Acta Cryst. (1995). D**51**, 536–540

Calculation of Relative Binding Affinities of Purine Nucleoside Phosphorylase Inhibitors

BY MIKE CARSON AND ZI YANG

Center for Macromolecular Crystallography, University of Alabama at Birmingham, Birmingham, AL 35294, USA

AND Y. S. BABU AND JOHN A. MONTGOMERY

BioCryst Pharmaceuticals, Inc., 2190 Parkway Lake Drive, Birmingham, AL 35244, USA

(*Received* 12 *July* 1994; *accepted* 9 *January* 1995)

Abstract

The competitive binding of inhibitors to purine nucleoside phosphorylase (PNP) has been experimentally measured. Fast and reliable computational methods to estimate binding would allow assessment of any proposed inhibitor before its synthesis. Binding-energy calculations with a representative set of PNP inhibitors were compared to the empirical values. Relatively simple and fast calculations were executed with *X-PLOR*, *DelPhi* and *SoftDock*. The computational results are mixed.

1. Introduction

The enzyme purine nucleoside phosphorylase (PNP) converts purine nucleosides plus phosphate into purine bases and sugar-phosphates. Potent competitive inhibitors have been designed for PNP, based on empirical crystallographic analysis (Ealick *et al.*, 1991). Some of the inhibitors are currently being tested as drugs in the treatment of arthritis and psoriasis. A summary of this work is given by Bugg, Carson & Montgomery (1993).

PNP is composed of three identical subunits of 289 amino-acid residues that undergo significant conformational change on binding substrates. The active site lies in a cleft at a subunit interface, and may be subdivided into distinct guanine, sugar and phosphate sites. The guanine site is very polar, but neutral overall. The phosphate site is positively charged. Both these sites are well defined by fairly rigid residues from a single subunit, and are linked by the flexible hydrophobic sugar site at the subunit interface. This site lies at the entrance of the active site, and is fully formed only upon substrate binding.

The drug-design strategy was to fill the active site with a molecule complementary in shape and chemical properties, starting with the guanine site. BioCryst determined the atomic coordinates of PNP and of its complexes with inhibitors. BioCryst also determined IC_{50} values as an empirical measure of each inhibitor's binding affinity. The Center for Macromolecular Crystallography (CMC) evaluated several ways of calculating this binding. Comparisons of the observed and calculated binding are presented.

2. Methods

2.1. *Inhibitors*

Most PNP inhibitors are based on the substrate guanine. All inhibitors in this study are 9-deazaguanines, where the guanine base to ribose sugar (N9—C1) bond has been replaced by a C—C bond. One guanine and seven guanine-based inhibitors are shown in Fig. 1. A 3-character code is employed: gua is guanine; ben has a benzyl ring occupying the sugar site; ocb, mcb, pcb are the *ortho*, *meta*, *para*-chlorinated ben; b04, b05 are stereoisomers with a CO_2^- for the PO_4^- site; bio is a BioCryst test compound with a pyridine ring.

The IC_{50} value is the inhibitor concentration required for 50% inhibition of the normal reaction. The IC_{50} value is taken as the empirical measure of inhibitor binding affinity. IC_{50} values were measured with 'high-phosphate' (50 m*M*) and 'physiological' (1 m*M*) concentrations of phosphate in the assay (Ealick *et al.*, 1991; Montgomery *et al.*, 1993; Erion *et al.*, 1993).

2.2. *Crystallography*

The atomic coordinates of PNP and of its complexes with inhibitors were determined by single-crystal X-ray diffraction methods. The native enzyme and guanine/enzyme complex were refined by standard crystallographic techniques. Native enzyme coordinates refined at 2.75 Å resolution to an *R* factor of 0.20 (Narayana *et al.*, 1995).

