Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

The structure of vanin 1: a key enzyme linking metabolic disease and inflammation

Ykelien L. Boersma,[a]‡ Janet Newman,[b]‡ Timothy E. Adams,[b] Nathan Cowieson,[c] Guy Krippner,[d] Kiymet Bozaoglu[d] and Thomas S. Peat[b]*

[a]Department of Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, A. Deusinglaan 1, 9713 AV Groningen, The Netherlands, [b]CSIRO Biosciences Program, 343 Royal Parade, Parkville, VIC 3052, Australia, [c]Australian Synchrotron, 800 Blackburn Road, Clayton, VIC 3168, Australia, and [d]Baker IDI, 75 Commercial Road, Melbourne, VIC 3004, Australia

‡ These authors contributed equally to this work.

Correspondence e-mail: tom.peat@csiro.au

Received 16 July 2014
Accepted 16 October 2014

PDB references: vanin 1, native, 4cyf; trigonal form, 4cyy; ligand-bound, 4cyg

Although part of the coenzyme A pathway, vanin 1 (also known as pantetheinase) sits on the cell surface of many cell types as an ectoenzyme, catalyzing the breakdown of pantetheine to pantothenic acid (vitamin B_5) and cysteamine, a strong reducing agent. Vanin 1 was initially discovered as a protein involved in the homing of leukocytes to the thymus. Numerous studies have shown that vanin 1 is involved in inflammation, and more recent studies have shown a key role in metabolic disease. Here, the X-ray crystal structure of human vanin 1 at 2.25 Å resolution is presented, which is the first reported structure from the vanin family, as well as a crystal structure of vanin 1 bound to a specific inhibitor. These structures illuminate how vanin 1 can mediate its biological roles by way of both enzymatic activity and protein–protein interactions. Furthermore, it sheds light on how the enzymatic activity is regulated by a novel allosteric mechanism at a domain interface.

1. Introduction

Inflammation and cellular stress are linked to a variety of metabolic pathologies that include obesity and type 2 diabetes (Hotamisligil, 2010). The impact of obesity alone as a global health epidemic is best captured by the fact that in 2013 about two billion people worldwide were obese (Ng *et al.*, 2014). In this context, it is critical that we obtain a deeper understanding of the interplay between the pathways that govern inflammatory and stress responses at a cellular level in order to develop more effective strategies to manage the spectrum of diseases that characterize what is known as metabolic syndrome (Huang, 2009).

Vanin 1, or pantetheinase, is a glycosylphosphatidylinositol (GPI)-anchored ectoenzyme that has been shown to be widely expressed in many tissues, including the liver, kidney and gut (Pitari *et al.*, 2000). Pantetheinase catalyzes the hydrolysis of pantetheine to pantothenic acid (vitamin B_5) and cysteamine, which together with cystamine participate in the regulation of cellular pathways involved in oxidative stress and inflammation (Maras *et al.*, 1999). Germline deletion of vanin 1 in mice results in an absence of tissue cysteamine (Pitari *et al.*, 2000). As such, vanin 1 has been shown to both protect tissues from and sensitize tissues to damage in different disease settings (Berruyer *et al.*, 2006; Jansen *et al.*, 2009; Kaskow *et al.*, 2012). Mice that lack vanin 1 display resistance to oxidative tissue damage induced by whole-body γ-irradiation, with a reduction in apoptosis and inflammation (Berruyer *et al.*, 2004). The absence of cellular cysteamine was associated with elevated levels of the potent antioxidant glutathione.

By a different mechanism, germline deletion of vanin 1 protects mice from chemically induced ulcerative colitis by preventing the vanin-mediated release of pro-inflammatory cytokines from the intestinal epithelium (Martin *et al.*, 2004; Berruyer *et al.*, 2006). A key target of vanin 1 is peroxisome proliferator-activated receptor γ (PPAR-γ), a potent regulator of anti-inflammatory activity in the intestine; vanin 1 regulates the level of expression of PPAR-γ and antagonizes ligand-induced activation of this receptor in epithelial cells, resulting in the release of a number of pro-inflammatory cytokines (Berruyer *et al.*, 2006).

Vanin 1-deficient mice are not only resistant to gut inflammation but also show increased resistance to stress (Berruyer *et al.*, 2004, 2006; Martin *et al.*, 2004; Gensollen *et al.*, 2013). Both vanin 1 and vanin 3 were upregulated (at both the mRNA and the protein levels) by psoriasis-associated pro-inflammatory cytokines in human keratinocyte cultures (Jansen *et al.*, 2009). One way in which vanin 1 promotes this inflammatory reaction and intestinal injury is by decreasing the activity of γ-glutamylcysteine synthetase and thereby decreasing the stores of reduced glutathione (Berruyer *et al.*, 2004; Martin *et al.*, 2004). Additionally, vanin 1 is strongly induced by fasting: Chen and coworkers have shown that it activates gluconeogenesis and increases glucose output in hepatic cells and that this is mediated through the Akt signalling pathway (Chen *et al.*, 2014).

In contrast, vanin 1 exerts a cytoprotective effect on pancreatic islet beta cells when treated with streptozotocin (STC), and its absence in non-obese diabetic (NOD) mice is associated with an increase in the incidence of diabetes and the severity of inflammatory insulitis (Roisin-Bouffay *et al.*, 2008). Cystamine prevented STC-induced islet-cell death *in vitro* and reduced the incidence of diabetes in two mouse models when administered *in vivo*. More recently, several studies have shed light on an emerging role for vanin 1 in the regulation of a number of key metabolic pathways. It was observed that fasting-induced elevation of hepatic vanin 1 mRNA required PPAR-α, a major transcriptional regulator of hepatic lipid metabolism, and could be induced by the administration of both fatty acids and agonists of PPAR-α (Rakhshandehroo *et al.*, 2010). In fact, the vanin 1 gene is the most highly up-regulated gene in the liver in response to PPAR-α manipulation (van Diepen *et al.*, 2014). Inhibition or deletion of vanin 1 enhanced the increase in hepatic triglycerides and was associated with the appearance of hepatic steatosis.

Thus, vanin 1 is of key importance both in metabolism and in the inflammatory response, and could potentially form a new therapeutic target. Whilst there has been investigation of vanin 1 and its possible role in different disease states, there are still gaps in our understanding of what vanin 1 does and how it functions. To better understand the role played by this enzyme in tissue homeostasis and metabolism, we have determined the crystal structure of human vanin 1 itself as well as with a synthetic pantetheine analogue bound. We show how the enzymatic activity can be regulated by a novel allosteric mechanism at a domain interface.

