Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Invariom refinement of a new monoclinic solvate of thiostrepton at 0.64 Å resolution

K. Pröpper,[a] J. J. Holstein,[a] C. B. Hübschle,[a] C. S. Bond[b] and B. Dittrich[a]*

[a]Institut für Anorganische Chemie, Georg-August-Universität, Tammannstrasse 4, D-37077 Göttingen, Germany, and [b]School of Biomedical, Biomolecular and Chemical Sciences, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

Correspondence e-mail: bdittri@gwdg.de

Received 4 January 2013
Accepted 18 April 2013

PDB Reference: thiostrepton, 4hp2

A new monoclinic solvate containing two molecules of the thiopeptide antibiotic thiostrepton in the asymmetric unit has been crystallized in space group $P2_1$. Single-crystal diffraction data to a resolution of 0.64 Å were collected at the SLS synchrotron, allowing structure solution by direct methods and resolution of the disorder present. Valence electron density can be observed in the Fourier residual density from refinement with the independent-atom model, which is a prerequisite for successful application of more sophisticated aspherical-atom scattering factors such as the invariom model when aiming to improve the structural model. Invariom refinement improves quality indicators such as $R_1(F)$ for thiostrepton, as previously demonstrated for small molecules. The nonspherical electron-density model also allows the direct derivation of a dipole moment and an electrostatic potential for the whole molecule, which is discussed in the context of antibiotic activity and molecular recognition.

1. Introduction

Macromolecular crystallography has been a major driving force for progress in modern structural biology, and this also holds for progress in methods development in single-crystal X-ray diffraction. While its achievements are undoubtedly impressive, single-crystal structure determination of macromolecules still faces several challenges, such as the further improvement of accuracy closer to the level usually seen in small-molecule crystallography. The precision (*i.e.* the standard uncertainty) with which a set of atomic positions can be determined from an experiment can be compromised by several factors. In an ideal situation with 'perfectly imperfect' single crystals, precision is principally dependent on experimental resolution. The crystals need to be perfect enough to scatter to high resolution, but less than ideal to avoid extinction. However, in the case of large-amplitude atomic motion even high resolution does not allow precise determination of all parts of a molecular structure. The example of thiostrepton with its known tetragonal and monoclinic structures is a fitting illustration of the problems that are encountered, since in both solvates the complete chain of the tail region could not be resolved (Anderson *et al.*, 1970; Bond *et al.*, 2001).

Thiostrepton, an oligopeptide that can be extracted from several *Streptomycetes* strains, *e.g. S. azureus*, has attracted attention for various reasons in recent decades. For a chemical diagram of the molecule, see Fig. 1. Its biological function as an antibiotic (Pagano *et al.*, 1956) has been shown to be caused by interaction with the 70S subunit of the bacterial ribosome (Walter *et al.*, 2012; Gonzalez *et al.*, 2007; Jonker *et al.*, 2007; Lee *et al.*, 2007). Anticancer properties in mammalian cancer cell lines have also been reported (Kwok *et al.*, 2008). The low

solubility of thiostrepton in water and its molecular size prohibit its clinical use despite its similar activity to penicillins towards Gram-positive bacteria (Bagley *et al.*, 2005). Current progress in understanding its biosynthesis (Arndt *et al.*, 2009), including total and active-fragment syntheses (Nicolaou, Zak, Rahimipour *et al.*, 2005; Nicolaou, Zak, Safina *et al.*, 2005; Nicolaou, 2012), allows chemical modification and might lead to a lower-weight pharmaceutical drug molecule derived from thiostrepton with better properties in the future.

The biomedical interest in thiostrepton provided a strong motivation to elucidate its solid-state structure. The aforementioned monoclinic structure in space group $C2$ was the first reported structure (Anderson *et al.*, 1970). Since deposition of structural data did not take place, the exact positional parameters of the $C2$ structure are not presently available. However, structural data for a tetragonal form of thiostrepton crystallized in space group $P4_32_12$ were made available by Bond *et al.* (2001) (PDB entry 1e9w). This latter structure determination confirmed the molecular connectivity and the overall shape of the main part of the molecule; structural differences between the $C2$ and the tetragonal form appear to occur mainly in the solvent and the tail region of the molecule, which demonstrates considerable conformational flexibility. In this region of the molecule, which consists mostly of planar dehydroalanine residues, large-amplitude motion and disorder were reported to occur in both known structures.

Here, we report an accurate structure determination of a new monoclinic form of thiostrepton. It was crystallized in space group $P2_1$ using vapour-diffusion methods with diethyl ether, which is incorporated into the structure. The tail region is resolved but is split into two components in each of the two independent thiostrepton molecules in the asymmetric unit. Apart from resolving disorder in this and other parts of the structure, our single-crystal diffraction data measured to 0.64 Å resolution allow structure solution by direct methods and modelling with nonspherical scattering factors. Based on the nonspherical electron density, we obtain an accurate molecular structure (corrected for asphericity shifts; Coppens *et al.*, 1969), derive the molecular dipole moment (Spackman, 1992) and report the molecular electrostatic potential (ESP). Possible mechanisms of molecular recognition with respect to antibiotic activity are discussed.

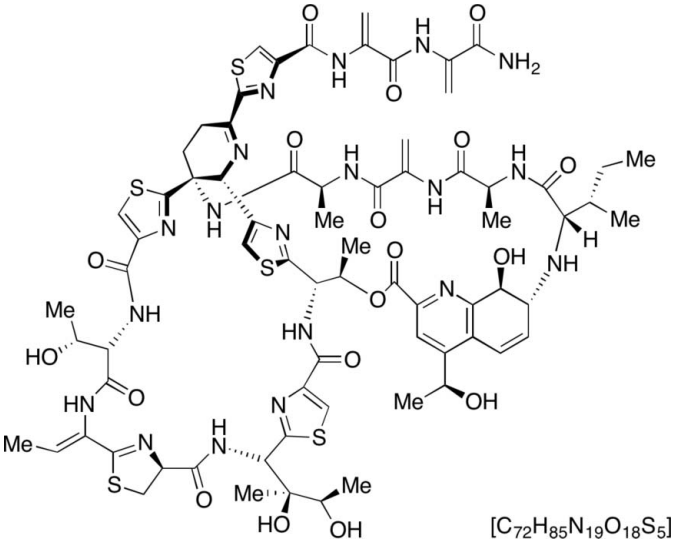

Figure 1
Chemical composition and structure of thiostrepton.

1.1. Aspherical-atom modelling of macromolecules

One way to improve the accuracy of structural parameters that is comparatively new to macromolecular crystallography is the use of a more sophisticated scattering model than the commonly used independent-atom model (IAM). Current approaches are either based on the Hansen and Coppens multipole model (Hansen & Coppens, 1978), which is a modification of the 'rigid pseudoatom' model devised by R. F. Stewart (Stewart, 1976, 1977), or use spherical scatterers to describe valence electron density (Hellner, 1977). The latter approach has seen recent use in modelling of bonding electron density in macromolecular structures (Afonine *et al.*, 2004, 2007), and the main advantage is the ease of software implementation. However, spherical scatterers require additional parameters[1] and their use is limited to structure refinement, whereas the electron density from the multipole model allows further analysis; it provides a more detailed description of electron density and can additionally be evaluated to describe molecular properties, as is exploited in this work. A fascinating alternative approach is restrained refinement with fixed aspherical Gaussian-type scattering factors that can take polarization into account (Schnieders *et al.*, 2009) and we look forward to seeing its future application.

