# Ultrahigh Resolution Drug Design I: Details of Interactions in Human Aldose Reductase–Inhibitor Complex at 0.66 Å

**E.I. Howard,[1] R. Sanishvili,[2] R.E. Cachau,[3] A. Mitschler,[1] B. Chevrier,[1] P. Barth,[4] V. Lamour,[1] M. Van Zandt,[5] E. Sibley,[5] C. Bon,[6] D. Moras,[1] T.R. Schneider,[7] A. Joachimiak,[2]* and A. Podjarny[1]***

[1]*Laboratoire de Génomique et de Biologie Structurales, UMR 7104 du CNRS, IGBMC, Illkirch, France*
[2]*Structural Biology Center, Biosciences Division, Argonne National Laboratory, Argonne, Illinois*
[3]*Advanced Biomedical Computing Centre, National Cancer Institute, SAIC, Frederick Maryland*
[4]*Laboratoire de Géochimie Biorganique UMR 7509 ECPM 25, Strasbourg, France*
[5]*Institute for Diabetes Discovery, Inc., Branford, Connecticut*
[6]*Institut de Pharmacologie et de Biologie Structurale, UMR5089, Toulouse, France*
[7]*Goettingen University, Department of Structural Chemistry, Goettingen Germany*

**ABSTRACT** The first subatomic resolution structure of a 36 kDa protein [aldose reductase (AR)] is presented. AR was cocrystallized at pH 5.0 with its cofactor $NADP^+$ and inhibitor IDD 594, a therapeutic candidate for the treatment of diabetic complications. X-ray diffraction data were collected up to 0.62 Å resolution and treated up to 0.66 Å resolution. Anisotropic refinement followed by a blocked matrix inversion produced low standard deviations (<0.005 Å). The model was very well ordered overall (CA atoms' mean B factor is 5.5 Å²). The model and the electron-density maps revealed fine features, such as H-atoms, bond densities, and significant deviations from standard stereochemistry. Other features, such as networks of hydrogen bonds (H bonds), a large number of multiple conformations, and solvent structure were also better defined. Most of the atoms in the active site region were extremely well ordered (mean B ∼3 Å²), leading to the identification of the protonation states of the residues involved in catalysis. The electrostatic interactions of the inhibitor's charged carboxylate head with the catalytic residues and the charged coenzyme $NADP^+$ explained the inhibitor's noncompetitive character. Furthermore, a short contact involving the IDD 594 bromine atom explained the selectivity profile of the inhibitor, important feature to avoid toxic effects. The presented structure and the details revealed are instrumental for better understanding of the inhibition mechanism of AR by IDD 594, and hence, for the rational drug design of future inhibitors. This work demonstrates the capabilities of subatomic resolution experiments and stimulates further developments of methods allowing the use of the full potential of these experiments. Proteins 2004;55:792–804. © 2004 Wiley-Liss, Inc.

**Key words: X-ray crystallography; subatomic resolution; drug design; diabetes complications**

## INTRODUCTION

The number of macromolecular structures solved each year at sub-Angstrom resolution increases continuously, due to improvements in expression systems, crystallogenesis, crystallographic methods, and instrumentation enabling the use of the high potential of the brilliant synchrotron radiation sources. Subatomic resolution (d < 0.8 Å[1,2]) structures of small- and medium-size biomolecules have become reality. However, because of stringent requirements in all steps including crystal growth and data acquisition and analysis, the number of such structures remains small, and is limited by the protein size, for example, crambin (0.54 Å diffraction, 46 residues, three S—S bonds),[3,4] trypsin (0.81 Å diffraction, 242 residues, three S—S bonds)[5] and subtilisin (0.78 Å diffraction, 269 residues).[6] For more details, see the reviews by Esposito et al.,[7] Dauter et al.,[8] and Schmidt and Lamzin.[2]

In the presented work, we were able to push the limits of resolution and molecular size one step further by studying the structure of the ternary complex of human aldose reductase, its cofactor $NADP^+$, and the inhibitor IDD 594 refined at 0.66 Å resolution, the highest resolution ever reported for a monomeric enzyme with a molecular weight of 36 kDa and no S—S bonds.

Aldose reductase (AR; EC 1.1.1.21) is a 316-amino acid enzyme that reduces a wide range of substrates, such as aldehydes, aldoses, and corticosteroids. As it reduces D-glucose into D-sorbitol, it is believed to be involved in degenerative complications of diabetes, because the accumulation of excess sorbitol resulting from diabetic hyperglycaemia has been shown to play a role in the development of diabetic complications such as retinopathy, neuropathy, and nephropathy.[9,10] The catalytic reaction involves a hydride transfer from the C4 atom in the nicotinamide ring of the coenzyme NADPH to the C atom of the substrate carbonyl and a proton donation from the enzyme to the O atom of the substrate carbonyl.[11] Kinetic studies of the reduction of DL-glyceraldehyde by pig muscle AR showed that at pH 7.0 it follows a sequentially ordered mechanism with NADPH binding first, leading to a conformational change in the enzyme, and the substrate binding afterwards.[12] Some aldehyde substrates, such as glycoaldehyde, displayed an inhibitory effect, shown by a decrease of the initial velocity with increasing substrate concentration.[13] Replacement of the aldehyde moiety of substrates by a carboxylate made potent inhibitors in the forward direction, as the enzyme[14] does not reduce them. However, some of these inhibitors show toxic effects mainly because of interference with aldehyde reductase, a key enzyme for degrading aldehydes in the kidney.[10,14] Therefore, the development of AR-specific, potent and nontoxic inhibitors is an important goal for diabetes treatment.

Several X-ray[15–19] site-directed mutagenesis[20–23] and modeling studies[24] of AR its enzymatic mechanism[25] and its inhibitor binding[26] have been carried out. AR folds as a $(\beta/\alpha)_8$ barrel with the $NADP^+$ nicotinamide ring buried at the bottom of the active site cleft of approximate dimensions 9 Å (diameter) and 11 Å (depth). Our previous crystallographic studies[17] showed that AR inhibitors of the carboxylic acid (e.g., tolrestat) and of the spirohydantoin (e.g., sorbinil) families bind at the catalytic site, making hydrogen bonds with Tyr 48, His 110, and Trp 111. Some of these inhibitors induce a conformational change that creates a nearby "specificity pocket"[17,19,27] lined by residues Trp 111, Thr 113, Phe 122, Ala 299, and Leu 300 [Fig. 1(a)]. The essential role played by Trp 111 and by Thr 113 in inhibitor binding has been confirmed by site-directed mutagenesis and modeling studies.[26,28] The residues forming the specificity pocket are not conserved between aldose and aldehyde reductases, and therefore, inhibitors bind to these two enzymes with different affinities, providing a key feature for developing discriminating inhibitors without toxic effects. An example is provided by epalrestat, used in drugs for treatment of diabetic neuropathy and reviewed by Miyamoto,[29] Constantino,[30] and El-Kabbani (submitted).

AR inhibitors containing carboxylic acids do not compete with aldehyde substrates for binding to the enzyme (for review, see ref. 14). This was confirmed for the inhibitor alrestatin.[32] Even though the lack of competitive inhibition in the forward reaction does not necessarily exclude binding of the inhibitor in the active site, it opened a discussion about the existence of a separate inhibitor-binding site.[32] However, such a site was not observed crystallographically. Furthermore, certain inhibitors are competitive for the reverse reaction (oxidation of the alcohol to the corresponding aldehyde), which is also catalyzed by AR.[33–35] To explain the inhibition characteristics of both the direct and reverse reactions, it was proposed[17,31] that charged inhibitors bind to the same site as do substrates, but only after the catalytic cycle has been completed; for example, they bind preferably the AR–$NADP^+$ form, whereas the substrates bind the AR–NADPH form. Such preferred binding of inhibitors to AR–$NADP^+$ has been confirmed by calorimetric measures of binding energies (Kraemer et al., manuscript in preparation). This interpretation fits the proposed model of "anionic site" for binding inhibitors and substrates. This site is situated between His 110, Tyr 48, and $NADP^+$ [Fig. 1(a)].[20,34] Following this model, the strong discrimination between $NADP^+$ and NADPH is more consistent with a neutral His 110 residue, because in this case the main electrostatic interaction of the inhibitor is formed with $NADP^+$. The presence of a neutral His 110 is consistent with the observation that the carboxylic acid inhibitor, alrestatin, still binds to the enzyme when His 110 is mutated to neutral Ala.[31] On the other hand, the charged inhibitor is expected to favor a doubly protonated His 110 making a salt bridge, as proposed by Lee et al.[24] It is clear that to resolve this controversy, the experimental observation of the protonation state of His 110 is essential. X-ray crystallography at ultrahigh resolution can provide this level of detail.

## RESULTS AND DISCUSSION

The crystal structure of a ternary complex of human AR (316 residues), its coenzyme $NADP^+$ and inhibitor IDD 594 [Fig. 1(b)] was solved and refined against diffraction data at 0.66 Å, as described in the Methods section, with a final $R$-factor of 9.38% for all reflections (Tables I–III).

### Inhibitor Preparation

The inhibitor IDD 594 was prepared as part of a cogeneric family of inhibitors sharing a common scaffold.

