Acta Cryst. (1996). D**52**, 1082–1091

Refined High-Resolution Structure of the Metal-Ion Dependent L-Fuculose-1-phosphate Aldolase (Class II) from *Escherichia coli*

MATTHIAS K. DREYER AND GEORG E. SCHULZ*

Institut für Organische Chemie und Biochemie, Albert-Ludwigs-Universität, Albertstrasse 21, D-79104 Freiburg im Breisgau, Germany. E-mail: schulz@bio5.chemie.uni-freiburg.de

(Received 29 April 1996; accepted 5 July 1996)

Abstract

The structure of the class II zinc-ion dependent L-fuculose-1-phosphate aldolase from *Escherichia coli* in its tetragonal crystal form has been established at 1.92 Å resolution. The homotetrameric enzyme has a molecular mass of 4×24 kDa and follows C_4 symmetry. The structure model is exactly symmetrical, which contradicts an observed birefringence anomaly of the crystals. The four catalytic centers are located in deep clefts at the interfaces of adjacent subunits. The zinc ion is coordinated by three histidines and one glutamate in an almost tetrahedral arrangement. In contrast to numerous other catalytically competent zinc ions, there is no water molecule in the ligand sphere. Replacement of zinc by a cobalt ion caused only small structural changes. A search through the Protein Data Bank indicated that the chain fold is novel. Sequence homology searches revealed a significant similarity to the bacterial L-ribulose-5-phosphate 4-epimerase.

1. Abbreviations

DHAP, dihydroxyacetone phosphate; FruA, fructose-1,6-bisphosphate aldolase; FucA, L-fuculose-1-phosphate aldolase; Fuc1P, L-fuculose-1-phosphate; RibE, L-ribulose-5-phosphate 4-epimerase; PCR, polymerase chain reaction.

2. Introduction

Several important reactions of the carbohydrate metabolism are catalyzed by aldolases. These enzymes have been subdivided into two groups, designated class I and class II (Horecker, Tsolas & Lai, 1972), with differing catalytic mechanisms. In class I enzymes, a lysine ε-amino group forms a Schiff base intermediate with the carbonyl group of the substrate which then functions as an electron sink for the cleavage reaction. In contrast, the electron sink of the class II aldolases is a divalent metal ion. While class I enzymes have mostly been observed in eukaryotes, those of class II seem to be more abundant in prokaryotic organisms (Marsh & Lebherz, 1992).

Several structures of class I enzymes are known, for instance the 2-keto-3-deoxy-6-phosphogluconate aldolase from *Pseudomonas putida* (Mavridis, Hatada, Tulinsky & Lebioda, 1982), the *N*-acetylneuraminate lyase from *Escherichia coli* (Izard, Lawrence, Malby, Lilly & Colman, 1994), and the fructose-1,6-bisphosphate aldolases (FruA) from rabbit muscle (Sygusch, Beaudry & Allaire, 1987), from human muscle (Gamblin *et al.*, 1990) as well as from *Drosophila melanogaster* (Hester *et al.*, 1991). All these class I enzymes possess a $(\beta\alpha)_8$-barrel as a central motif. Modeling of bound substrates together with mutagenetic studies gave rise to a detailed proposal for the enzymatic mechanism of class I FruA (Gamblin *et al.*, 1991; Littlechild & Watson, 1993).

Much less structural information is available on class II aldolases. There exists a preliminary account of the L-fuculose-1-phosphate aldolase (E.C. 4.1.2.17, FucA) structure (Dreyer & Schulz, 1993) as well as reports on the crystallization of class II FruA from *E. coli* (Naismith *et al.*, 1992; Kitagawa *et al.*, 1995). Mechanistic studies have been performed mostly on the class II FruA from yeast (Riordan & Christen, 1969; Belasco & Knowles, 1983; Kadonaga & Knowles, 1983) and from *E. coli* (Scamuffa & Caprioli, 1980; Szwergold, Ugurbil & Brown, 1995).

Aldolases have attracted attention as catalysts for stereoselective aldol additions in organic syntheses giving rise to novel carbohydrates and derivatives therefrom (Bednarski *et al.*, 1989; Ozaki, Toone, Osten, Sinskey & Whitesides, 1990; Fessner *et al.*, 1991; Wong, Halcomb, Ichikawa & Kajimoto, 1995*a*,*b*). These compounds become more and more important because carbohydrates are involved in a number of recognition processes such as cell adhesion, metastasis or immune response.

We are here concerned with FucA, catalyzing the cleavage of L-fuculose-1-phosphate (Fuc1P) to dihydroxyacetone phosphate (DHAP) and L-lactaldehyde during the bacterial degradation of fucose (Ghalambor & Heath, 1962). Like several other DHAP-converting aldolases the enzyme is highly specific for DHAP, but accepts a wide range of aldehyde components with varying stereoselectivity (Ozaki *et al.*, 1990; Fessner *et al.*, 1991). The enzyme is a homotetramer with subunits consisting of 215 amino-acid residues ($M_r = 23\,775$) and one zinc ion. We present the refined wild-type structure at 1.92 Å resolution together with the structure

after exchanging zinc for cobalt. The aim of this work is to provide a basis for protein-engineering studies broadening the substrate spectrum of this enzyme and thus opening new possibilities for organic syntheses.

3. Materials and methods

3.1. *Subcloning and overexpression*

The structural gene fucA was amplified by PCR from plasmid pfuc41a (Chen, Zhu & Lin, 1987) using two mismatching oligonucleotides in order to introduce an *Eco*RI restriction site 20 bp before the start codon and a *Pst*I site 15 bp behind the stop codon. The sequences of the oligonucleotides were 5′-CTGTTTT*GAATTC*AGAGAGAGGTAG-3′ and 5′-CTTCTCCTTG***CTGCAG***TACGAAATTAC-3′ (restriction sites italic, mismatches in bold). The amplified fragment was then ligated into the multicloning site of the expression vector pKK223-3 (Brosius & Holy, 1984). The resulting vector pKKFA2 was more efficient than a previous construct where only the Shine–Dalgarno sequence of pKK223-3, but not that of the fucA gene was used. After transforming the vector into *E. coli* strain JM105 a 1 l overnight bacterial cell culture grown on LB-broth yielded about 7.5 g cells containing 1.1 g of FucA. The high overexpression allowed us to take generously only the highest parts of the peaks during the two-step purification procedure on DEAE–sepharose and subsequent gel-permeation chromatography, resulting in about 150 mg crystallizable protein. The specific activity of the purified enzyme was 17 Units mg^{-1} (Schneider, 1994). All handling was carried out in 20 m*M* Tris–HCl buffer pH 7.6 containing 0.5 m*M* $ZnCl_2$ and 10 m*M* β-mercaptoethanol. The latter is necessary to maintain the activity.

