Science
AAAS

Emulating Membrane Protein Evolution by Rational Design

Mikaela Rapp, *et al.*

Science **315**, 1282 (2007);
DOI: 10.1126/science.1135406

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of September 22, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/cgi/content/full/315/5816/1282

Supporting Online Material can be found at:
http://www.sciencemag.org/cgi/content/full/1135406/DC1

A list of selected additional articles on the Science Web sites **related to this article** can be found at:
http://www.sciencemag.org/cgi/content/full/315/5816/1282#related-content

This article **cites 27 articles**, 11 of which can be accessed for free:
http://www.sciencemag.org/cgi/content/full/315/5816/1282#otherarticles

This article has been **cited by** 38 article(s) on the ISI Web of Science.

This article has been **cited by** 13 articles hosted by HighWire Press; see:
http://www.sciencemag.org/cgi/content/full/315/5816/1282#otherarticles

This article appears in the following **subject collections**:
Biochemistry
http://www.sciencemag.org/cgi/collection/biochem

Downloaded from www.sciencemag.org on September 22, 2010

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2007 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

23. S. F. Grant *et al.*, *Nat. Genet.* **38**, 320 (2006).
24. Y. Horikawa *et al.*, *Nat. Genet.* **17**, 384 (1997).
25. K. Yamagata *et al.*, *Nature* **384**, 458 (1996).
26. S. D. Brown *et al.*, *Biochem. Biophys. Res. Commun.* **248**, 879 (1998).
27. Materials and methods are available as supporting material on Science Online.

28. We thank members of the kindred studied for their invaluable participation in this project. Research supported in part by an NIH KO8 award (to A.M.) and grants P50 HL55007 and P01DK68229 from NIH (to R.L.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/315/5816/1278/DC1

Materials and Methods
Figs. S1 to S3
Tables S1 and S2
References

16 October 2006; accepted 1 February 2007
10.1126/science.1136370

Emulating Membrane Protein Evolution by Rational Design

Mikaela Rapp,[1]* Susanna Seppälä,[1]* Erik Granseth,[1,2] Gunnar von Heijne[1,2]†

How do integral membrane proteins evolve in size and complexity? Using the small multidrug-resistance protein EmrE from *Escherichia coli* as a model, we experimentally demonstrated that the evolution of membrane proteins composed of two homologous but oppositely oriented domains can occur in a small number of steps: An original dual-topology protein evolves, through a gene-duplication event, to a heterodimer formed by two oppositely oriented monomers. This simple evolutionary pathway can explain the frequent occurrence of membrane proteins with an internal pseudo–two-fold symmetry axis in the plane of the membrane.

Membrane protein evolution is often accomplished by gene-duplication and gene-fusion events (*1*), and high-resolution membrane protein structures have disclosed an unanticipated number of cases where homologous N- and C-terminal domains are related by an approximate two-fold symmetry axis either perpendicular to or in the plane of the membrane. In the former case, each domain has an even number of transmembrane helices and the two domains are oriented parallel to each other in the membrane, whereas in the latter case each domain has an odd number of transmembrane helices and the two domains are antiparallel. Representative examples of membrane proteins with parallel domains are LacY (*2*), GlpT (*3*), the Sav1866 ABC transporter (*4*), AcrB (*5*), EmrD (*6*), and the ADP/ATP carrier (*7*); among membrane proteins with antiparallel domains are LeuT (*8*), SecY (*9*), BtuCD (*10*), AQP1 (*11*), GlpF (*12*), AmtB (*13*), the ClC H^+/Cl^- exchange transporter (*14*), and NhaA (*15*).