The PNP/guanine complex refined at 2.7 Å resolution to an *R* factor of 0.21. The seven inhibitor structures were fit to difference maps; they have undergone little refinement (Y. S. Babu, BioCryst, personal communication). The guanine moiety is tightly bound, with hydrogen bonding to the enzyme resembling a G·C base pair of DNA (see Fig. 2).

2.3. *Semi-empirical methods*

The inhibitor coordinates came from small-molecule fragments. Semi-empirical calculations with *MOPAC* (Stewart, 1990) determined optimal geometries, dipole moments, and partial atomic charges. These were gas-phase calculations; solvation effects were ignored.

ISSN 0907-4449

2.4. *X-PLOR*

The geometries and partial charges from *MOPAC* allow integration with *X-PLOR* (Brünger, 1992). Topology/parameter files were constructed for each compound. Atom types were assigned by analogy to atoms present in the library of residues provided with *X-PLOR*. All PNP inhibitors, PNP/inhibitor complexes, and apoenzymes were energy minimized with *X-PLOR*. Formal charges were turned off, except for key residues in the active site. The inhibitor and all C_α coordinates were restrained to their original position for 100 cycles, followed by 150 cycles of unrestrained minimization.

A partial binding energy may be computed as,

$$E_{\text{binding}} = E_{\text{complex}} - (E_{\text{pnp}} + E_{\text{inhibitor}}).$$

These E values are potential energies (kcal) of the energy-minimized structures. However, *X-PLOR* includes no solvation term, known to be an important component of the true free energy of binding.

The method of Eisenberg & McLachlan (1986) estimates the free energy of solvation from the surface area of five basic atom types. Charged N, charged O, neutral N/O, C and S atoms all have a characteristic constant on the order of 25 cal $Å^{-2}$ mol. An *X-PLOR* script was written to calculate the solvent-accessible surface area (Lee & Richards, 1971) of each atom, multiply it by the appropriate constant, and sum the results to estimate the free energy. The organic Cl atoms were treated the same as C atoms.

The *X-PLOR* potential energy and the estimated solvation free energy will be arbitrarily combined into an 'energy',

$$E' = E_{X\text{-}PLOR}\ (\text{kcal}) + \text{scale} * \Delta G_{\text{solv}}\ (\text{kcal mol}^{-1}).$$