It has been shown that for refinements at resolutions limited to $d \geq 0.5$ Å multipole parameters should be fixed and cannot be freely refined in conjunction with positional and displacement parameters owing to severe correlation (Stewart, 1976; Dittrich *et al.*, 2009). Such a comparatively high resolution has so far only been reached in macromolecular work for the 46-residue peptide crambin (Jelsch *et al.*, 2000; Schmidt *et al.*, 2011) and a Z-DNA hexamer duplex (Brzezinski *et al.*, 2011). Only in very special circumstances does free refinement of multipole parameters seem possible in a protein, even when a subset of atoms with low B factors is selected. Studies of macromolecular refinement at high resolution have been reviewed by Petrova & Podjarny (2004). Pioneering work on modelling protein data with experimentally derived aspherical scattering factors include studies of the enzyme aldose reductase (Guillot *et al.*, 2008), crambin (Jelsch *et al.*, 2000) and a snake toxin (Housset *et al.*, 2000). This work has recently been summarized by Lecomte *et al.* (2008). Theoretically derived scattering factors have so far found use in the refinement of two larger oligopeptide structures. These are the undecapeptide cyclosporin (Johnas *et al.*, 2009) and the dimeric 18-residue peptaibol antibiotic trichotoxin A50E

[1] To illustrate this statement, detailed modelling of an oxygen lone pair would require eight additional (two times x, y, z and occupancy) parameters. Free refinement of numerous additional parameters is usually not possible for protein data owing to parameter correlation.

Table 1
Crystal and structure-refinement data for thiostrepton in space group $P2_1$ (PDB entry 4hp2).

Data collection	
Beamline	PXII – X10SA
Temperature (K)	100
Wavelength (Å)	0.6500
Space group, Z	$P2_1$, 4
Crystal system	Monoclinic
Empirical formula	$C_{85.55}H_{19.55}N_{21.83}O_{23.10}S_5.H_2O$
Formula weight (g mol^{-1})	1983.78
Unit-cell parameters	
a (Å)	21.393 (1)
b (Å)	22.870 (4)
c (Å)	22.780 (6)
β (°)	106.451 (16)
V (Å^3)	10689 (4)
Calculated density (g cm^{-3})	1.226
$F(000)$	4202
Absorbance coefficient μ (mm^{-1})	0.182
Maximum 2θ (°)	61.88
Resolution limit (Å), $(\sin\theta/\lambda)_{max}$ (Å^{-1})	0.64, 0.769
Ranges of h, k, l before merging	$-28\rightarrow h\rightarrow 28$, $-35\rightarrow k\rightarrow 33$, $-32\rightarrow l\rightarrow 35$
Reflections, measured	160148
Reflections, unique	72903
Multiplicity† (%)	3.61 (1.87)
Completeness† (%)	91.0 (79.2)
R_{merge}‡ (%)	5.05
$R_{r.i.m.}$‡ (%)	5.67
Average $I/\sigma(I)$†	16.88 (6.8)
Refinement	see Table 2

† Values in parentheses are for the outer shell. ‡ $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$. $R_{r.i.m.}$ is the multiplicity-weighted R_{merge} (Diederichs & Karplus, 1997).

(Dittrich *et al.*, 2010). Both systems were selected as disorder was almost completely absent, and both refinements led to significant improvements in figures of merit such as, for example, $R_1(F)$, as typically observed in small-molecule refinements.

The current consensus for work relying on the Hansen–Coppens multipole model appears to be that databases of fixed scattering factors are best suited for modelling protein structures at 'ultrahigh' resolution ($d \leq 0.85$ Å).[2] Three such databases exist. Scattering factors are obtained either experimentally by the averaging of multipole parameters from small-molecule crystallography (Zarychta *et al.*, 2007; Domagała *et al.*, 2012) or from theoretical calculations. Here, two approaches compete: scattering factors are derived either from averaging of (now theoretical) fragments from single-point calculations that are similar within a standard deviation after averaging (Volkov *et al.*, 2004; Dominiak *et al.*, 2007; Jarzembska & Dominiak, 2012) or are obtained from geometry-optimized model compounds that reproduce the local chemical environment of an atom (Dittrich *et al.*, 2004; Dittrich, Hübschle *et al.*, 2006) in the invariom approach. The latter database has been chosen for this work, since the coverage of chemical environments required for thiostrepton can currently only be provided in this database.

Table 2
Quality indicators for the two refinement models, independent-atom refinement (IAM) and invariom refinement (INV), of the $P2_1$ form of thiostrepton.

Program/model	*XD*/IAM	*XD*/INV
Resolution (Å)	0.64	0.64
No. of parameters	1486	1486
No. of reflections	62211	62211
$R_1(F)$, $I > 3\sigma(I)$† (%)	7.32	6.77
$R_w(F^2)$, $I > 3\sigma(I)$‡ (%)	21.20	20.01
$R_{all}(F)$§ (%)	8.08	7.55
GoF¶	1.79	1.68
Flack parameter	0.05 (4)	0.05 (3)
Minimum r.d.†† (e Å^{-3})	−0.580	−0.654
Maximum r.d. (e Å^{-3})	0.995	1.038

† $R_1(F) = \sum ||F_o| - |F_c||/\sum |F_o|$. ‡ $R_w(F^2) = [\sum w||F_o| - |F_c||^2/\sum w|F_o|^2]^{1/2}$, $w = 1/\sigma^2(I)$. § $R_{all}(F)$ was calculated using all reflections. ¶ GoF = $[(\sum ||F_o| - |F_c||^2)/(n_o - m_{var})]^{1/2}$. †† r.d. is the residual electron density.

The use of nonspherical scattering factors in macromolecular crystallography seems to be currently limited to a few selected 'well behaved' structures. Although improvements in accuracy might not manifest themselves in disordered structures or in structures with a resolution that does not fulfil the requirement of $d \geq 0.85$ Å, there is an increasing number of ultrahigh-resolution structures that contain unmodelled additional information. Scattering-factor databases could provide a helpful methodology in such cases. To illustrate the potential and current limitations of database approaches in the refinement of macromolecular structures, we present the results for a new monoclinic dimethyl ether solvate of the sulfur-rich hexadecapeptide thiostrepton.

2. Crystallization and data collection

The two known monoclinic and tetragonal structures of thiostrepton are unsuitable for obtaining accurate coordinates of all atoms owing to the aforementioned large-amplitude motion and disorder. The discovery of a new solvate structure changed the situation. Crystals were grown from wet dimethylformamide by vapour-diffusion methods with diethyl ether as an antisolvent and diffracted to high resolution. A favourable development for high-resolution crystallography in general is the advent of a new generation of area detectors with lower background noise,[3] fast read-out and a higher dynamic range compared with CCD area detectors. For rapid (the experiment lasted less than 15 min) and shutterless data collection (Broennimann *et al.*, 2006) of the new $P2_1$ solvate of thiostrepton, the new generation of DECTRIS PILATUS detectors (the 6M pixel detector) was used in combination with comparably hard X-rays of wavelength 0.6500 Å. The detector distance was 165 mm. The programs *XDS* (Kabsch, 2010) and *SADABS* (Sheldrick, 2007) were used for intensity integration and scaling, respectively. Although a correction for absorption is already part of the functionality of *XDS*, scaling for the two

[2] 'Ultrahigh' or better subatomic resolution in macromolecular refinement has been defined as a resolution better than 0.85 Å (Petrova & Podjarny, 2004). A useful definition of data quality relying on various indicators has been described by Urzhumtseva *et al.* (2009).