The structure–activity relations indicate that the carboxylate moiety, the ether oxygen of the phenoxyacetic acid, and the amide N—H are necessary for activity. The IDD 594 carboxylate group appears to be particularly important for efficient inhibition. Indeed, alkaloid inhibitors having hydrophobic bodies but no charged group, such as papaverin and berberin, bind AR weakly with much higher $IC_{50}$ values of 140,000 and 52,000 nM, respectively.[14] The aromatic side chain is critical for both potency and selectivity (Van Zandt et al., in preparation). The observed selectivity pattern is consistent with the major differences between human aldose reductase 2 (hALR2) and 1 (hALR1), which are in the C-terminal loop (adjacent to the IDD 594 side chain). The preferred substitution pattern for the benzyl side chain is either a 6-nitro group, or a fluorine in the 5-position with a bromide in the 7-position. Alternatively, the 10,11,13-trifluoro-

**TABLE I. Data Collection, Structure Refinement, and Model Status**

| | |
|---|---|
| Cell dimensions, space group | a = 49.28 Å, b = 66.59 Å, c = 47.26 Å $\beta$ = 92.4°, P2$_1$ |
| Data Collection: Resolution (Å) (last shell) | 20.0 to 0.66 (0.68 to 0.66) |
| No. of unique reflections(last shell) | 511265 (43107) |
| Completeness (%)(last shell) | 89.1 (75.4) |
| I/$\sigma$(I) (last shell) | 14.5 (2.4) |
| R(I)$_{merge}$ (%)(last shell) | 2.9 (27.9) |
| Refinement: Reflections with F>4$\sigma_F$ (for R$_{free}$) | 405150 (21261) |
| Isotropic: R$_{cryst}$/R$_{free}$ without H (%) | 16.30/16.99 |
| Anisotropic: R$_{cryst}$/R$_{free}$ without H (%) | 9.63/10.01 |
| R$_{cryst}$/R$_{free}$ with riding H (%) | 8.42/9.34 |
| Anisotropic using all reflections (for R$_{free}$) | 485678 (25587) |
| R$_{cryst}$/R$_{free}$ without H (%) | 10.56/10.94 |
| R$_{cryst}$/R$_{free}$ with riding H (%) | 9.36/10.32 |
| Model | |
| ESDs of fully occupied atomic positions | |
| Most ordered regions (B< 3 Å$^2$) | <0.005 Å |
| Rest of the molecule | 0.007 Å |
| RMSD from ideal values | |
| Bonds (Å) | 0.016 |
| Angles (Å) (comment: angles are restrained as 1–3 distances in SHELXL) | 0.034 |
| Ramachandran plot | |
| Residues in most favored regions (%) | 91.0 |
| Residues in additional allowed regions (%) | 9.0 |
| Mean B-factor (Å$^2$) (fully occupied, partially occupied) | |
| Protein main chain 1036/426 | 5.4/6.3 |
| Protein side chain 909/597 | 6.9/9.9 |
| NADP$^+$ 48/0 | 3.9/n/a |
| Inhibitor 24/0 | 4.2/n/a |
| Citrate 0/26 | n/a/6.9 |
| Solvent 415/198 | 19.7/12.5 |

It should be noted that at this resolution the deviation from ideal bond length values indicated here I (0.016 Å) corresponds mostly to true differences, as the typical standard uncertainty for bond lengths as derived from the blocked inversion of the refinement matrix (using the same methods as in small molecule crystallography) is on the order of 0.003 Å

benzothiazole side chain also provides excellent activity. The introduction of fluorine in the R 2 position provides optimal *in vitro* and *in vivo* activity. However, the thioamide is generally more efficacious *in vivo* than the amide. Thus, IDD 594, with a fluorine in the 5-position and a bromide in the 7-position in the general thioamide scaffold, improves both affinity and selectivity. This is shown by the IC$_{50}$ values for aldose reductase: 3300 nM for the scaffold without additions, 550 nM with an added bromine, 1300 nM with one chlorine substitution in the bromine position (IDD 594 scheme), and finally 30 nM with both Br and F (IDD 594). IDD 594 was tested against aldehyde reductase to establish selectivity. Aldehyde reductase is the main competitor for binding inhibition, as it performs the critical function of degrading aldehydes in the kidney. The measured IDD 594 IC$_{50}$ value against aldehyde reductase was 33,000 nM 1100 times higher than the one for AR. Kinetic and biological assays (Van Zandt et al., paper in preparation) show that IDD 594 reduces sorbitol concentration in the sciatic nerve by 87%, and that it does not compete with the substrate for binding.

## Protein Model

The final model shows the $(\beta/\alpha)_8$ barrel fold [Fig. 2(a)], including all the amino acids except for three C-terminal residues [Fig. 2(b)], and including all atoms of cofactor and inhibitor. Double occupancies have been assigned for 99 residues, which have at least one atom with more than one conformation. In the final model there are 613 water molecules with B-factors ranging from 2.8 to 56.9 Å$^2$ [Fig. 2(c)]; of these, 415 are fully occupied, 11 sites are alternatively occupied by water and citrate molecules, 49 water molecules have double sites, and the rest have partial occupations. Accuracy and resolution of the experiment lead to the observation of a number of fine details. In what follows we will describe some of these details.

## Electron Density and Hydrogen Atoms

The electron density maps show clear density for all nonhydrogen atoms. The atom types can be unambiguously assigned by the map volume around them in many regions of the protein, as well as the cofactor and inhibitor (Fig. 3). The hydrogen atoms were located in an F$_o$–F$_c$ difference map with hydrogens omitted (OMIT-H map). If the residual electron density had a value larger than 1 $\sigma$ at the expected hydrogen position, the latter was counted as observed. The percentage of observed hydrogens is linearly correlated with the temperature factor of the atoms to which the hydrogen is covalently linked (correlation coefficient of 0.97 for the B range 3–13 Å$^2$) (Fig. 4). Fifty-four percent of all the possible hydrogen atoms in the whole protein were observed [Fig. 5(a)], most of them riding on the atoms with the lowest B factors. In the active site region, which is particularly well ordered ($\langle B \rangle$ = 3.4 Å$^2$), 77% of all possible H-atoms are identified.

## Bond Lengths and Bond Densities

Residual electron density is observed around the highly ordered atoms, indicating departure from the spherical atom model. Such density is seen, for example, between atom centers in the zone 44–46 near the active site [Fig. 5(a)]. It corresponds to $\sigma$-bonds (involving s and p orbitals), as discussed in the refinement of AR using a multipolar model.[37] Bond densities in the peptide link 44–45 appear both in the C—N and C—O bonds, suggesting a resonance between them. On the other hand, in the link 45–46 the bond density appears mostly along the C—O bond but not along the C—N bond, suggesting a double bond C—O and a single bond C—N. Furthermore, these bond orders are confirmed by the C—O and C—N bond lengths, 1.241(3) and 1.330(3) Å for the 44–45 peptide link, and 1.220(3) and 1.355(3) Å for the 45–46 peptide link, respectively.

**TABLE II. List of Data Collection Scans for the Two Crystals Used**

| Scan | Res. Range | Frame width | Exposure time | Number of frames | 2θ |
|---|---|---|---|---|---|
| Very low (2 scans) | 20–2.15 Å | 2° | 1 s | 90 | 0° |
| Low (3 scans) | 20–2 Å | 1° | 1 s | 180 | 0° |
| Medium (2 scans) | 20–1 Å | 0.4° | 1 s | 450 | 0° |
| High (4 scans) | 20–0.65 Å | 0.2° | 15 s | 900 | 30° |

**TABLE III. Data Collection Statistics as a Function of Resolution Range**

| Lower shell limit (A) | Upper shell limit (A) | Completeness (%) | Average $I/\sigma(I)$ | Norm chi$^2$ | Linear R-mer |
|---|---|---|---|---|---|
| 20.00 | 1.79 | 99.2 | 18.78 | 1.028 | 0.022 |
| 1.79 | 1.42 | 99.6 | 13.45 | 1.026 | 0.032 |
| 1.42 | 1.24 | 99.8 | 12.35 | 1.023 | 0.035 |
| 1.24 | 1.13 | 99.8 | 11.41 | 1.025 | 0.038 |
| 1.13 | 1.05 | 99.6 | 11.28 | 1.019 | 0.041 |
| 1.05 | 0.99 | 98.9 | 11.52 | 1.025 | 0.045 |
| 0.99 | 0.94 | 96.6 | 10.61 | 1.024 | 0.053 |
| 0.94 | 0.90 | 94.4 | 10.11 | 1.022 | 0.060 |
| 0.90 | 0.86 | 92.3 | 9.59 | 1.029 | 0.067 |
| 0.86 | 0.83 | 90.1 | 8.64 | 1.024 | 0.079 |
| 0.83 | 0.81 | 88.4 | 7.93 | 1.025 | 0.089 |
| 0.81 | 0.78 | 86.6 | 7.26 | 1.023 | 0.099 |
| 0.78 | 0.76 | 84.9 | 6.51 | 1.028 | 0.110 |
| 0.76 | 0.74 | 83.3 | 5.74 | 1.027 | 0.126 |
| 0.74 | 0.73 | 81.8 | 4.97 | 1.025 | 0.144 |
| 0.73 | 0.71 | 80.1 | 4.41 | 1.029 | 0.161 |
| 0.71 | 0.70 | 78.5 | 3.72 | 1.029 | 0.191 |
| 0.70 | 0.68 | 77.5 | 3.19 | 1.028 | 0.220 |
| 0.68 | 0.67 | 76.0 | 2.66 | 1.020 | 0.260 |
| 0.67 | 0.66 | 74.7 | 2.22 | 0.990 | 0.301 |
| All Reflections | | 89.1 | 15.20 | 1.024 | 0.029 |
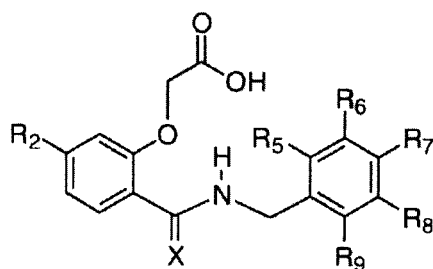
Scaffold                                    IDD 594
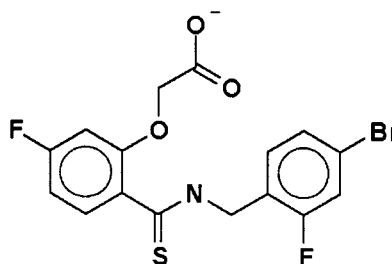


Scheme 1.