2. Materials and methods

2.1. Expression, purification and crystallization

Codon-optimized human vanin 1 cDNA was obtained from Qiagen. There is a single change from the native consensus sequence of human vanin 1 at position 26: instead of a threonine residue, the Qiagen-supplied sequence has an isoleucine residue. This is a natural SNP variant (rs2294757) found in the UniProt database. PCR was used to generate a template encoding amino acids 22–493, deleting the native signal sequence and the C-terminal GPI-anchor sequence. The insert was cloned downstream of the mouse interleukin-3 signal peptide in the mammalian expression vector pAPEX-3P (Evans *et al.*, 1995) and a FLAG tag (DYDDDDK). Transient expression was performed in suspension-adapted Freestyle 293 cells (Life Technologies) grown in Freestyle 293 serum-free medium (Life Technologies) in shaker flasks by the addition of 1 mg vector DNA complexed with 25 kDa polyethylenimine (Polysciences) to 1 l of cells. 5 µ*M* kifunensine (Sapphire Bioscience) was added to the cell-culture medium to reduce the amount of glycosylation on the protein. After 9 d, the cell supernatants were harvested and FLAG-tagged soluble protein was purified by immunoaffinity chromatography (Elleman *et al.*, 2001). The partially purified protein was loaded onto a Superdex 200 16/60 column pre-equilibrated with Tris-buffered saline solution (TBSA; 25 m*M* Tris pH 8.0, 140 m*M* NaCl, 2.5 m*M* KCl, 0.02% NaN_3), and the fractions containing vanin 1 (∼15.0 ml) were pooled and concentrated to 360 µl, giving a concentration of 17.7 mg ml^{-1}. The protein was transferred by dialysis into a lower pH formulation (50 m*M* bis-tris pH 6.5, 50 m*M* NaCl), which was shown by differential scanning fluorimetry (DSF) to increase the thermal stability of the protein by up to 15°C (see below).

Mutations were made using the megaprimer method to obtain the E79A, K178A, E249Q and E439Q mutants. All restriction and modifying enzymes were purchased from NEB. Positive colonies were subjected to colony PCR (cPCR) using GoTaqGreen Master Mix (Promega). All clones were subsequently sequenced (Micromon, Monash University). For small-scale protein production, adherent HEK293T cells were cultured in 1 ml DMEM (Life Technologies) supplemented with 10% fetal calf serum in a humidified incubator with 5% CO_2. Cells were transiently transfected with 1 µg DNA using FuGene6 (Promega) as the transfection reagent and were then grown for 3 d before harvesting the supernatant. All point mutants were soluble and expressed well as seen on Western blots using a horseradish peroxidase-conjugated FLAG M2 antibody (Sigma) for detection (Supplementary Fig. S1[1]). Large-scale production of the E249Q and E439Q mutants was performed as described above for the native protein.

Crystallization experiments set up with the protein in TBSA did not yield crystals. Crystallization experiments were set up at both 293 and 281 K with kifunensine-treated vanin 1 protein in 50 m*M* bis-tris pH 6.5, 50 m*M* NaCl and these

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: MH5151).

Table 1
Data-collection, phasing and refinement (SIRAS) statistics.

Values in parentheses are for the highest resolution shell.

	Vanin 1 + NaI	Native vanin 1	Vanin 1 + RR6	Vanin 1, trigonal
PDB code		4cyf	4cyg	4cyy
Data collection				
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P3_22$
Unit-cell parameters				
a (Å)	119.8	119.9	120.1	101.9
$b = c$ (Å)	222.8	221.9	221.9	133.7
$\alpha = \beta$ (°)	90	90	90	90
γ (°)	90	90	90	120
Resolution (Å)	2.90 (3.06–2.90)	2.25 (2.37–2.25)	2.30 (2.42–2.30)	2.89 (3.04–2.89)
$R_{\rm merge}$†	0.190 (0.755)	0.142 (1.18)	0.153 (1.24)	0.210 (1.12)
$R_{\rm p.i.m.}$‡	0.043 (0.226)	0.038 (0.316)	0.041 (0.326)	0.054 (0.286)
$CC_{1/2}$	0.997 (0.851)	0.998 (0.806)	0.998 (0.762)	0.997 (0.931)
$\langle I/\sigma(I)\rangle$	23.6 (7.8)	15.2 (2.7)	15.6 (2.6)	14.5 (2.7)
Completeness (%)	100 (100)	99.8 (99.3)	99.6 (99.3)	99.7 (97.9)
Multiplicity	43.0 (43.7)	14.8 (14.5)	14.9 (15.2)	16.3 (16.1)
Refinement				
Resolution (Å)		2.25	2.30	2.89
Unique reflections		73353 (11026)	68703 (10365)	18526 (2621)
$R_{\rm work}/R_{\rm free}$§ (%)		19.5/22.2	18.6/21.4	17.4/21.9
No. of atoms				
Total		7688	7820	3710
Protein		7448	7457	3699
Ligand			42	
Water		240	307	11
B factors (Å^2)				
Overall		40.4	36.9	62.3
Protein		41.0	37.3	62.3
Ligand			49.5	
Water		36.3	34.2	46.4
R.m.s deviations¶				
Bond lengths (Å)		0.007	0.006	0.010
Bond angles (°)		1.253	1.240	1.382

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{\rm p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$. § $R_{\rm work} = \sum_{hkl}\big||F_{\rm obs}| - |F_{\rm calc}|\big|/\sum_{hkl}|F_{\rm obs}|$ and is calculated using all data; $R_{\rm free}$ is the R factor based on 5% of the data excluded from refinement. ¶ R.m.s.d. is the root-mean-square deviation from ideal values (Engh & Huber, 1991).

experiments gave initial crystals after six weeks using the PACT crystallization screen (Newman *et al.*, 2005). The conditions were refined and the best crystals were grown at 281 K in 21%(*w*/*v*) PEG 1500, 10%(*v*/*v*) malate–MES–Tris buffer at pH 6.3; crystals also grew with protein that had 1 m*M* pantothenic acid added to it and which was treated *in situ* with PNGase. The protein was at 12 mg ml^{-1} in the bis-tris/NaCl formulation, and droplets were set up in sitting-drop vapor-diffusion experiments consisting of 200 nl protein solution and 200 nl reservoir solution. Crystals formed over the course of about 40 d, reaching full size by day 55, with the largest adopting a long rod-like shape with dimensions of ~20 × 50 × 350 µm (Supplementary Fig. S2). Crystals grew more reliably with microseeding, but still took over a month to form. A second crystal form was later found that grew in similar conditions [281 K in 25% (*w*/*v*) PEG 1500, 10%(*v*/*v*) succinate–phosphate–glycine buffer at pH 6.0] and belonged to space group $P3_22$.

2.2. Vanin structure solution

Crystals were harvested into mylar loops (MiTeGen) and cryocooled by rapid plunging into liquid nitrogen in a cold room (277 K). No additional cryoprotectant was used. Two crystals were harvested from a single droplet, and 1 µl of a 5% solution of saturated NaI (about 10 *M*) in 25% PEG 1500 was then added to the drop. Two more crystals were harvested (within 1 min of NaI addition) and cryocooled as above. Crystals harvested at room temperature showed sub-optimal diffraction patterns or melted during the harvesting process.