[3] The spot profiles do not change substantially at different detector distances, but the background area increases with the square of the distance. A comparatively large detector further away from the sample hence increases the signal-to-noise ratio.

independent runs with different exposure times and intensity filters was further improved with *SADABS*. Diffraction data were collected on beamline X10SA at the Swiss Light Source (SLS) synchrotron to a resolution of 0.64 Å. Data were cut at this resolution on the basis of the quality indicators the average $I/\sigma(I)$ ratio and $R_{\rm int}$ in the highest resolution shell. Owing to the measurement setup (φ-scans only) full data coverage could not be attained for the highest resolution shell. Full crystallographic details are given in Table 1.

3. Structure refinement

Initial high-resolution full-matrix refinement was carried out with the program *SHELXL* (Sheldrick, 2008) after structure solution with *SHELXD* (Schneider & Sheldrick, 2002). Within *SHELXL*, restraints for atomic displacement parameters and constraints for riding H atoms were used extensively and 3441 least-squares parameters were refined, giving an $R_1(F)$ of 7.3% in a refinement on F^2 using all reflections. Occupancies of split positions for solvent atoms and disordered regions of the two main molecules in the asymmetric unit were determined with free variables and subdivided into PARTs. PART instructions are a constituent of the command set in *SHELXL* and allow disordered atoms to be divided into two (or more) groups. Thus, each group represents one component of the disorder. The graphical user interface *SHELXLE* (Hübschle *et al.*, 2011) facilitated these tasks. A depiction of the molecule with anisotropic displacement parameters illustrating the disorder of the tail region is given in Fig. 2.

Since aspherical-atom refinement is not a feature available in *SHELXL*, the program *XD* was employed for subsequent full-matrix aspherical-atom refinement. The *XD* package (Koritsánszky *et al.*, 2003) was not coded for refinement of macromolecular structures, but rather for modelling non-disordered small-molecule structures at high resolution. Because of our own experience and software development connected to the *XD* package, we chose a reparameterized version of the least-squares refinement program *XDLSM* for refinement. This locally modified version of the 2003 release of *XD* allows a larger number of atoms (including dummy atoms), parameters and constraints as required, for example, for riding H-atom treatment of macromolecular structures. It also allows refinement of the Flack parameter (Flack, 1983) for absolute structure determination (see §4.3). Input files for *XD* were generated with the preprocessor program *Invariom-Tool* (Hübschle *et al.*, 2007).

An alternative to the *XD* suite of programs is the aspherical-atom refinement program *MoPro* (Guillot *et al.*, 2001; Jelsch *et al.*, 2005), which has seen more development for macromolecular refinement purposes than *XD*. This point becomes obvious for restraints, which have not been implemented in *XD*. However, *XD* allows treatment of atomic displacements as a rigid group. This functionality was successfully used for those disordered residues that are part of the main molecular core. Hence, *MoPro* was not used for refinement in this work.

Throughout full-matrix least-squares refinement with *XD* it was helpful to check progress with a graphical display of the residual electron density, as is well established in macromolecular crystallography and implemented in programs such as *Coot* (Emsley *et al.*, 2010). Such functionality is now also available in the program *MoleCoolQt* (Hübschle & Dittrich, 2011). Moreover, *MoleCoolQt* can display the local atomic coordinate systems that are required for aspherical-atom refinements. Coordinate systems need to be set correctly to orient multipole populations in order to successfully model valence electron density.

The multipole parameters used in the refinement were fixed to the values taken from a new generalized version of the invariom database (Dittrich, Hübschle *et al.*, 2013), thereby taking into account local bonding and lone-pair electron density. Both ordered and disordered atoms were described this way.[4] The level of the multipolar expansion was $l = 4$ for all atoms including hydrogen. For atoms heavier than carbon, κ' parameters were refined in the database. The occupancies of split sites obtained from *SHELXL* were preserved. Using calculated H-atom positions from *SHELXL*, bond distances to H atoms were set (with the RESET BOND command) to the values contained in the invariom database and were subsequently constrained to ride on their parent atom sites. Values of the X—H bond distances in the database were taken from geometry optimizations with the method/basis set combination B3LYP/D95++(3df,3pd). Isotropic displacement parameters of riding H atoms were constrained to a factor of 1.2 ($C_{\rm arom}$, CH, NH and CH_2) or 1.5 (CH_3 and H_2O) of those of the respective parent atoms. Positional and displacement parameters of disordered atoms of the solvent and the side

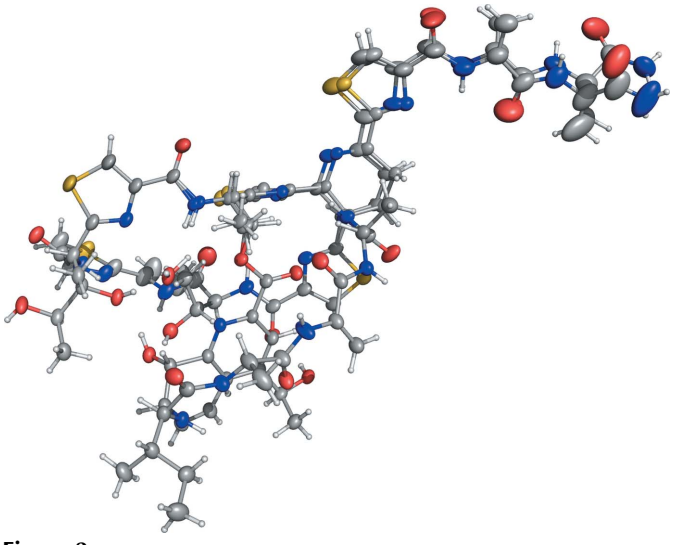

Figure 2
Structure of one of the thiostrepton molecules in the asymmetric unit, plotted with H atoms and anisotropic displacement parameters at 50% probability. The structure contains resolved and refined split positions owing to disorder which involves the bis-dehydroalanine tail and parts of the thiazoline macrocycle residues.

[4] Only asphericity of the major component was considered for the particular atom where disorder starts and that is connected to split sites but which is not disordered itself.

chain were not refined in *XDLSM*. In particular, anisotropic atomic displacement parameters (ADPs) of the disordered dehydroalanine-rich side chain were constrained to the RIGU restrained result (Thorn *et al.*, 2012) from *SHELXL* in order to prevent them from converging to physically implausible values. The RIGU restraint, as available in the 2013 version of *SHELXL*, is an extension of DELU restraints in the bonding direction that are based on the Hirshfeld test (Hirshfeld, 1976). RIGU allows three out of the six possible anisotropic atomic displacement parameters per atom to be restrained. It requires their calculation in an orthogonal *xyz* axis system and allows a physically reasonable behaviour of the relative motion of the atoms to be imposed not only in the bonding direction (*zz*) but also in the *xz* and *yz* directions, *i.e.* perpendicular to it.