There is negative correlation between the C—N and the C—O bond lengths; when the order of one bond increases, the other one should decrease, as the total valence of the C atom remains constant. This trend is apparent for all the residues where the main chain temperature factors are less than 3 Å$^2$ (Fig. 6), revealing the shifting balance of resonance states of the peptide bond.

The electron density in the best defined regions is sufficiently accurate to compute "experimental" electrostatic and electron density deformation maps. The analy-sis of these maps, which open the possibility of studying *in situ* reactivity, is presented elsewhere.[1,36]

**Deviations from Standard Stereochemistry**

We have found 27 instances of 10° or larger deviations of the ω angle from planarity. In all cases but one, the main-chain N atom was involved in a strong H-bond. In at least two cases (residues 44 and 105) this ω deviation was correlated with sp$^3$ coordination of the N atom. In 10 of the 27 cases, the residues involved were in the vicinity of the
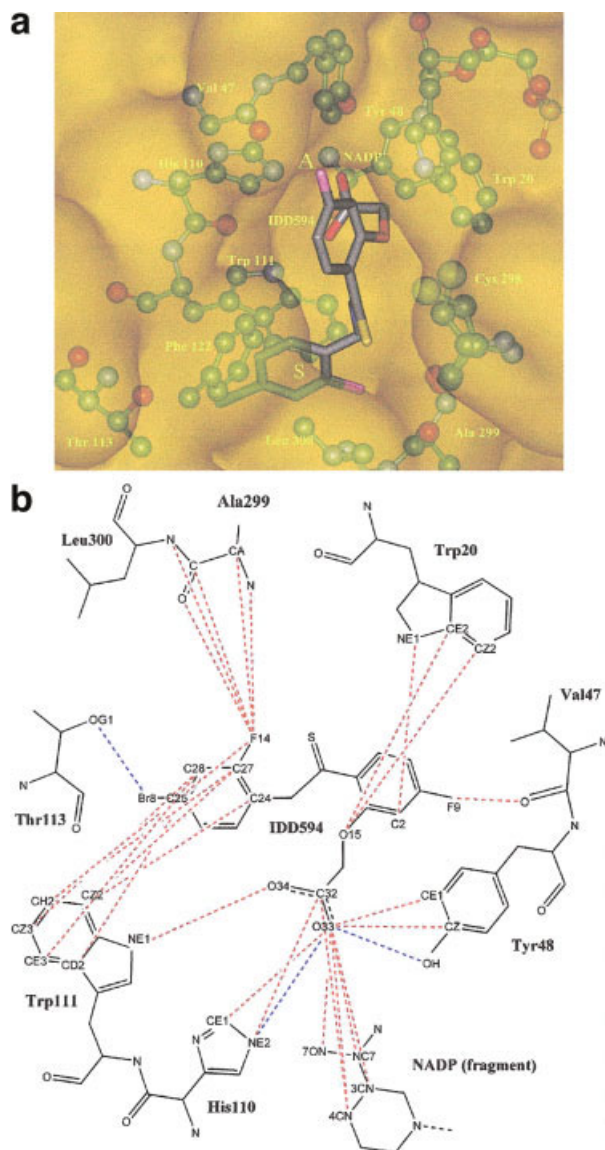
Fig. 1. Contacts of IDD 594 with AR and NADP$^+$. (**a**) 3D view of inhibitor contacts. View of the inhibitor-binding site down the barrel axis with the semitransparent representation of AR surface. The active site cleft (marked A), is occupied by the inhibitor (stick only; color code: C = gray, N = blue, O = red, F = pink, S = orange, Br = green) and surrounded by the catalytic residues His 110, Tyr 48, and the coenzyme NADP$^+$. The residues within 3.9 Å and NADP$^+$ are shown (balls and sticks). The specificity pocket between Leu 300, Phe 122, and Trp 111 (marked S) is occupied by the brominated aromatic ring of the inhibitor. (**b**) Scheme of IDD 594 and its contacts. Red dashed lines show contacts between 3.5 and 3.0 Å, and blue dashed lines show contacts <3.0 Å. The exact values are given in Table IV.



Fig. 2. Aldose reductase general features. (**a**) Cartoon scheme of the protein; (**b**) ORTEP plot of the protein CA atoms (50% probability ellipsoids). (**c**) ORTEP plot of the water molecules.

active site. One example can be found in the peptide bond 76–77, where the ω angle is −167.4(2)° [Fig. 5(b)]. This conformation is stabilized by an H-bond [2.681(3) Å] of the Lys 77 N with the OG atom of Ser 76, by a hydrophobic interaction of Ala 45 and the aliphatic chain of Lys 77 and by a salt bridge of Lys 77 with Asp 43. The involved hydroxyl (Ser 76) is conserved in all the enzymes of the aldo-ketoreductase family as part of Ser or Thr side chains.
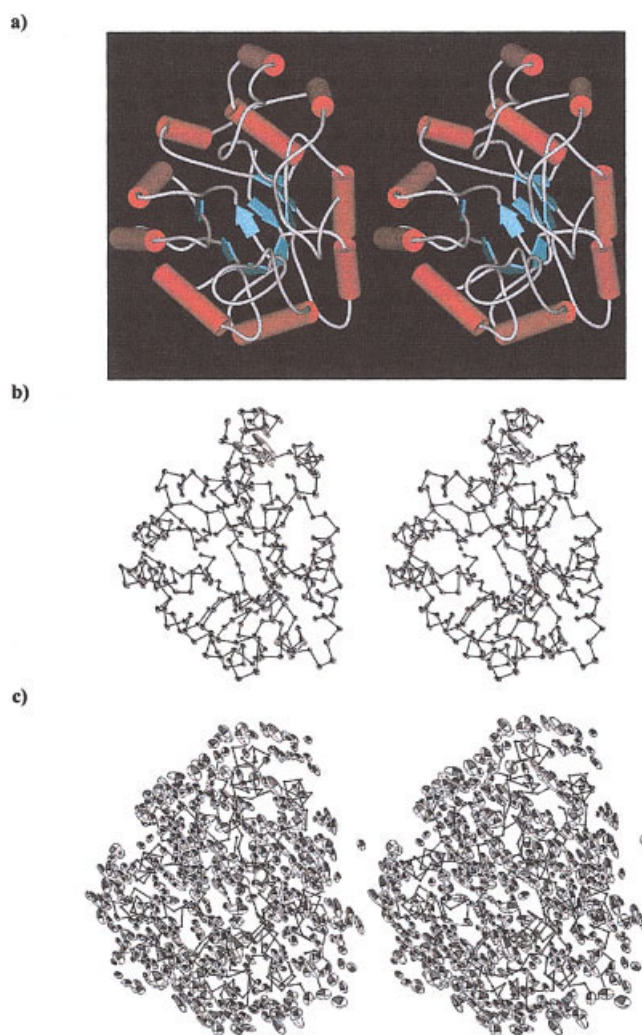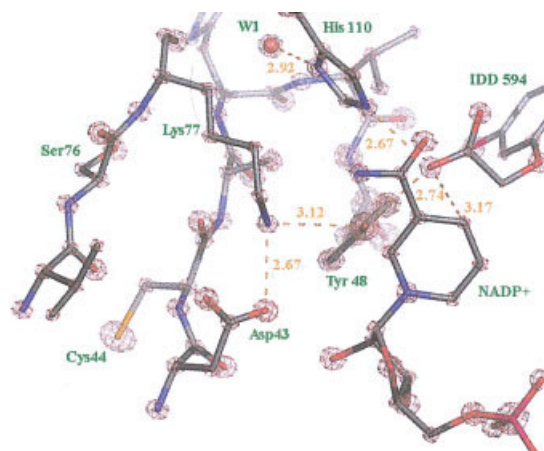


Fig. 3. Model of the active site, superimposed with a σA-weighted 2Fo–Fc map, contoured salmon at 8.0 e/Å$^3$ (8 σ). Note that the size of the electron density sphere around each atom is indicative of the atomic species. Note also the H-bond pattern between Lys 77, Tyr 48, His 110, and the inhibitor carboxylate head.
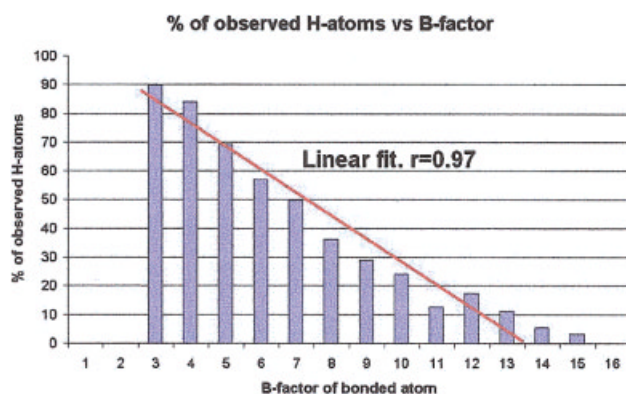
Fig. 4. Histogram of percentage of observed hydrogen atoms as a function of the B-factor of the bonded heavy atom.

