Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure and epimerase activity of anthocyanidin reductase from *Vitis vinifera*

Mahmoud Gargouri,[a,b] Claude Manigand,[a] Chloé Maugé,[a] Thierry Granier,[a] Béatrice Langlois d'Estaintot,[a] Olivier Cala,[c] Isabelle Pianet,[c] Katell Bathany,[c] Jean Chaudière[a] and Bernard Gallois[a]*

[a]Chimie et Biologie des Macromolécules et des Nanoobjets UMR CNRS 5248, Bâtiment B8, Avenue des Facultés, Université Bordeaux 1, 33405 Talence CEDEX, France, [b]Laboratoire de Physiologie Moléculaire de la Vigne, CBBC, BP 901, Hamman-Lif 2050, Tunisia, and [c]Chimie et Biologie des Membranes et des Nanoobjets UMR CNRS 5248, Institut Européen de Chimie et de Biologie, 2 Rue Robert Escarpit, 33607 Pessac, France

Correspondence e-mail: b.gallois@cbmn.u-bordeaux.fr

Received 12 May 2009
Accepted 29 June 2009

PDB References: anthocyanidin reductase, 2rh8, r2rh8sf; 3hfs, r3hfssf.

Together with leucoanthocyanidin reductase, anthocyanidin reductase (ANR) is one of the two enzymes of the flavonoid-biosynthesis pathway that produces the flavan-3-ol monomers required for the formation of proanthocyanidins or condensed tannins. It has been shown to catalyse the double reduction of anthocyanidins to form 2*R*,3*R*-flavan-3-ols, which can be further transformed to the 2*S*,3*R* isomers by non-enzymatic epimerization. ANR from grape (*Vitis vinifera*) was expressed in *Escherichia coli* and purified. Unexpectedly, RP-HPLC, LC-MS and NMR experiments clearly established that the enzyme produces a 50:50 mixture of 2,3-*cis* and 2,3-*trans* flavan-3-ols which have been identified by chiral chromatography to be 2*S*,3*S*- and 2*S*,3*R*-flavan-3-ols, *i.e.* the naturally rare (+)-epicatechin and (−)-catechin, when cyanidin is used as the substrate of the reaction. The first three-dimensional structure of ANR is described at a resolution of 2.2 Å and explains the inactivity of the enzyme in the presence of high salt concentrations.

1. Introduction

The most common group of flavonoids consumed in the diet consists of the flavan-3-ols and their polymeric condensation products the proanthocyanidins (PAs) or condensed tannins (Aron & Kennedy, 2008). They are present in the fruits, leaves and seeds of many plants, where their major function is to provide protection against microbes, fungi, insects or herbivores (Dixon *et al.*, 2005; Scarlbert, 1991; Peters & Constabel, 2002). When ingested, some of them may induce beneficial effects for human health by acting as cardioprotective or neuroprotective agents (Bagchi *et al.*, 2000; Lin *et al.*, 2002; Cos *et al.*, 2004; Rasmussen *et al.*, 2005; Collie, 2006). They also contribute to the astringency and taste of many plant products such as fruit juices and tea. In grapes, flavan-3-ols and PAs are important components of the organoleptic properties of wine as they not only strongly contribute to its gustatory impact but can also modify the colour stability when interacting with anthocyanins (Glories, 1988). The important implications of these compounds in human health and food quality call for a better understanding of the mechanisms by which they are synthesized.

The flavonoid-biosynthetic pathway has largely been investigated (Winkel-Shirley, 2001). Most of the enzymes that control each single step have been identified. Nevertheless, little is known about the way in which PAs are formed and about the origin of the 2,3-*cis* stereochemistry observed for some of the flavan-3-ol subunits. It is only recently that the branches of the pathway leading to the biosynthesis of flavan-3-ols have fully been elucidated. Leucoanthocyanidin

4-reductase (LAR; EC 1.17.1.3), which converts leucoanthocyanidins to the corresponding 2*R*,3*S*-flavan-3-ols such as (+)-catechin by NADPH reduction (Stafford, 1990; Joseph *et al.*, 1998; Tanner *et al.*, 2003), is not the only enzyme that is involved in flavan-3-ol biosynthesis. The recent biochemical characterization of the *BANYULS* gene from *Arabidopsis thaliana* and *Medicago truncatula* (Xie *et al.*, 2003, 2004), which was initially thought to be a *lar* gene (Devic *et al.*, 1999), showed that this gene encodes a new enzyme, anthocyanidin reductase (ANR; EC 1.3.1.77), which functions downstream of anthocyanidin synthase (Fig. 1) and catalyses the double NADPH reduction of anthocyanidins. It was shown that the reaction products of ANR from *A. thaliana* (At-ANR) and *M. truncatula* (Mt-ANR) included two distinct isomers, namely 2*R*,3*R*- and 2*S*,3*R*-flavan-3-ols. With cyanidin as substrate, the products are (−)-epicatechin and (−)-catechin, respectively. However, the minor presence of the 2*S*,3*R* isomer was assumed to result from a non-enzymatic C2-epimerization of the 2*R*,3*R* isomer, thereby suggesting that ANR only converts anthocyanidins to 2*R*,3*R*-flavan-3-ols.

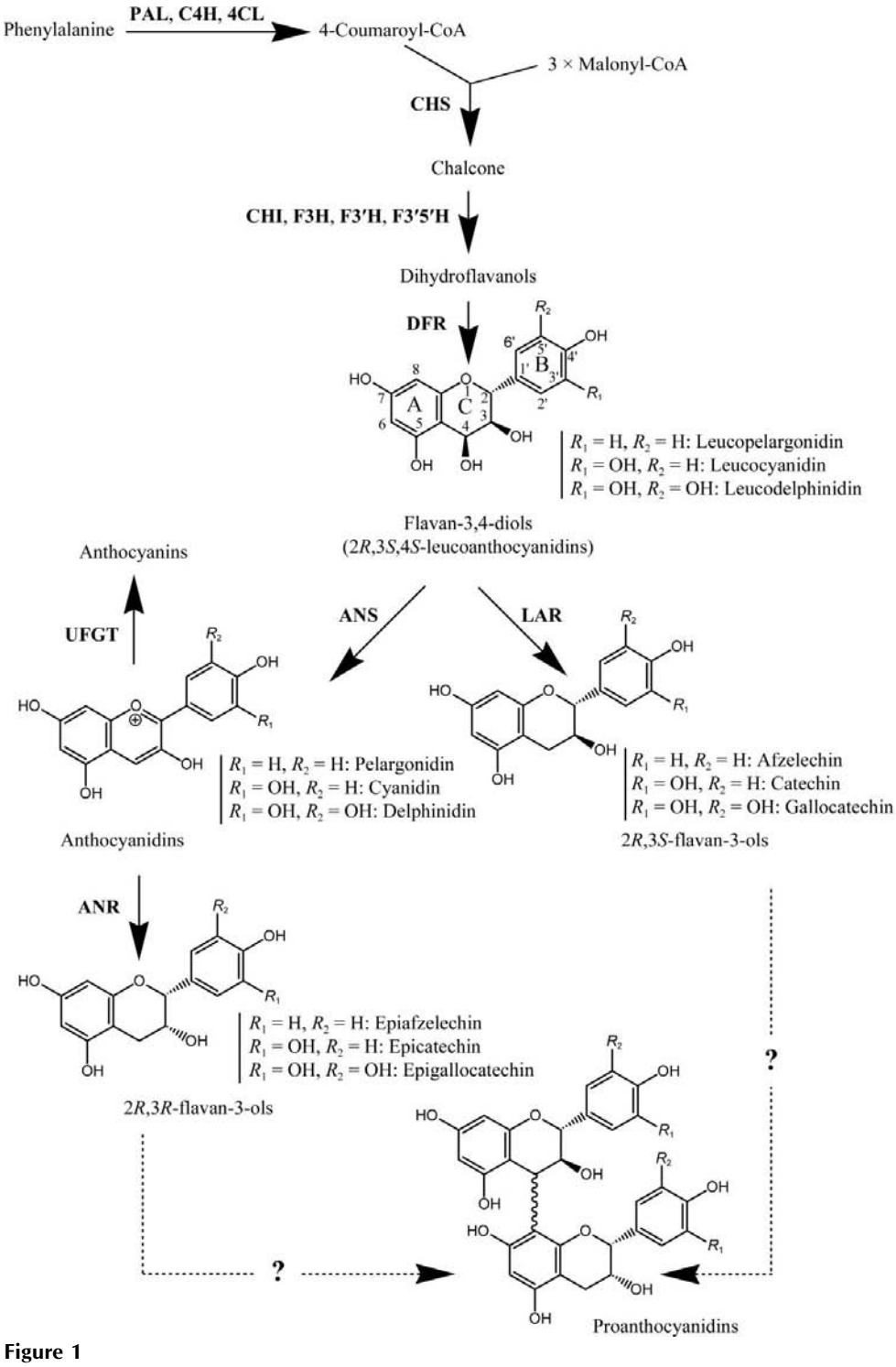

Figure 1
Schematic drawing of the biosynthetic pathway of anthocyanins and proantocyanidins. Enzyme abbreviations: ANR, anthocyanidin reductase; ANS, anthocyanidin synthase (also known as leucoanthocyanidin dioxygenase); CHI, chalcone isomerase; CHS, chalcone synthase; C4H, cinnamate 4-hydroxylase; DFR, dihydroflavonol 4-reductase; F3H, flavanone 3-hydroxylase; F3′H, flavanone 3′-hydroxylase; F3′5′H, flavanone 3′,5′-hydroxylase; LAR, leucoanthocyanidin reductase; PAL, L-phenylalanine ammonia-lyase; UFGT, UDP-glucose:flavonoid 3-*O*-glucosyltransferase; 4CL, 4-coumaroyl:CoA-ligase.

