elongation, a 10-fold molar excess of unlabelled G-actin in buffer B containing 100 mM KCl was added to the Alexa-Fluor-647-labelled filaments and the reaction was allowed to proceed for 30–120 min on ice. Finally, the double-labelled filaments were diluted to a final concentration of 5 to 20 nM, and the elongated part of the filaments was labelled and stabilized by the addition of rhodamine–phalloidin. Double-labelled filaments were only used fresh and were discarded after 12 h.

Received 3 September; accepted 19 December 2003; doi:10.1038/nature02303.

1. Toyoshima, Y. Y. *et al.* Myosin subfragment-1 is sufficient to move actin filaments *in vitro*. *Nature* **328,** 536–539 (1987).
2. Manstein, D. J., Ruppel, K. M. & Spudich, J. A. Expression and characterization of a functional myosin head fragment in *Dictyostelium discoideum*. *Science* **246,** 656–658 (1989).
3. Rayment, I. *et al.* Structure of the actin-myosin complex and its implications for muscle contraction. *Science* **261,** 58–65 (1993).
4. Berg, J. S., Powell, B. C. & Cheney, R. E. A millennial myosin census. *Mol. Biol. Cell* **12,** 780–794 (2001).
5. Wells, A. L. *et al.* Myosin VI is an actin-based motor that moves backwards. *Nature* **401,** 505–508 (1999).
6. Homma, K., Yoshimura, M., Saito, J., Ikebe, R. & Ikebe, M. The core of the motor domain determines the direction of myosin movement. *Nature* **412,** 831–834 (2001).
7. Kollmar, M., Dürrwang, U., Kliche, W., Manstein, D. J. & Kull, F. J. Crystal structure of the motor domain of a class-I myosin. *EMBO J.* **21,** 2517–2525 (2002).
8. Veigel, C. *et al.* The motor protein myosin-I produces its working stroke in two steps. *Nature* **398,** 530–533 (1999).
9. Prakash, B., Renault, L., Praefcke, G. J., Herrmann, C. & Wittinghofer, A. Triphosphate structure of guanylate-binding protein 1 and implications for nucleotide binding and GTPase mechanism. *EMBO J.* **19,** 4555–4564 (2000).
10. Uyeda, T. Q., Abramson, P. D. & Spudich, J. A. The neck region of the myosin motor domain acts as a lever arm to generate movement. *Proc. Natl Acad. Sci. USA* **93,** 4459–4464 (1996).
11. Anson, M., Geeves, M. A., Kurzawa, S. E. & Manstein, D. J. Myosin motors with artificial lever arms. *EMBO J.* **15,** 6069–6074 (1996).
12. Ruff, C., Furch, M., Brenner, B., Manstein, D. J. & Meyhöfer, E. Single-molecule tracking of myosins with genetically engineered amplifier domains. *Nature Struct. Biol.* **8,** 226–229 (2001).
13. Kliche, W., Fujita-Becker, S., Kollmar, M., Manstein, D. J. & Kull, F. J. Structure of a genetically engineered molecular motor. *EMBO J.* **20,** 40–46 (2001).
14. Kull, F. J. & Endow, S. A. Kinesin: switch I & II and the motor mechanism. *J. Cell Sci.* **115,** 15–23 (2002).
15. Vale, R. D. & Milligan, R. A. The way things move: looking under the hood of molecular motor proteins. *Science* **288,** 88–95 (2000).
16. Furch, M., Geeves, M. A. & Manstein, D. J. Modulation of actin affinity and actomyosin adenosine triphosphatase by charge changes in the myosin motor domain. *Biochemistry* **37,** 6317–6326 (1998).
17. Furch, M., Remmel, B., Geeves, M. A. & Manstein, D. J. Stabilization of the actomyosin complex by negative charges on myosin. *Biochemistry* **39,** 11602–11608 (2000).
18. Van Dijk, J. *et al.* Differences in the ionic interaction of actin with the motor domains of nonmuscle and muscle myosin II. *Eur. J. Biochem.* **260,** 672–683 (1999).
19. Knetsch, M. L. W., Uyeda, T. Q. & Manstein, D. J. Disturbed communication between actin- and nucleotide-binding sites in a myosin II with truncated 50/20-kDa junction. *J. Biol. Chem.* **274,** 20133–20138 (1999).
20. Tsiavaliaris, G. *et al.* Mutations in the relay loop region result in dominant-negative inhibition of myosin II function in *Dictyostelium*. *EMBO Rep.* **3,** 1099–1105 (2002).
21. Ponomarev, M. A., Furch, M., Levitsky, D. I. & Manstein, D. J. Charge changes in loop 2 affect the thermal unfolding of the myosin motor domain bound to F-actin. *Biochemistry* **39,** 4527–4532 (2000).
22. Reubold, T. F., Eschenburg, S., Becker, A., Kull, F. J. & Manstein, D. J. A structural model for actin-induced nucleotide release in myosin. *Nature Struct. Biol.* **10,** 826–830 (2003).
23. Ringler, P. & Schulz, G. E. Self-assembly of proteins into designed networks. *Science* **302,** 106–109 (2003).
24. Kron, S. J. & Spudich, J. A. Fluorescent actin filaments move on myosin fixed to a glass surface. *Proc. Natl Acad. Sci. USA* **83,** 6272–6276 (1986).
25. Lehrer, S. S. & Kerwar, G. Intrinsic fluorescence of actin. *Biochemistry* **11,** 1211–1217 (1972).
26. Lorenz, M., Poole, K. J., Popp, D., Rosenbaum, G. & Holmes, K. C. An atomic model of the unregulated thin filament obtained by X-ray fiber diffraction on oriented actin-tropomyosin gels. *J. Mol. Biol.* **246,** 108–119 (1995).
27. Dominguez, R., Freyzon, Y., Trybus, K. M. & Cohen, C. Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: visualization of the pre-power stroke state. *Cell* **94,** 559–571 (1998).
28. Schröder, R. R. *et al.* Three-dimensional atomic model of F-actin decorated with *Dictyostelium* myosin S1. *Nature* **364,** 171–174 (1993).
29. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18,** 2714–2723 (1997).
30. Herm-Götz, A. *et al. Toxoplasma gondii* myosin A and its light chain: a fast, single-headed, plus-end-directed motor. *EMBO J.* **21,** 2149–2158 (2002).