## Short Contacts

Unusually short interatomic distances, some involving CH—O contacts, are observed. Figure 7(a) shows a well-refined region with such contacts. The short contact between Pro252-CD and Leu248-O [2.955(4) Å] can be considered as a typical C—H...O hydrogen bond for a proline residue situated in an alpha helix as described by Chakrabarti and Chakrabarti.[37] In the same region, there is another short contact between CZ of buried Arg 250 [Fig. 7(a)] and the disordered carbonyl oxygen of Phe 278-O [2.85(2) and 2.91(2) Å for the two different conformations]. The plane of the guanidinium group is almost perpendicular to the carbonyl C—O bond. The energy balance of this interaction, which might involve both quantum (orbital overlap) and electrostatic (between charged guanidinium and partially charged carbonyl oxygen) contributions needs further studies. The position of the Arg 250 side chain is stabilized by hydrogen bonds of the guanidinium N-atoms with Gln 254-OE [2.806(4) Å] and Val 275-O [2.883(4) Å]. There is also a very short contact [2.973(4) Å] between the bromine atom of the inhibitor and the hydroxyl oxygen of Thr 113, which is discussed in "details of inhibitor binding."

## Multiple Conformations

The high-resolution maps allow us to distinguish alternative conformations for many residues, even when the corresponding atoms are close in space. They can propagate along the chain over several residues [e.g., helices 192–202 or residues 295–298, Fig. 7(b)], or can involve residues that are close in space but far in the sequence, and show correlated double conformations. These can involve intermolecular contacts; for example, Glu 29 and Ser 127 make an intermolecular H-bond in one conformation and none in the other. Some conformations [e.g., those involving Asp 102 and Arg 293; see Fig. 7(c)] have clear alternative densities, with both conformations very well defined with temperature factors around 10 Å². Asp 102 exhibits clear water contacts stabilizing each conformation, while the solvent structure around Arg 293 is more disordered; this is likely due to higher flexibility of the arginine side chain. Multiple conformations are also seen for main-chain carbonyl atoms. It is also noteworthy that
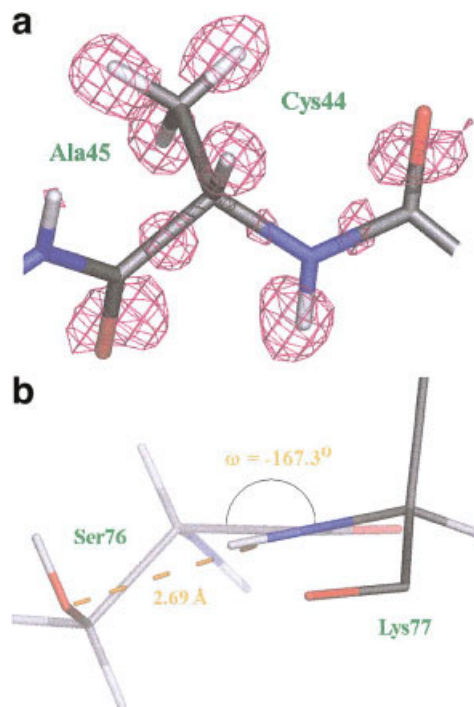


Fig. 5. Model of residues Cys 44 and Ala 45 superimposed with σA-weighted Fo–Fc map with omitted hydrogen atoms (**a**), contoured magenta at at 0.25 e/Å³ (2.5σ) showing the position of the hydrogen atoms and the bond densities. Model of residues Ser 76 and Lys 77(**b**), viewed down the C—N peptide bond showing the ω angle of −167.3° stabilized by an H-bond.
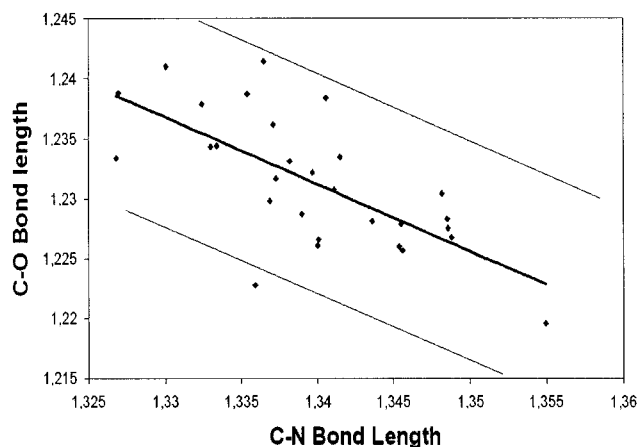


Fig. 6. Graph of the C—O vs. C—N bond lengths for the 29 residues with the lowest temperature factors. The thick line represents the regression curve [length(C—O) = 1.98 Å − 0.55 × length(CN), Correlation coefficient = 0.69]; the two thin lines represent the 2σ error margins assuming a RMSD value of 0.003 Å in both bond lengths.

Cys 298, whose mutation to serine diminishes catalytic efficiency,[38] has two clearly defined conformations, which are in strong hydrophobic interaction with the catalytic C4 carbon of $NADP^+$ [Fig. 7(d)].

## Solvent Structure

About two hydration sites per residue are found, while in a lower resolution refinement (∼2 Å) only one is usually
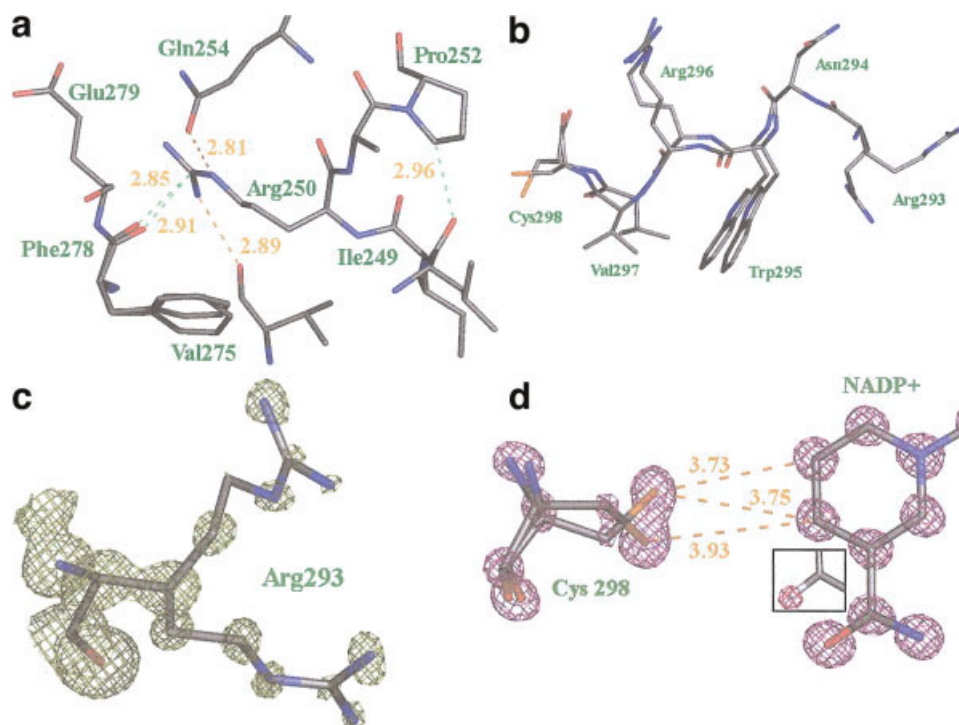
Fig. 7. Contacts of Arg 250 and Pro 252 (**a**). Note the CH—O contact between Pro 252-CD and Leu 248-O and the short C—O contact involving the buried side chain of Arg 250. (**b**) Double conformations of the residues 293–298. (**c**) Double conformation of Arg 293, superimposed with the 2Fo–Fc [σA-weighted, olive contours at 2.5 e/Å³ (2.5σ)]. (**d**) Details of the interaction NADP⁺-Cys 298, superimposed with the 2Fo–Fc [σA-weighted, violet contours at 3.11 e/Å³ (3.0σ)] showing the double conformation of the cysteine residue. The insert shows the σA-weighted Fo–Fc map with omitted hydrogen atoms contoured magenta at 0.25 e/Å³ (2.5σ). The C4 atom of the nicotinamide is singly protonated in the plane of the ring, confirming that the coenzyme is NADP⁺.

seen. The "extra" hydration sites are mostly in contact with the nonpolar parts of the protein (either in C—H. . .O hydrogen bond with the protein, or in hydrophobic contact at a distance of about 4 Å from the C of the protein). Multiple tetrahedral arrangements are seen (Fig. 8), with an O—O distance around 2.8 Å. Deformation of tetramers often correlates with higher B-factors. The tetrahedral arrangements resemble local ice structure, revealing a very high degree of order of the solvent in those regions. Pentameric clusters of hydration sites are also observed (Fig. 8), both in hydrophobic and hydrophilic contacts with the protein. The O—O distance in these arrangements lies within the usual values for water–water interactions. Some of these pentamers fuse and interact in arrangements resembling the polyhedral two-dimensional networks observed in clathrates and in the surface of hydrophobic moieties observed at atomic resolution.[4] A water channel is seen connecting the active site with the external solvent. Two water molecules, one near His 110 and another near Cys 298, are well positioned to play a functional role in the catalytic mechanism. The best refined solvent regions provide some insight into solvent dynamics. The temperature factor of water molecules (mean value 21 Å²) were refined anisotropically. For water molecules close to the protein, the major axis of the B-factor ellipsoid tends to be parallel to the protein surface [Fig. 2(c)]. This agrees with previous information from a multicopy analysis on aldose reductase[39] that placed mul-

tiple water copies in the crevices of the surface and with a molecular dynamics simulation study,[40] which has shown that water molecules around proteins and DNA have a higher diffusion rate parallel to the solute surface than perpendicular to it.