A particularly notable mode of gene duplication–based membrane protein evolution was suggested recently by an analysis of proteins in the small multidrug-resistance (SMR) family (*16*). The best-studied SMR protein is EmrE from *E. coli*, an inner-membrane drug-efflux pump with four transmembrane helices. EmrE likely has a dual topology with identical copies of the protein forming an antiparallel homodimer (or higher oligomer) composed of N_{in}-C_{in} and N_{out}-C_{out} monomers (*16*–*20*), although some data suggest a parallel N_{in}-C_{in} dimer (*21*, *22*). Membrane protein topology is largely governed by the positive-inside rule (*23*)—i.e., loops rich in Lys and Arg residues tend to orient toward the cytoplasm. EmrE has only a weak K+R bias (*24*), as would be expected for a dual-topology protein, and it is encoded by a singleton gene with no homologous genes nearby on the chromosome. In contrast, many genes encoding SMR proteins, in both *E. coli* and other bacteria, appear as pairs on the chromosome, and the two encoded proteins have large but opposite K+R biases and therefore likely insert into the membrane with opposite orientations (*16*). These findings immediately suggest an evolutionary scenario in which a singleton gene encoding a dual-topology protein undergoes a gene duplication, after which the two resulting proteins become fixed in opposite orientations by evolving opposite K+R biases. Possibly, oppositely oriented proteins occasionally mutate back to a dual topology (*25*). In a final step, the two oppositely oriented proteins may fuse into a single polypeptide that folds into an antiparallel structure

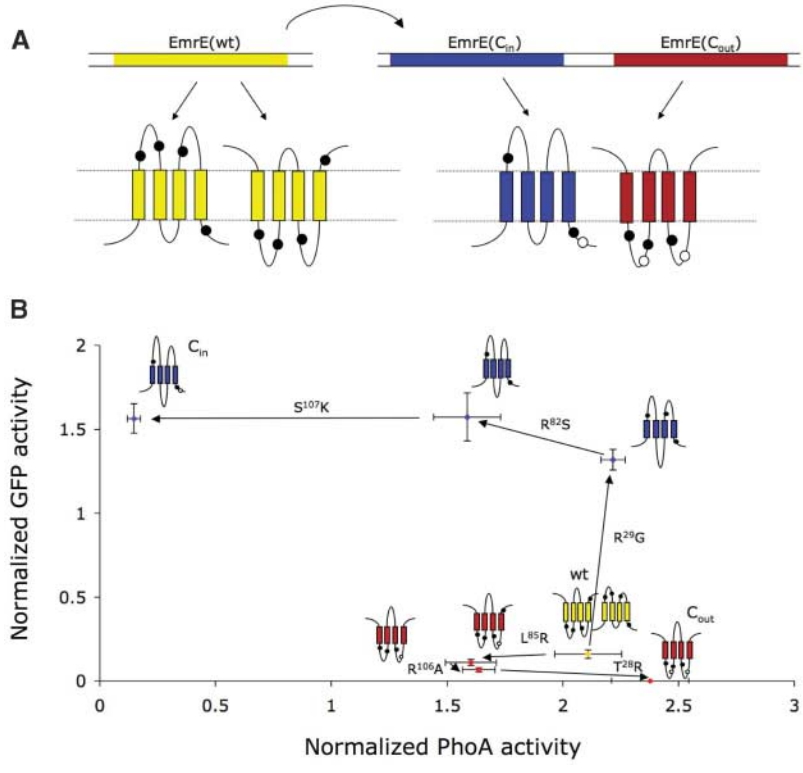

Fig. 1. (**A**) Conversion of the wild-type (wt) dual-topology protein EmrE to two oppositely oriented proteins, EmrE(C_{in}) and EmrE(C_{out}). Black circles indicate positively charged residues present in wild-type EmrE; white circles indicate positively charged residues added by mutagenesis. The mutations in EmrE(C_{in}) are R29G, R82S, and S107K, and the mutations in EmrE(C_{out}) are T28R, L85R, and R106A. (**B**) PhoA activities and GFP fluorescence levels for C-terminal PhoA and GFP fusions to the indicated EmrE mutants [normalized against the median PhoA activity or GFP fluorescence of a large set of *E. coli* inner-membrane protein PhoA and GFP fusions (*26*)]. High PhoA activity and low GFP fluorescence indicate a periplasmic localization of the C terminus, whereas low PhoA activity and high GFP fluorescence indicate a cytoplasmic localization. Error bars show standard errors in the mean value determinations.

[1]Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden. [2]Stockholm Bioinformatics Center, AlbaNova, SE-106 91 Stockholm, Sweden.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: gunnar@dbb.su.se

Downloaded from www.sciencemag.org on September 22, 2010

with a pseudo–two-fold symmetry axis in the plane of the membrane (*16*).