Data were collected from these crystals on the MX2 (microfocus) beamline of the Australian Synchrotron (Table 1). One of the crystals harvested initially provided a native data set to 2.54 Å resolution, and one of the NaI-soaked crystals proved to be a well diffracting, isomorphous iodide derivative. Starting at one end of the NaI-soaked, long rod-like crystal, several successive 180° data sets were collected to give four data sets in total. The data were indexed using *XDS* (Kabsch, 2010) and scaled using *SCALA* (Evans, 2006). The statistics for the first 180° data set from the NaI-soaked crystal showed that there was anomalous signal for the iodide at the 8500 eV (1.459 Å) wavelength used for data collection. The combination of iodide data sets 1 through 3 together with a native data set were used for phasing (SIRAS) using *AutoRickshaw* (Panjikar *et al.*, 2005). This gave robust phasing, and a good starting model was obtained using *Buccaneer* (Cowtan, 2006). Within *AutoRickshaw*, $F_{\rm A}$ values were calculated using *SHELXC* (Sheldrick, 2010), the heavy-atom sites were found using *SHELXD* (Schneider & Sheldrick, 2002) and the correct hand of the substructure was determined using *ABS* (Hao, 2004) and *SHELXE* (Sheldrick, 2002). Based on the analysis of the hand of the substructure, the space group of the data was determined to be $P4_32_12$. The occupancy of all substructure atoms was refined using *BP*3 (Pannu *et al.*, 2003). The twofold noncrystallographic symmetry (NCS) operator was found using *RESOLVE* (Terwilliger, 2000). Density modification, phase extension and NCS averaging were performed using *DM* (Cowtan & Main, 1996). *Phaser* (McCoy *et al.*, 2007) was used with the partial model from *AutoRickshaw* (2.54 Å) to obtain a solution from molecular replacement into the best native data set (2.25 Å resolution) and this model was rebuilt by hand using *Coot* (Emsley *et al.*, 2010) and further refined using *REFMAC* (Murshudov *et al.*, 2011). The structure with the compound RR6 was obtained by diluting a 50 m*M* solution of RR6 to 1 m*M* with the reservoir solution before adding 1 µl to the drop containing crystals. This was allowed to soak for approximately 24 h before harvesting the crystals for data

collection. The statistics from *MolProbity* (Chen *et al.*, 2010) show that the native structure (PDB entry 4cyf) has 15 poor rotamers and three Ramachandran outliers out of a total of 912 residues and a *MolProbity* score of 1.13 (100th percentile), while the RR6-bound structure (PDB entry 4cyg) has 11 poor rotamers and four Ramachadran outliers out of a total of 912 residues and a *MolProbity* score of 1.13 (100th percentile), and the structure in space group $P3_22$ has five poor rotamers and no Ramachandran outliers out of 444 residues and a *MolProbity* score of 1.79 (100th percentile). The differences between the two space groups are small: there is missing density in the $P3_22$ crystal form for one loop (between residues 112 and 116), which has higher *B* factors in the tetragonal space-group structure, confirming that it is a mobile loop, and all five Asn residues in the $P3_22$ space-group structure show saccharide density for glycosylation, whereas only four of the sites show this extra density in the tetragonal space-group structure. As the cells were treated with kifunensine, the glycosylation is limited to two *N*-acetylglucosamine residues near the Asn residue and about ten mannose saccharides beyond this. Only the *N*-acetylglucosamine residues are seen in the electron density of the structure.

Table 2
Experimental parameters relating to the collection and analysis of SAXS data.

AS, Australian Synchrotron.

	Wild type	E249Q	E439Q
Data-collection parameters			
Instrument	AS SAXS/WAXS beamline	AS SAXS/WAXS beamline	AS SAXS/WAXS beamline
Beam geometry	120 µm point source	120 µm point source	120 µm point source
Wavelength (Å)	1.127	1.127	1.127
q range (Å^{-1})	0.01–0.55	0.01–0.55	0.01–0.55
Exposures	10 × 1 s	10 × 1 s	10 × 1 s
Flow rate (µl s^{-1})	4	4	4
Temperature (°C)	25	25	25
Structural parameters			
$I(0)$ [from $P(r)$] (cm^{-1})	0.166	0.14	0.145
R_g [from $P(r)$] (Å)	28.28	29.58	28.97
$I(0)$ (from Guinier) (cm^{-1})	0.168	0.148	0.145
D_{max} (Å)	97.69	103.11	104.33
R_g (from Guinier) (Å)	28.136	28.94	28.136
Porod volume estimate (Å^3)	106695	112448	114399
Dry volume calculated from sequence (Å^3)	62359	62359	62359
Molecular-mass determination			
Partial specific volume (cm^3 g^{-1})	0.726	0.726	0.726
Contrast ($\Delta\rho \times 10^{30}$ cm^{-2})	2.998	2.998	2.998
Molecular mass M_r [from $I(0)$]	65.7	70.3	64.5
Calculated monomeric M_r from sequence (kDa)	65.4	65.4	65.4
Software employed			
Primary data reduction	*ScatterBrain* (AS)	*ScatterBrain* (AS)	*ScatterBrain* (AS)
Data processing	*PRIMUS*/*GNOM*	*PRIMUS*/*GNOM*	*PRIMUS*/*GNOM*
Computation of model intensities	*CRYSOL*	*CRYSOL*	*CRYSOL*

2.3. Vaninfluor assay

Vanin 1-mediated pantetheinase activity was measured using a fluorescent substrate, *N*-(4-methyl-2-oxo-2*H*-chromen-7-yl)-D-pantothenamide, which we call 'vaninfluor'. Hydrolysis of vaninfluor by vanin releases 7-amino-4-methylcoumarin, which emits at a wavelength of 430 nm, as has been shown previously (Ruan *et al.*, 2010). To determine the enzymatic activity of the generated mutants and recombinant vanin 1 protein itself, 5 µl cell supernatant containing secreted protein or purified protein was added to a 96-well plate in triplicate and was incubated in 100 µl of 10 µ*M* vaninfluor substrate in 0.5 *M* potassium phosphate buffer pH 8.0. Three wells each containing 0.5 *M* potassium phosphate buffer alone and 10 µ*M* Vaninfluor reagent alone were also added to determine the background vaninfluor activity. The plate was then inserted into a Perkin Elmer Victor 3V 1420 Multi-label counter (Perkin Elmer, Massachusetts, USA). Vanin 1 activity was determined *via* fluorescence with excitation at 350 nm and emission at 430 nm at 310 K for optimal enzymatic activity. Activity was measured every 5 min over 1.5 h.

2.4. RR6 inhibitor synthesis

The RR6 inhibitor was synthesized by Anthem Biosciences, Bangalore using the method described in Jansen *et al.* (2013).

2.5. Differential scanning fluorimetry

Stability studies were performed on native vanin 1 and the E249Q and E439Q mutants. In all cases, a series of different pH values/buffers/salt concentrations were tested in triplicate (Seabrook & Newman, 2013). All of the proteins were well behaved during protein purification and could easily be concentrated to over 15 mg ml^{-1} (Supplementary Fig. S3).

