL column (GE Healthcare) equilibrated with a 25 mM ammonium acetate buffer (pH 5.5) containing 120 mM NaCl. The samples were eluted with the same buffer.

Mitogenic assays. The ability of $FGF9^{WT}$ and $FGF9^{Eks}$ to transduce signals via FGFRs was analyzed by a mitogenic assay using BaF3 cells expressing specific FGFRs as described previously[17]. We plated 5,000 cells per well in a 96-well assay plate in media containing varying concentrations of FGF9 and heparin (Wako). $FGF9^{WT}$ or $FGF9^{Eks}$ were added to each well for a total volume of 200 μl per well. The cells were then incubated at 37 °C for 36 h. We added 1 μCi of [^{3}H]thymidine to each well in 20 μl of media. The cells were harvested after 4 h by filtration through glass fiber paper and the incorporated [^{3}H]thymidine was counted on a Wallac MicroBeta TriLux scintillation counter (PerkinElmer).

Analytical heparin affinity chromatography. We loaded 3 mg of purified $FGF9^{WT}$ and $FGF9^{Eks}$ onto a 1 ml HiTrap heparin HP column (GE Healthcare) equilibrated with 25 mM ammonium acetate buffer (pH 5.5) containing 120 mM NaCl. The bound $FGF9^{WT}$ or $FGF9^{Eks}$ were eluted with a linear gradient of NaCl (120 mM to 2.0 M) in the same buffer.

Surface plasmon resonance analysis of FGF9–heparin interactions. Surface plasmon resonance analysis for measurements of $FGF9^{WT}$–heparin and $FGF9^{Eks}$–heparin interactions were done using a BIAcore 3000 instrument (Biacore AB). In order to immobilize heparin (Wako) on the streptavidin-conjugated sensor chip SA, 100 μg/ml biotinylated heparin in HBS-EP buffer was injected at a flow rate of 10 μl/min and was immobilized to 63 response units (RU). All measurements were carried out at room temperature, and refractive index errors due to bulk solvent effects were corrected by subtracting responses on the noncoated sensor chip for the $FGF9^{WT}$ and $FGF9^{Eks}$ concentrations used. To obtain kinetic data, we injected different concentrations of analytes ($FGF9^{WT}$ and $FGF9^{Eks}$) in HBS-EP over the heparin sensor chip at a flow rate of 20 μl/min. At the end of each sample injection (120 s), HBS-EP buffer was passed over the sensor surface to monitor the dissociation phase. Following 120 s of dissociation, the sensor surface was regenerated by injection of 5 μl of 1 M NaCl in HBS-EP. We used five different analyte concentrations to determine the kinetic parameters for each interaction. Kinetic parameters were obtained by global fitting of the sensorgrams to a 1:1 model using BIAevaluation software.

Molecular-dynamics simulation. Starting structures of monomeric $FGF9^{WT}$, dimeric $FGF9^{WT}$, monomeric $FGF9^{WT}$-heparin, dimeric $FGF9^{WT}$-heparin and $FGF9^{Eks}$ for molecular-dynamics simulations were taken from the PDB (PDB ID: 1IHK)[14]. The structures of monomeric $FGF9^{Eks}$, dimeric $FGF9^{Eks}$ and heterodimeric $FGF9^{WT/Eks}$ were constructed based on $FGF9^{WT}$ using molecular modeling software MOE (Chemical Computing Group). A hexasasaccharide (UAP-SGN-IDU-SGN-IDU-SGN) is used as a heparin oligosaccharide. UAP is 1,4-dideoxy-5-dehydro-O2-sulfo-glucuronic acid, SGN is O6-disulfo-glucosamin and IDU is 1,4-dideoxy-O2-sulfo-glucuronic acid. For monomeric $FGF9^{WT}$–heparin and dimeric $FGF9^{WT}$–heparin simulations, heparin oligosaccharide was bound to $FGF9^{WT}$ structures obtained from molecular-dynamics simulations using the molecular docking program GOLD (version 3.0)[37]. In the docking protocol, the standard default setting of GA parameters was used and GoldScore was used as the scoring function. The structures of monomeric $FGF9^{Eks}$–heparin and dimeric $FGF9^{Eks}$–heparin were built in the same manner as $FGF9^{WT}$–heparin complexes. All the starting structures for molecular-dynamics simulations were surrounded by TIP3P water molecules[38] spherically. After energy minimizations, all molecular-dynamics simulations were carried out for 10 ns at 300 K using modified Amber 8.0 (ref. 39) for MDGRAPE3 system[40]. The Amber ff03 force field[41] was adopted, and the simulation time step was set at 1 s. We calculated the binding free energies by the molecular mechanics Poisson-Boltzmann/surface area (MM-PBSA) method[42] using the last 2 ns molecular-dynamics trajectories.

Retroviral misexpression. Mouse *Fgf9*^{WT} and *Fgf9*^{Eks} cDNAs were cloned into the RCASBP(A) vector[43]. The virus solutions were injected into the hind limb bud of chicken embryos at Hamburger-Hamilton stage 17. We examined the expression of mouse *Fgf9* transcripts and skeletal changes 2 and 5 d after injection, respectively.

Subcutaneous insertion of $FGF9^{Eks}$ beads in mouse fetal skulls. AffiGel-Blue beads (BioRad) soaked in 100 μg/ml $FGF9^{WT}$ or $FGF9^{Eks}$ were implanted onto E15 mouse skulls by *ex utero* surgery as previously described[24] Operated heads were collected 24 h later and *Spp1* transcripts were detected by whole-mount *in situ* hybridization. The area of *Spp1* expression was measured using NIH image software.

Implantation of $FGF9^{Eks}$ beads in mouse forelimb buds. AffiGel-Blue beads that had been soaked in 500 μg/ml $FGF9^{WT}$ or $FGF9^{Eks}$ were implanted into the dorsal and central region of E10.5 *Fgf9*$^{-/-}$ forelimb buds. Limb buds were subsequently cultured for 2 h on Transwell filters (Costar, Coaning) in serum-free medium (BGJb, 2 mg/ml BSA, penicillin (50 units/ml), streptomycin (50 μg/ml)), in a humid, 37 °C and 5% CO_2 environment. Explants were fixed in 4% paraformaldehyde and embedded in paraffin. Sections through the equator of the bead were analyzed for exogenous FGF9 using goat antibody to FGF9 (R&D Systems) and a cell and tissue staining kit HRP-AEC system (R&D Systems). We analyzed the signal area and intensity using NIH image software.

Immunoprecipitation and protein blot analysis. cDNA fragments encoding the full length mouse $FGF9^{WT}$ and $FGF9^{Eks}$ proteins were cloned into the p3xFlag-CMV-14 vector (Sigma-Aldrich) and into the pCMV-Tag3B vector (Stratagene) to allow expression of FGF9 proteins with either C-terminal Myc or 3xFlag tags. These vectors were transfected into COS7 cells, and 48 h later, culture supernatants were incubated with anti-Flag M2 affinity beads (Sigma-Aldrich) for 1 h at RT and washed three times with PBS and then subjected to protein blots with antibody to Flag M2 (Sigma-Aldrich) or antibody to Myc (Upstate).

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

This study was supported in part by the RIKEN Structural Genomics/Proteomics Initiative (RSGI) and the National Project on Protein Structural and Functional Analysis, Ministry of Education, Culture, Sports, Science and Technology of Japan (S.Y.) and US National Institutes of Health grant HD049808 (D.M.O.).

