

Type II-FAS	FabB FabF	ISSACATSA IATACTSGV	-126aa- -131aa-	YLNSHGTST YVNAHGTGT	-24aa- -26aa-	TKAMTGHSLGA TKSMTGHTLQA
Type I-PKS	DEBS1	VDTACSSSL	-126aa-	AVEAHGTGT	-27aa-	VKSNLGHITQAA
Type I-PKS	PikaIV	VDTACSSSL	-126aa-	VVEGHGTGT	-29aa-	LKSNIGHGTGT
Type II-PKS	Act K $\alpha$ Tcm K $\alpha$	VSTGCTSGL VSTGCTSGL	-131aa- -131aa-	YINAHGSGT YINAHGSGT	-26aa- -26aa-	IKSMVGHSLGA IKSMIGHSLGA
Type III-PKS	CHS2 RppA	YQQGCFAGG AQLGCAAGG	-130aa- -123aa-	FWIAHPGGP FFIVHAGGP	-22aa- -22aa-	VLSDYGNMSSA TLTERGNIASS
	NonK NonJ	VSCGCASSS VSGSCNVAL	-143aa- -122aa-	YVNGGEGGD FVNDYADGN	-26aa- -28aa-	QEACFGHSGAP QEAVFGHVAGT

**Fig. 4.** Alignments of the conserved catalytic residues of NonJK with KSs from FAS, type I and II PKS, and type III PKS. The conserved Cys residue for C–C or C–O bond formation and His-His or His-Asn residues for decarboxylation are highlighted. Protein accession numbers are given after the protein names: FabB, P14926; FabF, P39435; DEBS1, Q03131; PikaIV, AAC69332; Act K $\alpha$ , CAC44200; Tcm K $\alpha$ , AAA67515; CHS2, P30074; RppA, BAA33495; NonJ, AAD37451; and NonK, AAD37450.

mutated catalytic triad—Cys-Tyr-His for NonJ or Cys-Gly-His for NonK, suggesting that NonJK lack the decarboxylation activity (Fig. 4). This is consistent with the proposal that NonJK catalyze C–O bond formation by using the –OH as the nucleophile directly. Finally, to confirm that Cys plays a catalytic role in the C–O bond-forming step, we replaced the conserved Cys residue in NonJ or NonK, respectively, with Gly by site-directed mutagenesis (24). The resultant mutants completely lose their ability to biotransform ( $\pm$ )-3 into **1** (Fig. 2, B and C, entries IX and X). Thus, the NonJK KSs catalyze the C–O bond-forming step in nonactin biosynthesis, acting directly on ( $\pm$ )-2 and using the same active-site residue Cys that is used in KS catalysis of C–C bond formation. Given the high sequence homology and conserved active site between NonJK and other KSs, the structural plasticity and catalytic flexibility of KSs upon protein engineering (6, 9, 10), and now the C–O bond-forming ability of the NonJK KSs, further mechanistic and structural characterization of the NonJK KSs could allow us to rationally engineer the C–O bond-forming activity into other KSs. This could expand the size and diversity of polyketide library accessible by combinatorial biosynthesis.

**References and Notes**

- D. A. Hopwood, *Chem. Rev.* **97**, 2465 (1997).
- R. McDaniel, S. Ebert-Khosla, D. A. Hopwood, C. Khosla, *Science* **262**, 1546 (1993).
- B. Shen, C. R. Hutchinson, *Science* **262**, 1535 (1993).
- A. F. A. Marsden *et al.*, *Science* **263**, 378 (1994).
- R. Pieper, G. Luo, D. E. Cane, C. Khosla, *Nature* **378**, 263 (1995).
- R. McDaniel, S. Ebert-Khosla, D. A. Hopwood, C. Khosla, *Nature* **375**, 549 (1995).
- R. S. Gokhale, S. Y. Tsuji, D. E. Cane, C. Khosla, *Science* **284**, 482 (1999).
- B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, C. Khosla, *Science* **291**, 1790 (2001).
- D. E. Cane, C. T. Walsh, C. Khosla, *Science* **282**, 63 (1998).
- J. M. Jez *et al.*, *Chem. Biol.* **7**, 919 (2000).
- S. Eckermann *et al.*, *Nature* **396**, 387 (1998).
- J. Kennedy *et al.*, *Science* **284**, 1368 (1999).
- N. Funa *et al.*, *Nature* **400**, 897 (1999).
- J. G. Metz *et al.*, *Science* **293**, 290 (2001).
- W. Liu, S. D. Christenson, S. Standage, B. Shen, *Science* **297**, 1170 (2002).
- J. Ahlert *et al.*, *Science* **297**, 1173 (2002).
- J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **18**, 380 (2001).