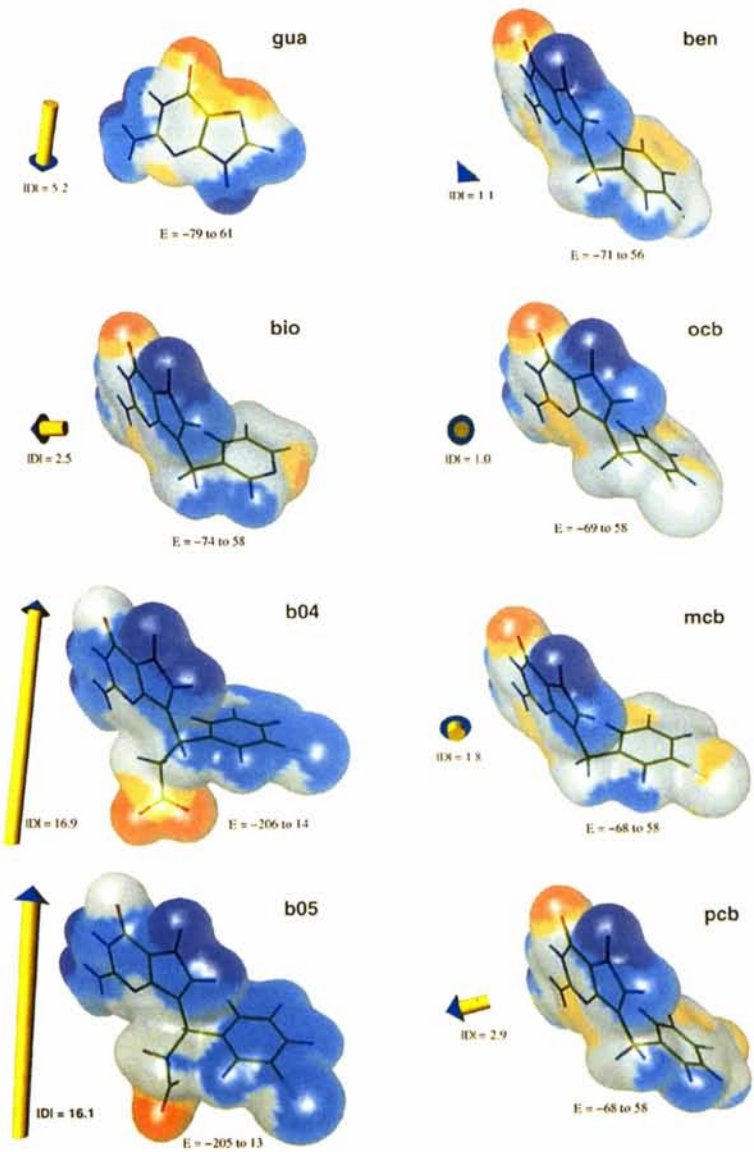

Fig. 1. The structures of guanine and the inhibitors are shown and labeled with a three-character code. The C, O, N, H and Cl atoms are colored green, red, blue, cyan and white. The electrostatic potentials are generated by *SPARTAN* (Henre, Burke, Shusterman & Pietro, 1993). Each compound is shown with a semi-transparent electron-density surface, color coded by the potential. The five colors (red, orange, gray, cyan, blue) map five equal ranges from the labeled minimum (negative, red) through the maximum (positive, blue) surface potentials in kcal mol^{-1}. The dipole vectors are drawn to scaled with their magnitudes labeled; blue points in the positive direction.

We also consider the two cases where,

$$E' = E_{X\text{-}PLOR}$$
$$E' = \Delta G_{solv}.$$

The binding energy is then given by,

$$E'_{binding} = E'_{complex} - (E'_{pnp} + E'_{inhibitor}).$$

2.5. *DelPhi*

The partial charges and coordinates allow calculations of electrostatic properties and solvent effects with *DelPhi* (Gilson & Honig, 1988). Another paper in this volume gives considerably more background on the theory and thermodynamic cycles employed (Honig *et al.*, 1995). Input parameters as suggested in the tutorial were used.

Solvation free energies were computed for each inhibitor as,

$$\Delta G_{solv} = \Delta G_{water}\ (D = 78) - \Delta G_{vacuum}\ (D = 1)$$

with inhibitor dielectric constant $D = 2$. All inhibitors were run at a grid spacing of 0.5 Å. Binding energies were computed in water at physiological ionic strength as,

$$\Delta G_{binding} = \Delta G_{complex} - (\Delta G_{pnp} + \Delta G_{inhibitor}).$$

The PNP trimer requires a grid spacing of > 2 Å due to the grid limitations in the finite element implementation. A shell of residues around the active site was extracted leading to a grid spacing of 0.8 Å, however, this truncation of the structure violated the spirit of *DelPhi*. A charged residue far from the active site can in principal make a significant contribution to the electrostatic potential at the active site.

2.6. *SoftDock*

The 'soft docking' method of Jiang & Kim (1991) has been implemented as the CMC program *SoftDock*. The method scores interactions based on shared surface area sampled in a grid of small cubes. The number of matching cubes, based on complementary surface normals, score the degree of geometric fit. A penalty for volume overlap is deducted. Atoms are divided into six classes: negative, positive, hydrogen-bond donor, hydrogen-bond acceptor, polar or hydrophobic. Pairwise interactions between inhibitor and enzyme are tallied, with each possible pair being assigned as favorable or unfavorable (1 or –1). These are summed to give an interaction energy score which is added to the geometric term.