Despite the limitations encountered in the free refinement of all positional and displacement parameters in *XDLSM*, the sum of all pseudoatom scattering factors allowed the reconstruction of the total molecular electron density based on all atoms. A full list of model compounds that were used to derive the scattering factors for the chemical environments present in thiostrepton is given in the Supplementary Material.[5]

4. Results and discussion

4.1. Quality indicators

It is expected that figures of merit (*i.e.* *R* factors, goodness of fit, values of the maximum negative/positive residual electron density) should improve when nonspherical scattering factors are used.[6] This has been shown numerous times for small-molecule structures. However, for disordered samples the improvement has been found to be less substantial (Dittrich, Hübschle *et al.*, 2006). Similar observations are made for the current structure: the $R_1(F)$ can be improved by 0.54%. With respect to the considerable manual work involved, this improvement might seem to be rather modest, since for nondisordered small molecules improvements of more than 1% are common (Dittrich, Hübschle *et al.*, 2013). However, considering the core scattering of the S atoms the improvement seen is consistent with the improvements seen for small-molecule structures. Quality indicators are given in Table 2.

Since the weighting scheme used is slightly different and no correction for disordered solvent ('SWAT') was performed in *XD* refinements, the values obtained from *SHELXL* and *XD* IAM refinement are not directly comparable. Therefore, the *SHELXL* results are not given. However, the figures of merit obtained were very similar to those from *XD* IAM refinement. The GoF is notably different in *SHELXL* compared with *XD*; it is calculated for modified weights using $w = 1/[(\sigma^2 F_o^2) + (aP)^2]$ with $P = [2F_c^2 + \max(F_o^2, 0)]/3$ in *SHELXL*, which gives a value close to one, whereas in *XD* refinements a classical weighting based on the counting statistics with $1/\sigma^2$ was used. The cutoff in *XD* for refinement was $I > 3\sigma(I)$. This rather high cutoff was chosen since σ seems to have been underestimated. This can be deduced from the *XD* GoF of 1.68 from invariom refinement, which is considerably larger than estimated by assuming an agreement of F_o^2 and F_c^2 in the range of σ^2.

4.2. Residual electron density

Fig. 3 shows residual bonding electron density that is present in the quinaldic acid residue in IAM (top) but that disappears after invariom refinement, as seen in the lower part of the figure. It was generated with the program *MoleCoolQt* (Hübschle & Dittrich, 2011) with model phases and represents an $F_o - F_c$ map. The strongest improvement in the signal when comparing the IAM and invariom models was obtained with a cutoff value of 0.23 e Å^{-3}. While for ordered parts of the structure the residual electron density is systematically reduced, it can even increase for disordered parts, as the values given in Table 2 show. Disorder is emphasized since it is the main remaining part of the observed structure that is in disagreement with the structural model. Such increased residual electron density is usually localized to a particular disordered region or to unmodelled features.

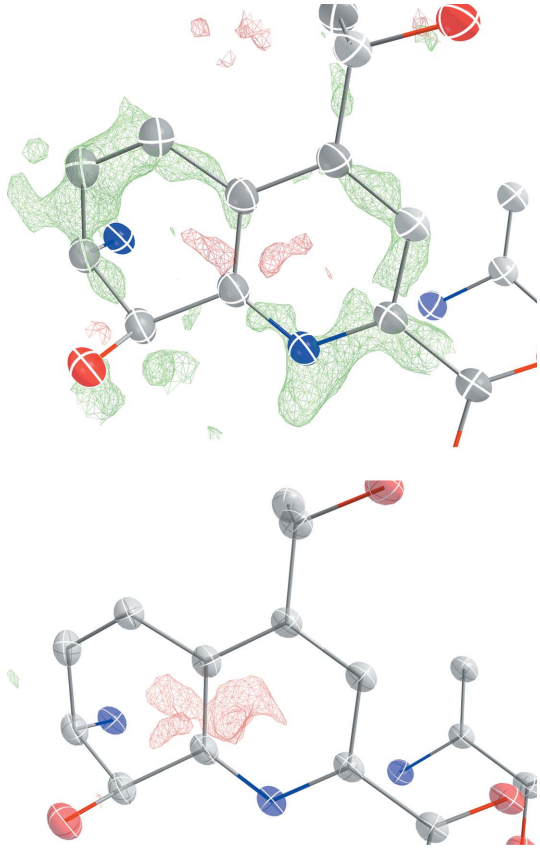

Figure 3
Residual electron-density map for the quinaldic acid residue of thiostrepton [top, independent atom model (IAM); bottom, invariom model] as generated with *MoleCoolQt* with values of ±0.23 e Å^{-3}.

[5] Supplementary material has been deposited in the IUCr electronic archive (Reference: DZ5276). Services for accessing this material are described at the back of the journal.
[6] The suitability of a compound also depends on the ratio S = volume/$\sum n_{core}^2$ (Stevens & Coppens, 1976), *i.e.* on valence electron density that can actually be modelled.

4.3. Absolute structure and the Flack parameter

For natural products derived from the naturally occurring amino acids the chirality of the centres is expected to be known. Visual inspection shows that the correct absolute structure of thiostrepton had already been assigned by Anderson *et al.* (1970). Nevertheless, we wanted to verify the absolute structure *via* the presence of the anomalous signal with our current data set by Flack parameter (Flack, 1983) refinement with invarioms (Dittrich, Strumpel *et al.*, 2006). In order to arrive at a correct result, the values for anomalous scattering at 0.6500 Å were extrapolated from a fit of the known values given in *International Tables for Crystallography* Volume C; Friedel mates were used in refinement as separate reflections. Modelling of the valence electron density in thiostrepton indeed improves the standard deviation of the Flack parameter, which is 0.05 (4) in *XD* IAM refinement and 0.05 (3) in *XD* invariom refinement.

4.4. The molecular electrostatic potential

One-electron properties such as the molecular electrostatic potential (ESP) can be rapidly calculated from the invariom database parameters. An ESP that is calculated in this way is superior to one derived from point-charge models and is directly comparable to results from quantum-chemical calculations, as has been proven from difference-electron densities in small-molecule crystal structures (Holstein *et al.*, 2012). The molecular ESP for thiostrepton is shown in Fig. 4 (top, front side view; bottom, back side view). For ESP calculation the program *XDPROP* from the 2006 version of the *XD* package (Volkov *et al.*, 2006) was used. Only the higher occupied positions of the disordered residues were chosen for property calculation, not taking into account sites with lower occupancy. The *XD* property program *XDPROP* assumes that split sites are fully occupied in the process. Multipole populations hence correspond to atoms with full occupancies, since derived properties such as the ESP do not make physical sense for a model with split occupancies. In order to derive a correct ESP it is advantageous that the correct quantum-mechanical X—H distances from the invariom database are used. Alternatively, averaged values from neutron diffraction (Allen & Bruno, 2010) would provide a similar but slightly less accurate result, since the number of individual bonding situations (causing differences in bond distances) available from neutron diffraction is naturally limited. Alternatively, only H-atom positions can be geometry-optimized by *ab initio* methods, keeping non-H-atom positions fixed. How to best derive H-atom positions will be reported in a subsequent paper.