- B. Shen, *Top. Curr. Chem.* **209**, 1 (2000).
- C. Bisang *et al.*, *Nature* **401**, 502 (1999).
- J. Schröder, in *Comprehensive Natural Products Chemistry*, D. Barton, K. Nakanishi, O. Meth-Cohn, Eds. (Elsevier, New York, 1999), vol. 1, pp. 749–771.
- W. C. Smith, L. Xiang, B. Shen, *Antimicrob. Agents Chemother.* **44**, 1809 (2000).

- H.-J. Kwon, W. C. Smith, L. Xiang, B. Shen, *J. Am. Chem. Soc.* **123**, 3385 (2001).
- M. E. Nelson, N. D. Priestley, *J. Am. Chem. Soc.* **124**, 2894 (2002).
- Supplementary figures and additional information about materials and methods are available on Science Online.
- W. F. Fleck, M. Ritzau, S. Heinze, U. Gräfe, *J. Basic Microbiol.* **36**, 235 (1996).
- M. He, M. Varoglu, D. H. Sherman, *J. Bacteriol.* **182**, 2619 (2000).
- J. G. Olsen *et al.*, *FEBS Lett.* **460**, 46 (1999).
- J. M. Jez, J.-L. Ferrer, M. E. Bowman, R. A. Dixon, J. P. Noel, *Biochemistry* **39**, 890 (2000).
- Supported in part by a CAREER award (MCB9733938) from NSF to B.S. B.S. is a recipient of an Independent Scientist Award (A151689) from NIH. A.J.S. was supported in part by NIH grant T32 GM08505.

**Supporting Online Material**

www.sciencemag.org/cgi/content/full/297/5585/1327/DC1  
Materials and Methods  
Figs. S1 and S2

23 April 2002; accepted 12 July 2002

# Structure of the Extracellular Region of HER3 Reveals an Interdomain Tether

Hyun-Soo Cho and Daniel J. Leahy\*

We have determined the 2.6 angstrom crystal structure of the entire extracellular region of human HER3 (ErbB3), a member of the epidermal growth factor receptor (EGFR) family. The structure consists of four domains with structural homology to domains found in the type I insulin-like growth factor receptor. The HER3 structure reveals a contact between domains II and IV that constrains the relative orientations of ligand-binding domains and provides a structural basis for understanding both multiple-affinity forms of EGFRs and conformational changes induced in the receptor by ligand binding during signaling. These results also suggest new therapeutic approaches to modulating the behavior of members of the EGFR family.

The epidermal growth factor receptor (EGFR) is the founding member of the ErbB family of receptor tyrosine kinases that in humans includes HER1 (EGFR, ErbB1), HER2 (Neu, ErbB2), HER3 (ErbB3), and HER4 (ErbB4) (1–4). These receptors respond to EGF and related ligands to mediate cellular growth and differentiation in multiple tissues in both the developing embryo and adult (5–8). Loss of any of the ErbB family members results in embryonic lethality in mice with defects observed in organs including the brain, heart, skin, lung, and gastrointestinal tract, depending on the receptor affected (7). Overexpression and activation of ErbB receptors, most notably HER1 and

HER2, are found in many human cancers and are critical factors in the development and malignancy of these tumors (9). Therapies that target these receptors have shown promise, and a monoclonal antibody against HER2, with the trade name Herceptin, is currently being used to treat breast cancer (10).