3.2. *Crystallization and crystal properties*

Crystals were grown at 293 K in hanging drops. Drops were prepared by mixing equal volumes of a protein solution of 10 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.6, 0.5 m*M* $ZnCl_2$, 10 m*M* β-mercaptoethanol and reservoir buffer. The reservoir buffer contained 1.62 *M* ammonium sulfate, 20 m*M* potassium phosphate pH 7.8, 4 m*M* $ZnCl_2$, 4 m*M* β-mercaptoethanol. Under these conditions we obtained one tetragonal (*T*) and one cubic (*K*) crystal form. Upon addition of 4% 2-methyl-2,4-pentanediol to the reservoir solution prior to drop preparation, another cubic crystal form (*W*) was observed. Only crystal form *T* diffracted to high resolution, while the cubic crystal forms showed only medium-resolution patterns. The crystals of form *T* were square platelets with a maximum size of 600 × 600 × 150 µm and were stable in reservoir buffer.

Metal ion exchange from zinc to cobalt in the tetragonal crystal form was achieved by washing a crystal twice with reservoir buffer lacking $ZnCl_2$, then adding EDTA to 15 m*M* and incubating for 15 h. The crystal was then transferred to a 10 m*M* EDTA solution for 7 h and finally soaked with reservoir buffer containing 10 m*M* $CoCl_2$ instead of zinc for 9 d.

Table 1. *Data-collection statistics*

	FucA (Zn)	FucA (Co)
Resolution (Å)	10.0–1.92	10.0–2.0
No. of crystals	3	1
Unique reflections	14541	11886
Redundancy factor	9	3
Overall		
Completeness (%)	97	89
R_{sym}* (%)	6.5	6.1
Last shell (Å)	1.96–1.92	2.04–2.00
Completeness (%)	93	80
R_{sym}* (%)	28.4	15.9

* $R_{sym} = \sum_{\mathbf{h}} \sum_i |I(\mathbf{h})_i - \langle I(\mathbf{h})\rangle| / \sum_{\mathbf{h}} \sum_i I(\mathbf{h})_i$.

3.3. *Data collection and refinement*

All data were collected with a multiwire area detector (Siemens X-1000) on a rotating anode (Rigaku RU200B). The refinement started from an earlier model containing residues 3–206 (Dreyer & Schulz, 1993). The resolution was extended to 1.92 Å, where still one third of the reflections were above 3σ. The data were processed with program *XDS* (Kabsch, 1988). Data statistics are given in Table 1. For model building we used program *O* (Jones, Zou, Cowan & Kjeldgaard, 1991). The refinement was carried out with *X-PLOR* (3.0)

Fig. 1. The tetragonal crystal form *T* in a plastic dish viewed exactly along the fourfold axis under the polarizing microscope where no birefringence should be expected. Still, the two pairs of opposite triangles change colors differently upon rotation. Under glass, the whole crystal is uniformly dark at 0 and 90° rotation and uniformly bright at 45°. At angles between these positions, the pairs of opposite triangles show different brightnesses (different colors under plastic). All faces of the crystals are optically flat, *i.e.* there is no pyramidal growth.

(Brünger, 1992), considering the zinc ligands His92, His94 and His155 as uncharged and applying no restrictions for the coordinative bonds. *X-PLOR* does not permit anisotropic refinement of a selected atom like the zinc ion. Amino-acid residues 78, 96, 107, 158 and 160 showed additional densities in the $(2F_o - F_c)$ and the $(F_o - F_c)$ maps that were fitted by assuming alternative side-chain conformations and refining relative occupancies. Water molecules were incorporated manually at positions with $(F_o - F_c)$ density above 3σ and satisfactory hydrogen-bond geometry.

The refined zinc-containing structure served as the model for the calculation of the initial $(2F_o - F_c)$ and $(F_o - F_c)$ density maps of the cobalt-containing structure, which was subsequently extended and refined. During refinement all parts including the water molecules were checked. It should be noted that the FucA(Co) structure was derived from one crystal whereas the FucA(Zn) structure is based on three crystals. If there were some crystal heterogeneity at the C-terminus, the chance of finding a more uniform structure would be higher with FucA(Co), which may explain the density for residues 207–210 in this structure.

Table 2. *Characteristics of the refined structures*

	FucA (Zn)	FucA (Co)
Resolution (Å)	10.0–1.92	10.0–2.0
R factor (%)	18.6	18.0
Residues in model	1–206	1–210
Non-H polypeptide atoms*	1597	1629
Water molecules†	119	108
Sulfate ions	2	2
Metal ion	Zn^{2+}	Co^{2+}
Luzzati (1952) coordinate error (Å)	0.22	0.23
Read (1986) σ_A coordinate error (Å)	0.23	0.22
R.m.s. deviations from standard		
For bond lengths (Å)	0.013	0.013
For bond angles (°)	1.7	1.7
Mean *B* factors		
Main chain (Å^2)	17	16
Side chain (Å^2)	21	21
Water molecules (Å^2)	37	33
R.m.s. *B*-factor difference		
Along bonds (Å^2)	2.0	2.3
Along angles (Å^2)	3.1	3.6

* The bound β-mercaptoethanol is included. † 91 water molecules were at equivalent positions in the zinc and cobalt enzymes, as defined by a displacement below 1.0 Å.

4. Results and discussion

4.1. *Crystallization and structure determination*

The homotetrameric enzyme crystallized in three forms: form *T* with space group $P42_12$ and axes $93.7 \times 93.7 \times 42.8$ Å, form *K* with space group *F*432 and axes of 183 Å (Dreyer & Schulz, 1996), and form *W* with space group *P*432 and axes of 144 Å. All forms grew in 1.6 *M* ammonium sulfate at pH 7.8. Crystal form *T* showed the diffraction pattern with the highest resolution. It contains one subunit of FucA per asymmetric unit and 38% solvent. The structure had been solved by multiple isomorphous replacement (Dreyer & Schulz, 1993). Data collection has now been extended to higher resolution, and the structure has been refined to convergence. The exchange of Zn^{2+} for Co^{2+} was performed in the crystals by washing with EDTA and subsequent soaking. The structural refinement of the cobalt-containing enzyme started from the closely resembling zinc enzyme. An overview of the structural results is given in Table 2.

