

Structure of thymidylate kinase reveals the cause behind the limiting step in AZT activation

Structural comparison of thymidylate kinase complexed with either dTMP or with AZTMP suggests that the low phosphorylation rate of AZTMP is due to an induced P-loop movement.

Thymidylate kinase (TmkK) is an essential enzyme for cell replication. TmkK (E.C. 2.7.4.9; ATP:dTMP phosphotransferase) phosphorylates thymidine mono-phosphate (dTMP) to the corresponding diphosphate (dTDP) in the presence of ATP and magnesium. dTDP is then further activated by the non-base-specific nucleoside diphosphate kinase to thymidine triphosphate (dTTP), which is the substrate for DNA polymerases. The HIV prodrug 3'-azido-3'-deoxythymidine (AZT) is converted by cellular enzymes of the salvage pathway to its active form, azidothymidine triphosphate (AZTTP), in an analogous fashion to thymidine activation. The most important antiviral effect of AZT is a consequence of efficient AZTMP incorporation into a growing DNA chain by HIV reverse transcriptase (RT) but to a much lesser extent by human DNA polymerases, resulting in viral DNA chain termination in HIV infected cells^{1,2}. A prerequisite for AZT to be a potent drug is that all phosphorylation steps to the triphosphate are catalyzed efficiently by the cellular enzymes, TmkK being one of them. However, azidothymidine monophosphate (AZTMP), the substrate of TmkK, accumulates in millimolar concentrations^{3,4} in cells treated with AZT, implicating TmkK to be the rate limiting kinase in AZT activation^{5,6}.

To understand the reasons behind the poor activation of AZTMP by TmkK, and with the aim of finding solutions to this bottleneck, we have set out to determine the previously unknown three dimensional structure of TmkK. We report the fully refined structures of the binary complexes TmkK-dTMP and TmkK-AZTMP, and deduce a probable structural cause behind the 200-fold reduction in the AZTMP phosphorylation rate⁷.

Overall structure

Thymidylate kinase from *Saccharomyces cerevisiae* (*cdc8* gene) has 216 amino acid residues and a relative molecular mass of 25,000 (M_r 25K). It has 44% amino acid identity with the human enzyme, suggest-

ing a very similar three dimensional structure. TmkK is a globular protein built up from eight α -helices surrounding a five-stranded parallel β -sheet (Fig. 1), a very similar topology to that of adenylate kinase, although there is no apparent amino acid sequence similarity. The most striking difference is that adenylate kinase is a monomer, while TmkK is a dimer (Fig. 1a) as is the herpes simplex virus thymidine kinase^{8,9}. A detailed structural comparison with the herpes simplex virus thymidine kinase, which has also a thymidylate kinase activity, is in progress since the coordinates became available only recently (accession code 1KIN).

Nucleoside monophosphate (NMP) kinases undergo a number of conformational changes depending upon the state of nucleotide binding¹⁰, from an open unoccupied enzyme, to a partially closed conformation induced by NMP or ATP binding, and finally to a fully closed conformation when both substrates are present. The structures that we present here correspond to the partially-closed conformation, with either dTMP or AZTMP bound. It is only in the fully-closed conformation that a so called 'LID' closes over the phosphate donor, and in keeping with this, the LID region (residues 135–144) is disordered in our partially closed complexes.

The nucleotide's base binding site is located in a groove formed by Asp 93, Arg 94, and Tyr 95. These residues (the latter two are the only ones occurring in disallowed regions of the Ramachandran plot) make up a conserved signature sequence in all thymidylate kinases. The residues form the connection between β -strand 3 and α -helix 5 (see Fig. 1b) in what seems to be a unique fold; the helix continues in the same direction as the preceding strand, with the DRY sequence forming a kink of nearly 90°. Asp 93 is involved in ligating the essential magnesium ion (preliminary data from the TP₅A-complex structure; TP₅A is a bisubstrate analog of ATP and dTMP). The NE of Arg 94 coordinates the phosphate of dTMP; the importance of the fol-