ErbB receptors consist of an ~620-amino acid extracellular region followed by a single transmembrane-spanning region and a cytoplasmic kinase domain. The extracellular regions of ErbB receptors are made up of four domains arranged as a tandem repeat of a two-domain unit consisting of an ~190-amino acid L domain followed by an ~120-amino acid cysteine-rich domain. The first three of these domains share 15 to 20% sequence identity with the first three domains of the type I insulin-like growth factor receptor (IGFR), for which a structure is known (11). Unlike the homologous region of HER1, an NH<sub>2</sub>-terminal three-domain frag-

Department of Biophysics and Biophysical Chemistry, Howard Hughes Medical Institute, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205, USA.

\*To whom correspondence should be addressed. E-mail: leahy@groucho.med.jhmi.edu

## REPORTS

ment of IGFR does not bind ligand with high affinity (11, 12), indicating that ErbB receptors and IGFR not only interact with different ligands but also use different binding modes.

The ErbB kinases become activated after ligand binding, which leads to autophosphorylation and recruitment of adaptor proteins that in turn activate downstream components of the signaling pathway (8). Ligand-induced oligomerization is thought to be the mechanism by which the kinases become activated (13, 14), although recent studies suggest that EGFR may exist as a dimer in the cell membrane and a ligand-induced conformational change stimulates kinase activation (15). More than 15 EGF-like ligands are known; many exhibit specificity for multiple ErbB receptors and can stimulate both homo- and heteromeric receptor combinations (6). It is through pairing with other ErbB family members that HER3, which has an inactive kinase domain, becomes phosphorylated

and capable of activating downstream signaling components (16).

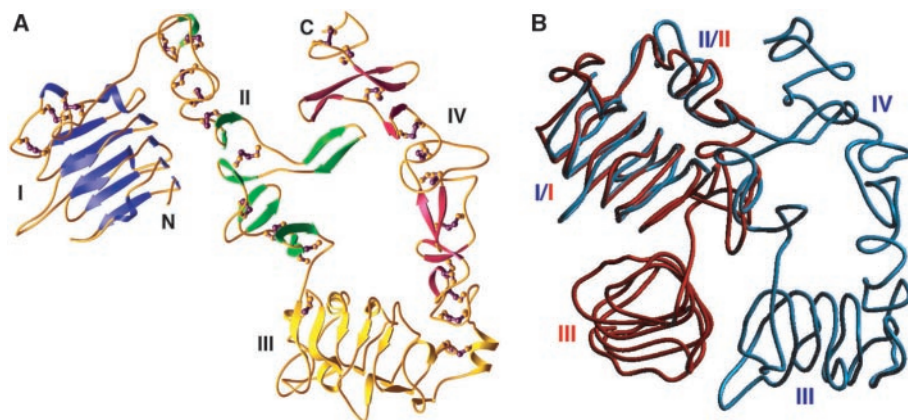
Isolated ErbB receptors and receptor fragments display binding and dimerization properties that are more complex than would be expected from a simple ligand-induced dimerization model. The active signaling complex appears to contain two EGF and two receptor molecules (17), but it has been difficult to detect stable dimers of receptor-ligand complexes after ligand binding to soluble extracellular regions of HER1 (sHER1) (17, 18). Biophysical studies indicate that ligand binding induces marked conformational changes in the absence of dimerization (19). If the transmembrane region is added to the extracellular portion of EGFR, ligand-induced receptor dimers become detectable at nanomolar receptor concentrations (20). This result led to the suggestion that unliganded receptors adopt a conformation that inhibits association of the transmembrane regions

(20), a model consistent with the observation that the transmembrane regions of ErbB receptors self-associate in membranes (21).

Ligand binding studies also indicate multiple sHER1 states. sHER1 exhibits a small (~5%) population with a dissociation constant ( $K_d$ ) for EGF of 2 to 20 nM and a larger population with a  $K_d$  of 400 to 550 nM (22). A small population of receptors with higher affinity is also observed on cells (23). These multiple-affinity states appear to be mediated in part by domain IV; deletion of domain IV converts sHER1 to a single high-affinity state with a  $K_d$  for EGF of 13 to 21 nM (12).