To date, putative *anr* sequences have been mentioned in *Phaseolus coccineus* and *Gossypium arboreum* (Tanner *et al.*, 2003), *Gingko biloba* (Shen *et al.*, 2006), *G. hirsutum* (Xiao *et al.*, 2007) and *Diospyros kaki* (Ikegami *et al.*, 2007) and ANR activity has been reported for a limited number of plants such as *Camellia sinensis* (Punyasiri *et al.*, 2004), *Vitis vinifera* L. and *Malus* × *domestica* Borkh. (Pfeiffer *et al.*, 2006) and *Lotus corniculatus* (Paolocci *et al.*, 2007). In grapes, the nucleotide sequence of the *anr* gene (*Vitis vinifera* cv. Cabernet Sauvignon; GenBank accession No. AB199315) has been published (Fujita *et al.*, 2005) and both the expression of *anr* genes and the accumulation of flavan-3-ols and PAs have been followed during berry development (Fujita *et al.*, 2005; Bogs *et al.*, 2005).

No three-dimensional structure of an ANR has been determined to date. The only structural information con-

cerning the enzyme is that deduced from sequence analysis. ANR sequences contain most of the motifs present in the short-chain dehydrogenase/reductase (SDR) superfamily, including the invariant Y*XXX*K motif (where *X* represents any amino-acid residue), a catalytic serine and a highly conserved N-terminal glycine-rich nucleotide-binding domain. Consequently, ANR folding should present a large α/β single conserved domain named the Rossmann fold as in other members of the superfamily.

Here, we report the successful heterologous expression of grape ANR (Vv-ANR), leading to the active enzyme from *V. vinifera* cv. Cabernet Sauvignon (Swiss-Prot accession No. Q5FB34). Optimization of purification conditions allowed us to characterize its activity, which is only observed at low salt concentrations. Under such conditions, ANR from *V. vinifera* behaves as a pro-*S* reductase/epimerase, producing a nearly 50:50 mixture of the naturally rare (+)-epicatechin and (−)-catechin. Together with the biochemical properties of the enzyme, we present the first ANR three-dimensional structure. Its description leads us to understand the enzyme inactivity in presence of high salt concentrations.

2. Materials and methods

2.1. Chemicals

Cyanidin chloride, pelargonidin chloride, delphinidin chloride, (+)-catechin, (−)-epicatechin and (−)-epigallocatechin were purchased from Extrasynthese (Genay, France). (−)-Catechin was purchased from Sigma. NADPH (from Sigma) solutions were freshly prepared in cold reaction buffer used for enzyme assay.

2.2. Cloning and expression of Vv-ANR in *Escherichia coli*

The open reading frame of the Vv-*ANR* gene was PCR-amplified with a *V. vinifera* cv. Cabernet Sauvignon post-veraison berry cDNA library as template. PCR was performed using the oligonucleotide primers GCTA**CCATGG**CCACCCAGCAC and CGTA**GGTACC**TTAATTCTGCAATAGCCCCT so that the PCR product contained *Nco*I and *Kpn*I restriction sites at the 5′- and 3′-termini, respectively. The PCR program started with an initial denaturation step of 1 min at 368 K, which was followed by 35 cycles at 368 K for 30 s, 331.6 K for 45 s and 345 K for 75 s and a final extension for 4 min at 345 K using *PfuTurbo* DNA polymerase (Stratagene). The amplified gene product was ligated into pGEM-T Easy (Promega). The plasmid was transformed into *E. coli* JM109 competent cells (Promega) for sequence analysis. The open reading frame was cloned between the *Nco*I and *Kpn*I sites of the pETGB_1a bacterial expression vector (kindly supplied by P. Zou and G. Stier, EMBL Hamburg) to generate an N-terminally His_6-GB1 fused ANR construct presenting a TEV-cleavable linker at the C-terminal part of the fusion immunoglobulin-binding domain of streptococcal protein G GB1. The resulting plasmid was transformed into *E. coli* strain BL21 (DE3). Cultures were grown at 310 K overnight on 1%(*w*/*v*) LB agar supplemented with 25 µg ml^{-1} kanamycin. A single colony was used to inoculate 1 l LB medium and the cells were grown to an OD_{600} of 0.6 at 310 K and 250 rev min^{-1}. The cells were chilled on ice for 10 min before IPTG was added to a final concentration of 1.0 m*M* to initiate overexpression. The culture was incubated for 24 h at 288 K and was shaken at 180 rev min^{-1}. The cells were harvested by centrifugation at 3300*g* at 277 K for 20 min and resuspended in ice-cold lysis buffer [50 m*M* NaH_2PO_4, 300 m*M* NaCl, 10 m*M* imidazole, 1 mg ml^{-1} hen egg-white lysozyme pH 8 and two protease-inhibitor cocktail tablets (Complete EDTA-free, Roche) per 100 ml of buffer]. Cells were sonicated for 10 min (20 s pulses with 10 s intervals using a Branson Digital Sonifier) and debris was removed by centrifugation at 9000*g* for 30 min at 277 K.

2.3. Enzyme extraction and purification

All procedures were performed at 277 K. The supernatant was loaded onto a HiTrap Chelated HP 5 ml column (Amersham Biosciences) equilibrated with at least five column volumes of washing buffer (10 m*M* tricine, 10 m*M* imidazole, 500 m*M* NaCl pH 7.5) and the His_6-GB1-ANR fusion protein was eluted with five column volumes of washing buffer supplemented with imidazole to 300 m*M*. The elution fraction was immediately supplemented with 5 m*M* EDTA and 10 m*M* DTT to avoid protein precipitation or aggregation. The fusion protein was concentrated to 30 mg ml^{-1} using a 10 000 molecular-weight cutoff Amicon Ultra (Millipore) membrane and then analysed by electrophoresis on a 10% SDS–polyacrylamide gel stained with 0.25% Coomassie Brilliant Blue. ANR protein was cleaved from the His_6-GB1-tagged domain using TEV protease. To separate oligomeric states from monomers, the cleaved ANR elution fraction was loaded onto a Superdex-200 10/300 GL gel-filtration column (Amersham Biosciences) equilibrated with 10 m*M* tricine, 500 m*M* NaCl pH 7.5. The final fraction of interest was concentrated and then dialyzed against 10 m*M* tricine, 500 m*M* NaCl, 10 m*M* DTT, 5 m*M* EDTA pH 7.5 (buffer *A*) and stored at 277 K or dialyzed against 10 m*M* tricine, 50 m*M* NaCl, 10 m*M* DTT, 5 m*M* EDTA pH 7.5, 2.5% glycerol (buffer *B*) and stored at 253 K. The tagged protein was also stable in these buffers. The protein concentration was determined using a theoretical molar extinction coefficient of 28 795 M^{-1} cm^{-1} at 280 nm and the purity of the sample was assessed by SDS–PAGE and matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. The observed mass of the native protein was 36 850 Da. Owing to the accuracy of MALDI spectrometry, this value agrees with the gene integrity previously confirmed by DNA sequencing (theoretical mass of 36 846.6 Da without the initiator methionine).

2.4. Enzyme assay and identification of ANR reaction products

Different substrates (cyanidin, pelargonidin and delphinidin) were freshly prepared at 5 m*M* concentration using 0.4 m*M* sodium methanesulfonate in methanol pH 2.0. In order to remove DTT and EDTA (contained in buffers *A* and

B), which interfered with the assay, the enzyme was dialysed against 500 m*M* NaCl, 100 m*M* Tris–HCl pH 7.0. Enzyme assays were carried out at 303 K in a total volume of 2 ml containing 100 m*M* Tris–HCl pH 7.0, 500 µ*M* cyanidin, pelargonidin or delphinidin chloride, 2 m*M* NADPH and 15 µ*M* enzyme. The reaction was initiated by the addition of 20 µl enzyme solution, which decreased the NaCl concentration in the reaction mixture to 5 m*M*. The reaction was stopped by the addition of 20 ml of ethyl acetate to the reaction mixture. The two-phase solution was vortexed and centrifuged for 1 min at 12 000*g*. The organic phase was dried under nitrogen at room temperature and the residue was resuspended in 2 ml HPLC-grade methanol.

Analytical RP-HPLC was performed using an Atlantis dC 18 Waters column (4.6 × 250 mm, 300 Å, 5 µm) protected with a similar guard column (4.6 × 20 mm) with a flow rate of 1.0 ml min^{-1} and a 20 min linear elution gradient from 10% to 90% acetonitrile in 0.08% TFA (gradient *A*) or from 10% to 20% acetonitrile in 0.08% TFA (gradient *B*). The wavelength used for detection was 214 nm (996 Photodiode Array Detector, Waters) because it offers the highest sensitivity for quantification of substrate and reaction products.

Liquid chromatography/mass spectrometry (LC-MS) analyses were carried out on an electrospray ionization-ion trap mass spectrometer (LCQ Advantage, Thermo Electron) operated in positive mode coupled to a Surveyor HPLC system (Thermo Electron), using the same column as above with gradient *A*. The 1.0 ml min^{-1} flow rate was split post-column, with a flow of only 0.2 ml min^{-1} being sent to the electrospray ionization source. The latter was operated at a capillary voltage of 32 V, a spray voltage of 4.5 kV and a capillary temperature of 543 K. For collision-induced dissociation in the ion trap, helium was used as the collision gas and the collision energy was set to 35% (Aimé *et al.*, 2008).