Examination of crystal form *T* under the polarizing microscope led to an unexpected and still unexplained observation: crystals that were viewed along their fourfold axis exhibited birefringence in a checkered pattern (Fig. 1), although a fourfold axis is optically isotropic (Blundell & Johnson, 1976). This indicates small deviations from perfect fourfold symmetry that are different in the two types of triangular domains of Fig. 1. No such indication, however, could be detected during structure analysis. The R_{sym} values, in particular of the cobalt-substituted enzyme (Table 1), do not increase excessively at high resolution. Moreover, the refinement worked well at high resolution yielding a symmetric model. Therefore, twinning can be excluded. The remaining asymmetry must either be located in the low-density regions (residues 23–27, 207–215), or it must have been accounted for by *B*-factor broadening. The mobile C-terminal segment 207–215, however, is not the major source of asymmetry because crystals of a mutant lacking this segment showed the same birefringence pattern. Therefore, we suggest that the optical asymmetry is

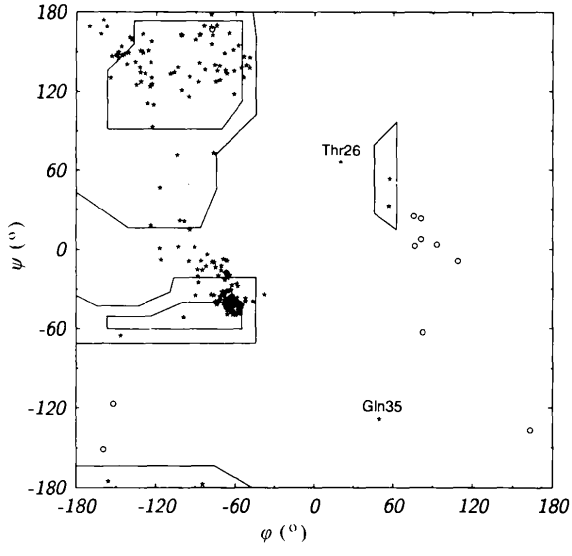

Fig. 2. Ramachandran plot showing the main-chain dihedral angles (φ,ψ) of the 192 non-Gly (⋆) and 12 Gly (○) residues in the model of zinc-containing FucA. Only Thr26 in a poorly defined loop and Gln35 in a well defined type II′ turn are in unfavored regions.

caused by displacements much below 1 Å hidden in the *B*-factor distribution of the X-ray analysis (see below). No birefringence was observed with the crystals belonging to the cubic forms *K* and *W*.

The zinc-containing model consists of residues 1–206, the nine C-terminal residues 207–215 are mobile. Except for loop 23–27 at the interface, which could not be fitted to our satisfaction, all modeled backbone atoms and most of the side-chain atoms are in well defined density. Additional density at the side chain of Cys14 could be clearly interpreted as a covalently bound β-mercaptoethanol from the solvent. The real-space correlation coefficient (Brändén & Jones, 1990) is 0.93 for the main chain, and the coordinate-error estimates are around 0.23 Å (Table 2). After zinc had been replaced by cobalt, appropriate density appeared for four more residues at the C terminus, resulting in a model for residues 1–210. As the r.m.s. $\Delta C\alpha$ deviation between the zinc and cobalt structures is only 0.15 Å, they are identical within the limits of error.

As a further model-quality index, the distribution of backbone dihedral angles is plotted in Fig. 2. Most residues are in favorable regions, only Thr26 and Gln35 adopt less favored conformations. Thr26 lies in the loop with poor density, whereas Gln35 is in well defined density participating in a type II′ turn between strands $\beta 1$ and $\beta 2$. Residue 190 is the only *cis*-proline. The side-chain dihedral angles $\chi 1$ and $\chi 2$ of leucines and isoleucines follow the common distribution (Ponder & Richards, 1987). The side chains of Met78, Val96, Ser107, Ile158 and Cys160 were refined in two conformations, which in all five cases were nearly equally occupied.

4.2. *Chain fold and secondary structure*

The secondary structures of FucA are given together with the sequence in Fig. 3. A topology sketch had been presented earlier (Dreyer & Schulz, 1993). The refined model now shows the short helix $3_{10}B'$ in addition to the previous assignments. Each subunit contains a central

```
          1        10        20        30          40        50
          .        .         .        ↓.           .   ↓     .
FucA      MERNKLARQIIDTCLEMTRLGLNQGTAGNVS-VRYQDGM-LITPTGIPYEKLT
                     αA                  β1       β2      3_10B
               * **     *      *   * **** *    *   * * *  *   *
RibE       MLEDLKRQVLEANLALPKHNLVTLTWGNVSAVDRERGVFVIKPSGVDYSVMT

                   60         70        80        90       100
                   .          .↓↓↓      .         . ↓ ↓     .
FucA      ESHIVFID-GNGKHEEG-KLPSSEWRFHMAAYQSRPDANAVVHNHAVHCTAV
          3_10B'β3     β4            αC           β5        αD
              *      *   ** * ***    *   **  *     ** *  * *
RibE      ADDMVVVSIETGEVVEGTKKPSSDTPTHRLLYQAFPSIGGIVHTHSRHATIW

                 110       120       130       140       150
                  .  ↓      .         .         .         .
FucA      SILNRSIPAIHYMIAAAGGNSIPCAPYATFGTRELSEHVALALKNRKATLLQ
                 β6    3_10E   β7             αF          β8
               ****     *      ***         *              *
RibE      AQAGQSIPATGTTHADYFYGTIPCTR--KMTDAEINGEYEWETGNVIVETFE

               160                   170       180        190
           ↓    .                     .         .          .
FucA      HHGLIACE-----VNLEKAL-W------LAHEVEVLAQL-YLTTLAITDPVP
              β9                        αG
            *  *       *       *        *   **    *          *
RibE      KQGIDAAQMPGVLVHSHGPFAWGKNAEDAVHNAIVLEEVAYMGIF-CRQLAP

                200        210    215
                 .          .      .
FucA      VLSDEEIAVVLEKF-KTYGLR--IEE
                   αH
           * *              *
RibE      QLPDMQQTLLDKHYLRKHGAKAYYGQ
```

Fig. 3. Amino-acid sequence of one FucA subunit (Chen, Lu & Lin, 1989) aligned to the sequence of L-ribulose-5-phosphate 4-epimerase (RibE) from *E. coli* comprising residues numbers 1 through 231 (Mineno *et al.*, 1990). It should be noted that a remark on this alignment by Dreyer & Schulz (1993) quoted incorrect residue numbers of RibE that had entered a sequence data bank. Secondary structures for FucA (underlined) were assigned manually according to the hydrogen-bond list of *X-PLOR* using the 3.5 Å distance criterion and N—H···O angles above 120°. The sequences were aligned with program *MSA* (Lipman, Altschul & Kececioglu, 1989); identities are indicated by *. Arrows point to all residues which bind the active center sulfate or the zinc ion. There are 34% identical amino acids within the N-terminal 125 residues of FucA. The alignment of the C-terminal part is of lower significance (13% identity).

