Structures of Four Crystal Forms of Decaplanin

by **Christopher Lehmann**[a]), **Judit É. Debreczeni**[b]), **Gábor Bunkóczi**[b]), **Miroslawa Dauter**[c]), **Zbigniew Dauter**[c]), **László Vértesy**[d]), and **George M. Sheldrick***[b])

[a]) Center for Advanced Research in Biotechnology, 9600 Gudelsky Drive, Rockville, MD 20850, USA
[b]) Lehrstuhl für Strukturchemie, Tammannstrasse 4, D-37077, Göttingen
(fax: +49(551)392582; e-mail: gsheldr@shelx.uni-ac.gwdg.de)
[c]) Synchrotron Radiation Research Station, National Cancer Research Institute, Brookhaven National Laboratory, Building 725A-X9, Upton, NY 11973, USA
[d]) *Aventis Pharma Deutschland* GmbH, D-65926 Frankfurt am Main

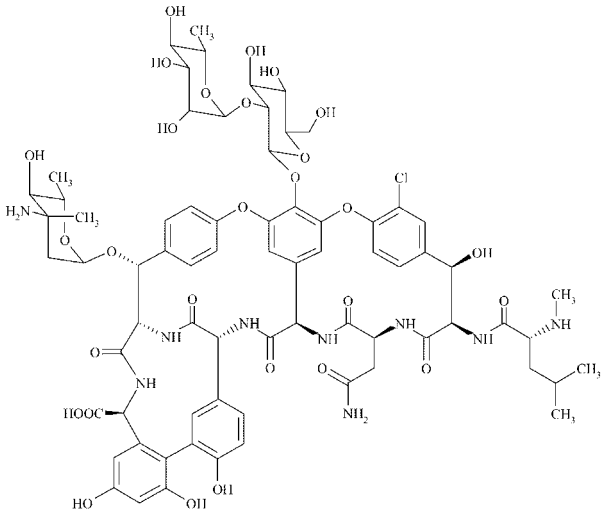
Dedicated to Professor *Jack D. Dunitz* on the occasion of his 80th birthday

The glycopeptide antibiotic decaplanin (**1**; formerly known as MM 47761 and M86-1410) crystallizes in two $P2_1$ and two $P6_122$ crystal forms, each with four monomers in the asymmetric unit, with solvent contents varying from 48 to 69%. Although with *ca.* 600 unique atoms, the structures are larger than typical small molecules, one was solved by direct methods. The other three were solved by typical macromolecular methods: single-wavelength anomalous diffraction (SAD) of the Cl-atoms present naturally in the structure, multiple-wavelength anomalous diffraction (MAD) at the Br absorption edge for a crystal soaked in NaBr solution, and molecular replacement. There is evidence of appreciable radiation damage with loss of 20–30% of the covalent and ionic halogens affecting the synchrotron datasets that may even have unintentionally facilitated the MAD structure solution. The structures contain the dimer units typical of antibiotics related to vancomycin, but, in addition, there are a variety of further intermolecular interactions responsible for the polymorphy leading to intertwined 6_1-helices in two of the crystal forms. Except for the sugars and some sidechains, the conformations of the 16 independent monomers are very similar.

Introduction. – The increasing threat presented by pathogens that have developed multiple resistance to antibiotics [1], in particular by methicillin-resistant *Staphylococcus aureus* (MRSA) strains, calls for new approaches in antibacterial chemotherapy. One such approach to combat bacterial infections involves the simultaneous application of several antibiotics. It has been shown that the antibiotic decaplanin (**1**) acts in synergy with cephalosporins [2]; low concentrations of **1** suffice to make cephalosporins effective in the treatment of MRSA infections that are caused by β-lactam-resistant pathogens.

Decaplanin (**1**) is a glycopeptide antibiotic related to vancomycin that was isolated from cultures of *Kibdellosporangium deccaensis* DSM 4763 in 1989 in India [2]. It was subsequently found to be identical to the antibiotic MM 47761 [3] produced by *Amycolatopsis orientalis*. Compared with typical polypeptides and also vancomycin, decaplanin (**1**) is unusually stable and retains its full activity after several hours at 80°. The decaplanin (**1**) molecule appears to bind anions strongly. The acetate, tartrate, and phosphate salts dissolve well in H_2O; both these and the less-soluble chloride and citrate crystallize fairly readily, and this has been exploited for the purification of **1** [2].