NMR spectroscopy analyses were performed on a Bruker Avance DSX-500 spectrometer (Bruker, France) operating at 500 MHz for protons and equipped with an HR-MAS probe. The two polyphenolic products were separated by semi-preparative RP-HPLC using an Atlantis dC-18 Waters column (10 × 250 mm, 300 Å, 5 µm) at a flow rate of 3 ml min^{-1} with gradient *B*. The ethyl acetate extract of each product was dissolved in 50 µl MeOD-d_4 with TSP in D_2O (2 m*M* final concentration of each flavan-3-ol) and transferred into a 4 m*M* HR-MAS rotor. 128 scans of 32 000 data points were acquired with a spinning frequency of 5000 Hz and with a spectral width of 5000 Hz (10 p.p.m.), a recycle delay of 2 s and a flip angle of 90° for 5.2 µs at 6 dB. Preliminary data processing was carried out with *TopSpin* v.2.0 software (Bruker). Spectra of authentic standards of catechin and epicatechin were recorded for reference.

Analytical chiral HPLC was achieved following the procedure described by Tanner *et al.* (2003) using a 250 × 4.6 mm, 5 µm Chiralcel OJ-H column (Daicel) protected with a similar guard column (10 × 4 mm) and eluted with a solvent consisting of hexane/ethanol [70/30(*v*/*v*)] at a flow rate of 0.5 ml min^{-1}, with detection at 214 nm. The choice of such a column was justified by its ability to separate (+)-catechin from (−)-catechin.

2.5. Synthesis of (*R*)-NADPD and (*S*)-NADPD

(*R*)-NADPD and (*S*)-NADPD were enzymatically synthesized using the method described by Pollock & Barber (2001). The final products were solubilized in 5 ml deionized water. Their purity and effective deuteration were checked by LC-MS experiments prior to lyophilization and storage at 353 K.

2.6. Crystallization, data collection and structure determination

The first crystallization trials were assayed in buffer *A* with the native and (N-terminally) His_6-GB1-tagged proteins in parallel. Different commercial screens were used. Positive hits were only obtained for native ANR. Of several conditions, two could be optimized. The first yielded poorly diffracting needles, whereas the second gave hexagonal-shaped crystals and these crystallization conditions were optimized at 293 K by vapour diffusion with hanging drops in Linbro plates. Drops were prepared on siliconized cover slides and equilibrated against 0.5 or 1.0 ml reservoir solution. The volume of the drops was 2 or 4 µl in total, with equal volumes of protein and reservoir solutions. The protein concentration was 17.33 mg ml^{-1}, containing five equivalents of $NADP^+$ or NADPH. ANR crystals appeared in drops equilibrated over reservoirs containing 100–250 m*M* ammonium acetate, 25% PEG 3350, 100 m*M* HEPES pH 7.5 within one week and grew to their final size within two weeks. The largest crystals measured 375 × 375 × 125 µm and were cryoprotected with reservoir solution supplemented with 17% glycerol prior to data collection. The crystals diffracted to 2.2 Å resolution.

Data collections were performed at 100 K using synchrotron radiation on beamline ID14-1 at the ESRF (European Synchrotron Radiation Facility) facility. Because of the long *c* unit-cell parameter (440.4 Å), the oscillation range was set to 0.5° per frame. Data from two crystals were merged in order to obtain full completeness. Data were indexed in space group $P6_122$, processed with the program *MOSFLM* (Leslie, 2006) and scaled with the program *SCALA* (Evans, 2006) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Data-collection statistics are given in Table 1. According to the unit-cell size, a Matthews coefficient of 2.13 Å^3 Da^{-1} and a solvent content of 47.8% were obtained for one protein molecule in the asymmetric unit. The ANR structure was solved by molecular replacement using the program *MOLREP* (Vagin & Teplyakov, 1997) with the atomic coordinates of the DFR molecule as a starting model (PDB code 2c29; Petit *et al.*, 2007), which shares 47% sequence identity with ANR. An early inspection of electron-density maps calculated from the *MOLREP* solution indicated that only 184 amino acids (belonging to the N-terminal domain) could be kept for further refinement. Subsequent model building and refinement were performed using the programs *Coot* (Emsley & Cowtan, 2004) and *REFMAC* (Murshudov *et al.*, 1997) iteratively. The model validity was

Table 1
Data-collection statistics.

Values in parentheses are for the last resolution shell.

Crystal form	Hexagonal	Monoclinic
Space group	$P6_122$	$P2_1$
Unit-cell parameters		
a (Å)	50.470	84.518
b (Å)	50.470	51.012
c (Å)	440.406	86.113
β (°)	90.0	110.3
Resolution range (Å)	43.73–2.22 (2.34–2.22)	79.31–3.17 (3.34–3.17)
Total reflections	201269 (30174)	39830 (5843)
Unique reflections	17253 (2484)	11952 (1725)
Completeness (%)	96.4 (98.8)	99.9 (99.9)
$R_{\rm merge}$† (%)	9.0 (35.6)	8.9 (38.9)
Multiplicity	11.7 (12.2)	3.3 (3.4)
$\langle I/\sigma(I)\rangle$	5.6 (2.0)	6.0 (2.0)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average of symmetry-related observations of unique reflections.

Table 2
Refinement statistics.

Values in parentheses are for the last resolution shell.

Crystal form	Hexagonal	Monoclinic
Resolution range (Å)	30.0–2.22 (2.28–2.22)	79.81–3.17 (3.25–3.17)
Reflections used for refinement	16263 (1189)	11368 (844)
Reflections used for $R_{\rm free}$	851 (55)	576 (37)
Completeness (%)	98.7 (95.7)	99.7 (99.7)
$R_{\rm work}$† (%)	19.9 (22.2)	21.8 (27.2)
$R_{\rm free}$‡ (%)	27.0 (29.1)	29.9 (36.1)
Protein atoms	2309	4288
Water molecules	124	1
Ions (Cl^-)	1	2
Average B factors (Å^2)		
Protein atoms	43.5	47.3
Solvent atoms	48.2	52.1
R.m.s.d. from ideal values		
Bond lengths (Å)	0.015	0.010
Bond angles (°)	2.14	1.95
Ramachandran plot		
Favoured regions (%)	98.0	94.4
Allowed regions (%)	100.0	100.0
Disallowed regions (%)	0.0	0.0

† $R_{\rm work} = \sum_{hkl}|F_{\rm obs} - F_{\rm calc}| / \sum_{hkl}|F_{\rm obs}|$, where $F_{\rm obs}$ and $F_{\rm calc}$ are the observed and calculated structure-factor amplitudes, respectively. ‡ $R_{\rm free}$ is the same as $R_{\rm work}$ but for 5% of the total reflections chosen at random and omitted from refinement.

checked with the programs *PROCHECK* (Laskowski *et al.*, 1993), *SFCHECK* (Vaguine *et al.*, 1999) and *WHAT IF* (Vriend & Sander, 1993). Protein stereochemical restraints were taken from Engh & Huber (1991). Water molecules were positioned in well defined positive ($mF_o - DF_c$) residual densities with a lower cutoff of 3σ if they participated in hydrogen bonds to the protein or other water molecules. The final protein model consists of residues 9–93, 103–155 and 163–336 plus 124 water molecules and one chloride ion.

New crystallization trials were assayed in buffer *B*, which contains only 50 m*M* NaCl, with the native and (N-terminally) His_6-GB1-tagged proteins in parallel. Crystals of the native protein were obtained with a reservoir solution containing 200 m*M* $MgCl_2$, 100 m*M* bis-tris pH 6.5 and 25% PEG 3350. They diffracted to a resolution of 3.17 Å (space group $P2_1$). The structure was solved by molecular replacement using the refined structure obtained from the first crystal type as a search model. The asymmetric unit consists of two protein chains. The refinement procedure was the same as that used for the hexagonal structure, except that tight noncrystallographic symmetry restraints were applied between chains. The final model consists of residues 7–92, 103–150 and 165–337 for chain *A*, residues 9–97, 102–152 and 163–337 for chain *B*, two chlorine ions and one water molecule.

For both of the structures, data-collection statistics and final refinement statistics are given in Tables 1 and 2, respectively. Figs. 4(*b*), 5 and 6 were drawn using the molecular-graphics program *PyMOL* (DeLano, 2002). Atomic coordinates have been deposited in the RCSB Protein Data Bank with accession codes 2rh8 and 3hfs.

3. Results

3.1. Production and stability of the enzyme

Several attempts to overexpress the enzyme using different bacterial expression vectors such as pET24+ (Novagen) or pQE-30Xa (Qiagen) failed. Owing to these difficulties, pETGB_1a expression vector was used to induce a high protein-expression level (Huth *et al.*, 1997) and to enhance the solubility and stability of the fusion protein (Zhou *et al.*, 2001). Significant amounts of the fusion protein (120–150 mg per litre of culture) were obtained with this vector when the protein expression was performed at 288 K. At higher temperatures (303 or 310 K) the protein of interest mainly precipitates into inclusion bodies.

Cleavage of the His_6-GB1 domain by TEV protease leads to a protein of poor stability. Different buffers, pH, salts and concentrations were assayed. The optimal stability conditions were found to be tricine pH 7.5 in the presence of 500 m*M* NaCl and other effectors (see §2). Good results were also obtained with bis-tris buffer pH 7.3 or HEPES buffer pH 7.5 with similar salt concentrations.

3.2. Enzyme activity

To assay the activity of grape ANR with each of the anthocyanidins, RP-HPLC and LC-MS analyses were performed on the MeOH-resuspended reaction products at a low salt concentration (5 m*M*). No activity was observed at concentrations of sodium chloride close to or higher than 200 m*M* (data not shown), a result which is consistent with that obtained for Mt-ANR (Xie *et al.*, 2004).