All calculations were run with a cube size of 1.0 Å^3. An average of two surface dot/normals per cube was generated. No sampling of rotational/translational space by the inhibitor was carried out; only a single calculation with the observed coordinates of the PNP/inhibitor complex was performed.

2.7. *Data fitting and averaging*

The calculated binding scores were assessed by their correlation with the negative log of the observed IC_{50} values. The standard correlation coefficient was computed. The optimal scale to weight the *X-PLOR* energy and solvation terms was chosen to give maximum correlation.

The root-mean-square and standard deviation of the eight calculated binding results from *X-PLOR*, *DelPhi* and *SoftDock* were computed. The results were then normalized to standard deviation units relative to the mean. These normalized results were simply averaged to give a score for various combinations of methods.

3. Results

The partial charges generated from the semi-empirical gas-phase calculations appear reasonable from inspection. The electrostatic potential and dipole moment of each inhibitor is shown in Fig. 1.

The inhibitors and complexes were well behaved during *X-PLOR* refinement. The resultant coordinate shifts from the starting model were relatively minor from

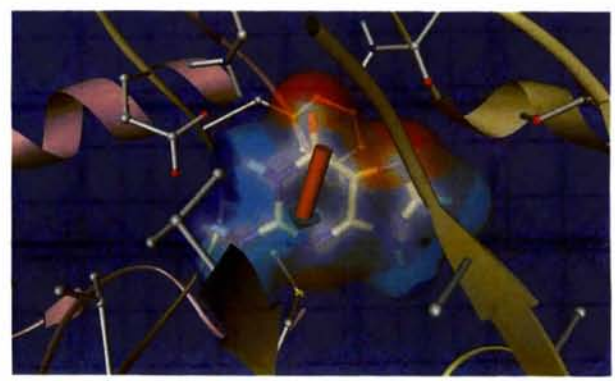

Fig. 2. Binding of guanine to PNP. The potential surface and dipole of Fig. 1 is superposed on the active site on PNP and rendered with the prototype *RIBBONS++*. Note the alignment of the dipole with the Glu and Asp (upper left corner). The mismatch of potential with the Asn (upper right corner) suggests a preference for the other tautomer of guanine.

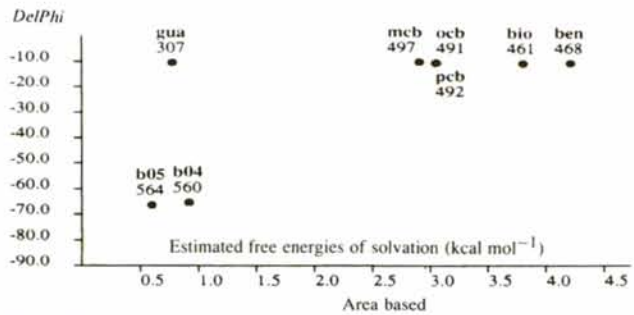

Fig. 3. Estimated free energies of solvation. The x axis plots the area-based values from *X-PLOR*. The y axis plots the results from *DelPhi*. The results are in kcal mol^{-1}. The area (Å^2) of each compound is also given. Note b04 and b05 are charged (–1).

visual inspection. The r.m.s. shift of atoms from the starting coordinates to the energy-minimized form was about 0.5 Å over the entire structure.

The electrostatic potential map at the active site generated by *DelPhi* from the PNP coordinates also appears reasonable from inspection. *DelPhi* runs at the two ionic strengths showed little difference.