To provide a molecular basis for understanding ligand-induced activation of ErbB receptors, we have determined the 2.6 Å crystal structure of the 621-amino acid extracellular region of human HER3 (sHER3). Extensive glycosylation of ErbB receptor extracellular domains—up to 30% of their molecular weight consists of carbohydrate—has hindered previous attempts to determine their structure. To circumvent this problem, we expressed sHER3 in Lec1 mutant Chinese hamster ovary (CHO) cells (24) and deglycosylated it with endoglycosidase H before crystallization (25). The structure was solved with multiwavelength anomalous diffraction data collected from crystals derivatized with mercurochrome (25); data collection and refinement statistics are listed in Table 1.

A ribbon diagram of the sHER3 structure is shown in Fig. 1A. Domains I and III of sHER3 exhibit the expected  $\beta$ -helical structure, and domains II and IV are extended repeats of seven small disulfide-containing modules. Domains I, II, and III of sHER3 may be individually superimposed with their IGFR counterparts with root mean square deviations (rmsds) of C $\alpha$  positions of 0.92, 1.55, and 1.11 Å for 91, 59, and 82 core residues, respectively. Domain II is slightly



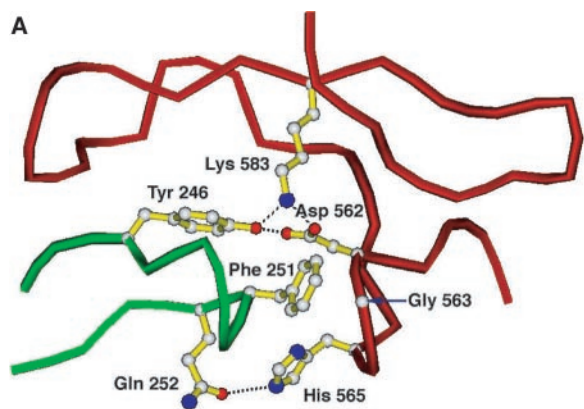
**Fig. 1.** The sHER3 structure. **(A)** Ribbon diagram of sHER3. Domain I (blue), domain II (green), domain III (yellow), and domain IV (red) are indicated. Disulfide bonds are shown in purple and gold, and the NH<sub>2</sub>- and COOH-termini are labeled. **(B)** Backbone traces of sHER3 (blue) and IGFR (red) are shown following superposition of the NH<sub>2</sub>-terminal L1 domains. **(A)** was made with RIBBONS (34), and **(B)** was made with SETOR (35).

**Table 1.** Data collection and refinement statistics. Statistics for all crystals except Native1 and Native2 were calculated considering  $F^+$  and  $F^-$  separately. The PDB accession code for atomic coordinates and structure factors is 1M6B.

Crystal	Native1	Native2	HgChrom1	HgChrom2	Hgchrom2	Hgchrom2	Hgchrom3	Hgchrom3
X-ray source	BNL X4A	Rigaku	Rigaku	BNL X4A	BNL X4A	BNL X4A	BNL X4A	BNL X4A
Wavelength (Å)	0.9199	1.5418	1.5418	0.9998	1.0088	1.0204	0.9998	1.0090
Resolution range (Å)	30–2.6	30–3.4	30–3.0	30–3.0	30–3.0	30–3.0	30–3.0	30–3.0
$R_{\text{merge}}^{*\dagger}$	0.067(0.521)	0.102(0.310)	0.076(0.321)	0.055(0.476)	0.050(0.410)		0.055(0.161)	0.053(0.147)
Completeness*	96.3 (79.2)	97.8 (98.7)	99.3 (97.7)	93.2 (87.9)	93.2 (87.9)		99.4 (100.0)	96.7 (98.3)
$\langle I/\sigma I \rangle^*$	15.2 (2.2)	9.9 (2.5)	17.1 (2.9)	16.8 (1.5)	19.3 (2.1)		21.6 (4.4)	20.2 (4.2)
Redundancy	3.0	2.2	4.1	3.0	3.0		3.3	3.1
Number of sites			4	2	2	2	4	4
Mean figure of merit (30–3.0 Å) = 0.50								
$R_{\text{crist}}^{\ddagger}$		0.235						
$R_{\text{free}}^{\ddagger}$		0.294						
rmsd bonds (Å)		0.013						
rmsd angles (°)		1.59						
rmsd $B$ values (Å <sup>2</sup> );								
Bonds/angles main chain		0.57/1.14						
Bonds/angles side chain		2.26/4.11						

