Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Catechol-*O*-methyltransferase in complex with substituted 3′-deoxyribose bisubstrate inhibitors

Manuel Ellermann,[a] Christian Lerner,[b] Guillaume Burgy,[b] Andreas Ehler,[b] Caterina Bissantz,[b] Roland Jakob-Roetne,[b] Ralph Paulini,[a] Oliver Allemann,[a] Heloïse Tissot,[a] Dan Grünstein,[a] Martine Stihle,[b] Francois Diederich[a]* and Markus G. Rudolph[b]*

[a]Laboratorium für Organische Chemie, ETH Zürich, Hönggerberg, HCI, CH-8093 Zürich, Switzerland, and [b]Pharma Division, Präklinische Forschung, F. Hoffmann-La Roche AG, CH-4070 Basel, Switzerland

Correspondence e-mail: diederich@org.chem.ethz.ch, markus.rudolph@roche.com

Received 11 November 2011
Accepted 10 January 2012

PDB References: catechol-*O*-methyltransferase–inhibitor complexes, 3u81; 3s68; 3nwb; 3nwe; 3r6t.

The biological activity of catechol neurotransmitters such as dopamine in the synapse is modulated by transporters and enzymes. Catechol-*O*-methyltransferase (COMT; EC 2.1.1.6) inactivates neurotransmitters by catalyzing the transfer of a methyl group from *S*-adenosylmethionine to catechols in the presence of Mg^{2+}. This pathway also inactivates L-DOPA, the standard therapeutic for Parkinson's disease. Depletion of catechol neurotransmitters in the prefrontal cortex has been linked to schizophrenia. The inhibition of COMT therefore promises improvements in the treatment of these diseases. The concept of bisubstrate inhibitors for COMT has been described previously. Here, ribose-modified bisubstrate inhibitors were studied. Three high-resolution crystal structures of COMT in complex with novel ribose-modified bisubstrate inhibitors confirmed the predicted binding mode but displayed subtle alterations at the ribose-binding site. The high affinity of the inhibitors can be convincingly rationalized from the structures, which document the possibility of removing and/or replacing the ribose 3′-hydroxyl group and provide a framework for further inhibitor design.

1. Introduction

Parkinson's disease is treated by oral administration of the prodrug L-DOPA, which is decarboxylated in the brain to the catechol-type neurotransmitter dopamine (Holm & Spencer, 1999; Katzenschlager & Lees, 2002; Keating & Lyseng-Williamson, 2005; Mercuri & Bernardi, 2005). Members of this important class of neurotransmitters are removed from the synapse either by active re-uptake into the neuron or are inactivated by chemical modification *via* amine oxidases or catechol-*O*-methyltransferase (COMT; Männistö & Kaakkola, 1999). COMT catalyzes the transfer of a methyl group from *S*-adenosylmethionine (SAM) to the catechol substrate in the presence of Mg^{2+} (Vidgren *et al.*, 1994; Guldberg & Marsden, 1975). The catechol neurotransmitter is inactivated by methylation and SAM is converted to *S*-adenosyl homocysteine (SAH). The inhibition of COMT greatly enhances peripheral L-DOPA levels as well as dopamine levels in the brain (Männistö & Kaakkola, 1999; Männistö *et al.*, 1992), which is relevant to Parkinson's disease and other disorders of the central nervous system such as schizophrenia and depression (McCarthy, 2001).

COMT is a monomeric enzyme with a central β-sheet flanked by α-helices on both sides (Fig. 1*a*). Crystal structures of human and rat COMT have been determined in complex with a plethora of catechol-type inhibitors (Ellermann *et al.*, 2009, 2011; Tsuji, Okazaki, Isaji *et al.*, 2009; Tsuji, Okazaki & Takeda, 2009; Rutherford *et al.*, 2008; Palma *et al.*, 2006; Lerner *et al.*, 2001; Bonifácio *et al.*, 2002; Vidgren *et al.*, 1994). The crystal structure of COMT in complex with SAM and catechol showed the close proximity of the activated methyl group and a hydroxyl group of the catechol, leading to the concept of bisubstrate inhibitors in which parts of the substrates are conjoined by a linker (Fig. 1*b*). Such inhibitors display a high affinity for COMT (with IC_{50} values as low as 9 n*M*) and are, by nature, competitive for both the catechol- and the SAM-binding sites (Masjost *et al.*, 2000; Lerner *et al.*, 2001, 2003; Paulini *et al.*, 2004, 2006; Ellermann *et al.*, 2009, 2011). Their principal binding mode was elucidated by crystallographic analysis (Lerner *et al.*, 2001), confirming the concept. In a next step, various sites of the bisubstrate inhibitors were modified in order to dissect their contribution to binding affinity. For instance, the N6 amino group of the SAM adenine base in the bisubstrate inhibitors tolerates alkyl substituents without loss of affinity and crystallographic analyses identified an energetically favourable expulsion of a single water molecule that is normally bound to the adenine moiety (Ellermann *et al.*, 2009). This study highlighted the stark influence of desolvation effects on COMT inhibitor binding. A more recent study further explored the SAM-binding site by incorporation of various adenine replacements into bisubstrate inhibitors. The binding modes of the most potent bisubstrate inhibitors were determined by cocrystal structure analyses, which defined the key interactions in the adenine binding site necessary for tight inhibitor binding (Ellermann *et al.*, 2011).

Here, we shift the focus from the adenine site to the ribose moiety (Fig. 1*c*). The molecular recognition of carbohydrates has attracted particular interest owing to their abundance and importance in biology (Dwek, 1996; Lis & Sharon, 1998; Davis & Wareham, 1999; Ferrand *et al.*, 2009). Earlier investigations of COMT inhibition by bisubstrate inhibitors revealed strong differences in the activities of compounds with a ribose or a deoxyribose motif (Fig. 2*a*; Paulini *et al.*, 2006). Compared with reference inhibitor **1**, which displayed an IC_{50} value of 9 n*M* in a radiochemical extraction assay (Zürcher & Da Prada, 1982), the IC_{50} value of the 2′-deoxyribose derivative **2** is increased by almost three orders of magnitude to 28 µ*M*. In contrast, the 3′-deoxyribose derivative **3** has only an approximately fourfold elevated IC_{50} value of 40 n*M* compared with **1**, clearly showing that the main energetic contribution for ribose binding originates from interactions of the 2′-hydroxyl group with the protein. The importance of the