## Details of Inhibitor Binding

Refinement of the ternary complex AR–NADP⁺–IDD 594 at subatomic resolution revealed a number of fine details at the active site region that helped deciphering mechanisms of inhibitor potency and selectivity. They are explained by protonation and polarization states and by short contacts. The inhibitor IDD 594 occupies part of the catalytic site cleft and the specificity pocket [Figs. 1(a), 1(b), and Table IV). The interactions in the catalytic site are mostly polar. The carboxylate group is firmly anchored in the active site, with hydrogen bonds to His 110, Tyr 48, and Trp 111, and with a strong electrostatic interaction with NADP⁺ (Figs. 1, 9, and Table IV).

The orientation of His 110 is unambiguously determined by the difference between the electron density peaks of the atoms CE1 (8.3 e/Å³) and NE2 (10.4 e/Å³). The Fo–Fc difference electron density map [Fig. 9(c)][15] shows that His 110 is protonated at the NE2 position. The W1 water molecule donates the hydrogen atom involved in the ND1–W1 H-bond. Furthermore, the C-NE2 bond length [1.344(3) Å] is significantly longer than the C-ND1 [1.325(3) Å]. All these observations indicate that His 110 is singly
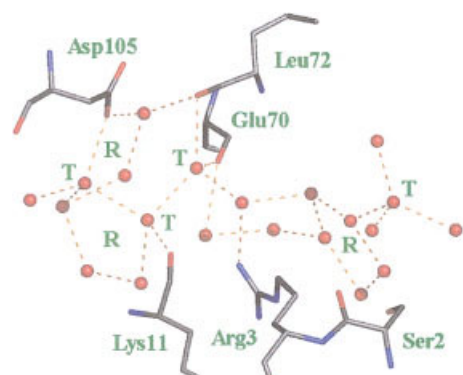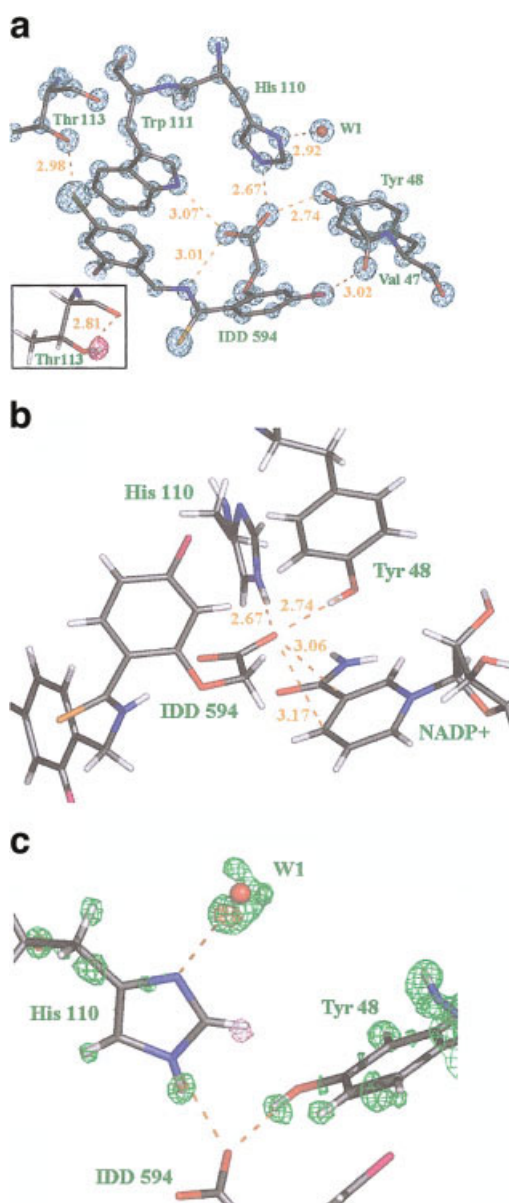
Fig. 8. Part of one water network linking two neighboring protein molecules. The tetrahedral water molecules are indicated by T, and the five member rings are indicated by R.

Fig. 10. Comparison of IDD 594 with other inhibitors. (**a**) Superposition (DD 594 (in yellow) and Statil. (**b**) Superposition IDD 594 (in yellow) and IDD 384

protonated at NE2, which is the proton donor in the H-bond with the carboxylate oxygen O33 of the inhibitor. The hydrogen of the Tyr 48 hydroxyl group is also a proton donor in another H-bond to this oxygen [Fig. 9(c) and Table IV]. The second carboxylate oxygen of IDD 594, O34, is at H-bond distance from Trp 111 NE1 and IDD 594 nitrogen [Table IV and Fig. 9(a)]. Even though no clear H-atom density appears in these contacts, both nitrogen atoms are *a priori* protonated for chemical considerations. Consequently, the second oxygen is also an acceptor of two hydrogen bonds, implying that the carboxylate is fully charged. Indeed, the carboxylate's C—O bond lengths
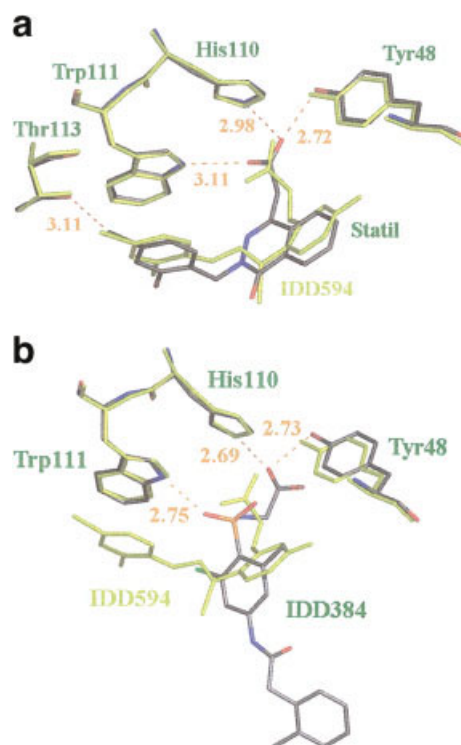
Fig. 9. Inhibitor contact details. (**a**) $\sigma$A weighted 2Fo–Fc maps, contoured blue at 4.15 e/Å$^3$ (4$\sigma$) together with color-coded atoms [same code as Fig. 2(b)]. The respective peak volumes correlate well with the atomic number. Insert: Close view of Thr 113, superimposed with the Fo–Fc (Omit-H, $\sigma$A-weighted, magenta contours at 0.25 e/Å$^3$ (2.5$\sigma$) showing the position of the hydrogen atom. (**b**) View of the interactions of the IDD 594 carboxylate, including the H-bonds with His 110 and Tyr 48 and the contact with NADP$^+$ [note that distances shown are between heavy (non-H) atoms]. (**c**) $\sigma$A-weighted Fo–Fc map of the active site region with omitted hydrogen atoms, contoured gold at 0.44 e/Å$^3$ (4.0$\sigma$), green at 0.31 e/Å$^3$ (2.8$\sigma$), and pink at 0.11 e/Å$^3$ (1.0$\sigma$). The single protonation of His 110 at NE2 is clearly visible, while the strong peak near the W1 water molecule indicates that W1 is the proton donor in the W1–His110 ND1 hydrogen bond. The hydrogen atoms (white) are observed in the maps, as interpreted by SHELXL. The 1.0$\sigma$ contour is restricted to the neighborhood of the H atom.