When cyanidin was used as a substrate, LC-MS chromatograms showed two major products (Fig. 2). The compounds eluting in peaks 1 and 2 were assigned as catechin and epicatechin, respectively, according to their m/z value of 291.0 and their retention times (11.7 and 12.3 min), which matched those of commercial standards. Two different reaction products were also synthesized when either pelargonidin or delphinidin was used as the substrate. Peaks 1′ and 2′ (retention times of 13.2 and 13.7 min) and 1″ and 2″ (retention times

of 9.4 and 11.1 min) were assigned to afzelechin and epiafzelechin (m/z 275.0) and to gallocatechin and epigallocatechin (m/z 306.9), respectively. Peaks 3, 3′ and 3″ indicate the presence of anthocyanidin dimers as minor products.

The evolution of the catechin and epicatechin peak areas (when cyanidin is used as a substrate) as a function of the incubation time is similar (Fig. 3*a*). The relative concentration of catechin *versus* epicatechin remains nearly constant and is close to 1. Similar results were obtained when alternative buffers such as MES pH 6.0, bis-tris pH 6.45, HEPES pH 7.5 or KH_2PO_4 pH 7.4 were used.

The chemical identification of the reaction products (peaks 1 and 2 in Fig. 2*a*) was confirmed by ^{1}H NMR. The spectra agree reasonably with those of catechin and epicatechin standards, as shown by each of the three expanded regions corresponding to the B, A and C rings (Fig. 3*b*), and with those published by Berregi *et al.* (2003) and Mirabel *et al.* (1999). The product associated with peak 1 is catechin (H2′ at 6.86 p.p.m., H6′ at 6.75 p.p.m., H8 at 5.89 p.p.m., $H4_1$ or $H4_2$ at 2.52 p.p.m.). The product associated with peak 2 is epicatechin (H2′ at 6.99 p.p.m., H6′ at 6.83 p.p.m., H8 at 5.96 p.p.m., $H4_1$ or $H4_2$ at 2.75 p.p.m.). For both catechin and epicatechin the signals generally assigned to H2 and H3 were difficult to detect because of their overlap with the resonance of residual water.

Further analysis of the reaction products was performed using chiral chromatography in order to identify the epimer stereochemistry. The chromatogram shows two distinct peaks (retention times of 37.70 and 39.70 min), the second peak coinciding with that of the (−)-catechin standard (Fig. 3*c*). The first peak reasonably matches that observed when the epicatechin ANR product isolated by RP-HPLC experiments (peak 2 in Fig. 2*a*) is eluted from the chiral column. It does not correspond to any of the three standards and is thus attributed to (+)-epicatechin.

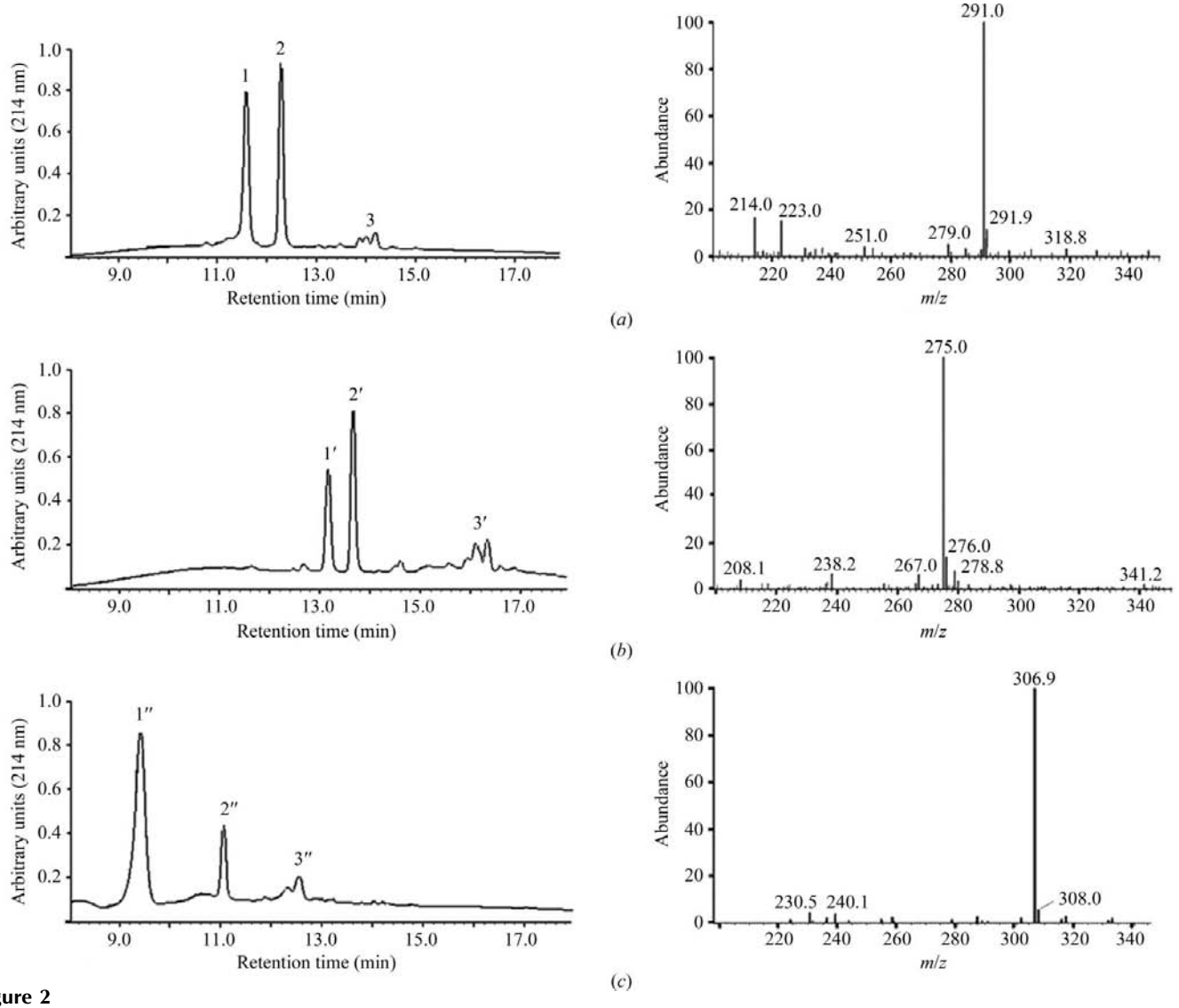

Figure 2
LC-MS analysis of the reaction products formed by the action of Vv-ANR on anthocyanidins in the presence of NADPH. The incubation time was set to 30 min and the pH was adjusted to 7.0. The substrate used in (*a*) was cyanidin, that in (*b*) was pelargonidin and that in (*c*) was delphinidin. For each substrate, the obtained chromatogram is shown as well as the ESI/MS spectrum associated with peak numbers 1, 1′ and 1″.

3.3. Non-enzymatic epimerization of flavan-3-ols

Since two isomers are produced during the reaction, the question arises as to whether they are independently produced by the enzyme or whether they result from a possible non-enzymatic epimerization. To answer this question, each of the commercially available standards [(−)-epicatechin, (−)-catechin and (+)-catechin] was individually incubated without the enzyme for 30 min at 303, 323 and 353 K in 100 m*M* Tris–HCl pH 7.0. In all cases epimerization was only observed at 353 K and the degree of conversion for (−)-epicatechin, (−)-catechin and (+)-catechin was found to be quite low and close to 11, 4 and 2.5%, respectively. No epimerization was observed at temperatures of ≤323 K. These results agree reasonably well with those reported by Komatsu *et al.* (1993) or more recently by Wang & Helliwell (2000).

Under the same conditions as above, no epimerization was evident using (+)-epicatechin. Since no commercial standard was available for this latter compound, Vv-ANR was used to produce it.

Moreover, neither (−)-catechin nor (+)-epicatechin (when used separately) was epimerized at 303 K in the presence of NADPH, whether Vv-ANR was present or not.

3.4. Stereospecificity of the enzyme

To determine the stereospecificity of the enzyme and to confirm that ANR uses NADPH as a direct hydride donor as usually found in the catalysis of enzymes belonging to the SDR family (Jörnvall *et al.*, 1995; Tanaka *et al.*, 1996; Oppermann *et al.*, 2003; Kavanagh *et al.*, 2008), the enzymatic reaction was initiated using cyanidin as a substrate and NADPH, (*R*)-NADPD or (*S*)-NADPD as a cofactor. Analysis of the resulting positive-ion mass spectra showed that the $(M + H)^+$ peak corresponding to the reaction products shifts from an *m*/*z* value of 291 when NADPH or (*R*)-NADPD is used as cofactor to a value of 293 when (*S*)-NADPD is used. The mass of 291 corresponds to the positive-ion mass of nondeuterated flavan-3-ols, whereas the mass of 293 corresponds to that of dideuterated flavan-3-ols. This result clearly establishes that ANR is strictly pro-*S* stereospecific and that the reaction mechanism involves two hydride transfers from two distinct NADPH molecules. Recently, we used the deuterium-labelled coenzyme NADPD and MS/MS analysis of deuterated products (Gargouri, Chaudière *et al.*, 2009) to demonstrate that the regiospecificity of hydride transfers was (C2, C4). This study also showed that no reduction intermediate was released in the medium during the catalytic cycle. In the second study, we demonstrated that the steady-state kinetic mechanism of the enzyme was hyperbolic sequential and more precisely rapid-equilibrium ordered Bi Uni Uni Bi (Gargouri, Gallois *et al.*, 2009).

