

of five strong lines. Two of these are resonance transitions, and if such lines were missing from the theoretical spectrum, an assignment of this region would be difficult. Note that the theoretical and experimental intensities of the discussed lines match well except for the peak at  $-8.53\text{ cm}^{-1}$ , which appears weaker in the observed spectrum. However, the discrepancy is only apparent because the experimental line is clearly broader, indicating a relatively wide resonance state. If the theoretical line were appropriately broadened, the heights would agree much better, but as noted above, a fixed broadening function was used for all theoretical lines, including resonance ones. (In principle, theory can provide these resonance widths, but this application is beyond the scope of the present work.)

Another example of how theory elucidated the spectra is manifested in the analysis of two wide observed peaks at  $-0.52$  and  $+0.14\text{ cm}^{-1}$ , and of a very strong one at  $9.61\text{ cm}^{-1}$ , the shapes of which could suggest superpositions of several weak lines. With theoretical insight, we can attribute this shape to their resonance character. On the other hand, there may be uncertainty in theoretical approach about whether a given level above the dissociation limit is a resonance or continuum state, and the presence of broadening in the experiment confirms their resonance character. Such

experimentally confirmed resonances are marked with daggers in table S2.

We have assigned an experimental infrared spectrum of the *ortho*-H<sub>2</sub>-CO complex in the region of the CO fundamental band. Also, theory allowed us to better understand some subtle aspects of the experimental spectrum. On the other hand, the experimental precision of about  $0.001\text{ cm}^{-1}$  provides a challenging test to evaluate the quality of theory and the adequacy of various approximations. For example, our work provides further confirmation that surfaces averaged over intramonomer motion are sufficiently accurate for applications such as the present one. The assigned spectrum and the calculated potential surface can now be used to interpret astronomical observations and to enable precise millimeter-wave measurements that need infrared guidance because of their limited frequency range scanning capabilities.

#### References and Notes

1. A. R. W. McKellar, *J. Chem. Phys.* **108**, 1811 (1998).
2. S. Green, P. Thaddeus, *Astrophys. J.* **205**, 766 (1976).
3. M. Wernli *et al.*, *Astron. Astrophys.* **446**, 367 (2006).
4. I. Pak *et al.*, *Chem. Phys. Lett.* **304**, 145 (1999).
5. A. V. Potapov *et al.*, *Astrophys. J.* **703**, 2108 (2009).
6. A. R. W. McKellar, *J. Chem. Phys.* **112**, 9282 (2000).
7. L. A. Surin, B. S. Dumesh, G. Winnewisser, I. Pak, *J. Chem. Phys.* **113**, 9351 (2000).

8. A. V. Potapov, V. A. Panfilov, A. A. Dolgov, L. A. Surin, B. S. Dumesh, *Opt. Spectrosc.* **106**, 655 (2009).
9. P. Jankowski, K. Szalewicz, *J. Chem. Phys.* **108**, 3554 (1998).
10. P. Jankowski, K. Szalewicz, *J. Chem. Phys.* **123**, 104301 (2005).
11. T. H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1989).
12. M. Jeziorska *et al.*, *Collect. Czech. Chem. Commun.* **68**, 463 (2003).
13. M. Jeziorska, P. Jankowski, K. Szalewicz, B. Jeziorski, *J. Chem. Phys.* **113**, 2957 (2000).
14. P. Jankowski, *J. Chem. Phys.* **121**, 1655 (2004).
15. P. Jankowski, *J. Chem. Phys.* **128**, 154311 (2008).
16. P. Jankowski, K. Szalewicz, *Chem. Phys. Lett.* **459**, 60 (2008).
17. R. Bukowski *et al.*, *J. Chem. Phys.* **110**, 3785 (1999).
18. J. M. Hutson, BOUND Computer Code, Version 5, distributed by Collaborative Computational Project no. 6 of the UK Science and Engineering Research Council, 1993.
19. K. P. Huber, G. Herzberg, *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).
20. M. Ziolkowski, thesis, Nicolaus Copernicus University (2006).