AUTHOR CONTRIBUTIONS

M.H., D.M.O. and H.K. developed the project and wrote the manuscript. M.H., S.H. and H.K. contributed to the purification of FGF9 proteins, mitogenic assays, analytical gel filtration chromatography, analytical heparin affinity chromatography, surface plasmon resonance analysis, skeletal preparation, histological analyses and *in situ* hybridization of sections, implantation of FGF9 beads in mouse forelimb buds and immunoprecipitation and protein blot analysis. H.M., A.O. and H.K. contributed to the identification of the *Eks* mutation. N.O., N.F. and M.T. contributed to the molecular-dynamics simulation. T.N. and S.I. contributed to the implantation of FGF9 beads in mouse fetal skulls. R.A., M.S. and S.Y. contributed to the analytical ultracentrifugation. Y.S. and A.K. contributed to the retroviral misexpression. Y.M.-K. contributed to the *in situ* hybridization of sections.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Ornitz, D.M. & Itoh, N. Fibroblast growth factors. *Genome Biol.* **2**, reviews 3005 (2001).
2. Wilkie, A.O.M. Bad bone, absent smell, selfish testes: the pleiotropic consequences of human FGF receptor mutations. *Cytokine Growth Factor Rev.* **16**, 187–203 (2005).
3. Su, N., Du, X. & Chen, L. FGF signaling: its role in bone development and human skeleton diseases. *Front. Biosci.* **13**, 2842–2865 (2008).
4. Ornitz, D.M. & Marie, P.J. FGF signaling pathways in endochondral and intramembranous bone development and human genetic disease. *Genes Dev.* **16**, 1446–1465 (2002).
5. Martin, G.R. The roles of FGFs in the early development of vertebrate limbs. *Genes Dev.* **12**, 1571–1586 (1998).
6. Ornitz, D.M. FGF signaling in the developing endochondral skeleton. *Cytokine Growth Factor Rev.* **16**, 205–213 (2005).
7. Montero, A. *et al.* Disruption of the fibroblast growth factor-2 gene results in decreased bone mass and bone formation. *J. Clin. Invest.* **105**, 1085–1093 (2000).
8. Hung, I.H., Yu, K., Lavine, K.J. & Ornitz, D.M. FGF9 regulates early hypertrophic chondrocyte differentiation and skeletal vascularization in the developing stylopod. *Dev. Biol.* **307**, 300–313 (2007).
9. Liu, Z., Xu, J., Colvin, J.S. & Ornitz, D.M. Coordination of chondrogenesis and osteogenesis by fibroblast growth factor 18. *Genes Dev.* **16**, 859–869 (2002).
10. Ohbayashi, N. *et al.* FGF18 is required for normal cell proliferation and differentiation during osteogenesis and chondrogenesis. *Genes Dev.* **16**, 870–879 (2002).
11. Garofalo, S. *et al.* Skeletal dysplasia and defective chondrocyte differentiation by targeted overexpression of fibroblast growth factor 9 in transgenic mice. *J. Bone Miner. Res.* **14**, 1909–1915 (1999).
12. Ornitz, D.M. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *Bioessays* **22**, 108–112 (2000).
13. Nybakken, K. & Perrimon, N. Heparan sulfate proteoglycan modulation of developmental signaling in *Drosophila*. *Biochim. Biophys. Acta* **1573**, 280–291 (2002).
14. Plotnikov, A.N. *et al.* Crystal structure of fibroblast growth factor 9 reveals regions implicated in dimerization and autoinhibition. *J. Biol. Chem.* **276**, 4322–4329 (2001).
15. Hecht, H.J. *et al.* Structure of fibroblast growth factor 9 shows a symmetric dimer with unique receptor- and heparin-binding interfaces. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 378–384 (2001).
16. Murakami, H. *et al.* Elbow knee synostosis (Eks): a new mutation on mouse Chromosome 14. *Mamm. Genome* **13**, 341–344 (2002).
17. Ornitz, D.M. *et al.* Receptor specificity of the fibroblast growth factor family. *J. Biol. Chem.* **271**, 15292–15297 (1996).
18. Hajihosseini, M.K. & Heath, J.K. Expression patterns of fibroblast growth factors-18 and -20 in mouse embryos is suggestive of novel roles in calvarial and limb development. *Mech. Dev.* **113**, 79–83 (2002).
19. Colvin, J.S., Feldman, B., Nadeau, J.H., Goldfarb, M. & Ornitz, D.M. Genomic organization and embryonic expression of the mouse fibroblast growth factor 9 gene. *Dev. Dyn.* **216**, 72–88 (1999).
20. Wang, Q., Green, R.P., Zhao, G. & Ornitz, D.M. Differential regulation of endochondral bone growth and joint development by FGFR1 and FGFR3 tyrosine kinase domains. *Development* **128**, 3867–3876 (2001).
21. Eswarakumar, V.P., Horowitz, M.C., Locklin, R., Morriss-Kay, G.M. & Lonai, P. A gain-of-function mutation of *Fgfr2c* demonstrates the roles of this receptor variant in osteogenesis. *Proc. Natl. Acad. Sci. USA* **101**, 12555–12560 (2004).
22. Storm, E.E. & Kingsley, D.M. Joint patterning defects caused by single and double mutations in members of the bone morphogenetic protein (BMP) family. *Development* **122**, 3969–3979 (1996).
23. Nalin, A.M., Greenlee, T.K. & Sandell, L.J. Collagen gene expression during development of avian synovial joints: Transient expression of types II and XI collagen genes in the joint capsule. *Dev. Dyn.* **203**, 352–362 (1995).
24. Iseki, S. *et al.* *Fgfr2* and *osteopontin* domains in the developing skull vault are mutually exclusive and can be altered by locally applied FGF2. *Development* **124**, 3375–3384 (1997).

25. Yoshida, T. *et al.* Twist is required for establishment of the mouse coronal suture. *J. Anat.* **206**, 437–444 (2005).
26. Flaumenhaft, R., Moscatelli, D. & Rifkin, D.B. Heparin and heparan sulfate increase the radius of diffusion and action of basic fibroblast growth factor. *J. Cell Biol.* **111**, 1651–1659 (1990).
27. Woo, H.J. & Roux, B. Calculation of absolute protein-ligand binding free energy from computer simulations. *Proc. Natl. Acad. Sci. USA* **102**, 6825–6830 (2005).
28. Schlessinger, J. *et al.* Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell* **6**, 743–750 (2000).
29. Dowd, C.J., Cooney, C.L. & Nugent, M.A. Heparan sulfate mediates bFGF transport through basement membrane by diffusion with rapid reversible binding. *J. Biol. Chem.* **274**, 5236–5244 (1999).
30. Johnson, D., Iseki, S., Wilkie, A.O.M. & Morriss-Kay, G.M. Expression patterns of Twist and *Fgfr1, -2* and *-3* in the developing mouse coronal suture suggest a key role for Twist in suture initiation and biogenesis. *Mech. Dev.* **91**, 341–345 (2000).
31. Tsang, M. & Dawid, I.B. Promotion and attenuation of FGF signaling through the Ras–MAPK pathway. *Sci. STKE* **2004**, pe17 (2004).
32. Mohammadi, M., Olsen, S.K. & Ibrahimi, O.A. Structural basis for fibroblast growth factor receptor activation. *Cytokine Growth Factor Rev.* **16**, 107–137 (2005).
33. Ornitz, D.M. *et al.* Heparin is required for cell-free binding of basic fibroblast growth factor to a soluble receptor and for mitogenesis in whole cells. *Mol. Cell. Biol.* **12**, 240–247 (1992).
34. Venkataraman, G., Shriver, Z., Davis, J.C. & Sasisekharan, R. Fibroblast growth factors 1 and 2 are distinct in oligomerization in the presence of heparin-like glycosaminoglycans. *Proc. Natl. Acad. Sci. USA* **96**, 1892–1897 (1999).
35. Kessel, M. & Gruss, P. Murine developmental control genes. *Science* **249**, 374–379 (1990).
36. Kessel, M. & Gruss, P. Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* **67**, 89–104 (1991).
37. Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W. & Taylor, R.D. Improved protein-ligand docking using GOLD. *Proteins* **52**, 609–623 (2003).
38. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
39. Case, D.A. *et al. AMBER 8* (University of California, San Francisco, 2004).
40. Narumi, T. *et al.* A 55 Tflops simulation of amyloid-forming peptides from yeast prion Sup35 with the special-purpose computer System MD-GRAPE3. <http://doi.acm.org/10.1145/1188455.1188506> (2006).
41. Duan, Y. *et al.* A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999–2012 (2003).
42. Srinivasan, J., Miller, J., Kollma, P.A. & Case, D.A. Continuum solvent studies of the stability of RNA hairpin loops and helices. *J. Biomol. Struct. Dyn.* **16**, 671–682 (1998).
43. Hughes, S.H., Greenhouse, J.J., Petropoulos, C.J. & Sutrave, P. Adaptor plasmids simplify the insertion of foreign DNA into helper-independent retroviral vectors. *J. Virol.* **61**, 3004–3012 (1987).

High-Performance Drug Discovery: Computational Screening by Combining Docking and Molecular Dynamics Simulations

Noriaki Okimoto[1,2☯*], Noriyuki Futatsugi[1,2☯], Hideyoshi Fuji[1,3], Atsushi Suenaga[1☯], Gentaro Morimoto[1,2], Ryoko Yanai[1], Yousuke Ohno[2], Tetsu Narumi[1,4], Makoto Taiji[1,2]

1 High-performance Molecular Simulation Team, Computational Systems Biology Research Group, Advanced Computational Sciences Department, RIKEN Advanced Science Institute, Yokohama, Kanagawa, Japan, **2** High-performance Computing Team, Integrated Simulation of Living Matter Group, Computational Science Research Program, RIKEN, Yokohama, Kanagawa, Japan, **3** Graduate School of Pharmaceutical Sciences, Chiba University, Chiba, Japan, **4** Department of Computer Science, The University of Electro-Communications, Chofu-shi, Tokyo, Japan

Abstract

Virtual compound screening using molecular docking is widely used in the discovery of new lead compounds for drug design. However, this method is not completely reliable and therefore unsatisfactory. In this study, we used massive molecular dynamics simulations of protein-ligand conformations obtained by molecular docking in order to improve the enrichment performance of molecular docking. Our screening approach employed the molecular mechanics/Poisson-Boltzmann and surface area method to estimate the binding free energies. For the top-ranking 1,000 compounds obtained by docking to a target protein, approximately 6,000 molecular dynamics simulations were performed using multiple docking poses in about a week. As a result, the enrichment performance of the top 100 compounds by our approach was improved by 1.6–4.0 times that of the enrichment performance of molecular dockings. This result indicates that the application of molecular dynamics simulations to virtual screening for lead discovery is both effective and practical. However, further optimization of the computational protocols is required for screening various target proteins.