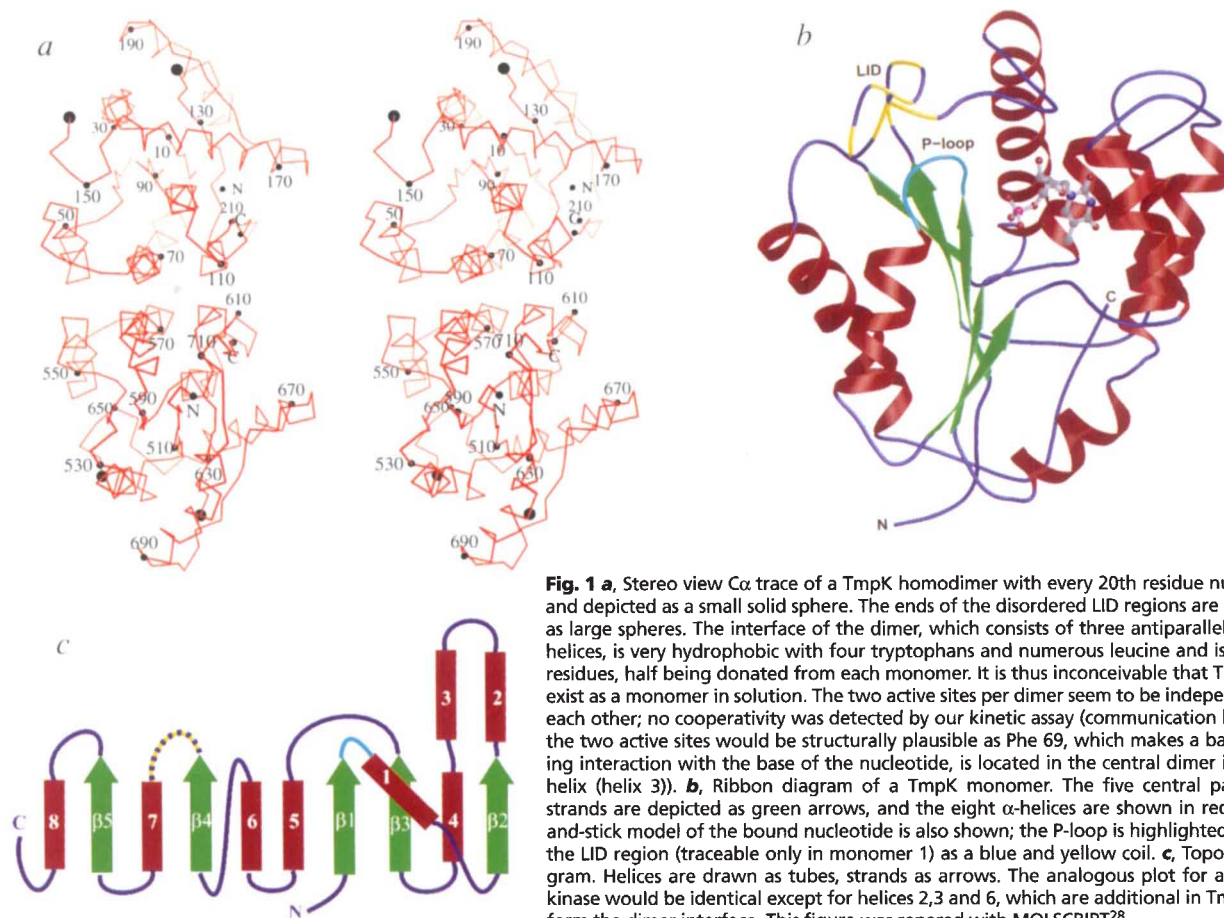
lowing tyrosine, which is on the other side of the kink, is unclear.