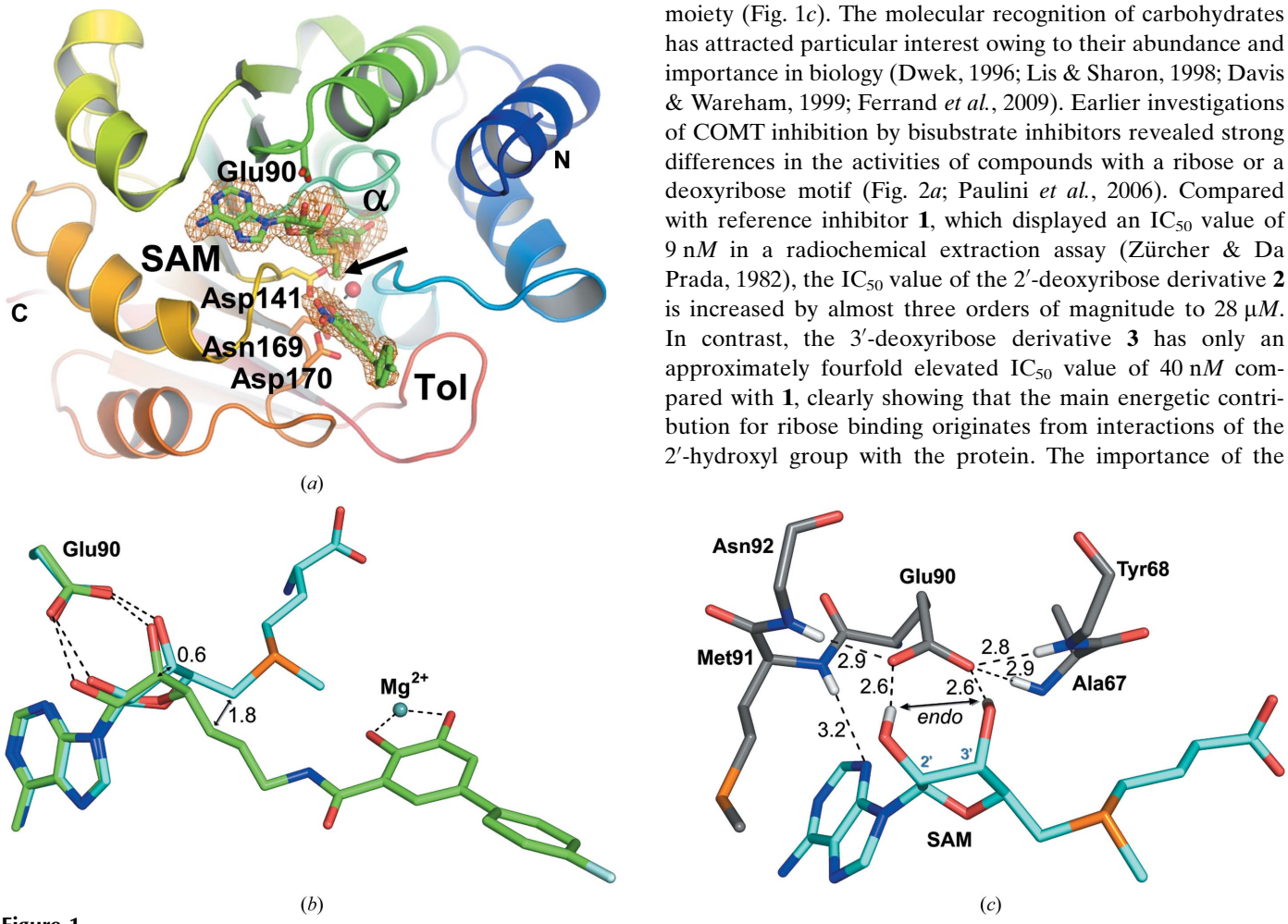
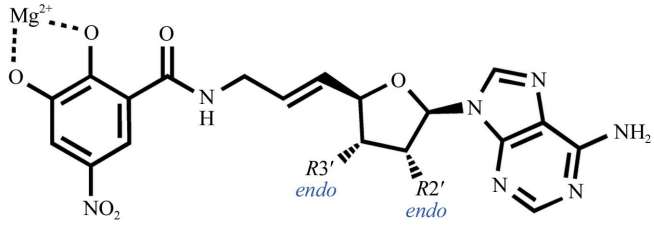


Figure 1
Overview of COMT and the concept of bisubstrate inhibitors. (*a*) Ribbon diagram of the COMT–SAM–Mg^{2+}–tolcapone complex, spectrally coloured from the N-terminus to the C-terminus. SAM and tolcapone are drawn as stick models with C atoms coloured green. Three aspartate side chains (yellow and orange) and a water molecule coordinating the Mg^{2+} are shown as stick models and as a sphere, respectively. $mF_o - DF_c$ difference electron density for the ligands is contoured as an orange mesh at 2.5 r.m.s.d. The arrow indicates the methyl group of SAM that is usually transferred to the substrate, marking the exit vector for connection with the catechol in bisubstrate inhibitors. (*b*) Superposition of SAM (cyan) and a prototype bisubstrate inhibitor (green; PDB entry 3nw9) shows a change in the C5′-pucker (arrow) and the shift of the ribose. The side chain of Glu90 is shown as a stick model and does not change its conformation appreciably. (*c*) Hydrogen-bonding environment of Glu90. Only polar H atoms that engage in interactions with Glu90 are shown as white stick models and are connected to their acceptors with black dashed lines. The H atoms of the two ribose OH groups adopt the *endo* conformation. Numbers represent heavy-atom distances in Å. Some amino-acid side chains are omitted for clarity.

2′-hydroxyl group is further corroborated by the observation that 2′-deoxy-SAM is inactive as a cofactor but 3′-deoxy-SAM still acts as a methyl donor during COMT catalysis: the K_m values for SAM and 3′-deoxy-SAM are 10 and 337 μ*M*, respectively (Borchardt *et al.*, 1976). Likewise, while 2′-deoxy-SAH is not a competitive inhibitor for COMT, 3′-deoxy-SAH and SAH inhibit COMT with inhibition constants of K_i = 138 μ*M* and K_i = 36 μ*M*, respectively (Borchardt & Wu, 1975). These data show a significant influence of the 3′-hydroxyl group on binding affinity, but an even more drastic effect of the 2′-hydroxyl group. The ratio of the K_i values for 3′-deoxy-SAH and SAH of 3.8 is remarkably close to the ratio of the IC_{50} values for **3** and **1** of 4.4. Owing to a lack of structural information on ribose-modified COMT inhibitors, a molecular explanation for the very different contributions of the ribose hydroxy groups has previously been unavailable. We report cocrystal structures of COMT–bisubstrate inhibitor complexes with substituted ribose moieties to explore the molecular interactions around the ribose-binding site.