Citation: Okimoto N, Futatsugi N, Fuji H, Suenaga A, Morimoto G, et al. (2009) High-Performance Drug Discovery: Computational Screening by Combining Docking and Molecular Dynamics Simulations. PLoS Comput Biol 5(10): e1000528. doi:10.1371/journal.pcbi.1000528

Editor: David Case, Rutgers University, United States of America

Received May 12, 2009; **Accepted** September 3, 2009; **Published** October 9, 2009

Funding: This work was supported by the contracted research "Protein 3000 Project" of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan and by the Research and Development of the Next-Generation Integrated Simulation of Living Matter, a part of the Development and Use of the Next-Generation Supercomputer Project of MEXT. In addition, this research was partially supported by MEXT through a Grant-in-Aid for Encouragement of Young Scientists (B), NO. 17790096 in 2005–2006. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: okimoto@gsc.riken.jp

☯ These authors contributed equally to this work.

Introduction

In early-phase drug development research, new lead compounds are detected by the computational screening of large compound libraries. Since the goal of computational screening is basically the same as that of experimental screening, i.e., high-throughput screening (HTS), it is expected that the integration and improvement of computational and experimental approaches will increase the productivity of drug discovery. HTS is currently widely adopted and is crucial to the generation of lead compounds. Despite the many successes achieved with HTS [1–5], there remain some problems regarding the cost, complexity of the assay procedure, and screening quality [5–8]. Thus, HTS alone may not improve lead productivity. Hence, computational screening methods, such as ligand- and structure-based screening, have become important. With the advancement of computer performance and calculation techniques, computational screening has become faster and less expensive than HTS. However, the ability of computational screening to enrich hit compounds remains unsatisfactory and less reliable.

Coupled with a rapidly rising number of structures for target proteins, structure-based screening has become prominent in drug discovery. Among the various structure-based computational methodologies adopted for compound screening, the principal one is molecular docking. When the three-dimensional structure of a target protein is available or can be modeled, molecular docking is often used for the screening of compound libraries. Molecular docking predicts the conformation of a protein-ligand complex and calculates the binding affinity. Most docking programs [9–15] involve two operations: "docking" and "scoring." The first involves the generation of multiple protein-ligand conformations, called "poses," or the sampling of the ligand's probable conformations in the binding pocket of the target protein. Most of these programs perform flexible ligand-rigid receptor docking, and some of them are highly capable of predicting poses that resemble the experimental structure for many target proteins [16]. Since such docking programs enable a fast conformational search of ligands in a short time, they are very attractive tools for compound screening. In the second operation, the affinity between the target protein and the ligand for each pose is calculated by

Author Summary

Lead discovery is one of the most important processes in rational drug design. To improve the rate of the detection of lead compounds, various technologies such as high-throughput screening and combinatorial chemistry have been introduced into the pharmaceutical industry. However, since these technologies alone may not improve lead productivity, computational screening has become important. A central method for computational screening is molecular docking. This method generally docks many flexible ligands to a rigid protein and predicts the binding affinity for each ligand in a practical time. However, its ability to detect lead compounds is less reliable. In contrast, molecular dynamics simulations can treat both proteins and ligands in a flexible manner, directly estimate the effect of explicit water molecules, and provide more accurate binding affinity, although their computational costs and times are significantly greater than those of molecular docking. Therefore, we developed a special purpose computer "MDGRAPE-3" for molecular dynamics simulations and applied it to computational screening. In this paper, we report an effective method for computational screening; this method is a combination of molecular docking and massive-scale molecular dynamics simulations. The proposed method showed a higher and more stable enrichment performance than the molecular docking method used alone.

using a scoring function. Then, multiple ligands are ranked according to these calculated binding affinities or docking scores. Many studies using docking programs have shown that these screenings have a higher enrichment of hits than random screening [17,18], but these screenings suffer from false positives and false negatives and are not sufficiently accurate to grade compounds according to the binding affinities [19]. This implies that the compounds with a higher rank include false positives and false negatives; thus, there is a practical difficulty with using docking. The problems of molecular docking as a screening tool have been widely discussed: the scoring functions are inaccurate and neglect the solvent-related terms, and protein flexibility is ignored. Furthermore, the docking score corresponding to binding free energy is less reliable because it is calculated using a single conformation even though the binding free energy is an ensemble property.

Molecular dynamics (MD) simulations can treat both proteins and ligands in a flexible manner, allowing the relaxation of the binding site around the ligand. In addition, they can directly estimate the effect of explicit water molecules. Further, more accurate MD-based computational techniques are available for estimating the binding free energy. These techniques include thermodynamic integration (TI) [20], free energy perturbation (FEP) [20], linear interaction energy (LIE) [21], and molecular mechanics/Poisson-Boltzmann and surface area (MM/PB-SA) [22] methods. The most rigorous computational techniques are the TI and FEP methods, but these are too expensive to be employed in computational screening. The computational cost of LIE is moderate, but it requires information regarding the binding affinities of experimentally known compounds. Hence, we focused on the MM/PB-SA method because many recent investigations have revealed that this method is highly capable of predicting the binding free energy [23]; further, its computational cost is lower than the computational costs of the FEP and TI methods by at least 10-fold, and its broad applicability is suitable for compound screening. In the MM/PB-SA method, the free energy is calculated using the snapshots of solute molecules obtained from explicit-solvent MD simulation. At this time, the explicit-solvent is replaced with implicit models (see Materials and Methods). These MD-based techniques can provide more accurate binding free energy, but their computational costs are considerably high, as compared to molecular docking. Further, the prediction of the optimal structures for protein-ligand complexes adds to the computational cost, even with extended-ensemble MD methods.

To solve the problem of molecular docking and MD simulations, a combination of molecular docking and MD simulations is effective because it can neutralize each other's defects. However, since the application of the MD technique to screening requires the execution of many MD simulations, the problem of the high computational cost of MD simulations remains unresolved. Because of this problem, most of the studies that have used MD-based computational techniques have reported only their ability to rank several ligands according to their experimental binding affinities [23,24]. Further, since the most important parameter for screening is the ability to distinguish true active compounds from a large number of inactive compounds, only a few researches have assessed the ability to enrich active compounds by virtual screening using MD-based computational techniques [25,26].

In order to reduce the significant computational cost of MD simulations, we used a special-purpose computer for MD simulations, "MDGRAPE-3," which functions with a high speed and accuracy [27,28]. In this study, we performed MD simulations of multiple protein-ligand conformations (multiple poses) rather than a single protein-ligand conformation (single pose). The multiple protein-ligand conformations were obtained from the result of molecular docking. Multiple poses were used so that the multiple local energy minima in the ligand's conformational space within the binding pocket could be sampled in the initial structures for MD simulations. Then, we performed massive MD simulations using multiple poses in a practically appropriate time for drug discovery.

In our screening approach, we adopted molecular docking and the MM/PB-SA method as the first and second filters for compound screening; this idea was inspired by the approach adopted by Kuhn and coworkers [25]. They made some important discoveries with respect to MD-based screening for lead generation. Their results showed that the application of the MM/PB-SA method to an energy-minimized complex structure is an adequate and more accurate approach than the calculation of the binding free energy using MD simulation. This is because the use of MD simulations introduces additional structural uncertainties and the free energy from the MD simulations leads to inaccuracy. Further, they reported that the strategy of using multiple poses cannot be recommended in general, and is useful only if the correct binding mode is contained within the higher-scored docking conformations but is not captured with a single pose. Their MD simulations were applied to the protein-ligand complexes for the top 200 compounds obtained by molecular docking, and the MD run for each complex was performed for 200 ps (with a time step of 1.5 fs). They concluded that a more sophisticated MD procedure involving an extended simulation time improved the results, although this time-consuming approach would not be of considerable interest as a tool for lead discovery.

In our study, we attempted to investigate whether a combination of molecular docking and massive-scale MD simulations would be effective in screening compound libraries. Furthermore, we evaluated which protocols for the MM/PB-SA method were effective for compound screening. In the basic MD strategy for our compound screening, a 700-ps MD simulation (with a time step of

0.5 fs) for each complex was performed for the top 1,000 compounds obtained by docking. With regard to the time resolution, simulation time, and number of protein-ligand complex structures, our MD runs were more massive and elaborate than those of previous MD-based screenings [25].

Overview of Our Approach

In our screening approach, we adopted molecular docking and the MM/PB-SA method based on MD simulations as the first and second filters, respectively. First, we performed molecular docking by using the conformations of a target protein and the compounds contained in the compound library. Additionally, the results of molecular docking were applied to the post-processing for the selection of successfully docked compounds and the classification of multiple binding poses (see Materials and Methods). Next, all of the conformations obtained from the molecular docking were energy-minimized using molecular mechanics (MM) force-field (hereafter we call this MM calculations). MD simulations were then applied to multiple conformations of the protein-ligand complexes. The binding free energies were calculated by the MM/PB-SA method using the coordinate sets obtained from the MM calculations and MD simulations. Finally, we assessed the enrichment of active compounds by using ranked lists of compounds graded on the basis of their binding free energies.

Results

To evaluate the ability of the MM/PB-SA method to act as a filter after molecular docking, we performed MD-based compound screening for four target proteins (trypsin, HIV-1 protease (HIV PR), acetylcholine esterase (AChE), and cyclin-dependent kinase 2 (CDK2)). These targets have been widely evaluated in structure-based computer-aided drug design [26,29–34]. For each target protein, we first assessed the enrichment of 12 types of binding free energies (Table 1). These 12 types of binding free energies were classified into four categories. G01–G03 in category 1 were the energies calculated from the MM calculations. The other categories 2–4, which contained the energies calculated from the MD simulations, were classified according to the combination of coordinate sets used for the enthalpy calculations; G04–G06, G07–G09, and G10–G12 belonged to categories 2, 3, and 4, respectively (a detailed explanation is given in the Materials and Methods section.). Analyses of the Receiver Operating Characteristic (ROC) curves [35] are given in Table 2. An ROC curve is closely related to an enrichment curve but is not exactly equivalent to it. This curve describes the tradeoff between sensitivity and specificity. Sensitivity is defined as the ability of the classifier to detect true positives, while specificity is the ability to avoid false positives. The area under an ROC curve, i.e., the ROC value, indicates the quality of enrichment. The ROC value of a random classifier is 0.5, while that of an excellent classifier is greater than 0.9.