Substrate specificity

The basis of the specificity of TmkK for dTMP is obvious upon examination of the interactions between substrate and enzyme (Fig. 2b,c). The cavity in which the base binds is too small to accommodate purine bases, thus reducing the specificity problem to the pyrimidines. O4 of the thymidine base makes a hydrogen bond with a conserved arginine (Arg 73), thus favoring thymidine or uracil over cytosine. Perhaps more important in the discrimination against cytosine bases is the 3.1 Å distance between O4 to the C α of the conserved Gly 99, which would cause a steric clash with the amino group of cytosine at the C4 position. Further interactions of the thymidine base are hydrogen bonds between N3 and O2 and water molecules, where the water interacting with O2 also interacts with the hydroxyl group of the conserved Tyr 151. The methyl group at the C5 position, however, which differentiates the base between uridine and thymidine, does not make any close interactions with the enzyme. Thus, the discrimination against UMP does not appear to be achieved at the level of the base, but rather at the sugar by the conserved Tyr 102, which is situated 3.7 Å from C2 of the ribose ring; a 2'-hydroxyl group, present in UMP, would be located too close. This implies that CMP, dCMP, and UMP should not be accepted as substrates, in contrast to dUMP. We confirmed this prediction by activity measurements (data not shown).

The 3'-hydroxyl of the ribose makes a bidentate 3 Å interaction with the carboxyl group of Asp 14. The interaction of the P-loop Asp 14 is almost unique to thymidylate kinases as most P-loop containing kinases do not make any direct P-loop sugar interactions with the monophosphate substrate. In most nucleotide binding proteins with the P-loop motif (GX₁X₂X₃X₄GKS/T), the two amino acids at positions X₂ and X₃ are either glycine or

correspondence



proline¹¹, however, in TmpK they are Asp 14 and Arg 15. While Asp 14 is observed to interact with the 3'-hydroxyl group of the deoxyribose, no clear electron density is visible for the side chain of the succeeding arginine in either the dTMP or AZTMP-enzyme complex structures.

dTMP and AZTMP binding

Where one could have surmised a movement of the nucleotide as a result of the substitution of the 3'-hydroxyl of dTMP by the 1.5 Å longer azido group (-N₃) of AZTMP, no displacement is observed. The location of dTMP and AZTMP is almost identical, with a nearly perfect overlay of all common atoms. This crystallographic result is in agreement with our kinetic data which show a very similar K_M value for dTMP and AZTMP⁷. The enzyme-AZTMP interactions are identical to those made with dTMP, including those with the 3' substituent. The close proximity of the azido group and the side chain of Asp 14 is surprising, since the carboxy group is expected to be deprotonated, so that this would form an unfavourable interaction

with the third nitrogen of the azido group, which can be drawn as bearing a formal negative charge. However, this constellation could result in a significant change in the pK_a of the carboxy group, due to destabilization of the deprotonated form of the side chain, so that the interaction may be a hydrogen bonding one between the protonated form of the aspartate group and the azido group. The identical TmpK-AZTMP interactions are made possible by a rigid body movement of the entire P-loop by about 0.5 Å away from the azido group (Fig. 2c). Even though the movement of the P-loop is only slightly larger than our coordinate error (0.3 Å), it can be regarded as a genuine movement, since it encompasses numerous atoms. While this observation must be evaluated with care since it could be a result of the different data collection temperatures used, it is clear that such a movement is necessary for proper positioning of Asp 14 (situated at the tip of the P-loop), which would otherwise be too close (2.3 Å) to the azido group.

The movement of Asp 14 results in the concomitant move of its neighboring

residues, most importantly Arg 15. Preliminary results from our TP₅A complex reveal that Arg 15 plays an important part in ATP binding (main chain) and presumably catalysis as the side chain folds back to interact with the γ -phosphate group of ATP (data not shown), the phosphoryl group to be transferred. The assignment of a catalytic role for Arg 15 is corroborated by the structural comparison of TmpK to adenylate kinase (PDB code 2aky), in which Arg 132, shown by mutational studies to have a catalytic role^{12,13}, occupies a similar position in the three dimensional structure.

We attribute the 200-fold reduced phosphorylation rate of AZTMP in comparison to dTMP to this P-loop movement. As it is the P-loop which binds the phosphoryl donor, its position determines the relative orientation of the donor (ATP) and acceptor (dTMP or AZTMP), and the relative position of Arg 15. The 0.5 Å shift of the P-loop does not affect binding, as reflected in the unchanged K_M values, but rather increases the activation energy for the phosphoryl transfer step, thus lowering the rate.