β-sheet that can be subdivided into seven antiparallel strands (β4, β3, β2, β1, β5, β9 and β8) and two additional strands (β7 and β6, Fig. 4). The antiparallel moiety fits the curvature of a β-barrel (Chothia, 1973) that is as regular as those of porins (Weiss, Wacker, Weckesser, Welte & Schulz, 1990), whereas strands β7 and β6 deviate strongly from this curvature.

As illustrated in Fig. 5, the β-sheet is sandwiched between two layers of helices. The resulting chain fold belongs to the less frequent type with an antiparallel (instead of a parallel) β-sheet covered by helices on both sides. The C-terminal helix αH protrudes from the compact structure and participates in the interface of the FucA tetramer. As a three-dimensional comparison with 558 representative protein structures from the Protein Data Bank using program *DALI* (Holm & Sander, 1993) showed no significant signal, we suggest that FucA is the first member of a new structural family.

FucA forms seven salt bridges within each chain. Most remarkable is the salt bridge between the extended side chains of Arg3 and Glu65 that fastens the N terminus across the 'top' of the molecule (Fig. 5) to the end of strand β4.

4.3. *Chain mobility*

The *B*-factor distributions of the zinc- and the cobalt-containing enzyme are given in Fig. 6. The poorly defined loop 23–27 assumes the highest values within the chain. Minimal values are found around residues 97 and 156 that are close to the metal-binding site and around residue 175 in the long helix αG. The N-terminal half of the protein is in general more mobile than the C-terminal

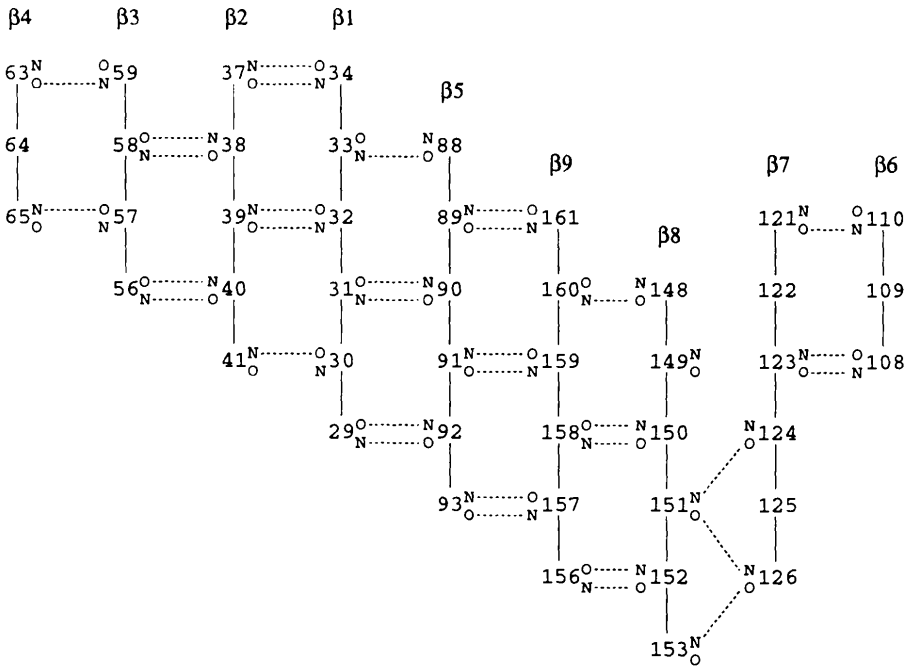

Fig. 4. Schematic representation of the central β-sheet. Dashed lines indicate all hydrogen bonds with donor–acceptor distances up to 3.5 Å and donor—H···acceptor angles above 120°. Except for β7 and β8, all neighboring strands are antiparallel.

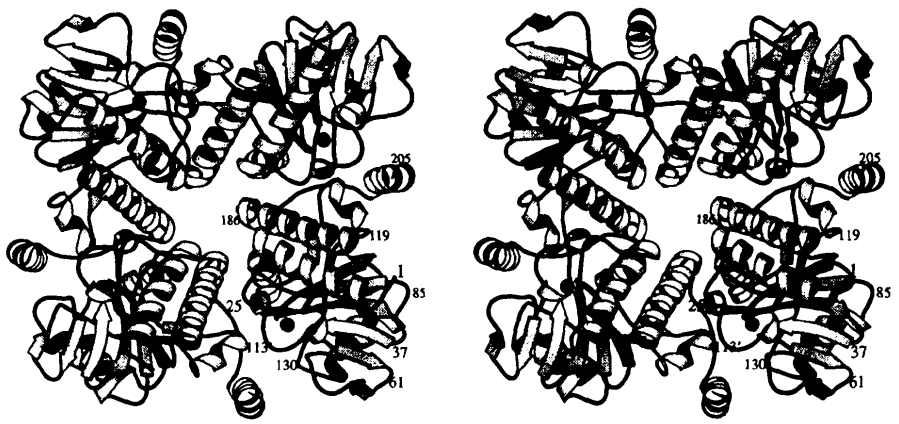

Fig. 5. Stereo ribbon diagram of the FucA tetramer, viewing approximately down the molecular fourfold axis. Black dots indicate the zinc ions at the subunit interfaces. Some residue positions are labeled.

half. The mean B factors of the $C\alpha$ atoms of residues 1–96 and 97–206 are 22 and 13 Å^2, respectively, the average for the whole chain is 17 Å^2. The C-terminal part is involved in the interface of the C_4 tetramer, whereas the N-terminal moiety is generally further apart from the tetramer center, except for loop 23–27 (Fig. 5). This mobile loop is at the interface and also near to the active site, where it may play a role in catalysis. As the B factor increases with the distance from the center, it is quite possible that the N-terminal moiety causes the anomalous birefringence along the fourfold axis (Fig. 1) by slightly asymmetric associations in the crystal.