**Acknowledgment:** Supported by NSF grant CHE-0848589.

#### Supplementary Materials

[www.sciencemag.org/cgi/content/full/336/6085/1147/DC1](http://www.sciencemag.org/cgi/content/full/336/6085/1147/DC1)  
Tables S1 to S3  
Figs. S1 and S2

23 February 2012; accepted 9 April 2012  
10.1126/science.1221000

## Secreted Kinase Phosphorylates Extracellular Proteins That Regulate Biomineralization

Vincent S. Tagliabracci,<sup>1</sup> James L. Engel,<sup>1</sup> Jianzhong Wen,<sup>1</sup> Sandra E. Wiley,<sup>1</sup> Carolyn A. Worby,<sup>1</sup> Lisa N. Kinch,<sup>2</sup> Junyu Xiao,<sup>1</sup> Nick V. Grishin,<sup>2,3</sup> Jack E. Dixon<sup>1,3\*</sup>

Protein phosphorylation is a fundamental mechanism regulating nearly every aspect of cellular life. Several secreted proteins are phosphorylated, but the kinases responsible are unknown. We identified a family of atypical protein kinases that localize within the Golgi apparatus and are secreted. Fam20C appears to be the Golgi casein kinase that phosphorylates secretory pathway proteins within S-x-E motifs. Fam20C phosphorylates the caseins and several secreted proteins implicated in biomineralization, including the small integrin-binding ligand, N-linked glycoproteins (SIBLINGs). Consequently, mutations in Fam20C cause an osteosclerotic bone dysplasia in humans known as Raine syndrome. Fam20C is thus a protein kinase dedicated to the phosphorylation of extracellular proteins.

**P**rotein phosphorylation is a nearly universal mechanism used by cells to regulate intracellular and extracellular processes (1). The majority of phosphoproteins are intracellular; however, numerous extracellular proteins are phosphorylated (2–4). The first evidence of pro-

tein phosphorylation was in 1883, when the secreted protein casein was shown to contain stoichiometric amounts of phosphate (5). Casein has been used as a model substrate for the detection of several protein kinases, including the first discovery of a protein kinase activity and the identification of casein kinase-1 and casein kinase-2 (6, 7). However, these enzymes are physiologically unrelated to casein because they are mainly cytosolic and nuclear proteins that would be unlikely to encounter casein in the secretory pathway and have therefore been renamed protein kinase CK1 and protein kinase CK2 (6). A

physiological casein kinase activity has been characterized from highly enriched Golgi fractions of lactating mammary gland, liver, brain, and kidney and named Golgi-enriched fraction casein kinase (GEF-CK) (8–10). The GEF-CK specifically recognizes the consensus S-x-E/pS (where x is any amino acid and E/pS can be Glu or phosphoserine) (11). This motif is phosphorylated in some 75% of human plasma and cerebrospinal fluid phosphoproteins (2–4, 12).

To identify candidates for the GEF-CK, we searched for protein kinases containing a signal peptide (SP) and no transmembrane helix. This architecture would orient the kinase in the lumen of the Golgi, in close proximity to proteins destined for secretion. Four-jointed is one such kinase that localizes within the Golgi and phosphorylates extracellular domains of cadherins in *Drosophila* (13). Therefore, we used the human four-jointed (Fjx1) sequence to identify, by means of Position-Specific Iterated- Basic Local Alignment Search Tool (PSI-BLAST), a family of eukaryotic proteins that are distantly related to the bacterial kinase HipA (Fig. 1A and fig. S1) (14). Members of the family with sequence similarity 20 (Fam20) and 198 (Fam198) have SPs and conserved residues required for protein kinase activity (fig. S2). To determine whether these proteins were secreted, we expressed C-terminal FLAG-tagged proteins in the human osteosarcoma cell line U2OS and analyzed FLAG immunoprecipitates from cell extracts and conditioned medium by means of protein immunoblotting. Some 90% of Fam20C was detected in the medium, and the

<sup>1</sup>Department of Pharmacology, University of California, San Diego, La Jolla, CA 92093-0721, USA. <sup>2</sup>University of Texas, Southwestern Medical Center, Dallas, TX 75390-9050, USA. <sup>3</sup>Howard Hughes Medical Institute, Chevy Chase, MD 20815-6789, USA.