Values for the ESP of thiostrepton range from +0.44 to −0.41 e Å^{-1}, a spread exceeding the typical range for neutral molecules (Politzer *et al.*, 2001; Holstein *et al.*, 2012) when using a ρ = 0.001 au (0.00067 e Å^{-3}) molecular surface. The molecule has a more hydrophobic front (Fig. 4, top) and a more hydrophilic back side (Fig. 4, bottom). The polarizable S atoms are more visible in the back part. Evidence has been provided that the flexible side chain of thiostrepton binds in between the ribosomal protein L11 and the 23S rRNA (Harms

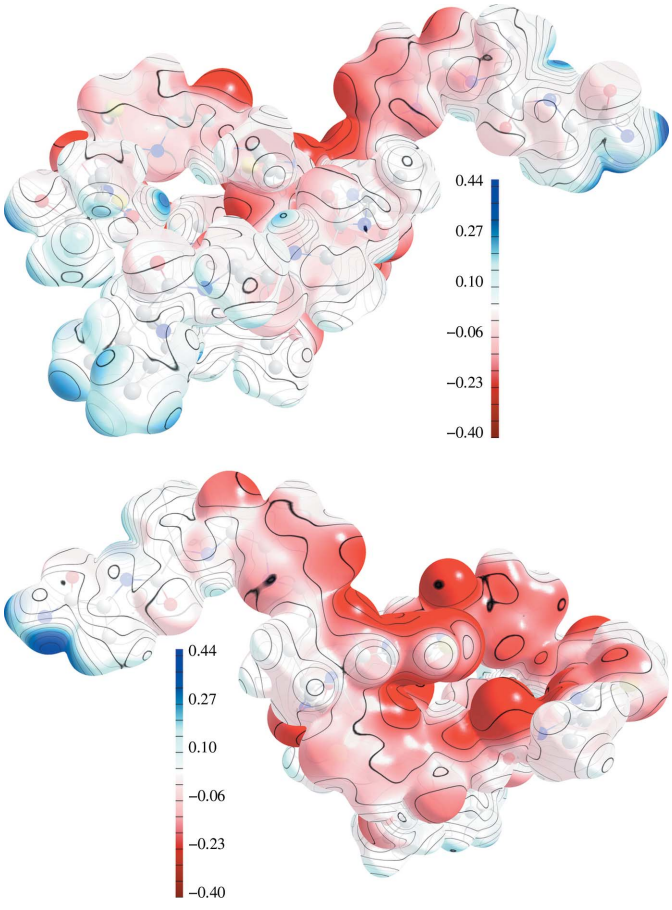

Figure 4
Electrostatic potential (e Å^{-1}) (top, front face; bottom, back face) of one of the two similar molecules in the asymmetric unit of the $P2_1$ form of thiostrepton mapped on an electron-density isosurface of ρ = 0.0067 e Å^{-3}. Solvent molecules were omitted for clarity.

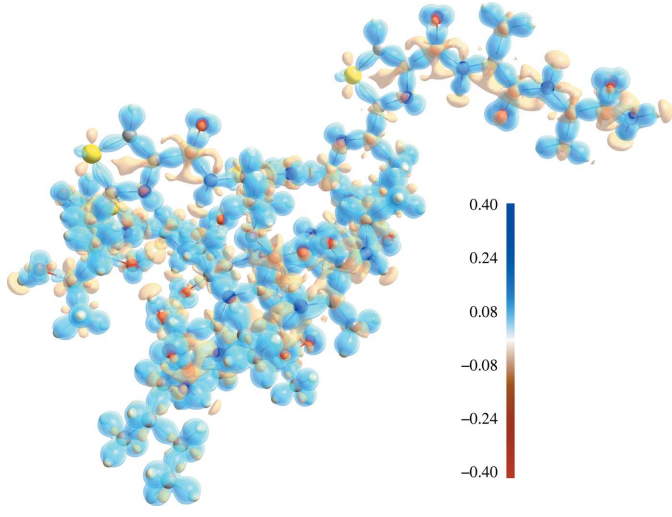

Figure 5
Deformation electron density of thiostrepton, illustrating the difference between the spherical IAM density and the aspherical valence electron density of the multipole model. Blue density depicts excess, whereas yellow regions show decreased electron density with respect to the IAM. Contour values are indicated in the legend in units of e Å^{-3}.

et al., 2008), thereby interfering with the 70S subunit of the ribosome. Hence, thiostrepton and other thiopeptides block protein translation in the ribosomal GTPase centre (Jonker *et al.*, 2007; Schoof *et al.*, 2009). With regard to biological activity, it would be interesting to investigate possible electrostatic complementarity with ribosomal L11 and 23S rRNA. Such complementarity has been shown to exist for a human aldose reductase–drug-molecule inhibitor complex (Muzet *et al.*, 2003). We expect similar drug–receptor interactions to also play an important role in molecular recognition of thiostrepton. Of particular interest could be the ESP of the flexible side chain containing the two planar dehydroalanine residues. This group shows a characteristic pattern of polarity with alternating carbonyl and dehydroalanine/amide functionalities, and future studies are required to determine whether electrostatic complementarity indeed plays a role in the recognition processes.

4.5. Deformation electron density and molecular dipole moment

Invariom database parameters were also used to generate a deformation electron-density plot.[7] The aim of such a plot in the context of invariom modelling is to verify that the orientation and assignment of the aspherical theoretical electron density described by the scattering factor is correct; a detailed picture of the deviation from the spherical scattering factors is obtained. These plots can either be obtained by fast Fourier transform of static multipole–IAM scattering-factor differences or from the static electron-density distribution as constructed from the invariom database parameters. For the generation of Fig. 5 we have chosen the latter approach, giving a more precise grid of deformation electron density. The rapid FFT calculation in *MoleCoolQt* enables a similar but instantaneous validation of the correct assignment of the aspherical scattering factor and its coordinate system and was used throughout refinement. A topological analysis according to Bader's quantum theory of atoms in molecules (Bader, 1990) is also feasible with the invariom density model. Since the electron density has been derived from theoretical calculations here, such analysis mainly reproduces the bonding situation of the model compounds and does not provide additional information. We therefore did not carry out a topological analysis.