4.4. *Solvent structure and cavities*

In the zinc enzyme we assigned 119 water molecules with $(2F_o - F_c)$ densities above 1σ and/or $(F_o - F_c)$ densities above 3σ at reasonable hydrogen-bond geometries. This amounts to 0.6 water molecules per residue. Their mean B factor is 37 Å^2. The highest observed $(2F_o - F_c)$ density for a water molecule is 4.0σ. Furthermore, there are two sulfate ions, one of which occupies the putative phosphate binding site in the active center, while the other one participates in a crystal contact.

A search for cavities with program *MSP* (Connolly, 1993; probe radius 1.2 Å) revealed two water-filled and a number of empty cavities. The biggest cavity is filled with four water molecules. It extends over a distance of 15 Å from Thr150 to Val202′ and has a volume of 138 Å^3. Another water-filled cavity neighbors Asn93 and has a volume of 36 Å^3, sufficient for one water molecule. These five buried water molecules have a mean B factor of 17 Å^2 like the backbone. The largest empty cavity has a volume of 70 Å^3 and is lined by a predominantly nonpolar surface at Ile123. The total volume of all cavities amounts to 530 Å^3 which is 2.1% of the molecular volume.

Table 3. *Interactions between the subunits*

Contacting residues*	Hydrogen bonds†	
	Atoms	Distance (Å)
24–26, 48–49, 73, 94, 96–97, 99–100, 103–104, 128, 130–133, 154–155, 182, 185–186	Gly25 N···His172′ NE2	3.4
	Tyr128 OH···Glu198′ OE1	3.1
	Thr130 OG1···Glu198′ OE1	2.8
111′–113′, 169′, 172′–173′, 176′, 179′–180′, 183′–184′, 187′–188′, 190′–194′, 198′, 201′–202′, 205′–206′	Phe131 O···Lys205′ NZ	3.4
	His154 ND1···Ile111′ O	2.6
	His155 O···His112′ NE2	2.9

*Considering all interactions up to 4.5 Å. †Donor–acceptor distances up to 3.5 Å.

4.5. *Quarternary structure*

Two types of homotetramers can be carved out of the crystal packing: a D_2 symmetry assembly with a total interface area (*i.e.* half the total buried surface area) of 2360 Å^2 and a C_4 symmetry assembly with 4200 Å^2. Because the total buried surface area of the C_4 tetramer is almost twice as large as that of the D_2 tetramer, we conclude that the rare C_4 symmetry is actually assumed in solution. This has been confirmed by the finding that Tyr113′ is important for catalysis (A. Jörger, personal communication), because only the C_4 arrangement brings this residue close to Zn^{2+}.

There are six hydrogen bonds connecting the subunits of the C_4 tetramer, four of them are between main- and side-chain atoms and two involve side chains only (Table 3). Altogether, 23 residues of each subunit form one

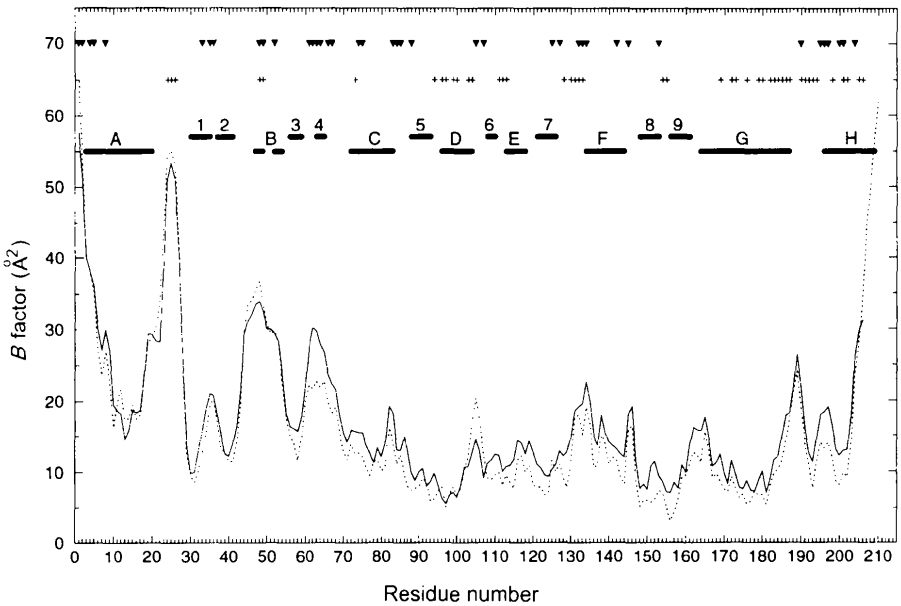

Fig. 6. B-factor distribution of a FucA subunit. Letters and numbers above the black bars indicate helices and β-sheet strands as defined in Fig. 3. Interface residues (+) and crystal contact residues (triangles) are marked. Note that the mobile loop 23–27 is at an interface contact, whereas most other interface contacts are in the C-terminal moiety.

Table 4. *Crystal packing*

Contacting subunits*	Contact area† (Å^2)		Contacting residues‡
I–II	700	*a–a*	*a* = 1, 33, 35–36, **61**, 83–84, **85**, **88**, 145, 195–196, **197**, 200–201, **204**
I–III§	490	*b–b*	*b* = 61–64, **66**, 67, 74–75, 132–133, **134**
I–IV	210	*c–d*	*c* = 2, 4–5, **8**, **52**
			d = **105**, 107, **125**, 127, 142, 153
I–V	160	*e–f*	*e* = 190; *f* = 48, 49
I–VI	80	*g–h*	*g* = **66**; *h* = **205**
I–VII	30	*i–i*	*i* = 61

* Specified in Fig. 7. † Buried solvent-accessible area (*X-PLOR*, probe radius 1.4 Å) per subunit. ‡ Contacting residues contain atoms at distances below 4.5 Å, residues in bold form hydrogen bonds. The contacts formed by residue groups *c–d*, *e–f* and *g–h* occur twice. § This contact contains two symmetry-related sulfates from the solvent.

intersubunit contact. Moreover, the interface contains the cavity with four water molecules close to the catalytic center (see below in Fig. 8).