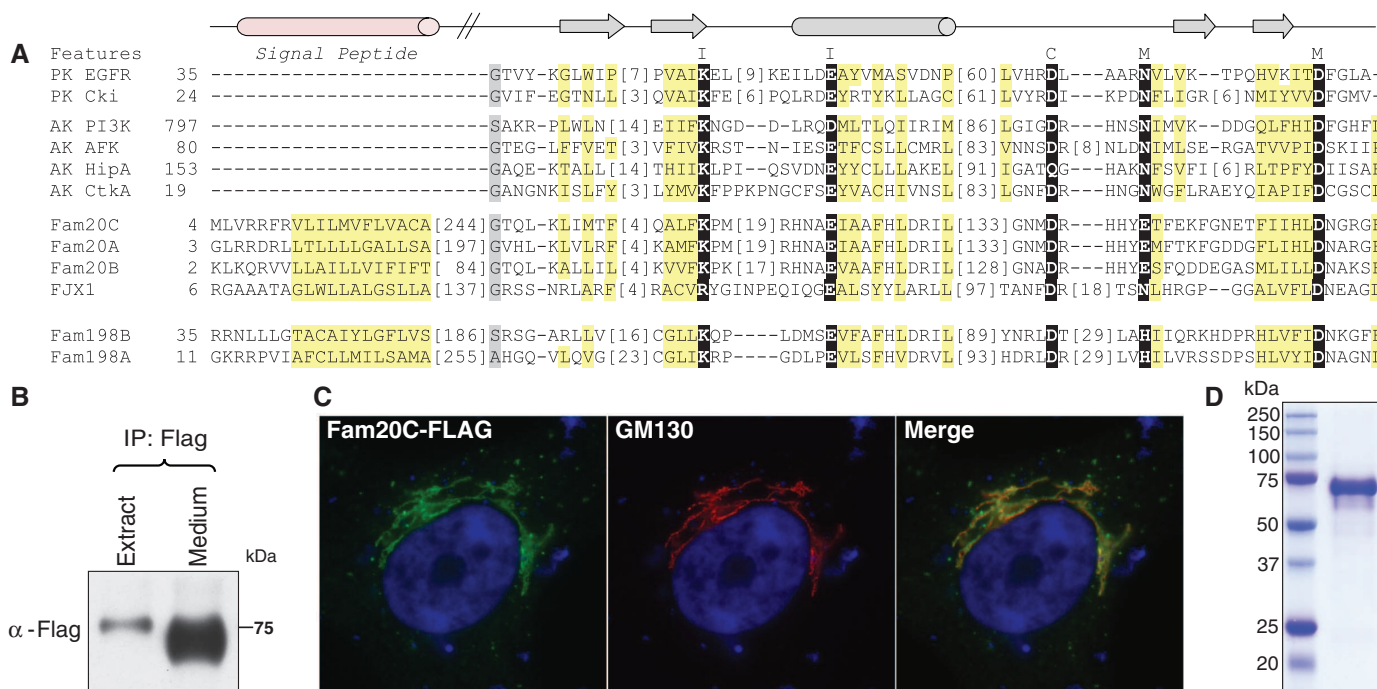
\*To whom correspondence should be addressed. E-mail: jedixon@ucsd.edu

intracellular protein colocalized with the Golgi resident protein GM130 (Fig. 1, B and C). Similarly, the other family members localized within the secretory pathway, and most were secreted (fig. S3, A and B). To test whether Fam20C was a protein kinase, we generated a human embryonic kidney (HEK) 293 T cell line stably expressing a C-terminal FLAG-tagged Fam20C and immunopurified the fusion protein to homogeneity from conditioned medium (Fig. 1D). The recombinant protein was N-linked glycosylated (fig. S4) and catalyzed phosphorylation of several peptides when incubated with [<sup>32</sup>P]ATP (ATP, adenosine 5'-triphosphate) and a Ser-Thr kinase substrate array (Fig. 1E). Many of the substrates contained an S-x-E motif (Fig. 1F). Therefore, we synthesized a substrate array consisting of peptides representing phosphorylation sites from secreted proteins (table S1) (2, 3, 12). Fam20C phosphorylated ~55% of the peptides and had

preference for peptides containing S-x-E motifs (fig. S5, A and B, and table S1). Fam20C phosphorylated Ser more readily than Thr and Tyr and tolerated only Glu at the *n*+2 position (fig. S5, C and D). Thus, Fam20C is a secreted protein kinase that specifically recognizes the consensus sequence S-x-E.