EmrE provides an ideal case to test this evolutionary scenario. According to the dual-topology model, two oppositely oriented EmrE monomers form the active homodimer. We reasoned that mutants of EmrE designed to insert into the inner membrane with a unique N_{in}-C_{in} or N_{out}-C_{out} orientation should therefore be nonfunctional when expressed alone but should complement each other when coexpressed, hence emulating the proposed gene-duplication and topology-evolution steps. In contrast, should the active protein be a parallel homodimer, as has been suggested (*21*), molecules with N_{in}-C_{in} and N_{out}-C_{out} orientations would not be expected to complement each other. Further mutation of functionally important residues in the N_{in}-C_{in} or N_{out}-C_{out} mutant might allow the design of active heterodimers incorporating single-site mutations that would inactivate wild-type homodimeric EmrE, reflecting a wider scope for evolutionary fine-tuning of oppositely oriented heterodimers as compared with dual-topology homodimers.

To push EmrE toward the N_{in}-C_{in} and N_{out}-C_{out} orientations, we manipulated the K+R bias (Fig. 1A). The orientation in the inner membrane of the various EmrE mutants was probed by making C-terminal fusions to the topology reporters green fluorescent protein (GFP) and alkaline phosphatase (PhoA) (*26*). The different fusion constructs all express to similar levels (fig. S1A), and the GFP and PhoA activities show that construct EmrE(C_{in}) indeed has an N_{in}-C_{in} topology, whereas construct EmrE(C_{out}) adopts the opposite orientation (Fig. 1B).

Expression of EmrE makes *E. coli* resistant to high levels of ethidium bromide (EtBr) and other cationic hydrophobic drugs (*27*). We found that cells expressing wild-type EmrE (that was not fused to GFP or PhoA) grow well in 0.25 mM EtBr (Fig. 2A and fig. S1B), whereas cells transformed with empty vector do not (see fig. S2 for growth at different concentrations of EtBr and fig. S3 for growth curves in 0.25 mM EtBr). The expression of EmrE(C_{in}) and EmrE(C_{out}) individually (in single or double copy) does not confer resistance to EtBr. Notably, however, coexpression of EmrE(C_{in}) and EmrE(C_{out}) restores EtBr resistance to the same level as seen with wild-type EmrE (Fig. 2A, green bar), strongly suggesting the formation of a functional, antiparallel heterodimer.

The EtBr-resistance levels conferred by expression of the mutants that are intermediate between wild-type EmrE and EmrE(C_{in}) or EmrE(C_{out}) are also consistent with the proposed evolutionary scenario (Fig. 2A). For cells expressing the two intermediates on the way to EmrE(C_{out})—EmrE($L^{85}\rightarrow R^{85}$) (hereafter, L85R) and EmrE(L85R, R106A)—the growth rate drops with the number of mutations, as expected. For cells expressing the intermediates on the way to EmrE(C_{in})—EmrE(R29G) and EmrE(R29G, R82S)—the growth rate remains close to that of wild-type EmrE and only with the addition of the final mutation (S107K) in EmrE(C_{in}) does the growth rate drop to near zero, mirroring the GFP/PhoA fusion protein results that indicate a mixed orientation of both intermediates.

Although these results are in full agreement with the proposed dual topology of wild-type EmrE, taking the activities of the wild-type EmrE GFP and PhoA fusions at face value suggests that these fusions have a predominantly N_{out}-C_{out} orientation (Fig. 1B). This prompted us to measure growth rates in the presence of EtBr also for cells expressing the GFP-reporter fusions. In contrast to the results for the nonfused constructs, cells expressing wild-type EmrE-GFP grow less well in EtBr than do cells expressing the C_{in} intermediates EmrE(R29G)-GFP and EmrE(R29G,