With *ca.* 600 unique non-H-atoms, these structures fall in the gap between typical small molecules and macromolecules, and their determination presented a challenge



Decaplanin (**1**)

that required the use of both small-molecule and macromolecular crystallographic techniques. The first crystal form, I, in space group $P2_1$ diffracted to 1.0 Å and could be solved by *ab initio* direct methods [4]; it is one of the largest structures solved by this essentially small-molecule method. The second $P2_1$ form, II, was solved with poor-quality data from a large but extremely mosaic crystal by molecular replacement [5], a predominantly macromolecular method, using a dimer from form I as a search fragment. The two $P6_122$ crystal forms were solved at lower resolution by typical macromolecular methods: multiple-wavelength anomalous diffraction (MAD [6]) of Br-ions incorporated by soaking (form III) and single-wavelength anomalous dispersion (SAD [7]) of the Cl-atoms present naturally in the crystal (form IV). Since these problems provide a good illustration of recent progress that has been made in macromolecular phasing, we have repeated our original somewhat inefficient structure determinations with up-to-date methodology as described below.

Results and Discussion. – *MAD Structure Determination of Form III*. The program SHELXD [4] with data truncated to 2.4 Å clearly revealed the presence of nine heavy atom sites with a range of occupancies that were assumed to be partially occupied bromides. These were input into SHELXE [8] without refinement of their coordinates or occupancies. After 50 cycles of density modification by the *sphere of influence* method, the map correlation coefficient relative to the final refined structure was 0.927, and the weighted mean phase error was 15.9°, and the map revealed the full structure. The only puzzling feature was that three of the putative bromides corresponded very closely to the positions of covalently bonded chlorines!

To resolve this puzzle, the data were reprocessed as two SAD experiments for the two wavelengths and one SIR experiment in which the high-energy remote data were

treated as the derivative and the inflection-point data as native. These data and the final phases were used to calculate three 'heavy-atom *Fourier* syntheses' [9]. The three covalent Cl sites gave strong peaks with the SIR data but were at the noise level for both anomalous datasets. The remaining six sites had similar peak heights in all three maps and occupied positions that would be chemically acceptable for bromides. Similar peak heights in all three maps would be expected for pure bromine MAD data because $2f''$ at both wavelengths and f'(remote) – f'(inflection) are all *ca.* +7 electrons. The only reasonable explanation is that the three 'covalent chlorine' peaks are caused by partial loss of these Cl-atoms during the data collection as a result of radiation damage, and the remaining peaks are indeed partially occupied bromides. Refining the occupancies of the sites in the model refinement proved difficult because four of the bromide sites appear to be mixed bromide/chloride sites with occupancies adding up to *ca.* 0.7 to 0.8 (the corresponding sites are occupied by chlorides in form IV), and because of correlations between occupancies and displacement parameters that are particularly large because of the low data resolution (1.9 Å). However, when a common occupancy was employed for all four covalent Cl-atoms, it refined to 0.8 for the remote wavelength (collected first) and 0.7 for the inflection-point data, which indicates that there was appreciable radiation damage. In addition, for all the synchrotron datasets employed in this project, large negative difference electron-density peaks were observed at the positions of the Cl-atoms when they were refined with unit occupancies, suggesting that radiation damage is a general problem for decaplanin (**1**). In retrospect, the additional sites caused by radiation damage probably provided useful phase information in the MAD experiment, and it may be possible to capitalize on this effect in future MAD structure determinations. The best strategy will be simply to collect the inflection-point data last so that the dispersive differences and radiation-damage effects have the same signs and reinforce one another rather than canceling out!