Supplementary Information accompanies this paper on **www.nature.com/nature**.

Acknowledgements We thank S. Zimmermann for excellent technical assistance, R. Fedorov for providing Fig. 1, H. Faulstich for providing rhodaminephalloidin, C. Herrmann for the hGBP-1 cDNA, R. S. Goody, K. C. Holmes, M. A. Geeves, F. J. Kull, R. Maytum and D. P. Mulvihill for comments and discussions, and K. C. Holmes for continuous support. The work was supported by grants from the Deutsche Forschungsgemeinschaft (to D.J.M.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.J.M. (manstein@bpc.mh-hannover.de).

Crystal structure and mechanism of a bacterial fluorinating enzyme

Changjiang Dong[1], Fanglu Huang[2], Hai Deng[1], Christoph Schaffrath[1], Jonathan B. Spencer[2], David O'Hagan[1] & James H. Naismith[1]

[1]*Centre for Biomolecular Sciences, The University of St Andrews, Fife KY16 9ST, UK*
[2]*University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK*

Fluorine is the thirteenth most abundant element in the earth's crust, but fluoride concentrations in surface water are low and fluorinated metabolites are extremely rare[1,2]. The fluoride ion is a potent nucleophile in its desolvated state, but is tightly hydrated in water and effectively inert. Low availability and a lack of chemical reactivity have largely excluded fluoride from biochemistry: in particular, fluorine's high redox potential precludes the haloperoxidase-type mechanism[3,4] used in the metabolic incorporation of chloride and bromide ions. But fluorinated chemicals are growing in industrial importance, with applications in pharmaceuticals, agrochemicals and materials products[5–7]. Reactive fluorination reagents requiring specialist process technologies are needed in industry and, although biological catalysts for these processes are highly sought after, only one enzyme that can convert fluoride to organic fluorine has been described[8]. *Streptomyces cattleya* can form carbon–fluorine bonds[9] and must therefore have evolved an enzyme able to overcome the chemical challenges of using aqueous fluoride. Here we report the sequence and three-dimensional structure of the first native fluorination enzyme, 5′-fluoro-5′-deoxyadenosine synthase, from this organism. Both substrate and products have been observed bound to the enzyme, enabling us to propose a nucleophilic substitution mechanism for this biological fluorination reaction.

When grown in the presence of F^- ions, *S. cattleya* secretes fluoroacetate and 4-fluorothreonine, demonstrating its ability to biosynthesize organofluorine metabolites[9]. This organism contains an enzyme with a relative molecular mass (M_r) of 32,200 that has been shown to catalyse the formation of a C–F bond by combining *S*-adenosyl-L-methionine (SAM) and F^- to generate 5′-fluoro-5′-deoxyadenosine (5′-FDA) and L-methionine (ref. 10 and Fig. 1). Purification[11] and now overexpression of 5′-fluoro-5′-deoxyadenosine synthase (5′-FDAS) have allowed a fuller characterization of activity: the enzyme has a catalytic rate constant (k_{cat}) of

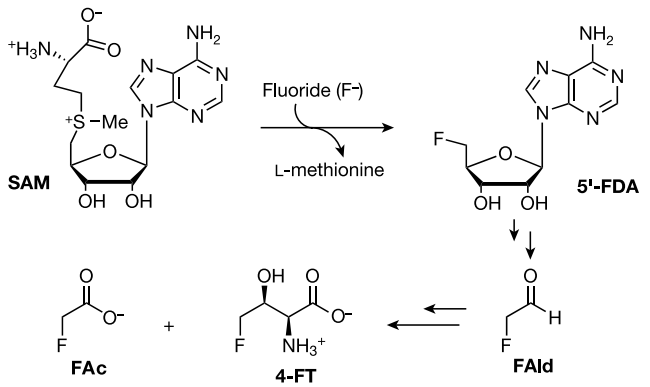

Figure 1 5′-FDAS from *S. cattleya* catalyses the formation of 5′-FDA from SAM and an F^- ion. 5′-FDA is the first-formed organofluorine metabolite, which is ultimately converted to fluoroacetate (FAc) and 4-fluorothreonine (4-FT) through fluoroacetaldehyde (FAld) by *S. cattleya*[30]. FAc is a toxin and 4-FT has antibiotic activity.