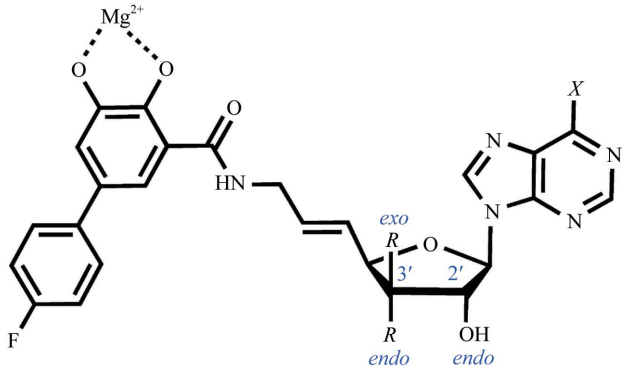
2. Materials and methods

General procedures, synthetic details and analytical data are collected in the Supplementary Material[1]. Cloning, production and purification of soluble N-terminally truncated rat and human COMT was performed as described previously (Ellermann *et al.*, 2009). COMT exists in membrane-bound and soluble forms. The soluble form lacking the N-terminal 43 residues, including an α-helical signal anchor typical for type II membrane proteins, was used in all studies described here. Residue numbering refers to the soluble form, *i.e.* is shifted by 43 compared with the membrane-bound form. IC_{50} values were determined for COMT present in rat liver homogenate using a radiochemical assay with ^{3}H-SAM as the probe (Zürcher & Da Prada, 1982; Lerner *et al.*, 2003; Paulini *et al.*, 2006; Ellermann *et al.*, 2009). Values for rat COMT may differ slightly from the human COMT data by up to a factor of two but follow the same trend, enabling meaningful comparison of crystal structural and enzyme-activity data. Crystallization and structure determination of rat COMT in complex with ligands was performed as described in Ellermann *et al.* (2009). Rat COMT crystallized more easily in our hands and appears to be more stable in solution than human COMT. At the outset of the project, rat COMT was used as a surrogate for human COMT. The proteins differ only by two residues in their active sites. Later, an active-site humanized rat COMT, *i.e.* the Met91Ile/Tyr95Cys double mutant, was used as a model system. For bisubstrate inhibitors, a 1 μ*M* solution of soluble rat COMT in 50 m*M* Tris–HCl pH 7.5, 50 m*M* NaCl, 10 m*M* DTT, 2 m*M* $MgCl_2$ was incubated at 273 K for 16 h with 10 μ*M* ligand from a 10 m*M* DMSO stock solution, concentrated to 13 mg ml^{-1} and cleared by centrifugation before crystallization at 295 K *via* sitting-drop vapour diffusion. Protein was mixed in a 2:1 volume ratio with reservoir solution consisting of 1.2 *M* ammonium sulfate, 0.1 *M* CHES pH 9.8, 0.2 *M* NaCl for ligand **4**, 1.4 *M* ammonium sulfate, 0.1 *M* CHES pH 9.2, 0.2 *M* NaCl for ligand **5** and 1.2 *M* ammonium sulfate, 0.1 *M* CHES pH 9.2, 0.2 *M* NaCl for ligand **6**. Crystals of the COMT–SAH semi-holo complex grew using a reservoir solution consisting of 1.3 *M* sodium malonate pH 7.0, 0.1 *M* HEPES–NaOH pH 7.5. The SAM–tolcapone complex was crystallized from 1.5 *M* ammonium sulfate, 0.1 *M* bis-Tris–HCl pH 6.5, 0.1 *M* NaCl. Crystals were cryoprotected in paraffin oil (Hampton) and vitrified in liquid N_2 for data collection. Diffraction data were measured at 100 K and a wavelength of



Compound	*R*3′	*R*2′	IC_{50} (n*M*)
1	OH	OH	9
2	OH	OH	28 000
3	H	OH	40

(*a*)



Compound	R_{endo}	R_{exo}	*X*	IC_{50} (n*M*)
4	H	H	Me	1368
5	H	F	NHMe	11
6	H	Me	NHPr	25
7	OH	H	NH_2	35
8	OH	H	NHEt	43
9	OH	H	NHPr	26
10	OH	H	NHcPr	82
11	OH	H	Et	236
12	H	H	NHMe	1858
13	OH	H	NHMe	11

(*b*)

Figure 2
Importance of the ribose hydroxyl groups for inhibitors. (*a*) Substitution of hydroxyl groups in 4-nitrocatechol bisubstrate inhibitors results in large differences in IC_{50} values. In contrast to the 3′-hydroxyl group, the 2′-hydroxyl group cannot be replaced by a H atom without a considerable loss in affinity. (*b*) Structures and activities of novel bisubstrate inhibitors with 3′-deoxyribose (**4**), 3′-*exo* fluoro (**5**) and 3′-*exo* methyl (**6**) moieties. Similar IC_{50} values are obtained for all three inhibitors. Me, Et, Pr and cPr denote methyl, ethyl, propyl and cyclopropyl, respectively (Ellermann *et al.*, 2009, 2011).

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: WD5172). Services for accessing this material are described at the back of the journal.

Table 1
Data collection, analysis and structure refinement.

Values in parentheses are for the highest resolution shell.

Data set	3u81, SAH, semi-holo	3s68, SAM, tolcapone	3r6t, deoxyribose **4**	3nwb, F-ribose **5**	3nwe, Me-ribose **6**
Data collection					
Resolution range (Å)	39.8–1.13 (1.23–1.13)	43.9–1.85 (1.95–1.85)	28.0–1.20 (1.30–1.20)	31.2–1.30 (1.33–1.30)	45.2–1.50 (1.53–1.50)
100% criterion† (Å)	1.14	1.85	1.20	1.31	1.50
Oscillation range (°)	0.5	1.0	1.0	1.0	1.0
Exposure time (s)	0.5	0.5	0.5	0.5	0.5
No. of frames	360	180	130	180	180
Crystal-to-detector distance (mm)	165	180	100	120	150
Mosaicity (°)	0.20	0.17	0.18	0.32	0.28
Space group	$P2_12_12_1$	$P3_221$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 33.4, b = 61.3, c = 104.7	a = 50.6, c = 168.6	a = 50.9, b = 56.0, c = 77.9	a = 50.4, b = 55.4, c = 79.4	a = 50.5, b = 55.2, c = 78.8
Unique reflections	79382 (16949)	22456 (3298)	68614 (14112)	54289 (2035)	35942 (1392)
Multiplicity	6.1 (5.8)	10.0 (9.5)	4.8 (4.3)	6.6 (3.1)	6.6 (2.7)
Completeness (%)	97.4 (93.8)	99.9 (99.9)	97.2 (94.1)	97.9 (80.2)	98.9 (81.5)
$R_{\rm merge}$‡ (%)	8.5 (71)	13.0 (83)	6.7 (24)	6.7 (61)	7.8 (40)
Average $I/\sigma(I)$	8.9 (1.5)	12.6 (1.3)	13.3 (5.3)	13.4 (1.6)	11.0 (1.9)
Refinement					
Resolution range (Å)	39.8–1.13 (1.14–1.13)	42.4–1.85 (1.93–1.85)	28.0–1.2 (1.23–1.20)	26.1–1.30 (1.33–1.30)	45.2–1.50 (1.53–1.50)
$R_{\rm cryst}$§ (%)	13.4 (31.4)	17.0 (35.0)	12.1 (16.2)	14.5 (27.7)	17.6 (22.8)
$R_{\rm free}$§ (%)	16.0 (35.8)	22.0 (39.3)	14.2 (17.6)	17.6 (32.2)	19.6 (29.0)
No. of residues	218	213	218	216	213
No. of waters	346	211	272	206	175
R.m.s.d. bonds (Å)	0.021	0.007	0.012	0.019	0.021
R.m.s.d. angles (°)	1.19	1.01	1.51	1.89	1.84
Ramachandran plot¶ (%)					
Core	93.2	91.0	91.5	91.5	91.5
Allowed	5.8	8.5	8.0	8.0	8.0
Disallowed	1.0	0.5	0.5	0.5	0.5
Average B value†† (Å^2)	13.7 ± 5.4	27.7 ± 11.5	11.3 ± 6.4	19.0 ± 6.5	19.4 ± 4.8