**TABLE IV. Inhibitor–Protein and Inhibitor–NADP⁺ Contacts within 3.5 Å**

| IDD594 atom | Holoenzyme atom | Residue or coenzyme | Distance (Å) |
|---|---|---|---|
| C2 | NE1 | Trp20 | 3.138 (4) |
| Br8 | OG1 | *Thr113* | 2.973 (4) |
| F9 | O | Val47 | 3.010 (4) |
| F14 | CH2 | *Trp111* | 3.225 (5) |
| | N | Ala299 | 3.263 (5) |
| | CA | | 3.059 (6) |
| | C | | 2.993 (6) |
| | O | | 3.467 (5) |
| | N | *Leu300* | 3.263 (5) |
| O15 | CZ2 | *Trp20* | 3.312 (5) |
| | CE2 | | 3.448 (4) |
| C24 | CZ2 | *Trp111* | 3.330 (5) |
| C25 | CD2 | *Trp111* | 3.423 (5) |
| C27 | CH2 | *Trp111* | 3.272 (5) |
| | CZ2 | | 3.337 (5) |
| C28 | CZ3 | *Trp111* | 3.361 (5) |
| | CE3 | | 3.431 (4) |
| C32 | NE2 | His110 | 3.373 (4) |
| | NC4 | NADP⁺ | 3.434 (4) |
| O33 | CE1 | Tyr 48 | 3.236 (4) |
| | CZ | | 3.401 (4) |
| | OH | | **2.733 (4)** |
| | CE1 | His110 | 3.383 (5) |
| | NE2 | | **2.658 (4)** |
| | NC3 | NADP⁺ | 3.054 (4) |
| | NC4 | | 3.159 (4) |
| | NC7 | | 3.232 (5) |
| | NO7 | | 3.168 (4) |
| O34 | NE1 | Trp111 | **3.058 (4)** |
| N17 | O15 | IDD594 | 2.720 (4) |
| | O34 | IDD594 | 3.002 (4) |

The hydrogen bonds are indicated in **bold.** The interactions in the specificity pocket are shown in *italic.*

[1.239(3) and 1.262(4) Å] indicate a deprotonated state.[41] This is also consistent with the fact that the pKa value of the IDD 594 inhibitor measured in solution is 2.90 ± 0.05 (see Methods) while the pH of the crystallization medium is 5.0.

The coenzyme is in the charged state NADP⁺, as shown by the planarity of the ring (RMSD deviation from plane = 0.0204 Å) and the presence of the hydrogen atom of C4 in the plane of the nicotinamide [Fig. 7(d), insert]. Thus, the negatively charged carboxylate of IDD 594 occupies the "anionic site"[34] defined by the Tyr 48, His 110, and the positively charged nicotinamide of NADP⁺ [Fig. 9(b)]. In addition, the network of polar residues surrounding the NADP⁺ (Fig. 3) helps to compensate the negative charge of the inhibitor carboxylate through local dipoles. Such a compensation also takes place in the neighborhood of one of the coenzyme phosphate oxygens O21, which is surrounded by backbone nitrogens of Leu 212 [2.817(4) Å] and Ser 214 [3.001(4) Å], while the closest positively charged atom (NZ of Lys 21) to either of the charged oxygens of the inhibitor's carboxylate is at 5.9 Å distance. Such charge compensation by dipoles has already been observed, for example, in the case of the sulphate-binding protein of

*Salmonella typhimurium.*[42] Observed interactions contribute to the potency of the carboxylic acid inhibitors of AR.

IDD 594 induces a conformational change upon binding to AR, creating a specificity pocket localized between Phe 122, Trp 111, Leu 300, and Ala 299 [Fig. 1(a)]. The interactions of IDD 594 in this pocket include the stacking between one of the inhibitor rings and Trp 111 [Fig. 9(a) and Table IV] and an unusually short [2.973(4) Å] contact between the IDD 594-Br and Thr 113-OG [Fig. 9(a)]. This Br-OG distance is closer to the one reported between bromine and carbonyl (~3.00 Å) than to that between bromine and alcohol, when the former is attached to an aromatic ring (minimum reported: 3.17 Å, Cambridge Data Bank[43]). The Br-OG close interaction is facilitated in our structure by the displacement of the hydrogen atom of the —OH group of Thr 113. This hydrogen atom, instead of lying on the Br-OG axis, is involved in an H-bond [2.798(4) Å] with the carbonyl of Thr 113 [Fig. 9(a), insert]. This H-bond enhances the negative partial charge of Thr 113-OG, while the polarization of the bromine atom[36] causes a positive partial charge in the direction of the Br—C bond of the inhbitor. A similar relative oriention of bromine and OH, with the OH bond nearly perpendicular to the O—Br axis, is observed in water OH groups hydrating halogen anions.[44] This observed short contact leads to selectivity against aldehyde reductase. Indeed, in the latter, Thr 113 is substituted by a much bulkier tyrosine, which would cause steric hindrance.[45]

The most important differences in the binding modes of different potent and selective inhibitors lie in the opening of, and binding to, the specificity pocket, as observed for IDD 594, statil, tolrestat, and zopolrestat. All these carboxylic acid inhibitors have a bulky halogenated moiety (Trifluoromethyl for tolrestat and zopolrestat, Br for IDD 594 and statil). Statil binding is overall very similar to IDD 594,[46] as shown in Figure 10(a). The Br-Thr 113 distance is longer (3.11 Å), but this could be an artifact of refinement restraints at lower resolution (2.2 Å). Zopolrestat binds in the same orientation, while in the case of tolrestat, the opening of the specificity pocket is perpendicular to that of IDD 594. With differences in overall binding details, all these inhibitors have a common feature, the contact with Thr 113. The interactions of the carboxylate head of IDD 594 are also similar to those observed in the complexes AR–tolrestat,[17] AR–zopolrestat,[19] and AR–statil.[46] Another inhibitor, IDD 384, which lacks halogenated moiety, has a different conformation of the carboxylate head [Fig. 10(b)], forced by the contact of its sulfonyl oxygen with Trp 111.[47] Although this inhibitor remains potent, its selectivity is considerably decreased. This is consistent with the fact that this inhibitor does not interact with Thr 113 because it has a different conformation and does not open the specificity pocket.

## CONCLUSIONS

Carboxylic acid inhibitors of AR containing the bromine atom are the result of screening for soluble compounds that bind to the enzyme strongly and selectively (see Methods). A number of factors contribute to the potency of

IDD 594, among which hydrophobic interactions of the inhibitor ring system, stacking with Trp 111, and numerous electrostatic interactions of the inhibitor with the protein and the cofactor all stand out as critical to the inhibitor behavior.

The negatively charged carboxylate head occupies the "anion well" defined by $NADP^+$, His 110, and Tyr 48.[34] It interacts with the protein through a direct contact with the positively charged $NADP^+$ moiety [distance IDD 594-O33–$NADP^+$–NC3: 3.067 Å; Fig. 9(b) and Table IV], reinforced by hydrogen bonds between Tyr 48 and the neutral His 110 and the same carboxylate O33-atom [Fig. 9(a) and Table IV]. Previous observations by site-directed mutagenesis show that binding of carboxylic acid inhibitors depends on Tyr 48 more than on His 110,[31] in addition to the general rule that OH—O hydrogen bonds are stronger than NH—O ones;[44] in our complex the H-bond between the inhibitor carboxylate and the Tyr 48 OH is further enhanced by the nearby Lys 77, as the positively charged Lys 77-NZ may polarize the hydroxyl group of Tyr 48 [distance NZ-O: 3.114(4) Å].[25] Note that the orientation of the Lys 77 side chain, necessary for this polarization effect, is determined by the unusual geometry of the peptide bond 76–77 [ω = −167.4(2), Fig. 5(b)]. Furthermore, the presence of a neutral histidine is consistent with the fact that negatively charged inhibitors interact with the AR–$NADP^+$ complex and not with AR–NADPH, explaining the noncompetitive character of IDD 594. Because the orientation of the carboxylate head is similar to the one observed in statil and tolrestat inhibitors [Fig. 10(a)], the same protonation states observed in the structure with IDD 594 can be assumed. The differences observed with IDD 384 [Fig. 10(b)] reinforces the hypothesis, already suggested[47] that the protonation state of the carboxylate head of IDD 384 might be different.

The electrostatic interaction between the Br atom of the inhibitor and the OG of Thr 113 has an unusually short distance of 2.973(4) Å. The short contact between Br and Thr 113 OG explains the selectivity of IDD 594 towards AR, because in aldehyde reductase the Thr residue is replaced by Tyr, leading to steric hindrance. Indeed, IDD 594 binds to AR 1100 times better than to aldehyde reductase. The IDD 594-Br/Thr 113-OG interaction also contributes to the potency of the inhibitor. Other halogens, such as chlorine, cannot engage in a similar interaction (due to its lower polarizability) and, consequently, increase the $IC_{50}$ value by a factor of 2.3 (Van Zandt et al., paper in preparation).

The fine details of interactions described above explain how IDD 594 can reach potency and specificity levels similar to those of other bulkier inhibitors (e.g., zopolrestat) with a much smaller scaffold. Although hydrophobic interactions may play a role in inhibitor binding, keeping these hydrophobic groups to a minimum in IDD 594 improves its solubility and therefore its pharmacological applicability.

## PERSPECTIVES

The role played by X-ray protein crystallography in drug design efforts is broadly recognized. Strength of the method lies in its ability to locate critical enzyme/drug interactions that can be exploited to simultaneously increase potency and specificity. These efforts have been successful at helping design drug scaffolds. However, these efforts have lacked, in general, the unambiguity required to determine the subtle details of the unusual interactions that differentiate a lead compound from a promising drug. Lack of resolution frequently results in the artificial standardization of geometries, driven by the use of standard restraints. Conversely, unusual geometries in lower resolution structures are frequently treated as defects because they cannot be trusted. It is the ability to characterize these unusual interactions that makes ultrahigh-resolution structural studies inherently powerful. The need for this extreme resolution lies partly in the nature of some of the more subtle interactions, frequently dominated by polarization or other low-strength interactions that are not properly accounted for by the restraint schemes used in protein refinement protocols at lower resolutions.