Table 2 shows the ROC values for all of the target proteins. From these values, we can observe three common features for three of the target proteins (trypsin, HIV PR, and AChE), excluding CDK2. It is obvious that the ROC values for all of the binding free energies (G01–G12) of multiple poses are higher than those of a single pose, suggesting that docking and its post-processing can sample potentially correct docking poses of active compounds. This implies that the potentially correct binding mode is contained within the top 10 highest-scored docking poses but is not always the highest-scored docking pose. In our study, after docking and post-processing, MD simulations were applied to an average of 5–6 docking poses for each compound in order to increase the efficiency of the sampling of a ligand's conformations. Although MD simulations of multiple poses are expensive, they are necessary for improving enrichment.

The second common feature is that the highest ROC value for each target protein was obtained for the energies calculated from the MD simulations, rather than for those calculated from the MM calculations. This implies that the introduction of protein flexibility and the effect of water molecules facilitated the refinement of the protein-ligand interactions and that the MD-based MM/PB-SA method provided a more reliable binding free energy. These two

Table 1. Computational strategies of 12 binding free energies.

Category	ΔG_{bind}	$H_{COMPLEX}$	$H_{PROTEIN}$	H_{LIGAND}	$TS_{COMPLEX}$	$TS_{PROTEIN}$	TS_{LIGAND}
1	G01	$MM_{COMPLEX}$	$MM_{COMPLEX}$	$MM_{COMPLEX}$	–	–	–
	G02	$MM_{COMPLEX}$	$MM_{COMPLEX}$	MM_{LIGAND}	–	–	–
	G03	$MM_{COMPLEX}$	$MM_{PROTEIN}$	MM_{LIGAND}	–	–	–
2	G04	$MD_{COMPLEX}$	$MD_{COMPLEX}$	$MD_{COMPLEX}$	$TS_{COMPLEX}$	$TS_{COMPLEX}$	TS_{LIGAND}
	G05	$MD_{COMPLEX}$	$MD_{COMPLEX}$	$MD_{COMPLEX}$	$TS_{COMPLEX}$	$TS_{PROTEIN}$	TS_{LIGAND}
	G06	$MD_{COMPLEX}$	$MD_{COMPLEX}$	$MD_{COMPLEX}$	–	–	–
3	G07	$MD_{COMPLEX}$	$MD_{COMPLEX}$	MD_{LIGAND}	$TS_{COMPLEX}$	$TS_{COMPLEX}$	TS_{LIGAND}
	G08	$MD_{COMPLEX}$	$MD_{COMPLEX}$	MD_{LIGAND}	$TS_{COMPLEX}$	$TS_{PROTEIN}$	TS_{LIGAND}
	G09	$MD_{COMPLEX}$	$MD_{COMPLEX}$	MD_{LIGAND}	–	–	–
4	G10	$MD_{COMPLEX}$	$MD_{PROTEIN}$	MD_{LIGAND}	$TS_{COMPLEX}$	$TS_{COMPLEX}$	TS_{LIGAND}
	G11	$MD_{COMPLEX}$	$MD_{PROTEIN}$	MD_{LIGAND}	$TS_{COMPLEX}$	$TS_{PROTEIN}$	TS_{LIGAND}
	G12	$MD_{COMPLEX}$	$MD_{PROTEIN}$	MD_{LIGAND}	–	–	–

We performed MM calculations (MM energy minimization) or MD simulations of a complex, a protein, and a ligand, and evaluated 12 types of binding free energies by combining the respective coordinate sets. The enthalpy contributions of $G_{protein}$ and G_{ligand} in equation 2 were calculated in the following two ways: (1) by using the coordinate sets of a protein (or ligand) obtained from the MD simulations (or MM calculations) of the protein (or ligand) and (2) by using the coordinate sets extracted from the MD simulation of a complex. Similar to the enthalpy contribution, the entropy contribution was also calculated by combining the respective MD coordinate sets. H indicates the sum of $<E_{MM}>$, $<G_{PB}>$, and $<G_{SA}>$ in equation 3, and TS indicates the entropy term in equation 3. $MD_{COMPLEX}$ ($TS_{COMPLEX}$), $MD_{PROTEIN}$ ($TS_{PROTEIN}$), and MD_{LIGAND} (TS_{LIGAND}) denote the use of MD coordinate sets for a complex, protein, and ligand, respectively. Similarly, $MM_{COMPLEX}$, $MM_{PROTEIN}$, and MM_{LIGAND} denote the use of MM coordinate sets for a complex, protein, and ligand, respectively.

doi:10.1371/journal.pcbi.1000528.t001

Table 2. Area under ROC curves.

ΔG_{bind}	Trypsin	HIV PR	AChE	CDK2	CDK2(l)
G01	0.754 (0.318)	0.775 (0.696)	0.655 (0.597)	0.719 (0.685)	0.719 (0.685)
G02	0.651 (0.323)	0.561 (0.550)	0.719 (0.627)	0.595 (0.652)	0.595 (0.652)
G03	0.491 (0.283)	0.538 (0.492)	0.747 (0.554)	0.586 (0.604)	0.586 (0.604)
G04	0.623 (0.321)	0.789 (0.435)	0.527 (0.409)	0.597 (0.599)	0.659 (0.627)
G05	0.539 (0.291)	0.775 (0.425)	0.506 (0.436)	0.565 (0.585)	0.636 (0.614)
G06	0.765 (0.391)	0.979 (0.550)	0.831 (0.603)	0.558 (0.568)	0.624 (0.558)
G07	0.543 (0.300)	0.514 (0.373)	0.509 (0.413)	0.625 (0.623)	0.647 (0.638)
G08	0.486 (0.265)	0.528 (0.383)	0.513 (0.435)	0.586 (0.606)	0.635 (0.626)
G09	0.694 (0.388)	0.778 (0.471)	0.843 (0.645)	0.606 (0.610)	0.622 (0.616)
G10	0.431 (0.369)	0.336 (0.373)	0.654 (0.458)	0.667 (0.615)	0.614 (0.606)
G11	0.422 (0.327)	0.326 (0.389)	0.678 (0.464)	0.653 (0.604)	0.605 (0.602)
G12	0.526 (0.406)	0.516 (0.352)	0.735 (0.541)	0.631 (0.605)	0.612 (0.606)

This table lists the ROC values obtained when the active compounds in the top 1,000 compounds are all considered to be as true positive. The values in parentheses denote the ROC values of a single pose, while the others denote those of multiple poses. The underlining indicates the highest ROC values in the respective categories. CDK2(l) indicates the values of longer MD simulations (1.4 ns) of CDK2.

doi:10.1371/journal.pcbi.1000528.t002

common features were clearly seen in the results for the active compounds. A typical successful example of MD simulations using multiple poses is shown in Figure 1. In the crystal structure of trypsin complexed with an inhibitor [36], the amidine group of the inhibitor (Figures 1 and S1; active compound (13) of trypsin) formed hydrogen bonds with the important residue Asp-180 in the binding pocket. Further, the highest-scored docking pose was so inaccurate that no important interactions were observed at all, but the 7th ranked docking pose was similar to that of the crystal structure. In addition, the application of the MD simulation to the 7th ranked docking pose appropriately improved the key hydrogen bonds and the position of the naphthalene group and G06 value of the 7th ranked docking pose was the lowest in all the poses.

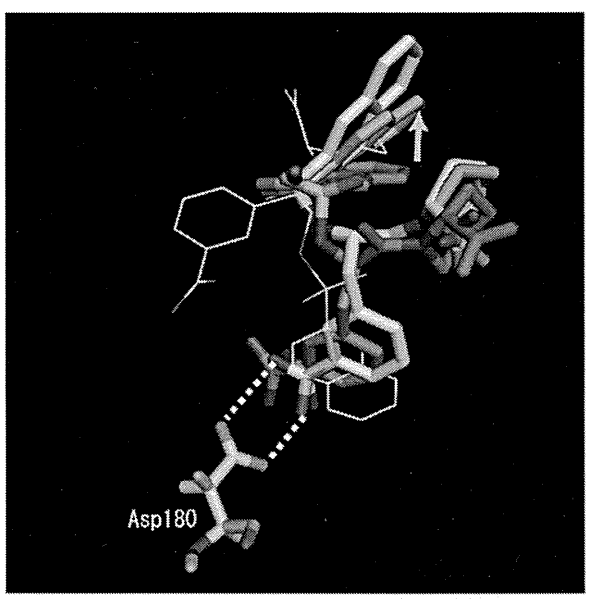

Figure 1. Successful example of MD simulations using multiple poses. The color codes for the stick models are as follows: yellow, conformation of the inhibitor in the crystal structure; purple, 7th-ranked docking pose; and pink, conformation of the 7th-ranked docking pose after the MD simulation. In addition, the highest-scored docking pose is shown by the wireframe model. The highest-scored docking pose (wireframe model) is inaccurate, but the 7th-ranked docking pose is similar to that of the crystal structure. MD simulation of the 7th-ranked pose improved the key hydrogen bonds and the position of the naphthalene group and G06 value of the 7th ranked docking pose was the lowest in all the poses.