The estimated solvation energies of each inhibitor based only on their surface areas calculated with *X-PLOR* are compared to the results computed by *DelPhi*

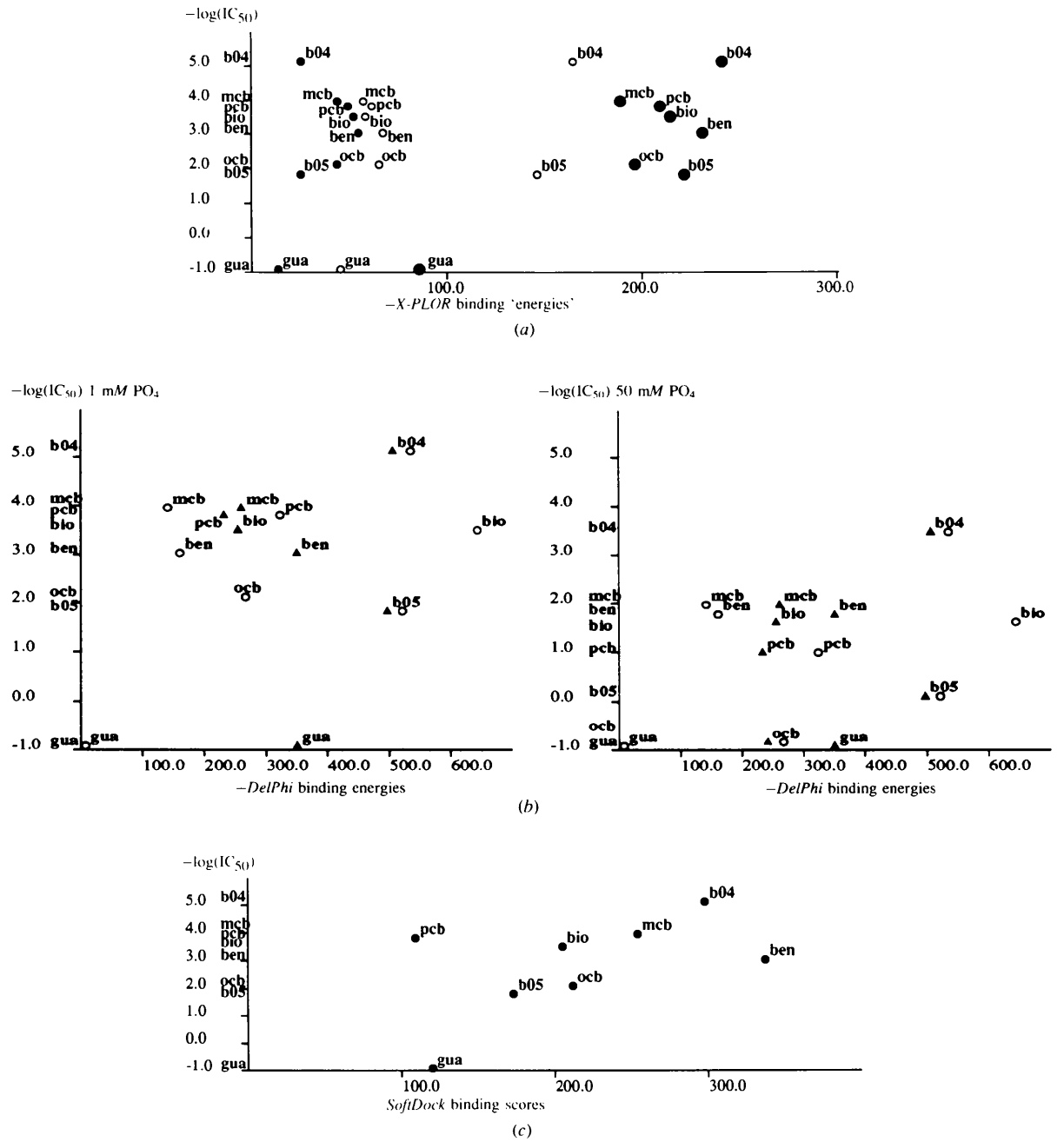

Fig. 4. (*a*) *X-PLOR* binding energy *versus* IC_{50}. The *x* axis plots the negative *X-PLOR* 'binding energy', E'. The *y* axis plots the negative logarithm of the IC_{50} value determined at physiological phosphate concentration. The small filled circles are $E' = \Delta G_{solv}$ (×10). The open circles are $E' = E_{X\text{-}PLOR}$. The large filled circles are $E' = E_{X\text{-}PLOR}$+ scale × ΔG_{solv}, where the optimal scale factor of 30.0 was employed. (*b*) *DelPhi* binding energy *versus* IC_{50}. The *x* axis plots the negative *DelPhi* free energy of binding, ΔG_b. The *y* axis plots the negative logarithm of the IC_{50} value. The left-hand plot is for physiological phosphate concentration; the right-hand plot is for high-phosphate concentration. The open circles are data using only a shell of residues around the active site. The filled triangles employed the entire trimer. (*c*) *SoftDock* interaction score *versus* IC_{50}. The *x* axis plots the *SoftDock* score. The *y* axis plots the negative logarithm of the IC_{50} value.

Table 1. *Correlation between experimental and calculated binding*

The $-\log(IC_{50})$ was used as the observed value. The program refers to *X-PLOR*, *DelPhi*, *SoftDock*, or a combination of the methods. The conditions list any special qualifiers and the phosphate concentration in the IC_{50} assay.

Correlation	Program, conditions
0.360	*X-PLOR*, energy terms only, low phosphate
0.529	*X-PLOR*, solvation terms only, low phospate
0.826	*X-PLOR*, energy + solvation, low phosphate
0.417	*X-PLOR*, energy terms only, high phosphate
0.277	*X-PLOR*, solvation terms only, high phosphate
0.645	*X-PLOR*, energy + solvation, high phosphate
0.520	*DelPhi*, shell of residues, low phosphate
0.404	*DelPhi*, shell of residues, high phosphate
0.006	*DelPhi*, entire trimer, low phosphate
0.268	*DelPhi*, entire trimer, high phosphate
0.530	*SoftDock*, low phosphate
0.657	*SoftDock*, high phosphate
0.792	*X-PLOR* + *SoftDock* + *DelPhi*, low phosphate
0.766	*X-PLOR* + *SoftDock*, low phosphate
0.717	*X-PLOR* + *DelPhi*, low phosphate
0.737	*SoftDock* + *DelPhi*, low phosphate