4.6. *Crystal packing*

In crystal form *T* (space group $P42_12$) C_4 tetramers are stacked along the fourfold axes with adjacent stacks running in opposite directions. The packing is fixed by four major contacts. The two strongest contacts, I–II and I–III, are around the 222 center of symmetry in the *x,y* plane (Table 4, Fig. 7). Contact I–III is stabilized by two symmetry-related sulfate ions that form hydrogen bonds to Trp74, Arg75 and Gly67′ as well as to water molecules 452, 460, 487, 489′ and 504′. Contacts I–IV and I–V are much weaker and maintain the *z*-stacking. The small contact I–VI involves the C-terminal end of the visible structure and forms the only intermolecular salt bridge, Glu66 OE1 to Lys205′ NZ.

These findings explain why the tetragonal crystals are unstable in buffers that do not contain sulfate ions. They also account for the crystal habit: the strong contacts in the *x,y* plane promote crystal growth in this plane, whereas the weak contacts in *z* direction give rise to slower growth, resulting in the observed square platelets.

4.7. *The active site*

The active center of FucA is located in a deep cleft at the interface between two subunits. The zinc ion at the bottom of this cleft is ligated by three histidines and one glutamate (Fig. 8). Tyr113′ OH of the adjacent subunit points to the zinc ion but is located at a distance of 3.5 Å. In contrast to most other ligand spheres of catalytically active zinc ions (Vallee & Auld, 1992) there is no associated water molecule. The nearest water molecule is at a distance of 4.0 Å and contacts Sul300 O4 at the putative phosphate site (Fig. 8). Quantitative data on the coordination sphere are given in Table 5. All three histidines contact Zn^{2+} with their NE2 atoms. In agreement with the observed distribution (Chakrabarti, 1990*a*,*b*) Zn^{2+} lies close to the planes of the imidazole rings and the carboxylate.

The average of all six ligand···Zn^{2+}···ligand angles (defined by the NE2 atoms and Glu73 CD) is 109° like in a symmetric tetrahedron, the mean deviation from this average is 9°. The bond lengths between Zn^{2+} and the NE2 atoms lie around the average values, while the 1.9 Å distance to OE1 is at the lower end of the observed range. Even after the final round of refinement, there is still positive difference density of approximately 3.3σ on both sides of the zinc ion along an axis through Zn^{2+} and His155 NE2. We consider this as a consequence of the isotropy in the *B*-factor refinement enforced by the limited resolution, that fails to model the apparently asymmetric mobility of Zn^{2+}, which is conceivably connected to catalysis. In any case, the additional Zn^{2+} density should not relate to the slight disorder indicated by the birefringence anomaly, because the environment of Zn^{2+} has very low *B* factors.

Upon replacing zinc by cobalt, no significant changes in the active-center geometry were observed, except for Glu73 becoming a bidental ligand (Table 5). All ligand···cobalt···ligand angles are essentially the same as with zinc. The crystals of the cobalt-containing enzyme, however, showed density for residues 207–210,

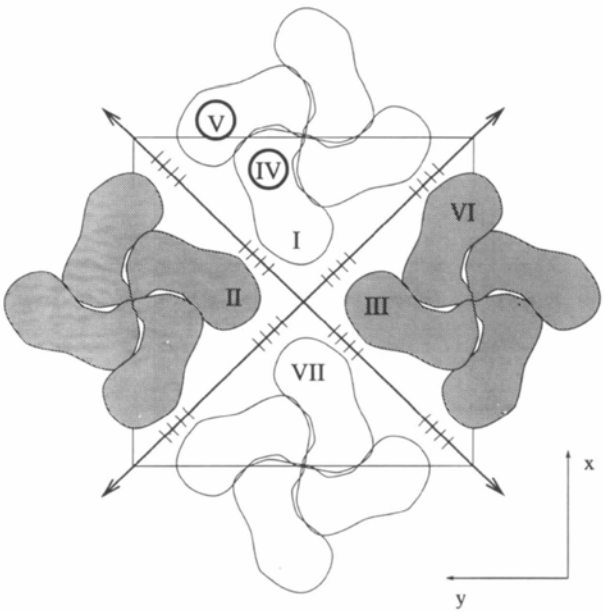

Fig. 7. Crystal packing of FucA. Blank and shaded tetramers are stacked around the crystallographic fourfold axes and run in opposite directions along *z*. Subunits IV and V (encircled numbers) are in the next layer along *z*. Thin lines represent the unit cell and arrows the twofold axes. The crystals form flat square platelets, where the square lies diagonally in the *x,y*-plane and the short dimension is along *z* towards the viewer. Taking subunit I as the reference at [*x*, *y*, *z*] the others are: II [*y* − 1, *x* + 1, −*z* + 1], III [−*y*, −*x*, −*z* + 1], IV [*x*, *y*, *z* + 1], V [$y-\frac{1}{2}$, $-x+\frac{1}{2}$, $z+1$], VI [$-x-\frac{1}{2}$, $y-\frac{1}{2}$, $-z+1$], VII [−*x* − 1, −*y* + 1, *z*].

Table 5. *Metal coordination in the active center*

All values in parentheses concern the cobalt-containing enzyme.

Metal or ligand	Distance (Å)	*B* factor (Å^2)	θ* ( )	φ* ( )
Zn (Co)		13 (7)		
Glu73 OE1	1.9 (2.2)	15 (15)	88 (84)	80 (90)
Glu73 OE2	2.4 (2.3)	14 (10)		
His92 NE2	2.1 (2.1)	8 (7)	89 (85)	6 (6)
His94 NE2	2.1 (1.9)	17 (11)	72 (79)	4 (8)
His155 NE2	2.0 (2.0)	6 (5)	87 (84)	2 (1)
Tyr113′ OH	3.5 (3.5)	11 (9)		

* θ and φ are defined according to Chakrabarti (1990*a*,*b*), *i.e.* θ is the deviation of the ligand···metal bond from the plane of the carboxylate or the imidazole ($\theta = 90°$ is in the plane); φ states the direction of the projection of the ligand···metal bond onto this plane ($\varphi = 0$ is along the CD—OE1 bond for glutamate and along the CD1—NE1—CD2 bisector for histidine).

elongating the C-terminal helix αH (Fig. 8). It should be noted that the oxidation state of cobalt at the time of the X-ray data collection is not known. Furthermore, the structures provide no explanation for the observed 30% higher activity of the cobalt-containing as compared to the zinc-containing enzyme (Schneider, 1994). Similar activity increases after substituting a catalytic zinc ion by a cobalt ion have been observed with thermolysin, dihydro-orotase and astacin (Holland, Hausrath, Juers & Matthews, 1995; Brown & Collins, 1991; Gomis-Rüth *et al.*, 1994).