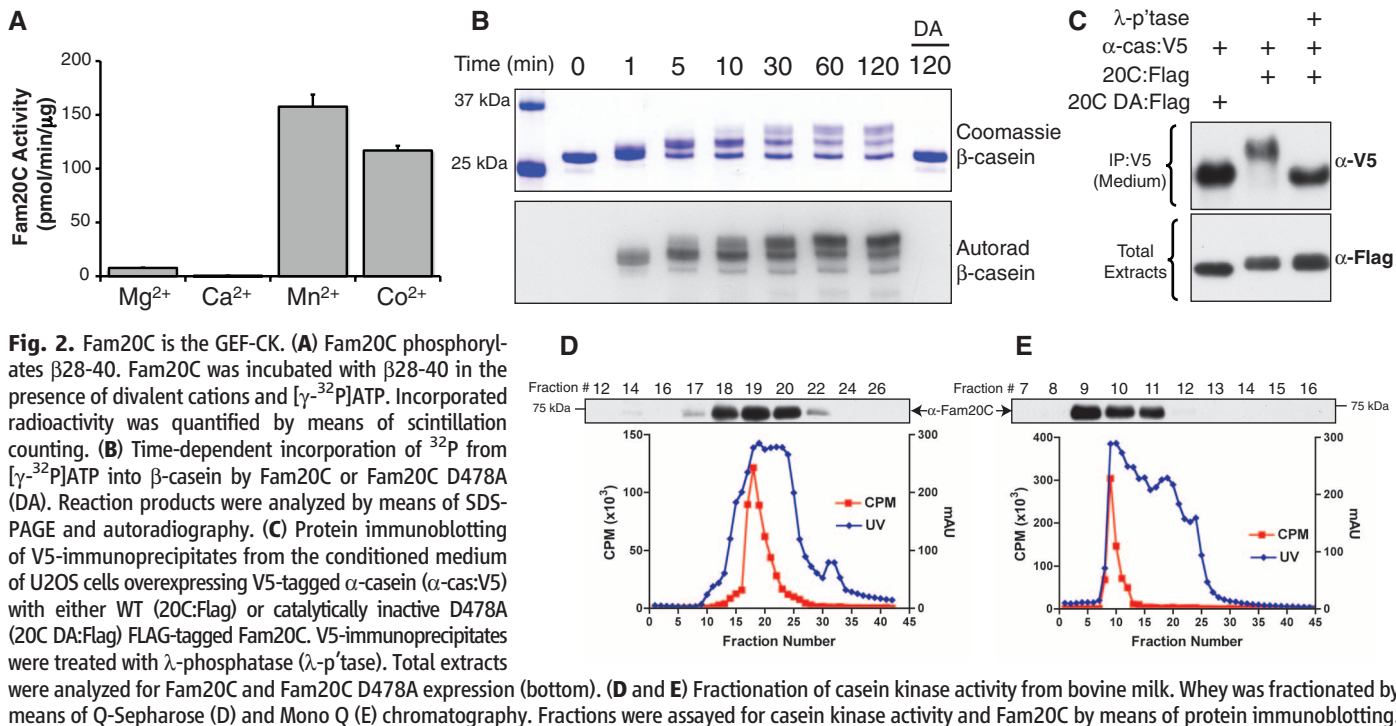
A peptide substrate reproducing one of the sites phosphorylated in β-casein, KKIEKQSE-EQQQ (β28-40), is selectively phosphorylated by GEF-CK (11). The GEF-CK consensus site partially overlaps with that of the polo-like kinase-2/3 (PLK-2/3) consensus site, E-x-x-S-x-E. However, β28-40 is not phosphorylated by PLK2 (15). Furthermore, the GEF-CK requires Mn<sup>2+</sup>, as an alternative to Mg<sup>2+</sup> as the activating cation (10, 16). Fam20C phosphorylated β28-40 in the presence of MnCl<sub>2</sub> and CoCl<sub>2</sub> more efficiently than in the presence of MgCl<sub>2</sub> and CaCl<sub>2</sub> (Fig. 2A). GEF-CK is insensitive to the kinase inhibitor staurosporine

(10). Staurosporine (200 μM) had no effect on Fam20C activity (fig. S6). Fam20C phosphorylated recombinant β-casein in a time-dependent manner, whereas the catalytically inactive D478A mutant, that is unable to coordinate Mn<sup>2+</sup>, did not (Fig. 2B and fig. S2). (In the mutants, other amino acids were substituted at certain locations; for example, D478A indicates that aspartic acid at position 478 was replaced by alanine). Protein immunoblotting of immunoprecipitates from conditioned medium of U2OS cells overexpressing V5-tagged α<sub>s1</sub>-casein with FLAG-tagged Fam20C revealed a mobility shift in α<sub>s1</sub>-casein that was absent when α<sub>s1</sub>-casein was coexpressed with the inactive Fam20C D478A mutant (Fig. 2C). Treatment with λ-phosphatase reversed the Fam20C-dependent mobility shift of α<sub>s1</sub>-casein, confirming a phosphorylation event (Fig. 2C). Furthermore, depletion of endogenous Fam20C in HEK293 T cells with small interfering RNA