However, it is possible to ask if, beyond the usefulness of ultrahigh-resolution studies, the structure here presented, together with the other few available at similar resolution, are just anomalies or if similar studies can be carried out in more general cases. An example of a protein structure solved at higher resolution than AR is that of crambin, with a nominal resolution very similar to that of AR. It has been suggested, in the original crambin paper, that the unusual high resolution of crambin crystals may have been the consequence of several factors including a very small unit cell, low mobility (associated to the presence of S—S bonds), strong packing contacts and a small solvent volume with the solvent almost perfectly packed. AR structure seems to violate every one of these conditions. AR has a much larger unit cell than that of crambin. AR does not have S—S bonds and the AR complex include three noncovalently linked moieties (AR, $NADP^+$, and IDD 594). The total solvent volume in AR crystals is much larger than that in crambin (AR solvent volume: 46,500 Å$^3$, solvent content 34.6%; crambin: 5,000 Å$^3$, solvent content 30.0%). Moreover, AR crystal packing does not reveal all water molecules nor does it show unusually tight protein–protein interactions. Nonetheless, the diffraction quality of AR crystals is nearly identical to that of crambin. Thus, it is worth asking why AR can diffract at nearly the same resolution as crambin. Part of the explanation may have to do with the lack of S—S bridges. Although S—S bridges can make the protein scaffold more rigid, S—S bridges are uncommonly sensitive to radiation damage.[48] Nonetheless, the conclusion we can draw from the availability of the AR structure at ultrahigh resolution is that it should be possible to obtain subatomic resolution quality diffracting crystals even in those cases that do not satisfy the requirements expressed in the crambin paper, greatly expanding the scope of application of subatomic protein crystallography. In particular, other AR–inhibitor complexes are currently being studied at subatomic resolution (papers in preparation). Careful protein crystallogenesis[1,49] and the availability of brilliant synchrotron radiation sources seem to be the true limiting conditions to

achieve subatomic resolution, both of which are within reach of modern protein crystallography.

## METHODS

The pKa of IDD 594 inhibitor was determined potentiometrically with a pH-sensitive electrode using the PCA 101 device from Sirius Analytical Instruments, England, in mixtures of water and DMSO; the Yasuda-Shedlovsky extrapolation was used to calculate the aqueous pKa value of IDD 594 ($2.90 \pm 0.005$).[50]

Human AR was expressed in *Escherichia coli* and cocrystallized with the oxidized form of the coenzyme $\beta$-NADP$^+$ and the inhibitor IDD 594 at pH 5.0 and at 277 K as described earlier.[51] The crystals belong to space group P2$_1$ with unit cell dimensions (at 100 K) a = 49.43 Å, b = 66.79 Å, c = 47.40 Å, and $\beta$ = 92.4°, with one complex per A.U. and a solvent content of 34.6%.

The crystal diffraction quality was significantly improved by cocrystallization with the cofactor and inhibitor, by microseeding and by optimization of the cryofreezing protocol.[49] The cocrystallization with IDD 594 was carried out at room temperature (ratios protein/coenzyme/inhibitor = 1/2/2). Previously equilibrated (against PEG 6000; 15%) hanging drops were seeded with stock seed solutions diluted 100 times. Similar experiments at 4°C or at lower dilution (1/10) gave crystals of smaller size that grew as clusters, from which monocrystals were difficult to retrieve. Cryofreezing was carried out through quick transfers into a stabilization solution (25% PEG 6000), then into a cryo-protecting solution (40% PEG 6000) and finally into either liquid nitrogen or ethane.

Diffraction experiments were conducted on two crystals of approximate size $0.3 \times 0.4 \times 0.6$ mm$^3$. The X-ray beam of the undulator line 19ID of SBC-CAT at the APS was very stable, highly collimated, and could be tailored to the crystal size. The relatively large ($210 \times 210$ mm$^2$) and efficient $3 \times 3$ tiled CCD area detector[52] with a very fast duty-cycle could be easily manipulated to adapt an optimal geometry of the experiment, required by the extremely high resolution. To avoid inclusion of poor-quality data the radiation damage of crystals was monitored by the stability of mosaicity (<0.3°), I/$\sigma$(I), $R_{merge}$ and B-factor values as a function of frame number. To overcome the limits in the sensitive area of the detector and in its dynamic range, large $2\theta$ angles were used. Because the detector dynamic range was less than that of the sample crystals, and because large $2\theta$ angles cause loss of redundancy and completeness, data from two crystals were measured in several scans with different X-ray doses. A total of 11 scans in different sample orientations and in four different resolution shells allowed the accumulation of sufficient completeness and redundancy to ensure accurate measurements (Podjarny et al., in press; Sanishvili et al., in preparation). The X-ray wavelength (0.653 Å), the flux ($4 \times 10^{11}$–$1 \times 10^{10}$ X-ray photons/s), the exposure time (1–15 s), the oscillation range (0.2–1°), the $2\theta$ angle (0–30°) and the crystal-to-detector distance (130–200 mm) were adjusted to optimize the data quality. Data from the two isomorphous crystals were reduced and scaled with HKL2000.[53] Relevant statistics of diffraction data are given in Table I. The outmost shell (0.66–0.62 Å) was discarded because of the low I/$\sigma$(I) (=2.1), high $R$(I)$_{merge}$ (>30%) and low completeness (<30%) values.

The scan limits were set as shown in Table II (more than two scans correspond to different crystal orientations). The redundancy of each scan ($0 < \omega < 180°$) in the P2$_1$ space group is 1.65. The final statistics as a function of resolution range are given in Table III. $R_{merge}$ compares reflections from different scans.

The structure was initially refined with CNS,[54] starting with the coordinates of the complex of human AR with another inhibitor, IDD 384, solved at 1.7 Å, using data collected on the laboratory source.[47] The starting refinement was isotropic, followed by an anisotropic refinement and an introduction of H-atoms, as described in Table I. The refinement with SHELXL[55] using 33,158 parameters and 46,316 restraints against 511,265 structure factor amplitude (corresponding to an observation/parameters ratio of 15.4) led to a final model comprising 313 out of the 316 residues, NADP$^+$, IDD 594, two citrate molecules, and 613 water molecules. Multiple conformations were observed for 99 residues; 198 water molecules have less than unit-occupancy. For the final model, the part of the full-matrix corresponding to the coordinates of the protein, cofactor, and inhibitor atoms was inverted. This allowed to calculate, using the same methods as in small molecule crystallography, standard uncertainties for coordinates and properties derived from them, such as bond lengths. At this resolution the deviation from ideal bond length values indicated in table I (0.016 Å) corresponds mostly to real differences, as the typical standard uncertainty for bond lengths derived as described above is on the order of 0.003 Å.

To obtain an independent measure of a typical standard uncertainty with a minimum of systematic errors, we calculated the standard deviation for a set of 174 C—CA single bond lengths restricted to well-ordered atoms (B-factor of CA < 4 Å$^2$). The resulting value of 0.006 Å is in qualitative agreement with the standard uncertainties estimated from the matrix inversion. The final model statistics are given in Table I.

Statistical estimates of the standard deviation of the electron density were calculated using the CCP4 suite of programs.[56] Maps were visualized and interpreted using the program O.[57] Absolute density levels were obtained with the program SHELXPRO.[55] Figures were prepared with the programs WEBLABVIEW (ACCELRYS), PYMOL[58] and ORTEP.[59] Contour levels were chosen to highlight the features shown, e.g. hydrogen atoms. Noise appears at lower contour levels in difference maps, notably near atomic positions. It is likely that a different model parameterization, that is, the multipolar one, will improve the difference maps for the ordered zones. A first multipolar refinement will be published elsewhere[36] (B. Guillot, Ph.D. thesis, Nancy University).

The atomic coordinates and structure factors of the AR–NADP$^+$–IDD 594 complex were deposited in the Protein Data Bank (1US0).