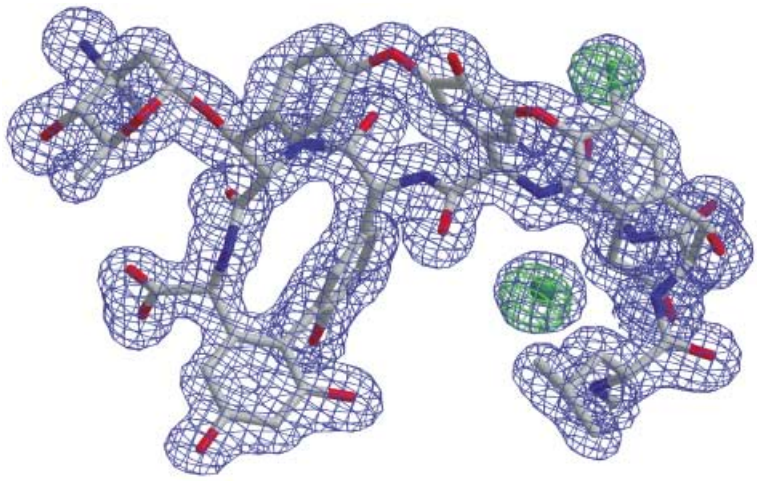


Fig. 1. *The model-free experimental electron-density map* (in blue) *after SHELXE processing of the SAD data for one monomer of form IV with the anomalous density* (green) *and superimposed final refined model*

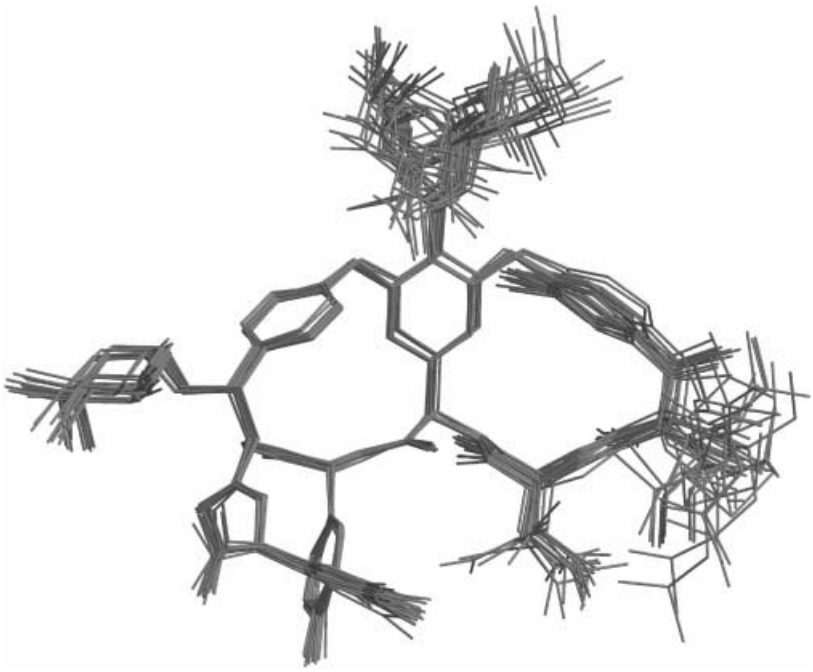


Fig. 2. *Superimposed structures of all 16 independent monomers based on a fit of the 41 best-fitting atoms.* Apart from the sugars and some flexible side chains, the fit is very close. The sugar at the top of the picture adopts two principal conformations and the side chains that can enclose the binding site are most flexible.

SAD Structure Determination of Form IV. The use of the anomalous scattering of Cl for solving macromolecular structures has already been demonstrated by *Loll* [10] using synchrotron data and *Lehmann et al.* [11] using in-house data but requires precise data (best achieved by a high redundancy); at a longer wavelength, f'' for Cl is larger. Although the overall redundancy (18.6) was not particularly high, SHELXD gave eight very clear sites when the SAD data collected at a wavelength of 1.542 Å were truncated to 2.1-Å resolution. Eight monomers with one covalently bound Cl-atom in each would have given an eminently reasonable solvent content of *ca.* 35%, but (as with form III) the self-rotation function showed only a 24-fold noncrystallographic symmetry axis parallel to the crystallographic 6_1-axis, which was more consistent with four molecules in the asymmetric unit and nearly 70% solvent. The *sphere of influence* method of density modification in SHELXE yielded a map correlation coefficient of 0.960 and weighted mean phase error of 11.8° relative to the final refined model and an essentially perfect map (*Fig. 1*). As with form III, the calculations were performed with standard settings for the programs and took only a couple of minutes.