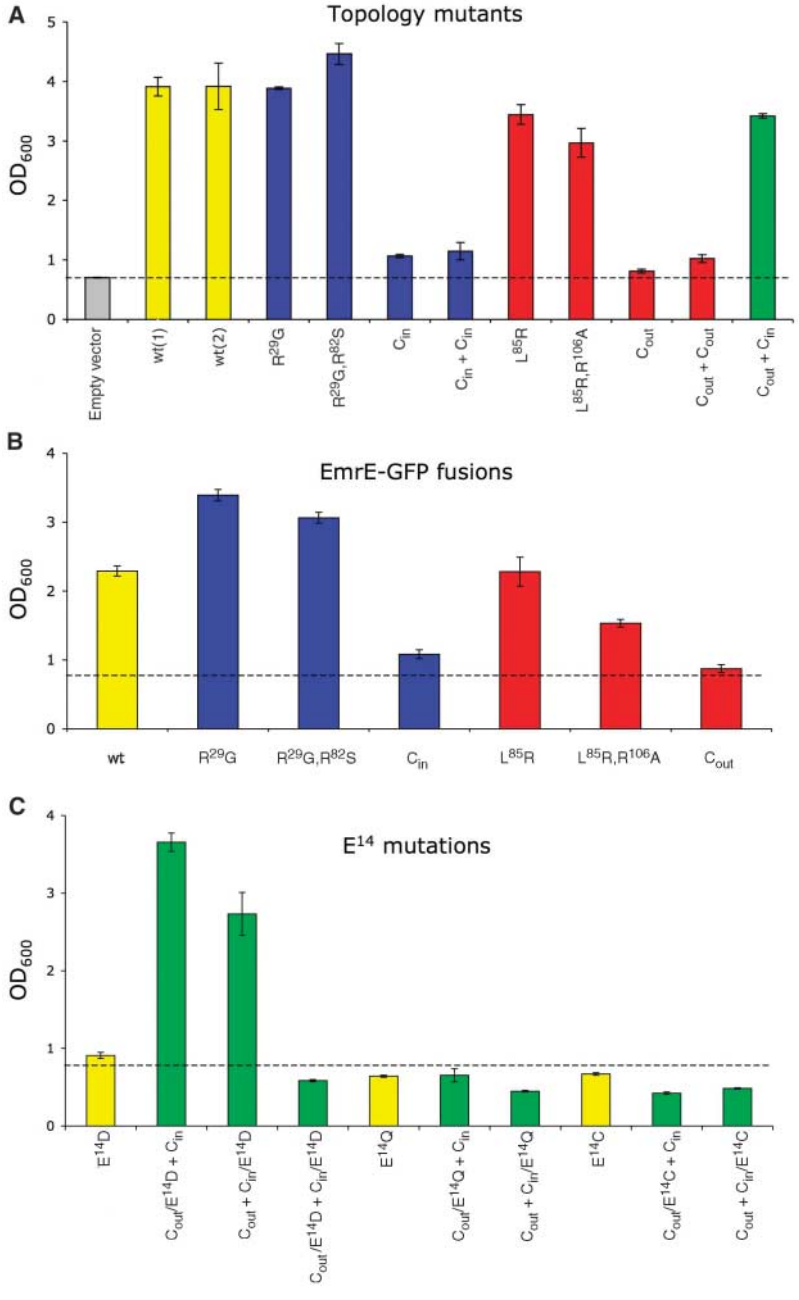

Fig. 2. (**A**) Optical density at 600 nm (OD_{600}) after 4 hours of growth in 0.25 mM EtBr of cells expressing the indicated EmrE mutants (not fused to GFP or PhoA). See Fig. 1 for color code. Green bar, coexpressed EmrE(C_{in}) + EmrE(C_{out}). The dashed line indicates the OD_{600} of cells transformed with empty pET Duet-1 vector (gray bar). Bars wt(1) and wt(2) show results for wild-type EmrE cloned into the first and second multiple-cloning site in the vector. (**B**) OD_{600} after 4 hours of growth in 0.25 mM EtBr of cells expressing the indicated EmrE-GFP fusions. (**C**) OD_{600} after 4 hours of growth in 0.25 mM EtBr of cells expressing the indicated EmrE mutants (not fused to GFP or PhoA). In all panels, error bars show standard errors in the mean value determinations.

Downloaded from www.sciencemag.org on September 22, 2010

R82S)-GFP; cells expressing EmrE(C_{in})-GFP do not grow, as expected (Fig. 2B). Among the C_{out} intermediates, EmrE(L85R)-GFP confers the same level of EtBr resistance as wild-type EmrE-GFP, whereas cells expressing EmrE(L85R, R106A)-GFP grow less well; those expressing EmrE(C_{out})-GFP grow no better than background.