Table 1 Data collection, phasing, and refinement statistics

	TMP1	TMP2	TMP UO ₂ Ac ₂	AZTMP
Data collection statistics				
Temperature (K)	100	277	277	277
Resolution (Å)	2.1	2.5	3.2	2.4
Observed reflections	51803	45824	15654	43762
Unique reflections	25398	14063	6633	17930
Completeness (% , overall/last shell)	88.5/70.1	80.6/48.5	81.1/75.2	93.3/89.4
R_{sym}^1 (% , overall/last shell)	6.5/28.1	5.0/11.9	6.1/9.6	8.3/26.8
Space group	P2 ₁	P2 ₁	P2 ₁	P2 ₁
Unit-cell				
a = (Å)	36.5	37.5	37.4	37.4
b = (Å)	141.2	144.2	144.2	144.2
c = (Å)	48.8	49.8	49.7	49.7
β = (°)	109.4	110.7	110.9	110.9
molecules / asymmetric unit	2	2	2	2
Phasing statistics				
Heavy atoms sites			2	
R_{cullis}^2			0.72	
Phasing power ³			1.2	
Refinement statistics				
$R_{\text{factor}}^4/R_{\text{free}}$ (%)	18.0/27.2			18.2/25.3
Resolution range (Å)	10–2.1			10.0–2.4
Reflections with $F > 0\sigma$ (working/test)	22386/1134			16820/850
Number of protein/water/nucleotide atoms	3301/282/41			3306/166/46
R.m.s. deviations				
bond lengths (Å)	0.011			0.008
bond angles (Å)	1.52			1.84
dihedral angles (°)	23.1			21.9
improper angles (°)	1.68			1.28
Average B-factor (Å ²)				
main chain	22			26
side chain	24			28
waters	31			33
nucleotides	17			19

$$^1 R_{\text{sym}} = \sum |I_j - \langle I_j \rangle| / \sum \langle I_j \rangle$$

$$^2 R_{\text{cullis}} = \sum |F_{\text{PH}} \pm F_{\text{PL}} - F_{\text{H}}| / \sum |F_{\text{PH}} - F_{\text{PL}}|$$

³ Phasing power = $F_{\text{H}} / E_{\text{RMS}} F_{\text{P}}$, F_{PH} , and F_{H} are the protein, derivative, and heavy atom structure factors respectively, and E_{RMS} is the residual lack of closure.

⁴ $R_{\text{factor}} = \sum (|F_{\text{obs}}| - |F_{\text{calc}}|) / \sum |F_{\text{obs}}|$ (5% randomly omitted reflections were used for R_{free})

Implications for drug development

There is suggestive evidence that the poor activation of AZT to AZTTP renders this chain terminator only partially successful in the treatment of HIV infected individuals⁷. Moreover, the low phosphorylation rate of AZT by TmpK does not only result in insufficient concentrations of the active AZTTP, but also in the accumulation of the toxic AZTMP^{14–16}. To correct this situation, there are two choices. One is to modify AZT so that it becomes a better substrate for TmpK (while keeping all the positive properties of AZT), or to modify TmpK so that it phosphorylates AZTMP more effectively. The first approach translates into finding a new chain terminating nucleoside with better properties than AZT. A promising direction would be the incorporation of non-ribose based nucleosides, such as those found in the highly successful drugs acyclovir or ganciclovir aimed against the her-

pes simplex virus (HSV). The 3'-azido analog of ganciclovir, but with the thymine base, may still maintain the antiviral effect of AZT, and with its added flexibility may bind without shifting the position of the P-loop. A second general approach is the introduction of a modified TmpK, or a viral kinase such as the HSV thymidine kinase, into infected cells, a strategy that is becoming more realistic with progress in the field of gene delivery¹⁷. The modified TmpK must be able to phosphorylate AZTMP at such a rate that it would not accumulate in the cell; this would allow a lower dose of AZT to achieve an increased antiviral effect with diminished cellular toxicity. Our TmpK structures form the basis for a structure based design of such TmpK mutants.