Comparison of the Monomers. The 16 independent monomers in the four structures were compared using the program ESCET [12]. Leaving out the disordered side-chain atoms near the N-terminus, 93 atoms remain. ESCET showed that there is a rigid core consisting of 41 of these atoms, with a mean deviation of 0.20 Å between corresponding atoms. The two most-similar molecules were two of the independent

molecules in form IV (mean deviation 0.05 Å), and the most-dissimilar were one molecule in form III and one in form IV (0.31 Å). *Fig. 2* shows a superposition of all 16 molecules after fitting these 41 atoms.

Crystal Packing. All four structures contain the decaplanin (**1**) homodimer shown in *Fig. 3* that is found for most glycopeptide antibiotics in the solid state and in solution [11][13]. There are a relatively limited number of contacts between molecules, accounting for the high solvent contents. *Fig. 4* shows a mixed aromatic and H-bonded interaction between two molecules not related by crystallographic symmetry seen in forms I and II. Both form III and form IV contain double intertwined helices about the 6_1 axis, with a large solvent channel through the middle. The nonpolar contact shown in *Fig. 5* is between different helices in both structures and involves a twofold crystallographic axis; the homodimer interaction (*Fig. 3*) and the polar interaction accommodating the halides (*Fig. 6*) are used to construct each helix. Thus, a single helix consists of two homodimers and the symmetry-related homodimers generated by the action of the 6_1 axis; the other helix is generated by the action of a twofold crystallographic axis from the first. From form III to form IV, the helices have been stretched, reducing their diameters a little in the process.

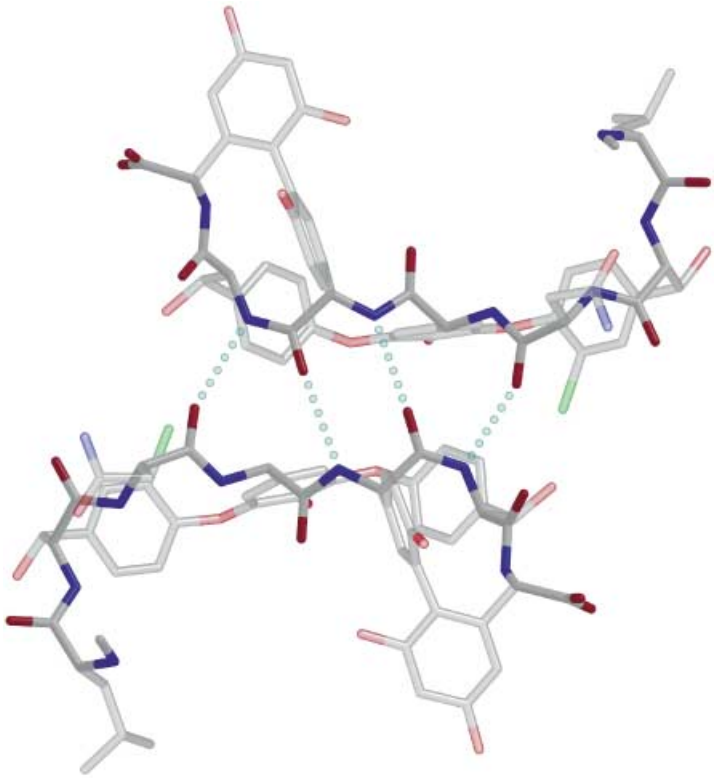


Fig. 3. *The decaplanin* (**1**) *homodimer present in all four crystal forms and in solution*