The simplest explanation for these growth patterns is that the C-terminal GFP moiety causes a shift toward more C_{out} orientation, which is "corrected" by the C_{in}-promoting mutations R29G and R82S, precisely as indicated by the GFP and PhoA activity measurements in Fig. 1B. It is not unexpected that C-terminal reporter fusions, although normally highly reliable indicators of C-terminal orientation (*28*), can have a subtle effect on the orientation of finely balanced dual-topology proteins. Indeed, the same reasoning might also explain the predominant C_{in} orientation seen for an EmrE-Myc-His_6 construct with reduced in vivo activity (*22*), in which positively charged His and Lys residues in the C-terminal tag may favor the C_{in} topology but leave enough C_{out}-oriented protein to confer a certain degree of drug resistance.

A small number of residues are conserved throughout the SMR family. To test whether such residues are required in both or only in one subunit of the dimer, we focused on Glu^{14} in EmrE, a residue intimately involved in the proton-driven extrusion of substrate from the cell (*29*). Glu^{14} is totally conserved in all SMR proteins with low K+R bias, but is replaced by Asp in a small number of N_{out}-C_{out} homologs that presumably form active heterodimers (fig. S4). Indeed, cells coexpressing constructs EmrE(C_{out}/E14D) and EmrE(C_{in}) grow in EtBr at the same rate as cells expressing wild-type EmrE (Fig. 2C and figs. S2 and S3). Cells coexpressing the alternate combination EmrE(C_{out}) and EmrE(C_{in}/E14D) are also EtBr resistant but grow more slowly than the first combination, possibly explaining why Asp^{14} has so far only been found in N_{out}-C_{out}–oriented EmrE homologs. As seen for the homodimeric EmrE(E14D) mutant (*29*), cells coexpressing EmrE(C_{in}/E14D) and EmrE(C_{out}/E14D) do not grow in EtBr. Thus, Glu^{14} can be replaced by Asp in one but not in both monomers in the EmrE dimer. Mutations in the EmrE(C_{out}) or EmrE(C_{in}) monomers that do not conserve the negative charge at residue 14 (E14Q and E14C) do not support growth in EtBr when coexpressed with either EmrE(C_{in}) or EmrE(C_{out}).

Our results show that an evolutionary path connecting a dual-topology protein to a pair of oppositely oriented homologs can be emulated by rational protein design based on the positive-inside rule, and that once a heterodimer of oppositely oriented subunits has appeared, additional mutations that render the original dual-topology protein nonfunctional can be selected in one of the two monomers. With such a readily accessible evolutionary pathway available, it is no surprise that many membrane proteins show signs of internal duplication and approximate two-fold in-plane symmetry in their three-dimensional structures.