The active center accomodates a sulfate ion bound by the polypeptide without direct charge compensation (Fig. 8), the distance between sulfur and Zn^{2+} being 7.4 Å. Presumably, the sulfate occupies the binding site for the phosphate of Fuc1P. This explains why a number of soaking experiments with substrates and substrate analogues failed. In particular, we soaked crystals with 5 m*M* Fuc1P ($K_M = 0.7$ m*M*; Ghalambor & Heath, 1966), 10 m*M* DHAP ($K_M = 0.6$ m*M*; Schneider, 1994), 250 m*M* fucitol-1-phosphate ($K_i = 70$ m*M*; Dreyer, 1995), 500 m*M* 2,5-anhydro-6-deoxy-L-talitol-1-phosphate ($K_i = 22$ m*M*; Schneider, 1994) and 60 m*M* D,L-glycerol-1-phosphate ($K_i = 6$ m*M*; Schneider, 1994), but found no significant difference density with any of these compounds. Unfortunately, sulfate participates in the strongest packing contact (Fig. 7) and therefore cannot be avoided in crystal formation. Moreover, cocrystallization with a ligand was not an attractive alternative because crystallization required several weeks whereas sugar phosphates are hydrolyzed within hours to days (Richard, 1984). Soaking with the inhibitor phosphoglycolohydroxamate destroyed the crystals immediately, but worked with the more complicated cubic crystal form *K* (Dreyer & Schulz, 1996).

4.8. *Sequence homology*

The recently published complete genome sequence of *Haemophilus influenzae* (Fleischmann *et al.*, 1995) included genes HI0611 and HI1012 which were both designated *fucA*, because they show 65 and 28% identical amino acids with the FucA from *E. coli*, respectively. As all residues in the active center are conserved and as its gene is located close to the other genes of the fucose metabolism, the product of HI0611 most likely functions as a FucA. In contrast, HI1012 carries the exchanges Ser72→Lys and Tyr113→Phe that should strongly diminish FucA activity, because Ser72 participates in binding the substrate phosphate and Tyr113 is at the zinc ion. Moreover, HI1012 is located far away from the other fucose metabolism genes. We therefore suggest that HI1012 is not a FucA, but fulfills a different function.

The sequence homology between FucA and bacterial L-ribulose-5-phosphate 4-epimerase (RibE; Mineno, Fukui, Ishino, Kato & Shinagawa, 1990; Lin,

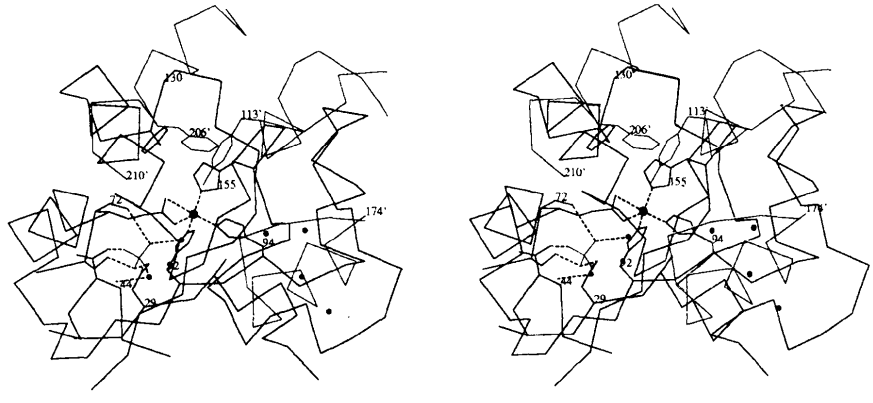

Fig. 8. Stereoview of the active center of FucA(Zn) at the subunit interface. Coordinative bonds to Zn^{2+} as well as hydrogen bonds to the sulfate ion (light blue) are represented by dashed lines. Residues of the two different subunits are colored black (residues 21–31, 40–46 and 70–162) and red (109′–123′, 159′–174′ and 198′–206′). The additional structured residues of the cobalt-containing enzyme are at the C-terminus and depicted in green. Small black dots represent water molecules. Some side chains at the active center are given. As detailed in the text, the zinc ion shows anisotropic mobility along an axis through Zn^{2+} and His155 NE2.

Lei, Studnicka & Wilcox, 1985) had been reported earlier (Dreyer & Schulz, 1993) quoting incorrect RibE numbers. In Fig. 3 the RibE sequence has now been aligned with FucA, and the correct numbers are stated. The sequences are 34% identical in the N-terminal moieties of the enzymes (residues 1–125 of FucA), but not significantly related in the C-terminal parts. It should be noted that Thr26 in, and Gly28 next to the disordered loop of FucA as well as Ser71, Ser72, Asn29 and Thr43 (as Ser) binding the sulfate ion at the putative phosphate site of FucA are conserved in RibE. Moreover, the zinc ligands Glu73 (as Asp), His92 and His94 are conserved. The fourth zinc ligand His155 is in the C-terminal moiety and has no direct counterpart in RibE. A metal dependence of RibE has been reported for the *Aerobacter aerogenes* enzyme (Deupree & Wood, 1972), but not yet for the *E. coli* enzyme.*

We are grateful to W.-D. Fessner (Aachen) for much appreciated organizational help and for discussions. We are also grateful to E. C. C. Lin, Boston, for providing the cloned gene of FucA and to A. Schneider, Freiburg, for supplying us with Fuc1P.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1FUA, 2FUA, R1FUASF, R2FUASF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: SE0198).

References

Bednarski, M. D., Simon, E. S., Bischofberger, N., Fessner, W. D., Kim, M.-J., Lees, W., Saito, T., Waldmann, H. & Whitesides, G. M. (1989). *J. Am. Chem. Soc.* **111**, 627–635.

Belasco, J. G. & Knowles, J. R. (1983). *Biochemistry*, **22**, 122–129.

Blundell, T. L. & Johnson, L. N. (1976). *Protein Crystallography*, pp. 97–104. New York: Academic Press.