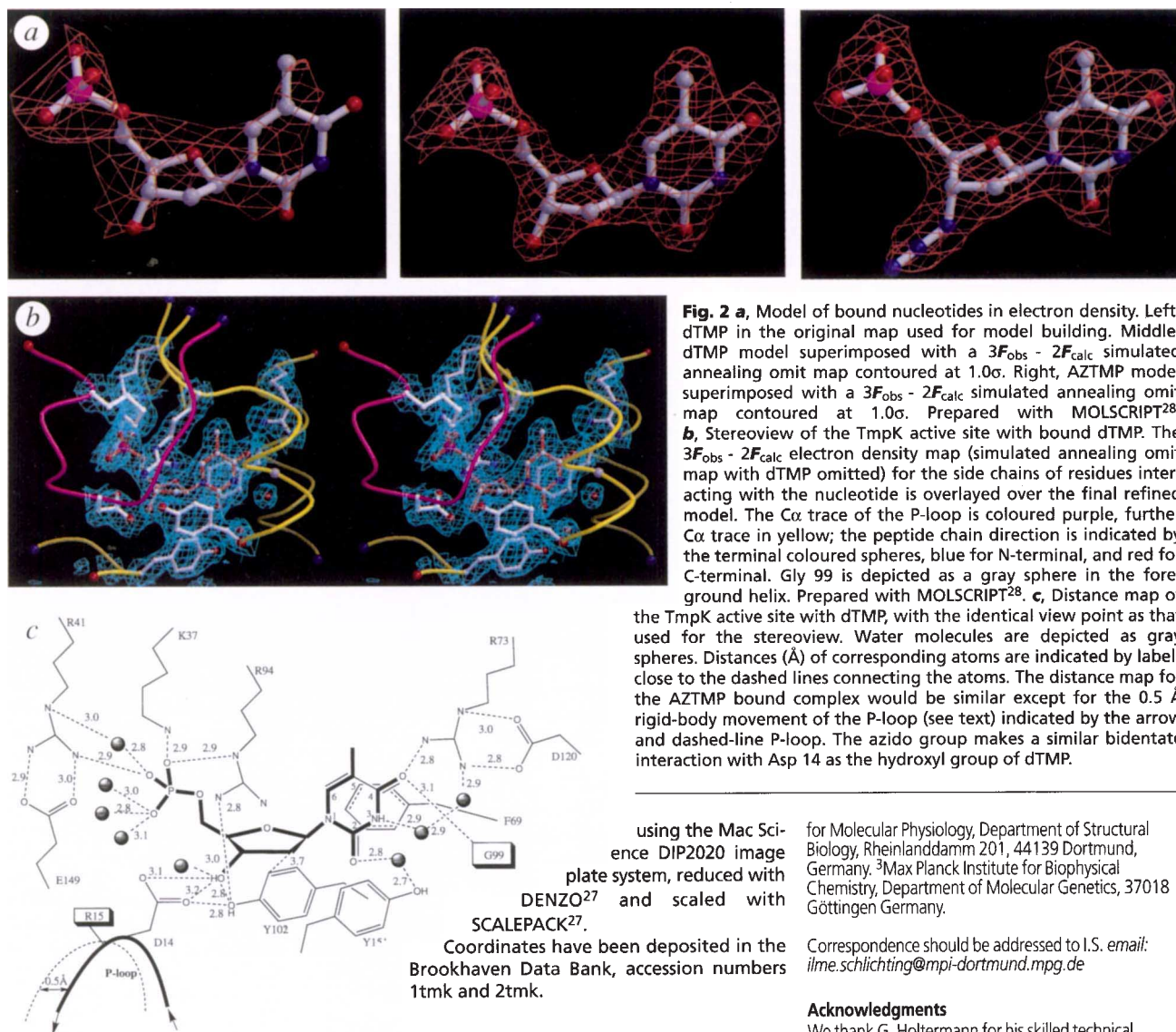
TmpK is no less important as an anti-cancer target. Numerous successful chemotherapeutic agents, such as methotrexate and 5-fluorouracil, target the