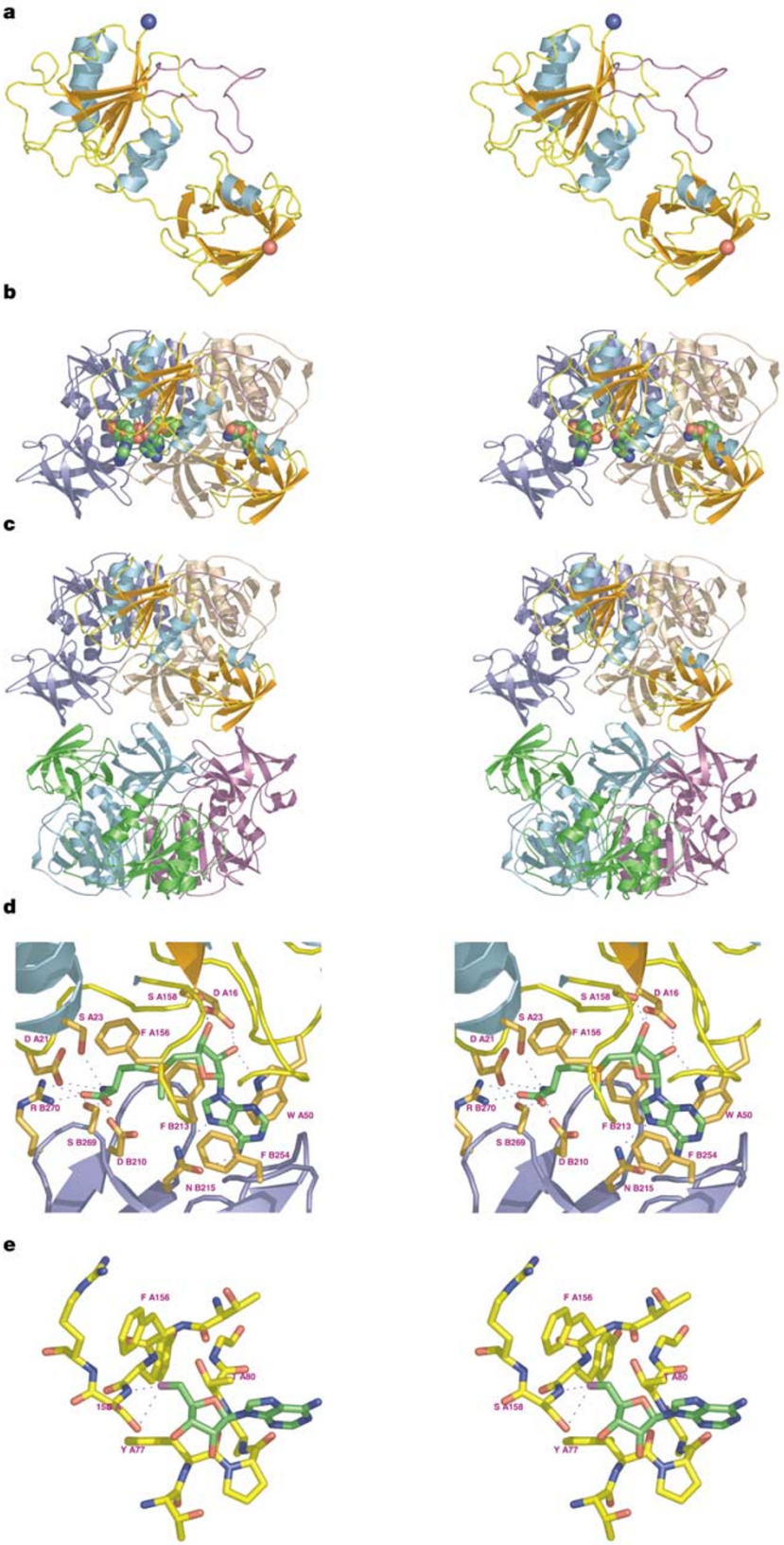

Figure 2 Structure of 5′-FDAS, drawn by PyMOL[31]. **a**, Stereoview of monomer unit (monomer A) in cartoon form, showing the N-terminal (top) and C-terminal (bottom) domains linked by a single polypeptide strand. The loop of 17 residues (98–114) unique to 5′-FDAS is shown in purple. **b**, The trimer formed by monomer A (coloured as in **a**), monomer B (blue) and monomer C (brown). Three molecules of SAM are shown in CPK format (carbon, green; oxygen, red; nitrogen, blue; sulphur, orange). The purple loop from the N-terminal domain sits in a cleft in the neighbouring N-terminal domain and in part forms the active site. SAM may act both as a substrate and in the structural integrity of the protein. **c**, The hexameric structure of 5′-FDAS shown in smoothed cartoon form. 5′-FDAS is composed of a dimer of trimers. **d**, Stereoview of the substrate complex. The colour scheme is the same as in **b**. Monomers A and B are both involved in extensive recognition of the substrate, which is completely buried. There is very little perturbation in the location of the adenosine, ribose or methionine between the product and substrate complexes, and the interactions are conserved. **e**, Stereoview of the 5′-FDA complex. The colour scheme is the same as in **b** and fluorine is shown in mauve. The organofluorine makes a hydrogen bond with amide nitrogen and a polar contact with the O_γ of Ser 158.