in Fig. 3. The optimal scale factor to combine *X-PLOR* potential energies and area-based solvation effects was 30, using data from either phosphate concentration. The calculated binding energies are compared to the observed binding energies in the scatter plots of Fig. 4. Results are plotted from the *X-PLOR*, *DelPhi* and *SoftDock* results. Table 1 lists the correlations between the calculated and observed values.

4. Discussion

Visual inspection and chemical intuition guided the original PNP drug design effort (Ealick *et al.*, 1991). The most effective use of computational methods was mapping the characteristics of the available volume in the active site. The calculations presented here were carried out after the fact to determine if computationally fast methods would reproduce the observed binding trends. Reliable and fast calculations would encourage the searching of databases for novel lead compounds.

The best correlation (0.83) was obtained from *X-PLOR* by combining the potential energy and the scaled area-based solvation term; however, the arbitrary scale (30) was adjusted to obtain the best result. It is not clear if this scale factor would transfer to other systems. The best correlations with the potential energy only, the area-based solvation term only, *DelPhi*, and *SoftDock* are of the order of 0.5. The shell of residues, with its finer grid sampling, gave consistently better results with *DelPhi*, compared to runs employing the entire trimer. Combining any combination of the scores produces correlations of over 0.7.

Examination of the scatter plots shows the general trend is indeed as anticipated. The poorest binding molecule in this study is the substrate guanine, almost always having the lowest score. The best binding molecule, b04, has the highest score for most methods. However, guanine is the smallest and b04 is the largest inhibitor, and molecules b04 and b05 are negatively charged. One might argue that the results simply divide the molecules into three classes: small neutral, medium neutral and large charged. This is certainly true for the *X-PLOR* and *DelPhi* results. The *SoftDock* method is the only one which correctly distinguishes between the stereoisomers b04 and b05. None of the methods, nor any combination, does a good job distinguishing the *ortho*, *meta* and *para*-substituted benzyl series of compounds.