doi:10.1371/journal.pcbi.1000528.g001

The last common feature was that the binding free energies with no entropy terms (i.e., G06, G09, and G12), which were obtained by using the trajectories of the MD simulations, showed the highest ROC values in the respective energy categories (2–4). Thus, the introduction of entropy terms tended to reduce enrichment. This is probably due to the difficulty of computing entropy values for the MM-PB/SA energy function. We will further discuss this problem in the Discussion section. Our MD simulations encouraged conformational relaxation, and the binding enthalpy from the MM-PB/SA method could satisfactorily increase the enrichment performance. However, the treatment of binding entropy terms involves certain unsolved problems. Here, we performed a statistical analysis using data on the ROC values to evaluate the differences between key classifiers, G01 (multiple poses), G06 (multiple poses), and molecular docking (Table 3). The program DBM MRMC version 2.1 was used in this analysis [37–41]. From this analysis, it was obvious that the differences in the ROC values between G06 and docking, and those between G06 and rescoring (docking), were statistically significant for trypsin and HIV PR, but the difference in the ROC values between G06 and docking for AChE was not statistically significant. On the other hand, the differences in the ROC values between G01 and docking were not statistically significant for trypsin, HIV PR, and AChE. An examination of the entire data set indicated that the binding free energies of multiple poses, especially G06, which was obtained from the MD trajectories of just the protein-ligand complexes with no entropies, showed a high and stable ability to enrich the active compounds.

This paper provides a detailed account of the ability of our approach to discriminate active compounds from inactive ones. Figure 2 shows the ROC curves for the respective target proteins.

Table 3. Statistical comparisons of ROC values.

	Trypsin	HIV PR	AChE	CDK2
G01#	0.754 (0.010)	0.775(0.054)	0.655(0.117)	0.719 (0.064)
G06#	0.765 (0.092)	0.979(0.004)	0.831(0.135)	0.558 (0.065)
GOLD/Rescore#	0.476(0.099)/0.477(0.081)	0.621(0.102)/0.636(0.117)	0.614(0.159)/0.268(0.098)	0.679 (0.067)/0.530 (0.038)
P values(G01/Docking)	0.054	0.542	0.922	0.915
p values(G01/Rescore)	0.055	0.320	0.019*	0.102
p values(G06/Docking)	0.045*	0.039*	0.330	-
p values(G06/Rescore)	0.046*	0.014*	0.001*	0.712

This analysis was analyzed by the program DBM MRMC 2.1 [37–41]. This program uses a jackknife method [37] to assess the statistical significance of the observed difference between two classifiers. The p-value (G06/Docking) in CDK2 is not shown because the ROC value of G06 was less than that of Docking. This analysis indicated that the differences between G01 and G06 were not statistically significant (data not shown). The analyses were performed using results of multiple poses. "Rescore" indicate the result of rescoring approach (docking) (see Materials and Methods).

#These values are ROC values. Standard errors are shown in parentheses.

*Differences are considered statistically significant at $p<0.05$.

doi:10.1371/journal.pcbi.1000528.t003

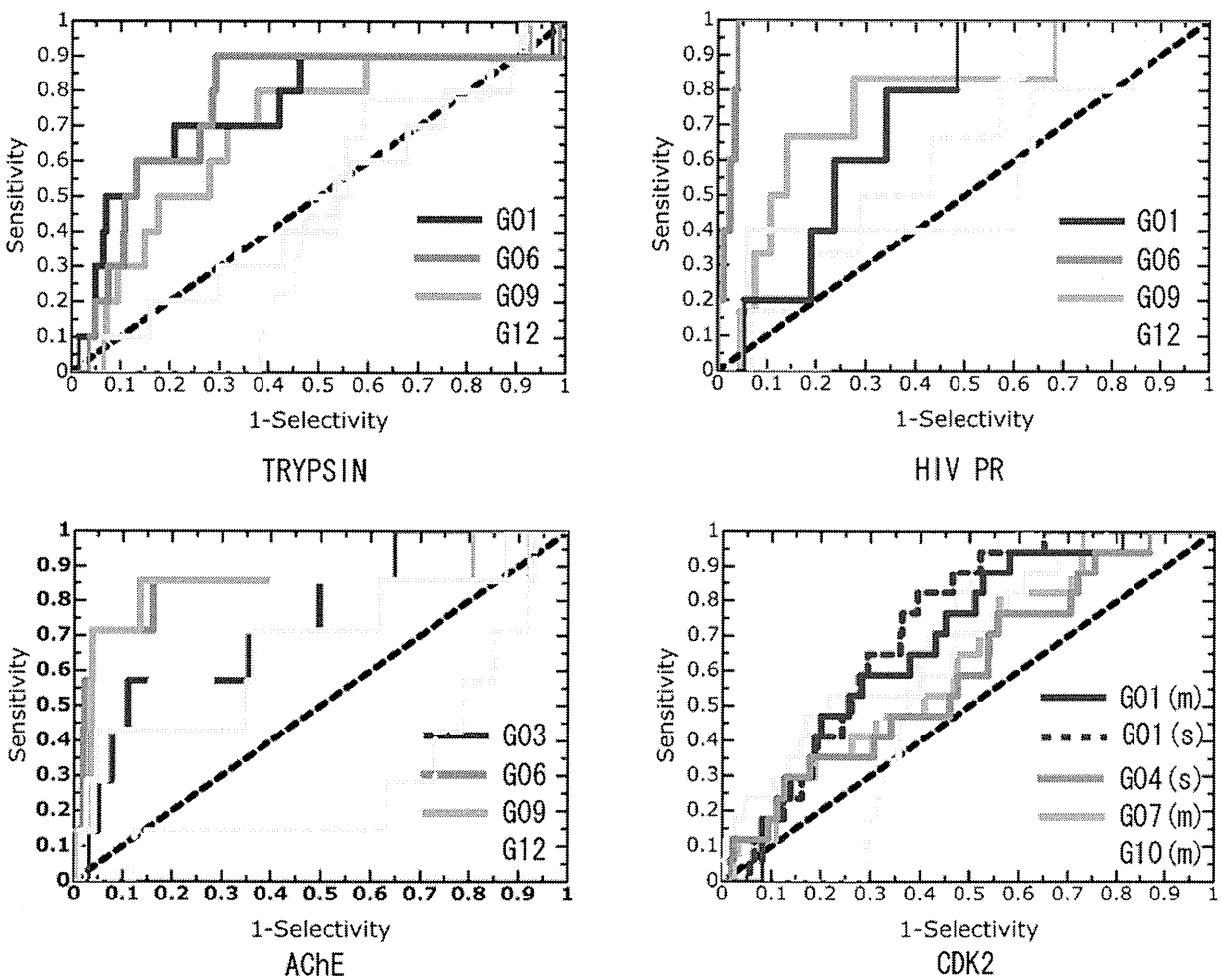

Figure 2. ROC curves for the four target proteins. Each graph shows the sensitivity versus 1-specificity. These indicate the ROC curves obtained when the active compounds in the top 1,000 compounds are all considered to be as true positive. The black dashed, sky blue solid, and sky blue dashed lines indicate random screening, molecular docking and rescoring (docking), respectively. For trypsin, HIV PR, and AChE, the ROC curves of the binding free energies (multiple poses) with the highest ROC values in the respective categories are shown. For CDK2, the curves of G01 (single and multiple poses), G04 (single pose), G07 (multiple poses), and G10 (multiple poses) are shown. (s) and (m) indicate single pose and multiple poses.

doi:10.1371/journal.pcbi.1000528.g002

The ROC curves of the binding free energies, with the highest ROC values in the respective categories, were observed for trypsin, HIV PR, and AChE. Table 4 shows the information on the enrichment factors to allow the abilities of the classifiers to be understood clearly.

For trypsin, 10 active compounds (out of 21) were ranked in the top 1,000 compounds (Figure S1). No significant difference was observed in the results of molecular docking and random screening for these top 1,000 compounds. Considerable improvement was observed in the results of the MM calculations and MD simulations. G01 and G06, in particular, showed high enrichment performances, and the enrichment factors for the top 100 compounds were 5.00 and 4.00, respectively (see Table 4). Furthermore, G06 detected no less than nine active compounds in the top 300.

For HIV PR, 6 active compounds (out of 8) were ranked in the top 1,000 compounds (Figure S2). A slightly better enrichment was achieved by docking than by random screening. As seen in the curves, we found drastically improved enrichment by G06. It detected 6 active compounds in the top 100 compounds, and the enrichment factor for the top 100 was 10.0 (Table 4).

For AChE, 7 active compounds (out of 14) were detected in the top 1,000 compounds (Figure S3). We found that the enrichments of G06 and G09 were considerably better than that of molecular docking, although the difference in the ROC values between G06 (or G09) and docking was not statistically significant. Because there was only a slight difference between G06 and G09, both of them detected five active compounds in the top 100.