† The 100% criterion was calculated using *SFTOOLS* (Winn *et al.*, 2011) and represents the resolution in Å of a 100% complete hypothetical data set with the same number of reflections as the measured data. ‡ $R_{\rm merge} = 100\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith measurement of reflection hkl and $\langle I(hkl)\rangle$ is the average value of the reflection intensity. § $R_{\rm cryst} = \sum_{hkl}||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl}|F_{\rm obs}|$, where $F_{\rm obs}$ and $F_{\rm calc}$ are the structure-factor amplitudes from the data and the model, respectively. $R_{\rm free}$ is $R_{\rm cryst}$ calculated using a 5% test set of structure factors. ¶ Calculated using *PROCHECK* (Laskowski *et al.*, 1993). †† Calculated using all atoms.

1.0 Å using a MAR 225 CCD detector on beamline X10SA at the Swiss Light Source. Data from a single crystal were integrated and scaled using *XDS* (Kabsch, 2010) and *SADABS* (Bruker-AXS), respectively. The structures were determined by molecular replacement with *Phaser* (Winn *et al.*, 2011) using an in-house COMT structure as the starting model. Models were built with *Coot* (Emsley *et al.*, 2010) and refined with *PHENIX* (Zwart *et al.*, 2008). Riding H atoms were generated once using *REDUCE* (Word *et al.*, 1999) and were allowed to change positions during refinement. The B values of non-H atoms were refined anisotropically for resolutions of <1.5 Å. The anisotropy of the SAM–tolcapone complex was modelled with nine TLS groups that were determined automatically by *PHENIX*. The structures of **4**, **6** and SAH were determined using the humanized rat COMT whereas the SAM/tolcapone and **5** complexes were determined using wild-type rat COMT. IC_{50} values do not differ appreciably between the wild-type and double mutant (2–5-fold higher for the mutant; data not shown). Data-collection and refinement statistics are summarized in Table 1. Coordinates and structure factors have been deposited in the Protein Data Bank [PDB codes 3u81 (SAH), 3s68 (SAM–tolcapone), 3r6t (**4**), 3nwb (**5**) and 3nwe (**6**)]. Modelling of ligands prior to synthesis and structural analysis was performed in *MOLOC* (Gerber & Müller, 1995).

3. Results and discussion

3.1. The SAM–tolcapone structure and the concept of bisubstrate inhibitors

As a reference state and to address any potential conformational changes associated with the bisubstrate inhibitors discussed below, the COMT–SAM–Mg^{2+}–tolcapone complex was chosen and its structure was determined to 1.85 Å resolution (Fig. 1*a*). Tolcapone is used as an adjunct agent in L-DOPA therapy against Parkinson's disease (Truong, 2009). The crystal structure of the COMT–SAM–Mg^{2+}–tolcapone complex shows an intact SAM co-substrate that entertains nine direct hydrogen bonds to protein atoms. The adenine base is hydrophobically encased by the side chains of Met91, His142 and Trp143. Notably, the backbone torsion angles of His142 are energetically disfavoured and lie in the disallowed region of the Ramachandran plot. In fact, this observation holds true for all complexes of COMT with ligands that occupy the adenine-binding site (Ellermann *et al.*, 2009, 2011; Lerner *et al.*, 2001). The strained conformation of His142 offers a possible explanation for the low affinity of SAM and related nucleotides, as some of their binding energy may be stored in the protein for facilitated catalysis. The carboxylate group of the Glu90 side chain binds ribose *via* hydrogen bonds to the 2′- and 3′-hydroxyl groups. The amino-acid part of SAM binds in a

hydrophilic pocket, in which the ammonium group is neutralized by Asp141 (Fig. 1*a*), which is also a ligand of the Mg^{2+} ion. The carboxylate group of SAM entertains two hydrogen bonds to backbone amide N atoms and its charge is partially neutralized by the positive end of an α-helix dipole (labelled 'α' in Fig. 1*a*). The SAM methyl group is poised for transfer to the hydroxyl group that is *ortho* to the nitro group of tolcapone (Me$\cdots$O distance of 2.5 Å; arrow in Fig. 1*a*). However, owing to the electron-withdrawing effect of the nitro group, the hydroxyl group in tolcapone is not nucleophilic enough and the SAM–tolcapone structure represents the Michaelis complex. The catechol moiety of tolcapone forms two dative bonds to Mg^{2+}, a motif that is retained in the bisubstrate inhibitors in which the methionine moiety is removed, and a two-atom linker connects C5′ of the ribose to an amide group on the catechol (Fig. 1*b*).