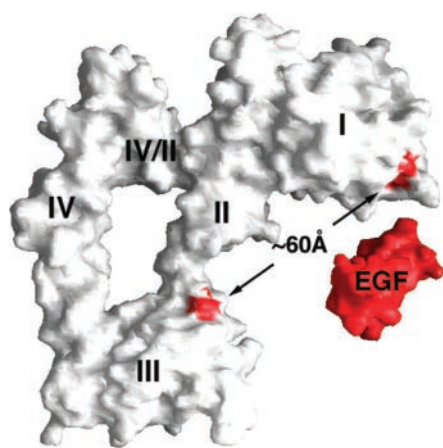
\*Values in parentheses apply to the high-resolution shell.  $\dagger R_{\text{merge}} = \frac{\sum_h \sum_i |I_i(h) - \langle I(h) \rangle|}{\sum_h \sum_i I_i(h)}$  where  $i$  is the  $i$ th measurement of reflection  $h$  and  $\langle I(h) \rangle$  is a weighted mean of all measurements of  $h$ .  $\ddagger R = \frac{\sum_h |F_{\text{obs}}(h) - F_{\text{calc}}(h)|}{\sum_h |F_{\text{obs}}(h)|}$ .  $R_{\text{crist}}$  and  $R_{\text{free}}$  were calculated from the working and test reflection sets, respectively. The test set constituted 5% of the total reflections not used in refinement.

## REPORTS



**B** HER3 (244) LVYNKLTFQLEP... (559) HFRDGPHCVSS... (581) IYKY  
 HER1 (244) MLYNPTTYQMDV... (560) HYIDGPHCVKT... (583) VWKY  
 HER2 (266) FVYNPTTFQLEH... (582) HFKDGPNCVEK... (604) IFKY  
 HER4 (272) VTYNTDTFESMP... (589) HYKDPPCVAR... (581) IWKF

**Fig. 2.** The domain II/IV contact. **(A)** Side chains of residues that mediate contacts between domain II (green) and domain IV (red) are shown. Hydrogen bonds and salt bridges are indicated by dashed lines. This figure was made with MOLSCRIPT (36). **(B)** Alignment of human ErbB receptor sequences from the domain II/IV contact regions. Residues with direct contacts to one another in sHER3 are highlighted in similar colors. Pro<sup>593</sup> and Phe<sup>595</sup> of HER2, which fail to conserve the pattern of contact residues, are underlined.



**Fig. 3.** Surface representations of sHER3 and EGF (32). sHER3 is rotated  $\sim 180^\circ$  about a vertical axis relative to its orientation in Fig. 1A. EGF and the sites in sHER3 (Y104 and V333) homologous to the sites in HER1 to which EGF cross-links (30, 31) are in red. Domains I to IV and the domain II/IV connection are labeled. This figure was made with the program GRASP (37).

kinked in IGFR relative to sHER3 with a bend following a region with insertions in both proteins (residues 242 to 259 in sHER3). Thirty-two core residues before this region and 27 after this region superimpose with rmsds of 0.60 and 1.16 Å, respectively. A surface area of 1000 Å<sup>2</sup> is buried between sHER3 domains I and II, and they retain an interdomain orientation roughly similar to that observed in the IGFR structure (Fig. 1B). A surface area of 326 Å<sup>2</sup> is buried between sHER3 domains II and III, and the relative positions of domain III of sHER3 and IGFR differ substantially when the NH<sub>2</sub>-terminal domains are superimposed (Fig. 1B). Interdomain flexibility between domains II and III may thus provide a focus for conformational change in the ErbB and insulin receptor families.