For CDK2, the ROC curves of the following representative binding free energies were drawn: G01 (single and multiple poses), G04 (single pose), G07 (multiple poses), and G10 (multiple poses). Seventeen active compounds (out of 26) were ranked in the top-scoring 1,000 compounds (Figure S4). The G01 of single and multiple poses as obtained from the MM calculations showed higher enrichment than random screening; however, there was no statistically significant difference between the results of G01 and molecular docking (see Table 3). The G01 of single and multiple poses detected 10 active compounds in the top 300 and showed only slightly higher enrichment factors than molecular docking (Table 4). In contrast, G04, G07, and G10, which were obtained from the MD simulations, remained unchanged or worsened as compared to docking, although they identified 6 or 7 active compounds in the top 200. Over all, the enrichments for CDK2 were not at all improved as mentioned above.

The ROC values for CDK2 showed a different tendency as compared to those for the other three proteins (Table 2). Among the 12 types of energies, the G01 of single and multiple poses showed the highest values (0.685 and 0.719, respectively), which implies that the enrichments of the MM calculations were better than those of the MD simulations. Moreover, for 7 types of energies (out of 12), the single pose results showed higher enrichment than those of multiple poses. In addition, the binding free energies with entropy terms showed slightly high enrichment performances in the respective categories (2–4), which were calculated from MD simulations. In particular, G04, G07, and G10, which included the binding entropy effects of the ligands, showed the highest ROC values in their respective categories.

We monitored the mobility of ligand molecules in MD simulations of CDK2. Figure 3 shows the cumulative percentages of positional displacements of ligand molecules relative to each protein between the docking and final MD structures. From this figure, it is clear that the docked ligands for CDK2 did not move very much in the MD simulations, as compared to the other target proteins, which implies that the protein-ligand interactions were not fully relaxed. Such insufficiency in conformational relaxation/refinement directly influences protein-ligand interactions. Partic-

Table 4. Enrichment factors for top 1,000 compounds.

	Trypsin	HIV PR	AChE	CDK2
EF (10%)	5.00 (G01)	3.33 (G01)	4.29 (G03)	1.76 (G01)
	4.00 (G06)	10.0 (G06)	7.14 (G06)	1.76 (G04)
	3.00 (G09)	3.33 (G09)	7.14 (G09)	1.43 (G07)
	0.00 (G12)	0.00 (G12)	4.29 (G12)	2.14 (G10)
	1.00 (Docking)	5.00 (Docking)	4.29 (Docking)	2.35 (Docking)
	0.00 (Rescoring)	1.67(Rescoring)	0.00 (Rescoring)	0.00 (Rescoring)
EF (30%)	2.33 (G01)	2.22 (G01)	1.90 (G03)	1.96 (G01)
	3.00 (G06)	3.33 (G06)	2.85 (G06)	1.17 (G04)
	2.00 (G09)	2.78 (G09)	2.85 (G09)	1.67 (G07)
	0.66 (G12)	0.55 (G12)	2.38 (G12)	1.90 (G10)
	1.00 (Docking)	2.78 (Docking)	1.43 (Docking)	1.76 (Docking)
	0.33 (Rescoring)	1.67 (Rescoring)	0.48 (Rescoring)	0.19 (Rescoring)

The enrichment factor (EF) can be defined as:

$$EF = (a/n)/(A/N)$$

where a is the number of active compounds in the n top-ranked compounds and A is the number of total N compounds. In this table, the n for EF (10%), n for EF (30%), and N were 100, 300, and 1,000, respectively. For the respective proteins, A, n for EF (10%), and n for EF (30%) indicate the numbers of active compounds in the top 1,000, top 100, and top 300, respectively. All of the EF values calculated from the result of screening using multiple poses. "Rescoring" indicate the result of rescoring approach (docking) (see Materials and Methods).

doi:10.1371/journal.pcbi.1000528.t004

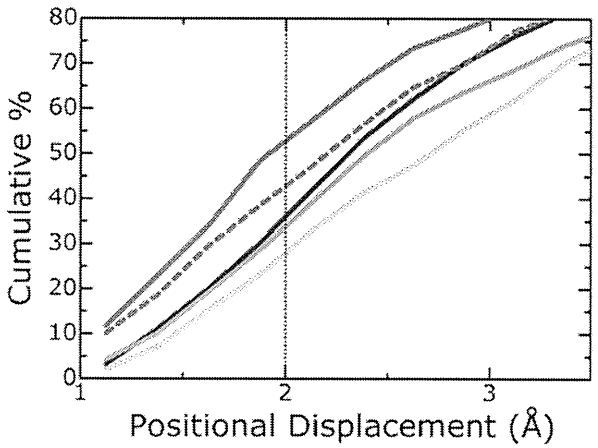

Figure 3. Mobility of ligand molecules in MD simulations. This graph indicates cumulative percentage graph of positional displacements of ligand molecules relative to the respective proteins between docking structures and final MD structures. The blue, green, orange, and red solid lines indicate cumulative curves for trypsin, HIV PR, AChE, and CDK2, respectively. The broken line indicates a curve for longer MD simulations (1.4 ns) of CDK2. From the red solid line, it is clear that the docked ligands for CDK2 did not move in the MD simulations since the positional displacements of approximately 50% of the ligand conformations were less than 2.0 Å.
doi:10.1371/journal.pcbi.1000528.g003

ularly for active compounds, the binding modes obtained by MD simulations were some different from those of experimental structures (refer to Figures S5 and S6). These data suggested that the use of MD simulations for CDK2 led to structural uncertainties for active compounds. In addition, the interactions of the inactive (decoy) compounds would not be refined fully in MD simulations. We think that such low mobility for ligand molecules and the improper conformational dynamics are due to an improper MD setup. This would be the reason why the enrichments of the MD simulations using multiple poses were worse than those of the MM calculations for CDK2.

Discussion

We evaluated the ability to enrich active compounds for four target proteins: trypsin, HIV PR, AChE, and CDK2. Our screening approach could improve the molecular docking results for all of the proteins except CDK2. For trypsin, HIV PR, and AChE, our results indicated that the use of multiple poses improved the enrichments of all the MM calculations and MD simulations. In addition, the binding free energies calculated from the MD simulations showed higher and more stable enrichments than those of the docking and MM calculations. In particular, the G06 using multiple poses was considered to be effective. This energy contained no entropy components. Further, the enthalpy components were calculated using the coordinate sets extracted from the MD simulation of a complex.

Kuhn and coworkers [25] reported that for the MM/PB-SA values of MM calculations, the strategy of using multiple poses could only show a high enrichment when the correct binding mode was contained within the higher-scored docking conformations, but was not captured with a single pose. In our study, we carefully selected multiple docking poses by the post-processing of docking results and used an average of five to six docking poses for each compound. As a result, the correct binding modes or potentially correct modes that could be refined by the MD simulations were sampled within the selected multiple docking poses, which did not often correspond to the top-scored pose. Therefore, the results using multiple poses showed a higher enrichment than those obtained using a single pose.

In addition, Kuhn et al. [25] showed that the use of MD simulations often leads to structural uncertainties and an inaccurate estimation of the binding free energy. The MM/PB-SA energies of the MM calculations and MD simulations in their study corresponded to G01 and G04 in our study. A comparison between G01 and G04 indicated that the enrichment of G04 was lower than that of G01, which is consistent with the results of Kuhn and coworkers [25], although there were large differences in the MD setup, MM/PB-SA setup, and target proteins. G01 contained only the enthalpy components that were calculated using the coordinate sets derived from the MM calculation of a complex. G04 contained the binding entropy effect of the ligand. Further, the enthalpy components were calculated using the coordinate sets from the MD calculation of a complex. The difference between G04 and G06 was the presence of the entropy effect. Therefore, we consider that for trypsin, HIV PR, and AChE, the structural refinement/relaxation by longer and higher time resolution MD simulations and the relatively accurate estimation of binding free energy (enthalpy) by the MM/PB-SA method led to increased enrichment, but the introduction of the entropy values induced an uncertainty in the binding free energies. On the other hand, we think that the use of MD simulations for CDK2 led to structural uncertainties and then an inaccurate estimation of the binding free energy (Figures S5 and S6). This would be due to an improper MD setup, as Kuhn and co-workers suggested in their paper [25].

Basically, it is well-known that it is difficult to calculate entropy values properly. In our work, the entropy values were calculated by principal component analysis (PCA). These values are sensitive to the data sampling frequency [42,43] and are likely to be overestimated [44]. Therefore, we believe that the entropy values were slightly unstable and not completely reliable. An alternative computational method is normal mode analysis. This may be stable to some extent, but it is known that conformations at different local energy minima provide rather similar entropy values even though there are differences in the finite temperature [42]. Moreover, the computational cost is significantly high to use for the calculation of many structures. Thus, even if we were to use normal mode analysis, the entropy values would induce an uncertainty in the binding free energies. Therefore, in order to achieve further improved enrichment, it is necessary to improve the calculations for the entropy terms.