$0.07\ min^{-1}$, a Michaelis constant (K_m) for F^- of 2 mM, and a K_m for SAM of 74 μM. Mutants of two glycosidase enzymes, which create a very potent electrophile, can catalyse carbon–halogen (F^-, Cl^- and Br^-) bond formation; however, the fluorinated compounds are reactive and have not been isolated[8].

The gene encoding 5′-FDAS was cloned by degenerate polymerase chain reaction (PCR) on the basis of the partial protein sequence of the wild-type enzyme[11] using genomic DNA from *S. cattleya* as a template. Sequence information obtained from the degenerate PCRs was used to determine the complete coding sequence (*flA*) for the enzyme. The *flA* gene is 897 base pairs (bp) in size and encodes a protein containing 299 amino acids. The *flA* gene amplified by PCR was inserted into the pET28a(+) plasmid and expressed in *Escherichia coli* BL21(DE3). The M_r of the purified recombinant product of *flA* (Glu-Ser_2-His_6-Ser_2-Glu-Leu-Val-Pro-Arg-Glu-Ser-His-FlA) was 34,402, in agreement with electronspray ionization mass spectrometry (ESI-MS) data for the wild-type enzyme[11].

Crystallographic phases were determined by multiwavelength anomalous diffraction (MAD) from the selenomethionine (SeMet)-containing enzyme. These phases were used to determine the structure of the wild-type enzyme purified from *S. cattleya* to 1.9 Å. The refined structure contains residues 8–298 from the total of 299 residues: the first seven and the last residue are not visible. The monomer is organized as an amino-terminal (residues 8–180) and a smaller carboxy-terminal (residues 195–298) domain (Fig. 2a). A distinctive, long, extended loop of 15 residues connects the two domains, which are otherwise linked by few interactions. No structural homologues were found for either domain by the SSM[12] and the DALI[13] queries, and thus there is no indication of an alternative or additional function.

The N-terminal domain has a central seven-stranded β-sheet, which combines parallel and antiparallel strands sandwiched between α-helices (Fig. 2a). The C-terminal domain is composed of a five- and a four-stranded antiparallel β-sheet. The Pfam database[14] assigns the enzyme to family PF01887, and the structure provides the fold for this superfamily. Notably, no member of this superfamily has a known function. The asymmetric unit of the crystal contains a trimer (Fig. 2b): the three N-terminal domains are arranged around a three-fold axis and make intimate contacts with each other; the three C-terminal domains make contacts with their neighbouring N-terminal domain but have no contact with each other.

Sequence searching[15] identifies several homologues from various bacteria and archaea. Sequence alignment of these homologues shows that an extended loop (98–114) in the N-terminal domain that is involved in the trimer contacts is missing in all homologues (Fig. 2a, b). In a model of the enzyme without this loop, the buried surface area is reduced by 25% and 8 of the 13 hydrogen bonds at each of the three trimer interfaces are deleted. Thus, although the superfamily shares the same monomeric fold, 5′-FDAS may well have a unique quarternary structure. Crystallographic symmetry generates a hexamer (Fig. 2c), which is a dimer of the asymmetric unit trimer. Gel filtration indicates a hexamer[11].

The enzyme crystallized from *S. cattleya* contains a bound SAM molecule (Figs 2d and 3a), and SAM clearly does not dissociate readily from the protein during purification. SAM is bound at the interface between the C-terminal domain of one monomer and the N-terminal domain of the neighbouring monomer (Fig. 2b, d), and three molecules are bound by the trimer. The SAM molecule is completely buried by the protein on binding, which means that the protein must have an alternative open conformation during the turnover cycle. The recognition of SAM is highly specific (Figs 2d and 4). For ease of discussion, SAM can be decomposed into three components: the adenosine ring, the ribose ring and methionine. Each of these three components is recognized by a combination of hydrogen bonds and van der Waals contacts. All three components of SAM are recognized by both monomers, which indicates that the precise and possibly unique quarternary structure of the enzyme is crucial for substrate recognition and catalysis.

It seems reasonable that the extensive contacts between SAM and the protein would drive closure of the domains to form the enveloped binding site. Notably, SAM is bound in a high-energy conformation, and the C2–O2 and C3–O3 bonds of the ribose ring are found in an eclipsed conformation (torsion angle 1°). A search of the non-disordered well-refined entries in the small-molecule database[16] with a furanose ring fragment query found only 1 ring from 264 possibilities that has a torsion angle for O2–C2–C3–O3 within 10° of the angle seen here. An analysis of well-refined high-resolution (<2.0 Å) protein structures[17] shows that this torsion angle is always greater than 19° for the 15 SAM molecules in these structures. Molecular energy calculations indicate that this conformation relaxes on minimization with any of the common force fields (AMBER, MM2 and MM3)[18].