2.6. SAXS measurements and modelling

SAXS measurements were made on the SAXS-WAXS beamline at the Australian Synchrotron, Melbourne, Australia (Kirby *et al.*, 2013; Table 2). For each SAXS measurement, 10 × 1 s exposures were measured and averaged together after verifying that there was no evidence of radiation damage (systematic change in the shape of the scattering curves as a function of exposure time). During data collection the sample was flowing through a 1.5 mm quartz capillary at a rate of 4 µl s^{-1} to further control radiation damage. Measurements were performed on a protein dilution series from 0.3 to 13 mg ml^{-1} in 10 m*M* sodium potassium phosphate pH 7.7, 137 m*M* NaCl, 2.5 m*M* KCl. The molecular weights of the scattering species were estimated from the total forward scatter of the SAXS measurements, which was normalized by comparison to water scatter and with reference to the measured protein concentrations. The partial specific volume

and scattering length density were calculated using *MULCH* (Whitten *et al.*, 2008). The monomeric state of the protein was inferred by comparison of the theoretical molecular weight of the protein sequence with the calculated molecular weight from the SAXS experiment. A 1.6 m camera was used with an X-ray energy of 11 KeV, giving a q range from 0.01 to 0.5 Å^{-1}. Data were collected on a PILATUS 1M detector (Dectris) and averaging of images, subtraction of blanks and radial integration was performed using the *ScatterBrain* beamline-control software (Australian Synchrotron). Measurements were made at 298 K. Interpretation of the SAXS data with respect to the crystal structure was performed using *CRYSOL* (Svergun *et al.*, 1995). The radius of gyration (R_g), total forward scatter [$I(0)$] and $P(r)$ functions were derived using the automated functions in *PRIMUS* (Konarev *et al.*, 2003) without manual intervention (Supplementary Fig. S4).

3. Results

We determined the X-ray crystal structure of the human vanin 1 protein using single isomorphous replacement with anomalous scattering to 2.25 Å resolution (Table 1). The X-ray data showed clear density for 462 residues (22–483), only lacking the signal peptide and the last eight residues of the mature protein. The protein has two domains (Fig. 1): a base domain and a nitrilase domain with pantetheinase activity (Brenner, 2002; Barglow *et al.*, 2008; see below). These data reveal that five of the six predicted glycosylation sites (Asn at positions 38, 130, 200, 315 and 353) have saccharide density (Fig. 2). Since the protein was produced from mammalian cell culture, it would be expected to have the native post-translational modifications associated with the naturally occurring protein.

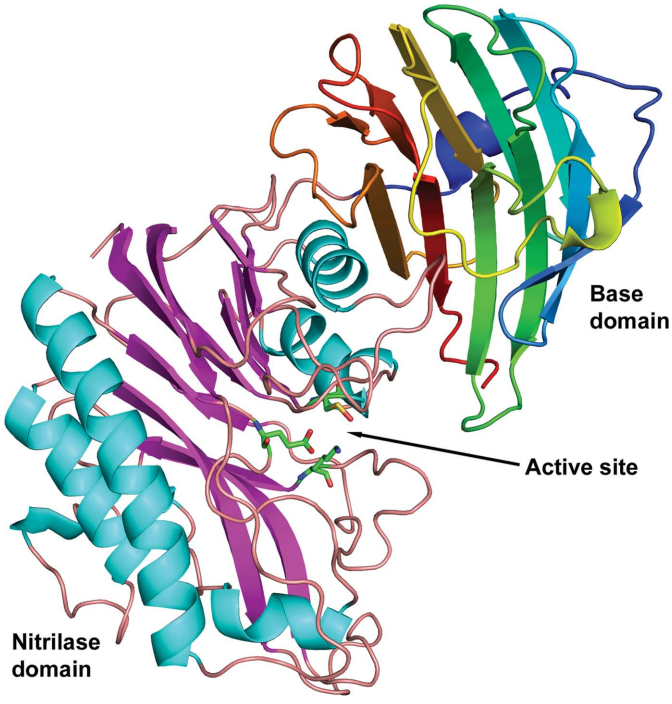


Figure 1
The structure of vanin 1. The nitrilase enzymatic domain of vanin 1 is represented by cyan α-helices and magenta β-strands and extends to Val313. The catalytic residues Glu79, Lys178 and Cys211 are shown in stick representation, showing that the active site is situated almost entirely in this domain. The base domain, which consists almost entirely of β-sheet secondary structure, is shown as a rainbow from dark blue (the N-terminus of this domain, starting at Val314) to dark red (the C-terminus). The interface of the two domains is made up of β-strands from the base domain and a single helix and loop from the nitrilase enzymatic domain.

We also determined the structure of vanin 1 bound to a recently described inhibitor, RR6 (Jansen *et al.*, 2013; Fig. 3). This shows the inhibitor in the active site and covalently bound to the active-site cysteine, residue 211 (Brenner, 2002; Rommelaere *et al.*, 2013). The inhibitor was soaked into the crystal, showing that the enzyme is still at least partially active in the crystalline environment. Two phenylalanine residues, 182 and 212, situated at one end of the active site show better electron density and lower B factors in the presence of the inhibitor. These phenylalanine residues, along with Phe338, Trp238 and Trp85, line much of the active-site pocket of vanin 1. Only Phe182 and Phe212 show significantly higher B factors in the empty pocket.

To determine whether the Glu79 and Lys178 residues in the active site of vanin 1 affect vanin activity, we generated mutations of these residues and tested their vanin activity using a fluorophore assay. As expected, the wild-type enzyme has high activity in the enzymatic assay (Fig. 4). The active-site residues Glu79 and Lys178 help to orient and activate the Cys211 residue for catalysis; not surprisingly, the Glu79Ala and Lys178Ala mutations ablated the activity of the enzyme and demonstrate the importance of these residues for catalytic activity. It has recently been shown that mutation of Cys211 ablates the activity of vanin 1 (Rommelaere *et al.*, 2013).

The vanin family of proteins is classified as subgroup 4 of the nitrilase superfamily, and members of this superfamily generally form homo-oligomeric complexes, with the basic building block being a homodimer (Brenner, 2002). The crystal structure of vanin 1 shows no sign of oligomerization either by inspection or by analysis using *PISA* (Krissinel & Henrick, 2007; the largest protein–protein interface is 664 Å^2). In addition, both the gel-filtration profile during purification and small-angle X-ray scattering experiments (SAXS) indicate that the protein is monomeric (Table 2). Sequence alignments show that the residues that should be important for oligomerization are Gln179–Phe182, Pro191–Glu193, Phe212–Asp218, Pro243–His251 and Leu326–Asn330. The first three regions form a groove on the surface of vanin 1, whereas Pro243–His251, although close to this groove, is buried by the second domain. The last set of residues (Leu326–Asn330) is on the surface of the base domain, far away from the previously described groove. In the mouse nitrilase structure (PDB entry 2w1v), two α-helices (Arg195–Ile202 and His225–Asp236) and the last 18 residues (295–312) make extensive interactions to form the dimer structure (Barglow *et al.*, 2008). The extended C-terminal structure of the mouse nitrilase protein deviates from the vanin 1 structure starting at residue 296 of PDB entry 2w1v and residue 312 of vanin 1 (Fig. 5*a*). When superposed (*PDBeFold*; r.m.s.d. = 1.8 Å, 223 aligned

residues, ten gaps, 23% sequence identity), this region sits at the interface between the vanin 1 domains (Fig. 5*a*). These data show that vanin 1, unlike the other known nitrilase enzymes, is a monomer and not an oligomer. In contrast, the active-site residues (Glu79, Lys178 and Cys211) of vanin 1 overlay well with the catalytic triads of other nitrilase structures, with the same orientations and approximately the same distances between side-chain atoms (less than 0.15 Å deviation; Fig. 5*b*). Additionally, many of the residues lining the active-site pocket are also conserved between vanin 1 and the nitrilases.