3.2. Different ribose conformations in SAM and bisubstrate inhibitors

The linker in bisubstrate inhibitors changes the ribose pucker compared with SAM, leading to a more pronounced 5′-*endo* conformation (Fig. 1*b*). The SAM ribose pucker is identical to that observed in a COMT–SAH semi-holo complex lacking both an Mg^{2+} ion and a catechol (Supplementary Fig. S1). While in the holo complex both the adenine-binding and the Mg^{2+}-binding sites are occupied, only the adenine-binding site is occupied in the semi-holo complex. Here, the side chains forming the Mg^{2+}-binding site are not oriented to allow octahedral coordination of the metal ion and consequently no catechol can bind. The distinct ribose pucker of SAM derivatives is independent of protein conformation and thus seems to be a salient feature of the nucleotide. Comparison of all bisubstrate-inhibitor structures determined to date (Ellermann *et al.*, 2009, 2011; Lerner *et al.*, 2001; this work) reveals the same pucker (but different from that of SAM), indicating that this conformation is a consequence of linking the ribose to the catechol while removing the methionine part of SAM. The distance between the ribose C5′ atoms in SAM and bisubstrate inhibitors is $\sim$1.8 Å, with a concomitant lateral shift of the ribose by $\sim$0.6 Å (arrow in Fig. 1*b*). In contrast, the side-chain position and conformation of Glu90 remain unaltered.

3.3. Molecular determinants for the importance of the 2′-hydroxyl group

As outlined above, the 3′-hydroxyl group can be omitted from nitrocatechol-based bisubstrate inhibitors, but the 2′-hydroxyl group is absolutely essential for high-affinity binding (Paulini *et al.*, 2006). Deoxy derivatives at the 2′- and 3′-positions display IC_{50} values that differ by a factor of 700, translating into a difference in free enthalpy of 16 kJ mol^{-1} at 298 K.

The importance of Glu90 for binding of the ribose has been noted previously (Lerner *et al.*, 2001). The question arises: what are the molecular determinants for the strong contribution of the 2′-hydroxyl group? In the SAM–tolcapone complex, the hydrogen-bond lengths between Glu90 and the ribose are 2.60 $\pm$ 0.06 Å for the hydroxyl groups, depending on the weighting scheme for refinement. Thus, bond lengths do not offer an explanation for the stronger hydrogen bond donated by the 2′-hydroxyl group. In six related higher resolution bisubstrate-inhibitor complexes described here and previously (Ellermann *et al.*, 2009, 2011), the hydrogen bond to the 2′-hydroxyl group varies in length between 2.49 and 2.63 Å [PDB entries 3hvh (1.30 Å resolution), 3hvi (1.20 Å resolution), 3hvj (1.79 Å resolution), 3hvk (1.30 Å resolution), 3nw9 (1.65 Å resolution) and 3oe4 (1.49 Å resolution)]. For the 3′-hydroxyl group, the extreme distances are 2.49 Å (PDB entry 3nw9) and 2.69 Å (PDB entry 3oe4). These distances may still be considered to be identical within coordinate error. Likewise, all ribose hydroxyl OH$\cdots$Glu90 angles are >150° and thus within the range of strong hydrogen bonds (130–180°; Desiraju & Steiner, 2006). Both H atoms of the ribose hydroxyl groups interact with the *syn* lone pairs of Glu90, which is the preferred orientation for OH$\cdots$carboxylate interactions in the Cambridge Structural Database (CSD; Bissantz *et al.*, 2010). However, a distinguishing feature between the OH$\cdots$Glu90 geometries is the angular deviation of the hydrogen bond from the acceptor plane π-system. Usually, in the absence of charges the hydrogen-bond donor should lie in the plane of the acceptor plane π-system. Angle deviations of above 25–30° of the carboxylate plane are supposed to result in a weaker hydrogen-bond interaction (Bissantz *et al.*, 2010). While for the 2′-hydroxyl group this angle is $\sim$5°, it amounts to 25° for the 3′-hydroxyl group, indicating a much weaker interaction.

An analysis of 146 small-molecule crystal structures of nucleosides in the CSD shows that the H atoms of the ribose hydroxyl groups are rarely both in the sterically strained *endo* conformation (3%; five examples: RFPNCX10, DAPJUA, HDTURD10, BRGUOS01 and BRINOS10) as is required for hydrogen bonding to the *syn* sp^2 oxygen lone electron pairs of Glu90 (Fig. 1*c*). (Here, the *endo* conformation describes the orientation of the hydroxyl O—H σ bond with respect to the cyclic hydrogen-bonding network towards Glu90. It does not refer to the exo/endo nomenclature describing the lower and upper hemispheres of saccharides.) The majority of nucleosides have at least one H atom in the *exo* conformation. Thus, the removal of one hydroxyl group could alleviate the steric and electrostatic strain of the double *endo* conformation. It does not explain the preference for the 2′-hydroxyl over the 3′-hydroxyl group, however.

Glu90 is buried beneath the ribose but does not have a counter-charge. This potentially destabilizing situation could be remedied by solvation with nearby hydrogen-bond donors such as protein backbone amide or ribose hydroxyl groups. For a greater stabilizing effect, the acidity of the hydroxyl groups should be high. The acidity of the 2′-hydroxyl group is increased relative to the 3′-hydroxyl group by inductive effects from the glycosidic bond and the ring O atom (Velikyan *et al.*, 2001). Thus, compared with the 3′-hydroxyl, the 2′-hydroxyl group is more prone to share a H atom with an acceptor and should form a stronger hydrogen bond to Glu90.

In contrast, the basicity of the Glu90 carboxylate O atoms should be high for a strong interaction with the ribose. The O atoms are solvated differently by the protein (Fig. 1*c*). The Glu90 side-chain O atom interacting with the 3′-hydroxyl group of the ribose is further solvated by two hydrogen bonds from backbone amide groups of Ala67 and Tyr68. The NH···O—Glu90 angles are ∼145° and ∼173° for positions 67 and 68, respectively, indicating strong interactions with the *anti* sp^2 lone electron pairs of Glu90. A similarly strong but single hydrogen bond exists between the backbone amide of Asn92 and the Glu90 carboxylate O atom contacting the 2′-hydroxyl group (Asn92—NH···O—Glu90 angle of ∼170°). Thus, COMT itself better solvates the Glu90 O atom contacting the 3′-hydroxyl than the other O atom, thereby increasing the basicity and hydrogen-bond acceptor strength of the latter. In the absence of a ribose-carrying ligand, Glu90 is likely to be solvated by water, but once a ligand is present, strong interaction of the 2′-hydroxyl group with Glu90 is essential. Together with the increased acidity of the 2′-hydroxyl group, the different solvation of the Glu90 carboxylate O atoms by COMT is in line with the higher energetic contribution of the 2′-hydroxyl compared with the 3′-hydroxyl group (Fig. 2).

In summary, four structural determinants that influence the hydrogen-bond network between ribose and Glu90 were identified: (i) a stronger deviation of the hydrogen-bond angle from the carboxylate plane for the 3′-hydroxyl group compared with the 2′-hydroxyl group, (ii) a statistically disfavoured *endo*-conformation of the ribose hydroxyl groups, (iii) a higher polarity of the 2′-hydroxyl group compared with the 3′-hydroxyl group owing to inductive effects and (iv) the different solvation of the Glu90 O atoms by the protein in the absence of a counter-charge.