These methods require about 1 h per compound on our Indigo workstations; the *SoftDock* method requires only minutes. This is fast enough; and c.p.u. speed will only increase. The question is the reliability of the result. For example, take the top four compounds ranked by the *X-PLOR* binding energy: b04, ben, b05, and bio. The top four ranked by the observed binding are b04, mcb, pcb and bio. Is this half-right or half-wrong?

The PNP system provides an ideal testing ground for the comparison of computational theory and experiment. We plan to investigate other scoring methods presented here (Blaney & Dixon, 1993; Watson *et al.*, 1995; Hol *et al.*, 1995) in our search to rationalize the relative binding affinities of inhibitors.

We gratefully acknowledge NASA grant NAGW-813 for support, Marek Jedrzejas for help with *DelPhi* and Charlie Bugg for advice and encouragement.

References

BLANEY, J. M. & DIXON, J. S. (1993). *Perspect. Drug Disc. Design*, **1**, 301–319.

BRÜNGER, A. T (1992). *X-PLOR. Version* 3.1. *A System for X-ray Crystallography and NMR*. New Haven: Yale Univ. Press.

BUGG, C. E., CARSON, W. M. & MONTGOMERY, J. A. (1993). *Sci. Am.* **269**, 92–98.

EALICK, S. E., BABU, Y. S., BUGG, C. E., ERION, M. D., GUIDA, W. C., MONTGOMERY, J. A. & SECRIST, J. A. III (1991). *Proc. Natl Acad. Sci. USA*, **88**, 11540–11544.

EISENBERG, D. & MCLACHLAN, A. D. (1986). *Nature (London)*, **319**, 199–203.

ERION, M. D., NIWAS, S., ROSE, J. D., ANANTHAN, S., ALLEN, M., SECRIST, J. A. III, BABU, Y. S, BUGG, C. E., GUIDA, W. C., EALICK, S. E. & MONTGOMERY, J. A. (1993). *J. Med. Chem.* **36**, 3771–3783.

GILSON, M. K. & HONIG, B. (1988). *Proteins Struct. Funct. Genet.* **4**,7–18.

HENRE, W. J., BURKE, L. D., SHUSTERMAN, A. J. & PIETRO, W. J. (1993). *Experiments in Computational Organic Chemistry*. Pomona, California: Wavefunction Inc.

JIANG, F. & KIM, S.-H. (1991). *J. Mol. Biol* **219**, 79–102.

LEE, B. & RICHARDS, F. M. (1971). *J. Mol. Biol.* **55**, 379–400.

MONTGOMERY, J. A., NIWAS, S., ROSE, J. D., SECRIST, J. A. III, BABU, Y. S, BUGG, C. E., ERION, M. D., GUIDA, W. C. & EALICK, S. E. (1993). *J. Med. Chem.* **36**, 55–69.

NARAYANA, S. V. L, BUGG, C. E. & EALICK, S. E. (1995). *Acta Cryst.* D**51**. Submitted.

STEWART, J. J. P. (1990). *MOPAC* 6.0, *A General Molecular Orbital Package. Quantum Chemistry Program Exchange*, No. 455, Vol. 10, pp. 86–86.

VELLIEUX, F. M. D., HAJDU, J. & HOL, W. G. J. (1995). *Acta Cryst.* D**51**, 575–589.

WATSON, K. A., MITCHELL, E. P., JOHNSON, L. N., CRUCIANI, G., SON, J. C., BICHARD, C. J. F., FLEET, G. W. J., OIKONOMAKOS, N. G., KNOTOU, M. & ZOGRAPHOS, S. E. (1995). *Acta Cryst.* D**51**, 458–472.