A notable feature of the sHER3 structure is a  $\beta$ -hairpin loop (residues 242 to 259) that extends nearly 20 Å from domain II to contact the COOH-terminal portion of domain

IV and creates a large pore in sHER3 bounded by domains II, III, and IV (Figs. 1 to 3). This contact, which is intramolecular and thus not a result of crystal packing, buries 810 Å<sup>2</sup> of surface and is mediated by a single main-chain/main-chain hydrogen bond and three sets of side-chain interactions (Y246 with D562 and K583, F251 with G563, and Q252 with H565) (Fig. 2) (26). Calculation of a shape complementarity parameter for this interface yields a value of 0.75, which is comparable to those observed for protease/protease inhibitor complexes and notably larger than is observed for antibody/antigen complexes (27).

Conservation of the residues at this contact interface in HER1 and HER4 and the appearance of a pore-containing structure markedly similar to that of sHER3 in electron micrographs of sHER1 (28) suggest that this contact is a conserved feature of HER1 and HER4 (Fig. 2B). However, the substitution of G563 and H565 with proline and phenylalanine, respectively, in HER2 suggest this contact may not be conserved in HER2. In particular, the positive  $\phi$  observed for G563 could not be preserved after substitution with proline, which would alter the main-chain as well as the side-chain configuration. HER2 is unique among ErbB receptors in not possessing a high-affinity ligand, and the presence of a domain II/IV contact may thus correlate with regulation of ligand binding.

The only points of communication between domain IV and the other domains of sEGFR are its attachment to domain III and the contact with domain II, which implicates these features in mediating the 20- to 40-fold decrease in affinity for ligand when domain IV is present (12). The sHER3 structure provides a simple model for how this inhibition might occur. The binding site for EGF and transforming growth factor- $\alpha$  (TGF- $\alpha$ ) on HER1 maps predominantly to domain III, with a lesser contribution from domain I (29), and EGF cross-links to specific residues in both domains in conditions where stable receptor dimers are

not observed (30, 31). Residues in sHER3 homologous to the cross-link sites (Y104 and V333) are separated by  $\sim 60$  Å, and a substantial domain rearrangement would be needed for simultaneous contact of these sites by EGF, which has a longest dimension of just over 30 Å (32) (Fig. 3). By constraining the movement of the COOH-terminus of domain III, the domain II/IV contact seems likely to provide a barrier to formation of the optimal arrangement of domains I and III for ligand binding. Providing such a constraint also couples ligand binding to a large conformational change that could promote dimerization or other modes of signaling through the presentation of new interaction surfaces, the removal of barriers to interactions, or both.

The presence of a low-population, high-affinity form of sHER1 with a  $K_d$  comparable to that observed when domain IV is deleted suggests an equilibrium between low- and high-affinity forms of the receptor (22). A simple basis for such an equilibrium could be the opening and closing of the domain II/IV contact or a domain shift that preserves the contact. The constitutive presence of the high-affinity form of sHER1 indicates that conversion to this form is by itself insufficient to initiate signaling.

The involvement of ErbB receptors in the development and severity of several human cancers has stimulated much interest in developing modulators of their activity. The sHER3 structure suggests at least two approaches to developing such modulators. Proteolysis experiments indicate that domain III of HER1 binds EGF and TGF- $\alpha$  with micromolar affinity (33). The lack of contiguity between the domain I and III ligand binding sites in the unliganded sHER3 structure suggests that a ligand that bound only to domain III (or domain I) would bind the receptor but fail to induce the domain arrangement that seems necessary for signaling. Therefore, it should be possible to identify molecules—for example, mutant forms of EGF or TGF- $\alpha$ —that function as dominant-negative inhibitors of

ErbB receptor function by binding receptor but failing to induce signaling.

The nature of the domain II/IV interaction, which is mediated entirely by a short hairpin loop from domain II, suggests another approach to modulating ErbB receptor behavior. Cyclic or linear peptides corresponding to this loop region (or the pocket on domain IV) or their analogs may be able to disrupt this interaction and potentiate ligand binding. Conversely, if ligand binding requires breaking of the domain II/IV interaction and the domain II loop then participates in interreceptor interactions, such peptides may antagonize signaling. Future experiments are needed to evaluate these strategies.