3.4. Variations of the ribose motif in novel COMT bisubstrate inhibitors

With the importance of the 2′-hydroxyl group established, bisubstrate inhibitors were modelled that retained the 2′-hydroxyl group but varied both the nature of the 3′-substituent and the chirality of its stereogenic centre. Conceptually, the simplest approach was to delete the 3′-hydroxyl group (*R* = H), as realised in compounds **4** and **12**. While a significant amount of biochemical data is available on the binding of 3′-deoxy-SAM analogues to COMT (Borchardt & Wu, 1975; Borchardt *et al.*, 1976; Paulini *et al.*, 2006), structural information on this class of compounds has been lacking. Modelling also suggested that the active site of COMT could accommodate small 3′-substituents *trans* to the 2′-hydroxyl group of the ribose moiety (*endo* with respect to C5′). The 3′-fluoro (**5**) and 3′-methyl (**6**) derivatives with inverted chirality at C3′ were synthesized and nanomolar IC_{50} values were determined (Fig. 2). The rationale for the fluorine substituent in **5** was to enhance affinity of the compound by its strong σ-inductive effect (Müller *et al.*, 2007), lowering the pK_a value of the 2′-hydroxyl group and strengthening the hydrogen bond to the side chain of Glu90. Despite its electron-donating effect, which should weaken this hydrogen bond, the 3′-methyl group in **6** was expected to engage in stabilizing van der Waals interactions with nearby hydrophobic side chains. Variation of the adenine base at the N6 position either by alkylation or substitution with a methyl group (**8**–**12**) did not strongly affect the IC_{50} values of the ligands (Ellermann *et al.*, 2009, 2011). Also, the nitro group on the catechol part was replaced by 4-fluorophenyl in all bisubstrate inhibitors discussed here, which increases COMT affinity but does not significantly alter the complex structure (Paulini *et al.*, 2004, 2006; Ellermann *et al.*, 2009, 2011).

3.5. Crystal structures of bisubstrate inhibitors

The first crystal structure of COMT in complex with a 3′-deoxyribose-containing ligand (**4**) repeats the salient features of earlier bisubstrate inhibitors containing an intact ribose (Fig. 3*a*). The N1 atom of the methylpurine moiety forms a strong hydrogen bond (2.9 Å) to the backbone NH of Ser119 and is encased by van der Waals interactions with Ile91 (3.5–3.9 Å), His142 (∼3.8 Å) and Trp143 (down to 3.3 Å). Of note, the edge-to-face interaction with Trp143 appears to be particularly short, indicating a strong energetic contribution. Substitutions at the 6-position will displace the Trp143 indole side chain to cover the ribose instead of the purine (see below). The stabilizing effect of the purine–His142 interaction might be offset by the fact that His142 locates to a disallowed region of the Ramachandran diagram, possibly explaining the overall weak affinity of SAM and SAH. A similarly destabilizing interaction is present in the SAM cocrystal structure of the 2′-*O*-methylcytidine-generating tRNA methyltransferase from *Pyrococcus horikoshii* (Kuratani *et al.*, 2008; PDB entry 2yy8). For COMT, this energetic penalty has to be paid by all inhibitors of the adenine-binding site. The catechol moiety completes the octahedral Mg^{2+} coordination and the 4-fluorophenyl substituent fills a hydrophobic cleft lined by Trp38, Val173, Pro174 and Leu198. Importantly, the 3′-desoxyribose moiety retains the canonical C3′-*exo* (C2′-*endo* or 'South') conformation (Altona & Sundaralingam, 1972) that is observed in all COMT crystal structures in complex with ribose derivatives (Fig. 3). Removal of the 3-hydroxyl group leaves a small cavity that is partially filled by a 0.3 Å shift of the C3′-atom towards Glu90.

The introduction of a 3′-fluoro substituent in compound **5** leads to subtle conformational changes around the ribose moiety when bound to COMT and Mg^{2+} (Fig. 3*b*). The ribose ring retains the 'South' conformation, indicating that inversion of configuration, similar to removal of the 3′-substituent, does not change the pucker in the ternary complex. The vicinal F atom is likely to reduce the pK_a value of the 2′-hydroxyl group, which forms a short 2.5 Å hydrogen bond to the carboxylate of Glu90. At the same time, the electron-withdrawing effect of the fluorine on the C3′ H atom may strengthen a weak CH···Glu90 hydrogen bond ($d_{CH\cdots O}$ = 3.4 Å). The 3′-fluorine substituent also favorably contacts the edges of the Tyr95 ($d_{F\cdots C\varepsilon}$ = 3.3 Å) and Tyr68 ($d_{F\cdots C\delta}$ = 3.9 Å) planes and the terminal methyl group of Met40 ($d_{F\cdots C\delta}$ = 3.3 Å) and is thus

located in a rather hydrophobic environment. Similar interactions are not possible in the 3′-deoxy compound **4**. Taken together, these minor adjustments arising from the introduction of a single atom in **5** may overcompensate for the destabilizing removal of the 3′-hydroxyl group in **4**, leading to a high-affinity compound with an IC_{50} value of 11 n*M*.

The binding mode and hydrophobic environment of inhibitor **6** with a 3′-methyl group of inverted configuration relative to ribose are similar to those of the 3′-fluoro compound **5** (Fig. 3*c*). The hydrogen bond of the 2′-hydroxyl group to Glu90 is slightly elongated by ∼0.1 Å compared with **5**, which might be explained by the lack of activation of the 2′-hydroxyl group by the 3′-methyl substituent. Indeed, the positive σ effect exerted by the methyl group should lead to a weaker hydrogen bond. In summary, the minute conformational differences between **5** and **6** in the co-crystal structures are reflected by their IC_{50} values (Fig. 2*b*), which differ by only a factor of two.

3.6. Comparison of the bisubstrate inhibitor binding modes

Superposition of all ribose-modified bisubstrate inhibitor complexes and comparison with a nonmodified bisubstrate inhibitor structure (PDB entry 3nw9; Ellermann *et al.*, 2011) reveals no large conformational changes in the protein apart from the side chains of Met40 in the case of **4** (see above) and Trp143. In compounds that are substituted at the adenine N6 position, the indole side chain of Trp143 is flipped away from the adenine to cover the ribose. Most conformational differences, however, are centred at the ribose, which slightly changes its pucker to minimize the cavity left by the removal of the 3-hydroxyl group. In contrast, the Glu90 side chain that binds to the ribose hydroxyl groups is fixed so strongly by interactions with protein backbone amide groups that its carboxylate group may only rotate by 7° ($C^{\beta}-C^{\gamma}$ torsion angles of −120° and −113° for **4** and **5**, respectively) to accommodate the changes in ribose conformation (Fig. 3*d*). The ribose conformation is neither influenced by the lack of a 3′-substituent nor the introduction of small substituents at C3′ with inverted chirality. Rather, it is the 2′-hydroxyl group that dictates the ribose conformation by forming a strong hydrogen bond to Glu90. The strength of this hydrogen bond may be modulated by the electronegativity of the C3′ substituent. An additional hydrogen bond exists between the 2′-hydroxyl group and a water molecule (Fig. 3*a*) that is conserved in all COMT–bisubstrate structures, thus saturating the hydrogen-bonding potential of the ribose.