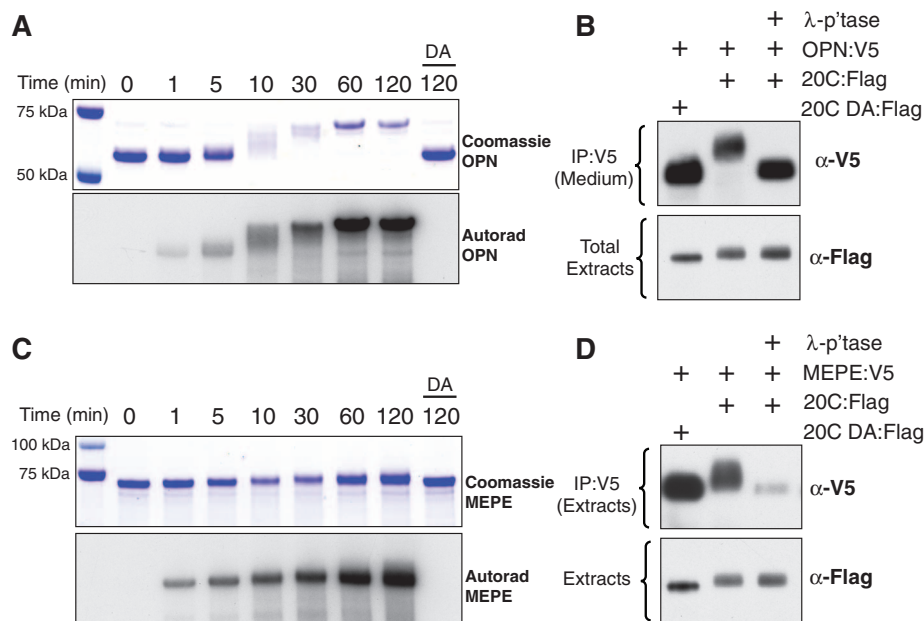
**Fig. 1.** Fam20C is a secreted protein kinase that phosphorylates S-x-E motifs. (A) Sequence alignment of members of the Fjx1 family of proteins with Canonical (PK) and atypical (AK) protein kinases highlighting conserved residues required for protein kinase activity (black), hydrophobic residues contributing to the structure (yellow), and the SP. Conserved protein kinase secondary structural elements and sequence features are indicated above. I, ion pair residues; C, catalytic residue; M, metal-binding residues. (Single letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.) (B) Protein immunoblotting of FLAG immunoprecipitates from the cell extract and conditioned medium of U2OS cells expressing FLAG-tagged Fam20C. (C) Immunofluorescence analysis of HeLa cells overexpressing FLAG-tagged Fam20C. The Golgi resident protein GM130 is shown. (D) SDS-polyacrylamide gel electrophoresis

(SDS-PAGE) and Coomassie staining of FLAG-tagged Fam20C immunopurified from the conditioned medium from HEK293 T cells. (E) Autoradiograph of a kinase substrate array depicting peptides (in duplicate) phosphorylated by Fam20C. (F) Sequences of peptides phosphorylated by Fam20C in (E).



**Fig. 2.** Fam20C is the GEF-CK. **(A)** Fam20C phosphorylates  $\beta$ 28-40. Fam20C was incubated with  $\beta$ 28-40 in the presence of divalent cations and  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ . Incorporated radioactivity was quantified by means of scintillation counting. **(B)** Time-dependent incorporation of  $^{32}\text{P}$  from  $[\gamma\text{-}^{32}\text{P}]\text{ATP}$  into  $\beta$ -casein by Fam20C or Fam20C D478A (DA). Reaction products were analyzed by means of SDS-PAGE and autoradiography. **(C)** Protein immunoblotting of V5-immunoprecipitates from the conditioned medium of U2OS cells overexpressing V5-tagged  $\alpha$ -casein ( $\alpha$ -cas:V5) with either WT (20C:Flag) or catalytically inactive D478A (20C DA:Flag) FLAG-tagged Fam20C. V5-immunoprecipitates were treated with  $\lambda$ -phosphatase ( $\lambda$ -p'tase). Total extracts were analyzed for Fam20C and Fam20C D478A expression (bottom). **(D and E)** Fractionation of casein kinase activity from bovine milk. Whey was fractionated by means of Q-Sepharose (D) and Mono Q (E) chromatography. Fractions were assayed for casein kinase activity and Fam20C by means of protein immunoblotting.