References and Notes

1. A. Ullrich *et al.*, *Nature* **309**, 418 (1984).
2. C. I. Bargmann, M. C. Hung, R. A. Weinberg, *Nature* **319**, 226 (1986).
3. M. H. Kraus, W. Issing, T. Miki, N. C. Popescu, S. A. Aaronson, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9193 (1989).
4. G. D. Plowman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4905 (1990).
5. S. Cohen, *In Vitro Cell Dev. Biol.* **23**, 239 (1987).
6. D. J. Riese II, D. F. Stern, *Bioessays* **20**, 41 (1998).
7. M. A. Olayioye, R. M. Neve, H. A. Lane, N. E. Hynes, *EMBO J.* **19**, 3159 (2000).
8. J. Schlessinger, *Cell* **103**, 211 (2000).
9. C. K. Tang, M. E. Lippman, in *Hormones and Signaling*, B. W. O'Malley, Ed. (Academic Press, San Diego, CA, 1998), vol. I, pp. 113–165.
10. S. Shak, *Semin. Oncol.* **26**, 71 (1999).
11. T. P. Garrett *et al.*, *Nature* **394**, 395 (1998).
12. T. C. Elleman *et al.*, *Biochemistry* **40**, 8930 (2001).
13. Y. Schechter, L. Hernaez, J. Schlessinger, P. Cuatrecasas, *Nature* **278**, 835 (1979).
14. Y. Yarden, J. Schlessinger, *Biochemistry* **26**, 1443 (1987).
15. T. W. Gadella Jr., T. M. Jovin, *J. Cell Biol.* **129**, 1543 (1995).
16. M. X. Sliwkowski *et al.*, *J. Biol. Chem.* **269**, 14661 (1994).
17. M. A. Lemmon *et al.*, *EMBO J.* **16**, 281 (1997).
18. P. M. Brown *et al.*, *Eur. J. Biochem.* **225**, 223 (1994).
19. C. Greenfield *et al.*, *EMBO J.* **8**, 4115 (1989).
20. K. G. Tanner, J. Kyte, *J. Biol. Chem.* **274**, 35985 (1999).
21. J. M. Mendrola, M. B. Berger, M. C. King, M. A. Lemmon, *J. Biol. Chem.* **277**, 4704 (2002).
22. T. Domagala *et al.*, *Growth Factors* **18**, 11 (2000).
23. J. Schlessinger *et al.*, *CRC Crit. Rev. Biochem.* **14**, 93 (1983).
24. P. Stanley, V. Caillibot, L. Siminovitch, *Cell* **6**, 121 (1975).
25. Supplementary materials and methods are available on Science Online.
26. 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.
27. M. C. Lawrence, P. M. Colman, *J. Mol. Biol.* **234**, 946 (1993).
28. I. Lax *et al.*, *J. Biol. Chem.* **266**, 13828 (1991).
29. I. Lax *et al.*, *Cell Regul.* **2**, 337 (1991).
30. D. G. Wu, L. H. Wang, Y. Chi, G. H. Sato, J. D. Sato, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3151 (1990).
31. R. L. Woltjer, T. J. Lukas, J. V. Staros, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7801 (1992).
32. G. T. Montelione *et al.*, *Biochemistry* **31**, 236 (1992).
33. D. Kohda *et al.*, *J. Biol. Chem.* **268**, 1976 (1993).
34. M. Carson, Ed., *Ribbons* (Academic Press, San Diego, CA, 1997), vol. 277.
35. S. Evans, *J. Mol. Graphics* **11**, 134 (1993).
36. P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).

37. A. Nicholls, K. A. Sharp, B. Honig, *Proteins* **11**, 281 (1991).
38. We thank A. Ullrich for supplying the Her-3 cDNA; C. Ogata and M. Becker for assistance at beamlines X-4A and X-25, respectively, at Brookhaven Laboratory NSLS; P. Longo and P.-Y. Wu for technical assistance; D. Denney for helpful discussions; W. Yang for assistance with refinement and figures; and M. Amzel, J. Berg, J. Nathans, P. Beachy, G. Rose, J. Lorsch, P. Cole, M. Lemmon, K. Ferguson, and S.

Bouyain for comments on the manuscript. This work was supported by the Howard Hughes Medical Institute and NIH.