Our strategy could not significantly improve the molecular docking results for CDK2. It is well known that, as compared to the binding pockets of the other three proteins, the binding pocket of CDK2 is more flexible and hydrophobic. We compared the binding pockets in two different X-ray crystal structures of CDK2 [45,46] (Figure 4). This figure indicates that the shape of the binding pocket is very flexible and that the hydrophobic region covers the surface of the binding pocket. In addition, a study on molecular docking using different CDK2 crystal structures reported that the volume (flexibility) of the binding site is a key factor for predicting docking poses [29]. Although only one CDK2 structure was used in this study, we applied MD simulations to protein-ligand structures obtained from molecular docking to facilitate the relaxation of protein-ligand interactions. Unfortunately, our MD simulations were insufficient to relax the protein-ligand conformations in the binding pockets (see Figure 3). Such insufficiency is believed to be due to the MD setup. To improve

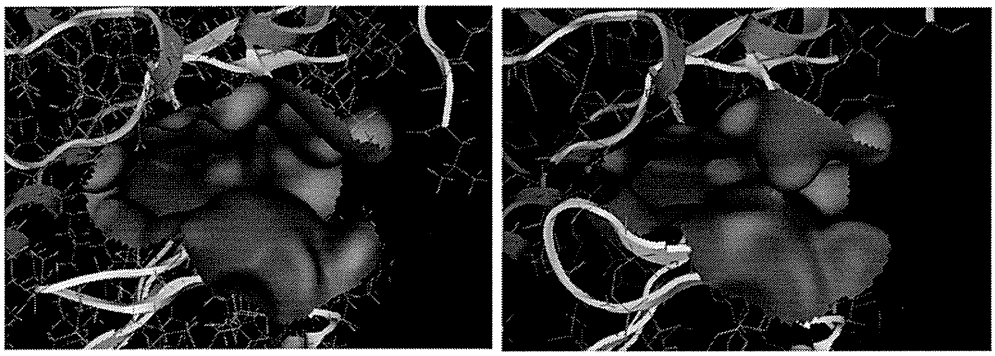

Figure 4. Binding pockets in two different X-ray crystal structures for CDK2. The left figure indicates the binding pocket in the CDK2-Oxindole inhibitor complex (PDB Id: 1FVV). This structure was used in our study. The right figure indicates the binding pocket in the CDK2-NU2058 inhibitor complex (PDB Id: 1H1P). The hydrophobic regions in the binding pocket are drawn in blue, and the hydrophilic regions are drawn in red. From this figure, it is clear that large hydrophobic regions were located in both binding pockets. In contrast, the shapes of the two binding pockets were considerably different, which means that the binding pocket of CDK2 was flexible.
doi:10.1371/journal.pcbi.1000528.g004

the insufficient relaxation, we applied the MD simulations for 1.4 ns to each configuration; this simulation time was twice that of the initial MD simulation time. These MD simulations effected some relaxation/refinement of the ligand conformations (Figure 3); the enrichments of the multiple poses were found to be higher than those of the single pose (Table 2). Despite this, G04–G09 showed only small improvements in the enrichment performance. These results suggest that further improvement of the MD setup was necessary. To obtain information about how to improve the MD setup, we attempted to maximize the ROC values by using an approach based on the linear response (LR) [47] and MM/PB-SA methods (LR-MM/PB-SA approach [48]) The LR-MM/PB-SA equation was derived from equations 2–4 (see Materials and Methods):

$$\Delta G_{bind} = a\Delta E_{\mathrm{int}} + b\Delta E_{ele} + c\Delta E_{vdW} + d\Delta G_{PB} + e\Delta G_{SA} - fT\Delta S(1)$$

where a, b, c, d, e, and f are weighting factors ranging from 0.5 to 1.5. The terms on the right side of equation 1 represent the energy difference between the complex and protein plus ligand. This approach is usually used for estimating the binding affinity by combining an empirical MM/PB-SA energy calculation with an LR optimization of coefficients against the experimental binding affinities of several compounds. The optimized free energy model is used for interpreting the binding model and predicting the binding affinity of unknown molecules. In our study, we optimized the weighting factors to maximize the ROC value, that is, the enrichment performance, using a genetic algorithm (GA). We applied the LR-MM/PB-SA approach to the G10 of multiple poses obtained from the initial MD simulations, because G10 showed the highest ROC value among those of the binding free energies calculated from the MD simulations (Table 2). As a result, when the weighting factors of a–f were 1.12, 0.91, 1.47, 1.01, 0.87, and 1.49, respectively, a maximum ROC value of 0.812 was obtained (Figure 5). This result suggested that the LR-MM/PB-SA approach was effective at improving the enrichment performance, and these weighting factors indicated an improvement plan for the MD setup. The weighting factor of the entropy term, 1.49, would contribute to dilute the percentage of inactive (decoy) compounds. This would be related to the fact that the ligands in the binding pocket of CDK2 could not move largely (Figure 3). In addition, it is conceivable that the weighting factor of ΔE_{vdW}, i.e., 1.47, enriched the active compounds because they include a hydrophobic region and formed comparatively strong hydrophobic interactions with the binding pocket (See Figure S4). As the binding modes obtained through MD simulations were some different from the experimentally observed binding modes, the conformational refinement was considered insufficient or improper to accurately predict the binding free energy. This information also suggests that fully conformational relaxation/refinement is required to improve the enrichment performance. In this study, the ligand, water molecules, and protein residues around the binding pocket were allowed to move, but other protein residues were restrained to the X-ray structure in all of the MM calculations and MD simulations (a detailed explanation is given in the Materials and Methods section). Hence, to achieve the

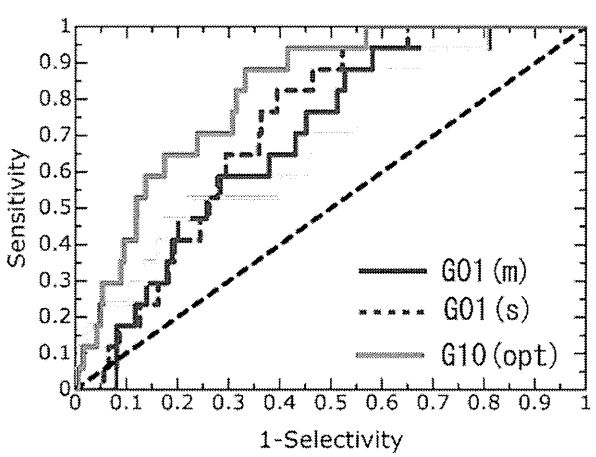

Figure 5. ROC curves of the LR-MM/PB-SA approach for CDK2. This graph shows the sensitivity versus 1-specificity. This indicates the ROC curves obtained when the active compounds in the top 1,000 compounds are all considered to be as true positive. The black dashed, sky blue solid, and sky blue dashed lines indicate random screening, molecular docking and rescoring (docking), respectively. The G01(m), G01(s), molecular docking and rescoring(docking) are the same as in Figure 2. G10 (opt) is the ROC curve of G10 (multiple poses) obtained by the LR-MM/PB-SA approach. The ROC value is 0.812.
doi:10.1371/journal.pcbi.1000528.g005

conformational relaxation of the binding pocket of CDK2, allowing wider protein residues to move in MD simulations and longer MD simulations are required. The former is an especially important parameter for improving the enrichment performance, although it would increase the computational cost. In our next study, which will focus on the extent of the mobility of protein residues, along with simulation time and force-field parameters for organic small molecules, we will attempt to optimize the MD setup using a recent widely used dataset of decoy compounds [49].

The computational screening of large compound libraries involves the use of hierarchical multiple filters, such as ligand- and structure-based approaches. Molecular docking plays the primary role in these filters. With advancements in computer performance and computational chemistry, docking programs have become more accurate, but their ability to enrich hit compounds remains unsatisfactory. In order to improve the enrichment performance of molecular docking, we attempted to use the MM/PB-SA method [50] as a post-molecular docking filter. The basis of our approach was to perform massive MD simulations of protein-ligand conformations obtained from molecular docking, aim at the refinement/relaxation of protein-ligand conformations after docking, and predict more accurate binding free energies using the MM/PB-SA method in a practical time for lead discovery. Combining molecular docking and MD simulations basically allows each of them to neutralize the other's defects, but certain problems remain even with MD simulations, particularly with regard to compound screening applications. The major drawback of MD simulations is insufficient sampling due to the significant computational cost involved. To solve this problem, we performed MD simulations using various docking conformations obtained by molecular docking. However, the computational cost of this technique was approximately five to six times that of MD simulations using single docking conformations, such as the top-scored docking conformation. The enormous computational time needed for MD simulations is a serious problem. Here, we solved this problem by accelerating most of the time-consuming operations of the MD simulation using a high-performance special-purpose computer for MD simulations, "MDGRAPE-3" [27,28]. Accordingly, our approach could be performed in a practical time (about a week) for lead discovery. The evaluation in this study provides valuable information on in-silico drug design. Further, a more rigorous MD-based filter is under consideration for further improving the enrichment performance. This technique will also be applied to the lead optimization stage of drug development research.

In conclusion, our approach could improve the enrichment of virtual screening by molecular docking. Among the 12 types of binding free energies, G06, which was obtained from the MD simulations using multiple poses, showed the highest and most stable ability to enrich the active compounds. The strategy of multiple poses can be used to sample the potentially correct poses of active compounds; thus, it increases the enrichment performance. Since the G06 enrichment factors for the top 100 compounds ranged from 4 to 10 (see Table 4), which indicates approximately 1.6–4.0 times higher values than the enrichment performance of molecular docking, with the exception of CDK2, it is obvious that a stable and high enrichment can be achieved after molecular docking. In addition, G06 is suitable for compound screening because its computational cost is the least among those of the other MM/PB-SA energies obtained from the MD simulations. We also confirmed that G01, which was obtained from the MM calculations, showed good enrichment ability despite its low computational cost. This result agreed with that of the previous study [25]. The ability of G01 to enrich active compounds was lower and less stable than that of G06, but we believe that G01 acted as an effective filter between molecular docking and the MD-based MM/PB-SA method. From this study, we conclude that the application of MD simulations to virtual screening for lead discovery is effective and practical, but that further optimization of the MD simulation protocols is required for the screening of various target proteins, including kinases.