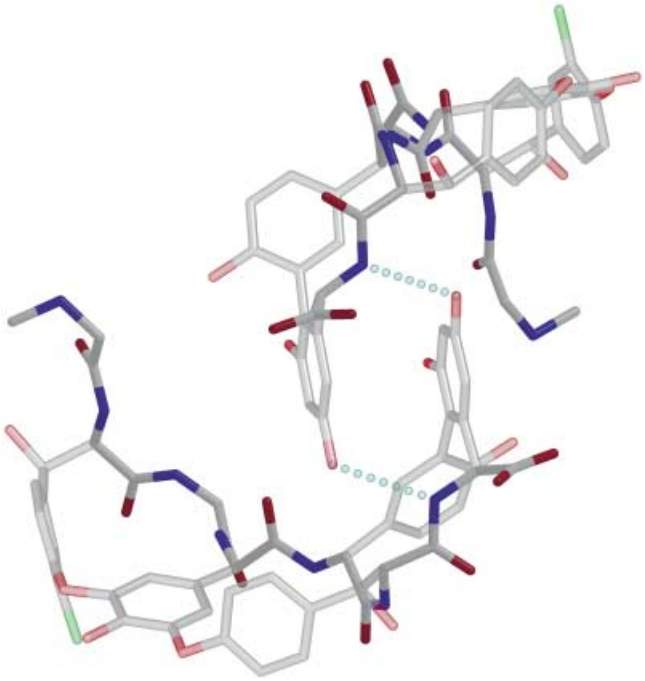


Fig. 4. *Additional intermolecular interactions observed in forms I and II*

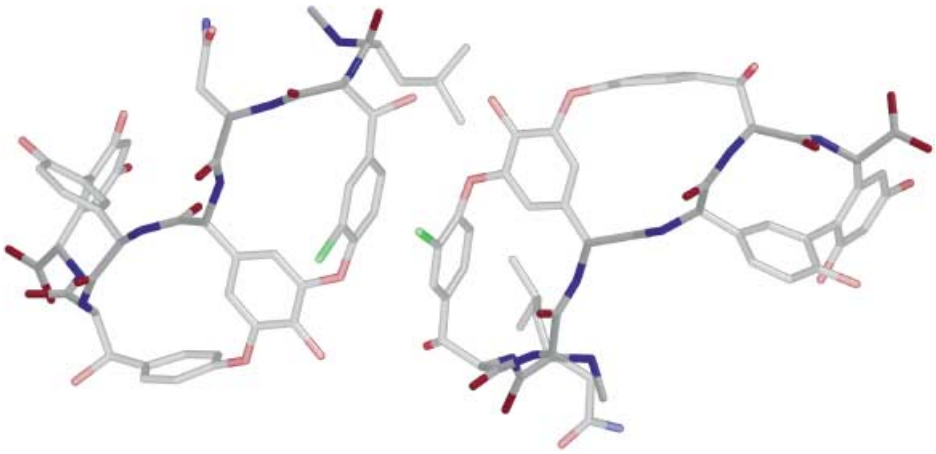


Fig. 5. *Nonpolar intermolecular interactions between the two intertwined* 6_1 *helices in form III generated by the action of a crystallographic twofold axis.* There are similar interactions between the two helices in form IV.

Conclusions. – The high solvent content was almost certainly partly responsible for the very high quality of the model-free experimental maps for the two hexagonal forms. In fact, atomic resolution data and *ab initio* direct methods could achieve only a

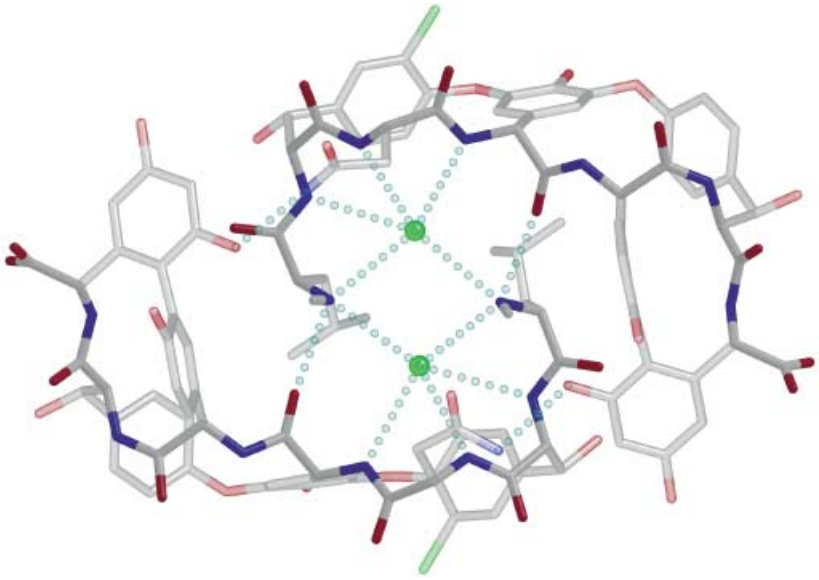


Fig. 6. *Two chloride sites in the binding pocket in form IV.* The other two chlorides have a similar environment; the corresponding sites are mixed bromide/chloride sites in form III.