**Fig. 3.** Phosphorylation of SIBLINGs by Fam20C. **(A)** Time-dependent incorporation of  $^{32}\text{P}$  into OPN by Fam20C or the D478A mutant. Reaction products were analyzed by means of SDS-PAGE and autoradiography. **(B)** Phosphorylation of OPN by Fam20C. Protein immunoblotting of V5-immunoprecipitates from the medium of U2OS cells overexpressing V5-tagged OPN (OPN:V5) with either WT (20C:Flag) or catalytically inactive D478A (20C DA:Flag) FLAG-tagged Fam20C. V5-immunoprecipitates were treated with  $\lambda$ -phosphatase ( $\lambda$ -p'tase). Total extracts were analyzed for Fam20C and Fam20C D478A expression (bottom). **(C)** Phosphorylation of MEPE by Fam20C as in (A). **(D)** Protein immunoblotting of V5-immunoprecipitates from cell extracts of U2OS cells overexpressing V5-tagged MEPE (MEPE:V5) with either WT (20C:Flag) or catalytically inactive D478A (20C DA:Flag) FLAG-tagged Fam20C. Samples were analyzed as in (B).



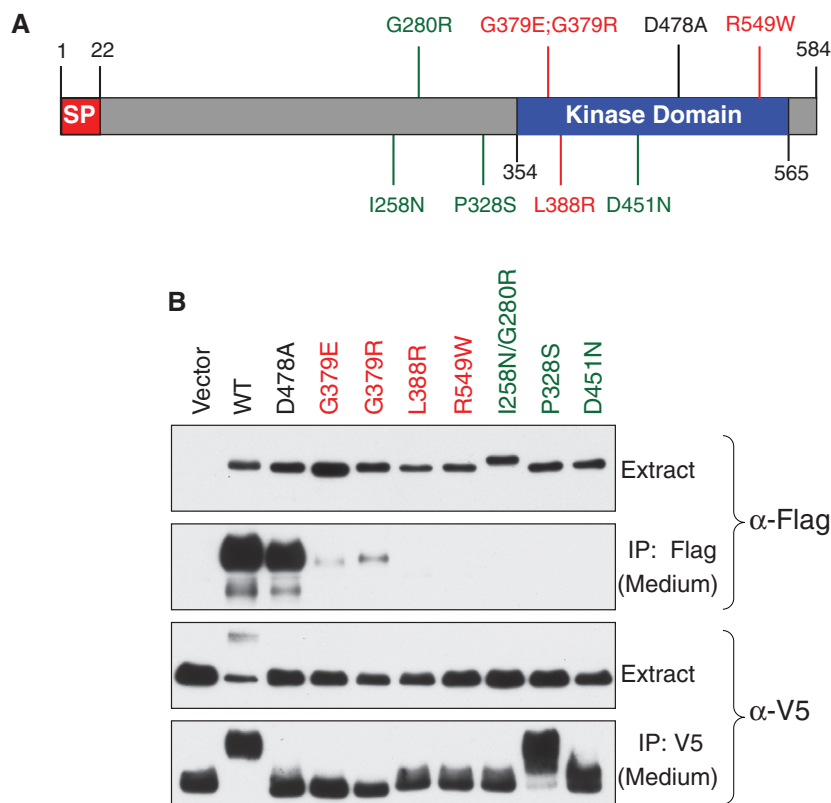
increased the mobility of overexpressed  $\alpha_{s1}$ -casein (fig. S7, A and B). Thus, Fam20C appears to phosphorylate casein in vitro and in mammalian cells.

GEF-CK is secreted and has been purified ~80,000-fold from bovine milk (17). We therefore prepared whey from nonpasteurized, non-homogenized bovine milk and fractionated it by means of Q-Sepharose and Mono Q chromatography (GE Healthcare, Pittsburgh, PA). The resulting fractions were assayed for casein kinase activity and Fam20C protein by means of immunoblotting. Casein kinase activity was detected

primarily in fractions containing Fam20C protein (Fig. 2, D and E). These results support our conclusion that Fam20C is the GEF-CK.