4. Conclusions

Three high-resolution cocrystal structures of ribose-modified bisubstrate inhibitors in complex with COMT and Mg^{2+} were determined that include the first 3′-deoxyribose-based ligands. The 3′-*endo* fluorine and methyl substituents do not sterically interfere with the protein, but display favourable effects with respect to binding affinity. Both the fluorine and the methyl substituents point into a similar hydrophobic pocket. While

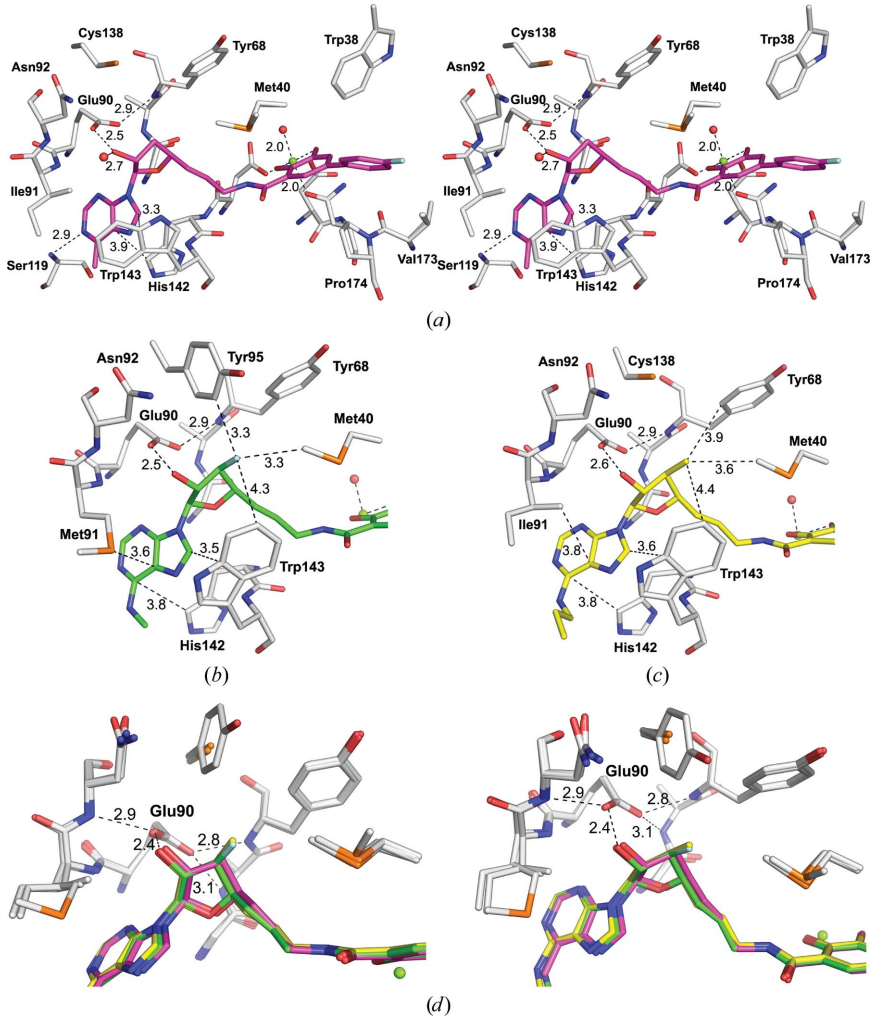


Figure 3
Crystal structures of ribose-modified compounds bound to COMT and Mg^{2+}. (*a*) Stereoview of the 3′-deoxyribose bisubstrate inhibitor **4**. A key hydrogen bond is formed between the nucleobase and the backbone NH of Ser119 ($d_{N1\cdots N}$ = 2.9 Å). The nucleobase also engages in electrostatic edge-to-face interactions with the strained imidazole ring of His142 (d = 3.9 Å) and makes van der Waals contacts to Ile91. For a detailed discussion of the adenine site, see Ellermann *et al.* (2009, 2011). (*b*) Close-up of the nucleosidic interactions of the 3′-fluoro compound **5**. A sulfur–π interaction with Met91 ($d_{C5\cdots S}$ = 3.6 Å) replaces the otherwise very similar environment of the adenine. The F atom (cyan) contacts Tyr95, Met40 and the indole side chain of Trp143, which has flipped compared with (*a*). (*c*) The same view as in (*b*) for the C3′-methyl compound **6**. The methyl group is located in a very similar environment to the fluorine. (*d*) Two representations of a superposition of the structures in (*a*)–(*c*) focusing on the ribose-binding site.

such an environment is clearly favourable for alkyl groups, it can also constitute a favourable location for a fluorine, provided that the halogen is oriented towards CH, CH_2 and CH_3 groups of the protein side chains and that it does not point into aromatic π-systems (Müller *et al.*, 2007). Both of these prerequisites are met in the complex of **5** with COMT. The interaction of fluorine with Tyr95 is side on, not face on, and there is a favourable interaction with the terminal methyl group of Met40. Quite high resolutions of the structures were needed in order to allow the detection of the subtle shifts of the ligand and rotations of the Glu90 side chain that optimize hydrogen bonding to the 2′-hydroxyl group. In summary, this study shows that slight adjustments of the environment of small substituents, aided by improved ligand solvation by protein side chains, can provide a structural explanation for high affinity despite the disruption of an apparently strong hydrogen bond. Such strategies may also help to remove potentially labile hydroxyl groups during lead optimization of other ligand classes.

We thank the staff at beamline PX-II of the Swiss Light Source for support during data collection, Valérie Goetschy, Doris Roth and Pia Warga (Roche) for comments on the inhibition measurements and D. Schlatter (Roche) for purification of COMT. Research at ETH Zürich was supported by F. Hoffmann-La Roche AG, Basel.

References

Altona, C. & Sundaralingam, M. (1972). *J. Am. Chem. Soc.* **94**, 8205–8212.