weighted mean phase error of 21.8° for form I, inferior to the two hexagonal forms despite their much lower resolution. Although, for form IV, separate datasets were used for finding the anomalous differences and for phase expansion (the latter at a shorter wavelength to enable data to be collected to higher resolution), it is clear that a single data set would have sufficed, making the anomalous scattering of Cl potentially a very attractive approach to the solution of the phase problem [10][11]. After the solution of this structure with synchrotron data, we decided to take a closer look at the in-house data that we had collected to 2.0 Å to characterize the crystal with a rotating anode generator, multiplayer optics, and MAR345 image-plate detector (redundancy 15, mean (I/σ) 40). SHELXD found the eight Cl sites routinely after truncating the data to 2.6 Å; the sphere of influence density modification in SHELXE produced an easily interpretable map with a map correlation coefficient (against an isotropically refined model for the same data) of 0.908 and a weighted mean phase error of 17.5°.

The four crystal forms show remarkable variation in solvent content. The two hexagonal forms contain unusual interpenetrating double helices that enclose wide solvent channels, which may help to explain the stability of the crystals although they consist predominantly of liquid water.

We are grateful to the *Deutsche Forschungsgemeinschaft* (SFB416), to the *Fonds der Chemischen Industrie* and to the *EU* (contract HPRI-CT-1999-00017 to EMBL/DESY) for support.

Experimental Part

Monoclinic Form I. Crystals were obtained from a saturated decaplanin soln. containing 33% Li_2SO_4, 20% glycerol and 0.2M *Tris* buffer at pH 8.5, and data were collected on beamline *X9B* at *Brookhaven National Laboratory*. Integration and scaling were performed with *Denzo/Scalepack* [14]. Routine application of the *ab initio* direct methods program SHELXD [4] revealed an almost complete structure consisting of two homodimers; only the disordered isobutyl side chains were not well-defined. In addition to the four decaplanin (**1**) monomers, the final asymmetric unit contained three molecules of glycerol (with occupancies fixed at 1.0, 0.7 and 0.7) and 187H_2O molecules.

Monoclinic Form II. A sat. soln. of decaplanin in 1M sodium citrate at pH 7 was centrifuged at 10000 rpm to remove solid material, then allowed to stand in a 10-ml flask for several weeks at 0°. Colorless crystals with maximum dimensions up to 3 mm were obtained, but under a polarizing microscope all appeared severely fractured. Despite the very high mosaic spread, it proved possible to collect a data set with a *Bruker 1K SMART-CCD* detector mounted on an offset *Huber/Stoe* four-circle goniometer with a graphite monochromated MoK_α sealed tube as X-ray source and to process it with the *Bruker* programs SAINT, SADABS, and XPREP. Despite the relatively high resolution, the quality of the data was too poor for the application of direct methods, so the structure was solved by molecule replacement with the program EPMR [5] with a homodimer from form I as the