cellular nucleotide pools. Both of the above drugs target the thymidine *de-novo* phosphorylation pathway; methotrexate inhibits thymidylate synthase indirectly by inhibiting dihydrofolate reductase, while 5-fluorouracil does so directly by binding to thymidylate synthase¹⁸. In principle, TmpK is a better target for inhibition, since it functions at the junction of the salvage and *de novo* pathways of thymidine activation, the step following thymidylate synthase. Moreover, since TmpK expression is cell cycle controlled¹⁹, some measure of specificity could be expected at this level, discriminating between resting (normal) and dividing (cancerous) cells. The structures presented here provide a basis for developing tightly binding specific inhibitors for use as chemotherapeutic agents.

Methods

S. cerevisiae thymidylate kinase was expressed and purified as described⁷. dTMP–TmpK complex crystals were grown by equilibrating a 1:1 mixture of 10 mg ml⁻¹ TmpK, 2 mM dTDP, with 19% PEG 4000, 25 mM MgAc, and 100 mM HEPES buffer pH 7.5, using the hanging drop method. Although the drops were set up with dTDP, no electron density for the β-phosphate was observed in the refined structure, consistent with the HPLC analysis of dissolved crystals that showed the presence of dTMP. The dTMP complex crystallizes in space group P2₁ with a dimer in the asymmetric unit. Of more than 50 heavy atom reagents screened only uranyl acetate resulted in an isomorphous derivative. The positions and occupancies of the two uranils bound per asymmetric unit were refined using MLPHARE²⁰, and phases were calculated. The resulting single isomorphous replacement map clearly showed the molecular boundary of each monomer. A mask was generated around one monomer (using MAMA²¹), and the matrix relating the two monomers was refined (with RAVE²¹). The refined matrix and the initial phases calculated by MLPHARE were optimized with DM²² by solvent flattening, histogram matching, and most importantly, non-crystallographic symmetry averaging. The phases were extended from 3.2 Å to 2.8 Å, and the resulting averaged electron density map was used for model building with O²³. Poly-alanine α-helices and β-strands were fit as rigid bodies into the electron density. The partially traced poly-alanine model was refined with X-PLOR²⁴ and used for phase combination with SIGMAA²⁵. This process was reiterated. The substrate density (see Fig. 2a) was modelled after most of the protein residues were visible. The data set labelled TMP2 was used for the structure determination and initial stages of refinement until the higher resolution data set TMP1 was collected, with which refinement was continued (Table 1). All dTMP–TmpK complex data sets were collected using a Siemens/Hi-STAR multi-wire area detector mounted on a Mac Science rotating anode operating at 45 kV, 100 mA, and were

correspondence



using the Mac Science DIP2020 image plate system, reduced with DENZO²⁷ and scaled with SCALEPACK²⁷.

Coordinates have been deposited in the Brookhaven Data Bank, accession numbers 1tmk and 2tmk.

Arnon Lavie¹, Ingrid R. Vetter²,
Manfred Konrad³, Roger S. Goody¹,
Jochen Reinstein¹ and Ilme
Schlichting¹

¹Max Planck Institute for Molecular Physiology,
Department of Physical Biochemistry, Rheinlanddamm
201, 44139 Dortmund, Germany. ²Max Planck Institute

for Molecular Physiology, Department of Structural
Biology, Rheinlanddamm 201, 44139 Dortmund,
Germany. ³Max Planck Institute for Biophysical
Chemistry, Department of Molecular Genetics, 37018
Göttingen Germany.

Correspondence should be addressed to I.S. email:
ilme.schlichting@mpi-dortmund.mpg.de

Acknowledgments

We thank G. Holtermann for his skilled technical assistance, P. Herde and S. Hönig for help in enzyme purification and kinetic assays, and A. Beste for help in the synthesis of AZTMP. A.L. was supported in part by a fellowship from the Alexander von Humboldt Stiftung and in part from the Human Frontier Science Program.

Received 30 April 1997; accepted 25 June 1997.

reduced with XDS²⁶ and scaled with XSCALE.