Observations from isotopic labelling[19] in whole-cell experiments indicate that the fluorination reaction occurs with a stereochemical 'inversion of configuration' at the C5′ carbon of SAM to generate

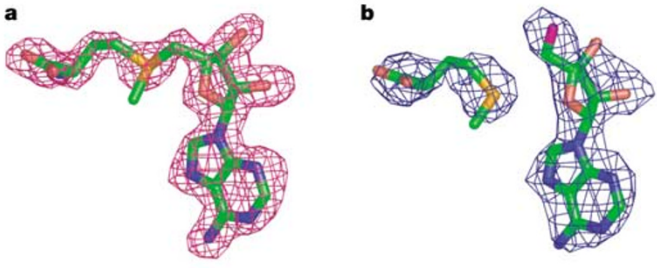

Figure 3 Fo–Fc electron density maps with phases calculated from models that do not include ligand. The colour scheme is the same as in Fig. 2. **a**, Map contoured at 3σ for SAM, shown in magenta chicken wire. **b**, Map contoured at 2.6σ for 5′-FDA and methionine, shown in blue chicken wire.

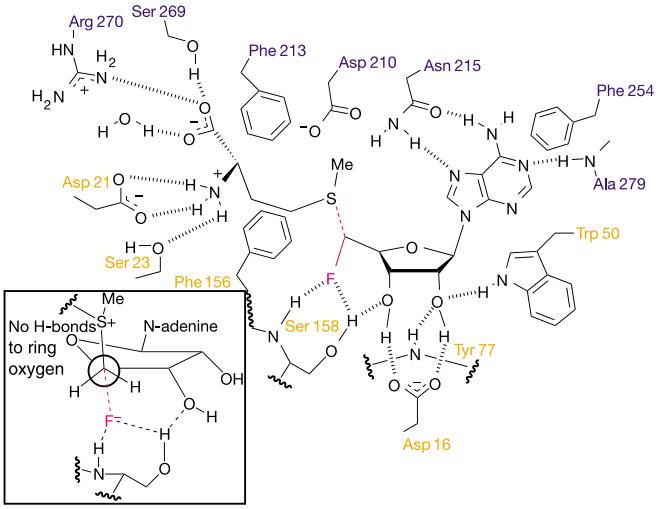

Figure 4 Representation of 5′-FDA and methionine bound to the enzyme, showing hydrogen-bonding to the fluoromethyl group from Ser 158, and the *anti* relationship between the C–F bond (red) and the disconnected C–S bond (dotted red) of SAM that is indicative of an S_N2 reaction course. Key residues are shown (monomer A, orange; monomer B, blue). Inset, the trajectory of the S_N2 and the conformation of ribose that minimizes negative stereoelectronic effects.

5′-FDA consistent with a nucleophilic substitution (S_N2) process. To define the molecular mechanism, we grew crystals of the enzyme with the two products of the reaction. This was done by incubating SeMet protein with 50 mM KF and 10 mM SAM for 4 h before crystallization. The resulting crystals diffracted to 2.7 Å and have essentially the same unit cell as the wild-type (SAM-bound) protein. The protein structure is identical to that observed for the SAM complex. It is immediately obvious from the electron density map that the SAM molecule is broken into two fragments (Fig. 3b). We built methionine and 5′-FDA into the density. The molecules are most clearly defined in the A subunit and reported distances relate to this subunit; however, the locations are essentially identical in all three subunits.

Methionine is bound in a manner very similar to that observed in the SAM complex; the only difference is in the torsion angles (maximum 16°) of the side chain, which results in a shift in the S and Cε atoms of 1.0 Å between the two structures. A similar observation is made when comparing 5′-FDA and SAM, where the positions of the adenosine and the ribose ring are unchanged. The planarity of the ribose ring refines to a different value in 5′-FDA, because the C2–O2 and C3–O3 bonds are no longer eclipsed; however, the resolution of this structure precludes detailed analysis. The C5′ atom of the ribose ring in the 5′-FDA complex has shifted by 0.8 Å from its position in the SAM complex. The electron density clearly shows that the F atom is located in a pocket defined by a main chain stretch from Phe 156 to Ser 158 and the side chains of Phe 156, Tyr 77, Thr 80 and Ser 158 (Fig. 2e).

There is a clear hydrogen bond (N–F, 3.1 Å) between the amide NH of Ser 158 and the F atom of the fluoromethyl group, which also makes a second polar contact with the side chain of Ser 158. It is well known that organic, bound fluorine is a relatively poor hydrogen-bonding acceptor[20]; however, the short contact here is very clear, suggesting an optimal NH–F contact of 2.0 Å (ref. 21). The pocket is unoccupied in the SAM complex and is hydrophobic in nature. There are no water molecules in the vicinity of the fluorine, which suggests that the F^- ion is fully dehydrated before C–F bond formation. The radius of this pocket is roughly 1.4–1.6 Å (assuming hydrogen atoms at their calculated positions with a van der Waals radius of 1.0 Å), which would preferentially select the F^- ion and possibly water over the other halogens. This would explain the observation that 5′-FDAS does not use other halogens.