The molecular dipole moment provides

> a compact summary of the molecular charge distribution (*i.e.* the distribution of both positive and negative charge)

(Spackman, 1992). It can alternatively be obtained from a point-charge model, but evaluation of the population parameters of the multipole model is expected to provide better accuracy, since contributions from the aspherical electron density are considered. The calculation again required focusing on the nondisordered and the major components of the disordered parts of the molecule and solvent molecules were omitted. Both molecules have a very similar structure and dipole moment (55.0 and 50.0 D). To obtain dipole moments for comparison from theory we have performed a single-point energy calculation as well as initiated a geometry optimization with the program *GAUSSIAN*09 (Frisch *et al.*, 2009) of one isolated molecule with the conformation found in the crystal. Owing to the size of the molecule the method/basis set combination [B3LYP/6-31G(d)] of the optimization was rather modest, whereas it was more extended for the single-point calculation [B3LYP/D95++(3df,3pd)]. The optimized geometry is rather similar to the solid-state conformation and differs mainly in the orientation of the disordered dehydroalanine-rich side chain. The dipole moment of the single-point calculation (7.6 D) is strikingly lower than the multipole-model result. The dipole moment is artificially increased to 45.4 D after a projection of the quantum-chemical electron density from optimization *via* structure factors. We ascribe this to difficulties of the Hansen–Coppens multipole model in accurately modelling the diffuse electron density of the S atom (Dominiak & Coppens, 2006).[8] These difficulties might also partially explain the higher numerical spread between the minimum and maximal values observed in the ESP. Taking into account the effort in generating an accurate model for this macromolecule this result is unsatisfactory. Either increasing the order l in combination with m-dependent radial functions of an improved rigid pseudoatom model (Koritsánszky *et al.*, 2012) or moving to a basis-set description (Jayatilaka & Grimwood, 2001) can be expected to provide a remedy. While for strongly dipolar molecules containing only C, H, N, O and F atoms a satisfactory agreement between theory and multipole model was found (Holstein *et al.*, 2012), not only

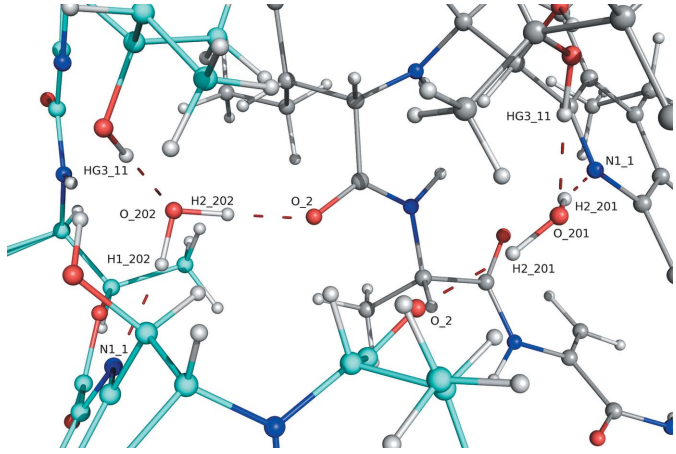

Figure 6
Intermolecular interactions between water and thiostrepton molecules in the new $P2_1$ crystal structure, applying the notation used in the PDB file. Thiostrepton molecules in the asymmetric unit are shown in grey.

[7] The deformation density, or more precisely the static deformation electron density, is calculated using the equation $\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}_{\rm model} - \mathbf{r}_{\rm reference})$, with the reference model being the IAM. The model electron density is calculated from the electron-density expression of the Hansen–Coppens multipole model (Hansen & Coppens, 1978). The term static indicates that only electron-density parameters are taken into account, not smearing from ADPs.

[8] Modifying the exponents of the S deformation radial functions using $n_l = 2, 4, 6$ and 8 as suggested by these authors gave a very similar result of 44.3 D rather than 45.4 D.

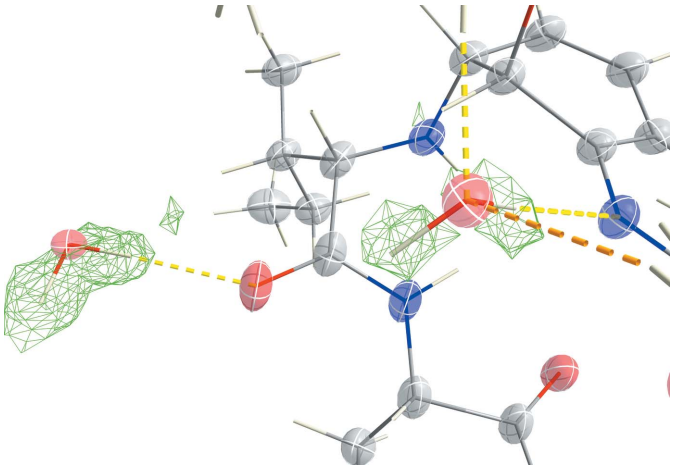

Figure 7
$F_o - F_c$ OMIT map of the water H atoms with the green mesh showing isosurfaces of 0.28 e Å^{-3} in a similar orientation as in Fig. 6.

sulfur-rich compounds would benefit from more accurate electron-density models in the future (Volkov & Coppens, 2001; Dittrich *et al.*, 2012; Bąk *et al.*, 2012).

4.6. Hydrogen bonding and intermolecular interactions

Intramolecular and intermolecular hydrogen-bond interactions will contribute to and be influenced by the polarization observed in the ESP. They are likely to be involved in the molecular-recognition processes and antibacterial action of thiostrepton. The new high-resolution $P2_1$ structure of thiostrepton allows full characterization of both intermolecular and intramolecular interactions (Fig. 6). This includes the split atomic positions of the bis-dehydroalanine tail region. A water molecule has been found to play a special role in influencing the conformation of thiostrepton in the tetragonal structure of thiostrepton (Bond *et al.*, 2001). This water molecule is again observed in the new $P2_1$ form. We initially attempted to crystallize thiostrepton from dry dimethylformamide (DMF) solvent. These attempts were unsuccessful, and only when wet DMF was used were crystals obtained. We conclude that this water molecule is essential for crystallizing thiostrepton in that it causes the rigid conformation required for packing. It is conceivable that chemical modification to replace the water molecule by a covalent bond would enforce the conformation, thereby influencing antibiotic activity. It might also be related to the problems encountered with poor solubility that limit the bioavailability of thiostrepton. With respect to the dynamics of the interaction with ribosomal L11 and 23S rRNA, removing the tightly bound water would cause an energetic penalty but would also increase the molecular conformational flexibility.

In the new $P2_1$ form a second well resolved water molecule is found. It connects two molecules related to each other by symmetry. Both water molecules and their hydrogen-bonding environment are shown again in Fig. 7. In this figure an $F_o - F_c$ OMIT map highlights the position of the H atoms, thereby illustrating the presence of signal for H atoms of the non-disordered parts of the structure in the diffraction data.

5. Discussion: macromolecular refinement with aspherical scattering factors

The example of thiostrepton highlights the challenges and the opportunities for aspherical-atom refinement of macromolecules. Such refinements require the highest possible data quality, coverage and resolution that can be achieved. Equally important are accurate low-order reflections that carry most of the information on valence electron density. Hence, a compromise in the measurement strategy is required that allows the measurement of accurate low-order and high-order reflections (Dauter, 2003) while also minimizing radiation damage. Not many macromolecular crystals scatter to high resolution. Such studies can hence only be carried out for these very few favourable peptide, protein or DNA structures, which often have efficient packing and low solvent content. However, thiostrepton shows that high resolution in itself is not sufficient, since large-amplitude vibrations of the dehydroalanine-rich tail region lead to continuous disorder in the previously characterized $C2$ and $P4_32_12$ (Bond *et al.*, 2001) solvate structures. Such continuous or large-amplitude disorder cannot be successfully modelled with the combination of atomic positions and displacement parameters without extensive use of restraints or external information. The reason for the failure of the scattering model in such cases is that parameters for electron density and thermal motion become increasingly correlated and ill-determined and that the fundamental assumption of sharp atomic peaks becomes invalid as atoms 'fade away'.[9] Main-chain disorder in thiostrepton also shows that ensemble models (MacArthur & Thornton, 1999; Lang *et al.*, 2010) might be required. We clearly need to go beyond the established methodology to model these and other high-resolution structures. Even ultralow temperatures do not provide an easy solution, since the disorder is usually frozen, with different orientations in different unit cells or mosaic blocs (as we learned from an unpublished 5 K high-resolution data collection from the $P4_32_12$ solvate of thiostrepton). Only the crystallization of a new solvate ultimately led to a data set for thiostrepton that was suitable for aspherical-atom refinement. Modelling required that the starting model from *SHELXL* includes as many resolved solvent molecules as possible. Correction terms for disordered solvent were therefore not required. Such corrections are counterproductive for modelling valence electron density anyway, since they modify the scattering contribution in the low-order reflections that we try to evaluate. Ultimately, we wanted to reach a point at which the electron density and the atomic displacements owing to thermal motion or disorder were deconvoluted. This was achieved for the new solvate of thiostrepton and is rewarded by accurate atomic positions and molecular properties[10]