AZTMP complex crystals were grown using the hanging drop method by equilibrating equal volumes of a 10 mg ml⁻¹ TmpK solution and 2 mM AZTMP with a solution of 22% PEG 5000 monomethylether, 200 mM ammonium sulfate, and 100 mM HEPES buffer pH 7.5. Data were collected

- Balzarini, J., Herdewijn, P. & De, C.E. *J. Biol. Chem.* **264**, 6127–6133 (1989).
- Huang, P., Farquhar, D. & Plunkett, W. *J. Biol. Chem.* **267**, 2817–2822 (1992).
- Frick, L.W., Nelson, D.J., St. Clair, M.H., Furman, P.A. & Krenitsky, T.A. *Biochem. Biophys. Res. Commun.* **154**, 124–129 (1988).
- Fridland, A., Connelly, M.C. & Ashmun, R. *Mol. Pharmacol.* **37**, 665–670 (1990).
- Furman, P.A., et al. *Proc. Natl. Acad. Sci. U S A* **83**, 8333–8337 (1986).
- Qian, M., Bui, T., Ho, R.J. & Unadkat, J.D. *Biochem. Pharmacol.* **48**, 383–389 (1994).
- Lavie, A., et al. *Nature Med.* **3**, 922–924 (1997).
- Wild, K., Bohner, T., Aubry, A., Folkers, G. & Schulz, G.E. *FEBS Lett.* **368**, 289–292 (1995).
- Brown, D.G., et al. *Nature Struct. Biol.* **2**, 876–881 (1995).
- Vonrhein, C., Schlauderer, G.J. & Schulz, G.E. *Structure*

- 3**, 483–490 (1995).
- Saraste, M., Sibbald, P.R. & Wittinghofer, A. *Trends Biochem. Sci.* **15**, 430–434 (1990).
- Kim, H.J., et al. *Biochemistry* **29** 1107–1111 (1990).
- Tsai, M.-D. & Yan, H. *Biochemistry* **30**, 6806–6818 (1991).
- Yan, J.P., et al. *J. Biol. Chem.* **270**, 22836–22841 (1995).
- Bridges, E.G., Faraj, A. & Sommadossi, J.P. *Biochem. Pharmacol.* **45**, 1571–1576 (1993).
- Harrington, J.A., Reardon, J.E. & Spector, T. *Antimicrob. Agents Chemother.* **37**, 918–920 (1993).
- Caruso, M. & Klatzmann, D. *Proc. Natl. Acad. Sci. USA* **89**, 182–186 (1992).
- Gilman, A.G., Ralli, T.W., Nies, A.S. & Taylor, P. *Goodman and Gilman's the Pharmacological Basis of Therapeutics* (McGraw-Hill, New York; 1990).
- White, J.H., Green, S.R., Barker, D.G., Dumas, L.B. & Johnston, L.H. *Exp. Cell Res.* **171**, 223–231 (1987).
- Otwinowski, Z. in *Isomorphous Replacement and*

- Anomalous Scattering: Proceedings of the CCP4 Study Weekend* (eds Wolf, W., Evans, P.R. & Leslie, A.G.W.) 80–87 (SERC Daresbury Laboratory, Warrington, UK; 1991).
- Kleywegt, G.J. & Jones, T.A. in *From Map to Final Model* (eds Bailey, S., Hubbard, R. & Waller, D.) 59–66 (SERC Daresbury Laboratory, Warrington, UK; 1994).
- Cowtan, K. Joint CCP4 and ESF-EACBM Newsletter on Prot. Crystallogr. **31**, 24–28 (1994).
- Jones, T.A., Zhou, J.-Y., Cowan, S.W. & Kjeldgaard, M. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Brünger, A.T. *X-PLOR: a system for X-ray crystallography and NMR* (Yale University Press, New Haven, Connecticut; 1993).
- Read, R.A. *Acta Crystallogr. A* **42**, 140–149 (1986).
- Kabsch, W. *J. Appl. Crystallogr.* **24**, 795–800 (1993).
- Otwinowski, Z. & Minor, W. *Meth. Enz.* **276**, 307–326 (1997).
- Kraulis, P.J. *J. Appl. Crystallogr.* **24**, 946–950 (1991).