Table 1. *Crystal Data and Refinement Statistics*

Crystal form	I	II	III (MAD)	III (MAD)	IV (SAD)	IV (Native)
Crystal system	Monoclinic	Monoclinic	Hexagonal	Hexagonal	Hexagonal	Hexagonal
Space group	$P2_1$	$P2_1$	$P6_122$	$P6_122$	$P6_122$	$P6_122$
T [K]	100	133	100	100	100	100
Wavelength [Å]	0.9800	0.7107	0.9200[a])	0.9180[b])	1.5420	0.8345
Source	BNL X9B	MoK_α	BNL X9B	BNL X9B	BNL X9B	EMBL BW7B
Crystal dimensions [mm]	0.1,0.1,0.1	0.8,0.8,0.8	0.2,0.2,0.3	0.2,0.2,0.3	0.2,0.2,0.3	0.2,0.25,0.3
Resolution range [Å]	30.85 – 1.00	18.08 – 1.15	23.23 – 1.90	23.23 – 1.90	29.29 – 1.60	48.39 – 1.47
a [Å]	25.60(5)	28.95(3)	64.23(5)	64.36(5)	60.46(5)	60.10(6)
b [Å]	38.60(8)	31.91(3)	64.23(5)	64.36(5)	60.46(5)	60.10(6)
c [Å]	31.71(6)	34.73(4)	83.96(7)	84.13(7)	133.1(2)	131.4(2)
β [°]	105.88(3)	109.87(5)	90	90	90	90
V [Å^3]	30144	30170	299989	301826	421352	411004
Z	8	8	48	48	48	48
Reflections measured	104761	161613	86334	79699	367077	272910
Independent reflections	31754	21506	8544	8592	19774	24558
Reflections with $I > 2\sigma(I)$	27724	17395	7387	7718	18934	21985
Mean (I/σ)	12.2	12.3	21.5	18.5	36.2	23.6
R_{merge} (based on I)	4.9	9.8	7.9	11.2	5.4	5.8
Completeness [%]	98.5	99.3	99.7	99.5	99.6	99.1
Restraints	7612	6603		1847		6587
Number of variables	5658	5685		2175		4895
Final R_1 [%] ($I > 2\sigma$)	10.64	18.87		20.44		17.29
Final R_1 [%]	11.59	22.07		21.60		17.87
wR_2 [%]	27.37	45.63		53.20		43.46
Goodness-of-fit	2.154	3.759		3.015		2.277
Reflections for free R	1535	1056		449		1244
Free R [%]	13.81	26.32		24.69		21.22
$\Delta\rho$ (min, max) [e Å^{-3}]	– 0.33, 0.74	– 0.59, 0.92		– 0.05, 0.06		– 0.40, 0.50
Solvent content [%]	52	48		58		69
PDB Deposition code	1HH3	1HHC		1HHA		1HHF

[a]) Inflection point. [b]) High-energy remote wavelength.

search fragment. The program located two dimers with a correlation coefficient of 42.4% that were optimized by refinement as eight rigid groups. The final asymmetric unit contained four decaplanin (**1**) molecules, two complete citrate anions, one three-atom fragment (possibly of a citrate) and 164 H_2O molecules.

Hexagonal Form III. Crystals were obtained from a sat. soln. of **1** containing 44% $MgSO_4$ and 0.1M *Tris* buffer at pH 8.5, and soaked briefly in 1M NaBr to incorporate Br^- ions. Data were collected at two wavelengths (high-energy remote and inflection point) close to the Br absorption edge and processed with *Denzo/Scalepack* followed by XPREP to prepare the MAD analysis that was discussed above. The final asymmetric unit contained four monomers of **1**, four mixed bromide/chloride sites, a glycerol molecule disordered across a twofold axis and 182 H_2O molecules.

Hexagonal Form IV. A sat. soln. of **1** was centrifuged at 10000 rpm to remove insoluble material, then a 2M soln. of $NH_4H_2PO_4$ in 0.1M *Tris* buffer at pH 8.5 added. Hexagonal bipyramidal crystals grew over the course of 1–3 months, and further crystals could be obtained more quickly by micro-seeding. A native data set was collected to 1.47 Å on beamline *BW7B* at EMBL/DESY at a wavelength of 0.8345 Å, and a data set for exploiting the anomalous scattering of the Cl naturally present in **1** was collected to 1.60 Å at a wavelength of 1.542 Å on *X9B* at *Brookhaven National Laboratory*. After structure solution by the SAD method (see above) and refinement, it was possible to localize four monomers of **1**, four Cl^- ions, two phosphates with occupancies fixed at 0.4 and 93 H_2O molecules in the asymmetric unit. In view of the solvent content of nearly 70%, this means that a large fraction of the scattering matter was present in the form of disordered (liquid) solvent not accounted for directly in the refined model.