The base domain has no sequence homology to anything in the PDB. Interestingly, it does have structural homology to a lectin-binding domain of a *Streptococcus pneumoniae* glycoside hydrolase (PDB entry 2wmk), a protein involved in specific recognition of the Lewis antigen (Higgins *et al.*, 2009; Fig. 5*c*). In general, the base domain overlaps well with the 2wmk structure for ∼150 residues: residues ∼327–478 of vanin 1 and ∼843–1005 of 2wmk (Fig. 5*c*). There is one major divergent insertion that occurs at residues 402–421 of vanin 1 and residues 902–949 of 2wmk. Otherwise, ten of the 11 β-strands overlap reasonably well, with the loops connecting them showing some similarity. As these domains are structurally similar, we hypothesize that the base domain plays a role in binding to other proteins (either through saccharide binding or through protein–protein interactions) and in signalling, functions that have previously been described for vanins but without an obvious explanation as to how they occurred (Aurrand-Lions *et al.*, 1996; Yamashita *et al.*, 1999). These binding events may also regulate the enzymatic activity of vanins, as movement of the base domain relative to the enzymatic domain can change the activity of the protein (see below).

Structural analysis reveals two unusually positioned glutamic acid residues, Glu249 and Glu439, one from each domain, which are buried within 4 Å of each other in the interface between the two domains, an interface which is mostly hydrophobic (Fig. 6). This would appear to be energetically unfavorable and there must be a reason to have these residues in this configuration. These two glutamic acid residues were individually mutated to glutamine residues to determine how they might affect the function of the protein. Using an enzymatic fluorophore assay (Fig. 4*b*), it was shown that these subtle mutations completely abolished the activity. We have shown that this is not owing to unfolding of the protein (SAXS and differential scanning fluorimetry; Supporting Information), but must be due to a domain rearrangement that affects the activity of the catalytic domain. SAXS experiments were performed to ascertain whether

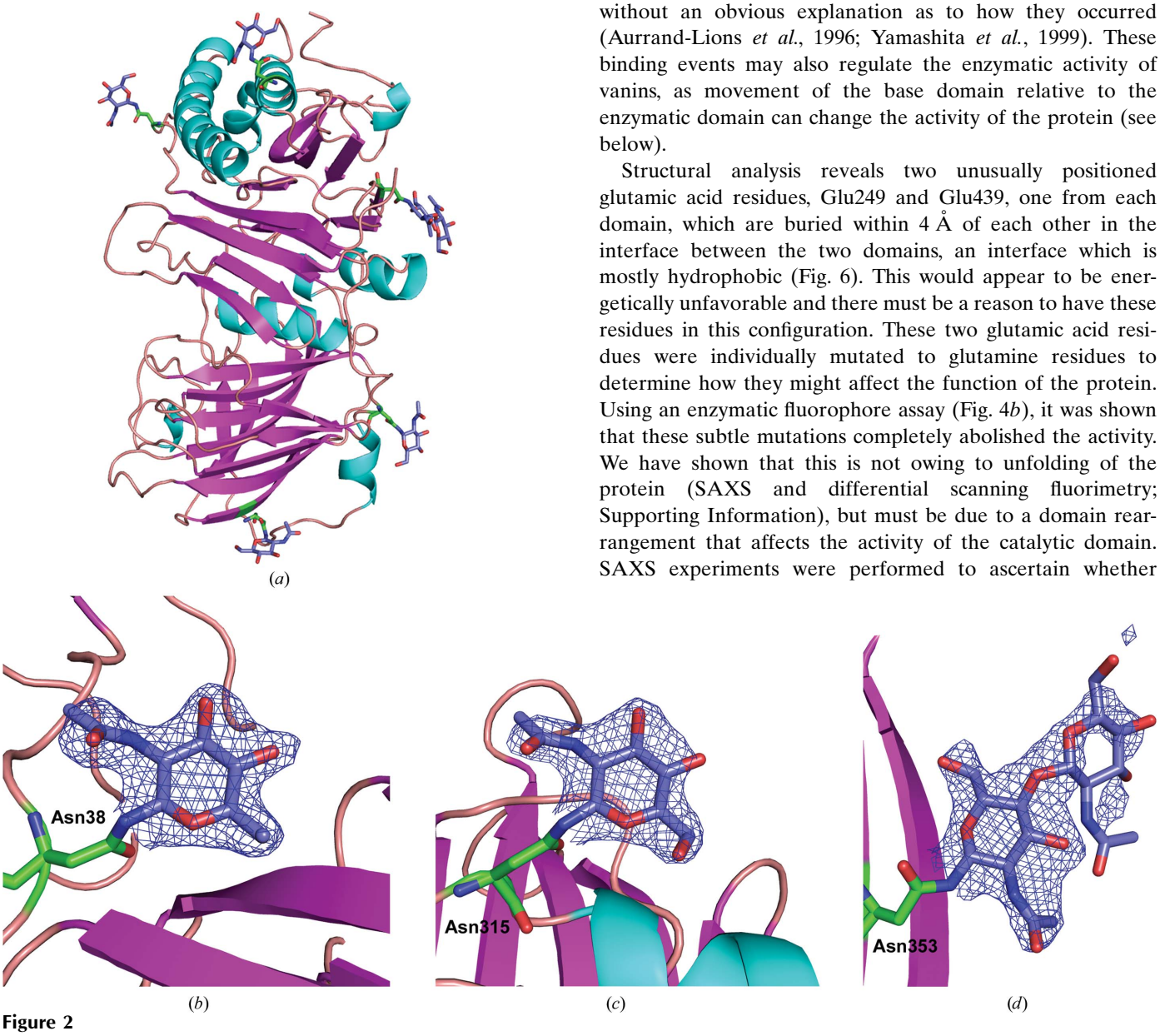


Figure 2
The vanin 1 structure highlighting the glycosylation sites at residues Asn38, Asn130, Asn200, Asn315 and Asn353. (*a*) The structure is shown as a cartoon highlighting the secondary structure, with the Asn residues with green C atoms and the NAG sugar residues with blue C atoms. (*b*) Representative composite OMIT map density (blue mesh) shown at 1.5σ around the NAG on Asn38. (*c*) Representative composite OMIT map density (blue mesh) shown at 1.5σ around the NAG on Asn315. (*d*) Representative composite OMIT map density (blue mesh) shown at 1.5σ around the two NAGs on Asn353.

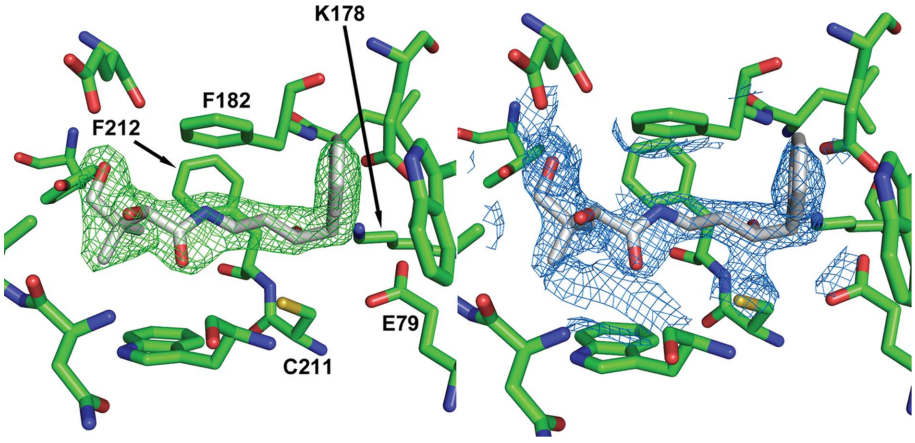


Figure 3
The inhibitor RR6 (C atoms in grey) shown with difference density ($mF_o - DF_c$ map, green mesh, contoured at 3σ) and standard density ($2mF_o - DF_c$ map, blue mesh, contoured at 1.5σ) in the catalytic site of vanin 1. Active-site residues are highlighted (Glu79, Lys178 and Cys211) as well as two phenylalanine residues (Phe182 and Ph212) which become more ordered when the inhibitor is bound.