[9] Neither can anharmonic models of thermal motion provide a remedy in such cases.

[10] The question arises as to how much derived properties change when comparing results from a more accurate/precise structure from full-matrix least-squares refinement with those from a lower resolution structure refined with the usual geometry restraints. This question cannot be answered with the methodology used at the current stage.

derived from the electron-density model of this macromolecular structure.

6. Conclusion

A new $P2_1$ solvate of the thiopeptide antibiotic thiostrepton was crystallized and diffraction data were measured to d = 0.64 Å with synchrotron radiation using a new pixel-array detector. The structure was refined with invariom, namely nonspherical scattering factors. This led to a drop in $R_1(F)$ and to physically more meaningful ADPs and avoids asphericity shifts as previously observed for small molecules. However, the effort involved is considerable, especially for the disordered parts of the molecule. The effort is rewarded by obtaining a molecular electron-density distribution as reconstructed from pseudoatomic fragments. Evaluation of the molecular electron density provides one-electron properties that can be directly and rapidly calculated after refinement. We report and discuss the deformation electron density, the electrostatic potential mapped on the electron-density isosurface and the molecular dipole moment. Comparisons with theoretical dipole moments show that for this sulfur-rich compound the reproducibility is worse than, for example, the fluoroquinolone class of compounds studied previously.

Compared with the known tetragonal and monoclinic structures of thiostrepton, the results of the present study suggest that high-resolution data are certainly required, but in themselves are not sufficient, to derive accurate atomic positions of macromolecules when extensive disorder is present. To obtain reliable electrostatic properties, accurate geometries are mandatory. This limits the application of aspherical-atom refinements to a small number of well suited (well behaving) macromolecular structures. For thiostrepton a detailed electrostatic potential could only be obtained after the new monoclinic solvate was crystallized, since the flexible dehydroalanine tail region is thus far only resolved in the new $P2_1$ structure.

Funding for this work within the Deutsche Forschungsgemeinschaft (DFG; grant DI 921/3-2) is gratefully acknowledged. We thank the staff at beamline X10SA of the SLS synchrotron for their valuable support before the measurements. We are grateful to the referees for their recommendations for improving the manuscript.

References

Afonine, P. V., Grosse-Kunstleve, R. W., Adams, P. D., Lunin, V. Y. & Urzhumtsev, A. (2007). *Acta Cryst.* D**63**, 1194–1197.
Afonine, P. V., Lunin, V. Y., Muzet, N. & Urzhumtsev, A. (2004). *Acta Cryst.* D**60**, 260–274.
Allen, F. H. & Bruno, I. J. (2010). *Acta Cryst.* B**66**, 380–386.
Anderson, B., Hodgkin, D. C. & Viswamitra, M. A. (1970). *Nature (London)*, **225**, 233–235.
Arndt, H.-D., Schoof, S. & Lu, J.-Y. (2009). *Angew. Chem. Int. Ed.* **48**, 6770–6773.
Bader, R. F. W. (1990). *Atoms in Molecules: A Quantum Theory*, 1st ed. Oxford: Clarendon Press.
Bagley, M. C., Dale, J. W., Merritt, E. A. & Xiong, X. (2005). *Chem. Rev.* **105**, 685–714.
Bąk, J. M., Czyżnikowska, Ż. & Dominiak, P. M. (2012). *Acta Cryst.* A**68**, 705–714.
Bond, C. S., Shaw, M. P., Alphey, M. S. & Hunter, W. N. (2001). *Acta Cryst.* D**57**, 755–758.
Broennimann, C., Eikenberry, E. F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T., Toyokawa, H. & Wagner, A. (2006). *J. Synchrotron Rad.* **13**, 120–130.
Brzezinski, K., Brzuszkiewicz, A., Dauter, M., Kubicki, M., Jaskolski, M. & Dauter, Z. (2011). *Nucleic Acids Res.* **39**, 6238–6248.
Coppens, P., Sabine, T. M., Delaplane, G. & Ibers, J. A. (1969). *Acta Cryst.* B**25**, 2451–2458.
Dauter, Z. (2003). *Methods Enzymol.* **368**, 288–337.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Dittrich, B., Bond, C. S., Spackman, M. A. & Jayatilaka, D. (2010). *Cryst. Eng. Commun.* **12**, 2419–2423.
Dittrich, B., Hübschle, C. B., Holstein, J. J. & Fabbiani, F. P. A. (2009). *J. Appl. Cryst.* **42**, 1110–1121.
Dittrich, B., Hübschle, C. B., Luger, P. & Spackman, M. A. (2006). *Acta Cryst.* D**62**, 1325–1335.
Dittrich, B., Hübschle, C. B., Pröpper, K., Dietrich, F., Stolper, T. & Holstein, J. J. (2013). *Acta Cryst.* B**69**, 91–104.
Dittrich, B., Koritsánszky, T. & Luger, P. (2004). *Angew. Chem. Int. Ed.* **43**, 2718–2721.
Dittrich, B., Strumpel, M., Schäfer, M., Spackman, M. A. & Koritsánszky, T. (2006). *Acta Cryst.* A**62**, 217–223.
Dittrich, B., Sze, E., Holstein, J. J., Hübschle, C. B. & Jayatilaka, D. (2012). *Acta Cryst.* A**68**, 435–442.
Domagała, S., Fournier, B., Liebschner, D., Guillot, B. & Jelsch, C. (2012). *Acta Cryst.* A**68**, 337–351.
Dominiak, P. M. & Coppens, P. (2006). *Acta Cryst.* A**62**, 224–227.
Dominiak, P. M., Volkov, A., Li, X., Messerschmidt, M. & Coppens, P. (2007). *J. Chem. Theory Comput.* **3**, 232–247.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Flack, H. D. (1983). *Acta Cryst.* A**39**, 876–881.
Frisch, M. J. *et al.* (2009). *GAUSSIAN*09, revision A.02. Gaussian Inc., Wallingford, Connecticut, USA.
Gonzalez, R. L., Chu, S. & Puglisi, J. D. (2007). *RNA*, **13**, 2091–2097.
Guillot, B., Jelsch, C., Podjarny, A. & Lecomte, C. (2008). *Acta Cryst.* D**64**, 567–588.
Guillot, B., Viry, L., Guillot, R., Lecomte, C. & Jelsch, C. (2001). *J. Appl. Cryst.* **34**, 214–223.
Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.
Harms, J. M., Wilson, D. N., Schluenzen, F., Connell, S. R., Stachelhaus, T., Zaborowska, Z., Spahn, C. M. & Fucini, P. (2008). *Mol. Cell*, **30**, 26–38.
Hellner, E. (1977). *Acta Cryst.* B**33**, 3813–3816.
Hirshfeld, F. L. (1976). *Acta Cryst.* A**32**, 239–244.
Holstein, J. J., Hübschle, C. B. & Dittrich, B. (2012). *Cryst. Eng. Commun.* **14**, 2520–2531.
Housset, D., Benabicha, F., Pichon-Pesme, V., Jelsch, C., Maierhofer, A., David, S., Fontecilla-Camps, J. C. & Lecomte, C. (2000). *Acta Cryst.* D**56**, 151–160.
Hübschle, C. B. & Dittrich, B. (2011). *J. Appl. Cryst.* **44**, 238–240.
Hübschle, C. B., Luger, P. & Dittrich, B. (2007). *J. Appl. Cryst.* **40**, 623–627.
Hübschle, C. B., Sheldrick, G. M. & Dittrich, B. (2011). *J. Appl. Cryst.* **44**, 1281–1284.
Jarzembska, K. N. & Dominiak, P. M. (2012). *Acta Cryst.* A**68**, 139–147.
Jayatilaka, D. & Grimwood, D. J. (2001). *Acta Cryst.* A**57**, 76–86.
Jelsch, C., Guillot, B., Lagoutte, A. & Lecomte, C. (2005). *J. Appl. Cryst.* **38**, 38–54.
Jelsch, C., Teeter, M. M., Lamzin, V., Pichon-Pesme, V., Blessing, R. H. & Lecomte, C. (2000). *Proc. Nat. Acad. Sci. USA*, **97**, 3171–3176.