Examination with the program HOLE[22] does not find a convincing channel that is wide enough to allow the F^- anion to diffuse into the active site when SAM is bound, indicating that F^- may be bound first. Comparing the product and substrate complexes gives a distance between F (5′-FDA) and C5′ (SAM) of 1.9 Å, aligns the C–F bond of the product (5′-FDA) *anti* (164 °) to the C5′–S bond of the substrate (SAM) and reveals only very small structural changes between substrate and complex. We regard this as extremely strong structural evidence for an S_N2 mechanism (Fig. 4), with these structures representing the 'before and after'. This is consistent with previous stereochemical data[19]. Implicit in this is the requisite that the F atom of the fluoromethyl group in 5′-FDA is bound at the F^--binding site and secured by the same hydrogen-bonding contacts used to anchor F^-.

It is our interpretation that the high-energy conformation of the ribose ring is optimal for the nucleophilic substitution reaction mediated by the enzyme. The pucker in the ring will act as a force on C5′, weakening the C5′–S bond. The conformation forces a F^- trajectory orthogonal to the C4–O4 bond, minimizing electrostatic repulsion during the S_N2 reaction. The inductive influence of the ring oxygen will strengthen the C5′–S bond (Fig. 4). O4 is prevented, however, by the ribose conformation from involvement in an anomeric interaction with the adenine ring and has no hydrogen bonds. Both these factors would substantially increase the inductive influence of O4 hindering the reaction. We note that O3 of SAM is hydrogen-bonded to Ser 158 OH, which contacts F^-. Therefore, the position of O3 may be important for binding F^-.

SAM is extensively recognized, and its co-purification with the protein suggests that it is tightly bound. We predict that the SAM binding must drive the dehydration of F^- as it pushed into its binding pocket. Full desolvation of F^- would require roughly 400 kJ mol^{-1}: a very high activation energy for the enzyme. Extrapolating from the fluoromethyl of the product structure, it seems that F^- makes polar contacts with the enzyme through a strong hydrogen bond to the backbone amide and a bifurcated hydrogen bond to the serine OH. It is clear that the enzyme uses these contacts to avoid having to pay the full free energy of desolvation. Thus 5′-FDAS specifically dehydrates the F^- ion, making it a very potent nucleophile. The F^- ion is collinear with the C–S bond of SAM, which the enzyme forces into a conformation designed to promote nucleophilic cleavage of the C–S bond. Although no

Table 1 **Experimental data**

Data collection	MAD (methionine and 5′-FDA)			SAM0.932
Wavelength (Å)	Peak 0.9786	Inflection 0.9783	Remote 0.8984	
Resolution (highest shell,Å)	53–2.7 (2.74–2.67)	53.5–3.1 (3.18–3.10)	45.9–3.1 (3.18–3.10)	65–1.9 (1.98–1.90)
Space group		$C222_1$		
Cell constants (Å) $\alpha = \beta = \gamma = 90°$		$a = 76.2$, $b = 129.8$, $c = 183.9$		$a = 75.9$, $b = 130.3$, $c = 183.4$
Unique reflections	26,303	17,635	16,925	67,204
Average redundancy	6.3 (3.5)	3.0 (2.8)	3.4 (3.3)	10 (7.7)
I/σ	8.5 (4.5)	11.9 (10.3)	14.1 (11.3)	4.8(1.7)
Completeness (%)	97 (81)	84 (72)	82 (69)	94(93)
Anom complete (%)*	97 (81)	84 (72)	82 (69)	–
R_{merge}†	6.4 (13)	4.2 (6.8)	3.7 (6.3)	11.0 (44.0)
Refinement				
R	17.6	–	–	16.7
R_{free}	23.9	–	–	21.7
r.m.s. deviation bonds (Å)/angles (°)	0.014/1.34	–	–	0.017/1.65
B-factor deviation bonds/angles ($Å^2$):	0.51 / 0.87	–	–	1.15/1.8
main chain side chains	1.35/2.28			2.53/3.85
Residues in Ramachandran core (%)	84	–	–	88
Protein atoms	6,660	–	–	6,660
Water atoms	158			719
Ligand atoms	84			81
Average B-factor ($Å^2$)	23.1	–	–	18.4
PDB accession code	1RQR	–	–	1RQP

*Anomalous completeness corresponds to the fraction of possible acentric reflections for which an anomalous difference has been measured.
†$R_{merge} = \Sigma_{hkl}\Sigma_i|I_i - \langle I\rangle|/\Sigma_{hkl}\Sigma_i\langle I\rangle$, where I_i is an intensity for the ith measurement of a reflection with indices hkl, and $<I>$ is the weighted mean of the reflection intensity.

structural data are available, in the most efficient glycosidase mutant capable of C–F bond formation, a hydrogen bond with serine has been proposed to direct a dehydrated F^- ion[8].