Supporting Online Material

www.sciencemag.org/cgi/content/full/297/5585/1130/DC1  
Materials and Methods  
Figs. S1 and S2

3 June 2002; accepted 18 July 2002

# Variant of SCN5A Sodium Channel Implicated in Risk of Cardiac Arrhythmia

Igor Splawski,<sup>1\*</sup> Katherine W. Timothy,<sup>2</sup> Michihiro Tateyama,<sup>3</sup> Colleen E. Clancy,<sup>3</sup> Alka Malhotra,<sup>2</sup> Alan H. Beggs,<sup>4</sup> Francesco P. Cappucco,<sup>5</sup> Giuseppe A. Sagnella,<sup>6</sup> Robert S. Kass,<sup>3</sup> Mark T. Keating<sup>1\*</sup>

Every year, ~450,000 individuals in the United States die suddenly of cardiac arrhythmia. We identified a variant of the cardiac sodium channel gene *SCN5A* that is associated with arrhythmia in African Americans ( $P = 0.000028$ ) and linked with arrhythmia risk in an African-American family ( $P = 0.005$ ). In transfected cells, the variant allele (Y1102) accelerated channel activation, increasing the likelihood of abnormal cardiac repolarization and arrhythmia. About 13.2% of African Americans carry the Y1102 allele. Because Y1102 has a subtle effect on risk, most carriers will never have an arrhythmia. However, Y1102 may be a useful molecular marker for the prediction of arrhythmia susceptibility in the context of additional acquired risk factors such as the use of certain medications.

Cardiac arrhythmias are a common cause of morbidity and mortality (1). Myocardial infarction, cardiac ischemia, cardiomyopathy, and many medications are common risk factors for life-threatening cardiac arrhythmias. However, not all individuals with a specific risk factor develop arrhythmias, and the reasons for this variability in response are not understood. One possibility is that genetic factors modulate arrhythmia risk in the setting of common, extrinsic factors (2).

The *SCN5A* gene encodes  $\alpha$  subunits that form the sodium channel responsible for initiating the cardiac action potential (3). Mutations in *SCN5A* have been implicated in rare, familial

forms of cardiac arrhythmia, including long QT syndrome (4, 5), idiopathic ventricular fibrillation (6), and cardiac conduction disease (7–9). To identify common polymorphisms that increase the risk of arrhythmia in the general population, we screened DNA samples obtained from individuals with nonfamilial cardiac arrhythmias. In one case, a 36-year-old African-American woman (individual 5, table S1) with idiopathic dilated cardiomyopathy and hypokalemia developed prolongation of the corrected QT (QTc) interval and torsade de pointes ventricular tachycardia while on the anti-arrhythmic agent amiodarone (Fig. 1A). Prolongation of the QT interval is associated with an increased risk of life-threatening ventricular tachyarrhythmias (9). Single-strand conformation polymorphism (SSCP) and DNA sequence analyses (10) revealed a heterozygous transversion of C to A in codon 1102 of *SCN5A*, causing a substitution of serine (S1102) with tyrosine (Y1102). S1102 is a conserved residue located in the intracellular sequences that link domains II and III of the channel (fig. S1).

To determine the frequency of Y1102 in the general population, we used SSCP to screen DNA samples obtained from controls (10). Y1102 was observed in 19.2% of West Africans and Caribbeans (90/468). Eighty-

<sup>1</sup>Department of Cardiology, Children's Hospital, and Departments of Pediatrics and Cell Biology, Harvard Medical School and Howard Hughes Medical Institute, Boston, MA 02115, USA. <sup>2</sup>Department of Human Genetics, University of Utah, Salt Lake City, UT 84112, USA. <sup>3</sup>Department of Pharmacology, College of Physicians and Surgeons of Columbia University, New York, NY 10032, USA. <sup>4</sup>Division of Genetics, Children's Hospital, and Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA. <sup>5</sup>Department of General Practice and Primary Care, <sup>6</sup>Department of Physiological Medicine, St. George's Hospital Medical School, London SW17 0RE, UK.

\*To whom correspondence should be addressed. E-mail: igor@enders.tch.harvard.edu (I.S.); mkeating@enders.tch.harvard.edu (M.T.K.)