Brändén, C.-I. & Jones, T. A. (1990). *Nature (London)*, **34**, 687–689.

Brosius, J. & Holy, A. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 6929–6933.

Brown, D. C. & Collins, K. D. (1991). *J. Biol. Chem.* **266**, 1597–1604.

Brünger, A. T. (1992). *X-PLOR* Manual Version 3.0, Yale University, New Haven, Connecticut, USA.

Chakrabarti, P. (1990*a*). *Protein Eng.* **4**, 49–56.

Chakrabarti, P. (1990*b*). *Protein Eng.* **4**, 57–63.

Chen, Y.-M., Lu, Z. & Lin, E. C. C. (1989). *J. Bacteriol.* **171**, 6097–6105.

Chen, Y.-M., Zhu, Y. & Lin, E. C. C. (1987). *Mol. Gen. Genet.* **210**, 331–337.

Chothia, C. (1973). *J. Mol. Biol.* **75**, 295–302.

Connolly, M. L. (1993). *J. Mol. Graphics*, **11**, 139–141.

Deupree, J. & Wood, W. A. (1972). *J. Biol. Chem.* **247**, 3093–3097.

Dreyer, M. (1995). PhD thesis, University of Freiburg, Germany.

Dreyer, M. K. & Schulz, G. E. (1993). *J. Mol. Biol.* **231**, 549–553.

Dreyer, M. K. & Schulz, G. E. (1996). *J. Mol. Biol.* **259**, 458–466.

Fessner, W. D., Sinerius, G., Schneider, A., Dreyer, M., Schulz, G. E., Badia, J. & Aguilar, J. (1991). *Angew. Chem. Int. Ed. Engl.* **30**, 555–558.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., Fitzhugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Sproggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhermann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O. & Venter, C. (1995). *Science*, **269**, 496–512.

Gamblin, S. J., Cooper, B., Millar, J. R., Davies, G. J., Littlechild, J. A. & Watson, H. C. (1990). *FEBS Lett.* **262**, 282–286.

Gamblin, S. J., Davies, G. J., Grimes, J. M., Jackson, R. M., Littlechild, J. A. & Watson, H. C. (1991). *J. Mol. Biol.* **219**, 573–576.

Ghalambor, M. A. & Heath, E. C. (1962). *J. Biol. Chem.* **237**, 2427–2433.

Ghalambor, M. A. & Heath, E. C. (1966). *Methods Enzymol.* **10**, 538–542.

Gomis-Rüth, F.-X., Grams, F., Yiallouros, I., Nar, H., Kosthardt, U., Zwilling, R., Bode, W. & Stöcker, W. (1994). *J. Biol. Chem.* **269**, 17111–17117.

Hester, G., Brenner-Holzach, O., Rossi, F. A., Struck-Donatz, M., Winterhalter, K. H., Smit, J. D. G. & Piontek, K. (1991). *FEBS Lett.* **292**, 237–242.

Holland, D. R., Hausrath, A. C., Juers, D. & Matthews, B. W. (1995). *Protein Sci.* **4**, 1955–1965.

Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.

Horecker, B. L., Tsolas, O. & Lai, C. Y. (1972). *The Enzymes*, 3rd ed., Vol. 7, edited by P. D. Boyer, pp. 213–258. New York: Academic Press.

Izard, T., Lawrence, M. C., Malby, R. L., Lilly, G. G. & Colman, P. (1994). *Structure*, **2**, 361–369.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

Kadonaga, J. T. & Knowles, J. R. (1983). *Biochemistry*, **22**, 130–136.

Kitagawa, Y., Leonard, G. A., Harrop, S. J., Peterson, M. R., Hunter, W. W., Qamar, S. & Berry, A. (1995). *Acta Cryst.* D**51**, 833–834.

Lin, H. C., Lei, S. P., Studnicka, G. & Wilcox, G. (1985). *Gene*, **34**, 129–134.

Lipman, D., Altschul, S. F. & Kececioglu, J. D. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 4412–4415.

Littlechild, J. A. & Watson, H. C. (1993). *Trends Biochem. Sci.* **18**, 36–39.

Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

Marsh, J. J. & Lebherz, H. G. (1992). *Trends Biochem. Sci.* **17**, 110–113.

Mavridis, M., Hatada, M. H., Tulinsky, A. & Lebioda, L. (1982). *J. Mol. Biol.* **162**, 419–444.

Mineno, J., Fukui, H., Ishino, Y., Kato, I. & Shinagawa, H. (1990). *Nucleic Acids Res.* **18**, 6722–6722.

Naismith, J. H., Ferrara, J. D., Bailey, S., Marshall, K., Dauter, Z., Wilson, K. S., Habash, J., Harrop, S. J., Berry, A. & Hunter, W. N. (1992). *J. Mol. Biol.* **225**, 1137–1141.

Ozaki, A., Toone, E. J., van der Osten, C. H., Sinskey, A. & Whitesides, G. M. (1990). *J. Am. Chem. Soc.* **112**, 4970–4971.

Ponder, J. W. & Richards, F. M. (1987). *J. Mol. Biol.* **193**, 775–791.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Richard, J. D. (1984). *J. Am. Chem. Soc.* **106**, 4926–4936.

Riordan, J. F. & Christen, P. (1969). *Biochemistry*, **8**, 2381–2386.

Scamuffa, M. D. & Caprioli, R. M. (1980). *Biophys. Biochim. Acta*, **614**, 583–590.

Schneider, A. (1994). PhD thesis, University of Freiburg, Germany.

Sygusch, J., Beaudry, D. & Allaire, M. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 7846–7850.

Szwergold, B. S., Ugurbil, K. & Brown, T. R. (1995). *Arch. Biochem. Biophys.* **317**, 244–252.

Vallee, B. L. & Auld, D. (1992). *Faraday Discuss.* **93**, 47–65.

Weiss, M. S., Wacker, T., Weckesser, J., Welte, W. & Schulz, G. E. (1990). *FEBS Lett.* **267**, 268–272.

Wong, C. H., Halcomb, R. H., Ichikawa, Y. & Kajimoto, T. (1995*a*). *Angew. Chem. Int. Ed. Engl.* **34**, 412–432.

Wong, C. H., Halcomb, R. H., Ichikawa, Y. & Kajimoto, T. (1995*b*). *Angew. Chem. Int. Ed. Engl.* **34**, 521–546.