In summary, our data have revealed details of biological fluorination, whereby nature converts inorganic fluoride to organic fluorine. opening up an area of research focusing on the production of highly valuable organofluorine compounds by biotechnological means. □

Methods

Cloning and overexpression of FlA in *E. coli*

Three degenerate primers, pFN (5′-AAAAGGATCCGCSGCSAACWSSACSCGSCGSCC SATCATC-3′), PFI-1 (5′AAAAGAATTCSCGYTRGAARCCYGCRCCSSWRCCSGCCCA-3′) and PFI-2 (5′-NCKNSWRTARTARAANGTNGGYTCNGGYTGYTC-3′), were designed on the basis of the partial amino acid sequence of the wild-type enzyme[11] synthesized, and used for amplification of the corresponding gene fragments by PCR. All primers were synthesized by MWG-Biotech. The PCR products were subcloned into pUC18 and sequenced. We used two primers, pSCF1 (5′-TGACGTCCGGGCAGATGCTG TACATGAGCCCCTTGCACT-3′) and pSCF2 (5′-GAGGAGCACGGCTACCTGGAGGC GTACGAGGTCACCTCG-3′), in further gene walking to identify the rest of the *flA* gene sequence. *flA* was amplified by PCR from genomic DNA of *S. cattleya* using the primers pSCF3 (5′-GCAGGAGGAATTC*ATATG*GCTGCCAACAGC-3′) and pSCF4 (5′CCCTCACGCCG*CTCGAG*GGTACGTCGTCGC-3′). Three bases were changed (indicated by italics) in the primers to create *Nde*I and *Xho*I restriction sites (underlined) for cloning. The ATG in the *Nde*I site is the start codon of the gene. The pSCF4 primer is located 50-bp downstream of the stop codon of the *flA* gene. The PCR product was inserted into the pET28a(+) vector (Novagen) between the *Nde*I and *Xho*I sites to generate pET28-flA. This plasmid directs the expression of FlA with a small His-tag-containing peptide (Met-Glu-Ser_2-His_6-Ser_2-Glu-Leu-Val-Pro-Arg-Glu-Ser-His) fused to the N-terminal of the enzyme.

E. coli BL21(DE3) transformed with pET28-flA was grown in Luria broth containing 100 μg ml^{-1} kanamycin at 37 °C until an absorbance of 0.4–0.6 at 600 nm was reached. We induced overexpression of FlA by adding isopropylthiogalactoside (IPTG) to 0.2 mM and continued the incubation at 16 °C overnight. Cells were collected and lysed by sonication. After centrifugation, the cell lysate was applied onto a column packed with Ni^{2+}-charged His-Bind resin (Novagen). Recombinant protein bound on the resin was eluted with 20 mM Tris-HCl (pH 7.9), 250 mM imidazole, 0.5 M NaCl and 10% glycerol. The protein was further purified by gel filtration fast protein liquid chromatography (FPLC) on a Superdex S-200 (HR16/60) column (PharmaciaBiotech) and concentrated using a Vivaspin concentrator (Vivascience). We used Bradford reagent (BioRad) to measure protein concentrations. The M_r of the protein was determined by polyacrylamide gel electrophoresis, gel-filtration FPLC and ESI-MS analysis.

Assay of the recombinant fluorination enzyme

In a total reaction volume of 100 μl, 15 μg of FlA was incubated with 0.8 mM SAM, 10 mM NaF and 50 mM Tris-HCl (pH 7.9) at 37 °C for 3 h. The reaction product 5′-FDA was confirmed by liquid chromatograpy mass spectrometry (LC-MS) using synthetic 5′-FDA[10] as a standard.

Crystallography

Full details of the purification[11] and crystallization of the enzyme from *S. cattleya* have been published elsewhere[23]. The 1.9-Å data set was recorded on ID14-2 of the European Synchrotron Radiation Facility (ESRF). The protein crystals were thought to be wild type, but on structure elucidation it became clear that SAM was bound. SeMet-labelled protein was produced by overexpression in *E. coli* by the methionine inhibition protocol[24] and confirmed by mass spectroscopy. Crystals of 5′-FDA and methionine complex were obtained for SeMet protein under similar conditions to those used for the 'native' SAM crystals, except that the enzyme was first incubated with 50 mM KF and 10 mM SAM for 4 h.

Data were collected at three wavelengths on BM14 UK at the ESRF and reduced with MOSFLM/SCALA[25]; full details are given in Table 1. SOLVE[26] located 18 Se positions and generated phases. Automatic interpretation of electron density was done by combining the phases from the SeMet crystals with the isomorphous 1.9 Å in Arp/wARP[27]. This procedure led to an essentially complete chain trace of the monomer. Both structures were refined by using REFMAC[28] and statistics are given in Table 1. Dictionaries were derived from the PRODG server[29]; however, all torsion restraints were removed from the ribose ring to ensure a non-dictionary-biased conformation was obtained from refinement. Experimental data and structures have been deposited in the relevant data banks.

Received 2 October; accepted 9 December 2003; doi:10.1038/nature02280.