References and Notes

1. T. Shimizu, H. Mitsuke, K. Noto, M. Arai, *J. Mol. Biol.* **339**, 1 (2004).
2. J. Abramson *et al.*, *Science* **301**, 610 (2003).
3. Y. Huang, M. J. Lemieux, J. Song, M. Auer, D. N. Wang, *Science* **301**, 616 (2003).
4. R. J. Dawson, K. P. Locher, *Nature* **443**, 180 (2006).
5. S. Murakami, R. Nakashima, E. Yamashita, A. Yamaguchi, *Nature* **419**, 587 (2002).
6. Y. Yin, X. He, P. Szewczyk, T. Nguyen, G. Chang, *Science* **312**, 741 (2006).
7. E. Pebay-Peyroula *et al.*, *Nature* **426**, 39 (2003).
8. A. Yamashita, S. K. Singh, T. Kawate, Y. Jin, E. Gouaux, *Nature* **437**, 215 (2005).
9. B. van den Berg *et al.*, *Nature* **427**, 36 (2004).
10. K. P. Locher, A. T. Lee, D. C. Rees, *Science* **296**, 1091 (2002).
11. H. Sui, B. G. Han, J. K. Lee, P. Walian, B. K. Jap, *Nature* **414**, 872 (2001).
12. D. Fu *et al.*, *Science* **290**, 481 (2000).
13. S. Khademi *et al.*, *Science* **305**, 1587 (2004).
14. R. Dutzler, E. B. Campbell, M. Cadene, B. T. Chait, R. MacKinnon, *Nature* **415**, 287 (2002).
15. C. Hunte *et al.*, *Nature* **435**, 1197 (2005).
16. M. Rapp, E. Granseth, S. Seppälä, G. von Heijne, *Nat. Struct. Mol. Biol.* **13**, 112 (2006).
17. P. J. Butler, I. Ubarretxena-Belandia, T. Warne, C. G. Tate, *J. Mol. Biol.* **340**, 797 (2004).
18. I. Ubarretxena-Belandia, J. M. Baldwin, S. Schuldiner, C. G. Tate, *EMBO J.* **22**, 6175 (2003).
19. S. J. Fleishman *et al.*, *J. Mol. Biol.* **364**, 54 (2006).
20. Y. Elbaz, S. Steiner-Mordoch, T. Danieli, S. Schuldiner, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1519 (2004).
21. M. Soskine, S. Mark, N. Tayer, R. Mizrachi, S. Schuldiner, *J. Biol. Chem.* **281**, 36205 (2006).
22. S. Ninio, Y. Elbaz, S. Schuldiner, *FEBS Lett.* **562**, 193 (2004).
23. G. von Heijne, *EMBO J.* **5**, 3021 (1986).
24. 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.
25. T. Kikukawa, T. Nara, T. Araiso, S. Miyauchi, N. Kamo, *Biochim. Biophys. Acta* **1758**, 673 (2006).
26. Materials and methods are available as supporting material on *Science* Online.
27. M. Putman, H. W. van Veen, W. N. Konings, *Microbiol. Mol. Biol. Rev.* **64**, 672 (2000).
28. D. O. Daley *et al.*, *Science* **308**, 1321 (2005).
29. H. Yerushalmi, S. Schuldiner, *J. Biol. Chem.* **275**, 5264 (2000).

30. Supported by grants from the Swedish Research Council, the Marianne and Marcus Wallenberg Foundation, and the Swedish Foundation for Strategic Research to G.v.H.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1135406/DC1
Materials and Methods
Figs. S1 to S4
References

21 September 2006; accepted 16 January 2007
Published online 25 January 2007;
10.1126/science.1135406
Include this information when citing this paper.

Local Interactions Select for Lower Pathogen Infectivity

Michael Boots* and Michael Mealor

Theory suggests that the current rapid increase in connectivity and consequential changes in the structure of human, agricultural, and wildlife populations may select for parasite strains with higher infectivity. We carried out a test of this spatial theory by experimentally altering individual host movement rates in a model host/pathogen system by altering the viscosity of their environment. In our microevolutionary selection experiments, the infectivity of the virus was, as predicted by the theory, reduced in the most viscous populations. We therefore provide empirical support for the theory that population structure affects the evolution of infectious organisms.

Because of the importance of the evolution of parasites and pathogens to human, agricultural, and wildlife systems, there is a well-developed theory that focuses on how transmission and increased mortality due to infection (virulence) may evolve (*1–4*). The classic theory of parasite evolution shows that natural selection will act to maximize the epidemiological basic reproductive number R_0: the number of secondary infections resulting from one infected host in a naïve host population (*1–4*). Clearly, parasite transmission acts to increase R_0, whereas virulence acts to decrease it by reducing the infectious period. One component of transmission that may evolve in the parasite is the probability of successfully infecting a susceptible host upon contact (infectivity). In the absence of constraints on selection, we would predict that the parasite would evolve maximal infectivity and zero virulence (*1–4*). This classic theory makes the assumption that the host populations are homogeneously mixed, and therefore each susceptible host has an equal probability of being infected by any infectious individual within the population. However, many natural host/parasite systems can be characterized by both localized transmission, where secondary infections are more frequent in individuals neighboring infected hosts, and patchy host distributions. Theory that examines the importance of local interactions (*5–9*) has predicted that spatial structure in host populations can constrain the evolution of parasite infectivity. When

Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN, UK.

*To whom correspondence should be addressed. E-mail: m.boots@sheffield.ac.uk

Downloaded from www.sciencemag.org on September 22, 2010