The Fam20C S-x-E consensus motif is found in several secreted proteins, including a family of secretory calcium-binding phosphoproteins (SCPP) that cluster to chromosome 4 in humans (fig. S8) (18). Members of the SCPP family (including the caseins) are secreted phosphoproteins, have a high affinity for calcium, and regulate biomineralization. The small integrin-binding ligand, N-linked glycoproteins (SIBLINGs) are SCPPs encoded by five identically oriented tandem genes

clustered within an ~375-kb span of nucleotides on human chromosome 4 (fig. S8). The genes encode osteopontin (OPN), dentin matrix protein-1 (DMP1), bone sialoprotein (BSP), matrix extracellular phosphoglycoprotein (MEPE), and dentin sialophosphoprotein (DSPP). The SIBLINGs are highly phosphorylated proteins (DSPP has ~200 pSer) and contain multiple phosphorylated S-x-E/S motifs (fig. S9). To explore whether the SIBLINGs are substrates for Fam20C, we purified recombinant OPN, MEPE, and DMP1 from *Escherichia coli* and used them in in vitro kinase assays. Fam20C phosphorylated OPN,



**Fig. 4.** Fam20C mutations and Raine syndrome. **(A)** Schematic representation of the Fam20C protein depicting missense mutations associated with Raine syndrome. SP, signal peptide. Lethal and nonlethal mutants are in red and green, respectively. **(B)** Secretion of Fam20C mutants. Protein immunoblotting of FLAG and V5 immunoprecipitates from conditioned medium of U2OS cells cotransfected with V5-tagged OPN and FLAG-tagged mutants of Fam20C. Cell extracts were also analyzed for intracellular protein levels.

MEPE, and DMP1 in a time-dependent manner, whereas the inactive Fam20C D478A mutant did not (Fig. 3, A and C, and fig. S10A). Overexpression of V5-tagged OPN or V5-tagged MEPE with FLAG-tagged Fam20C and subsequent protein immunoblotting of immunoprecipitates revealed a mobility shift of tagged OPN and MEPE that was sensitive to  $\lambda$ -phosphatase and absent in the presence of Fam20C D478A mutant (Fig. 3, B and D). Fam20C activity was dependent on a functional SP. Deletion of the SP prevented OPN phosphorylation and Fam20C secretion (fig. S11, A and B). Depletion of Fam20C by using lentiviral-based short hairpin RNA also prevented OPN phosphorylation (fig. S12). Extracellular substrates for Fam20C are not restricted to the SCPP family of proteins. Salivary acidic proline-rich phosphoprotein-1 (PRP1)—a secreted phosphoprotein found in saliva and bone morphogenic protein-15 (BMP15), which is a transforming growth factor- $\beta$  superfamily member secreted by oocytes—were effectively phosphorylated by overexpressed Fam20C but not by the D478A Fam20C mutant (fig. S10, B and C). Consistently, OPN, PRP1, and BMP15 are known substrates of GEF-CK (19–21).

Mutations in Fam20C cause Raine syndrome, a bone dysplasia characterized by osteosclerosis and ectopic calcifications that often causes death

in the neonatal period (22, 23). We reasoned that abnormal phosphorylation of the SIBLINGs might account for the biomineralization phenotype in Raine patients. Phosphorylated S-x-E motifs bind calcium and regulate calcium phosphate precipitation as hydroxyapatite (HA) (24). For example, OPN inhibits HA formation, and this inhibition is dependent on the degree of phosphorylation (25). We therefore generated Fam20C missense mutants associated with lethal and nonlethal forms of Raine syndrome (Fig. 4A). The FLAG-tagged mutants were overexpressed with V5-tagged OPN in U2OS cells. Fam20C secretion and OPN phosphorylation were monitored by means of protein immunoblotting. The Fam20C mutants phosphorylated OPN less efficiently than the wild-type (WT) protein, as judged by the OPN mobility shift (Fig. 4B), even though several mutants colocalized with OPN (fig. S13). Most mutations prevented Fam20C secretion, despite the fact that many localized within the secretory pathway (Fig. 4B and fig. S14). The nonlethal P328S and D451N mutants phosphorylated OPN, albeit not as efficiently as the WT protein. Thus, mutations in Fam20C resulting in Raine syndrome appear to affect Fam20C kinase activity and secretion.