1. O'Hagan, D. & Harper, D. B. Fluorine-containing natural products. *J. Chem.* **100,** 127–133 (1999).
2. Xu, X.-H. *et al.* 5-Fluorouracil derivatives from the sponge *Phakellia fusca*. *J. Nat. Prod.* **66,** 285–288 (2003).
3. vanPee, K. H. Biosynthesis of halogenated metabolites by bacteria. *Annu. Rev. Microbiol.* **50,** 375–399 (1996).
4. Littlechild, J. Haloperoxidases and their role in biotransformation reactions. *Curr. Opin. Chem. Biol.* **3,** 28–34 (1999).
5. Sandford, G. Organofluorine chemistry. *Phil. Trans. R. Soc. Lond. A* **358,** 455–471 (2000).
6. Mann, J. Modern methods for the introduction of fluorine into organic molecules—an approach to compounds with altered chemical and biological activities. *Chem. Soc. Rev.* **16,** 381–436 (1987).
7. Hutchinson, J. & Sandford, G. Elemental fluorine in organic chemistry. *Top. Curr. Chem.* **193,** 1–43 (1997).
8. Zechel, D. L. *et al.* Enzymatic synthesis of carbon–fluorine bonds. *J. Am. Chem. Soc.* **123,** 4350–4351 (2001).
9. Sanada, M. *et al.* Biosynthesis of fluorothreonine and fluoroacetic acid by the thienamycin producer, *Streptomyces cattleya*. *J. Antibiotics* **141,** 259–265 (1986).
10. O'Hagan, D., Schaffrath, C., Cobb, S. L., Hamilton, J. T. G. & Murphy, C. D. Enzyme catalysed organofluorine synthesis. *Nature* **416,** 279 (2002).
11. Schaffrath, C., Deng, H. & O'Hagan, D. Isolation and characterisation of 5′-fluorodeoxyadenosine synthetase, a fluorination enzyme from *Streptomyces cattleya*. *FEBS Lett.* **547,** 111–114 (2003).
12. Boutselakis, H. *et al.* E-MSD: the European Bioinformatics Institute Macromolecular Structure Database. *Nucleic Acids Res.* **31,** 458–462 (2003).
13. Holm, L. & Sander, C. Protein folds and families: sequence and structure alignments. *Nucleic Acids Res.* **27,** 244–247 (1999).
14. Bateman, A. *et al.* The Pfam Protein Families Database. *Nucleic Acids Res.* **30,** 276–280 (2002).
15. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 (1997).
16. Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. B* **58,** 380–388 (2002).
17. Berman, H. M. *et al.* The Protein Data Bank. *Acta Crystallogr. D* **58,** 899–907 (2002).
18. Kollman, P. A. *et al.* Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33,** 889–897 (2000).
19. O'Hagan, D. *et. al.* An assay for the enantiomeric assay of [$^2H_1$]-fluoroacetic acid: Insight into the stereochemical course of fluorination during fluorometabolite biosynthesis in *Streptomyces cattleya*. *J. Am. Chem. Soc.* **125,** 379–387 (2003).
20. Dunitz, J. & Taylor, R. Organic fluorine hardly ever accepts hydrogen bonds. *Chem. Eur. J.* **3,** 89–98 (1997).
21. Howard, J. A. K., Hoy, J. V., O'Hagan, D. & Smith, G. T. How good is fluorine as a hydrogen bond acceptor? *Tetrahedron* **38,** 12613–12622 (1996).
22. Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. P. HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph. Model* **14,** 354–360 (1996).
23. Dong, C. *et al* Crystallization and X-ray diffraction of the 5′-fluoro-5′-deoxyadenosine synthase, a fluorination enzyme from *Streptomyces cattleya*. *Acta Crystallogr. D* **60,** 760–763 (2003).
24. Doublie, S. Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* **276,** 523–530 (1997).
25. Bailey, S. The CCP4 Suite—programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
26. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55,** 849–861 (1999).
27. Morris, R. J., Perrakis, A. & Lamzin, V. S. ARP/wARP's model-building algorithms. I. The main chain. *Acta Crystallogr. D* **58,** 968–975 (2002).
28. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D* **55,** 247–255 (1999).
29. van Aalten, D. M. F. *et al* PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J. Comput. Aided Mol. Des.* **10,** 255–262 (1996).
30. Schaffrath, C., Cobb, S. L. & O'Hagan, D. Cell-free biosynthesis of fluoroacetate and 4-fluorothreonine in *Streptomyces cattleya*. *Angew. Chem. Int. Edn Engl.* **41,** 3913–3915 (2002).
31. DeLano, W. L., *The PyMOL Molecular Graphics System* ⟨http://www.pymol.org/⟩ (2003).

Acknowledgements We thank M. Dorward for technical assistance and G. Leonard for help with data collection. J.H.N. is a Biotechnology and Biological Sciences Research Council (BBSRC) Career Development Fellow; H.D. was supported by the BBSRC; D.O.H. thanks the BBSRC for financial support; J.H.N. thanks the Wellcome Trust.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.H.N. (naismith@st-andrews.ac.uk). The nucleotide sequence of *flA* reported here is deposited in the DDBJ, EMBL and GenBank nucleotide sequence databases under accession number AJ581748.