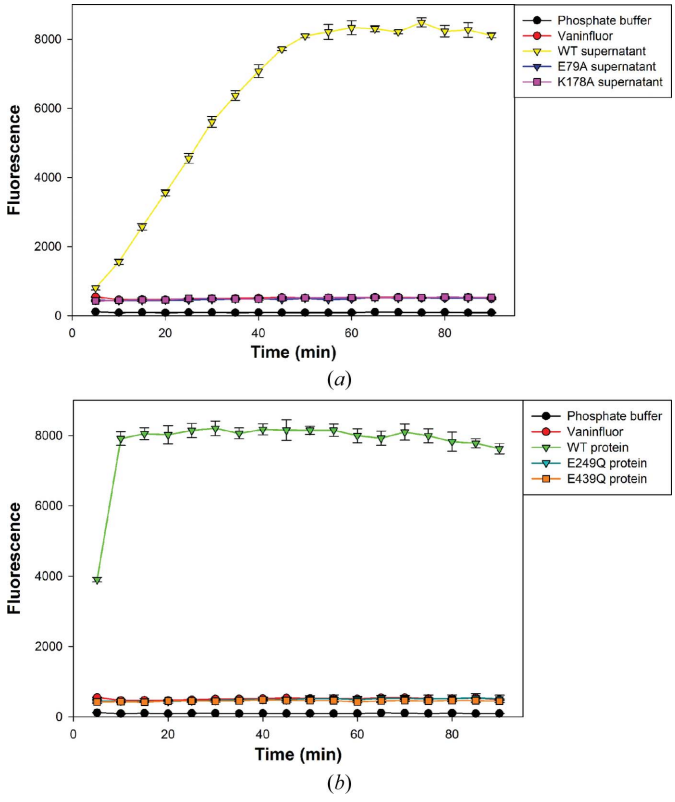


Figure 4
Vanin 1 activity in the Vaninfluor assay. (*a*) The phosphate buffer and Vaninfluor samples are controls and show baseline activity. The E79A, K178A and WT samples are cell-supernatant samples (partially purified protein). The wild-type protein shows full activity, whereas the active-site mutants show no activity in this assay. (*b*) The Vaninfluor assay with purified protein. This assay was performed with the same purified proteins as used for the SAXS analysis; the wild type shows full activity and both domain-interface mutants show no activity in this assay. Data are expressed as the mean ± SEM. All samples were measured in triplicate.

the domain movement between the native protein and the mutated proteins was on a time scale long enough to observe clear differences in the overall solution structure. These data (Table 2 and Supplementary Fig. S4) show that the two mutated proteins adopt an identical structure in solution and that this is similar to the native structure, although subtly but significantly more extended. The measured difference between the wild-type and mutant structures is strong evidence of a change in the inter-domain interface. Taken together with the effect of these mutations on catalytic activity, we hypothesize that the mutations lock the structure of the two domains together and, as for many other enzymes, catalytic activity is dependent on a dynamic structure (Eisenmesser *et al.*, 2005).

Although uncommon, there are multiple structures in the PDB and in the literature that show that similarly charged residues can be close to each other in three-dimensional space (Magalhaes *et al.*, 1994; Flocco & Mowbray, 1995; Kajander *et al.*, 2000). In all of the cases that we could find, these were either solvent-exposed or had other compensating (oppositely charged) residues nearby to balance out the energetically unfavorable interaction (*e.g.* PDB entries 1a3q, 1hxw, 2gst, 2or3, 2p74, 3f33 and 3hyr). For the vanin 1 structure, we find no solvent in the buried interface and no compensatory charge interactions to modulate the charge density in this hydrophobic environment. It has been observed that charged residues can have anomalous pK_a values, shifting the pK_a towards the neutral form of the acid/base, in different environments (Kim *et al.*, 2005; Harms *et al.*, 2009). We suspect that this may be the case and that at least one of the two glutamic acid residues may be protonated. Domain movement may change the orientation and hence the solvent accessibility and environment of these residues and their respective pK_a values.

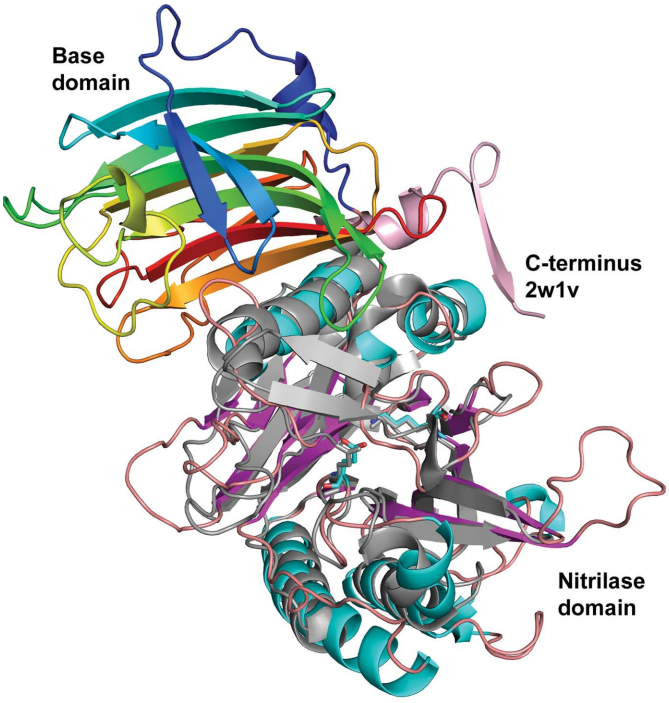
4. Discussion

The structure of vanin 1 was unknown and is essential information that will allow us to develop a better understanding of the activity of vanin 1 in different disease states. Through sequence analysis, it was hypothesized that vanins had a nitrilase-like domain (Brenner, 2002) and an extension at the carboxy-terminus of about 180 residues. The oligomeric state of the protein and the nature of this extension were unknown. More recently, an inhibitor of the protein has been developed (Jansen *et al.*, 2013), but how this inhibitor bound to the protein and caused its effects was also unknown. Here, we not only demonstrate that vanin 1 has a base domain as well as a nitrilase domain, but we also show that its activity is dependent on a potential high-energy Glu–Glu charged pair