Structure Refinements. For all refinements, standard restraints were employed for bond lengths, 1,3-distances, planarity, and displacement parameters using SHELXL-97 [15]. For crystal form III, only the covalent Cl-atoms were refined anisotropically, the remaining atoms were isotropic; for the other three crystal forms, all non-H-atoms were refined anisotropically. A two-parameter bulk-solvent model [16] was refined in all cases. Further details of the refinements are given in the *Table*. The *R* values *etc.* are typical of macromolecules and would, of course, be unacceptable for small molecules; however, it should be noted that the unit-cell volumes, solvent contents, and data resolution are in the normal range for medium-sized protein structures! In view of the solvent contents in the range 50–70% (high even for proteins), a large fraction of the scattering power is not accounted for in the refined model, so it is not possible to determine a meaningful density *etc.* from the refined model. The appreciable radiation damage must also have been detrimental to the *R* values. The refined coordinates have been deposited with the *Protein Data Bank* (www.rcsb.org/pdb), and the entry codes are given in the *Table*.

REFERENCES

[1] H. C. Neu, *Science* **1992**, *257*, 1064.
[2] Hoechst AG, Europ. Pat. 356894, 1989; *Hoechst AG*, Ger. Offen. DE 3909056 A1, 1990; *Hoechst AG*, Europ. Pat. 437846, 1990.
[3] S. J. Box, A. L. Elson, M. L. Gilpin, D. J. Winstanley, *J. Antibiotics* **1990**, *43*, 931.
[4] I. Usón, G. M. Sheldrick, *Curr. Opin. Struct. Biol.* **1999**, *9*, 643; G. M. Sheldrick, H. A. Hauptman, C. M. Weeks, R. M. Miller, I. Usón, in 'International Tables for Crystallography', Eds. E. Arnold and M. Rossmann, Kluwer Academic Publishers, Dordrecht, 2001, Vol. F, p. 333; T. R. Schneider, G. M. Sheldrick, *Acta Crystallogr, Sect. D.* **2002**, *58,* 1772.
[5] C. R. Kissinger, D. K. Gelhaar, D. B. Fogel, *Acta Crystallogr., Sect. D* **1999**, *55*, 484; C. R. Kissinger, D. K. Gehlhaar, B. A. Smith, D. Bouzida, *Acta Crystallogr., Sect. D* **2001**, *57*, 1474.
[6] W. A. Hendrickson, *Science* **1991**, *254*, 51; J. L. Smith, in 'Direct Methods for Solving Macromolecular Structures', Ed. S. Fortier, Kluwer Academic Publishers, Dordrecht, 1998, p. 211.
[7] W. A. Hendrickson, M. M. Teeter, *Nature* **1981**, *290*, 107; Z. Dauter, M. Dauter, E. de La Fortelle, G. Bricogne, G. M. Sheldrick, *J. Mol. Biol*. **1999**, *289*, 83.
[8] G. M. Sheldrick, *Z. Kristallogr*. **2002**, *217*, 644.
[9] G. Strahs, J. Kraut, *J. Mol. Biol.* **1968**, *35*, 503.
[10] P. J. Loll, *Acta Crystallogr., Sect. D* **2001**, *57*, 977.
[11] C. Lehmann, G. Bunkóczi, L. Vértesy, G. M. Sheldrick, *J. Mol. Biol.* **2002**, *318*, 723.
[12] T. R. Schneider, *Acta Crystallogr., Sect. D* **2000**, *56*, 714.

[13] M. Schäfer, T. R. Schneider, G. M. Sheldrick, *Structure* **1996**, *4*, 1509; M. Schäfer, G. M. Sheldrick, T. R. Schneider, L. Vértesy, *Acta Crystallogr., Sect. D* **1998**, *54*, 175; J. P. Mackay, U. Gerhard, D. A. Beauregard, M. S. Westwell, M. S. Searle, D. H. Williams, *J. Am. Chem. Soc.* **1994**, *116*, 4581.
[14] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307.
[15] G. M. Sheldrick, T. R. Schneider, *Methods Enzymol.* **1997**, *277*, 319.
[16] P. C. Moews, R. H. Kretsinger, *J. Mol. Biol*. **1975**, *91*, 201.

Received March 18, 2003