## ACKNOWLEDGMENTS

## REFERENCES

1. Podjarny A, Howard E, Mitschler A, Chevrier B, Lecompte C, Guillot B, Pichon-Pesme V, Jelsch C. X-ray crystallography at subatomic resolution. Europhys News 2002;33.
2. Schmidt A, Lamzin VS. Veni, Vidi, Vici— atomic resolution unraveling the mysteries of protein function. Curr Opin Struct Biol 2002;12:698–703.
3. Jelsch C, Teeter MM, Lamzin V, et al. Accurate protein crystallography at ultra-high resolution: valence electron distribution in crambin. Proc Natl Acad Sci USA 2000;97:3171–3176.
4. Teeter MM, Roe SM, Heo NH. Atomic resolution (0.83 A) crystal structure of the hydrophobic protein crambin at 130 K. J Mol Biol 1993;230:292–311.
5. Rypniewski WR, Oestergaard P, Noerregaard-Madsen M, Dauter M, Wilson KS. Fusarium oxysporum trypsin at atomic resolution at 100 and 283 K: a study of ligand binding. Acta Crystallogr D 2001:57:8–19.
6. Kuhn P, Knapp M, Soltis M, et al. The 0.78 Å structure of a serine protease: *Bacillus lentus* subtilisin. Biochemistry 1998;37:13446–13452.
7. Esposito L, Vitagliano L, Mazzarella L. Recent advances in atomic resolution protein crystallogrpahy. Protein Peptide Lett 2002;9:95–105.
8. Dauter Z, Lamzin VS, Wilson KS. The benefits of atomic resolution. Curr Opin Struct Biol 1997;7:681–688.
9. Oates PJ, Mylari BL. Aldose reductase inhibitors: therapeutic implications for diabetic complications. Expert Opin Investig Drugs 1999;8:2095–2119.
10. Yabe-Nishimura C. Aldose reductase in glucose toxicity: a potential target for the prevention of diabetic complications. Pharmacol Rev 1998;50:21–33.
11. Wermuth B, Münch JDB, von Wartburg JP. Stereospecificity of hydrogen transfert of aldehyde reductase. Experientia 1979;35:1288–1289.
12. Kubiseski TJ, Hyndman DJ, Morjana NA, et al. Studies on pig muscle aldose reductase. Kinetic mechanism and evidence for a slow conformational change upon coenzyme binding. J Biol Chem 1992;267:6510–6517.
13. Grimshaw CE, Shahbaz M, Putney CG. Mechanistic basis for nonlinear kinetics of aldehyde reduction catalyzed by aldose reductase. Biochemistry 1990;29:9947–9955.
14. Humber L. The medicinal chemistry of aldose reductase inhibitors. Prog Med Chem 1987;24:299–343.
15. Rondeau JM, Tete-Favier F, Podjarny A, et al. Novel NADPH-binding domain revealed by the crystal structure of aldose reductase. Nature 1992;355:469–472.
16. Tete-Favier F, Barth P, Mitschler A, et al. Aldose reductase from pig lens. Eur J Med Chem 1995;30S:589s–603s.
17. Urzhumtsev A, Tete-Favier F, Mitschler A, et al. A "specificity" pocket inferred from the crystal structures of the complexes of aldose reductase with the pharmaceutically important inhibitors tolrestat and sorbinil. Structure 1997;5:601–612.
18. Wilson DK, Bohren KM, Gabbay KH, et al. An unlikely sugar substrate site in the 1.65 A structure of the human aldose reductase holoenzyme implicated in diabetic complications. Science 1992;257:81–84.
19. Wilson DK, Tarle I, Petrash JM, et al. Refined 1.8 A structure of human aldose reductase complexed with the potent inhibitor zopolrestat. Proc Natl Acad Sci USA 1993;90:9847–9851.
20. Bohren KM, Grimshaw CE, Lai CJ, et al. Tyrosine-48 is the proton donor and histidine-110 directs substrate stereochemical selectivity in the reduction reaction of human aldose reductase: enzyme

21. kinetics and crystal structure of the Y48H mutant enzyme. Biochemistry 1994;33:2021–2032.
21. Schlegel BP, Jez JM, Penning TM. Mutagenesis of 3 alpha-hydroxysteroid dehydrogenase reveals a "push-pull" mechanism for proton transfer in aldo-keto reductases. Biochemistry 1998;37:3538–3548.
22. Schlegel BP, Ratnam K, Penning TM. Retention of NADPH-linked quinone reductase activity in an aldo-keto reductase following mutation of the catalytic tyrosine. Biochemistry 1998;37:11003–11011.
23. Tarle I, Borhani DW, Wilson DK, et al. Probing the active site of human aldose reductase. Site-directed mutagenesis of Asp-43, Tyr-48, Lys-77, and His-110. J Biol Chem 1993;268:25687–25693.
24. Lee YS, Chen Z, Kador PF. Molecular modeling studies of the binding modes of aldose reductase inhibitors at the active site of human aldose reductase. Bioorg Med Chem 1998;6:1811–1819.
25. Cachau R, Howard E, Barth P, et al. The subatomic resolution structure of human aldose reductase shows the catalytic mechanism. J Phys IV France 2000;10:3–13.
26. Singh SB, Malamas MS, Hohman TC, et al. Molecular modeling of the aldose reductase-inhibitor complex based on the X-ray crystal structure and studies with single-site-directed mutants. J Med Chem 2000;43:1062–1070.
27. Harrison DHL, Bohren KM, Petsko GA, et al. The alrestatin double-decker: binding of two inhibitor molecules to human aldose reductase reveals a new specificity determinant. Biochemistry 1997;36:16134–16140.
28. Rastelli G, Costantino L., Gamberini MC, Del Corso A, Mura U, Petrash JM, et al. Binding of 1-benzopyran-4-one derivatives to aldose reductase: a free energy perturbation study. Bioorg Med Chem 2002;10:1427–1436.
29. Miyamoto S. Recent advances in aldose reductase inhibitors: potential agents for the treatment of diabetic complications. Expert Opin Ther Patents 2002;12:621–631.
30. Costantino L, Rastelli G, Gamberini MG, Barlocco D. Pharmacological approaches to the treatment of diabetic complications. Expert Opin Ther Patents 2000;10:1245–1262.
31. Ehrig T, Bohren KM, Prendergast FG, et al. Mechanism of aldose reductase inhibition: binding of NADP+/NADPH and alrestatin-like inhibitors. Biochemistry 1994;33:7157–7165.
32. Kador PF, Lee YS, Rodriguez L, et al. Identification of an aldose reductase inhibitor site by affinity labeling. Bioorg Med Chem 1995;3:1313–1324.
33. Griffin BW, McNatt LG. Characterization of the reduction of 3-acetylpyridine adenine dinucleotide phosphate by benzyl alcohol xatalyzed by aldose reductase. Arch Biochem Biophys 1986;246:75–81.
34. Harrison DH, Bohren KM, Ringe D, Petsko GA, Gabbay KH. An anion binding site in human aldose reductase: mechanistic implications for thebinding of citrate, cacodylate, and glucose 6-phosphate. Biochemistry 1994;33:2011–2020.
35. Liu SQ, Bhatnagar A, Srivastava SK. Does sorbinil bind to the substrate binding site of aldose reductase? Biochem Pharmacol 1992;44:2427–2429.
36. Muzet N, Guillot B, Jelsch C, et al. Electrostatic complementarity in an aldose reductase complex from ultra-high-resolution crystallography and first-principles calculations. Proc Natl Acad Sci USA 2003;100:8742–8747.
37. Chakrabarti P, Chakrabarti S. C—H. . . O hydrogen bond involving proline residues in alpha-helices. J Mol Biol 1998;284:867–873.
38. Petrash JM, Harter TM, Devine CS, et al. Involvement of cysteine residues in catalysis and inhibition of human aldose reductase-site-directed mutagenesis of Cys-80, Cys-298, and Cys-303. J Biol Chem 1992;267:24833–24840.
39. Podjarny AD, Howard EI, Urzhumtsev A, et al. A multicopy modeling of the water distribution in macromolecular crystals. Proteins Struct Funct 1997;28:303–312.
40. Makarov VA, Feig M, Andrews BK, et al. Diffusion of solvent around biomolecular solutes: a molecular dynamics simulation study. Biophys J 1998;75:150–158.
41. Borthwick PW. Some relationships between bond lengths and angles in —COO—, COOH and —COOCH$_3$ groups. Acta Crystallogr 1980;B36:628–630.
42. Pflugrath J, Quiocho F. Sulphate sequestered in the sulphate-binding protein of *Salmonella typhimurium* is bound solely by hydrogen bonds. Nature 1985;314:257–260.

43. Cambridge crystallographic data base. Cambridge, UK: Cambridge Crystallographic Data Center; 2001.
44. Jeffrey GA, Saenger W. Hydrogen bonding in biological structures. Heidelberg: Springer-Verlag; 1991.
45. El-Kabbani O, Carper DA, McGowan MH, et al. Studies on the inhibitor-binding site of porcine aldehyde reductase: crystal structure of the holoenzyme-inhibitor ternary complex. Proteins Struct Funct Genet 1997;29:186–192.
46. El-Kabbani O, Ramsland P, Darmanin C, et al. Structure of human aldose reductase holoenzyme in complex with statil: an approach to structure-based inhibitor design of the enzyme. Proteins Struct Funct Genet 2003;50:230–238.
47. Calderone V, Chevrier B, Van Zandt M, et al. The structure of human aldose reductase bound to the inhibitor IDD384. Acta Crystallogr D 2000;56:536–540.
48. Weik M, Ravelli RB, Kryger G, et al. Specific chemical and structural damage to proteins produced by synchrotron radiation. Proc Natl Acad Sci USA 2000;97:623–628.
49. Howard E, Lamour V, Mitschler A, et al. Resolution improvement to 0.9A in crystals of human aldose reductase. Acta Crystallogr 1999;A55(Suppl):P09.OB.001.
50. Shedlovsky T. In: The behavior of carboxylic acids in mixed solvents. Pesce B, editor. Electrolytes. New York: Pergamon Press; 1962. p 146–151.

51. Lamour V, Barth P, Rogniaux H, et al. Production of crystals of human aldose reductase with very high resolution diffraction. Acta Crystallogr D 1999;D55:721–723.
52. Westbrook EM, Naday I. Charge-coupled device-based area detectors. Methods Enzymol 1997;276:233–268.
53. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
54. Brunger AT, et al. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr 1998;D54:905–921.
55. Sheldrick GM, Schneider TR. Shelxl—high-resolution refinement. Methods Enzymol 1997;277:319–343.
56. Collaborative Computational Project N. The CCP4 suite: programs for protein crystallography. Acta Crystallogr 1994;D50:760–763.
57. Jones TA, Cowan S, Zou J-Y, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Cryst 1991;A47:110–119.
58. DeLano WL. PyMOL: an open-source molecular graphics tool. Ccp4 Newslett Protein Crystallogr 2002;40:11.
59. Burnett MN, Johnson CK. ORTEP-III: Oak Ridge Thermal Ellipsoid Plot Program for crystal structure illustrations. Oak Ridge, TN: Oak Ridge National Laboratory; 1996.