Materials and Methods

Preparation of the Target Protein

We applied our approach to four target proteins: trypsin, HIV PR, AChE, and CDK2. These structures with crystallographic resolutions of less than 3.0 Å, were retrieved from the Protein Data Bank (PDB) because the conformations of residues in the binding pocket affect the molecular docking results (PDB Id: 1C5S (trypsin) [51], 1HWR (HIV PR) [52], 1E66 (AChE) [53], and 1FVV (CDK2) [46]). All of the bound crystal water molecules, ligands, and other organic compounds were removed from each protein. Hydrogen atoms were added, and energy minimizations on the hydrogen atoms were performed using the Molecular Operating Environment (MOE) program (Chemical Computing Group Inc. [54]).

Seeded Compound Library for Docking

For each target protein, we prepared a test set of compounds that included 10,000 randomly selected compounds, or decoys, from the Maybridge library of compounds and experimentally known active compounds. It was confirmed that 95.5% of the selected decoy compounds obeyed the Lipinski rule of 5 [55]. The active compounds, which had binding affinities (K_i, K_d, or IC_{50}) below 30 μm, were selected from the PDBbind database [56,57] and by referring to the literatures [26,58]. Most of the active compounds also obeyed the Lipinski rule of 5. The numbers of active compounds selected for each of the respective target proteins was as follows: 21 (trypsin), 8 (HIV PR), 14 (AChE), and 26 (CDK2) (see Figure S1, S2, S3, S4). For each compound of the test set, a 3D conformation was generated, ionized, and energy minimized using LigPrep (Schrödinger Inc. [59]), assuming a pH of 7.0.

Docking

Molecular dockings were performed using the Genetic Optimisation of Ligand Docking (GOLD) version 3.1 [9,10]. This program employs a GA to explore the possible binding modes. The standard default settings for the GA parameters were used. The binding site radius was 12 Å. We performed the docking run three or four times using the GoldScore or ChemScore function for each target protein and selected the result that showed the best enrichment. GoldScore (default settings) was used as the scoring function for trypsin and HIV PR. In contrast, ChemScore (default settings) was used for AChE and CDK2 because docking runs using GoldScore can detect few of the successfully docked active compounds for AChE and CDK2. For AChE alone, the torsional rotations of Phe-330 (chi1 and chi2) were treated as flexible in the docking process. For each docking run, the 10 highest-scoring docking poses were saved to obtain a variety of binding modes.

Post-processing of the Docking Results

First, among the 10 highest-scoring docking poses saved for each compound, those in which the compound did not occupy the binding pocket or did not interact with the important residues were removed. The latter was used only for trypsin and HIV PR. The important residues were Asp180 for trypsin and Asp24 in

each monomer for HIV PR. These treatments had the effect of reducing the false positives for molecular docking. The docked compounds were then arranged in descending order from the highest score with respect to the multiple docking poses, and the top 1,000 compounds were selected from the test set. Finally, for the top 1,000 compounds, the docking poses of each compound were clustered using the root mean square deviation of 0.9 Å (complete link method [60]). After post-processing, approximately 6,000 docking poses were selected for the 1,000 compounds, which were then used as the initial conformations for MD simulations. Some active compounds were not ranked in the top 1,000. The numbers of active compounds in the top-scoring 1,000 were 10, 6, 7, and 17 for trypsin, HIV PR, AChE, and CDK2, respectively. In addition, the compounds in the top-scoring 1,000 were rescored with ChemScore (trypsin and HIV PR) or GoldScore (AChE and CDK2) because it is known that the rescoring approach increases the enrichment performance [61]. Furthermore, we analyzed ROC curves using molecular weight as classifier (Figure S7). From statistical analysis, it is obvious that the differences in the ROC values between G06 and molecular weight were statistically significant for trypsin, HIV PR, AChE.

MD Simulation Protocols

We performed MD simulations of each complex (ligand-bound protein), protein, and ligand to obtain various types of binding free energies (see the following subsection). The active sites of the protein-ligand complexes were immersed in an approximately 28–30 Å sphere of transferable intermolecular potential 3 point (TIP3P) water [62] molecules. The radius of the water droplet was selected such that the distance of the atoms of all the docked compounds from the water wall was greater than 15 Å (see Figure 6). The total number of atoms in the respective systems was approximately 8,000–12,000. On the solvent boundary, a half-harmonic potential (1.5 kcal/mol-Å^2) was applied to prevent the evaporation of the water molecules. The ligand, water molecules, and protein residues that were approximately 12 Å of the active center were allowed to move, but other protein residues were restrained to the X-ray structure by the harmonic energy term (1.5 kcal/mol-Å^2) in all of the MM calculations, namely the MM energy-minimization, and MD simulations. For the simulations of the ligands, each ligand was immersed in a water droplet, and this structure was used as the initial structure for the MD simulation of the ligand. In addition, the simulation of each protein (trypsin, HIV PR, AChE, and CDK2) was performed in the same manner as that of the complex.

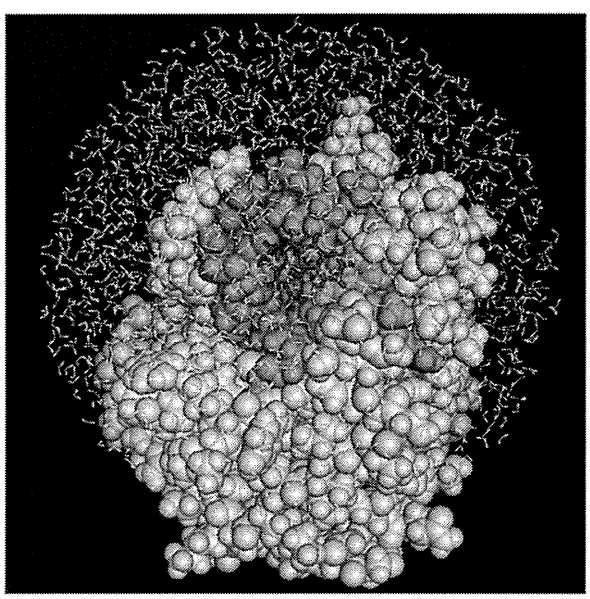

Figure 6. System for MD simulation of trypsin. The protein is shown by the space-filled model, and the ligand is colored blue. The peripheral residues around the active center (red region), a ligand, and water molecules were allowed to move in the MD simulation. The protein residues (grey) were restrained to the X-ray structure by a harmonic energy term. Similar systems were used for the other target proteins.
doi:10.1371/journal.pcbi.1000528.g006

All of the simulations were performed using AMBER 8.0 [63] modified for MDGRAPE-3 [27,28]. The ff03 force field [64] was adopted, and the time step was set at 0.5 fs. To carefully consider the motion of hydrogen atoms in the interactions between the ligands and protein residues, no bond length constraint was applied to solute atoms. The temperature of each system was gradually increased to 300 K during the first 25 ps, and additional MD simulations were performed for 700 ps for equilibration. The temperature was maintained at 300 K by using the method described by Berendsen et al. [65], and the system was coupled to a temperature bath with coupling constants of 0.2 ps. The parameters and charges for the ligands were determined using the *antechamber* module version 1.27 of AMBER 8.0 [63] by utilizing the general atom force field (GAFF) [66] and the AM1-BCC charge method [67,68]. Although the computational cost of the AM1-BCC charge method is low, a some difference between the charge and that of ff03 was noticeable. Since the original GAFF parameters were insufficient to cope with the parameters of all the ligands, we filled the missing parameters on the basis of the information on regarding atom types, bonds, valences, angles, and dihedrals by using an in-house program (see Text S1). (Note: these parameters for proteins and small organic molecules are very important to calculate the binding free energies between proteins and ligands)

Our MDGRAPE-3 system is a cluster of personal computers, each equipped with two MDGRAPE-3 boards. Each board contains 12 MDGRAPE-3 chips and has a peak speed of approximately 2 Tflops. The computations of non-bonded forces and energies for MD simulations were accelerated by MDGRAPE-3, and the other calculations were performed by the host central processing unit (CPU). In this study, we used 50 host computers equipped with 100 MDGRAPE-3 boards. The calculations for an MD simulation and the estimation of the binding free energies by the MM/PB-SA method were performed simultaneously. The average computational time for a single protein-ligand complex was 2.5 h, and the computations for approximately 6,000 protein-ligand conformations obtained by docking for each protein were completed in a week. The total simulation time for each protein was 4 μs, which corresponded to an 8-μs MD simulation with a standard time step of 1 fs. A single MD simulation for the system (Figure 6), without using MDGRAPE-3, requires more than 10 times the abovementioned computational time. Thus, in the current state, it would be quite difficult to use our screening approach without the MDGRAPE-3 system in a practically appropriate time for lead discovery. Therefore, our study can provide important information for MD-based screening.

Calculation of Binding Free Energy by the MM/PB-SA Method

The production MD trajectory was collected for the last period of 210 ps. In the calculation of the binding free energies by the