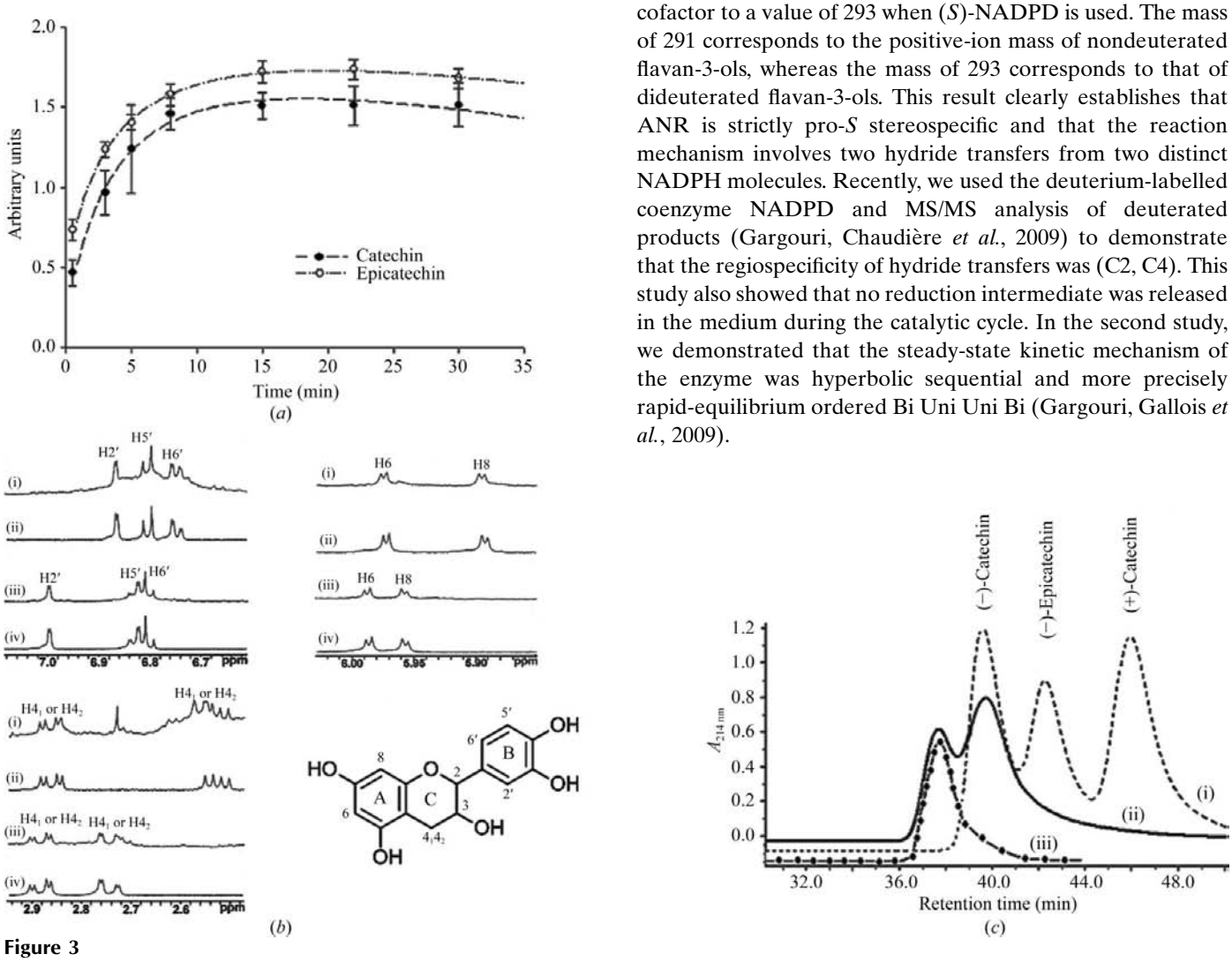

Figure 3
Time-course of formation and identification of the reaction products. (*a*) Evolution of the reaction products in Tris–HCl buffer as a function of the incubation time. Each point represents the RP-HPLC peak-area average of four independent replicate experiments. (*b*) Expanded regions of the ^{1}H NMR spectra for the ANR reaction products (cyanidin used as substrate) and for epicatechin and catechin standards. The 6.60–7.10, 5.85–6.02 and 2.46–2.96 p.p.m. regions correspond to the protons of rings B, A and C, respectively. (i) ANR product of peak 1 in Fig. 2(*a*), (ii) (−)-catechin standard, (iii) ANR product of peak 2 from Fig. 2(*a*), (iv) (−)-epicatechin standard. (*c*) Chiral HPLC analysis of the reaction products obtained by incubation of Vv-ANR with cyanidin and NADPH. (i) Mixture of the three standards (−)-catechin, (−)-epicatechin and (+)-catechin. The retention times are 39.7, 42.2 and 45.9 min, respectively. (ii) ANR reaction products. (iii) Product eluted in peak 2 of Fig. 2(*a*).

3.5. Structure description

Two crystal forms have been obtained. The structure described here is that of the hexagonal form at 2.2 Å resolution. The monoclinic crystal form (3.2 Å) is very similar and will only be referred to in passing. The structure corresponds to the apo form of the protein, crystals of which grew in the presence of a high salt concentration. Although NADPH was added to the crystallization conditions, it did not bind to the protein.

The ANR subunit contains 12 α-helices and 12 β-strands (Fig. 4*a*). Its overall structure is consistent with that found for all members of the SDR family. It presents two distinct domains. The N-terminal end adopts a Rossmann-fold motif, with a twisted parallel seven-stranded β-sheet flanked on both sides by a total of six α-helices (α1–4, α6 and α9). The C-terminal end is shorter and composed of six α-helices and five β-strands, of which three form a β-sheet ($\beta2'$, $\beta3'$ and $\beta4'$).

A three-dimensional alignment search using the protein structure-comparison service *SSM* (Krissinel & Henrick, 2004) at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/ssm) showed that the best superimposition of the ANR final model was obtained with two proteins of the SDR family: dihydroflavonol reductase (DFR) from *V. vinifera* (Petit *et al.*, 2007; PDB code 2c29) and vestitone reductase (VR) from alfalfa (Shao *et al.*, 2007; PDB code 2p4h). When ANR was superimposed on DFR or VR, an r.m.s. deviation of 1.60 or 1.70 Å was obtained, with 253 or 248 aligned C^{α} positions (20/22 or 20/24 secondary-structure elements) and a sequence identity of 47% or 42%, respectively.

Nevertheless, several differences are noted between the secondary-structure elements. At the N-terminal end, the small α-helix present in DFR and VR between helix α2 and strand β3 is not observed in ANR. The protein segment located between β4 and α4 could not be modelled in ANR. Helices α4 and α6 appear to be seven and eight amino acids shorter than in DFR. In the C-terminal domain, the ANR structure exhibits a α-helix ($\alpha1'$) that is longer than the short 3_{10}-helix present in DFR and that is absent in VR. $\beta1'$ builds a short sheet with a symmetry-related $\beta5'$ strand, whereas in DFR it builds a sheet with the intramolecular $\beta5'$ strand. This $\beta1'$–$\beta5'$ interaction between symmetry-related subunits is also observed in the monoclinic 3.2 Å resolution structure. The 156–163 region corresponding to the α5 helix observed in both DFR and VR is disordered. Finally, the α7 and α8 ANR helices are longer than in DFR. The largest r.m.s. deviations involve the C-terminal part of the proteins. Whereas most of the secondary-structure elements are conserved, their relative positions have shifted (Fig. 4*b*). Since the DFR and VR structures are