We identified a family of secreted protein kinases and identified Fam20C as the GEF-CK that phosphorylates proteins destined for secretion on

S-x-E (fig. S15). This may have broad biological consequences because S-x-E motifs are phosphorylated in extracellular proteins, including pepsin (26) and fibrinogen (27), as well as some biologically active peptide hormones, including adrenocorticotropin (28), progastrin (29), and many others.

#### References and Notes

- P. Cohen, *Nat. Cell Biol.* **4**, E127 (2002).
- J. M. Bahl, S. S. Jensen, M. R. Larsen, N. H. Heegaard, *Anal. Chem.* **80**, 6308 (2008).
- W. Zhou *et al.*, *J. Proteome Res.* **8**, 5523 (2009).
- M. Carrascal *et al.*, *J. Proteome Res.* **9**, 876 (2010).
- O. Hammarsten, *Zeitschrift für Physiologische Chemie* **7**, 227 (1883).
- J. E. Allende, C. C. Allende, *FASEB J.* **9**, 313 (1995).
- G. Burnett, E. P. Kennedy, *J. Biol. Chem.* **211**, 969 (1954).
- A. Moore, A. P. Boulton, H. W. Heid, E. D. Jarasch, R. K. Craig, *Eur. J. Biochem.* **152**, 729 (1985).
- E. W. Bingham, H. M. Farrell Jr., J. J. Basch, *J. Biol. Chem.* **247**, 8193 (1972).
- M. Laso, O. Marin, L. A. Pinna, *Eur. J. Biochem.* **243**, 719 (1997).
- M. Laso-Benito, O. Marin, F. Meggio, L. A. Pinna, *FEBS Lett.* **382**, 149 (1996).
- M. Salvi, L. Cesaro, E. Tibaldi, L. A. Pinna, *J. Proteome Res.* **9**, 3335 (2010).
- H. O. Ishikawa, H. Takeuchi, R. S. Haltiwanger, K. D. Irvine, *Science* **321**, 401 (2008).
- M. A. Schumacher *et al.*, *Science* **323**, 396 (2009).
- M. Salvi *et al.*, *Biochem. Biophys. Res. Commun.* **418**, 156 (2012).
- E. W. Bingham, H. M. Farrell Jr., *J. Biol. Chem.* **249**, 3647 (1974).
- J. S. Duncan, M. C. Wilkinson, R. D. Burgoyne, *Biochem. J.* **350**, 463 (2000).
- K. Kawasaki, K. M. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 4060 (2003).
- E. Tibaldi *et al.*, *FEBS Lett.* **584**, 801 (2010).
- A. M. Brunati, O. Marin, A. Bisinella, A. Salviati, L. A. Pinna, *Biochem. J.* **351**, 765 (2000).
- M. Laso, P. L. Chang, C. W. Prince, L. A. Pinna, *Biochem. Biophys. Res. Commun.* **240**, 602 (1997).
- J. Raine, R. M. Winter, A. Davey, S. M. Tucker, *J. Med. Genet.* **26**, 786 (1989).
- M. Fradin *et al.*, *Clin. Genet.* **80**, 177 (2011).
- A. George, A. Veis, *Chem. Rev.* **108**, 4670 (2008).
- G. K. Hunter, C. L. Kyle, H. A. Goldberg, *Biochem. J.* **300**, 723 (1994).
- J. Tang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3437 (1973).
- B. Blombaeck, M. Blombaeck, P. Edman, B. Hessel, *Nature* **193**, 833 (1962).
- H. P. Bennett, C. A. Browne, S. Solomon, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4713 (1981).
- G. J. Dockray *et al.*, *J. Biol. Chem.* **262**, 8643 (1987).

**Acknowledgments:** This work was supported in part by NIH grants DK018849-36 and DK018024-37 (to J.E.D.), GM094575 (to N.V.G.), NIH/National Cancer Institute Training Grant T32 CA009523 (to V.S.T.), and the Welch Foundation I-1505 (to N.V.G.). We thank J. Jewell, G. Taylor, X. Guo, S. Mattoo, A. Newton, S. Taylor, A. DePaoli-Roach, P. Roach, and members of the Dixon lab for insightful discussions and comments regarding the manuscript.

#### Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1217817/DC1  
Materials and Methods

Figs. S1 to S15  
Tables S1 and S2  
References (30–41)

13 December 2011; accepted 26 April 2012  
Published online 10 May 2012;  
10.1126/science.1217817