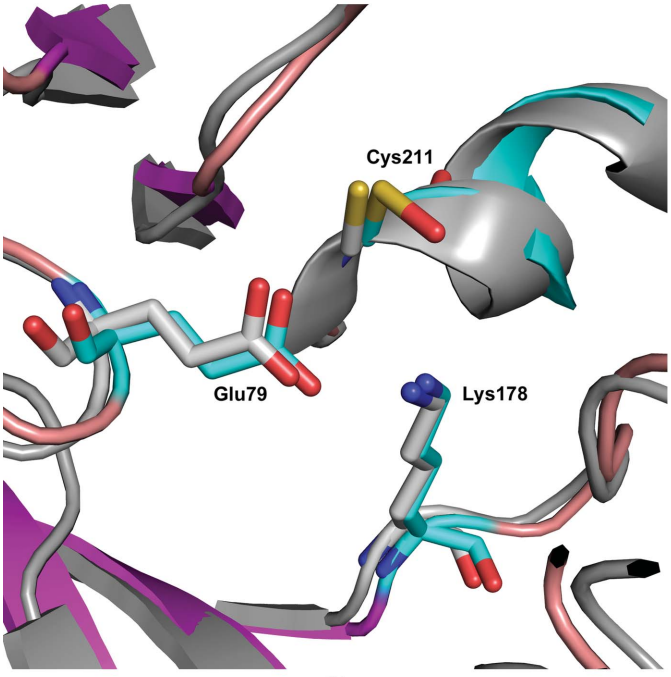
between the two domains. We were able to manipulate the activity of vanin 1 by generating mutations of these Glu residues and by making use of the recently identified vanin 1 inhibitor RR6. These novel data have unraveled an important mechanism of vanin 1 in which it can mediate its biological effects by either its enzymatic activity or by protein–protein interactions (Aurrand-Lions *et al.*, 1996; Yamashita *et al.*, 1999).

The most surprising feature of the new structure is the Glu–Glu pair in the hydrophobic interface between the two domains of vanin 1. These residues are by and large conserved between the various vanin 1 sequences in other species, so we expect the whole family to look and behave in a similar manner. After manually parsing through hundreds of structures with closely packed charged residues, we have not been able to find anything similar in terms of a buried charge to what is found in vanin 1. It was equally surprising to find that a very subtle change of either residue from Glu to Gln could completely ablate the enzymatic function of the protein without either substantially affecting its solution structure or its stability (SAXS and DSF analysis).

Our data provide new evidence which confirms that vanins are indeed comprised of two domains: an enzyme domain which is part of the nitrilase superfamily and which contributes to the redox balance through the release of cysteamine and pantothenic acid, and a second domain with structural features that are consistent with a protein-binding function. Furthermore, we demonstrate that the enzymatic activity of vanins is modulated by way of an embedded pair of glutamic acid



(*a*)



(*b*)

(*c*)

Figure 5
Comparison of vanin 1 to other structures. (*a*) Superposition of the mouse nitrilase domain (PDB entry 2w1v) with vanin 1. 223 residues overlap with an r.m.s.d. of 1.8 Å with 23.3% sequence identity and ten gaps (superposition was performed with *Coot*). The C-terminus of 2w1v, residues 295–312, makes extensive interactions with a second monomer to form a dimer in solution. The C-termini diverge between the two structures at about residue 312 of vanin 1 (residue 296 of 2w1v). (*b*) Superposition of the mouse nitrilase structure (PDB entry 2w1v) with the vanin 1 structure. The active-site residues, Glu79, Lys178 and Cys211, are in almost the exact same location in the two structures. Cys211 is shown as oxidized in the vanin 1 structure as there was extra density for this O atom in the native (non-inhibitor bound) structure. (*c*) A superposition of the vanin 1 base domain (rainbow-coloured with the N-terminus in dark blue and the C-terminus in red) and the 2wmk domain (in grey). 107 residues superpose with an r.m.s.d. of 3.6 Å and eight gaps with a sequence identity of 7.5%. It is clear that the β-sheets overlay well and that some of the loops do not superpose well; several are of different lengths. The 2wmk structure is missing residues 906–946 of one loop (near the top of the figure).

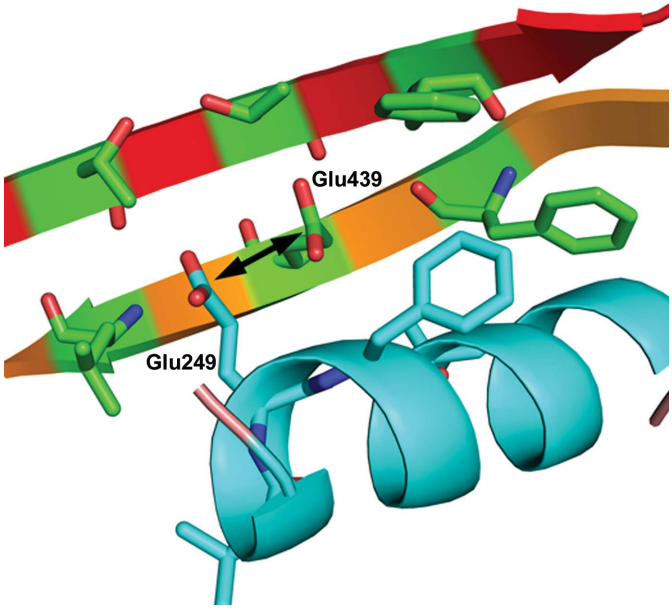


Figure 6
Orientation of the glutamic acid residues in the domain interface. The vanin 1 domain interface highlighting the two glutamic residues, Glu249 and Glu439, which are very close in space and in a hydrophobic environment. The base-domain residues are highlighted with green C atoms and protrude from the two last β-strands of the base structure, whereas the nitrilase-domain residues have cyan C atoms and originate from the α-helix buried in this interface. The arrow between the two glutamic acid residues is there to indicate that the distance between these is less than 4 Å (3.5–3.9 Å depending on the atoms chosen).

residues between the signalling and enzyme domains, providing an exquisitely sensitive mechanism for allosteric communication between one function and the other.

In conclusion, we have determined the first atomic structure of vanin 1, which verifies its proposed domains, and show how vanin 1 could operate as both an enzyme and a homing/signalling protein, as previously suggested. Our observations may thus provide insight into the fine balance of actions of vanins in metabolic diseases and the inflammation often associated with metabolic pathologies such as obesity (Maachi *et al.*, 2004; Hotamisligil, 2006). The probe compound RR6 will be a valuable tool to tease apart these two different functions, and our structure with a bound inhibitor will allow the rational exploration of chemical space for specific inhibitors of this enzyme.

We thank the beamline scientists at the Australian Synchrotron for their help with data collection. Crystals were grown in the CSIRO Collaborative Crystallisation Centre (http://www.csiro.au/C3). We also thank Xiaowen Xiao, John Bentley, Tram Phan and Lindsay Sparrow for help with molecular biology, protein purification, cell-culture work and mass spectrometry, Del Lucent, Nat Echols and Marilyn Gunner for discussions and OpenEye Scientific Software for a license for *AFITT*. This research was supported by a Grant-In-Aid (G 10M 5182) from the National Heart Foundation of Australia, a HFSP long-term fellowship to YLB (LT001131/2011), the Science and Industry Endowment Fund Special Research Program – Synchrotron Science and CSIRO internal funds. We thank Associate Professor Matthew Piggott (School of Chemistry & Biochemistry, The University of Western Australia) for the vaninfluor compound used in the activity assays.