Bissantz, C., Kuhn, B. & Stahl, M. (2010). *J. Med. Chem.* **53**, 5061–5084.

Bonifácio, M. J., Archer, M., Rodrigues, M. L., Matias, P. M., Learmonth, D. A., Carrondo, M. A. & Soares-Da-Silva, P. (2002). *Mol. Pharmacol.* **62**, 795–805.

Borchardt, R. T., Shiong, Y., Huber, J. A. & Wycpalek, A. F. (1976). *J. Med. Chem.* **19**, 1104–1110.

Borchardt, R. T. & Wu, Y. S. (1975). *J. Med. Chem.* **18**, 300–304.

Davis, A. P. & Wareham, R. S. (1999). *Angew. Chem. Int. Ed.* **38**, 2978–2996.

Desiraju, G. R. & Steiner, T. (2006). *The Weak Hydrogen Bond in Structural Chemistry and Biology*, p. 13. Oxford University Press.

Dwek, R. A. (1996). *Chem. Rev.* **96**, 683–720.

Ellermann, M., Jakob-Roetne, R., Lerner, C., Borroni, E., Schlatter, D., Roth, D., Ehler, A., Rudolph, M. G. & Diederich, F. (2009). *Angew. Chem. Int. Ed.* **48**, 9092–9096.

Ellermann, M., Paulini, R., Jakob-Roetne, R., Lerner, C., Borroni, E., Roth, D., Ehler, A., Schweizer, W. B., Schlatter, D., Rudolph, M. G. & Diederich, F. (2011). *Chemistry*, **17**, 6369–6381.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Ferrand, Y., Klein, E., Barwell, N. P., Crump, M. P., Jiménez-Barbero, J., Vicent, C., Boons, G.-J., Ingale, S. & Davis, A. P. (2009). *Angew. Chem. Int. Ed.* **48**, 1775–1779.

Gerber, P. R. & Müller, K. (1995). *J. Comput. Aided Mol. Des.* **9**, 251–268.

Guldberg, H. C. & Marsden, C. A. (1975). *Pharmacol. Rev.* **27**, 135–206.

Holm, K. J. & Spencer, C. M. (1999). *Drugs*, **58**, 159–177.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Katzenschlager, R. & Lees, A. J. (2002). *J. Neurol.* **249**, 19–24.

Keating, G. M. & Lyseng-Williamson, K. A. (2005). *CNS Drugs*, **19**, 165–184.

Kuratani, M., Bessho, Y., Nishimoto, M., Grosjean, H. & Yokoyama, S. (2008). *J. Mol. Biol.* **375**, 1064–1075.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lee, J.-C., Chang, S.-W., Liao, C.-C., Chi, F.-C., Chen, C.-S., Wen, Y.-S., Wang, C.-C., Kulkarni, S. S., Puranik, R., Liu, Y.-H. & Hung, S.-C. (2004). *Chemistry*, **10**, 399–415.

Lerner, C., Masjost, B., Ruf, A., Gramlich, V., Jakob-Roetne, R., Zürcher, G., Borroni, E. & Diederich, F. (2003). *Org. Biomol. Chem.* **1**, 42–49.

Lerner, C., Ruf, A., Gramlich, V., Masjost, B., Zurcher, G., Jakob-Roetne, R., Borroni, E. & Diederich, F. (2001). *Angew. Chem. Int. Ed.* **40**, 4040–4042.

Lis, H. & Sharon, N. (1998). *Chem. Rev.* **98**, 637–674.

Männistö, P. T. & Kaakkola, S. (1999). *Pharmacol. Rev.* **51**, 593–628.

Männistö, P. T., Ulmanen, I., Lundström, K., Taskinen, J., Tenhunen, J., Tilgmann, C. & Kaakkola, S. (1992). *Prog. Drug Res.* **39**, 291–350.

Masjost, B., Ballmer, P., Borroni, E., Zürcher, G., Winkler, F. K., Jakob-Roetne, R. & Diederich, F. (2000). *Chemistry*, **6**, 971–982.

McCarthy, J. J. (2001). *Trends Biotechnol.* **19**, 283–284.

Mercuri, N. B. & Bernardi, G. (2005). *Trends Pharmacol. Sci.* **26**, 341–344.

Mort, C. J., Migaud, M. E., Galione, A. & Potter, B. V. (2004). *Bioorg. Med. Chem.* **12**, 475–487.

Müller, K., Faeh, C. & Diederich, F. (2007). *Science*, **317**, 1881–1886.

Palma, P. N., Rodrigues, M. L., Archer, M., Bonifácio, M. J., Loureiro, A. I., Learmonth, D. A., Carrondo, M. A. & Soares-da-Silva, P. (2006). *Mol. Pharmacol.* **70**, 143–153.

Paulini, R., Lerner, C., Jakob-Roetne, R., Zürcher, G., Borroni, E. & Diederich, F. (2004). *Chembiochem*, **5**, 1270–1274.

Paulini, R., Trindler, C., Lerner, C., Brändli, L., Schweizer, W. B., Jakob-Roetne, R., Zürcher, G., Borroni, E. & Diederich, F. (2006). *ChemMedChem*, **1**, 340–357.

Posner, G. H. & Haines, S. R. (1985). *Tetrahedron Lett.* **26**, 5–8.

Rutherford, K., Le Trong, I., Stenkamp, R. E. & Parson, W. W. (2008). *J. Mol. Biol.* **380**, 120–130.

Truong, D. D. (2009). *Clin. Interv. Aging*, **4**, 109–113.

Tsuji, E., Okazaki, K., Isaji, M. & Takeda, K. (2009). *J. Struct. Biol.* **165**, 133–139.

Tsuji, E., Okazaki, K. & Takeda, K. (2009). *Biochem. Biophys. Res. Commun.* **378**, 494–497.

Velikyan, I., Acharya, S., Trifonova, A., Földesi, A. & Chattopadhyaya, J. (2001). *J. Am. Chem. Soc.* **123**, 2893–2894.

Vidgren, J., Svensson, L. A. & Liljas, A. (1994). *Nature (London)*, **368**, 354–358.

Vorbrüggen, H. & Ruh-Polenz, R. (2001). *Handbook of Nucleoside Synthesis*. New York: Wiley.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Word, J. M., Lovell, S. C., Richardson, J. S. & Richardson, D. C. (1999). *J. Mol. Biol.* **285**, 1735–1747.

Yakambram, P., Puranik, V. G. & Gurjar, M. K. (2006). *Tetrahedron Lett.* **47**, 3781–3783.

Zürcher, G. & Da Prada, M. (1982). *J. Neurochem.* **38**, 191–195.

Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C. & Adams, P. D. (2008). *Methods Mol. Biol.* **426**, 419–435.