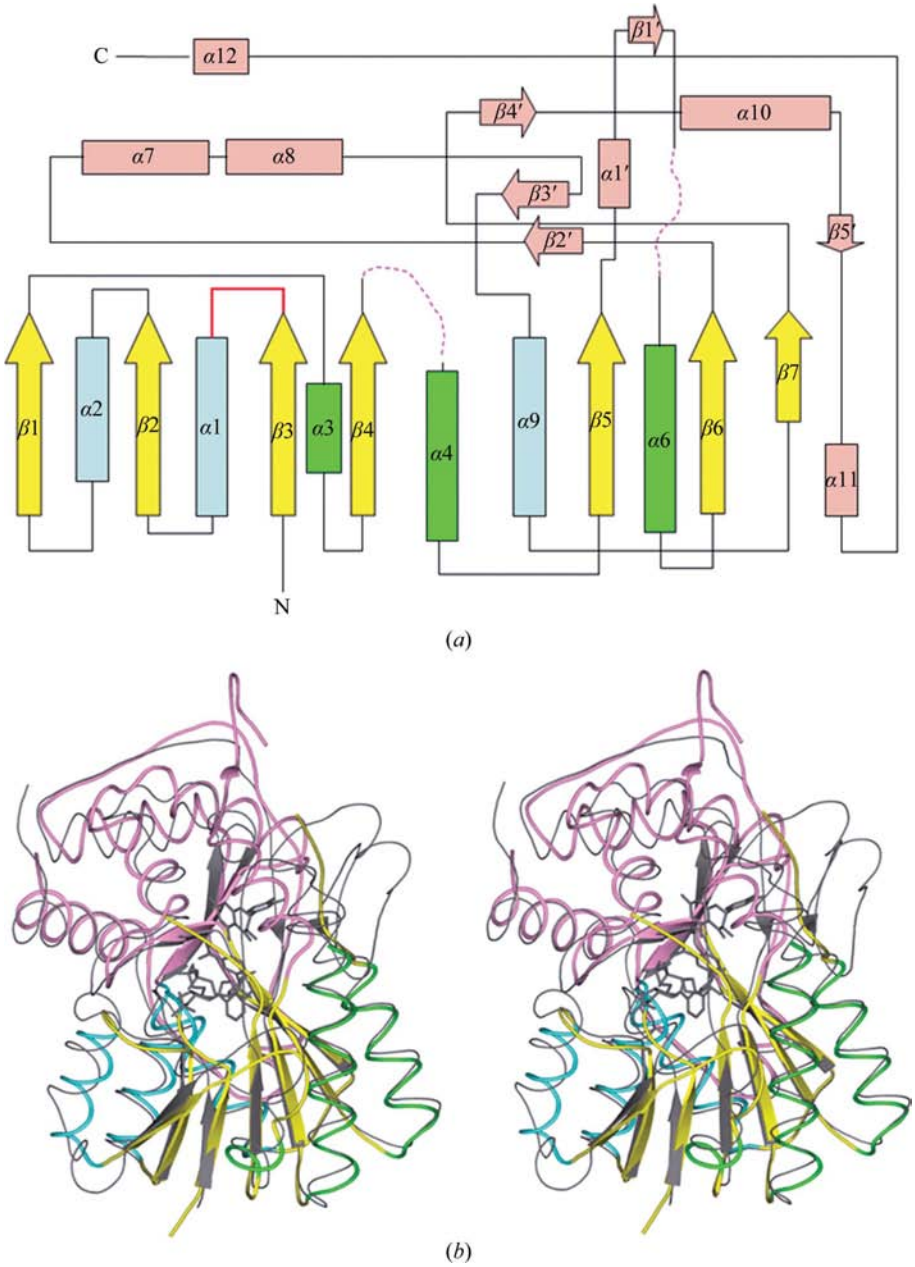

Figure 4
ANR topology and three-dimensional structure. (*a*) ANR topology. The NADPH-binding domain (N-terminal domain) is composed of a seven-strand β-sheet (yellow) surrounded by cyan (above) and green (below) α-helices. The C-terminal region is shown in pink. Next to the N-terminus, the glycine-rich motif constituting the NADPH-binding site observed in DFR is shown in red. Pink dotted lines represent disordered regions. (*b*) Superimposition of the DFR–NADPH–DHQ ternary complex and ANR structures. DFR, NADPH and DHQ are coloured dark grey. ANR is coloured according to the topology diagram.

very similar to one another (23/25 aligned secondary-structure elements), ANR will be compared with DFR in the following.

3.5.1. NADPH-binding site. The $\beta1\alpha1$ motif includes the glycine-rich loop Gly16-Gly17-Thr18-Gly19-Phe20-Val21-Ala22-Ser23 to which the cofactor is anticipated to bind. A superimposition of the corresponding regions of the DFR–NADP$^+$–DHQ ternary complex and apo ANR shows that the conformations of the glycine-rich loops are very similar (Fig. 5*a*). In the ANR structure the expected position of the NADPH central diphosphate group is occupied by a chloride ion which is coordinated by the main-chain amino groups of residues Val21 and Val91 and three water molecules. One of three waters is in the same position as a highly conserved water molecule that is observed in Rossmann-fold dinucleotide-binding domains (Bottoms *et al.*, 2002). Residue Val91 belongs to a loop which links strand $\beta4$ to helix $\alpha4$. The C-terminal part of this loop (residues 94–102) is highly disordered in the ANR structure, whereas its N-terminal part is well defined. It lies much closer to the glycine-rich loop than is observed in the DFR ternary complex.

3.5.2. Catalytic site. The ANR putative catalytic triad Ser131, Tyr168 and Lys172 adopts a three-dimensional arrangement which is very different from the arrangement observed in the ternary complex DFR–NADP$^+$–DHQ. This is illustrated in Figs. 5(*b*) and 5(*c*), where the structures of the concerned regions are depicted for the ANR and DFR ternary complex structures. The first amino acid of the triad, Ser131, is located at the C-terminal end of strand $\beta5$ and lies at the same position in both structures. This is not the case for the two other amino acids, Tyr168 and

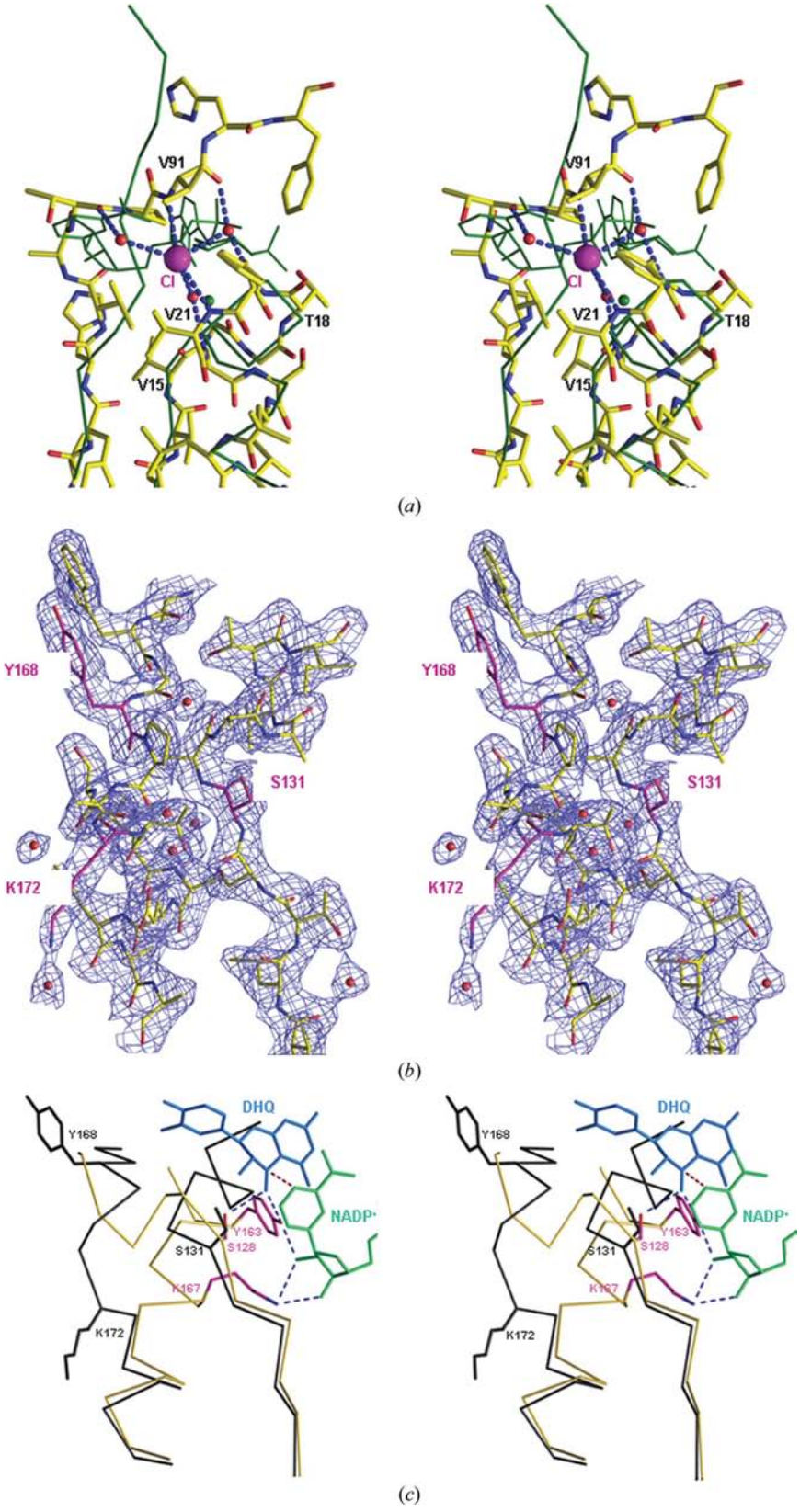

Figure 5
Stereoviews of the glycine-rich loop region and the catalytic site. (*a*) Stereoview of the glycine-rich loop region in the apo ANR and DFR ternary complex structures. ANR C atoms are shown in yellow and all atoms of the DFR complex are shown in dark green. The chlorine ion is shown in pink. The conserved water molecules are shown in red for ANR and in dark green for DFR. (*b*) The electron-density map contoured at 1.2σ in the vicinity of the catalytic triad in the ANR structure. The C atoms of residues Ser131, Tyr168 and Lys172 are shown in pink. (*c*) Superimposition of the catalytic sites of ANR and DFR. The ribbons of the ANR and DFR structures are shown in yellow and dark grey, respectively. The C atoms of the catalytic residue side chains of ANR and DFR are pink and cyan, respectively. Dihydroquercetin and NADP$^+$ molecules from the DFR structure are shown in blue and green, respectively.

Lys172. Contrary to what is observed in the DFR ternary complex, these residues do not belong to helix α6, which is shortened at its C-terminal end in ANR. Instead, they belong to a long loop which relates strand β1′ to helix α6 and their side chains are turned away from the nucleotide-binding site and are largely exposed to solvent.

4. Discussion

As previously observed in *M. truncatula* and *A. thaliana* (Xie *et al.*, 2004), two different reaction products are synthesized by grape ANR. These are 2,3-*cis* and 2,3-*trans* flavan-3-ols. The stereochemistry of the 2,3-*trans* compounds is unchanged whatever the enzyme (Mt-ANR, At-ANR or Vv-ANR). In each case the 2*S*,3*R* isomer is produced. This is no longer the case for 2,3-*cis* flavan-3-ol. In grapes, the 2*S*,3*S* absolute configuration has been identified by chiral chromatography, whereas the 2*R*,3*R* isomer is synthesized by Mt-ANR and At-ANR. When cyanidin is used as substrate, (−)-catechin and (+)-epicatechin are produced by Vv-ANR and (−)-catechin and (−)-epicatechin are produced by Mt-ANR and At-ANR.

For Mt-ANR and At-ANR (Xie *et al.*, 2004), the minor presence of the 2*S*,3*R* isomer was assumed to result from a non-enzymatic C2-epimerization of the 2*R*,3*R* isomer. For Vv-ANR, we conclude from the lack of chemical epimerization at 303 K and the time-invariant ratio of (−)-catechin and (+)-epicatechin that the two products are actually synthesized by the enzyme.