References

Aurrand-Lions, M., Galland, F., Bazin, H., Zakharyev, V. M., Imhof, B. A. & Naquet, P. (1996). *Immunity*, **5**, 391–405.
Barglow, K. T., Saikatendu, K. S., Bracey, M. H., Huey, R., Morris, G. M., Olson, A. J., Stevens, R. C. & Cravatt, B. F. (2008). *Biochemistry*, **47**, 13514–13523.
Berruyer, C., Martin, F. M., Castellano, R., Macone, A., Malergue, F., Garrido-Urbani, S., Millet, V., Imbert, J., Duprè, S., Pitari, G., Naquet, P. & Galland, F. (2004). *Mol. Cell. Biol.* **24**, 7214–7224.
Berruyer, C., Pouyet, L., Millet, V., Martin, F. M., LeGoffic, A., Canonici, A., Garcia, S., Bagnis, C., Naquet, P. & Galland, F. (2006). *J. Exp. Med.* **203**, 2817–2827.
Brenner, C. (2002). *Curr. Opin. Struct. Biol.* **12**, 775–782.
Chen, S., Zhang, W., Tang, C., Tang, X., Liu, L. & Liu, C. (2014). *Diabetes*, **63**, 2073–2085.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.
Cowtan, K. D. & Main, P. (1996). *Acta Cryst.* D**52**, 43–48.
Diepen, J. A. van *et al.* (2014). *J. Hepatol.* **61**, 366–372.
Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E. & Kern, D. (2005). *Nature (London)*, **438**, 117–121.
Elleman, T. C. *et al.* (2001). *Biochemistry*, **40**, 8930–8939.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Evans, M. J., Hartman, S. L., Wolff, D. W., Rollins, S. A. & Squinto, S. P. (1995). *J. Immunol. Methods*, **184**, 123–138.
Flocco, M. M. & Mowbray, S. L. (1995). *J. Mol. Biol.* **254**, 96–105.
Gensollen, T. *et al.* (2013). *Inflamm. Bowel Dis.* **19**, 2315–2325.
Hao, Q. (2004). *J. Appl. Cryst.* **37**, 498–499.
Harms, M. J., Castañeda, C. A., Schlessman, J. L., Sue, G. R., Isom, D. G., Cannon, B. R. & García-Moreno, E. B. (2009). *J. Mol. Biol.* **389**, 34–47.
Higgins, M. A., Whitworth, G. E., El Warry, N., Randriantsoa, M., Samain, E., Burke, R. D., Vocadlo, D. J. & Boraston, A. B. (2009). *J. Biol. Chem.* **284**, 26161–26173.
Hotamisligil, G. S. (2006). *Nature (London)*, **444**, 860–867.
Hotamisligil, G. S. (2010). *Cell*, **140**, 900–917.
Huang, P. L. (2009). *Dis. Mod. Mech.* **2**, 231–237.
Jansen, P. A., Kamsteeg, M., Rodijk-Olthuis, D., van Vlijmen-Willems, I. M., de Jongh, G. J., Bergers, M., Tjabringa, G. S., Zeeuwen, P. L. & Schalkwijk, J. (2009). *J. Investig. Dermatol.* **129**, 2167–2174.
Jansen, P. A. M., van Diepen, J. A., Ritzen, B., Zeeuwen, P. L. J. M., Cacciatore, I., Cornacchia, C., van Vlijmen-Willems, I. M. J. J., de Heuvel, E., Botman, P. N. M., Blaauw, R. H., Burghout, P., van Galen, P. M., Mouton, J. W., Rutjes, P. J. T. & Schalkwijk, J. (2013). *ACS Chem. Biol.* **8**, 530–534.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kajander, T., Kahn, P. C., Passila, S. H., Cohen, D. C., Lehtiö, L., Adolfsen, W., Warwicker, J., Schell, U. & Goldman, A. (2000). *Structure*, **8**, 1203–1214.
Kaskow, B. J., Proffitt, J. M., Blangero, J., Moses, E. K. & Abraham, L. J. (2012). *Biochem. Biophys. Res. Commun.* **417**, 653–658.
Kim, J., Mao, J. & Gunner, M. R. (2005). *J. Mol. Biol.* **348**, 1283–1298.
Kirby, N. M., Mudie, S. T., Hawley, A. M., Cookson, D. J., Mertens, H. D. T., Cowieson, N. & Samardzic-Boban, V. (2013). *J. Appl. Cryst.* **46**, 1670–1680.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Maachi, M., Piéroni, L., Bruckert, E., Jardel, C., Fellahi, S., Hainque, B., Capeau, J. & Bastard, J.-P. (2004). *Int. J. Obes. Relat. Metab. Disord.* **28**, 993–997.
Magalhaes, A., Maigret, B., Hoflack, J., Gomes, J. N. F. & Scheraga, H. A. (1994). *J. Protein Chem.* **13**, 195–215.
Maras, B., Barra, D., Duprè, S. & Pitari, G. (1999). *FEBS Lett.* **461**, 149–152.
Martin, F., Penet, M.-F., Malergue, F., Lepidi, H., Dessein, A., Galland, F., de Reggi, M., Naquet, P. & Gharib, B. (2004). *J. Clin. Invest.* **113**, 591–597.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J. L., Stuart, D. I. & Perrakis, A. (2005). *Acta Cryst.* D**61**, 1426–1431.
Ng, M. *et al.* (2014). *Lancet*, **384**, 766–781.
Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* D**61**, 449–457.
Pannu, N. S., McCoy, A. J. & Read, R. J. (2003). *Acta Cryst.* D**59**, 1801–1808.
Pitari, G., Malergue, F., Martin, F., Philippe, J. M., Massucci, M. T., Chabret, C., Maras, B., Duprè, S., Naquet, P. & Galland, F. (2000). *FEBS Lett.* **483**, 149–154.
Rakhshandehroo, M., Knoch, B., Müller, M. & Kersten, S. (2010). *PPAR Res.* **2010**, 612089.
Roisin-Bouffay, C., Castellano, R., Valéro, R., Chasson, L., Galland, F. & Naquet, P. (2008). *Diabetologia*, **51**, 1192–1201.
Rommelaere, S., Millet, V., Gensollen, T., Bourges, C., Eeckhoute, J., Hennuyer, N., Baugé, E., Chasson, L., Cacciatore, I., Staels, B., Pitari, G., Galland, F. & Naquet, P. (2013). *FEBS Lett.* **587**, 3742–3748.
Ruan, B. H., Cole, D. C., Wu, P., Quazi, A., Page, K., Wright, J. F., Huang, N., Stock, J. R., Nocka, K., Aulabaugh, A., Krykbaev, R., Fitz, L. J., Wolfman, L. M. & Fleming, M. L. (2010). *Anal. Biochem.* **399**, 284–292.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Seabrook, S. A. & Newman, J. (2013). *ACS Comb. Sci.* **15**, 387–392.
Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.
Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, 479–485.
Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Whitten, A. E., Cai, S. & Trewhella, J. (2008). *J. Appl. Cryst.* **41**, 222–226.
Yamashita, Y. A., Sendo, F., Fujita, H. T., Takeda, Y., Sato, Y., Ohtake, T. K., Kinoshita, T., Araki, A., Suzuki, T. K. & Watanabe, T. (1999). *J. Immunol.* **162**, 4277–4284.