Johnas, S. K. J., Dittrich, B., Meents, A., Messerschmidt, M. & Weckert, E. F. (2009). *Acta Cryst.* D**65**, 284–293.
Jonker, H. R., Ilin, S., Grimm, S. K., Wöhnert, J. & Schwalbe, H. (2007). *Nucleic Acids Res.* **35**, 441–454.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Koritsánszky, T., Richter, T., Macchi, P., Volkov, A., Gatti, C., Howard, S., Mallinson, P. R., Farrugia, L., Su, Z. W. & Hansen, N. K. (2003). *XD – A Computer Program Package for Multipole Refinement and Topological Analysis of Electron Densities from Diffraction Data.* Freie Universität Berlin, Germany.
Koritsánszky, T., Volkov, A. & Chodkiewicz, C. (2012). *Struct. Bond.* **147**, 1–26.
Kwok, J. M.-M, Myatt, S. S., Marson, C. M., Coombes, R. C., Constantinidou, D. & Lam, E. W.-F. (2008). *Mol. Cancer Ther.* **7**, 2022–2032.
Lang, P. T., Ng, H.-L., Fraser, J. S., Corn, J. E., Echols, N., Sales, M., Holton, J. M. & Alber, T. (2010). *Protein Sci.* **19**, 1420–1431.
Lecomte, C., Jelsch, C., Guillot, B., Fournier, B. & Lagoutte, A. (2008). *J. Synchrotron Rad.* **15**, 202–203.
Lee, D., Walsh, J. D., Yu, P., Markus, M. A., Choli-Papadopoulou, T., Schwieters, C. D., Krueger, S., Draper, D. E. & Wang, Y.-X. (2007). *J. Mol. Biol.* **367**, 1007–1022.
MacArthur, M. W. & Thornton, J. M. (1999). *Acta Cryst.* D**55**, 994–1004.
Muzet, N., Guillot, B., Jelsch, C., Howard, E. & Lecomte, C. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 8742–8747.
Nicolaou, K. C. (2012). *Angew. Chem. Int. Ed.* **51**, 12414–12436.
Nicolaou, K. C., Zak, M., Rahimipour, S., Estrada, A. A., Lee, S. H., O'Brate, A., Giannakakou, P. & Ghadiri, M. R. (2005). *J. Am. Chem. Soc.* **127**, 15042–15044.
Nicolaou, K. C., Zak, M., Safina, B. S., Estrada, A. A., Lee, S. H. & Nevalainen, M. (2005). *J. Am. Chem. Soc.* **127**, 11176–11183.
Pagano, J. F., Weinstein, M. J., Stoup, H. A. & Donovick, R. (1956). *Antibiotics Annual 1955–1956*, edited by H. Welch & F. M. Ibañez, pp. 554–559. New York: Medical Encyclopedia Inc.
Petrova, T. & Podjarny, A. (2004). *Rep. Prog. Phys.* **67**, 1565–1605.
Politzer, P., Murray, J. S. & Preralta-Inga, Z. (2001). *Int. J. Quantum Chem.* **85**, 676–684.
Schmidt, A., Teeter, M., Weckert, E. & Lamzin, V. S. (2011). *Acta Cryst.* F**67**, 424–428.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Schnieders, M. J., Fenn, T. D., Pande, V. S. & Brunger, A. T. (2009). *Acta Cryst.* D**65**, 952–965.
Schoof, S., Baumann, S., Ellinger, B. & Arndt, H.-D. (2009). *Chem. Phys. Chem.* **10**, 242–245.
Sheldrick, G. M. (2007). *SADABS v.2008/2. A Program for Area Detector Absorption and Other Corrections.* University of Göttingen, Germany.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Spackman, M. A. (1992). *Chem. Rev.* **92**, 1769–1797.
Stevens, E. D. & Coppens, P. (1976). *Acta Cryst.* A**32**, 915–917.
Stewart, R. F. (1976). *Acta Cryst.* A**32**, 565–574.
Stewart, R. F. (1977). *Isr. J. Chem.* **16**, 124–131.
Thorn, A., Dittrich, B. & Sheldrick, G. M. (2012). *Acta Cryst.* A**68**, 448–451.
Urzhumtseva, L., Afonine, P. V., Adams, P. D. & Urzhumtsev, A. (2009). *Acta Cryst.* D**65**, 297–300.
Volkov, A. & Coppens, P. (2001). *Acta Cryst.* A**57**, 395–405.
Volkov, A., Li, X., Koritsánzky, T. & Coppens, P. (2004). *J. Phys. Chem. A*, **108**, 4283–4300.
Volkov, A., Macchi, P., Farrugia, L. J., Gatti, C., Mallinson, P., Richter, T. & Koritsánszky, T. (2006). *XD2006 – A Computer Program Package for Multipole Refinement, Topological Analysis of Charge Densities and Evaluation of Intermolecular Energies from Experimental or Theoretical Structure Factors.* University at Buffalo, USA.
Walter, J. D., Hunter, M., Cobb, M., Traeger, G. & Spiegel, P. C. (2012). *Nucleic Acids Res.* **40**, 360–370.
Zarychta, B., Pichon-Pesme, V., Guillot, B., Lecomte, C. & Jelsch, C. (2007). *Acta Cryst.* A**63**, 108–125.