Another noticeable point of this investigation concerns the nature of the enantiomers synthesized by Vv-ANR. To date, these have only rarely been encountered in nature, in contrast to (+)-catechin and (−)-epicatechin, which are widely distributed in plants. To the best of our knowledge, (−)-catechin has only been isolated from *Chamaebatia foliolosa* (Nahrstedt *et al.*, 1987) and from *Centaurea maculosa* (Bais *et al.*, 2002) and (+)-epicatechin has only been isolated from *Palmae* (Delle Monache *et al.*, 1972) and *Paullinia cupana* (Yamaguti-Sasaki *et al.*, 2007). Both of them have been demonstrated to have phytotoxic activity (Bais *et al.*, 2003).

In grapes, the monomeric, oligomeric or polymeric flavan-3-ol composition of different cultivars, as well as their changes during berry development, has been widely studied by several authors (Souquet *et al.*, 1996; de Freitas & Glories, 1999; Monagas *et al.*, 2003; Serratosa *et al.*, 2008). (+)-Catechin, (−)-epicatechin, (+)-gallocatechin and some of their derivatives such as (−)-epicatechin gallate are the major monomeric flavan-3-ols that were unanimously reported. To our knowledge, neither (−)-catechin nor (+)-epicatechin have been mentioned in grapes. It is noteworthy that in most of these investigations identification of the flavan-3-ols was achieved thanks to the retention times of standards using reverse-phase HPLC; that is to say, using a technology that is able to separate *cis* and *trans* isomers, such as epicatechin and catechin, but is not able to discriminate enantiomers. The lack of stereospecific assays in

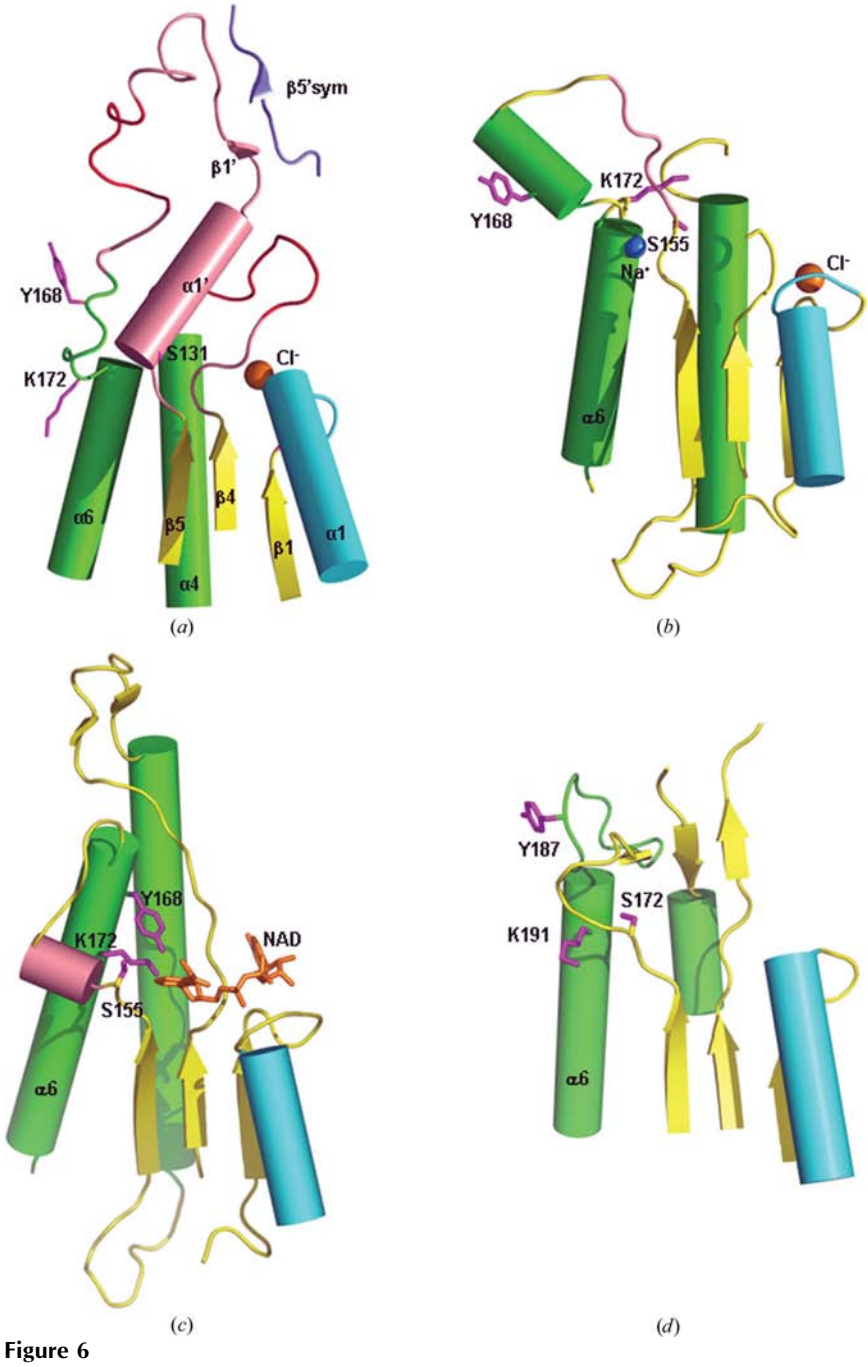

Figure 6
Comparison of the catalytic triad geometry in (*a*) Vv-ANR, (*b*) ABAD, (*c*) the ternary complex of the ABAD C214R variant and (*d*) Cp-SCR. The secondary-structure elements are coloured as in the ANR topology diagram. Catalytic residue side chains are coloured magenta.

the literature could be one of the reasons why the two ANR products have not been identified to date.

During the course of a recent investigation of the structural and functional properties of recombinant LAR1 from *V. vinifera*, we confirmed by chiral chromatography that this enzyme catalyses the formation of (+)-catechin (Maugé *et al.*, work to be published) as previously demonstrated by Tanner *et al.* (2003) in *Desmodium uncinatum*. It is therefore not surprising to find the presence of (+)-catechin in grape extracts. In contrast, according to our results, no valid explanation could justify the presence of (−)-epicatechin as a major product and the mechanism of its formation remains to be understood. Is there another biosynthetic pathway that is responsible for the production of (−)-epicatechin? This suggestion has been put forward by Szankowski *et al.* (2009), who have demonstrated an increase of epicatechin in transgenic plants when anthocyanidin synthase is silenced.

As far as the conformation of the ANR is concerned and as mentioned above, the active site is wide open with the side chains of the catalytic residues Tyr168 and Lys172 turned away from the nucleotide-binding site (Fig. 6*a*). This geometry should be compared with those described for two other SDRs: the human amyloid-β peptide (Aβ) binding alcohol dehydrogenase (ABAD) and the carbonyl reductase from *Candida parapsilosis* (Cp-SCR; Zhang *et al.*, 2008). For ABAD, one structure was obtained by cocrystallizing the protein in the presence of NADH and the Aβ peptide (Lustbader *et al.*, 2004) using a high concentration of sodium chloride (2.5 *M*) in the reservoir solution and the second was obtained by co-crystallizing the C214R variant with NAD^+ and the inhibitor AG18051 (Kissinger *et al.*, 2004) with 0.2 *M* sodium chloride in the protein solution. In the first structure, neither NADH nor the Aβ peptide were observed. The catalytic site appeared to be wide open with the catalytic residues Ser155, Tyr168 and Lys172 far apart from each other (Fig. 6*b*). Two ions were positioned, a sodium cation near the glycine-rich loop and a chloride ion near residues 173 and 175. In the second structure both NAD and inhibitor were observed, with the catalytic residues oriented towards them (Fig. 6*c*) in a manner similar to that observed in the DFR ternary complex. The difference in the catalytic site geometry between these two ABAD structures is directly related to the shortening of the α6 helix at its N-terminal end when the site is wide open, as observed in the ANR structure.

The crystals of Cp-SCR were obtained with 150 m*M* sodium chloride in the reservoir solution. This protein forms a tetramer, with four inactive subunits showing two different active-site conformations. In one of them the entrance to the NADPH pocket is blocked by a surface loop, but the other conformation exhibits the same nonfunctional catalytic site (Fig. 6*d*) with the α6 helix one turn shorter at its N-terminal end and the catalytic residues Tyr187 and Lys191 arranged far away from each other, leading to a wide-open site as for ANR.

To our knowledge, no other SDR structures presenting an alteration of the catalytic triad geometry are available in the Protein Data Bank. ANR is the third protein for which a nonfunctional catalytic site has been described. In each case, the crystals used for structure determination were always obtained in the presence of high salt concentrations. These observations led us to consider that the presence of a high concentration of sodium chloride (500 m*M*) could be a factor responsible for the alteration of the protein conformation and consequently for the inactivity of the enzyme. Is this factor the only one preventing enzyme activity? The absence of any structural data for the ABAD protein crystallized without peptide prevents us from coming to a proper conclusion, as the Aβ peptide (even if it is not observed in the structure) is considered by the authors to alter the catalytic site geometry when it binds to the protein (Lustbader *et al.*, 2004).

In the monoclinic crystal form of ANR, it turns out that the protein exhibits the same features as those observed in the hexagonal form, *i.e.* no NADPH in the active site, a chloride ion near the glycine-rich loop, a comparable geometry of the catalytic site and similar intermolecular interactions which involve strands $\beta1'$ and $\beta5'$ of symmetry-related molecules. For the monoclinic crystals, the initial concentration of chloride ions in the crystallization drop is estimated to be greater than 200 m*M*, a concentration close to that at which enzyme inactivation is observed. Therefore, again it cannot be ruled out that a high salt concentration could alter the conformation of the Vv-ANR active site and thus inhibit the enzyme.

We thank the European Synchrotron Radiation Facility (Grenoble, France) for provision of synchrotron-radiation facilities and the staff of beamline ID14-1 for their kind assistance. We are grateful to Claire Le Henaff for her help during the molecular-biology and expression experiments. This work was supported by the CIVB (Conseil Interprofessionnel du Vin de Bordeaux).

References

Aimé, C., Plet, B., Manet, S., Schmitter, J. M., Huc, I., Oda, R., Sauers, R. R. & Romsted, L. S. (2008). *J. Phys. Chem. B*, **112**, 14435–14445.
Aron, P. M. & Kennedy, J. A. (2008). *Mol. Nutr. Food Res.* **52**, 79–104.
Bagchi, D., Bagchi, M., Stohs, S. J., Das, D. K., Ray, S. D., Kuszynski, C. A., Joshi, S. S. & Pruess, H. G. (2000). *Toxicology*, **148**, 187–197.
Bais, H. P., Walker, T. S., Kennan, A. J., Stermitz, F. R. & Vivanco, J. M. (2003). *J. Agric. Food Chem.* **51**, 897–901.
Bais, H. P., Walker, T. S., Stermitz, F. R., Hufbauer, R. A. & Vivanco, J. M. (2002). *Plant Physiol.* **128**, 1173–1179.
Berregi, I., Santos, J. I., del Campo, G. & Miranda, J. I. (2003). *Talanta*, **61**, 139–145.
Bogs, J., Downey, M. O., Harvey, J. S., Ashton, A. R., Tanner, G. J. & Robinson, S. P. (2005). *Plant Physiol.* **139**, 652–663.
Bottoms, C. A., Smith, P. E. & Tanner, J. J. (2002). *Protein Sci.* **11**, 2125–2137.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Collie, A. (2006). *Agro Food Ind. Hi-Tech*, **17**, XVII–XIX.
Cos, P., de Bruyne, T., Hermans, N., Apers, S., Berghe, D. V. & Vlietinck, A. J. (2004). *Curr. Med. Chem.* **11**, 1345–1359.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, Palo Alto, USA. http://www.pymol.org.
Delle Monache, F., Ferrari, F., Poce-Tucci, A. & Marini-Bettolo, G. B. (1972). *Phytochemistry*, **11**, 2333–2335.

Devic, M., Guilleminot, J., Debeaujon, I., Bechtold, N., Bensaude, E., Koornneef, M., Pelletier, G. & Delseny, M. (1999). *Plant J.* **19**, 387–398.
Dixon, R. A., Xie, D.-Y. & Sharma, S. B. (2005). *New Phytol.* **165**, 9–28.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Freitas, V. A. P. de & Glories, Y. (1999). *J. Sci. Food Agric.* **79**, 1601–1606.
Fujita, A., Soma, N., Goto-Yamamoto, N., Shindo, H., Katuka, T., Koizumi, T. & Hashizume, K. (2005). *Am. J. Enol. Vitic.* **56**, 336–342.
Gargouri, M., Chaudière, J., Manigand, M., Maugé, C., Bathany, K., Schmitter, J. M. & Gallois, B. (2009). Submitted.
Gargouri, M., Gallois, B. & Chaudière, J. (2009). Submitted.
Glories, Y. (1988). *Plant Flavonoids in Biology and Medicine II: Biochemical, Cellular, and Medicinal Properties*, edited by V. Cody, E. Middleton, J. B. Harborne & A. Beretz, pp. 123–134. New York: Alan R. Liss.
Huth, J. R., Bewley, C. A., Jackson, B. M., Hinnebusch, A. G., Clore, G. M. & Gronenborn, A. M. (1997). *Protein Sci.* **6**, 2359–2364.
Ikegami, A., Eguchi, S., Kitajima, A., Inoue, K. & Yonemori, K. (2007). *Plant Sci.* **172**, 1037–1047.
Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzàles-Duarte, R., Jeffery, J. & Ghosh, D. (1995). *Biochemistry*, **34**, 6003–6013.
Joseph, R. G., Tanner, G. J. & Larkin, P. J. (1998). *Aust. J. Plant Physiol.* **25**, 271–278.
Kavanagh, K. L., Jörnvall, H., Persson, B. & Oppermann, U. (2008). *Cell. Mol. Life Sci.* **65**, 3895–3906.
Kissinger, C. R., Rejto, P. A., Pelletier, L. A., Thomson, J. A., Showalter, R. E., Abreo, M. A., Agree, C. S., Margosiak, S., Meng, J. J., Aust, R. M., Vanderpool, D., Li, B., Tempczyk-Russel, A. & Villafranca, J. E. (2004). *J. Mol. Biol.* **342**, 943–952.
Komatsu, Y., Suematsu, S., Hisanobu, Y., Saigo, H., Matsuda, R. & Hara, K. (1993). *Biosci. Biotechnol. Biochem.* **57**, 907–910.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Leslie, A. G. W. (2006). *Acta Cryst.* D**62**, 48–57.
Lin, L. C., Kuo, Y. C. & Chou, C. J. (2002). *J. Nat. Prod.* **65**, 505–508.
Lustbader, J. W. *et al.* (2004). *Science*, **304**, 448–452.
Mirabel, M., Glories, Y., Pianet, I. & Dufourc, E. J. (1999). *J. Chem. Phys.* **96**, 1629–1634.
Monagas, M., Gómez-Cordovés, C., Bartolomé, B., Laureano, O. & Ricardo Da Silva, J. M. (2003). *J. Agric. Food Chem.* **51**, 6475–6481.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nahrstedt, A., Proksch, P. & Conn, E. E. (1987). *Phytochemistry*, **26**, 1546–1547.
Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B. & Jörnvall, H. (2003). *Chem. Biol. Interact.* **143**, 247–253.
Paolocci, F., Robbins, M. P., Madeo, L., Arcioni, S., Martens, S. & Damiani, F. (2007). *Plant Physiol.* **143**, 504–516.
Peters, D. J. & Constabel, C. P. (2002). *Plant J.* **32**, 701–712.
Petit, P., Granier, T., Langlois d'Estaintot, B., Manigand, C., Bathany, K., Schmitter, J. M., Lauvergeat, V., Hamdi, S. & Gallois, B. (2007). *J. Mol. Biol.* **368**, 1345–1357.
Pfeiffer, J., Kühnel, C., Brandt, J., Duy, D., Punyasiri, P. A. N., Forkmann, G. & Fischer, T. C. (2006). *Plant Physiol. Biochem.* **44**, 323–334.
Pollock, V. V. & Barber, M. J. (2001). *Biochemistry*, **40**, 1430–1440.
Punyasiri, P. A. N., Abeysinghe, I. S. B., Kumar, V., Treutter, D., Duy, D., Gosch, C., Martens, S., Forkmann, G. & Fischer, T. C. (2004). *Arch. Biochem. Biophys.* **431**, 22–30.
Rasmussen, S. E., Frederiksen, H., Krogholm, K. S. & Poulsen, L. (2005). *Mol. Nutr. Food Res.* **49**, 159–174.
Scarlbert, A. (1991). *Phytochemistry*, **30**, 3875–3883.
Serratosa, M. P., Lopez-Toledano, A., Merida, J. & Medina, M. (2008). *J. Agric. Food Chem.* **56**, 2810–2816.
Shao, H., Dixon, R. A. & Wang, X. (2007). *J. Mol. Biol.* **369**, 265–276.
Shen, G., Pang, Y., Wu, W., Liu, X., Zhao, L., Sun, X. & Tang, K. (2006). *J. Plant Physiol.* **163**, 224–227.
Souquet, J. M., Cheynier, V., Brossaud, F. & Moutonet, M. (1996). *Phytochemistry*, **43**, 509–512.
Stafford, H. A. (1990). *Flavonoid Metabolism*, pp. 63–100. Boca Raton: CRC Press.
Szankowski, I., Flachowsky, H., Li, H., Halbwirth, H., Treutter, D., Regos, I., Hanke, M. V., Stich, K. & Fischer, T. C. (2009). *Planta*, **229**, 681–692.
Tanaka, N., Nonaka, T., Nakanishi, M., Deyashiki, Y., Hara, A. & Mitsui, Y. (1996). *Structure*, **4**, 33–45.
Tanner, G. J., Francki, K. T., Abrahams, S., Watson, J. M., Larkin, P. J. & Ashton, A. R. (2003). *J. Biol. Chem.* **278**, 31647–31656.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.
Vriend, G. & Sander, C. (1993). *J. Appl. Cryst.* **26**, 47–60.
Wang, H. & Helliwell, K. (2000). *Food Chem.* **70**, 337–344.
Winkel-Shirley, B. (2001). *Plant Physiol.* **126**, 485–493.
Xiao, Y.-H., Zhang, Z.-S., Yin, M.-H., Luo, M., Li, X.-B., Hou, L. & Pei, Y. (2007). *Biochem. Biophys. Res. Commun.* **358**, 73–78.
Xie, D.-Y., Sharma, S. B. & Dixon, R. A. (2004). *Arch. Biochem. Biophys.* **422**, 91–102.
Xie, D.-Y., Sharma, S. B., Paiva, N. L., Ferreira, D. & Dixon, R. A. (2003). *Science*, **299**, 396–399.
Yamaguti-Sasaki, E., Ito, L. A., Dias Canteli, V. C., Ushirobira, T. M. A., Ueda-Nakalura, T., Dias Filho, B. P., Nakamura, C. V. & Palazzo de Mello, C. P. (2007). *Molecules*, **12**, 1950–1963.
Zhang, R., Zhu, G., Zhang, W., Cao, S., Ou, X., Li, X., Bartlam, M., Xu, Y., Zhang, X. C. & Rao, Z. (2008). *Protein Sci.* **17**, 1412–1423.
Zhou, P., Lugovskoy, A. A. & Wagner, G. (2001). *J. Biomol. NMR*, **20**, 11–14.